KAM theory near multiplicity one resonant surfaces in perturbations of a-priori stable Hamiltonian systems

Michael Rudnev
Applied Mathematics 217-50
Caltech
Pasadena, CA 91125

Stephen Wiggins
Departments of Applied Mechanics and
Control and Dynamical Systems 104-44
Caltech
Pasadena, CA 91125 *

May 14, 1997

Abstract

We consider a near-integrable Hamiltonian system in action-angle variables with analytic Hamiltonian. For a given resonant surface of multiplicity one we show that near a Cantor set of points on this surface, whose remaining frequencies enjoy the usual diophantine condition, the Hamiltonian may be written in a simple normal form which, under certain assumptions, may be related to the class which, following Chierchia and Gallavotti [1994], we call *a-priori unstable*. For the a-priori unstable Hamiltonian we prove a KAM-type result for the survival of whiskered tori under the perturbation as an infinitely differentiable family, in the sense of Whitney, which can then be applied to the above normal form in the neighborhood of the resonant surface.

1 Introduction

This paper is concerned with certain aspects of multiplicity-one resonant dynamics for a generic $n-$degree of freedom a-priori stable system, i.e. a near-

*This research was partially supported by NSF Grant DMS-9403691.

integrable non-degenerate Hamiltonian system expressed in action-angle variables whose Hamiltonian is real-analytic in some open domain. The number of degrees of freedom n is greater than 2. We consider a Cantor set of points on a codimension-one resonant plane $\omega_1 = 0$ in the frequency space, whose $n-1$ remaining frequencies satisfy a classical diophantine condition. This Cantor set remains fixed throughout our analysis. The main theorem we prove establishes local integrability near this set, if several assumptions are satisfied. We show that one may take a comparably large neighborhood of this Cantor set, namely a union of open n-balls centered at every point of this set, whose size will be defined by the norm of the perturbation, the parameters of the problem, and the number of degrees of freedom, where a simple normal form may be derived. The latter under certain assumptions may be regarded as a perturbed a-priori unstable system, namely a system whose integrable part possesses a separatrix. The a-priori stable systems allowing such a construction we call *hyperbolic in the first order.*

For this reason at the beginning we devote a separate section to a-priori unstable systems. We present a KAM-type proof to establish the existence of a differentiable, in the sense of Whitney, family of whiskered tori which survive a perturbation. Generally we follow the well established iterative scheme due to Arnold. Our approach relies heavily on Pöschel [1982], [1993]. Further we apply this result to the neighborhood of the resonant surface for the initial a-priori stable system. There we argue that the one-parameter families of $n-1$-dimensional invariant tori, corresponding to the multiplicity-one resonances for the unperturbed problem, break up under perturbation and yield hyperbolic $n-1$-tori with their stable and unstable manifolds (whiskers). In this way we build up a global structure in the neighborhood of the entire resonant surface as a differentiable family of whiskered tori, thus relating different whiskered tori to each other, which is further expected to enable us to advance in the search for the heteroclinic intersection of their whiskers and to establish locally a possibility of drift along the resonant surface, where the density of the whiskered tori is sufficiently high.

Related results can be found in Graff [1974], where a novel *non-singular* extension of the KAM theory to systems with hyperbolicity was set up and much more, and later in de la Llave and Wayne [1990], Treschev [1991], and Chierchia and Gallavotti [1994], where the terms "a-priori stable" and "a-priori unstable" were coined.

However, our work differs from the forementioned references both in its spirit and technique. We consider it the first builiding block in our research project concerning the so-called Arnold's diffusion problem, or rather *Arnold's mechanism*, suggested in Arnold [1964], the latter being a much less general and simpler problem than the former, as pointed out by Lochak [1995], though far from being completely understood. The main purpose of this piece is to suggest a useful Normal form, globally valid in the neighborhood of a multiplicity-one resonant surface that describes simultaneously all the weakly hyperbolic tori with frequencies satisfying some diophantine condition, and straightens out there local stable and unstable manifolds, the kind of Normal form that can further be used as the starting point for analysis of the homoclinic splitting. In this connection all the quoted references except for Chierchia and Gallavotti [1994] deliver only local results that explicitly pertain to only one single torus, and are consequently insufficient for being immediately used in search for the heteroclinic orbits, the last one (namely the part of it which proves the analogous result for the a-priori unstable systems) is close in its philosophy to what we suggest, but we believe that the technique we use, in particular Lie transforms and what we call the Absolute Norm (see Appendix A), a modification of the weighted norm introduced by Pöschel [1993], makes it less unwieldy and essentially easier to conceive and to use. Withal, we are a bit more concerned with the sharpness of estimates, performed in the course of the proof, which becomes important for further considering the splitting of separatrices, since the latter problem is exclusively sensitive with respect to different parameters and their relation to each other; generally we believe that most of the dependencies we suggest are in a way optimal.

We want to emphasize that although in the section devoted to the a-priori unstable systems where we pay tribute to the tradition founded by Poincaré in his "Méthodes nouvelles.." by introducing an "independent" perturbation parameter μ apart from ε, throughout the whole presentation we are dealing with the *singular* problem (see also Treschev [1991]), which turns out to be very similar to a non-singular one as far as the construction of the Normal form is concerned with all the modifications included into the smallness condition for the perturbation parameter.

In conclusion of this section we also want to point out that our approach differs from the one used by Graff [1974] and Treschev [1991] in about the same sense as different are the ways to prove the standard KAM theorem, suggested by Kolmogorov [1954] and Arnold [1963]. The Whitney-smoothness of the family of the perturbed whiskered tori along the resonant surface that we establish is the essential fact for further study of the heteroclinic intersections between their whiskers.

2 Main result

2.1 Set-up and notation

We study a Hamiltonian system whose Hamiltonian has the form

$$H(I,\varphi,\varepsilon) = H^0(I) + \varepsilon H^1(I,\varphi,\varepsilon) \tag{1}$$

in the action-angle variables $(I,\varphi) \in D \times T^n$, where D is an open subset of R^n, T^n is an n-torus, obtained from R^n by identifying points whose coordinates differ by integer multiples of 2π, and is analytic in some complex neighborhood of the above domain, and smooth enough in the small real non-negative perturbation parameter ε. We will assume that H^1 is uniformly bounded for all small values of ε and that its Taylor series in ε has a nontrivial zero order term. H^1 is assumed to be 2π-periodic in each of the angle variables $\varphi_i,\ i = 1, 2, .., n$, where $n > 2$.

The standard KAM theorem argues that for the Hamiltonian (1) in the action-angle variables, which for $\varepsilon = 0$ describes a simple dynamics on tori with constant frequencies $\omega(I) = H^0_I$, the introduction of the perturbation does not affect most of these tori (having asymptotically full measure as $\varepsilon \to 0$), causing only their slight deformation, unless their frequencies are near resonances, which are described by condition that $k \cdot \omega = 0$ for some non-trivial integer vector $k \in Z^n$.

Here we consider a multiplicity one resonance, i.e., there is only one independent integer vector satisfying $k \cdot \omega = 0$. Geometrically, this relation defines a hyperplane in the frequency space which we refer to as a resonant plane. By a suitable linear symplectic transformation of the coordinate system this resonant plane can be transformed to $\omega_1 = 0$. We will assume this has been done. Hence a hyperplane

$$P \equiv \{\omega : \omega_1 = 0\} \tag{2}$$

in the *frequency space* Ω we call the *resonant plane.* Its image in the action space

$$\Pi \equiv \{I : \omega_1(I) = 0\} \tag{3}$$

we call the *resonant surface.* Further we'll need several non-degeneracy assumptions on the unperturbed *frequency map* $\Psi_0 :\ \omega(I) = H^0_I$. Qualitatively, we assume that it is *locally invertible* near the whole resonant surface (otherwise we would have to restrict ourselves to only a part of it). This will allow us to consider along with the a ction-angle variables (I, φ) also the frequency-angle variables (ω, φ), though non-canonical. The frequency-angle variable formulation used by Pöschel [1982] allows certain conceptual simplifications that prove convenient.

Henceforth we will extend the action of the frequency map Ψ and its inverse Ψ^{-1} over all the other variables and parameters apart from I upon which the Hamiltonian may depend as the identity transformation:

$$\Psi(I, \cdot) = (\omega(I), \cdot), \quad \Psi^{-1}(\omega, \cdot) = (I, \cdot).$$

For functions of the frequency-angle variables we will use calligraphic letters. In particular, for the Hamiltonian (1), written in these variables, we write

$$\mathcal{H}(\omega, \varphi, \varepsilon) = H \circ \Psi^{-1}. \tag{4}$$

This will be helpful, since in the frequency space the resonant condition holds on a hyperplane; besides, the construction we make is quite simple in this representation.

We shall sometimes admit a slight abuse of rigour, simply saying "the frequency variables" or "the action variables" without referring to the other (angle or hyperbolic) variables, that the Hamiltonian in consideration may depend on, basically meaning transiting back and forth under the action of the just generalized transformations Ψ and Ψ^{-1}.

On a resonant plane P we identify a Cantor set Ω_γ of "good" diophantine frequencies, let I^0_γ be its image through the inverse of the frequency map. Quantitatively,

$$\Omega_\gamma \equiv \{\omega:\ \omega_1 = 0,\ |\tilde{\omega}\cdot k| \geq \gamma |k|^{-\tau},\ \forall k \in Z^{n-1}\setminus\{0\}\}, \tag{5}$$

where $\tilde{\omega} \equiv (\omega_2, .., \omega_n)$ and $\gamma > 0$, $|k| = \sum_{i=1}^{n-1} |k_i|$; $\tau \geq n-2$ is kept fixed. For a small γ the relative measure of Ω_γ on the resonant plane is big if $\tau > n-2$. Further we will often write ω, instead of $\tilde{\omega}$, just keeping in mind that for the given a-priori stable system we always stay close to the plane P.

Remark: In fact, we shall consider the intersection of D and a neighborhood of Π which is the image in the action space of some full-dimensional neighborhood of Ω_γ through the inverse of the frequency map, or, equivalently, a subset $\tilde{\Omega}_\gamma \subseteq \Omega_\gamma$, such that the image of some its full-dimensional neighborhood lies in D. To avoid the unnecessary notational complications we will always assume that D itself and its complex extension will contain, respectively, all the real sets in the action space and their complex extensions that we will be dealing with, including Π.

At this point we will introduce a necessary piece of notation. Here we assume that apart from the action-angle variables (I, φ) the Hamiltonian also depends on two extra variables (p, q), to which we refer as the *hyperbolic variables*, given in an open real ball $\bar{B}^2_\kappa \equiv \{(p,q) \in R^2 : |p|^2 + |q|^2 < \kappa^2\}$ with radius $\kappa > 0$ near the origin, namely,

$$H = H(I, \varphi, p, q). \tag{6}$$

We will need this in the following discussion about the a-priori unstable systems. If there is no dependence on (p, q), it must be just ignored in the notation for domains and norms in an obvious way by skipping the indices, corresponding to these variables.

Further, $r, s, \kappa, \rho, \sigma, \eta$ will stand for some positive parameters.

Henceforward, the notation $\bar{B}^l_\kappa$ (B^l_κ) will stand for a real (complex) ball in R^l (C^l) of radius κ. For instance,

$$B^2_\kappa \equiv \{(p,q) \in C^2 : |p|^2 + |q|^2 < \kappa^2\}.$$

Also if S is an arbitrary set in R^l, which stands for either the action or the frequency space, then for an arbitrary dimension l we write the complex extension of the domain of definition of (6) as:

$$S+(r,s,\kappa) \equiv S+(r) \times W_sT^l \times B^2_\kappa \subseteq C^l \times C^l \times C^2,$$

where

$$S+(r) \equiv \{z \in C^l : |z-S| < r\}$$

in the $sup-$norm, and

$$W_sT^l \equiv \{\varphi \in C^l : |\varphi| \leq s\}$$

in the $sup-$norm. Equivalently sometimes we write $S+(r,s,\kappa)$ as $S_{r,s,\kappa}$. Also we reserve the notation $S+(r,\kappa)$, or $S_{r,\kappa}$ for the direct product

$$S+(r,\kappa) \equiv S+(r) \times B^2_\kappa \subseteq C^l \times C^2.$$

For a complex set $\mathcal{S} \subseteq C^{l_0}$ which, like $S_{r,s,\kappa}$ is the direct product of three complex bounded sets $\mathcal{S}_1 \subseteq C^{l_1},\ \mathcal{S}_2 \subseteq C^{l_2},\ \mathcal{S}_3 \subseteq C^{l_3}$, where $l_1+l_2+l_3 = l_0$ we will write $\mathcal{S}-(\rho,\sigma,\eta)$ when we shrink $\mathcal{S}$, cutting off the border layers of thickness ρ,σ,η from the corresponding sets $\mathcal{S}_i,\ i=1,2,3$ of $\mathcal{S}$ and taking the direct product.

Now we will define the Absolute norm we will be using, for more information a reader is referred to Appendix A which contains a summary of the properties of this norm and its relation to a more traditional $sup-$norm. Given the function $u(\cdot,p,q)$, analytic in $p,q \in B^2_\kappa$ we define as $\overline{u(\cdot,p,q)}$ the absolute sum of its Taylor series in p,q coordinates near $(p,q)=(0,0)$. Namely if

$$u(\cdot,p,q) = \sum_{i,j=0}^{\infty} u(\cdot)_{ij}p^iq^j,$$

then

$$\overline{u(\cdot,p,q)} \equiv \sum_{i,j=0}^{\infty} |u(\cdot)_{ij}p^iq^j|.$$

Then, given an analytic in $S_{r,s,\kappa}$ function $u(x,p,q,\varphi)$, where $x \in S+(r)$ stands for either the action or the frequency variables, with a Fourier series

$$u(\cdot,\varphi) = \sum_{k \in Z^{n-1}} u^k(\cdot)e^{ik\cdot\varphi},$$

we define its absolute norm as

$$|u|_{S_{r,s,\kappa}} \equiv \sup_{(x,p,q)\in S_{r,\kappa}} \sum_{k\in Z^{n-1}} \overline{u^k(x,p,q)}e^{|k|s}.$$

Also, for norms of functions given on the complex extension $\mathcal{S}$ of the fixed domain C in the action or the frequency space $\mathcal{S} = S_{r,s,\kappa}$ we will often write $|\cdot|_{\mathcal{S}}$, or $|\cdot|_{r,s,\kappa}$ instead of $|\cdot|_{S,r,s,\kappa}$. For an arbitrary vector-function v the corresponding norm will be $|v|_{r,s,\kappa} \equiv \max_i |v_i|_{r,s,\kappa}$, for an arbitrary matrix A the corresponding norm $|A|_{r,s,\kappa} \equiv \max_{i,j} |A_{ij}|_{r,s,\kappa}$. Often in the notation for domains and norms we will omit the indices, in which we won't be interested at the moment, where it does not lead to confusion.

To finish with the introduction, below we give several more definitions and pieces of notation we will be using. For an analytic function $u(\cdot, p, q)$ we will write $\mathcal{D}u$ to denote the part of its Taylor series in (p, q), which depends only on the product pq, namely

$$\mathcal{D}u(\cdot, p, q) \equiv \sum_{i=0}^{\infty} u(\cdot)_{ii}(pq)^i.$$

We will use the $\mathcal{P}$ symbol for a norm $|I, \varphi, p, q|_{\mathcal{P}} \equiv \max(|I|, |\varphi|, |p|, |q|)$ and for the induced operator norm.

If $f = \sum_{k \in Z^n} f_k e^{ik\cdot\varphi}$ denotes the Fourier expansion of a function, then we define $T_K f \equiv \sum_{k \in Z^n_K} f_k e^{ik\cdot\varphi}$, where $Z^n_K \equiv \{k \in Z^n \mid |k| \leq K\}$. We refer to $T_K f$ as the *ultraviolet cutoff* of f.

We assume that for some open domain $D \subseteq R^n$ in the action space the Hamiltonian (1) is real analytic in $D + (r_0, s_0)$ for some positive parameters r_0, s_0. Besides, suppose, for all small positive values of ε for $|H^1|_{r_0,s_0} \leq 1$.

2.2 Main theorem and General Strategy

Let's consider the unperturbed resonant surface Π and the image I^0_γ of the Cantor set Ω_γ on it through the inverse of the frequency map Ψ^{-1}. For the unperturbed problem each point of this set corresponds to an $n-$dimensional multiplicity-one resonant torus, which is foliated into a one-parameter family of tori of dimension $n-1$. Under certain assumption, as we'll show, for this family the perturbation brings in hyperbolicity in the first order in ε. We will call perturbed systems of this kind *hyperbolic in the first order.* More technically speaking, by applying once a canonical near-identity transformation $\mathcal{R}_0$ which will be provided by our Main Lemma, we derive a simple normal form for the Hamiltonian (1), valid in some full-dimensional open neighborhood of the set I^0_γ, whose truncation up to the first-order terms in ε for $\varepsilon \neq 0$ is integrable, and may be considered as an a-priori unstable system. We will show that (I_1, φ_1) may be considered as the "hyperbolic variables". We will redefine them as (p, q) and reserve the notation (I, ω, φ) for the rest of the actions, frequencies and angles.

A well-known result (Moser [1956], Chierchia and Gallavotti [1994]) says that locally in the neighborhood of a hyperbolic equilibrium point the integrable

Hamiltonian of the a-priori unstable system by applying some canonical transformation $\mathcal{R}$ may be written in some new canonical coordinates (I', φ', p', q') such that $I = I'$, $J = p'q'$ simply as a function of (I, J). So we can apply this result to the truncation of the above normal form and if we proceed thinking of it as the "unperturbed" Hamiltonian, and the remainder as a perturbation, hence the question basically boils down to the KAM theory for the a-priori unstable systems.

Part of discussion devoted to the a-priori unstable systems will be presented in a self-contained fashion, in particular the notation will be entirely independent. The basic strategy of the iterative KAM-type proof for the a-priori unstable systems will be as follows. At every step, numbered by $i = 0, 1, ..$, the Hamiltonian H_i will be the sum of two components: the "unperturbed Hamiltonian" $h(I, J)_i$ plus a perturbation $f(I, \varphi, p, q)_i$ which we target to make smaller and smaller. The way we position the index i is also to point out that on every step we are actually dealing with different coordinate systems. The "frequency map" on each step will be defined as $\Psi_i : \ (I, \cdot) \to (\omega = \partial_I h_i, \cdot)$; as mentioned in Section 2.1 we consider its action upon all the other variables and parameters as an identity. Then we can write the Hamiltonian in the frequency variables simply as $\mathcal{H}(\omega, \varphi, p, q)_i = H_i \circ \Psi_i^{-1}$. When we apply to H_i the Main lemma, which will be our technical tool to seek the canonical changes of variables, thus reducing the size of the perturbation, the unperturbed Hamiltonian picks up a part $\mathcal{D}(T_K \bar{f})_i$ of the Fourier series for the perturbation, truncated with some ultraviolet cutoff parameter K, growing bigger from step to step, averaged in the "fast" angles, which depends only on the actions I and the product pq. So, the new unperturbed Hamiltonian becomes $h_{i+1} = h_i + \mathcal{D}(T_K \bar{f})_i$, it defines a new frequency map Ψ_{i+1}, which is close to Ψ_i.

To accomplish this we use Lie transforms and the canonical variables (I, φ, p, q). In the Main lemma that we use on each step we construct a Hamiltonian flow $X^t_{\phi_i}$ with Hamiltonian ϕ_i and define the transformation we are looking for in the *canonical* variables as $\Xi_i = X^1_{\phi_i}$. Then in the frequency variables the corresponding transformation will be $\Phi_i = \Psi_i \Xi_i \Psi_{i+1}^{-1}$. It is in terms of this transformation that we formulate the result of the Main lemma and work out all the estimates for the Lie transform Hamiltonian, using the standard Cauchy inequalities. We do it to make sure that the projection of the range of Φ_i onto the frequency space lies within a union of the open balls $\Omega_\gamma + (r_i)$ with radius r_i around Ω_γ, the latter being fixed once and for all, then the projection of its domain must be the union of balls of smaller size r_{i+1}.

Thus we construct a sequence of open domains $\{\Omega_\gamma + (r_i)\}$, where the sequence $\{r_i\}$ rapidly goes to zero, then the limit for the above sequence of the domains is a Cantor set Ω_γ. The transformation $\Phi = \Phi_0 \circ \Phi_1 \circ ...$, then acts from $\overline{\Omega_\gamma + (0, s_*, \kappa/2)}$ into $\Omega_\gamma + (r, s, \kappa)$ for some initial values (r, s, κ) of the analyticity parameters, and $s_* < s$; it is C^∞ in the frequency variables in the sense of Whitney and analytic in (φ, p, q). Finally we define the limit frequency map as

$\Psi_\infty : (I, \cdot) \to (\lim_{i\to\infty} \partial_I h_i, \cdot)$ and then the transformation $\Xi \equiv \Psi_0^{-1}\Phi\Psi_\infty$ will be canonical on a closed set $\Psi_\infty^{-1}\Omega_\gamma$. The latter set is certainly not as palpable as Ω_γ.

Finally taking the composition $\Xi_* = \mathcal{R}_0 \circ \mathcal{R} \circ \Xi$ we arrive at our main result which is formulated in our Main Theorem.

At this point we want to quote the Whitney Extension Theorem which allows us to speak about the differentiable family of the surviving tori. Using the fact that Ω_γ is closed, Whitney's theorem enables one to extend the above transformation Φ to the open domain in the frequency space containing the Cantor set Ω_γ in such a way that being restricted to Ω_γ, this extension $\mathcal{E}\Phi$ will boil down to the transformation Φ itself.

Theorem 2.1 (The Extension Theorem) *For any closed set $F \in R^n$ there exists a linear extension operator $\mathcal{E} : C^\infty(F) \to C^\infty(R^n)$, for $u \in C^\infty(F)$ $u \to U \equiv \mathcal{E}u$, such that for all the derivatives of u in the sense of Whitney $D^kU|_F = u^{(k)}$, $k = 0, 1, \ldots$ and $|U|_{R^n} \leq c|u|_F$ in the C^∞-norm.*

Remark: The extension $\mathcal{E}\Xi$, will be canonical *only* being restricted to a Cantor set $\Psi_\infty^{-1}\Omega_\gamma$ by construction.

For definitions of Whitney differentiation the reader is referred to Pöschel [1982].

During our treatment, following Pöschel [1982], we will fix the value of the parameter γ in the diophantine condition (5) equal to 1, thus dealing with the Cantor set Ω_1 rather than with Ω_γ. The result we prove then can be easily reformulated for any $\gamma < 1$, because Ω_1 may be obtained from Ω_γ by "blowing up" the frequencies, by a factor γ^{-1} and multiplying the Hamiltonian by γ^{-2}.

Applying the KAM theory for the a-priori unstable systems to the neighborhood of the multiplicity-one resonant surface of the a-priori stable system, which is hyperbolic in the first order, we arrive at our main result, which we formulate now; the rest of this paper is devoted to its proof. The formulation we give here is at one point qualitative, as far as the exact definition of hyperbolicity in the first order is concerned. Putting all rigorously will require the quantitative formulation of the assumptions we've mentioned, that we do in Section 4.

Theorem 2.2 (Main Theorem) *For the non-degenerate Hamiltonian system (1) hyperbolic in the first order there exists ε_0 small enough depending only on the analyticity and non-degeneracy parameters of the problem and parameter $\tau \geq 1$, proportional to γ^2, such that if*

$$|H^1|_{r_0,s_0} \leq 1$$

for all $0 < \varepsilon \leq \varepsilon_0$, there exists a C^∞-differentiable family in the sense of Whitney of invariant whiskered tori $\mathcal{T}_\varepsilon$, whose projection on the action space is ε-close to the Cantor set $I_\gamma^0 = \Psi_0^{-1}\Omega_\gamma$.

Moreover, there exists a transformation Ξ_*, such that the projection of its range on the action space lies in the small neighborhood of the resonant surface Π, and whose action on ε is trivial:

$$\Xi_* : \ (p, I', q, \varphi', \varepsilon) \to (I, \varphi, \varepsilon),$$

where (I', φ') are the new $n-1$ action-angles and (p, q) are the hyperbolic variables.

The transformation Ξ_* casts the Hamiltonian (1) into the form

$$H \circ \Xi_* = H'(I', pq, \sqrt{\varepsilon}), \tag{7}$$

which is as smooth in $\sqrt{\varepsilon}$ as (1) is in ε for $0 \le \varepsilon \le \varepsilon_0$.

The new Hamiltonian (7) defines the new frequency map $\Psi_* : \ (I', \cdot) \to (\omega' = \partial_{I'} H'(I', J, \sqrt{\varepsilon}), \cdot)$, such that its inverse $\Psi_*^{-1} : \ (\omega', \cdot) \to (I'(\omega', J, \sqrt{\varepsilon}), \cdot)$ is real analytic in J for $J = pq$, smooth in $\sqrt{\varepsilon}$ and C^∞ in the $n-1$ frequencies ω'.

The maps Ξ_* and $\Phi_* = \Psi_0 \circ \Xi_* \circ \Psi_*^{-1}$ are real analytic in variables $\varphi' \in W_{s_*} T^{n-1}$, $(p, q) \in B^2_{\kappa_*}$ for some small positive parameters $s_* < s_0$ and $\kappa_* = O(\sqrt{\varepsilon})$. Besides, the transformation Φ_* is C^∞ in the $n-1$ frequency variables ω' on the whole resonant plane P, also for real J the transformation Ξ_* will be C^∞ in the $n-1$ action variables I'.

Moreover, the restriction of Ξ_* in the action space to the Cantor set $\Psi_*^{-1}\Omega_\gamma$ is canonical with respect to the standard symplectic structure in the new variables (I', φ', p, q):

$$\omega^2 = dI' \wedge \varphi' + dp \wedge dq.$$

In the new variables the equations of motion for $I' \in I'_\gamma \equiv \Psi_*^{-1}\Omega_\gamma$ are given by:

$$\begin{aligned} I' &= I'_0, \\ \varphi' &= \varphi'_0 + \omega' t, \ \ \omega' \in \Omega_\gamma, \\ p &= p_0 e^{-\lambda' t}, \\ q &= q_0 e^{\lambda' t}, \end{aligned} \tag{8}$$

where $\lambda' = \partial_J H'(I', J, \sqrt{\varepsilon}) = O(\sqrt{\varepsilon})$ is a real analytic function of J, smooth in $\sqrt{\varepsilon}$, such that its real part at $J = 0$ is strictly positive for $\varepsilon > 0$.

The following corollary is the restatement of the part of this theorem describing the transformation Ξ_*, so its proof is immediate.

Corollary 1 *The tori $\mathcal{T}_\varepsilon$ are a part of invariant $n-$dimensional surfaces, which can be analytically parametrized by $(p, q) \in B^2_{\kappa_*}$ and $\varphi' \in W_{s_*} T^{n-1}$ as:*

$$I_i = I'_i(pq, \omega', \sqrt{\varepsilon}) + \varepsilon \Theta_1(I', \varphi', p, q, \sqrt{\varepsilon}), \ \ i = 2, .., n;$$

$$
\begin{aligned}
\varphi_i &= \varphi_i' + \Theta_2(I', \varphi', p, q, \sqrt{\varepsilon}),\ i = 2, .., n; \\
I_1 &= \Theta_3(I', \varphi', p, q, \sqrt{\varepsilon}); \\
\varphi_1 &= \Theta_4(I', \varphi', p, q, \sqrt{\varepsilon}),
\end{aligned} \tag{9}
$$

where $I'(pq, \omega', \sqrt{\varepsilon}) \in \Psi_*^{-1}\Omega_\gamma$, and the smoothness class of functions I', Θ_i, $i = 1, .., 4$ in $\sqrt{\varepsilon}$ is the same as originally for H in (1) in ε.

Remark: In fact one can say more about functions $\Theta_2, \Theta_3, \Theta_4$ in the statement of the above Corollary, since as we've mentioned before the transformation Ξ_* is constructed in several steps. The reader can readily retrieve these details, that we choose not to mention in the general formulation, after studying the paper.

3 A-priori unstable systems

This section contains the proof of the self-contained result, namely the KAM-type theorem on local integrability of the perturbed a-priori unstable system on a Cantor set Ω_γ in the frequency space. The proof consists of the infinite repetitive use of the Main lemma, which will be our principal technical tool. This result will subsequently be applied to the neighborhood of the resonant plane P of the a-priori stable system from our earlier discussion.

Nevertheless, the following discussion may be regarded independently. In particular, functions and quantities that stand behind the notation H, n, r, s, ε, etc. are generally different from those of the previous section.

3.1 The set-up for a-priori unstable systems

The focus of this section is a system with Hamiltonian, which apart from the action-angle variables (I, φ), depends on an extra pair of variables (p, q), which we refer to as the *hyperbolic variables*:

$$
H(p, I, q, \varphi, \varepsilon, \mu) = h_0(I, \varepsilon) + \tilde{P}(p, I, q, \varepsilon) + \tilde{f}(p, I, q, \varphi, \varepsilon, \mu). \tag{10}
$$

This Hamiltonian will be real analytic on some complex domain $G + (r, s, \kappa)$ in the notation of Section 2.1. Here G is some open real domain in the action space: $I \in G \subseteq R^{n-1}$, functions of the angles φ are assumed to be $2\pi-$periodic in these variables; $(p, q) \in B^2_\kappa$. We will consider ε, μ as two independent small parameters such that μ is complex, $|\mu| \ll \varepsilon$ and ε is real and positive. We will assume that the Hamiltonian (10) is analytic in μ for its values in the complex ball $B^1_{\mu_0}$ near the origin with a positive radius μ_0, and smooth in ε. Mostly in this section we will omit the dependence of (10) in ε, which is not important unless the conditions of the lemmata developed below are violated. However, we shall keep in mind that all the non-degeneracy and analyticity parameters of the unperturbed (corresponding to $\mu = 0$) problem are in fact $\varepsilon-$dependent.

Speaking about the coordinate transformations, we will imply that their action on parameters ε, μ is trivial.

We will assume that the Taylor series for $\tilde{f}$ in μ begins with the first-order terms. Also for $\mu = 0$ we assume that for all I, ε $(p, q) = (0, 0)$ is a hyperbolic equilibrium point, with the real part of its Lyapunov exponent λ:

$$\lambda^2(I, \varepsilon) \equiv (\partial^2_{pq}\tilde{P})^2 - \partial^2_{pp}\tilde{P}\partial^2_{qq}\tilde{P}\Big|_{(p,q)=(0,0)}$$

being strictly positive for all I, ε. Without loss of generality we may also assume that for all I, ε $\tilde{P}(0, I, 0, \varepsilon) = 0$. Otherwise we can always achieve this by adjusting h_0 We will often refer to the hyperbolic part of the system (10), given by $\tilde{P}$ as a "pendulum" and the remaining action-angle variables as "rotors". Following Chierchia and Gallavotti [1994], if the Hamiltonian (10) satisfies the basic non-degeneracy assumptions, which we will shortly formulate quantitatively, the Hamiltonian system (10), is called *a-priori unstable*. In this case we can bring such a system into a Normal Form near this equilibrium, see, for example, Moser [1956] or Chierchia and Gallavotti [1994]. In fact, the proof of the following lemma is very much similar to the proof of the Theorem 3.1 that follows, although much simpler, since the unperturbed Hamiltonian does not depend on the angles, thus depending on I as a parameter, and the reader can simply synthesize it from the proof of the Theorem 3.1. Besides, in the case when P does not contain I, which is usually considered as the application, for instance the so-called Thirring model, see e.g. Gallavotti [1994] for an alternative consideration, the analysis would not contain the I variable at all.

Lemma 3.1 *For every $I^* \in G$, any $\varepsilon > 0$, and $|\mu|$ small enough there is a complex ball $B^{n-1}_{r'}$ centered in I^* in the action space with radius $r'(I^*)$ such that there exists a real analytic canonical transformation $\mathcal{R}$: $(I', \varphi', p', q') \in B^{n-1}_{r'} \times W_{s'}T^{n-1} \times B^2_{\kappa'} \to (I, \varphi, p, q) \in B^{n-1}_r \times W_s T^{n-1} \times B^2_\kappa$, for some smaller values of the analyticity parameters $r' \leq r$, $s' < s$, $\kappa' < \kappa$, where κ' may be made proportional to the minimum of the real part of the Lyapunov exponent $\lambda(I, \varepsilon)$. $\mathcal{R}$ is smooth in ε, and its action on I, ε is trivial. Moreover, the transformed Hamiltonian becomes:*

$$\begin{aligned} H \circ \mathcal{R} &= h_0(I', \varepsilon) + P(I', p'q', \varepsilon) + f(p', I', q', \varphi', \varepsilon, \mu) \\ &\equiv h(I', J', \varepsilon) + f(p', I', q', \varphi', \varepsilon, \mu), \end{aligned} \tag{11}$$

where $h = h_0 + P$.
The change of variables, brought by $\mathcal{R}$ may be written as follows:

$$\begin{aligned} I &= I', \\ \varphi &= \varphi' + \Theta^0_2(I', \varphi', p', q', \varepsilon), \\ p &= \Theta^0_3(I', \varphi', p', q', \varepsilon), \\ q &= \Theta^0_4(I', \varphi', p', q', \varepsilon). \end{aligned} \tag{12}$$

Besides, the Lyapunov exponent remains unchanged under the transformation, and one has $\lambda(I',\varepsilon) = \partial'_J h|_{J'=0}$ *for* $J' = p'q'$.

Remark: If $\tilde{P}$ does not contain the $I-$dependence, then obviously the result of this Lemma holds globally for all I and $s' = s$.

For simplicity we will further assume that a transformation $\mathcal{R}$ exists globally for $I \in G$, otherwise the domain G may be taken as small as $B^{n-1}_{r'}$ above. We will further take the Hamiltonian (11) for granted as the initially given Hamiltonian and omit primes in the notation for variables and parameters.

So, in fact we start with the Hamiltonian

$$\begin{aligned} H &= h_0(I,\varepsilon) + P(I,pq,\varepsilon) + f(p,I,q,\varphi,\varepsilon,\mu) \\ &\equiv h(I,J,\varepsilon) + f(p,I,q,\varphi,\varepsilon,\mu), \end{aligned} \tag{13}$$

where $J = pq$, analytic in $G_{r,s,\kappa}$ for some positive values of analyticity parameters r, s, κ.

The unperturbed Hamiltonian equations, given by (13) for $\mu = 0$ imply that for every I, ε there exists an invariant $n-1$-torus

$$I = const,\ \varphi = \varphi_0 + \omega t,\ (p,q) = (0,0),$$

where $\omega = \partial_I h(I,0,\varepsilon)$. This torus is a part of an invariant $n-$dimensional surface, described by the equations above plus the equations for the stable and unstable manifolds, or whiskers:

$$p = p_0 e^{-\lambda t},\ q = q_0 e^{\lambda t}.$$

In the next section, applying the results of this section to a-priori stable systems, we will be dealing with the situation when $\mu = \varepsilon^p$ for $1 < p < 2$ and will show that the assumptions of the theorem may be satisfied for any $p \geq \frac{3}{2}$ if $\varepsilon > 0$ is small enough.

3.2 Survival of the Quasiperiodic Tori under Perturbations

In section 2 we defined a Cantor set of "good" frequencies Ω_γ. We will use the same notation for the set with the same diophantine properties in the frequency space of the rotors. The domain where the Hamiltonian is defined and analytic in the action space of the rotors must include the image of the union of open balls with some fixed positive radius, centered at points of Ω_γ. To formulate the result we need several assumptions:

Assumption 1

- *Uniformly on $G_{r,\kappa}$ for all positive ε and some positive quantity $R(\varepsilon)$ the sup-norms of the Hessian matrix $A = \partial^2_{II} h(I,J,\varepsilon)$ and its inverse are bounded as*

$$|A|^\infty,\ |A^{-1}|^\infty \leq R(\varepsilon). \tag{14}$$

- *The real part of the Lyapunov exponent $\lambda \equiv \partial_J h(I,J,\varepsilon)|_{J=0}$ of the pendulum is bounded away from zero uniformly in $G_{r,\kappa}$ for all positive ε by a positive quantity $\Lambda(\varepsilon)$.*

In particular, from definition of R it's clear that $R \geq 1$, since it's a maximum of a norm of a matrix and its inverse.

Throughout this section we'll be mostly performing our construction in the frequency space of the rotors rather than in their action space. We'll keep in mind though that the frequency-angle coordinates are generally non-symplectic.

Remark: As in the previous section, we fix the value of the parameter γ in the diophantine condition (5) equal to 1 to ease the proof. indeed, if we stretch actions and hyperbolic variables as $I \to \gamma I,\ p \to \gamma p,\ q \to \gamma q$ and multiply the Hamiltonian by γ^{-2}, Ω_γ gets blown up into Ω_1 whereas the Hessian matrix and the Lyapunov exponent remain the same.

By redefining r as $\frac{r}{R}$ we insure that the image of $\Omega_1 + (r,s,\kappa)$ through the inverse of the frequency map lies within the initial domain of analyticity.

The Hamiltonian (13), written in the frequency-angle variables of the rotors we'll denote as:

$$\mathcal{H}(\omega,p,\varphi,q,\varepsilon,\mu) \equiv H \circ \Psi^{-1} = \mathcal{H}^0(\omega,J,\varepsilon) + \mathcal{F}(\omega,p,\varphi,q,\varepsilon,\mu), \tag{15}$$

The Hamiltonian (15) will then be analytic in $\Omega_1 + (r,s,\kappa)$, as well as (13) in the image of this domain through the inverse of the frequency map.

Henceforward, for functions of the frequencies, denoted by calligraphic letters, $|\cdot|_{r,s,\kappa}$ will stand for the Absolute norm in $\Omega_1 + (r,s,\kappa)$ for some positive values of analyticity parameters r,s,κ.

Besides, we shall assume for convenience that $r,s,\kappa \leq 1$, although the proof shows that this certainly is not a crucial necessity.

Now we will formulate and prove the main theorem of this section. This formulation will be more technical, in particular, it deals with $\gamma = 1$. Further, regarding our Remark, we'll point to the amendments one has to make to generalize the result for the set Ω_γ for an arbitrary $\gamma < 1$ and reformulate the theorem in more general terms.

Theorem 3.1 *Suppose the Hamiltonian (13) written in the action variables or, equivalently, (15), written in the frequency variables, is analytic in* $\Omega_1 + (r, s, \kappa)$ *and* $\mu \in B^1_{\tilde{\mu}_0}$ *for some positive* $\tilde{\mu}_0$*, and the Assumption 1 holds.*

If in (15)

$$|\mathcal{F}|_{r,s,\kappa} \leq |\mu|$$

for all $\mu \in B^1_{\tilde{\mu}_0}$*, then given* $s_* < s$ *one can find some small constant* $r_0 = r_0(r, s - s_*, \tau, R, \kappa, \Lambda)$*, so that there exists a small positive* $\mu_0 = \mu_0(R, \Lambda, r_0, s - s_*, \kappa, \tilde{\mu}_0, \tau, \varepsilon)$*, such that for all* $\mu : \ |\mu| \leq \mu_0$ *there exists a symplectic near identity map* Ξ*, casting the Hamiltonian (13) into the form*

$$H \circ \Xi = H_+(I, J, \varepsilon, \mu),$$

which is defined on a Cantor set in the action space $\Psi_+^{-1}\Omega_1$*, where* $\Psi_+^{-1} : (\omega, \cdot) \to (I(\omega, J, \mu), \cdot)$ *is the inverse of the new frequency map* Ψ_+ *given as* $\Psi_+ : (I, \cdot) \to (\partial_I H_+, \cdot)$*.*

The map Ξ *is real analytic in* φ, p, q, μ *for* $\varphi \in W_{s_*}T^{n-1}$*,* $p, q \in B^2_{\frac{\kappa}{2}}$*,* $\mu \in B^1_{\mu_0}$ *and for real* J, μ *it is* C^∞ *in the sense of Whitney on a set* $\Psi_+^{-1}\Omega_1$*.*

The transformation Ξ *induces a transformation* $\Phi = \Psi \circ \Xi \circ \Psi_+^{-1}$ *in the frequency variables* $\Phi : \Omega_1 + (0, s_*, \kappa/2) \to \Omega_1 + (r, s, \kappa)$*, real analytic in variables* (φ, p, q) *and* C^∞ *- in the sense of Whitney - in the frequency variables, casting the Hamiltonian (15) on its new domain into the form*

$$\mathcal{H} \circ \Phi = \mathcal{H}_+(\omega, J, \varepsilon, \mu).$$

In particular, one can take

$$r_0 = \min\left(\frac{1}{2}(4\tau + 20)^{-(\tau+1)}\left(\frac{s - s_*}{3}\right)^{\tau+1}, \frac{1}{2}\Lambda^{\frac{\tau+1}{\tau}}, r\right) \tag{16}$$

and

$$\mu_0 \leq \min\left(\frac{2^{-2\tau-13} r_0^{\frac{2\tau+1}{\tau+1}}(s - s_*)}{27R^2}, \frac{2^{-2\tau-14} r_0^{\frac{\tau}{\tau+1}}\kappa_0^2}{27}, \frac{2^{-11}r_0^2}{9R^5}, \frac{\Lambda_0\kappa^2}{300}, \tilde{\mu}_0\right). \tag{17}$$

In particular, with such a choice of parameters the norm of the Hessian matrix $A_+ = \partial^2_{II}H_+$*, and the real part of the Lyapunov e xponent* $\lambda_+ = \partial_J H_+|_{J=0}$ *for Hamiltonian* H_+ *are bounded as in Assumption 1 but with* $1 + \frac{|\mu|}{\mu_0}R$ *instead of* R *and* $\Lambda(1 - \frac{|\mu|}{2\mu_0})$ *instead of* Λ*.*

Remark: This theorem as well as Theorem 2.2 and the following Theorem 3.2 can in fact be formulated in terms of existence of transformations Ψ_+^{-1} and Φ, as it was done in Chierchia and Gallavotti [1994], but we prefer a more traditional formulation, since the whole iterative procedure underlying the proof has the canonical nature.

Proof: The proof is inductive and consists in the repetitive application of the Main lemma. We construct Φ as a limit of successive transformations $\Phi = \Phi_0 \circ \Phi_1 \circ \Phi_2 \circ .. \circ \Phi_i \circ ....$

In order to be consistent we shall write s_0, R_0, Λ_0 for respectively s, R, Λ.

Take a sequence

$$\sigma_i = \frac{s_0 - s_*}{3} \cdot 2^{-i}, \; i = 1, 2, .., \tag{18}$$

$$s_i = s_0 - 3\sum_{j=1}^{i} \sigma_j, \; i = 1, 2, .., \tag{19}$$

and

$$\eta_i = \frac{\kappa_0}{6} \cdot 2^{-i}, \; i = 1, 2, .., \tag{20}$$

and

$$\kappa_i = \kappa_0 - 3\sum_{j=1}^{i} \eta_j, \; i = 1, 2, .., \tag{21}$$

and

$$r_i = K_r \sigma_{i+1}^{\tau+1} = r_0 2^{-i(\tau+1)}, \; i = 0, 1, 2, .. \tag{22}$$

with

$$r_0 = 2^{-\tau-1} K_r \left(\frac{s_0 - s_*}{3}\right)^{\tau+1}, \tag{23}$$

$$\rho_{i+1} = \frac{r_i}{6}, \; i = 0, 1, 2, .. \tag{24}$$

and

$$\mu_i = \mu_0 2^{-i(2\tau+7)} \; i = 0, 1, 2, ... \tag{25}$$

where K_r, μ_0 are some positive constants, to be determined.

All the above sequences vanish geometrically, their denominators being some powers of 2.

We will formally define four more sequences.

$$K_i = \left(\frac{1}{2r_i}\right)^{\frac{1}{\tau+1}}, \quad i = 0, 1, 2, .. \tag{26}$$

$$Y_i = \frac{1}{2K_i^{\tau}} = \frac{1}{2}(2r_i)^{\frac{\tau}{\tau+1}}, \quad i = 0, 1, 2, .. \tag{27}$$

Also we define

$$R_i = R_0 \cdot 2^i, \quad i = 0, 1, 2, .. \tag{28}$$

and

$$\Lambda_i = \Lambda_0 \cdot \left(\frac{3}{4}\right)^{-i}, \quad i = 0, 1, 2, .. \tag{29}$$

We want to show by induction that given $s_0, \kappa_0 \leq 1$ for some choice of the initial values of $r_0 < 1$, or equivalently parameter K_r, and μ_0 the above sequences will satisfy the conditions of the Main Lemma and we can apply it with parameters $r_i, s_i, \kappa_i, \mu_i, R_i, \Lambda_i, K_i$ and $\rho_{i+1}, \sigma_{i+1}, \eta_{i+1}$ for $i = 0, 1, ..$ to yield a transformation Ψ_i, and show that the size of the remainder will enjoy the estimate

$$|\mathcal{F}|_{r_{i+1}, s_{i+1}, \kappa_{i+1}} \leq \mu_{i+1}$$

Intuitively the above choice of the last four sequences is rather clear: the sequence (26) will be our choice of the ultraviolet cutoff parameter on each step to obey (41) in the Main Lemma; after each application of the Main Lemma the norm of the frequency map cannot exceed twice its original value and the bound on the real part of the Lyapunov exponent cannot decrease by more than one fourth of itself.
To satisfy for $i = 0$ the claim (42) of the Main Lemma on the Lyapunov exponent we will need

$$\Lambda_0 \geq 2Y_0 = (2r_0)^{\frac{\tau}{\tau+1}},$$

so we shall require

$$r_0 \leq \frac{1}{2}\Lambda_0^{\frac{\tau+1}{\tau}}, \tag{30}$$

with our choice of sequences (22), (29) it's clear that (30) ensures that the corresponding condition of the Main Lemma will be satisfied for all i.

Checking the condition (43) on the smallness of the perturbation μ_0, for all i with

$$\xi_i^2 = \min\left(\frac{r_i \sigma_{i+1}}{12R_i^2}, \frac{\kappa_i^2}{144}\right), \tag{31}$$

we see that it will be satisfied for all i if satisfied for $i = 0$, i.e.

$$\mu_0 \leq \frac{Y_0 \xi_0^2}{2} \leq \min\left(\frac{r_0^{\frac{2\tau+1}{\tau+1}}(s_0 - s_*)}{144R_0^2}, \frac{r_0^{\frac{\tau}{\tau+1}}\kappa_0^2}{576}\right). \tag{32}$$

Besides, we need $\mu_0 \leq \frac{\rho_1^2}{8R_0^5}$, to satisfy the condition on the frequency map in the Main Lemma, so regarding (32) we demand

$$\mu_0 \leq \min\left(\frac{r_0^{\frac{2\tau+1}{\tau+1}}(s_0 - s_*)}{144R_0^2}, \frac{r_0^2}{288R_0^5}, \frac{r_0^{\frac{\tau}{\tau+1}}\kappa_0^2}{576}\right). \tag{33}$$

By our choice of denominators for geometric series (22), (28), (25) the fact that (33) holds for $i=0$ will ensure that the analogous condition of the Main Lemma will be fulfilled for all i.

So we can apply the Main Lemma at least for the first time for $i=0$, and we can apply it as many times as we please if we prove that the size of the remainder is consistent with the choice of μ_i in (25). So, suppose we are doing the ith application, then the remainder on ith step will be given by

$$|\mathcal{F}_+|_{r_i-3\rho_{i+1},s_{i+1},\kappa_{i+1}} \le \left(1-\frac{2mu_i}{\xi_i^2 Y_i}\right)^{-1}\left(\frac{3\mu_i}{\xi_i^2 Y_i}+e^{-K_i\sigma_{i+1}}\right)\mu_i. \quad (34)$$

Evidently, $r_{i+1} \le r_i - 3\rho_{i+1} = \frac{r_i}{2}$. By the choice of K_i we have $K_i > 1$. Moreover, from (18), (22), (26) we have

$$K_i\sigma_{i+1} = (2K_r)^{-\frac{1}{\tau+1}}$$

and we may choose K_r so that $e^{-K\sigma_{i+1}} \le e^{-2\tau-10} \le 2^{-2\tau-10}$ by letting

$$K_r \le \frac{1}{2}(2\tau+10)^{-(\tau+1)}. \quad (35)$$

Suppose, the *second* term in the second bracket is larger than the first one.

Then (34) will imply:

$$|\mathcal{F}_+|_{r_{i+1},s_{i+1},\kappa_{i+1}} \le 8e^{-K\sigma_{i+1}}\mu_i \le 2^{-2\tau-7}\mu_i = \mu_{i+1}$$

Suppose, the *first* term dominates in the second bracket. Then using (18), (22), (25), (27), (28), (21), (31) we will have for all $i=0,1,2,..$:

$$\frac{\mu_i}{\xi_i^2 Y_i} \le \frac{\mu_0}{\xi_0^2 Y_0},$$

then (34) implies

$$|\mathcal{F}_+|_{r_{i+1},s_{i+1},\kappa_{i+1}} \le \frac{12\mu_i}{\xi_i^2 Y_i}. \quad (36)$$

That's why if we require

$$\min\left(\frac{12\cdot 288\mu_0}{\kappa_0^2 r_0^{\frac{\tau}{\tau+1}}}, \frac{12\cdot 36\cdot 4\mu_0 R_0^2}{r_0^{\frac{2\tau+1}{\tau+1}}(s_0-s_*)}\right) \le 2^{-2\tau-7},$$

where the left-hand side is just an estimate for (34) using (36), this will imply

$$|\mathcal{F}_+|_{r_{i+1},s_{i+1},\kappa_{i+1}} \le \mu_{i+1}.$$

So, to be consistent with our assumptions it's enough to ask

$$\mu_0 \le \min\left(\frac{2^{-2\tau-13} r_0^{\frac{2\tau+1}{\tau+1}}(s_0-s_*)}{27R_0^2}, \frac{2^{-2\tau-14}\kappa_0^2 r_0^{\frac{\tau}{\tau+1}}}{27}\right)$$

which is a stronger constraint compared to (33).

Adding together (30), (35) and recalling that

$$r_0 = K_r \left(\frac{s_0 - s_*}{3} \right)^{\tau+1} \cdot 2^{-\tau-1}$$

we choose

$$r_0 = \min \left(\frac{1}{2}(4\tau + 20)^{-(\tau+1)} \left(\frac{s_0 - s_*}{3} \right)^{\tau+1}, \frac{1}{2}\Lambda_0^{\frac{\tau+1}{\tau}}, r \right)$$

to establish one of the claims of this theorem.

At the end we will go back and compute how much the parameters R_i and Λ_i that one has after applying the Main Lemma i times really differ from their initial values R_0, Λ_0 for $i = 1, 2, ..$ (don't mix this up with (28) and (29) which were the formal sequences and virtually delivered the rough upper bounds for R_i and the lower bounds for Λ_i). We can assume inductively $R_i < 2R_0$ for all i. For $i = 1$ this is true, according to the Main Lemma. Then (see in the Proof of Main Lemma below the estimates (50), (51) for the norm of the Hessian matrix and its inverse) we will have the following estimate (with the inductive assumption that this is true for step $0, 1, .., i$:

$$R_{i+1} \leq \frac{R_i}{1 - \frac{4R_i^5 \mu_i}{\rho_{i+1}^2}} \leq R_0 \prod_{i=0}^{\infty} \left(1 + 4 \cdot 2 \cdot 32 \cdot 36 R_0^5 \frac{\mu_i}{r_i^2} \right)$$

$$\leq R_0 \exp \left(\frac{2^{10} \cdot 9 R^5 \mu_0}{r_0^2} \sum_{i=0}^{\infty} 32^{-i} \right),$$

where we've used that for $x \geq 0$ $e^x \geq 1 + x$. If we want the exponent to be not bigger than $\log 2 > \frac{16}{31}$ to prove our inductive assumption, then

$$\mu_0 \leq \frac{r_0^2}{2^{11} \cdot 9 R_0^5}, \tag{37}$$

and it is more strict than the second term in brackets in (33).

For the Lyapunov exponent the same inductive procedure can be used to prove that for any $i = 1, 2, ..$ $\Lambda_i \geq \frac{\Lambda_0}{2}$. Indeed, for $i = 1$ this is true by the Main Lemma. Besides, one evidently has

$$\Lambda_{i+1} \geq \Lambda_0 - \sum_{i=0}^{\infty} \frac{\mu_i}{\eta_{i+1}^2} \leq \Lambda_0 - \frac{144 \mu_0}{\kappa_0^2} \sum_{i=0}^{\infty} 32^{-i},$$

so $\Lambda_i \geq \frac{\Lambda_0}{2}$ if

$$\mu_0 \leq \frac{\Lambda_0 \kappa_0^2}{300}. \tag{38}$$

The limit $\Phi = \Phi_0 \circ \Phi_1 \circ \Phi_2 \circ ...$ exists on $\Omega_1 + (0, s_*, \kappa_0/2)$ (one can easily check that the above sequence of transformations is fundamental) and real analytic in (φ, p, q, μ). For $\mu = 0$ it becomes identity. It gives us a C^∞ - in the sense of Whitney - family of the surviving tori, corresponding to frequencies Ω_1, since each single transformation Φ_i is analytic, so C^∞, in the frequency variables on its domain. The fact that Φ is near identity simply follows from our choice of the series and the estimates on $\mathcal{P}$-norm of the Main lemma. Indeed, we just have to calculate the geometric sums $\sum_i \frac{\mu_i}{r_i}$, $\sum_i \frac{\mu_i}{\sigma_i}$, $\sum_i \frac{\mu_i}{\kappa_i}$, which by our choice of the sequences all converge not slower than 32^{-i}.

The last remark to make is that Φ has actually been constructed as a sequence of *symplectic* transformations in the action-angle variables: $\Phi_i = \Psi_i \Xi_i \Psi_{i+1}^{-1}$ (see the proof of the Main lemma). Then there exists a near identity transformation $\Xi = \Psi_0^{-1} \Phi \Psi_\infty$, where $\Psi_\infty = \lim_{i\to\infty} \Psi_i$, analytic in (φ, p, q, μ), which acts on some (generally complex) Cantor set $I_1' = \Psi_\infty^{-1}\Omega_1$, $\mu-$close to the image I_1 in the action variables of Ω_1 through the unperturbed inverse frequency map: $\Psi(I_1) \to \Omega_\gamma$, and depending on the perturbation and J. The map Ξ is C^∞ in I in the sense of Whitney for real J, μ, and the action of this map on I_1' is symplectic (since one can pass to a limit in the Poisson bracket operation, then the limit of a sequence of symplectic transformations is symplectic as well).

After all, we can use the Whitney extension theorem to extend the inverse of the frequency map Ψ_+^{-1} over an open domain in the frequency space, say $\Omega_1 + (r_0)$, since the estimates we have for its norm are formally valid on a Cantor set. Then the right-hand side in (17) should be multiplied by some constant, resulting from the Extension theorem, which will be close to 1 for small r_0 and which we shall omit. This "smoothening" makes differentiation of H_+ by I more palpable in the sense that it all can be expressed in terms of differentiation by ω through the frequency map Φ_+ and its inverse. $\square$.

Remark: The difference $s - s_*$ may be rather small, say, depend on ε, that becomes important, as pointed out by Jorba and Simo [1994], for rigorous treatment of homoclinic splitting with more than two degrees of freedom; in our proof it explicitly figures in the smallness condition for μ, namely $\mu_0 \sim (s - s_*)^{2\tau+2}$ and does not cause any extra difficulties.

Theorem 3.1 can be reformulated for a fixed $\gamma < 1$. The only changes one has to make is to include γ in the smallness conditions.

Theorem 3.2 *Suppose the Hamiltonian (13), written in the action variables, or, equivalently, (15) in the frequency variables is analytic in $\Omega_\gamma + (r, s, \kappa)$, and $\mu \in B^1_{\tilde{\mu}_0}$ for some positive $\tilde{\mu}_0$, and the Assumption 1 holds.*

Given $r, s, s_* < s, \kappa, R, \Lambda, \tau, \varepsilon$, one can find some $\mu_0(r, s - s_*, \kappa, R, \Lambda, \tau, \varepsilon) \leq \tilde{\mu}_0$, proportional to γ^2, such that if

$$|\mathcal{F}|_{r,s,\kappa} \leq \mu \leq \mu_0,$$

then there exists a near identity map Ξ

$$\Xi: \ (I', \varphi', p', q', \varepsilon, \mu) \to (I, \varphi, p, q, \varepsilon, \mu)$$

real analytic in (φ', p', q', μ), *acting trivially on* ε, μ, *that transforms the Hamiltonian function (13) into the form*

$$H \circ \Xi = H_+(I', J', \varepsilon, \mu),$$

with $J' = p'q'$, *thus defining a new frequency map* $\Psi_+ : \ (I', \cdot) \to (\omega' = \partial_{I'} H_+, \cdot)$, *such that its inverse* $\Psi_+^{-1} : (\omega', \cdot) \to (I'(\omega', J', \mu), \cdot)$ *is* C^∞ *in the frequency variables* $\omega' \in \Omega$ *for some open real domain* Ω, *such that* $\Omega_\gamma \subseteq \Omega \subseteq R^{n-1}$.

The induced transformation $\Phi = \Psi \circ \Xi \circ \Psi_+^{-1}$, *acts as:*

$$\Phi: \ \Omega + (0, s_*, \frac{\kappa}{2}) \to \Omega + (r, s, \kappa).$$

In particular,

$$\Omega_\gamma + (0, s_*, \frac{\kappa}{2}) \to \Omega_\gamma + (r, s, \kappa).$$

The transformation Φ *transforms the Hamiltonian function (15) written in the frequency variables into*

$$\mathcal{H} \circ \Phi = \mathcal{H}_+(\omega', J', \varepsilon, \mu).$$

It is real analytic in variables $(\varphi', p', q') \in W_{s_*} T^{n-1} \times B^2_{\frac{\kappa}{2}}$, $\mu \in B^1_{\mu_0}$ *and* C^∞ *in* $\omega' \in \Omega$.

The restriction of the transformation Ξ *on the set* $\Psi_+^{-1}\Omega_\gamma$ *is symplectic.*

Equivalently, the restriction of the transformation Φ *on the set* $\Omega_\gamma + (0, s_*, \kappa_*)$ *defines a* C^∞*-family of* $n-$*dimensional invariant surfaces, the equations of motion whereupon are simply:*

$$\omega' = const \in \Omega_\gamma, \ \varphi' = \varphi'_0 + \omega' t, \ p' = p'_0 e^{-\lambda' t}, \ q' = q'_0 e^{\lambda' t}, \tag{39}$$

where $\lambda' = \partial_{J'} H_+$. *These surfaces are* C^∞ $\mu-$*close to the unperturbed whiskered tori in the coordinates of Lemma 3.1, and may be parametrized in the initial coordinates as:*

$$\begin{aligned}
I &= I'(\omega', p'q', \varepsilon, \mu) + \Theta_1^1(I', \varphi', p', q', \varepsilon, \mu),\\
\varphi &= \varphi' + \Theta_2^0(I', \varphi', p', q', \varepsilon) + \Theta_2^1(I', \varphi', p', q', \varepsilon, \mu),\\
p &= \Theta_3^0(I', \varphi', p', q', \varepsilon) + \Theta_3^1(I', \varphi', p', q', \varepsilon, \mu),\\
q &= \Theta_4^0(I', \varphi', p', q', \varepsilon) + \Theta_4^1(I', \varphi', p', q', \varepsilon, \mu),
\end{aligned} \tag{40}$$

where $I'(\omega', J', \varepsilon, \mu) \in \Psi_+^{-1}\Omega_\gamma$, the functions $\Theta_i^0,\ i = 1,..,4$ are as in (12) of Lemma 3.1, and the Taylor series in μ of the functions $\Theta_i^1,\ i = 1,..,4$ begin with terms of the first order in μ.

Moreover $|\Theta_i^1 \circ \Psi_+^{-1}|_{\Omega,0,s_*,\kappa_*} \leq 2\mu,\ i = 1,..,4$. For the new Hamiltonian H_+ the norm of the frequency map Ψ_+ and its inverse is bounded by $(1+\frac{|\mu|}{\mu_0})R$ in the sup$-$norm, the real part of the new Lyapunov exponent $\lambda'|_{J'=0}$ is not smaller than $(1-\frac{|\mu|}{2\mu_0})\Lambda$.

Proof: The way we've included γ follows from the previous remarks. Indeed, for $0<\gamma<1$ if we scale $I\to\gamma I$, $p\to\gamma p$, $q\to\gamma q$ and multiply the Hamiltonian by γ^{-2}, which blows up Ω_γ into Ω_1. This modified Hamiltonian to which we apply Theorem 3.1 has the same non-degeneracy parameters as before rescaling, besides we will have to divide r and κ by γ, thus increasing them since $\gamma<1$. So, choice (16) will still remain the same and the right-hand side of (17) should be multiplied by γ^2.

Also, the only actual smallness parameter in the proof of Theorem 3.1 was r_0, or equivalently K_r, and it played the main role in the smallness condition for μ_0. Since we are free to take r_0 as small as we please, then according to the Theorem 3.1 for $\mu \leq \mu_0(r_0)$ we can construct a transformation Φ, acting on a Cantor set Ω_γ in the frequency space.

By construction the transformation Φ in Theorem 3.1 will be defined on the *closed set* $\Omega_\gamma \times \overline{W_{s_*}T^{n-1} \times B^2_{\frac{\kappa}{2}} \times B^1_{\mu_0}}$. So we can use the Extension theorem and consider a C^∞-extension $\mathcal{E}\Phi$ of Φ over $\Omega + (0, s_*, \frac{\kappa}{2})$ for some open real domain Ω in the frequency space, which contains the Cantor set Ω_γ: $\Omega_\gamma \subset \Omega \subseteq R^{n-1}$ this extension $\mathcal{E}\Phi$ will be analytic in the variables (φ', p', q') and C^∞ in ω'; if restricted to Ω_γ, it will give us the surviving tori.

Moreover, the induced transformation $\mathcal{E}\Xi = \Psi^{-1}\circ\mathcal{E}\Phi\circ\Psi_+$ will be defined on the set $\Phi_+^{-1}\Omega$, and for real J', μ will be C^∞ in the action variables $I' \in \Psi_+^{-1}\Omega$.

The only change in the smallness condition for μ_0 of Theorem (3.1) that the above extension will incur will be multiplying the right-hand side of (17) by some constant, coming from the Extension theorem. $\Box$.

3.2.1 The Main lemma

This part is purely technical. It contains the formulation and proof of the Main lemma which is used to establish Theorem 3.2.

Lemma 3.2 (The Main Lemma) *Suppose we are given positive parameters $r, s, \kappa < 1, \rho, \sigma, \eta$ such that $3\rho < r,\ 3\sigma < s,\ 3\eta < \kappa$. Let $\xi^2 \equiv \min(\frac{\rho\sigma}{2R^2}, \eta^2)$, $Y \equiv \frac{1}{2K^\tau}$. Suppose Assumption 1 holds, and also suppose,*

$$r \leq \frac{1}{2K^{\tau+1}}, \tag{41}$$

$$\Re\lambda \geq \frac{1}{K^{\tau}}. \tag{42}$$

If

$$|f|_{r,s,\kappa} \leq \min\left(\frac{Y\xi^2}{2}, \frac{\rho^2}{8R^5}\right), \tag{43}$$

then there exists a real analytic symplectic transformation Ξ *with range in the domain of the Hamiltonian (13), which casts this Hamiltonian into the form* $H \circ \Xi = h + \mathcal{D}\bar{f} + f_+ \equiv h_+ + f_+$.
The new frequency map is defined by $\Psi_+ : (I, \cdot) \to (\partial_I h_+, \cdot)$ *with the new Hessian matrix* $A_+ = \partial^2_{II} h_+$, *the new Lyapunov exponent as* $\lambda_+ = \partial_J h_+|_{J=0}$ *and they both satisfy Assumption 1 with constants* $2R, \frac{3}{4}\Lambda$. *The transformation* $\Phi = \Psi \circ \Xi \circ \Psi_+^{-1}$ *is real analytic and acts as follows:* $\Phi : \Omega_1 + (r - 3\rho, s - 3\sigma, \kappa - 3\eta) \to \Omega_1 + (r, s, \kappa)$. *It casts the Hamiltonian (15) into the form* $\mathcal{H} \circ \Phi = \mathcal{H}^0 + \mathcal{D}\bar{\mathcal{F}} + \mathcal{F}_+$ *and*

$$|\mathcal{F}_+|_{r-3\rho,s-3\sigma,\kappa-3\eta} \leq \left(1 - \frac{2|\mathcal{F}|_{r,s,\kappa}}{Y\xi^2}\right)^{-1}\left(\frac{3|\mathcal{F}|_{r,s,\kappa}}{Y\xi^2} + e^{-K\sigma}\right)|\mathcal{F}|_{r,s,\kappa}.$$

Moreover, $|W(\Phi - id)|_{\mathcal{P}} \leq \frac{Y}{\xi^2}|\mathcal{F}|_{r,s,\kappa}$ *uniformly on* $\Omega_1 + (r - 3\rho, s - 3\sigma, \kappa - 3\eta)$, *where* $W = diag(\rho^{-1}I_{n-1}, \sigma^{-1}I_{n-1}, \eta^{-1}I_2)$, I_l *denoting the* $l \times l$ *identity matrix,* $l = n - 1, 2$.

Proof. Let $\mathcal{S}$ be the image of $\Omega_1 + (r, s, \kappa)$ through the inverse Ψ^{-1} of the frequency map. We use the Lie transform method to seek the transformation from some subset $\mathcal{S}'$ into $\mathcal{S}$ which, being interpreted in the frequency variables, has the desired properties. Let X^t_ϕ be the Hamiltonian flow generated by the Hamiltonian ϕ. Ξ denotes the time-1 map obtained from X^1_ϕ. We consider the ultraviolet cutoff of the Fourier expansion of f with parameter K , then $f = T_K f + (f - T_K f)$. We denote $H_0 = h + T_K f$. From Taylor's formula we then have:

$$H_0 \circ \Xi = h + \{h, \phi\} + T_K f + \int_0^1 \{(1 - t)\{h, \phi\} + T_K f, \phi\} \circ X^t_\phi dt,$$

where as usual $\{\cdot, \cdot\}$ stands for the Poisson bracket of two smooth functions u, v of (I, φ, p, q)

$$\{u, v\} = (\partial_I u, \partial_\varphi v) - (\partial_I v, \partial_\varphi u) + \partial_p u \partial_q v - \partial_p v \partial_q u,$$

and $(\cdot, \cdot)$ denotes the standard inner product. We define $g \equiv \{h, \phi\} + T_K f$. To calculate the Poisson bracket we use the relations $\omega = \partial_I h$ and $\partial_p h = q\partial_J h = q\lambda$, in the same fashion $\partial_q h = p\lambda$. Then

$$g = -(\omega, \phi_\varphi) - \lambda(q\phi_q - p\phi_p) + T_K f. \tag{44}$$

Then the transformed Hamiltonian $H \circ \Xi = H_0 \circ \Xi + (f - T_K f) \circ \Xi$ will be given by

$$H \circ \Xi = h + g + f_+, \tag{45}$$

with

$$f_+ = \int_0^1 \{f_t, \phi\} \circ X_\phi^t dt + (f - T_K f) \circ X_\phi^1, \tag{46}$$

and

$$f_t \equiv (1-t)g + tT_K f.$$

We represent the perturbation as a Taylor series in the (p, q) variables near $(0, 0)$ and a Fourier series in the angle variables:

$$f = \sum_{i,j \geq 0,\ k \in Z^{n-1}} f_{ij}^k p^i q^j e^{\iota(k,\varphi)},$$

where ι stands for the imaginary unit.

A formal solution of (44) which puts the Hamiltonian in the desired form is given by:

$$\begin{aligned} g &= \mathcal{D}\bar{f}, \\ \phi &= \sum_{\substack{i,j \\ k \in Z_K^{n-1} \setminus \{0\}}} \frac{f_{ij}^k p^i q^j e^{\iota(k,\varphi)}}{\iota(k,\omega) + \lambda(j-i)} + \sum_{i \neq j} \frac{f_{ij}^0 p^i q^j}{\lambda(j-i)}, \end{aligned} \tag{47}$$

where g clearly depends only on I and the product $J = pq$ of the hyperbolic variables. Indeed, $\omega,\ \lambda$ are functions of $I,\ J$ only, so they are in the kernel of the Hamilton-Jacobi type operator $p\partial_p - q\partial_q$, that's why (47) formally satisfies (44).

In order to obtain the estimate on the norm of ϕ we need to estimate the small divisors in the sum given in (47). To do this we use the inequality for r given in the statement of the lemma and notice that for the case $i = j$ we have by our choice of r:

$$|(\omega, k)| \geq |(\omega^*, k)| - K|\omega - \omega^*| \geq K^{-\tau} - Kr \geq \frac{1}{2K^\tau},$$

where $\omega^* \in \Omega_1$ is the center of the ball. Otherwise, since ω is real on Ω_1, then $|\Im(\omega, k)| \leq Kr \leq \frac{1}{2K^\tau}$. For the case $i \neq j$ we use the assumption (42) on $\Re\lambda$ given in the statement of the lemma. Denoting $\delta = |f|_D = |\mathcal{F}|_{r,s,\kappa}$, we then easily obtain the following uniform estimates for ϕ and g:

$$|g|_D \leq \delta, \tag{48}$$

$$|\phi|_D \leq \frac{\delta}{Y}, \tag{49}$$

where $Y = \frac{1}{2K^\tau}$, as defined in the statement of the lemma.

Using Lemma 3 of Appendix A we obtain the following estimates for the derivatives of ϕ:

$$|\phi_I|^{\infty}_{D-(\bar{\rho},0,0)} \leq \frac{\delta}{Y\bar{\rho}}, \ |\phi_{p,q}|^{\infty}_{D-(0,0,\eta)} \leq \frac{\delta}{Y\eta}, \ |\phi_{\varphi}|^{\infty}_{D-(0,\sigma,0)} \leq |\phi|^{1}_{D-(0,\sigma,0)} \leq \frac{\delta}{Y\sigma},$$

where σ, η are as in the formulation of the lemma, and $\bar{\rho}$ is a parameter for denoting shrinking in the domain of the action variables, which we will define shortly. Here and further on we use the simple fact that the Absolute norm of a function is always not smaller than its infinity-norm.

Furthermore, if $\Xi : \ \mathcal{S}' \to \mathcal{S}$, then in the frequency variables we can define an induced transformation $\Phi = \Psi \circ \Xi \circ \Psi_+^{-1}$, where Ψ_+ is the new frequency map defined as $\Psi_+ \equiv \partial_I h_+$ by the new unperturbed Hamiltonian $h_+ = h + g$. Using Cauchy inequalities we get the following bounds on its Hessian matrix $A_+ = \partial^2_{II} h_+$:

$$|A_+ - A|^{\infty} \leq \frac{\delta}{\bar{\rho}^2}. \tag{50}$$

Using the equality

$$A^{-1} - A_+^{-1} = A_+^{-1}(A_+ - A)A^{-1},$$

from which by rearranging the terms and transiting to an arbitrary norm we derive

$$|A_+^{-1}|(1 - |A_+ - A|)|A^{-1}| \leq |A^{-1}|,$$

in particular, in the infinity-norm:

$$|A_+^{-1}| \leq \frac{|A^{-1}|}{1 - |A_+ - A||A^{-1}|} \leq \frac{R}{1 - \frac{\delta}{\bar{\rho}^2}R}, \tag{51}$$

provided that the denominator there is positive, what we will ensure by our choice of $\bar{\rho}$. Since one always has $R \geq 1$ fulfillment of (51) will guarantee (50), so to justify our claim for the norm of the new frequency map it's enough to ask

$$\delta \leq \frac{\bar{\rho}^2}{2R}.$$

So, if define $\bar{\rho} = \frac{\rho}{2R^2}$, we can always be sure that the induced loss of analyticity in the frequency space is not greater than ρ, and the above inequality will transfigure into

$$\delta \leq \frac{\rho^2}{8R^5},$$

which is part of (43).

Besides for the new Lyapunov exponent $\lambda_+ = \partial_J h_+|_{J=0}$ we will have

$$\Re|\lambda - \lambda_+| \leq \frac{\delta}{\eta^2} \leq \frac{\Lambda}{4},$$

what easily follows from (42) and (43). This proves the claim on the new Lyapunov exponent.

Now we want to restrict the domain of the Lie transform Hamiltonian ϕ in (49) to make sure that Φ maps $\Omega_1 + (r - 3\rho/2, s - 3\sigma/2, \kappa - 3\eta/2)$ into $\Omega_1 + (r - \rho, s - \sigma, \kappa - \eta)$, where the estimate (49) is valid. So we require:

$$\frac{2R^2\delta}{Y\rho} \le \sigma/2,\ \frac{\delta}{Y\eta} \le \eta/2,\ \frac{2R^2\delta}{Y\sigma} \le \rho/2.$$

Then if we define $\xi^2 \equiv \min(\frac{\rho\sigma}{2R^2}, \eta^2)$, these three requirements boil down to

$$\delta \le \frac{\xi^2 Y}{2} \tag{52}$$

To put it in other terms, if $W = diag(\rho^{-1}I_{n-1}, \sigma^{-1}I_{n-1}, \eta^{-1}I_2)$, I_l denoting the $l \times l$ identity matrix, $l = n-1, 2$, then $|W(\Phi - \mathrm{id})|_{\mathcal{P}} \le \frac{Y}{\xi^2}|f|_{r,s,\kappa}$ uniformly on $\Omega_1 + (r - 3\rho/2, s - 3\sigma/2, \kappa - 3\eta/2)$. This proves the claims for Φ.

It remains to estimate f_+. Evidently for $0 \le t \le 1$,

$$|f_t|_{r,s,\kappa} \le \delta.$$

Using this estimate along with (49), Lemma 4 of Appendix A gives

$$|\{f_t, \phi\}|_{r-\bar{\rho}, s-\sigma, \kappa-\eta} \le \frac{3}{\xi^2 Y}\delta^2.$$

Lemma 5 of Appendix A gives us a general estimate for the time-t map X_ϕ^t for $0 \le t \le 1$:

$$|u \circ X_\phi^t|_{r-2\rho, s-2\sigma, \kappa-2\eta} \le \left(1 - \frac{2}{\xi^2}|\phi|_{r,s,\kappa}\right)^{-1} |u|_{r,s,\kappa}.$$

We also have the estimate $|f - T_K f|_{r,s-\sigma,\kappa} \le e^{-K\sigma}|f|_{r,s,\kappa}$. Assembling all these pieces together we obtain the estimate:

$$|f_+|_{r-3\rho, s-3\sigma, \kappa-3\eta} \le \left(1 - \frac{|f|_{r,s,\kappa}}{\xi^2 Y}\right)^{-1} \left(\frac{3|f|_{r,s,\kappa}}{\xi^2 Y} + e^{-K\sigma}\right)|f|_{r,s,\kappa},$$

which will be the same for $\mathcal{F}_+ = f_+ \circ \Psi_+^{-1}$, as stated in the lemma. $\square$

Remark. This lemma and its estimates may be used in an obvious way when there is no hyperbolicity, simply by ignoring all that involves dependence of the Hamiltonian (13) on (p, q). Moreover, the proof can be carried out if the elliptic part of the system has more degrees of freedom than $n-1$. Indeed, suppose the Hamiltonian (13) depends on m more action-angle pairs, and Ω_1 is then a point set on the $n-1$-dimensional hyperplane in the $n-1+m$-dimensional frequency space, defined, for example, by m linearly independent resonant relations. Then g, being the average only in $n-1$ angles whose corresponding frequencies are known to be non-resonant at points of Ω_1, will also depend on these additional action-angle variables. The only difference is that in this case $\Omega_1 + (r)$ would mean a full-dimensional neighborhood of a lower dimensional Cantor set Ω_1.

4 Multiplicity-one resonant dynamics for the a-priori stable systems

In this section we will apply the theory of the a-priori unstable systems to the system (1). To introduce hyperbolicity we bring our system into a simple normal form near the resonant surface, and define what we call hyperbolicity in the first order. Then we apply the KAM theory of the a-priori unstable systems, developed in the previous section to this normal form and give the proof of Theorem 2.2.

4.1 A simple Normal form

Here we present a simple Normal Form valid near the resonant surface, which can be constructed under rather general assumptions. This is the transition step to introduce the hyperbolic variables thus reducing the problem to an a-priori unstable system. The normal form is obtained by a single application of the Main lemma of the previous section in a union of open balls of full dimension in the frequency space near Ω_1. At this point we come up with the first set of assumptions, necessary to go on with our discussion.

Assumption 2 (Non-degeneracy)

- *Suppose, for any $I \in D_{r_0}$ the infinity-norm of the Hessian matrix $Q = \partial^2_{II} H^0$ and its inverse are uniformly bounded by some positive constant R_0:*

$$|Q|^\infty, |Q^{-1}|^\infty \leq R_0. \tag{53}$$

- *Suppose, the resonant surface Π can be represented as a graph over variables $I^* \equiv (I_2, .., I_n)$, namely*

$$\Pi : \; I_1 = F(I^*),$$

where $F(I^*)$ is analytic in D_{r_0}.

For convenience we shall assume that $r_0, s_0 < 1$.

Consider a complex set $\Omega_\gamma + (r, s)$, where $r \leq \frac{r_0}{R_0}$, $s \leq s_0$. Then the Hamiltonian (1), considered as a function of the frequency-angle variables,

$$\mathcal{H} = H \circ \Psi^{-1} = \mathcal{H}^0(\omega) + \varepsilon \mathcal{H}^1(\omega, \varphi, \varepsilon), \tag{54}$$

is analytic in this set. Also suppose that uniformly $\varepsilon|\mathcal{H}^1|_{r,s} \leq \varepsilon \ll 1$.

As we mentioned before, we start by "blowing up" the frequency space, namely we rescale $I \to \gamma I$ and multiply the Hamiltonian by γ^{-2}. Then Ω_γ becomes Ω_1, and instead of r we will have $r\gamma^{-1}$.

Let's suppose that the perturbation H^1 in (1) is such that its zero-order coefficient in its Taylor series expansion in ε is not identically zero. In the above domain we can apply the version of the Main lemma proved above for some ultraviolet cutoff parameter K. Following the Remark after its proof, the averaging in the Main lemma must be made only throughout the *fast* angles $\varphi_2, .., \varphi_n$, and the part of the perturbation that is added to H^0 will be $\varepsilon T_K P_1 H^1(I, \varphi, \varepsilon)$, where P_1 stands for the projection of the Fourier series onto the first angle variable. Our purpose is to choose the values of parameters in the Main Lemma such that the remainder will be of the order $O(\varepsilon^p)$ for $1 < p < 2$. Recall the notation and constraints of the Main lemma. We can choose e.g. $\sigma = s/6$, $K = -\frac{6 \log \varepsilon}{s}$. To make sure that $K \geq 1$ it's enough to ask $\varepsilon \leq e^{-1}$. Besides, we require that ε is small enough to make sure that $r \geq \frac{1}{2} K^{-\tau-1}$ to obey (41), e.g.

$$\varepsilon \leq e^{-\frac{R_0}{12 r_0}},$$

since we've agreed that $r, s \leq 1$. Then we can redefine r as:

$$r = \frac{1}{2} \left(\frac{s}{6 \log \varepsilon^{-1}} \right)^{\tau+1}. \tag{55}$$

Hence with $\xi^2 = \frac{\sigma \rho}{2 R_0^2}$, $Y = \frac{1}{2K^\tau}$ the remainder $\mathcal{H}^1_+$ will be bounded as

$$|\mathcal{H}^1_+|_{r/2, s/2} \leq \left(1 - \frac{2\varepsilon}{Y \xi^2} \right)^{-1} \left(\frac{3\varepsilon}{Y \xi^2} + e^{-K\sigma} \right) \varepsilon. \tag{56}$$

If we express r, Y in terms of K, then to achieve our goal we must obey (43), for which it will be enough to ask $12 \cdot 8 \cdot 36 s^{-1} R_0^5 K^{2\tau+2} \varepsilon^{2-p} \leq 1$ what can certainly be achieved by taking ε small enough, e.g.

$$\varepsilon \leq B^{-1} \left(\frac{s}{6} \right)^{2\tau+2} \frac{1}{576 R_0^5} \frac{1}{\max_{0<x<1} x^{p-1} \log^{2\tau+2} x^{-1}} \tag{57}$$

for some $B > 1$.

With the above choice of ε the norm of the remainder is guaranteed to be smaller than $B^{-1}\varepsilon^p$. Needless to say that neither here, nor elsewhere, do we pursue the sharpest threshold smallness estimates for ε. So, in the new variables the Hamiltonian will be analytic in a smaller domain (namely in $\Omega_\gamma + (r/2, s/2)$) and in the action-angles it will look as follows:

$$H = H^0(I) + \varepsilon T_K P_1 H^1(I, \varphi_1, \varepsilon) + O(\varepsilon^p). \tag{58}$$

This Hamiltonian, truncated at the first order in ε, may be written as:

$$\tilde{h}(I,\varphi,\varepsilon) = H^0(I) + \varepsilon \tilde{H}^1(I,\varphi_1), \tag{59}$$

where $\tilde{H}^1 = P_1 H^1(I,\varphi,0)$. Here we've attempted to make the dependence of P in ε the simplest possible. By our choice of cut-off parameter K we can do it if ε is small, relegating the difference $T_K P_1 H^1(I,\varphi_1,\varepsilon) - P_1 H^1(I,\varphi,0)$ to the $O(\varepsilon^p)$ term.

Henceforth we'll assign $p = 3/2$, then if originally $H^1(I,\varphi,\varepsilon)$ is analytic in ε, then the transformed Hamiltonian (58) will be analytic in $\sqrt{\varepsilon}$.

At this point we'll change our notation. We will use symbols (p,q) instead of (I_1,φ_1), and (I,ω,φ) will stand for the rest of the actions, frequencies, and angles in an obvious way. Then we may rewrite (59) as:

$$\tilde{h}(I,p,q,\varepsilon) = H^0(I,p) + \varepsilon \tilde{H}^1(I,p,q), \tag{60}$$

With this notational change (59) describes the integrable system whose equations of motion are:

$$\begin{aligned} \dot{I} &= 0 \\ \dot{\varphi} &= \omega, \quad \omega \in \Omega_1 + (r/2) \\ \dot{p} &= -\varepsilon \tilde{H}^1_q(I,p,q) \\ \dot{q} &= H^0_p(I,p) + \varepsilon \tilde{H}^1_p(I,p,q). \end{aligned} \tag{61}$$

Besides, regarding (60) as the new "unperturbed Hamiltonian" we get the new frequency map $\omega = \tilde{h}_I$ which acts as an identity upon the rest of the variables, and the *sup*-norm of its Hessian matrix and its inverse will by construction be bounded by a constant $R = 2R_0$. We will call the perturbed Hamiltonian (1) *hyperbolic in the first order* if Equations (61) describe an a-priori unstable system. The following assumption formalizes this notion.

Assumption 3 (Hyperbolicity in the first order)

- *For $\varepsilon \neq 0$ suppose, for any I from the analyticity domain of the Hamiltonian (60) that the two last equations in (61) possess a hyperbolic equilibrium point (p^*,q^*), $p^* = p^*(I)$, $q^* = q^*(I)$, and the real part of the Lyapunov exponent $\lambda(I,\varepsilon)$*

  $$\lambda^2(I,\varepsilon) = (\partial^2_{pq}\tilde{h})^2 - \partial^2_{pp}\tilde{h}\partial^2_{qq}\tilde{h}\Big|_{p=p^*,q=q^*}$$

  at this point is bounded away from zero for all I and some positive constant λ_0 as $\lambda(I,\varepsilon) \leq \sqrt{\varepsilon}\lambda_0$.

In fact, this assumption requires a certain type of non-degeneracy from the perturbation H^1, namely that it shall remove the degeneracy on the resonant surface Π in the system of equations (61) for $\varepsilon = 0$. Also it becomes clear from the view of $\tilde{h}$ that we can't introduce hyperbolicity in the first order for complex ε in a ring near the origin, because of our claim on the real part of λ. Nevertheless one can see from (61) that this Assumption is naturally satisfied in plenty of examples. In particular the genericity of this Assumption becomes clear if one does the more traditional local analysis near a fixed point on a resonant surface in the action space, whose remaining frequencies are diophantine (see e.g Treschev[1991]), what can actually be done with (61) by fixing some I^* and Taylor expanding inside its small neighborhood.

In fact, from equations (61) we can easily derive that

$$p^*(I) = F(I) + O(\varepsilon),$$

where $F(I)$ is the same as in Assumption 2, and

$$q^*(I) = G(I) + O(\varepsilon),$$

where $G(I)$ is defined implicitly by

$$\partial_q \tilde{H}^1(I, F(I), G(I)) = 0.$$

Besides, the square of the Lyapunov exponent of the truncated "unperturbed" Hamiltonian (60) equals:

$$\lambda^2(I,\varepsilon) = \varepsilon^2 \partial^2_{pq} \tilde{H}^1(I, p^*, q^*) -$$

$$\varepsilon \partial^2_{qq} \tilde{H}^1(I, p^*, q^*) \cdot \partial^2_{pp} \left(H^0(I, p^*) + \varepsilon \tilde{H}^1(I, p^*, q^*) \right) = O(\varepsilon)$$

if in (1) the term $H^1(I, \varphi, \varepsilon)$ contains φ_1-dependence in its lower Fourier modes.

If we attribute all we've done with the original Hamiltonian (1) to the action of some "normalizing transformation" $\mathcal{R}_0$ which will certainly be symplectic and well-defined, then finally we arrive at Hamiltonian

$$H \circ \mathcal{R}_0 = H^0(I, p) + \varepsilon \tilde{H}^1(I, p, q) + \varepsilon^{\frac{3}{2}} B^{-1} \tilde{f}(I, \varphi, p, q, \sqrt{\varepsilon}), \tag{62}$$

which satisfies Assumption 2 with $R = 2R_0$ and Assumption 3, so represents an a-priori unstable system with $R = 2R_0$, $\Lambda = \sqrt{\varepsilon}\lambda_0$. From the way we've constructed the above Normal form, $|\tilde{f}| \leq 1$ in the Absolute norm on the domain $\mathcal{R}_0$.

Then without loss of generality (see Chierchia and Gallavotti [1994]) the equilibrium may be chosen as $(p^*, q^*) = (0, 0)$. Indeed, it may be brought into this form by a canonical transformation, defined by the generating function

$$(q - q^*(I'))p' + p^*(I')\sin(q - q^*(I')) + \varphi \cdot I' =$$

$$(q - G(I'))p' + F(I')\sin(q - G(I')) + \varphi \cdot I' + O(\varepsilon), \tag{63}$$

where the $O(\varepsilon)$ term does not contain φ-dependence, so the transformation $\mathcal{R}_1$ defined by (63) does not affect the actions.

Also by adjusting the (p,q)-independent part of it, we may rewrite (60) as

$$\tilde{h}(I,p,q,\varepsilon) = h_0(I,\varepsilon) + \tilde{P}(I,p,q,\varepsilon) \tag{64}$$

with the additional assumption that $\tilde{P}(I,0,0,\varepsilon) = 0$.

Finally, as for our estimate for the size of the remainder, we may introduce a new small parameter $\mu = B^{-1}\varepsilon^{\frac{3}{2}}$, so (62) may be rewritten as:

$$H(I,p,q,\sqrt{\varepsilon}) = h_0(I,\varepsilon) + \tilde{P}(I,p,q,\varepsilon) + \mu\tilde{f}(I,\varphi,p,q,\sqrt{\varepsilon}), \tag{65}$$

which describes an a-priori unstable system and will be analytic for $I \in D^* + (\frac{r}{2})$, $\varphi \in W_{\frac{s_0}{2}}T^{n-1}$, $p \in B^1_{\frac{r}{2}}$, $q \in B^1_{\frac{s_0}{2}}$ in notation of Section 2.1, where r is given by (55) with s_0 instead of s.

To finish with this section, we'll redefine $\mathcal{R}_0$ as $\mathcal{R}_1 \circ \mathcal{R}_0$ and call it a Normal form transformation. By construction, its impact on the action variables is near-identity.

4.2 Proof of the Main Theorem

At this point we have all the necessary building blocks to prove Theorem 2.2. To begin with, after applying the "simple Normal form" transformation $\mathcal{R}_0$, Assumption 3 has quantified the notion of hyperbolicity in the first order, and the entire previous subsection has prepared the grounds for using the KAM theory of the a-priori unstable systems.

Our first step will be to apply Lemma 3.1, which establishes the transformation $\mathcal{R}$ that casts the "unperturbed" Hamiltonian (64) into the Normal form

$$\tilde{h} \circ \mathcal{R} = h(I,pq,\sqrt{\varepsilon}) = h_0(I,\sqrt{\varepsilon}) + P(I,pq,\sqrt{\varepsilon}) \tag{66}$$

and the full Hamiltonian (65) into

$$H(I,p,q,\sqrt{\varepsilon}) = h(I,pq,\sqrt{\varepsilon}) + \mu f(I,\varphi,p,q,\sqrt{\varepsilon}). \tag{67}$$

Evidently, if ε is small enough then by construction of the above Normal form and by Lemma 3.1 this Hamiltonian will be analytic in some domain $G_{r',s',\kappa}$ for some $0 < r' < r, 0 < s' < \frac{s_0}{2}$, where s' does not depend on ε, and in particular we can take $\kappa = \Lambda = \lambda_0\sqrt{\varepsilon}$.

Besides for any $B > 1$ we can find ε small enough as in (57) such that

$$|f|_{r,s,\kappa} \leq B^{-1}\varepsilon^{\frac{3}{2}} = \mu.$$

We only have to check whether all the conditions of Theorem 3.1 will be satisfied with the choice of some large enogh value of a constant B. If we assign, for instance, $s_* = s'/2$ and assume $\tau \geq 2$ in the statement of Theorem 3.1, then the condition (17) boils down to choosing

$$B > B_1(r_0, s_0, R_0, \lambda_0, \tau),$$

where B_1 does not depend on ε, since in (17) (see also (16)) we will have at most $\mu_0 = B_1^{-1}\varepsilon^{\frac{3}{2}}$. That's why we always can find B large enough to exceed B_1; then (57) will give us the smallness condition on ε.

So, applying Theorem 3.1 and its generalization Theorem 3.2, we in fact arrive at the statement of the Main Theorem and the Corollary following it. Indeed, Theorems 3.1, 3.2 give us a symplectic transformation Ξ that establishes integrability of the corresponding a-priori unstable system on a Cantor set. Thus we can define $\Xi_* = \mathcal{R}_0 \circ \mathcal{R} \circ \Xi$ which transforms the Hamiltonian (1) in the way it is stated in the Main Theorem. Since Ξ differs from identity by $O(\varepsilon^{\frac{3}{2}})$ and $\mathcal{R}$ does not affect the action variables, then Ξ_* itself is near identity projected on the actions $I_2, .., I_n$.

The spurious element is that we've had to require $\tau \geq 2$ to be able to apply Theorem 3.1 (see (16), (17)) to the Normal form of Section 4.1. The reason for it was our choice of $p = \frac{3}{2}$ in Section 4.1. By choosing $\frac{1}{2} < p < 2$ we can extend our result for $\tau > 1$. To make it valid for all $\tau \geq 1$ which may be of interest, we should have applied the Main Lemma in Section 4.1 not once but twice, thus bringing the remainder down in size to $O(\varepsilon^2)$, which would be enough to apply Theorem 3.1. But this would not change almost anything as far as Assumption 3 or the system of equations (61) are concerned, since the added terms would be of higher order in ε. With this final remark we conclude the proof of the Main Theorem. $\Box$

Appendix A. The Absolute Norm.

We call the norm we use here the *absolute norm*. We introduce more notation for an arbitrary function $u(x, \varphi, p, q)$, analytic in the domain $C_{r,s,\kappa}$ for some open set $C \subseteq R^l$. As before $x \in C$ stands for either the action or the frequency variables, which are related through the globally invertible frequency map.

For $(x, p, q) \in C_{r,\kappa}$ we set

$$|u|_{x,p,q,s} \equiv \sum_{k \in Z^{n-1}} \overline{u_k(x, p, q)} e^{|k|s},$$

in the notation of Section 2.1,

$$|u|_{x,\kappa,s} \equiv \sup_{B^2_\kappa} |u|_{x,p,q,s}.$$

Lemma 1 (Lemma A1) *For the sup−norm* $|\cdot|^\infty_{r,s,\kappa}$ *on* $C_{r,s,\kappa}$ *for all* $\sigma > 0,\ \eta > 0$ *one has:*

$$|u|^\infty_{r,s,\kappa} \le |u|_{r,s,\kappa} \le \frac{(\kappa+\eta)^2}{\eta^2} \coth^{n-1}\sigma |u|^\infty_{r,s+2\sigma,\kappa+\eta}.$$

Proof. The first inequality is obvious. To prove the second one let's consider the absolute sum of the Taylor series of a function $u(\cdot,p,q)$ analytic in p,q in $B^2_{\kappa+\eta}$:

$$u(\cdot,p,q) = \sum_{p,q=0}^{\infty} u_{ij}(\cdot)p^i q^j$$

Clearly for a monomial $u_{ij}p^iq^j$ the coefficient equals:

$$u_{ij} = \frac{1}{i!j!}\frac{\partial^{i+j}}{\partial p^i \partial q^j}u(0,0) \le \frac{1}{(\kappa+\eta)^{i+j}}|u|^\infty_{r,s,\kappa+\eta}$$

here we skip the $(\cdot)$-dependence of u, and so

$$\overline{u(p,q)} \le |u|^\infty_{\kappa+\eta}\left(1+\frac{p}{\kappa+\eta}+\frac{p^2}{(\kappa+\eta)^2}+...\right)\cdot\left(1+\frac{q}{\kappa+\eta}+\frac{q^2}{(\kappa+\eta)^2}+...\right)$$

$$\le \frac{(\kappa+\eta)^2}{\eta^2}|u|^\infty_{\kappa+\eta}.$$

Then using the familiar estimate for the Fourier coefficients $|u_k(x,p,q)| \le e^{-|k|s}|u|^\infty_{r,s,\kappa}$ of the analytic functions and the identity $\sum_{k\in Z^{n-1}} e^{-2|k|\sigma} = \coth^{n-1}\sigma$, we come to the claimed result.□

Lemma 2 *For an arbitrary* $x \in C_r$ $|uv|_{x,s,\kappa} \le |u|_{x,s,\kappa}|v|_{x,s,\kappa}$ *for arbitrary functions* u, v*, analytic in* $C_{r,s,\kappa}$*.*

Proof. For any $(x,p,q) \in C_{r,s,\kappa}$ we have

$$|uv|_{x,p,q,s} = \sum_k \overline{(uv)_k(x,p,q)}e^{|k|s} \le \sum_k\sum_l \overline{u_{k-l}(x,p,q)}e^{|k-l|s}\overline{v_l(x,p,q)}e^{|l|s}$$

$$= |u|_{x,p,q,s}|v|_{x,p,q,s}.\square$$

Lemma 3 *For $0 < \rho < r,\ 0 < \sigma < s,\ 0 < \eta < \kappa$, and $(x,p,q) \in C_{r-\rho,\kappa-\eta}$*

$$|u_\varphi|^1_{x,p,q,s-\sigma} \equiv \sum_{i=1}^{n-1} |u_{\varphi_i}|_{x,p,q,s-\sigma} \leq \frac{1}{e\sigma}|u|_{r,s,\kappa}, \quad \max_{1\leq i\leq n-1} |u_{x_i}|_{x,p,q,s} \leq \frac{1}{\rho}|u|_{r,s,\kappa},$$

$$\max(|u_p|_{x,p,q,s}, |u_q|_{x,p,q,s}) \leq \frac{1}{\eta}|u|_{r,s,\kappa}.$$

Proof. For the $\varphi-$derivatives we have:

$$\sum_{i=1}^{n-1} |u_{\phi_i}|_{x,p,q,s-\sigma} = \sum_{i=1}^{n-1}\sum_k |k_i|\overline{u_k(x,p,q)}e^{|k|(s-\sigma)}$$

$$= \sum_k |k|\overline{u_k(x,p,q)}e^{|k|(s-\sigma)} \leq \sup_{t\geq 0} te^{-t\sigma} \cdot |u|_{x,p,q,s} = \frac{1}{e\sigma}|u|_{x,p,q,s}$$

The inequalities for the x,p,q-derivatives immediately follow from Cauchy formula. Here is how it follows for the x-derivatives, for the p,q-derivatives it's just the same argument.

If (x,p,q) is a point in $C_{r-\rho,\kappa-\eta}$, then:

$$u_{x_i} = \frac{1}{2\pi i}\int_\gamma \frac{u(x+\zeta,\cdot)}{\zeta}d\zeta,$$

where γ is a circle with radius ρ in the x-space. Then it follows that

$$|u_{x_i}|_{x,s,\kappa} \leq \frac{1}{2\pi}\int_\gamma \frac{|u|_{x+\zeta,s,\kappa-\eta}}{|\zeta|}|d\zeta| \leq \frac{1}{\rho}|u|_{r,s,\kappa}$$

uniformly for all $x \in C_{r-\rho}$ and all i. $\square$

Lemma 4 *For $0 < r-\rho < r_0,\ 0 < s-\sigma < s_0,\ 0 < \kappa-\eta < \kappa_0$ for u analytic in C_{r_0,s_0,κ_0} and v analytic in $C_{r,s,\kappa}$:*

$$|\{u,v\}|_{r-\rho,s-\sigma,\kappa-\eta} \leq$$

$$\left(\frac{1}{e}\left(\frac{1}{(r_0-r+\rho)\sigma} + \frac{1}{(s_0-s+\sigma)\rho}\right) + \frac{2}{(\kappa_0-\kappa+\eta)\eta}\right)|u|_{r_0,s_0,\kappa_0}|v|_{r,s,\kappa}.$$

Proof. Directly follows from the previous lemma. We just write for any $(x,p,q) \in C_{r-\rho,\kappa-\eta}$:

$$|\{u,v\}|_{r-\rho,s-\sigma,\kappa-\eta} \leq$$

$$\sum_{1\le i\le n-1} |u_{x_i}|_{x,p,q,s-\sigma} |v_{\varphi_i}|_{x,p,q,s-\sigma} + \sum_{1\le i\le n-1} |v_{x_i}|_{x,p,q,s-\sigma} |u_{\varphi_i}|_{x,p,q,s-\sigma} +$$

$$\frac{2}{(\kappa_0-\kappa+\eta)\eta} |u|_{x,p,q,s} |v|_{x,p,q,s} \le$$

$$\sum_{i=1}^{n-1} |u_{\varphi_i}|_{x,p,q,s-\sigma} \max_{1\le i\le n-1} |v_{x_i}|_{x,p,q,s} + \sum_{i=1}^{n-1} |v_{\varphi_i}|_{x,p,q,s-\sigma} \max_{1\le i\le n-1} |u_{x_i}|_{x,p,q,s} +$$

$$\frac{2}{(\kappa_0-\kappa+\eta)\eta} |u|_{x,p,q,s} |v|_{x,p,q,s} \le$$

$$\left(\frac{1}{e}\left(\frac{1}{(r_0-r+\rho)\sigma} + \frac{1}{(s_0-s+\sigma)\rho}\right) + \frac{2}{(\kappa_0-\kappa+\eta)\eta}\right) |u|_{r_0,s_0,\kappa_0} |v|_{r,s,\kappa}.$$

Here we've used the fact that u is analytic in a larger domain. $\square$

Lemma 5 *If $|\phi|_{r_0,s_0,\kappa_0} < \xi^2/2$, where ξ comes from the Main lemma, then*

$$|u\circ X_\phi^1|_{r-2\rho,s-2\sigma,\kappa-2\eta} \le \left(1 - \frac{2}{\xi^2}|\phi|_{r_0,s_0,\kappa_0}\right)^{-1} |u|_{r,s,\kappa}$$

for $0<\rho<r\le r_0-\rho,\ 0<\sigma<s\le s_0-\sigma,\ 0<\eta<\kappa\le\kappa_0-\eta$.

Proof. Consider the Lie series expansion:

$$u\circ X_\phi^1 = \sum_{h\ge 0} \frac{1}{h!} ad_\phi^h u,$$

where

$$ad_\phi^0 u = u,\ ad_\phi^h u = \{ad_\phi^{h-1} u, \phi\},\ h\ge 1$$

For $h\ge 1$ let's introduce $\tilde\rho = 2\rho/h,\ \tilde\sigma = 2\sigma/h,\ \tilde\eta = 2\eta/h,\ \tilde\xi^2 = 4\xi^2/h^2$. Let $|\cdot|_i \equiv |\cdot|_{r-i\tilde\rho,s-i\tilde\sigma,\kappa-i\tilde\eta}$ for $1\le i\le h$. Then we'll have:

$$|ad_\phi^i u|_i \le$$

$$\left(\frac{1}{e\tilde\rho(s_0-s+i\tilde\sigma)} + \frac{1}{e\tilde\sigma(r_0-r+i\tilde\rho)} + \frac{2}{\tilde\eta(\kappa_0-\kappa+i\tilde\eta)}\right) |\phi|_{r_0,s_0,\kappa_0} |ad_\phi^{i-1} u|_{i-1}$$

$$\leq \frac{3}{\tilde{\xi^2}} \frac{1}{h+i} |\phi|_{r_0,s_0,\kappa_0} |ad_\phi^{i-1} u|_{i-1}$$

So,

$$|ad_\phi^h u|_{r-2\rho,s-2_\sigma,\kappa-2\eta} \leq \left(\frac{3}{\tilde{\xi^2}}\right)^h \frac{h!}{(2h)!} |\phi|_{r_0,s_0,\kappa_0} |u|_{r,s,\kappa}.$$

Using a version of Stirling's formula: $(2h)! \leq (2h)^{2h} e^{-2h}$ we arrive at:

$$\left(\frac{3}{\tilde{\xi^2}}\right)^h \frac{1}{(2h)!} \leq \left(\frac{3h^2}{4\xi^2}\right)^h \frac{e^{2h}}{(2h)^{2h}} \leq \left(\frac{3e^2}{16\xi^2}\right)^h \leq \left(\frac{2}{\xi^2}\right)^h.$$

So, we end up with:

$$|u \circ X_\phi^1|_{r-2\rho,s-2\sigma,\kappa-2\eta} \leq \sum_{h\geq 0} \frac{1}{h!} |ad_\phi^h u|_{r-2\rho,s-2\sigma,\kappa-2\eta} \leq$$

$$\sum_{h\geq 0} \left(\frac{2}{\xi^2 |\phi|_{r_0,s_0,\kappa_0}}\right)^h |u|_{r,s,\kappa} = \left(1 - \frac{2}{\xi^2} |\phi|_{r_0,s_0,\kappa_0}\right)^{-1} |u|_{r,s,\kappa}. \square$$

References

Arnold, V.I. [1963] Proof of A. N. Kolmogorov's theorem on the preservation of quasiperodic motions under small perturbations of the Hamiltonian, *Russ. Math. Surveys*, **18**(5), 9-36.

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

Chierchia, L., Gallavotti, G. [1994] Drift and Diffusion in Phase Space. *Ann. de la Inst. H. Poincaré Phys. Th.*, **60**(91), 1-144.

Gallavotti, G [1994] Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. *Rev. Math. Phys.* **6**, 343-411.

Graff S. [1974] On the construction of hyperbolic invariant tori for Hamiltonian systems, J. Differential Equations **15** (1974), 1-69.

Jorba A., Simo C. [1994] On Quasiperiodic Perturbations of Elliptic Equilibrium Points [1994]. #94-310 Deposited in the archive mp_arc. To get a TeX version of it, send an empty e-mail message to **mp_arc@math.utexas.edu**.

de la Llave, R. and Wayne, C.E. [1990] Whiskered and low dimensional tori in nearly integrable Hamiltonian systems, University of Texas, Austin preprint.

Kolmogorov, A. N. [1954] On conservation of conditionally periodic motions under small perturbations of the Hamiltonian. *Dokl. Akad. Nauk. USSR*, **98**(4), 527-530.

Lochak, P. [1995] Arnold diffusion; a compendium of remarks and questions, Report.

Moser, J. [1956] The analytic invariants of an area-preserving mapping near a hyperbolic fixed point. *Commun. in Pure and Applied Math.* **9** (1956), 673-692.

Pöschel, J. [1982] Integrability of Hamiltonian Systems on Cantor Sets. *Comm. Pure App. Math.*, **35**, 653-695

Pöschel, J. [1993] Nekhoroshev estimates for quasi-convex Hamiltonian systems. *Math. Z.*, **213**, 187-216.

Treshchev, D.V. [1991] The mechanism of destruction of resonant tori of Hamiltonian systems, *Math. USSR Sb.*, **68**, 181.