Orbits of nearly integrable systems accumulating to KAM tori

M. Guardia, V. Kaloshin

July 29, 2014

Contents

1 Introduction 3
   1.1 The main result: time-periodic setting . . . . . . . . . . . . . . . . . . . 3
   1.2 Whitney KAM topology . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
   1.3 The quasi-ergodic hypotheses and Arnold diffusion . . . . . . . . . . .. . . 5

2 Outline of the proof of Theorem 2 6
   2.1 Net of Dirichlet resonances . . . . . . . . . . . . . . . . . . . . . . . . . . 9
   2.2 Deformation of the Hamiltonian and a tree of invariant cylinders . . . . .13
      2.2.1 Pöschel normal form and the molification of the Hamiltonian . . . .14
      2.2.2 Different regimes along resonances . . . . . . . . . . . . . . . . . . .15
      2.2.3 Normal form and invariant cylinders along single resonances . . . .17
      2.2.4 Normal form and invariant cylinders in the core of double resonances 19
   2.3 Localization of the Aubry sets . . . . . . . . . . . . . . . . . . . . . . . . .25
      2.3.1 Aubry sets in the transition zones . . . . . . . . . . . . . . . . . . .26
      2.3.2 Aubry sets in the core of the double resonances . . . . . . . . . . .26
   2.4 Shadowing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .27

3 First step of the proof: Selection of resonances 27
   3.1 Approximation of Diophantine vectors . . . . . . . . . . . . . . . . . . . .28
   3.2 Proof of Theorem 9 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .29
   3.3 Selection of resonances for all Diophantine frequencies: proof of Key Theorem 31

4 The deformation procedure 34
   4.1 The mollification of the Hamiltonian: proof of Lemma 2 . . . . . . . . . . . .34
   4.2 Perturbation along resonant segments . . . . . . . . . . . . . . . . . . . . .36
      4.2.1 The perturbation along the resonance of $S_0$ . . . . . . . . . . . . .36
      4.2.2 The perturbation along resonances of $S_n$ for $n > 1$ . . . . . . . . . .37
5 Different regimes and normal forms while approaching Diophantine frequencies

5.1 Covering the Dirichlet segments by neighborhoods of strong double resonances
5.2 Normal form in the transition zones
5.2.1 Proof of Theorem
5.3 Boumenoura Normal Form near in the core of the double resonances
5.3.1 Application of Corollary
5.3.2 Proof of Theorem
5.4 End of the proof of Key Theorem

6 The normally hyperbolic cylinders in the transition zones

6.1 Change to slow-fast variables
6.2 Existence of the cylinders
6.3 Proof of Theorem

7 Aubry sets in transition zones: proof of Key Theorem

7.1 Vertical estimates: proof of Theorem
7.2 Horizontal estimates: proof of Theorem
7.2.1 Localization for the slow angle: proof of Proposition
7.2.2 Localization for the slow action: proof of Proposition

8 Equivalent forcing classes and shadowing

9 Perturbation of single averaged potentials

9.1 Perturbation of families of functions on the circle
9.2 Condition on local minimum
9.3 Bifurcation of global minima
9.4 Deformation of single averaged potential along all Dirichlet resonant segments
9.5 Hölder norms and approximations
9.6 Smoothness of compositions and the inverse

A A time-periodic KAM

A.1 Regularity of straightening of the union of KAM tori: a simplified version
A.2 Anisotropic norms
A.3 Regularity of straightening of the union of KAM tori
A.4 Application of Pöschel theorem to our setting

B Generic Tonelli Hamiltonians and a Generalized Mapertuis Principle

B.1 A generalized Mapertuis principle
B.2 Proof of Lemma
1 Introduction

1.1 The main result: time-periodic setting

Let $U$ be a convex bounded open subset of $\mathbb{R}^n$. Consider a $C^2$ smooth strictly convex Hamiltonian $H_0(I)$, $I \in U$, i.e. for some $D > 1$ we have

$$D^{-1}\|v\| \leq \langle \partial^2 H_0(I)v, v \rangle \leq D\|v\| \text{ for any } I \in U \text{ and } v \in \mathbb{R}^n.$$  \hfill (1)

Fix $r \geq 2$ and consider the space of $C^r$-perturbations: $C^r(U \times \mathbb{T}^n \times \mathbb{T}) \ni H_1(I, \varphi, t)$. Denote the unit sphere with respect to the standard $C^r$ norm, given by maximum of all partial derivative of order up to $r$, by

$$S^r = \{H_1 \in C^r : \|H_1\|_{C^r} = 1\}.$$  

For a non-integer $r$ be use the standard Hölder norm (see Section 9.5).

In this paper we study dynamics of nearly integrable systems

$$H_\varepsilon(I, \varphi, t) = H_0(I) + \varepsilon H_1(I, \varphi, t).$$  \hfill (2)

Assume that for some $r \geq 2$,

$$\|H_0\|_{C^{r+3}} \leq 1, \|H_1\|_{C^r} \leq 1.$$  \hfill (3)

Consider small $\eta > 0$, $\tau > 0$. A vector $\omega \in \mathbb{R}^n$ is called $(\eta, \tau)$-Diophantine if

$$|\omega \cdot k + k_0| \geq \eta |(k, k_0)|^{-n-\tau} \text{ for each } (k, k_0) \in (\mathbb{Z}^n \setminus \{0\}) \times \mathbb{Z}.$$  

Denote by

$$\mathcal{D}_{\eta, \tau} = \{\omega : |\omega \cdot k + k_0| \geq \eta |(k, k_0)|^{-n-\tau}\}$$  \hfill (4)

this set of frequencies. Denote $U' = \nabla H_0(U)$ the set of values of the gradient of $H_0$. It is a bounded open set in $\mathbb{R}^n$. Let $U'_\eta \subset U'$ be the set of points whose $\eta$-neighborhoods belong to $U'$. Denote by $\text{Leb}$ the Lebesgue measure on $\mathbb{R}^n \times \mathbb{T}^n \times \mathbb{T}$ and $\mathcal{D}_{\eta, \tau}' := \mathcal{D}_{\eta, \tau} \cap U'_\eta$.

**Theorem 1.** (time-periodic KAM) Let $\eta, \tau > 0$, $r > 2n + 2\tau + 2$. Assume $\|H_0\|_{C^{r+3}} \leq 1$, $\|H_1\|_{C^r} \leq 1$, and $D_{\eta, \tau}'$. Then for each $\varepsilon > 0$ (1) $H_0(0, \eta, \tau) > 0$ and $c_0 = c_0(H_0, r, n) > 0$ such that for any $\varepsilon$ with $0 < \varepsilon < \varepsilon_0$ and any $\omega \in \mathcal{D}_{\eta, \tau}'$ the Hamiltonian $H_0 + \varepsilon H_1$ has a $(n+1)$-dimensional (KAM) invariant torus $\mathcal{T}_\omega$ and dynamics restricted to $\mathcal{T}_\omega$ is smoothly conjugate to the constant flow $(\dot{\varphi}, \dot{t}) = (\omega, 1)$ on $\mathbb{T}^{n+1}$. Moreover,

$$\text{Leb}(\cup_{\omega \in \mathcal{D}_{\eta, \tau} \cap U'_\eta} \mathcal{T}_\omega) > (1 - c_0\eta) \text{Leb}(U' \times \mathbb{T}^n \times \mathbb{T}).$$

This theorem was essentially proven by Pöschel [Pöss82]. We add the actual derivation in Appendix A. One can enlarge the set of KAM tori by lowering $\eta$ and improve the lower bound to $1 - c_0\sqrt{\varepsilon}$, but we do not rely on this improvement. In this paper we study the Arnold diffusion phenomenon for these systems in the case $n = 2$ and consider only the tori with frequencies in $\mathcal{D}_{\eta, \tau}$. Denote the set of all such KAM tori by

$$\text{KAM}_{\eta, \tau}^U := \cup_{\omega \in \mathcal{D}_{\eta, \tau}'} \mathcal{T}_\omega.$$  \hfill (5)
Theorem 2. *(The First Main Result)* Let $\eta, \tau > 0$. Then there is $r_0 > 0$ such that for any $r \geq r_0$ and any $C^{r+3}$ smooth strictly convex Hamiltonian $H_0$ there exists a $C^r$ dense set of perturbations $D \subset S^r$ such that for any $H_1 \in D$ there is $\varepsilon = \varepsilon(\eta, \tau, r, H_0, H_1) \in (0, \varepsilon_0)$ (where $\varepsilon_0$ is the constant defined in Theorem 1) with the property that the Hamiltonian $H_0 + \varepsilon H_1$ has an orbit $(I_\varepsilon, \varphi_\varepsilon)(t)$ accumulating to all KAM tori from $\text{KAM}_{\eta, \tau}^U$, i.e.

$$\text{KAM}_{\eta, \tau}^U \subset \bigcup_{t \in \mathbb{R}} (I_\varepsilon, \varphi_\varepsilon)(t).$$

An autonomous version of this result can be obtained using the standard energy reduction (see e.g. [Arn89, Sect 45]).

This result can be considered a weak form of the so-called quasi-ergodic hypothesis. The quasi-ergodic hypothesis was posed by Ehrenfest and Birkhoff and asserts that a typical Hamiltonian has a dense orbit in a typical energy surface. The result presented in this paper does not obtain full dense orbits but orbits dense in a set of large measure, and it only deals with nearly integrable systems. Note that the quasi-ergodic hypothesis is not always true. Herman showed a counterexample in $\mathbb{T}^2 \times [-\delta, \delta]^2$. He obtained open sets of Hamiltonian systems with a KAM persistent tori of codimension 1 in the energy surface (see [Yoc92]).

Theorem 2 is a considerable improvement of the result from [KZZ09], where we give a construction of a Hamiltonian of the form $\frac{1}{2} \sum_{j=1}^3 I_j^2 + \varepsilon H_1(I, \varphi)$ having an orbit accumulating to a positive set of KAM tori. An example of this form with an orbit accumulating to a fractal set of tori of maximal Hausdorff dimension is constructed in [KS12].

For the union of KAM tori $\text{KAM}_{\eta, \tau}^U$ one of the basic questions, studied in this paper, is the question of Lyapunov stability of these tori. In particular, we show that

*For a class of nearly integrable Hamiltonian systems all KAM tori in fixed diophatine class $\text{KAM}_{\eta, \tau}^U$ are Lyapunov unstable.*

Lyapunov stability of a KAM torus is closely related to a question of Lyapunov stability of a totally elliptic fixed point. An example of a 4-dimensional map with a Lyapunov unstable totally elliptic fixed point was constructed by P. Le Calvez and R. Douady [LCD83]. For resonant elliptic points of 4-dimensional maps instability was established in [KMV04]. A class of examples of nearly integrable Hamiltonians having an unstable KAM torus was recently obtained by J. Zhang and C. Q. Cheng [CZ13].

Douady [Dou88] proved that the stability or instability property of a totally elliptic point is a flat phenomenon for $C^\infty$ mappings. Namely, if a $C^\infty$ symplectic mapping $f_0$ satisfies certain nondegeneracy hypotheses, then there are two mapping $f$ and $g$ such that

- $f_0 - f$ and $f_0 - g$ are flat mappings at the origin and
- the origin is Lyapunov unstable for $f$ and Lyapunov stable for $g$.

This shows that Lyapunov stability is **not** an open property. Thus, the only chance to have robustness in the Main Theorem is to support perturbations of $H_1$ away from KAM tori. Moreover, the closer we approach to KAM tori the smaller is size of those perturbations. This naturally leads us to a Whitney topology relative to the union of KAM tori.
1.2 Whitney KAM topology

Denote $C^s(B^2 \times \mathbb{T}^3, \text{KAM}^U_{\eta, \tau})$ — the space of $C^s$ functions with the natural $C^s$-topology such that they tend to 0 as $(I, \varphi)$ approaches $\text{KAM}^U_{\eta, \tau}$ inside the complement.

Let $s$ be a positive integer. Let $M$ be one of $U \setminus D^U_{\eta, \tau} \subset \mathbb{R}^2$, $(U \setminus D^U_{\eta, \tau}) \times \mathbb{T}^3$, or $U \times \mathbb{T}^3 \setminus \text{KAM}^U_{\eta, \tau}$. If $f$ is a $C^s$ real valued function on $M$, the $C^s$-norm of $f$

$$\|f\|_{C^s} = \sup_{x \in M, |\alpha| \leq s} \|\partial^\alpha f(x)\|,$$

where the supremum is taken over the absolute values of all partial derivatives $\partial^\alpha$ of order $\leq s$. Definition of $C^s$-norm for non-integer $s$ is in section 9.5.

Introduce a strong $C^s$-topology. We endow it with the strong $C^s$-topology on the space of functions on a non-compact manifold or the $C^s$ Whitney topology. A base for this topology consists of sets of the following type. Let $\Xi = \{\varphi_i, U_i\}_{i \in \Lambda}$ be a locally finite set of charts on $M$, where $M$ is as above. Let $K = \{K_i\}_{i \in \Lambda}$ be a family of compact subsets of $M$, $K_i \subset U_i$. Let also $\varepsilon = \{\varepsilon_i\}_{i \in \Lambda}$ be a family of positive numbers. A strong basic neighborhood $N_s(f, \Xi, K, \varepsilon)$ is given by

$$\forall i \in \Lambda \quad \|(f \varphi_i)(x) - (g \varphi_i)(x)\|_s \leq \varepsilon_i \quad \forall x \in K_i.$$

The strong topology has all possible sets of this form. Let $\varepsilon_0$ be small positive. Endow $C^s(U \times \mathbb{T}^3, \text{KAM}^U_{\eta, \tau})$ with the strong topology.

**Theorem 3.** (The Second Main result) Let $\eta, \tau > 0$. Then there is $r_0 > 0$ such that for any $r \geq r_0$ and any $C^{r+3}$ smooth strictly convex Hamiltonian $H_0$ there exists a $C^r$ dense set of perturbations $D \subset S^r$ such that for any $H_1 \in D$ there is $\varepsilon = \varepsilon(\eta, \tau, H_0, H_1) > 0$ with the property that there is a set $W$ open in $C^r$-Whitney KAM topology such that for any $\Delta H_1 \in W$ the Hamiltonian $H_0 + \varepsilon(H_1 + \Delta H_1)$ has an orbit $(I_\varepsilon, \varphi_\varepsilon)(t)$ such that

$$\text{KAM}^U_{\eta, \tau} \subset \bigcup_{t \in \mathbb{R}} (I_\varepsilon, \varphi_\varepsilon)(t).$$

Theorem 3 certainly implies Theorem 2.

1.3 The quasi-ergodic hypotheses and Arnold diffusion

The quasi-ergodic hypothesis asks for the existence of a dense orbit in a typical energy surface of a typical Hamiltonian system. Instead Arnol’d diffusion asks whether a typical nearly integrable Hamiltonian has orbits whose actions make a drift with size independent of the perturbative parameter.

The study of Arnol’d diffusion was initiated by Arnol’d in his seminal paper [Arn64], where he obtained a concrete Hamiltonian system with an orbit undergoing a small drift in action. In the last decades there has been a huge progress in the area. The works of Arnold diffusion can be classified in two different groups, the ones which deal with *a priori unstable systems* and the ones dealing with *a priori stable systems* (as defined in [CG94].
The first ones are those whose first order present some hyperbolicity and have been studied both by geometric methods [DdlLS06, DdlLS08, DH09, Tre04, Tre12, DdlLS13] and variational methods [Ber08, CY04, CY09]. A priori stable systems are those who are close to an integrable Hamiltonian system, whose phase space is foliated by quasiperiodic tori. The existence of Arnold diffusion in a priori stable systems was only known in concrete examples [Dou88] until the recent works [Mat03, BKZ11, KZ12, Che12].

Acknowledgement: We would like to thank Ke Zhang, Georgi Popov, Jacques Féjoz, Kostya Khanin, Jinxin Xue, Yong Zheng, Jianlu Zhang, Hakan Eliasson, John Mather for useful conversations. The authors thank the Institute for Advanced Study for its hospitality, where a part of the work was done. The first author is partially supported by the Spanish MCyT/FEDER grant MTM2012-31714 and the Catalan SGR grant 2009SGR859. The second author acknowledges support of a NSF grant DMS-1157830.

2 Outline of the proof of Theorem 2

The proof of Theorem 2 has several parts which we describe in this section. The proofs of each part are given in Sections 3–8. We start by listing these parts. Recall that $U_{\eta} \subset \nabla H_0(U)$ is the set of points whose $\eta$-neighborhoods belong to $\nabla H_0(U)$ and $D_{\eta,\tau} := D_{\eta,\tau} \cap U'$. Fix $R_0 \gg 1$ (to be determined) and a sequence of radii $\{R_n\}_{n \in \mathbb{Z}^+}$, where $R_{n+1} = R_n^{1+2\tau}$ for each $n \in \mathbb{Z}_+$. Consider another “reciprocal” sequence $\{\rho_n\}_{n \in \mathbb{Z}^+}$, $\rho_n = R_n^{-3+5\tau}$ for each $n \in \mathbb{Z}_+$. For each $k \in (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$ denote $S(\Gamma_k) = \{\omega \in U' : k \cdot (\omega, 1) = 0\}$ the corresponding resonant segment.

1. A tree of Dirichlet resonant segments:
   - We construct a sequence of grids of Diophantine frequencies in $D_{\eta,\tau}^n \subset D_{\eta,\tau}^{U'}$ such that $3\rho_n$-neighborhood of $D_{\eta,\tau}^n$ contains $D_{\eta,\tau}^{U'}$ and $\rho_n$-neighborhoods of points of $D_{\eta,\tau}$ are pairwise disjoint.
   - To each grid $D_{\eta,\tau}^n$ we associate a collection of pairwise disjoint Voronoi cells (see Figure 1): open neighborhoods
     $$\{\text{Vor}_n(\omega_n)\}_{\omega_n \in D_{\eta,\tau}^n} \text{ such that } B_{\rho_n}(\omega_n) \subset \text{Vor}_n(\omega_n) \subset B_{3\rho_n}(\omega_n).$$
    We construct a collection of generations of resonant segments
    $$S = \bigcup_{n \in \mathbb{Z}_+} S_n, \quad S_n = \bigcup_{k_n \in J_n} \{S_{k_n}^{\omega_n}\},$$
    where $J_n \subset (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$ is a collection of resonances, $S_{k_n}^{\omega_n} \subset S(\Gamma_{k_n})$. We choose this collection such that most intersection between the different resonant segments are empty and if nonempty we have quantitative estimates.
   - By construction each segment of generation $n$ is contained in a Voronoi cell $\text{Vor}_n(\omega_n)$ for some $\omega_n \in D_{\eta,\tau}^n$, i.e. $S_{k_n}^{\omega_n} \subset \text{Vor}_n(\omega_n)$. We say that segments $S_{k_n}^{\omega_n}$ from $S_n$ belong to the generation $n$. 

6
We call a resonance segment from $S$ a **Dirichlet resonant segment**.

By construction Dirichlet segments accumulates to all the Diophantine frequencies in $D^{\eta,\tau}$ and satisfy several properties (see Key Theorem 1 for details).

These Dirichlet segments have a tree structure except that segments of the same generation can have **multiple intersections** (junctions)! See Section 3. Notice that this part is **number theoretic** and independent of the Hamiltonian $H_0$!

2. **Global (Pöschel) normal form and mollification**: In order to construct diffusion orbits consider the Hamiltonian $H_\varepsilon$ in the Pöschel normal form

$$N_\varepsilon = H_\varepsilon \circ \Phi_{\text{Pöschel}} = H'_0(I) + \varepsilon R(I, \varphi, t)$$

(see Theorem 3). Our main result does not have loss of derivatives. To compensate the loss of derivatives in the Pöschel normal form we mollify our Hamiltonian $H_\varepsilon$ in the complement to KAM tori $\text{KAM}^{U,\tau}_{\eta}$ to $H'_\varepsilon$ so that it is $C^\infty$ away from $\text{KAM}^{U,\tau}_{\eta}$ (see Section 4.1). Let $N'_\varepsilon := H'_\varepsilon \circ \Phi_{\text{Pöschel}}$.

It is convenient to identify resonances in the action space. To each Dirichlet segment $S_{k_n}^{\omega_n} \subset S_n$ with $k_n \in J_n$ we associated a Diophantine frequency $\omega_n \in D^{\eta,\tau}$. For each $\omega_n \in D^{\eta,\tau}$ denote $I_n := (\partial_t H'_0)^{-1}(\omega_n)$ and

$$\mathcal{I}_{k_n}^{\omega_n} := \{I \in U_\eta : \partial_t H'_0(I) \in S_{k_n}^{\omega_n}\}.$$ 

Similarly, we can define Voronoi cells in action space:

$$\text{Vor}_n(I_n) = \{I \in U_\eta : \partial_t H'_0(I) \in \text{Vor}_n(\omega_n)\}.$$ 

\[\text{1}\text{Certainly high norms of } H'_\varepsilon \text{ blow up to infinity as we approach } \text{KAM}^{U,\tau}_{\eta,\tau}\]
By construction we have $S^\omega_{k_n} \subset \text{Vor}_n(\omega_n)$. In the proof we always perturb away from KAM tori KAM$^U_{\eta,\tau}$ and this does not affect $H'_0$. It is convenient to describe perturbations as well as diffusing conditions in Pöschel coordinates. Notice that in Pöschel coordinates all KAM tori are flat, i.e.

$$\Phi_{\text{Pösch}}(\text{KAM}^U_{\eta,\tau}) = \{(I, \varphi, t) : \partial_t H'_0(I) \in \mathcal{D}^U_{\eta,\tau}\}.$$  

3. **Deformation of the Hamiltonian**: We modify the Hamiltonian $N'_\varepsilon$ to $N''_\varepsilon$ by a small $C^r$ perturbation supported away from KAM tori KAM$^U_{\eta,\tau}$ so that the resulting Hamiltonian $N''_\varepsilon$ is non-degenerate (in a way that we specify later). This perturbation is done in Section 4. Notice that this modification procedure is also done by induction:

$$N''_\varepsilon(I, \varphi, t) = N'_\varepsilon(I, \varphi, t) + \sum_{n \in \mathbb{N} \cup \{0\}} \Delta N''_\varepsilon(I, \varphi, t),$$

where

- the $0$-th generation perturbation $\Delta N''_\varepsilon$ is supported in a $O(\sqrt{\varepsilon})$-neighborhood of all horizontal and vertical resonant lines $S^\omega_k$ with $|k| \leq R_0$.
- the $n$-th generation perturbation $\Delta N''_\varepsilon$ is supported in $n$-th order Voronoi cells, i.e. $\Delta N''_\varepsilon(I, \varphi, t) = 0$ for any $I \not\in \text{Vor}_n(I_n)$ for any $I_n$ with $\partial_t H'_0(I_n) \in \mathcal{D}^n_{\eta,\tau}$.

The zero step is done as in [KZ12] and it allows us to construct a net of normally hyperbolic invariant cylinders along the resonances in $S_0$.

The $n$-th step we modify our Hamiltonian in Voronoi cells of order $n$ close to Dirichlet resonant segments of order $n$ so that it satisfies some non-degeneracy conditions, while non-degeneracy conditions of orders $k < n$ from previous steps hold true.

4. **Resonant Normal Forms**: Fix a generation $n$ and a Dirichlet segment $I^\omega_{k_n}$ of this generation. In Section 5, we derive Resonant Normal Forms for $N''_\varepsilon$ for points $(I, \varphi, t)$ such that $I$ is close to $I^\omega_{k_n}$. We obtain normal forms of two different types:

- A single resonant one

$$N''_{\varepsilon,k_n} \circ \Phi^{k_n}(\psi, J, t) = \mathcal{H}^{k_n}_0(J) + \mathcal{Z}^{k_n}(\psi^s, J) + \mathcal{R}^{k_n}(\psi^s, \psi^f, J, t),$$

where $\mathcal{Z}^{k_n}$ only depends on the slow angle $\psi^s = k_n \cdot (\psi, t)$, $(\psi^s, \psi^f, t)$ form a basis and $\mathcal{R}^{k_n}$ is small relatively to non-degeneracy of $\mathcal{Z}^{k_n}$. This normal form is similar to [BKZ11].

- A double resonant one

$$N''_{\varepsilon} \circ \Phi^{k_n,k'_n}(\psi, J, t) = \mathcal{H}^{k_n,k'_n}_0(J) + \mathcal{Z}^{k_n,k'_n}(\psi, J)$$

where $\mathcal{Z}^{k_n,k'_n}$ only depends on the two slow angles $\psi^s_1 = k_n \cdot (\psi, t)$, $\psi^s_2 = k'_n \cdot (\psi, t)$. We follow the normal form procedure explained in [Bou10].
(core of the double resonance) Thanks to the deformation procedure we can completely remove the time dependence in the normal form in the core of double resonances.

(transition zones to a single resonance) In [KZ12] the double resonant normal form Hamiltonian is a small perturbation of a two degrees of freedom mechanical system. Here we cannot make this reduction as a “non-mechanical” remainder becomes large. To resolve the issue we describe the leading order by a more general two degrees of freedom system. To study such systems we apply a generalized Maupertuis principle (see Appendix B) and use the technique developed in Appendix A, [KZ12].

5. **A tree of normally hyperbolic invariant cylinders**: In Section 6, using the non-degeneracy obtained in the Deformation step and Resonant Normal Forms from the previous step, we build a tree of normally hyperbolic invariant cylinders along all Dirichlet resonant segments.

In the single resonant regime, we obtain cylinders along the resonant segments $\mathcal{C}_{k_n}^{\omega_n}$ (see Key Theorem 3). We denote them by $\mathcal{C}_{k_n}^{\omega_n}$. In the double resonance regime we obtain several cylinders by analogy with [KZ12]. This is explained in Key Theorems 5 and 6 and Corollary 2.

6. **Localization of Aubry sets**: We prove the existence of certain Aubry sets localized inside the cylinders obtained in the previous step. We also establish a graph property, which essentially says that these Aubry sets have the same structure as Aubry-Mather sets of twist maps. This is done in Section 7.

7. **Shadowing**: In order to construct diffusing orbits along a chain of Aubry sets we use the notion of c-equivalence proposed by Bernard. It essentially consists of three parts: apriori unstable (diffusion along one cylinder), bifurcations (jump from one cylinder to another in the same homology), a turn jump from a cylinder with one homology to a cylinder with a different one. The first two regimes are essentially done in [BKZ11] using a lemma from [CY04, CY09]. The last one is similar to [KZ12].

Now we describe each step and split the proof of Theorem 2 in several Key Theorems.

### 2.1 Net of Dirichlet resonances

The first step is to construct a tree of generations of special resonant segments

$$S = \{S_n\}_{n \in \mathbb{Z}}, \quad S_n = \{\mathcal{C}_{k_n}^{\omega_n}\}_{k_n \in \mathcal{F}_n}$$

in the frequency space $U' \subset \mathbb{R}^2$ such that the following properties hold
each segment $S_{k_n}^{\omega_n}$ belongs to the intersection of the resonant segment and the corresponding Voronoi cell $S_{k_n}^{\omega_n} \subset S(\Gamma_{k_n}) \cap \text{Vor}_{n-1}(\omega_{n-1})$ i.e. there are $k_n \in (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$ and $\omega_{n-1} \in D_{\eta,\tau}^{n-1}$ such that this inclusion holds.

Any segment $S_{k_n}^{\omega_n}$ of generation $n$ intersects only segments of the generation $n-2$ and $n-1$ and the segments of generation $n+1$ and $n+2$ associated to frequencies $D_{\eta,\tau}^{n+1} \cap \text{Vor}_{n-1}(\omega_{n-1})$ and $D_{\eta,\tau}^{n+2} \cap \text{Vor}_{n-1}(\omega_{n-1})$ respectively. However, it does not intersect any segment of any previous generation generation $n+k$ or $n-k$, $k \geq 3$.

We have quantitative information about properties of intersections of segments of generation $n$ inside of the same Voronoi cell. The union $\bigcup_{n} \bigcup_{k_n \in F_n} S_{k_n}^{\omega_n}$ is connected. The closure $\overline{\bigcup_{n} \bigcup_{k_n \in F_n} S_{k_n}^{\omega_n}}$ contains all Diophantine frequencies in $D_{\eta,\tau}^{U}$ (see (4)).

We shall call these segments Dirichlet resonant segments. We define them to satisfy quantitative estimates on speed of approximation. Recall that an elementary pigeon hole principle show that for any bounded set $U \subset \mathbb{R}^2$ and any any $\omega = (\omega_1, \omega_2) \in U$ there is a sequence $k_n(\omega) \in (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$ such that

$$|k_n(\omega) \cdot (\omega, 1)| \leq |k_n(\omega)|^{-2} \quad \text{and} \quad |k_n(\omega)| \to \infty \text{ as } n \to \infty.$$  

We choose our segments so that

- for each $S_{k_n}^{\omega_n} \subset S(\Gamma_{k_n})$ there is a Diophantine number $\omega_n \in D_{\eta,\tau}^{n}$ such that

  $$|k_n \cdot (\omega, 1)| \leq |k_n|^{-2+3\tau}.$$  

- if $S_{k_n}^{\omega_n} \cap S_{k_n'}^{\omega_n'} \neq \emptyset$, then we have an upper bound on ratio $|k_n| / |k_n'|$ as well as a uniform lower angle of intersection between $\Gamma_{k_n}$ and $\Gamma_{k_n'}$.

The tree of Dirichlet resonances is obtained in two steps.

- in Section 3.3 we fix a Diophantine $\omega^* \in D_{\eta,\tau}^{U}$ and construct a connected “zigzag” approaching $\omega^*$.

- in Section 3.3 we define Dirichlet resonant segments and show how to deal with all frequencies in $D_{\eta,\tau}^{U}$ simultaneously.

It turns out that we cannot just construct the “zigzag” for each frequency and then consider the union over all frequencies in $D_{\eta,\tau}^{U}$, because we need estimates on the complexity of the intersection among resonant segments. Now we define the sequence of discrete sets $\{D_{\eta,\tau}^{n}\}_{n \in \mathbb{Z}^+}$.

---

2 Too many segments, too many intersections to control
Recall that $R_0 \gg 1$ (to be determined) and we denote

$$R_{n+1} = R_n^{1+2\tau}, \quad \rho_n = R_n^{-(3-5\tau)}, \quad \forall n \in \mathbb{Z}_+$$

(6)

and $B_r(\omega)$ be an $r$-ball centered at $\omega$. Note that the sequence $\rho_n$ also satisfies $\rho_{n+1} = \rho_n^{1+2\tau}$. Define a sequence of discrete sets $\mathcal{D}_n = \mathcal{D}_{\eta,\tau}^n \subset \mathcal{D}_{\eta,\tau}^U$ with $n \geq 1$.

By the Vitali covering lemma from any finite cover of $\mathcal{D}_{\eta,\tau}^U$ by $\rho_n$-balls, there is a subcover $\mathcal{B}_n$ having the following two properties:

1. Replacing $\rho_n$-balls of subcover $\mathcal{B}_n$ by cocentric $3\rho_n$-balls we cover $\mathcal{D}_{\eta,\tau}^U$.
2. The $\rho_n$-balls of $\mathcal{B}_n$ are pairwise disjoint.

Moreover, by the Besicovitch covering theorem from the cover by balls of radius $3\rho_n$, one can choose a finite subcover such that each point is covered by at most $\kappa$ balls for some $\kappa$ which depends only on dimension. In [Sul94] it is shown that in the 2-dimensional case $\kappa = 19$. Denote the set of the centers of the balls of this cover by $\mathcal{D}_{\eta,\tau}^\kappa$. For each point $\omega_n \in \mathcal{D}_{\eta,\tau}^n$ consider all neighbors, i.e. points $\leq 6\rho_n$ away from $\omega_n$. Define the set of points that are closer to $\omega_n$ than to any other neighbor and at most $3\rho_n$ away from $\omega_n$. Denote this set by $\text{Vor}_n(\omega_n)$ and, following the standard terminology, call it Voronoi cells. The properties of the coverings by balls and the Voronoi cells, are summarized in the following

**Lemma 1.** Consider the set of frequencies $U'_\eta \subset U'$. Then, there exists a sequence of discrete sets $\mathcal{D}^n_{\eta,\tau} \subset \mathcal{D}^U_{\eta,\tau} \subset U'_\eta$, $n \geq 1$, satisfying

$$\mathcal{D}_{\eta,\tau} = \bigcup_{n \geq 1} \mathcal{D}^n_{\eta,\tau}$$

such that for the Voronoi cell $\text{Vor}_n(\omega^*) \subset U'$ associated to each frequency $\omega^*$ of the set $\mathcal{D}^n_{\eta,\tau}$ we have

$$\omega^* \in B_{\rho_n}(\omega^*) \subset \text{Vor}_n(\omega^*) \subset B_{3\rho_n}(\omega^*).$$

Moreover, each $\omega \in \mathcal{D}^U_{\eta,\tau}$ belongs to at least one and at most 19 balls $B_{3\rho_n}(\omega^*)$, and

$$B_{\rho_n}(\omega^*) \cap B_{\rho_n}(\omega') = \emptyset$$

for any $\omega^* \neq \omega' \in \mathcal{D}^n_{\eta,\tau}$.

We use this sequence of discrete sets $\mathcal{D}^n_{\eta,\tau}$ to construct a tree of resonances. The key theorem of this first step is the following.

**Key Theorem 1.** Consider the sequence of sets $\{\mathcal{D}^n_{\eta,\tau}\}_{n \leq 1}$ and the associated Voronoi cells $\text{Vor}_n(\omega^*_n)$, $\omega^*_n \in \mathcal{D}^n_{\eta,\tau}$. Then, there are constants $c_1, c_2 > 0$, $R^* \gg 1$ and $0 < \tau^* \ll 1$, such that, for any $R_0 \in [R^*, +\infty)$ and $\tau \in (0, \tau^*)$ and taking $R_{n+1} = R_n^{1+2\tau}$, $\rho_n = R_n^{-(3-5\tau)}$, for each $n \in \mathbb{Z}_+$ and $\omega^*_n \in \mathcal{D}^n_{\eta,\tau}$, there exists $k^*_n \in (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$ which satisfies the following properties:

1. The vectors belong to the following annulus $\frac{R_n}{4} \leq |k^*_n| \leq R_n$.  

2. The vectors satisfy
\[ \eta R_n^{-(2+\tau)} \leq |k_n^* \cdot (\omega^*, 1)| \leq R_n^{-(2-3\tau)} \]

3. The associated resonant segment \( S(\Gamma_{k_n^*}) \) satisfies \( S(\Gamma_{k_n^*}) \cap B_{\rho_n}(\omega_n^*) \neq \emptyset \).

4. Let \( \omega_{n-1}^* \in D_{n-1}^{n-1} \) be a frequency in the previous generation so that \( \omega_{n-1}^* \in \text{Vor}_{n-1}(\omega_{n-1}^*) \) and \( k_{n-1}^* \) the associated resonant vector, i.e. \( S(\Gamma_{k_{n-1}^*}) \cap \text{Vor}_{n-1}(\omega_{n-1}^*) \neq \emptyset \). Then, the angle \( k_n^* \parallel k_{n-1}^* \) satisfies \( \angle(k_n^*, k_{n-1}^*) \geq \frac{\pi}{10} \).

5. The resonances \( S(\Gamma_{k_n^*}) \) and \( S(\Gamma_{k_{n-1}^*}) \) intersect in \( \text{Vor}_{n-1}(\omega_{n-1}^*) \) at only one point.

6. Let \( S_{k_n}^{\omega_n^*} \subset S(\Gamma_{k_n^*}) \) be the minimal segment, which contains the above intersection and also all the intersections of \( S(\Gamma_{k_n^*}) \) with the resonant lines associated to all frequencies in \( D_{n-1}^{n-1} \cap \text{Vor}_{n}(\omega_n^*) \). Then,

\[ S_{k_n}^{\omega_n^*} \subset B_{\rho_{n-1}}(\omega_{n-1}^*). \]

7. Any other resonant line \( S(\Gamma_k) \) intersecting \( B_{3\rho_{n}}(\omega_n^*) \), satisfies \( \frac{|k_n|}{|k|} \leq R_n^{c_1\tau} \).

8. Let \( F_n \subset (\mathbb{Z}^2 \setminus \{0\}) \times \mathbb{Z} \) be the set of integer vector \( k_n \)'s selected above. Let \( \omega' \) be an intersection point of \( S_{k_n}^{\omega_n} \) and \( S_{k_n'}^{\omega_n} \) for some \( \omega_n, \omega_n' \in D_{n-1}^{n-1} \) and \( k_n, k_n' \in F_n \). Then, there are at most to \( \rho_{n}^{c_2\tau} \) resonant lines \( S_{k_n}^{\omega_n} \) from the \( n \) generation passing through \( \omega' \).
This theorem is proved in Section 3.

As a consequence of this theorem we obtain a set of Dirichlet resonant vectors and a tree of resonances in the frequency space. Nevertheless the tree of resonances formed by the Dirichlet resonant segments, obtained in Theorem 1, is not connected since the segments associated to the first generation belong to different Voronoi cells and therefore they do not intersect. To connect them, we consider a 0-generation.

**Definition 4.** Fix \( R_0 \gg 1 \) and \( \rho_0 = R_0^{-3-5\tau} \). Define sets \( \mathcal{F}_n \subset (\mathbb{Z}^2 \setminus \{0\}) \times \mathbb{Z} \) as follows.

\[
\mathcal{F}_0 = \{ k \in (\mathbb{Z}^2 \setminus \{0\}) \times \mathbb{Z} : k \cdot (1,0) = 0 \text{ or } k \cdot (0,1) = 0, |k| \leq R_0 \}
\]

and its associated net of resonances in frequency space \( \mathcal{S}_0 \subset U_\eta' \subset \mathbb{R}^2 \) as

\[
\mathcal{S}_0 = \bigcup_{k \in \mathcal{F}_0} \{ \omega \in U_\eta' \subset \mathbb{R}^2 : (\omega,1) \cdot k = 0 \}.
\]

For \( n \geq 1 \) we define the set \( \mathcal{F}_n \) as the set of Dirichlet resonant vectors \( k \in (\mathbb{Z}^2 \setminus \{0\}) \times \mathbb{Z} \) obtained in Theorem 1 and

\[
\mathcal{S}_n = \bigcup_{k \in \mathcal{F}_n} \mathcal{S}_{k_n}^{\omega_n}, \quad \mathcal{S} = \bigcup_{n \geq 0} \mathcal{S}_n,
\]

where \( \mathcal{S}_{k_n}^{\omega_n} \) are the resonant segments obtained in Key Theorem 1.

By the definition of \( \mathcal{S}_0 \) and Theorem 1 we have constructed a net of resonances \( \mathcal{S} \) which satisfies the needed conditions. They are summarized in the next corollary of Theorem 1.

**Corollary 1.** The resonance net \( \mathcal{S} \subset U_\eta' \subset \mathbb{R}^2 \) is connected, \( \mathcal{D}_{\eta,\tau}^U \subset \mathcal{S} \). Moreover, each \( \mathcal{S}_n = \{ \mathcal{S}(\Gamma_k) : k \in \mathcal{F}_n \} \) satisfies all the properties of Key Theorem 1.

### 2.2 Deformation of the Hamiltonian and a tree of invariant cylinders

Step 2 of the proof is to deform the Hamiltonian. The deformation has two reasons.

- Away from KAM tori recover the loss of regularity due to normal form procedures.
- Obtain a properly non-degenerate Hamiltonian.

Using this non-degeneracies we prove the existence of the tree of invariant cylinders in Step 3. In this section we explain both Steps 2 and 3.

Fix small \( \eta, \tau > 0 \). Denote \( a_\tau := a + \mathcal{O}(\tau) \), where \( |\mathcal{O}(\tau)| < C\tau \) for some \( C > 0 \) independent of \( \varepsilon, \) all \( \rho_n \) and \( n \). Since \( \tau \) is small, the notation is used if a multiple in front of \( \tau \) is not important.
Consider the nearly integrable Hamiltonian (2). Assume that it satisfies conditions (1) and (3). We construct a sufficiently small $C^r$-perturbation $\Delta H$, so that for small enough $\varepsilon$ the Hamiltonian

$$H_0 + \varepsilon H_1 + \varepsilon \Delta H$$

satisfies the required properties. This perturbation consists of four parts:

$$\Delta H = \Delta H^{\text{mol}} + \Delta H^{\text{sr}} + \Delta H^{\text{dr}} + \Delta H^{\text{shad}}. \quad (7)$$

To perform these deformations we consider the net of resonances obtained in Section 3.

1. **Mollification $\Delta H^{\text{mol}}$:** Throughout the proof of Theorem 2 we need to perform normal forms, which decrease regularity. To stay in the $C^r$ regularity class, we mollifying the Hamiltonian $H_\varepsilon$ so that it is $C^\infty$ away from KAM tori. See Section 4.1.

2. **Perturbation $\Delta H^{\text{sr}}$:** It supported near single resonances. In Section 5 we construct invariant cylinders along single resonances. Normal forms show to construct these cylinders it is sufficient to have hyperbolicity of the first order approximation. Using the deformation we do attain required hyperbolicity. See Section 4.2.

3. **Perturbation $\Delta H^{\text{dr}}$:** It is supported in the neighborhoods of double resonances. Analogous to the previous one but for the double resonances. In double resonances, we need also some hyperbolicity properties to construct certain cylinders. This third deformation is done simultaneously with $\Delta H^{\text{sr}}$ in Sections 4.2.

4. **Perturbation $\Delta H^{\text{shad}}$:** It is supported near chosen resonant lines. This deformation allows us to shadow chosen collection of (Aubry) invariant sets. See Section 8.

### 2.2.1 Pöschel normal form and the mollification of the Hamiltonian

We start by performing the Pöschel normal form and mollifying $H_1$ away from KAM tori. Pöschel normal form was obtained in [Pö82] for autonomous Hamiltonian Systems. In Section 4.2, it is explained how to adapt it to the non-autonomous setting.

**Theorem 5.** Let $\eta, \tau > 0$, $r > 6 + 2\tau$. There exist $\varepsilon_0 > 0$ such that there exists a $C^{r-6-2\tau}$ change of coordinates $\Phi_{\text{Pö}}$ so that the Hamiltonian $N_\varepsilon = H_\varepsilon \circ \Phi_{\text{Pö}}$ is of the form

$$N_\varepsilon(I, \varphi, t) = H'_0(I) + R(I, \varphi, t)$$

where $H'_0$ satisfies

$$(2D)^{-1} \|v\| \leq \langle \partial^2 H'_0(I)v, v \rangle \leq 2D \|v\|$$

with the constant $D$ introduced in (1), and $R$ vanishes on the set

$$E_{\eta, \tau} = \{ (\varphi, I, t) : \partial_I H'_0(I) \in D_{\eta, \tau} \}, \text{ i.e. } R|_{E_{\eta, \tau}} \equiv 0.$$
Moreover, if $U_\lambda(E_{\eta,\tau})$ is the set of points $\lambda$-close to $E_{\eta,\tau}$, we have the following estimates

$$\| R \|_{C^j(U_{\rho_n}(E_{\eta,\tau}))} \leq C\varepsilon \rho_n^{-6-2\tau-j}$$

where $C > 0$ is a constant independent of $\varepsilon$ and $\rho_n$ and $j = 0, \ldots, \lfloor r - 6 - 2\tau \rfloor$.

Note that $E_{\eta,\tau} = \Phi^{-1}_{\text{P"os}}(\text{KAM}_{\eta,\tau})$ is the union of KAM tori $\text{KAM}_{\eta,\tau}$ with fixed Diophantine frequency. In the P"oschel coordinates all these tori are flat, $I = \text{constant}$. Moreover, $R$ being zero at $E_{\eta,\tau}$ implies that it is flat at these points. Indeed, each such a point is a Lebesgue density point.

Now we mollify the Hamiltonian $H_\varepsilon \circ \Phi_{\text{P"os}}$ away from $E_{\eta,\tau}$ or, equivalently, away from $\text{KAM}_{\eta,\tau}$. One possibility is just to mollify $H_1$. However, in the future we would like to state an open dense result with respect to KAM Whitney topology. This requires to make perturbations away from the KAM tori. We denote the mollification of $H_\varepsilon$, by $H_1^{\text{mol}}$ and the set of points in phase space $\rho_n$ far away from $\text{KAM}_{\eta,\tau}$ by $\Upsilon_{\rho_n}(\text{KAM}_{\eta,\tau})$.

**Lemma 2.** For any $\gamma > 0$ there exists a $C^r$ small perturbation $\Delta H_1^{\text{mol}}$ vanishing on $\text{KAM}_{\eta,\tau}$ and $c = c(H_0, r)$ such that

$$\| \Delta H_1^{\text{mol}} \|_{C^r} \leq \gamma \quad \text{and} \quad \| H_1 + \Delta H_1^{\text{mol}} \|_{C^{r+n}(\Upsilon_{\rho_n}(\text{KAM}_{\eta,\tau}))} \leq c \rho_n^{-n(1+4\tau)}$$

for any $\kappa > 0$.

From now on, we deal with the Hamiltonian

$$N'_\varepsilon(I, \varphi, t) = H_\varepsilon \circ \Phi_{\text{P"os}}(I, \varphi, t) + \Delta H_1^{\text{mol}} \circ \Phi_{\text{P"os}}(I, \varphi, t)$$

$$= H'(I) + R(I, \varphi, t) + \Delta H_1^{\text{mol}} \circ \Phi_{\text{P"os}}(I, \varphi, t).$$

(8)

In Step 1 we have constructed a tree of resonances segments in the frequency space. Now we need to pull it back to action space. We do it using the frequency map associated to the integrable Hamiltonian $H'_0$ obtained in Theorem 5. That is, we define the frequency map $\Omega(I) = \partial_t H'_0(I)$. Since $H'_0$ is $\varepsilon$-close to the original $H_0$ in (2), $H'_0$ is strictly convex and $\Omega$ is a global diffeomorphism. As explained in Definition 3, we have two type of resonances:

- horizontal and vertical resonant segments of low order from $\mathcal{F}_0$,
- the resonances from $\mathcal{F}_n$, $n \geq 1$, approaching the Diophantine frequencies.

In both cases we pull back these segments to the frequency space by $\Omega$. For each $k \in \mathcal{F}_0$ we define $\mathcal{I}_k$ as the pullback of the resonant segment $\mathcal{S}_k$. For each $\omega_n \in \mathcal{D}_{\eta,\tau}^{n, r}$ the corresponding resonant segment $\mathcal{I}_k^{\omega_n}$ is the pull back of $\mathcal{S}_k^{\omega_n}$.

### 2.2.2 Different regimes along resonances

To prove the existence of invariant cylinders along the resonant segments $\mathcal{I}_k^{\omega_n}$ in action space, we need to divide the tree of resonant segments in two different regimes. We call
these regimes transition zones and core of double resonances. The difference between them is whether the points are very close to double resonances whose intersecting resonant lines are of comparable order or are not close to such intersections.

Fix a generation \( n \), a segment \( T_{\omega_n} \) and \( \theta > 0 \). Define the \( \theta \)-strong resonances in \( T_{\omega_n} \).

**Definition 6.** Let \( k', k'' \in (\mathbb{Z}^2 \times 0) \times \mathbb{Z} \). Define

\[
\langle k', k'' \rangle := \{ k \in (\mathbb{Z}^2 \times 0) \times \mathbb{Z} : k = pk' + qk'', p, q \in \mathbb{Z} \}
\]

the lattice generated by integer combinations of \( k' \) and \( k'' \) and call

\[
\mathcal{N}(k', k'') = \min\{|K| : K \neq 0, K \in \langle k', k'' \rangle\}
\]

the order of the double resonance.

**Definition 7.** A double resonance \( T_{\omega_n} \cap \Gamma_{k'} \) is called \( \theta \)-strong if there exists \( k'' \in \mathcal{N}(\langle k_n, k' \rangle \setminus \langle k_n \rangle) \) such that

\[
|k'| \leq |k|^\theta.
\]

Otherwise, the double resonance is called \( \theta \)-weak.

By Lemma 9 the segment \( T_{\omega_n} \) can be covered by \( \mu \) balls \( B_\mu(I_i) \) where

\[
C^{-1} \rho^{(2\theta - 1) \frac{1+2\tau}{3-5\tau}} \leq \mu \leq C \rho^{(2\theta - 1) \frac{1+2\tau}{3-5\tau}}
\]

for some constant \( C > 0 \), and \( I_0 \in T_{\omega_n} \) is a \( \theta \)-strong double resonance.

Figure 3: Different regimes along a resonances. The larger balls have radii \( \sim C \rho^{(2\theta - 1) \frac{1+2\tau}{3-5\tau}} \) and centered at strong double resonances. The smaller balls represent the core of these double resonances. The rectangles are where the single resonant normal form are studied.

Now we divide each of these balls \( B_\mu(I_0) \) in transition zones and core of the double resonances. The core of the double resonance is defined as

\[
B_{C \rho^{m/2}}(I_0), \quad \text{where} \quad m = \frac{r}{10}.
\]
The transition zone is the annulus around the core of the double resonance, 

\[ \mathcal{A}(k_n, k') = \{ I \in B_u(I_0) : c \rho^{m/2} \leq |I - I_0| \leq \mu \} \]  

(10)

where \( c \) is chosen so that \( c > C \) to have overlap with (9).

In each of these zones we will obtain different invariant cylinders. The dynamics we obtain in the core of the double resonances is the same obtained in the double resonances in \( [KZ12] \). Nevertheless, here we face some additional difficulties since in this regime we are not close to a mechanical system anymore, but we have to deal with a more general 2 degrees of freedom Hamiltonian systems. This requires modifying the proof in \( [KZ12] \).

In the transition zones the dynamics is the same as in the transition zones studied in \( [KZ12] \). Note that here we do not need the purely single resonance regime which is used in \( [KZ12] \). The reason is that the double resonances are so abundant that the whole resonance can be covered by core of double resonances and transition zones.

In the next two sections we deal with the two different regimes. In both cases we first need to perform a normal form procedure to the Hamiltonian \( [K] \) obtained in Lemma \( [2] \). The normal form leads to new Hamiltonians which now can be treated as a perturbation of a new first order and is “easier” to analyze.

### 2.2.3 Normal form and invariant cylinders along single resonances

The transition zones \( (10) \) correspond to points “far enough” from strong double resonances. This implies that we can treat them, roughly speaking, as single resonance zones. To this end, we consider two overlapping transition zones and a tubular neighborhood inside them along the resonant segment \( \mathcal{I}_{k_n}^{\omega_n} \) such that it overlaps the core of two consecutive double resonances. Then, we prove that in such neighborhood there is a normally hyperbolic invariant cylinder. To this end, we need to

- perform a change of coordinates and show that our system is a small enough perturbation of a system which only depends on actions and on the slow angle \( k_n \cdot \varphi \).
- perturb the Hamiltonian so that this first order has sufficient nondegeneracy to imply the existence of a normally hyperbolic cylinder (see Key Theorem \( [2] \)).

Then, in Key Theorem \( [3] \) we use an isolating block argument, as done in \( [BKZ1] \), to prove that the full system has a normally hyperbolic invariant cylinder.

First, we perform a preliminary change of coordinates which separates the fast and slow angles. Thus, we define \( \varphi^s = k_n \cdot \varphi \) and \( \varphi^f = \varphi_2 \) (if the change is singular, take \( \varphi^f = \varphi_1 \)) and conjugated actions \( I^s \) and \( I^f \) so that the change of coordinates is symplectic. To simplify notation, we still call \( N'_{\epsilon} \) to the Hamiltonian \( [5] \) after the change of coordinates. Note that, this change of coordinates changes the \( C^1 \) (and higher regularity) norms of \( N'_{\epsilon} \) since \( k_n \) is large with respect to \( \rho_n \). The corresponding estimates are done in Section \( 5.2 \).

After this change of coordinates, the resonant segment \( \mathcal{I}_{k_n}^{\omega_n} \) can be parameterized by \( I^f \in [a_{k_n}^-, a_{k_n}^+] \) for some values \( a_{k_n}^\pm \).
Next theorem shows the existence of the normal form. A more precise statement is done in Section 5.2. Recall that $m = r/10$ (see (9)).

**Key Theorem 2.** There exists a $C^r$ function $\Delta H^{sr}$ supported in the transition zones and satisfying $\|\Delta H^{sr}\|_{C^r} \leq C\rho_n^{1+4r}$ in $\text{Vor}_n(\omega_n)$ for all $\omega_n \in D^n_{n,\tau}$, with $C, D > 0$ independent of $\rho_n$, so that, for each $k_n$, there exists a conformally symplectic change of coordinates $\Phi_k$ which transforms the Hamiltonian $(H + \Delta H^{mol} + \Delta H^{sr}) \circ \Phi_{\text{Pos}}$ into a Hamiltonian

$$
\mathcal{H}^k_n(\psi, J, t) = (H + \Delta H^{mol} + \Delta H^{sr}) \circ \Phi_{\text{Pos}} \circ \Phi^k_n(\psi, J, t)
$$

such that

- $\mathcal{H}^k_0$ is strictly convex. Moreover, $D^{-1}\rho_n^{m+1} \text{Id} \leq \partial^2 \mathcal{H}^k_0 \leq D\rho_n^{m-1} \text{Id}$.

- $\|\mathcal{Z}^k_n\|_{C^{r-6-2r-3m+\kappa}} \leq C\rho_n^{m/5-(1+4r)\kappa}$ and $\|\mathcal{R}^k_n\|_{C^{r-6-2r-3m+\kappa}} \leq C\rho_n^{q(r+1)-(1+4r)\kappa}$ for any $0 \leq \kappa \leq r - (r - 6 - 2\tau - 3m) = 6 + 2\tau + 3m$.

- Consider $b^-_{k_n} < b^+_{k_n}$, so that $J^f \in [b^-_{k_n}, b^+_{k_n}]$ parameterizes the single resonance $\mathcal{I}^\omega_n$ intersected with the transition zone in the new coordinates. Then, there exists a sequence of action values $\{J^f_i\}_{i=1}^N \subset [b^-_{k_n}, b^+_{k_n}]$, such that

  - for each $J^f_i \in (J^f_i, J^f_{i+1})$, the Hamiltonian $\mathcal{H}^k_n + \mathcal{Z}^k_n$ has a unique hyperbolic critical point which is a minimizer (the same happens in the intervals $[b^-_{k_n}, J^f_i]$ and $(J^f_{i+1}, b^+_{k_n}]$). We denote it by $(\psi^*_{s,i}, J^*_{s,i})$.

  - for $J^f = J^f_i$, the Hamiltonian $\mathcal{H}^k_n + \mathcal{Z}^k_n$ has two hyperbolic critical points which are minimizers. We denote them by $(\psi^*_{s,i}, J^*_{s,i})$ and $(\psi^*_{s,i+1}, J^*_{s,i+1})$.

In both cases, the eigenvalues of the hyperbolic periodic orbits $\lambda, -\lambda$ satisfy

$$
\lambda \geq C\rho_n^{dr+m+1}
$$

for some constant $C > 0$ independent of $\rho_n$.

The change of coordinates involved in this theorem is just the composition of a rescaling and a canonical transformation. The interval $[b^-_{k_n}, b^+_{k_n}]$ is just the former $[a^-_{k_n}, a^+_{k_n}]$ interval expressed with respect to the new fast action coordinate.

Theorem 2 implies that $\mathcal{H}^k_n + \mathcal{Z}^k_n$ has a sequence of normally hyperbolic invariant cylinders

$$
\{ (\psi, J, t) : (\psi^*, J^*) = (\psi^*_{s,J^f}, J^*_{s,J^f}), J^f \in [J^f_i, J^f_{i+1}] \}
$$

Thanks to hyperbolicity these cylinders can be actually slightly extended to intervals $[J^f_i - \delta, J^f_{i+1} + \delta]$ for small ($\rho_n$-dependent) $\delta > 0$. For the first and last one can take the intervals $[b^-_{k_n}, J^f_i + \delta]$ and $[J^f_{i+1} - \delta, b^+_{k_n}]$ respectively.
Note that the hyperbolicity of the cylinder is very weak since the eigenvalues of the hyperbolic fixed point of $H^{k_n} + Z^{k_n}$ are of size $\rho^{(dr+m+1)/2}$. The weak hyperbolicity is the reason why we need to reduce the size of the perturbation $R$ to $\|R\|_{C^2} \leq \rho^{(r+1)}$ to ensure that these cylinder persist for the full system. We state the persistence of the cylinders in the intervals $[J^f_i - \delta, J^f_{i+1} + \delta]$ to simplify the statement. The same theorem holds for the first and last interval.

**Key Theorem 3.** There exist $C^1$ maps

$$(\Psi^s_i, J^s_i)(\psi^f, J^f, t) : \mathbb{T} \times [J^f_i - \delta/2, J^f_{i+1} + \delta/2] \times \mathbb{T} \to \mathbb{T} \times \mathbb{R}$$

such that the cylinders

$$C^{\omega_n,i}_{k_n} = \{(\psi^s, J^s) = (\Psi^s_i, J^s_i)(\psi^f, J^f, t); \ (\psi^f, J^f, t) \in \mathbb{T} \times [J^f_i - \delta/2, J^f_{i+1} + \delta/2] \times \mathbb{T}\}$$

are weakly invariant with respect to the vector field associated to Hamiltonian (III), in the sense that the vector field is tangent to $C^{\omega_n,i}_{k_n}$. The cylinders $C^{\omega_n,i}_{k_n}$ are contained in the set

$$V_i := \{(\psi, J); J^f_i \in [J^f_i - \delta/2, J^f_{i+1} + \delta/2], \ ||\psi^s - \psi^s(J^f)|| \leq C\rho_n^{m-1+2dr}, \ ||J^s - J^s(J^f)|| \leq C\rho_n^{m-1+2dr}\},$$

for some constant $C > 0$ independent of $\rho$ and they contain all the full orbits of $\psi$ contained in $V_i$. We have the estimates

$$\|\Psi^s_i(\psi^f, J^f) - \psi^s(J^f)\| \leq C\rho_n^{m-1+5dr}, \ \|J^s_i(\psi^f, J^f, t) - J^s(J^f)\| \leq C\rho_n^{m-1+5dr},$$

and

$$\left\|\frac{\partial \Psi^s}{\partial J^f}\right\| \leq C\rho_n^{3dr-m+1/2}, \ \left\|\frac{\partial \Psi^s}{\partial \psi^f}\right\| \leq C\rho_n^{3dr-m+1/2} .$$

This theorem is proved in Section 6.

### 2.2.4 Normal form and invariant cylinders in the core of double resonances

To study the core of the double resonances we need to proceed as in the transition zone analysis and first perform a normal form procedure. We follow [Bou10]. This is explained in Section 5.3.

**Key Theorem 4.** There exists a $C^r$ function $\Delta H^{dr}$ supported in the core of double resonances satisfying $\|\Delta H^{dr}\|_{C^r} \leq C\rho^{q(r+1)}$ in $\text{Vor}_n(\omega_n)$ for all $\omega_n \in D_n$, with $C > 0$ independent of $\rho_n$, so that, for each double resonance $I_0 \in I^{\omega_n}_{k_n} \cap S(\Gamma_{k'})$ with $|k'| \leq R_\theta$, there exists a conformally symplectic change of coordinates $\Phi^{k_n,k'}_{\text{Fos}}$ which transforms the Hamiltonian $(H + \Delta H^{\text{mol}} + \Delta H^{\text{st}} + \Delta H^{dr}) \circ \Phi^{k_n,k'}_{\text{Fos}}$ into a Hamiltonian

$$\mathcal{H}_{k_n,k'}(\psi, J, t) = (H + \Delta H^{\text{mol}} + \Delta H^{dr}) \circ \Phi^{k_n,k'}_{\text{Fos}} \circ \Phi^{k_n,k'}(\psi, J, t)$$

$$= \mathcal{H}_0^{k_n,k'}(J) + Z^{k_n,k'}(\psi, J)$$

(13)
where $H_{0}^{k_{n},k'}$ only depends on actions and satisfies
\[ D^{-1}p^{m+1}_n \leq \partial J \leq Dp^{m-1}_n \]
for some $D > 1$ independent of $\rho_n$, in the sense of quadratic forms, and $Z^{k_{n},k'}$ is independent of time and satisfies
\[ \left\| Z^{k_{n},k'} \right\|_{C^{r-6-2\tau-m}} \leq C\rho^{m+(1+2\tau)(r-6-2\tau-m)/3}_n. \]

**Remark 8.** The perturbation $\Delta H^{dr}$ is nonzero in the overlapping zone between the core of the double resonances and the transition zones. Nevertheless satisfies $\left\| \Delta H^{dr} \right\|_{C^{r}} \leq C\rho^{dr}_n$. One can easily see that Key Theorems 2 and 3 remain true after this perturbation.

Note that the new Hamiltonian is autonomous and, therefore, has two degrees of freedom. In particular, its energy is conserved. We study dynamics within energy levels
\[ S_{E} = \left\{ (\psi, J) : H^{k_{n},k'}(\psi, J) = E \right\}. \]

The perturbation $\Delta H^{dr}$ makes the Hamiltonian $H^{k_{n},k'}$ nondegenerate. This implies that we can prove the existence of several hyperbolic objects in the different levels of energies. Recall that $H_{0}^{k_{n},k'}$ is convex and, making a translation if necessary, we have that $\nabla H_{0}^{k_{n},k'}(0,0) = 0$, $H_{0}^{k_{n},k'}(\psi, J) \geq 0$ and $H_{0}^{k_{n},k'}(0,0) = 0$.

The point $(\psi, J) = (0,0)$ is a critical point. In Appendix B, we show that due to nondegeneracy given by $\Delta H^{dr}$ this point $(0,0)$ is hyperbolic with distinct eigenvalues $-\lambda_1 < -\lambda_2 < 0 < \lambda_2 < \lambda_1$. As done in [KZ12], we split the study into two parts. First we deal with case of high energy, i.e. we study the dynamics in with $E \in (E_0, E_{*})$ with $E_0 > 0$ small and $E_{*}$ large. Then, we analyze the low energy regime, i.e. $E \in [0, e_0]$. We take $e_0 > 2E_0$ so that we have some overlap between the two regions.

**High energy regime** In the high energy regime we obtain similar results to the ones obtained in [KZ12]. Nevertheless note that now we do not deal with a mechanical system as in that paper since now the potential also depends on the actions. The proof in [KZ12] relies on identifying certain periodic orbits of the Hamiltonian with orbits of the geodesic flow on $\mathbb{T}^2$ given by the Mapertuis metric associated to the mechanical system restricted to a fixed (regular) energy level. This is not possible in the present case. However, following [CIPP98], one can obtain a Finsler metric $\delta_E$ associated to the Hamiltonian $H^{k_{n},k'}$ in the regular energy levels $S_{E}$. We call $\ell_E$ the length associated to the metric $\delta_E$. This metric allows us to obtain similar results to those in [KZ12]. This is explained in Appendix B.

We fix a homology class $h \in H^1(\mathbb{T}^2, \mathbb{Z})$ and we look for the shortest geodesics in this homology class with respect to the length $\ell_E$.

**Key Theorem 5.** Fix small $E_0 > 0$, large $E_{*} > 0$ and a homology class $h \in H^1(\mathbb{T}^2, \mathbb{Z})$. Then, there exists energy values $\{E_j\}_{j=1}^N$ such that

---

20
• For each $E \in \bigcup_{j=0}^{N}[E_j, E_{j+1}] \cup [E_N, E_*]$, there exists a unique shortest geodesic $\gamma^E_h$ with respect to the length $\ell_E$. Moreover, it is non-degenerate in the sense of Morse, i.e. the corresponding periodic orbit is hyperbolic.

• For $E = E_j$, there exists two shortest geodesics $\gamma^E_h$ and $\sigma^E_h$ with respect to the length $\ell_E$. Moreover, they are non-degenerate in the sense of Morse, i.e. the corresponding periodic orbit is hyperbolic. Moreover they satisfy

$$\left.\frac{d(\ell_E(\gamma^E_h))}{dE}\right|_{E=E_j} \neq \left.\frac{d(\ell_E(\sigma^E_h))}{dE}\right|_{E=E_j}.$$

This theorem is proved in Appendix B. Thanks to the hyperbolicity of the periodic orbits, $\gamma^E_h$ has a unique smooth local continuation to an energy interval $[E_j - \delta, E_{j+1} + \delta]$ for small $\delta > 0$. For the energies outside of $[E_j, E_{j+1}]$ the orbit is still hyperbolic but it is not the shortest geodesic anymore. We consider its union

$$\mathcal{M}^{E_j, E_{j+1}}_h = \bigcup_{E \in [E_j - \delta, E_{j+1} + \delta]} \gamma^E_h, \quad i = 1, \ldots, N.$$ 

For the boundary intervals we take $[E_0, E_1 + \delta]$ and $[E_N - \delta, E_*]$. The hyperbolicity of $\gamma^E_h$ implies that $\mathcal{M}^{E_j, E_{j+1}}_h$ is a normally hyperbolic cylinder of $\mathcal{H}^{k_a, k'}$.

Low energy regime  Now we deal with the low energy regime, that is, energy close to the energy $E = 0$ the saddle $(\phi, J) = (0, 0)$ belongs to. In $\mathcal{S}_0$, the Finsler metric is singular at $\phi = 0$. Let $\gamma^0_h$ be the shortest geodesic in the homology $h$ with respect to the Finsler metric induced on $\mathbb{T}^2$ from $\mathcal{S}_0$. In Appendix B is proved the following. The result for mechanical systems was done by Mather in [Mat03].

**Lemma 3.** The geodesic $\gamma^0_h$ satisfies one of the following statements:

• $0 \in \gamma^0_h$ and $\gamma^0_h$ is not self-intersecting. We call such homology class $h$ simple non-critical and the corresponding geodesic a simple loop.

• $0 \in \gamma^0_h$ and $\gamma^0_h$ is self-intersecting. We call such homology class $h$ non-simple and the corresponding geodesic a non-simple.

• $0 \not\in \gamma^0_h$ and $\gamma^0_h$ is a regular geodesic. We call such homology class $h$ simple critical.

As explained in [KZ12] in the mechanical systems setting, there are at least two homology classes which are simple non-critical, that is, with associated geodesics which contain the saddle and are non-intersecting. In the phase space, such geodesics are homoclinic orbits to the saddle at $(0, 0)$. The next lemma describes what happens generically in the self-intersecting case. It was proved by Mather in the mechanical systems setting in [Mat03] (see also Appendix B).

Recall that a property of a Tonelli Hamiltonian is Mané’s $C^r$ generic if there is a $C^r$ generic set of potentials $U$ on $\mathbb{T}^2$ such that $\mathcal{H}_U = \mathcal{H} + U$ satisfies this property.
Lemma 4. Suppose the Jacobi-Finsler metric $\delta_0$ satisfies properties of Corollary 3, i.e. it has exactly one singular point $\varphi^*$ such that for the associated Hamiltonian flow this point is a saddle with real eigenvalues. Consider $h \in H_1(\mathbb{T}^2,\mathbb{Z})$ such that the associated geodesic $\gamma^0_h$ of $\delta_0$, is non-simple. Then for a Mañé generic Hamiltonian $H$, there exist homology classes $h_1, h_2 \in H_1(\mathbb{T}^2,\mathbb{Z})$, integers $n_1, n_2$ with $h = n_1 h_1 + n_2 h_2$ such that the associated geodesics $\gamma^0_{h_1}$ and $\gamma^0_{h_2}$ are simple critical.

For $E > 0$, the geodesic has no self-intersections. Therefore, this lemma implies that there is a unique way of writing $\gamma^0_h$ as a concatenation of $\gamma^0_{h_1}$ and $\gamma^0_{h_2}$. Denote $n = n_1 + n_2$.

Lemma 5. There exists a sequence $\sigma = (\sigma_1, \ldots, \sigma_n) \in \{1, 2\}^n$, unique up to cyclical permutation, such that

$$\gamma^0_h = \gamma^0_{h_{\sigma_1}} * \gamma^0_{h_{\sigma_2}} * \cdots * \gamma^0_{h_{\sigma_n}}.$$ 

These lemmas describe what happens at the critical energy level. From this information we describe what happens for small energy. Assume the following.

[A1] The Hamiltonian $\mathcal{H}$ on the Mañé critical energy surface $\mathcal{H} = \alpha(0)$ has only one critical point $(\psi^*, J^*)$, it is a saddle with distinct real eigenvalues $-\lambda_1 < -\lambda_2 < 0 < \lambda_2 < \lambda_1$.

See Section B.1 for definition of Mañé critical value $\alpha(0)$. In a neighborhood of $(0, 0)$ there is a local system of coordinates $(s, u) = (s_1, s_2, u_1, u_2)$ such that the $u_i$ axes correspond to the eigendirections of $\lambda_i$ and the $s_i$ axes correspond to the eigendirections of $-\lambda_i$ for $i = 1, 2$.

Now we need to impose some assumptions on the geodesics $\gamma^0_h$. They are different depending on the properties stated in Lemma 3 satisfying the geodesic. We denote by $\gamma^+ = \gamma^0_{h,+}$ be a homoclinic to $(0, 0)$ of homology $h$ and $\gamma^- = \gamma^0_{h,-}$ the symmetric one with respect to the involution $I \mapsto -I$, $t \mapsto -t$. Assume that

[A2] The homoclinics $\gamma^\pm$ are not tangent to the $u_2$ or $s_2$ axis at $(0, 0)$.

Since the eigenvalues are all different, this implies that they are tangent to the $u_1$ and $s_1$ axis. Assume without loss of generality, that $\gamma^+$ approaches $(0, 0)$ along $s_1 > 0$ in forward time and along $u_1 > 0$ in backward time. Using reversibility, $\gamma^-$ approaches $(0, 0)$ along $s_1 < 0$ in forward time and along $u_1 < 0$ in backward time.

For the nonsimple case we assume the following. We consider two homoclines $\gamma_1$ and $\gamma_2$ which are in the same direction instead of being in the opposite direction.

[A2'] The homoclinics $\gamma_1$ and $\gamma_2$ are not tangent to the $s_2$ and $u_2$ axis at $(0, 0)$. Both approach $(0, 0)$ along $s_1 > 0$ in forward time and along $u_1 > 0$ in backward time.

In the case the homology $h$ is simple and $0 \notin \gamma_h$ we assume the following.

[A2''] The closed geodesic $\gamma_h^0$ is hyperbolic.
Fix any $d > 0$ and $\delta \in (0, d)$. Let $B_d$ be the $d$-neighborhood of $(0, 0)$ and let
\[ \Sigma^s_\pm = \{ s_1 = \pm \delta \} \cap B_d, \quad \Sigma^u_\pm = \{ u_1 = \pm \delta \} \cap B_d \]
be four local sections transversal to the flow of (13). We define the local maps
\[
\Phi_{\text{loc}}^{++} : U^{++} \subset \Sigma^s_+ \to \Sigma^u_+, \quad \Phi_{\text{loc}}^{--} : U^{--} \subset \Sigma^u_- \to \Sigma^s_-
\]
These maps are defined by the first time the flow of (13) hit the sections $\Sigma^u_\pm$. These maps are not defined in the whole sets $U^{\pm \pm}$ since some orbits might escape. For such points we consider that the maps are undefined.

For the simple case, by assumption [A1] we know that for $\delta$ small enough, $\gamma^+$ and $\gamma^-$ intersect the local stable and unstable manifolds $\Sigma^u_\pm$ and $\Sigma^s_\pm$. Let $p^+$ and $q^+$ (respectively $p^-$ and $q^-$) the intersection of $\gamma^+$ (resp. $\gamma^-$) with $\Sigma^u_+$ and $\Sigma^u_-$ (resp. $\Sigma^s_+$ and $\Sigma^s_-$). Then, for small neighborhoods $V^\pm \ni q^\pm$ there are well defined Poincaré maps
\[ \Phi_{\text{glob}}^\pm : V^\pm \to \Sigma^s_\pm. \]
In the nonsimple case, by [A2′] implies that $\gamma^i$, $i = 1, 2$, intersect $\Sigma^u_\pm$ at $q_i$ and $\Sigma^s_\pm$ at $p_i$. The associated global maps are denoted
\[ \Phi_{\text{glob}}^i : V^i \to \Sigma^s_+, \quad i = 1, 2. \]
Analyzing the composition of these maps at energy surfaces of small energy, we show that there exist periodic orbits shadowing the homoclinics and concatenation of homoclinics.

We now assume that the global maps are “in general position”. We will only phrase our assumptions [A3] for the homoclinic $\gamma^+$ and $\gamma^-$. The assumptions for $\gamma^1$ and $\gamma^2$ are identical, only requiring different notations and will be called [A3′]. Let $W^s$ and $W^u$ denote the local stable and unstable manifolds of $(0, 0)$. Note that $W^u \cap \Sigma^u$ is one-dimensional and contains $q^\pm$. Let $T^{uu}(q^\pm)$ be the tangent direction to this one dimensional curve at $q^\pm$. Similarly, we define $T^{ss}(p^\pm)$ to be the tangent direction to $W^s \cap \Sigma^s_\pm$ at $p^\pm$.

[A3 ] Image of strong stable and unstable directions under $D\Phi_{\text{glob}}^\pm(q^\pm)$ is transverse to strong stable and unstable directions at $p^\pm$ on the energy surface $S_0 = \{ H^s = 0 \}$.

For the restriction to $S_0$ we have
\[ D\Phi_{\text{glob}}^\pm(q^\pm)|_{T_{S_0}T^{uu}(q^\pm)} \cap T^{ss}(p^\pm). \]

[A3′ ] Suppose condition [A3] hold for both $\gamma_1$ and $\gamma_2$.

In the case that the homology $h$ is simple and $0 \notin \gamma_h^0$, we assume

[A3′′ ] The closed geodesic $\gamma_h^0$ is hyperbolic.
Figure 4: The cylinders by Corollary 2 in the low energy regime in the core of double resonances. They are foliated by the periodic orbits obtained in Key Theorem 6. The two cylinders are tangent to the same plane defined by the weak stable and unstable directions of the saddle $(0,0)$. Such cylinders are called kissing cylinders in [KZ12].

**Key Theorem 6.** 1. Fix one homology class $h \in H_1(\mathbb{T}^2, \mathbb{Z})$ (resp. two homology classes $h_1, h_2$). Then for an $C^r$ open $C^r$ dense set of Tonelli Hamiltonians assumptions [A1-A3] holds. For Hamiltonians satifying these assumptions we have:

2. Consider the geodesics $\gamma^\pm$ associated to the Hamiltonian (13). Then, there exists $e_0 > 0$ such that

- For each $E \in (0, e_0]$, there exists a unique periodic orbit $\gamma^+_E$ corresponding to a fixed point of the map $\Phi^+_{\text{glob}} \circ \Phi^{++}_{\text{loc}}$ restricted to the energy surface $S_E$.
- For each $E \in (0, e_0]$, there exists a unique periodic orbit $\gamma^-_E$ corresponding to a fixed point of the map $\Phi^-_{\text{glob}} \circ \Phi^{-+}_{\text{loc}}$ restricted to the energy surface $S_E$.
- For each $E \in [-e_0, 0)$, there exists a unique periodic orbit $\gamma^c_E$ corresponding to a fixed point of the map $\Phi^-_{\text{glob}} \circ \Phi^{-+}_{\text{loc}} \circ \Phi^+_{\text{glob}} \circ \Phi^{++}_{\text{loc}}$ restricted to the energy surface $S_E$.

3. Consider the geodesics $\gamma^1$ and $\gamma^2$ associated to the Hamiltonian (13). Then, there exists $e_0 > 0$ such that for each $E \in (0, e_0]$, the following hold. For any $\sigma = (\sigma_1, \ldots, \sigma_n) \in \{1, 2\}^n$, there is a unique periodic orbit, denoted by $\gamma^E_\sigma$, corresponding to a unique fixed point of the map

$$\Pi_{i=1}^n (\Phi^\sigma_{\text{glob}} \circ \Phi^{++}_{\text{loc}})$$

restricted to the energy surface $S_E$. (Product stands for composition of maps.)
The second part of this theorem is proved in [KZ12], Theorem 6. Namely, if conditions [A1-A3] holds then conclusions 2 and 3 holds. A proof that condition 1 is Mañé generic is Theorem 4, [KZ12].

As a corollary of Key Theorem 6, we have normally hyperbolic invariant cylinders.

**Corollary 2.** 1. Fix a simple noncritical homology class \( h \in H_1(\mathbb{T}^2, \mathbb{Z}) \). Consider the homoclinic orbits \( \gamma_\pm \) and family of periodic orbits \( \gamma^+_E, \gamma^-_E, \gamma^c_E \) obtained in Key Theorem 6. Then, the set

\[
\mathcal{M}_E^{e_0} = \bigcup_{0<E\leq e_0} \gamma^+_E \cup \gamma^+_E \bigcup_{-e_0\leq E<0} \gamma^c_E \cup \gamma^- \bigcup_{0<E\leq e_0} \gamma^-_E
\]

is a \( C^1 \) normally hyperbolic invariant manifold with boundaries \( \gamma^+_{e_0} \) and \( \gamma^-_{e_0} \) for the flow of Hamiltonian \( (13) \).

2. Fix a nonsimple homology class \( h \in H_1(\mathbb{T}^2, \mathbb{Z}) \). Consider the periodic orbits \( \gamma^E_\sigma \) obtained in Key Theorem 6 where \( \sigma \) is the sequence determined by Lemma 4. Then, for any \( 0 < e < e_0 \), the set

\[
\mathcal{M}_h^{e, e_0} = \bigcup_{e < E \leq e_0} \gamma^E_\sigma
\]

is a \( C^1 \) normally hyperbolic invariant manifold with boundary for the flow of Hamiltonian \( (13) \).

All cylinders associated to simple homology classes are tangent along a two dimensional plane at the original formed by the weak hyperbolic directions \((s_1, u_1)\). For this reason, in [KZ12], the cylinders \( \mathcal{M}_h^{E_0} \) are called kissing cylinders.

Taking the constants \( E_0 \) and \( e_0 \) introduced in Key Theorems 5 and 6 respectively, such that \( E_0 < e_0 \) implies that there is overlapping between the low and the high energy regime. Take a simple non-critical homology class \( h \). Then, \( \mathcal{M}_h^{E_0, E_1} \), the last of the cylinders given in Key Theorem 5 can be continued by \( \mathcal{M}_h^{e_0} \) thanks to Corollary 2 and the uniqueness given by Key Theorem 6. That is,

\[
\mathcal{M}_h^{E_0, E_1} \cup \mathcal{M}_h^{e_0} \cup \mathcal{M}_h^{E_0, E_1}
\]

is a normally hyperbolic invariant manifolds.

**2.3 Localization of the Aubry sets**

In this section we localize the Aubry sets of certain cohomologies and show that they belong to the cylinders obtained in Section 2.2. First in Section 2.3.1 we deal with the transition zones and then, in Section 2.3.2 we deal with the core of double resonances.
2.3.1 Aubry sets in the transition zones

We proceed as in [BKZ11]. Nevertheless, since the hyperbolicity of the cylinder $C_{k_n}^I$ is very weak the prove needs to use more accurate estimates. First we state it for the Hamiltonian obtained in Theorem 2 by a normal form procedure. Then we deduce analogous results for the Hamiltonian $\tilde{H}$ in (8).

**Key Theorem 7.** Fix a resonant segment $T_{k_n}^\omega$. Consider the Hamiltonian $H_{k_n}$ obtained in Key Theorem 2. For each cohomology class $c = (c^s, c^f) = (0, c^f)$ with $c^f \in [\tilde{a}^-, \tilde{a}^+]$, the Mane set $\tilde{N}(c)$ is contained in the cylinder $C_{k_n}^\omega$ obtained in Key Theorem 3. Moreover, let $\pi_{\varphi^f}$ be the projection onto the fast component $\varphi^f$. Then, we have that $\pi_{\varphi^f}|_{A(c)}$ is one-to-one and the inverse is Lipschitz.

The proof of this theorem is done in Section 7. It follows similar lines as the results obtained in [BKZ11] but additional estimates are required to deal with the extremely weak hyperbolicity.

From this theorem we can deduce analogous result for the Hamiltonian $\tilde{H}$ in (8).

**Corollary 3.** Fix a resonant segment $T_{k_n}^\omega$ and Consider the Hamiltonian $\tilde{H}$ in (8).

2.3.2 Aubry sets in the core of the double resonances

In the double resonance regime, after a normal form procedure, we deal with the Hamiltonian (13). In Corollary 2 we have constructed a variety of Normally Hyperbolic Invariant Manifolds. Now, we need to see that certain Aubry sets belong to those NHIMs. We proceed as in [KZ12]. Nevertheless, recall that now, at first order, we do not have a mechanical system.

The cylinders obtained in Key Theorem 5 and Corollary 2 are related to the integer homology classes whereas the Aubry sets are related to the cohomology classes. Therefore, the first step is to relate them. Each minimal geodesic $\gamma_E^h$ corresponds to a minimal measure of the system associated to Hamiltonian (13). Moreover, it has associated a cohomology class. Assume that $\gamma_E^h$ is parameterized so that it satisfies the Euler-Lagrange equation and call $T = T(\gamma_E^h)$ to its period. Then, the probability measure supported on $\gamma_E^h$ is a minimal measure and its rotation number is $T^{-1}$. Then, the associated cohomology class is the convex subset

$$L\mathcal{F}_\beta(h/T(\gamma_E^h)),$$

of $H^1(\mathbb{T}^2, \mathbb{R})$, where $L\mathcal{F}_\beta$ is the Legendre-Fenchel transform of the $\beta$-function defined by Mather. Recall that in Theorems 5 and 6 we have seen that for $E \in (0, E_0)$ or $E \in (E_j, E_{j+1})$ $j = 0, \ldots, N - 1$, there exists a unique minimal geodesic $\gamma_E^h$ for energy $E$. In this case, we define

$$\lambda_E^h = \frac{1}{T(\gamma_E^h)}.$$

For the bifurcation values $E = E_j$ there are two minimal geodesics $\gamma_E^h$ and $\gamma_E^h$. We write $\lambda_E^h = 1/T(\gamma_E^h)$, where the choice of geodesic is arbitrary. In [KZ12] it is shown that
the set $\mathcal{LF}_\beta(\lambda^E_h)h$ is independent of the choice of geodesic in the setting of mechanical systems. This fact is still true for Hamiltonians of the form (13). The set $\mathcal{LF}_\beta(\lambda^E_h)h$ is a well defined set function of $E$.

We call the union

$$\bigcup_{E>0} \mathcal{LF}_\beta(\lambda^E_h)$$

the channel associated to the homology $h$ and we choose a curve of cohomologies along this channel. For cohomologies in such curve, the Aubry sets will belong to the cylinders.

**Key Theorem 8.** Consider the Hamiltonian (13). There exists a continuous function $c_h : [0, E] \rightarrow H^1(T^2, \mathbb{R})$ satisfying $\mathcal{F}_h(E) \in \mathcal{LF}_\beta(\lambda^E_h)h$ with the following properties

1. For $E \in (0, E_0)$ or $E \in (E_j, E_{j+1}), j = 0, \ldots, N - 1$,
   $$A(\tau_h(E)) = \gamma^E_h.$$

2. For $E = E_j, j = 0, \ldots, N - 1$,
   $$A(\tau_h(E_j)) = \gamma^E_h \cup \gamma^E_h.$$

3. If $h$ is simple and critical, then,
   $$A(\tau_h(0)) = \gamma^0_h$$
   and for $\lambda \in [0, 1]$,
   $$A(\lambda \tau_h(0)) = \{\psi = 0\}.$$

4. If $h$ is simple and non-critical, then,
   $$A(\tau_h(0)) = \gamma^0_h \cup \{\psi = 0\}$$
   and for $\lambda \in [0, 1]$,
   $$A(\lambda \tau_h(0)) = \{\psi = 0\}.$$

### 2.4 Shadowing

#### 3 First step of the proof: Selection of resonances

In this section we prove Key Theorem 8. That is, we construct a net of resonances $\{S_n\}_{n \in \mathbb{Z}}$ in the frequency spaces $U'_\eta$ such that the following properties hold

- each segment $S_n$ belongs to a resonant segment

$$S_n \subset S(\Gamma_{k_n}) = \{\omega \in U'_\eta : k_n \cdot (\omega, 1) = 0\}$$

for some $k_n \in (\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$. 

27
The union $\bigcup_n S_n$ is connected.

The closure $\overline{\bigcup_n S_n}$ contains all Diophantine frequencies in $D_{\eta,\tau}$ (see (1)).

Moreover, we need these resonances satisfy quantitative estimates on speed of approximation. We choose our segments so that

- for each $S_n \subset S(\Gamma_{k_n})$ there is a Diophantine number $\omega \in D_{\eta,\tau}$ such that $$|k_n \cdot (\omega, 1)| \leq |k_n|^{-2+3\tau}.$$
- if $S_n \cap S_m \neq \emptyset$, then we have a bound on ration $|k_n|/|k_m|$ as well as angle of intersection.

### 3.1 Approximation of Diophantine vectors

In this section we explain how to construct a sequence of resonant segments approaching one single Diophantine frequency. Later, in Section 3.3 we explain how to deal with all $(\eta, \tau)$-Diophantine frequencies at the same time and we prove Key Theorem 1.

Take $\omega = (\omega_1, \omega_2) \in D_{\eta,\tau} \subset \mathbb{R}^2$. For each $k \in (\mathbb{Z}^2 \times 0) \times \mathbb{Z}$ denote $\Gamma_k = \{ \omega : (\omega, 1) \cdot k = 0 \}$ the resonant line. We distinguish the resonant line $\Gamma_k \in \mathbb{R}^2$ from the corresponding resonant segment $S(\Gamma_k) = \{ I : k \cdot (\partial H_0(I), 1) = 0 \}$. In the case $H_0(I) = (I_1^2 + I_2^2)/2$ they coincide, but in general, are different.

Fix a $(\eta, \tau)$-Diophantine frequency $\omega \in D_{\eta,\tau}$. We construct a zigzag of curves $\{\Gamma_n^\omega\}_{n \geq 1}$ approximating $\omega$ and satisfying certain quantitative estimates.

**Theorem 9.** Fix $\omega^* \in D_{\eta,\tau}$ and a sequence of radii $\{R_n\}_{n \in \mathbb{Z}}$ defined as $$R_{n+1} = R_n^{1+2\tau}, \quad R_0 \gg 1, \kappa \geq 1.$$ Then, there exist $\xi > 0$ and $c > 0$ and a sequence of $(\mathbb{Z}^2 \setminus 0) \times \mathbb{Z}$-vectors $\{k_n\}_{n \in \mathbb{N}}$, such that for each positive integer $n$ we have

1. The $n$-th vector $k_n$ belong to the following annulus $R_n^{-\frac{1}{2}} \leq |k_n| \leq R_n$.
2. The vectors satisfy $$\eta R_n^{-(2+\tau)} \leq |k_n \cdot \omega| \leq 16\eta^{-2}R_n^{-(2-2\tau)}.$$ 3. Consecutive vectors satisfy $k_n \not\parallel k_{n+1}$ and the angles between them satisfy $$\angle(k_n, k_{n+1}) \geq \frac{\pi}{6}.$$
4. Let
\[ \mathbb{D}_n(\omega^*) = \{ |\omega^* - \omega| \leq \xi \eta^{-7/2} R_n^{-(3-4\tau)} \}. \] (14)
Then \( \omega_{n+1} := \Gamma_{k_n} \cap \Gamma_{k_{n+1}} \subset \mathbb{D}_n(\omega^*) \). Denote \( \Gamma^\omega_n \subset \Gamma_{k_n} \) the segment between \( \omega_n \) and \( \omega_{n+1} \). Then \( \Gamma^\omega_n(\omega_n, \omega_{n+1}) \subset \mathbb{D}_n(\omega^*) \).

5. Any other resonance \( \Gamma_k \) intersecting \( \mathbb{D}_n(\omega^*) \), satisfies
\[ \frac{|k_n|}{|k|} \leq c_2 \eta^{-2} R_n^{2\tau}. \]

This theorem is proved in Section 3.2. From item 2 in this theorem, we can obtain upper and lower bounds for the distance between \( \omega^* \) and any point in \( \Gamma^\omega_n \).

Corollary 4. For any point \( \omega \in \Gamma^\omega_n \) we have
\[ \eta R_n^{3-\tau} \leq |\omega - \omega^*| \leq \xi \eta^{-7/2} R_n^{-(3-4\tau)}. \]

Proof. The second estimate is given in the theorem. For the first one, it is enough to use the Diophantine property 4 to obtain
\[ |\omega - \omega^*| \geq \frac{|k_n \cdot \omega^*|}{|k_n|} \geq \frac{\eta}{|k_n|^{3+\tau}} \geq \eta R_n^{3-\tau}. \]

3.2 Proof of Theorem 9

The key point of the proof of Theorem 9 is a non-homogeneous Dirichlet Theorem, which can be found in [Cas57, Chapter 5, Theorem VI]. We state a simplified version, which is sufficient for our purposes. In this section, the norm \( \| \cdot \| \) measures the maximal distance to the integers, where the maximum is also taken over each component, i.e. for \( v = (v_1, \ldots, v_n) \in \mathbb{R}^n \) we denote
\[ \|v\| := \min_{(m_1, \ldots, m_n) \in \mathbb{Z}^n} \max_{1 \leq j \leq n} |v_j - m_j|. \]

Theorem 10. Let \( L(x), x = (x_1, \ldots, x_n) \) be a linear form and fix \( C, X > 0 \). Suppose that there does not exist any \( x \in \mathbb{Z}^n \setminus 0 \) such that
\[ \|L(x)\| \leq A \quad \text{and} \quad |x_i| \leq X. \]
Then, for any \( \alpha \in \mathbb{R} \), the equations
\[ \|L(x) - \alpha\| \leq A_1 \quad \text{and} \quad |x_i| \leq X_1. \]
have an integer solution, where
\[ A_1 = \frac{1}{2}(h+1)A, \quad X_1 = \frac{1}{2}(h+1)X \quad \text{and} \quad h = X^{-n}A^{-1}. \]
This theorem allows us to prove Theorem 9.

**Proof of Theorem 9.** Let \( \omega^* \in D^U_{\eta, \tau} \), i.e. \( |(\omega^*, 1) \cdot k| \geq \eta|k|^{-2-\tau} \) for each \( k \in \mathbb{Z}^2 \times (\mathbb{Z} \setminus 0) \). Consider the sequence of \( \{R_n\}_{n \in \mathbb{N}} \). We prove Theorem 9 by induction. Assume that we have \( k_n \) and we need to obtain \( k_{n+1} \) satisfying items 1–5.

Consider a point \( \tilde{k} \) at distance \( R_{n+1}/2 \) in the direction perpendicular to \( k_n \) along the plane \( T_{\omega^*} = \{ x : x \cdot \omega^* = 0 \} \). Take a ball of radius \( R_{n+1}/4 \) around it. We want to find a good approximating vector in this ball. This vector must be of the form \( \tilde{k} + x \) with \( |x| \leq R_{n+1}/4 \). This implies that for any \( k_{n+1} \) from this ball \( k_n \) and \( k_{n+1} \) make an angle at least \( \pi/6 \) and item 3 holds.

Define the linear form \( L(x) = \omega^* \cdot x \) and \( \alpha = \omega^* \cdot \tilde{k} \). Then,

\[
\omega^* \cdot k_{n+1} = L(x) + \alpha.
\]

Now we apply Theorem 10 taking

\[
X = \left( \frac{\eta R_{n+1}}{4} \right)^{1/(1+\tau)}.
\]

Then, using the Diophantine condition we know that taking

\[
A = \frac{\eta}{X^{2+\tau}},
\]

the equation \( \|L(x)\| \leq A \) has no solution for \( |x| \leq X \). Thus, we get,

\[
h = \eta^{-1} X^\tau.
\]

Then, using the formula for \( X \), we get a vector \( x \) satisfying \( |x| \leq X_1 \) with

\[
X_1 = \frac{1}{2} (h + 1) X \leq \eta^{-1} X^{1+\tau} = \frac{R_{n+1}}{4}
\]

which is solution of

\[
|k_{n+1} \cdot \omega^*| = \|L(x) - \alpha\| \leq hA \leq \frac{1}{X^2} \leq \frac{16}{\eta^2 R_{n+1}^{2-2\tau}}.
\]

Moreover, \( |k_{n+1}| = |\tilde{k} + x| \geq R_{n+1}/4 \). Thus, \( k_{n+1} \) satisfies items 1 and 2.

In order to obtain the first \( k_1 \), it is enough to apply the classical Dirichlet Theorem. This \( k_1 \) satisfies better estimates.

Now, to prove that the resonant segment \( \Gamma_{k_{n+1}}^{\omega^*} \) intersects \( \Gamma_{k_n}^{\omega^*} \). We start by pointing out that, using the already obtained estimates from item 4

\[
\text{dist} \left( \Gamma_{k_n}^{\omega^*}, \omega \right) = \frac{|k \cdot \omega^*|}{|k|} \leq \eta^{-7/2} R_{n+1}^{-(3-4\tau)}.
\]
Then, taking into account that we have a uniform bound of the angle between $k_n$ and $k_{n+1}$, one can take $\xi$ large enough independent of $R_{n+1}$ so that the two resonant segments intersect.

It only remains to see that any other resonance crossing $\Gamma_{k_n}^*\cap\gamma$ is of almost the same order. Assume that there is a resonance $\Gamma_k$ intersects the disk

$$\mathbb{D}(\omega^*) = \{ |\omega - \omega^*| \leq \xi \eta^{-7/2} R_{n+1}^{-(3-4\tau)} \}.$$ 

This implies that the distance to $\omega^*$ has to satisfy

$$\text{dist}(\Gamma_k, \omega^*) = \frac{|k \cdot \omega^*|}{|k|} \leq \xi \eta^{-7/2} R_{n+1}^{-(3-4\tau)}.$$ 

which, using the Diophantine condition (4) implies

$$|k| \geq \xi^{-1} \eta^{7/2} R_{n+1}^{3-4\tau} |k \cdot \omega^*| \geq \xi^{-1} \eta^{9/2} R_{n+1}^{3-4\tau} |k|^{-2(2+\tau)}$$

and then

$$|k|^{3+\tau} \geq \xi^{-1} \eta^{9/2} R_{n+1}^{3-4\tau}.$$ 

Then, taking $\tau$ small enough, one can get the last estimate of Theorem 9. 

3.3 Selection of resonances for all Diophantine frequencies: proof of Key Theorem 1

In Section 3.1 we have obtained a sequence of Dirichlet resonances approaching a single Diophantine frequency $\omega^* \in \mathcal{D}_{\eta,\tau}$. Now we generalize this result to obtain sequences of resonances approaching all frequencies in $\mathcal{D}_{\eta,\tau}^n$ and we prove Theorem 1. Since the number of Dirichlet resonant segments grows we need to control the complexity of possible intersections. We modify the strategy of the proof of Theorem 9.

In Section 2.4 we have defined the sequence of sets of Diophantines frequencies $\mathcal{D}_{\eta,\tau}^n$, which are $3 \rho_n$-dense subsets of $\mathcal{D}_{\eta,\tau}^n$, such that no two points in $\mathcal{D}_{\eta,\tau}^n$ are closer than $\rho_n$. In Lemma 1 we have given some properties of these sets. Consider $\omega_n^* \in \mathcal{D}_{\eta,\tau}^n$ and its Voronoi cell $\text{Vor}_n(\omega_n^*)$ (see Section 2.1). Assume that at the stage $n$ we have obtained a resonance $\Gamma_{k_n^*}$ which intersects $B_{\rho_n}(\omega_n^*) \subset \text{Vor}_n(\omega_n^*)$ and we call $S(\Gamma_{k_n^*})$ a segment of it. Now, at the following step, we consider the frequencies of $\mathcal{D}_{\eta,\tau}^{n+1}$ which are contained in $\text{Vor}_n(\omega_n^*)$,

$$\{\omega_{n+1}^*\}_{j \in J} = \mathcal{D}_{\eta,\tau}^{n+1} \cap \text{Vor}_n(\omega_n^*),$$

which have associated Voronoi cells $\text{Vor}_n(\omega_{n+1}^*)$. From Lemma 1 and using the definition of the sequence $\{\rho_n\}_{n \in \mathbb{N}}$ one can easily deduce the following lemma.

**Lemma 6.** Take $\rho_0 > 0$ small enough. Then, for all $n \in \mathbb{N}$ and any $\omega_n^* \in \mathcal{D}_{\eta,\tau}^n$. The set of points in $\omega_n^* \in \mathcal{D}_{\eta,\tau}^{n+1}$ such that $B_{3\rho_{n+1}}(\omega_n^*) \cap \text{Vor}_n(\omega_n^*) \neq \emptyset$ contains at most $K \rho_n^{-4\tau}$ points, for some constant $K > 0$ independent of $\rho_n$. 

31
Now at each step $n$ we need to place resonant segments that intersect the corresponding Voronoi cells. We place them using the results in Section 3.1 in such a way that the placed resonant segments do not intersect resonant segments which do not belong to the neighboring generations. This, jointly with Lemma 6, will give a good bound of the possible multiple intersections of Dirichlet resonant lines.

Now we are ready to prove Key Theorem 1. Recall that we are looking for Dirichlet resonant lines $\Gamma_{k_{n+1}}^{*,j}$ such that

- $\Gamma_{k_{n+1}}^{*,j}$ intersects $B_{\rho_{n+1}}(\omega_{n+1}^{*,j}) \subset \text{Vor}_{n}(\omega_{n+1}^{*,j})$.
- $\Gamma_{k_{n+1}}^{*,j}$ intersects the previous resonant segment $S(\Gamma_{k_{n}})$ inside $B_{3\rho_{n}}(\omega_{n}^{*})$.
- Does not intersect resonant segments which do not belong to the neighboring generations.

**Proof of Key Theorem 1.** As in the proof of Theorem 9, the norm $\| \cdot \|$ measures the distance to the integers. We prove the theorem inductively. Assume that we have placed the Dirichlet resonant segments for the frequencies of the first $n$ generations and fix a frequency $\omega_{n}^{*} \in D_{n,\eta,\tau}$. Now we place the Dirichlet resonant segments associated with the frequencies $D_{n+1,\eta,\tau} \cap \text{Vor}_{n}(\omega_{n}^{*})$. Assume, by inductive hypothesis, that all the already placed Dirichlet resonant lines satisfy the properties stated in Theorem 1.

Now, proceeding as in the proof of Theorem 9, for each $\omega_{n+1}^{*,j} \in D_{n+1,\eta,\tau} \cap \text{Vor}_{n}(\omega_{n}^{*})$ we can obtain a Dirichlet resonant integer satisfying

$$\frac{R_{n+1}}{4} \leq |k_{n+1}^{*,j}| \leq R_{n+1}$$

and

$$\eta R_{n}^{-2(2+\tau)} \leq |k_{n+1}^{*,j} \cdot \omega_{n+1}^{*,j}| \leq 16\eta^{-2}R_{n}^{-(2-2\tau)}.$$ 

Thus, taking $R_{0}$ large enough (independently of $n$), we obtain the first two properties of the vectors $k_{n+1}^{*,j}$. Statements 4 and 5, 6 and 7 are also proved as in Theorem 9.

Now we prove the last statement. We prove that the segment $S(\Gamma_{k_{n}})$ can only intersect segments which belong to the previous, next to previous and future generations. We first show that $S(\Gamma_{k_{n}})$ cannot intersect any $\rho_{m}$ ball around any frequency of $D_{n,\eta,\tau}$ with $m \geq n + 2$. Take $\omega_{m}^{*} \in D_{n,\eta,\tau}$. Then, the distance

$$\text{dist}(\Gamma_{k_{n}}, \omega_{m}^{*}) = \frac{|k_{n} \cdot \omega_{m}^{*}|}{|k_{n}|} \geq \frac{\eta}{|k_{n}|^{3+\tau}} \geq \frac{1}{R_{n}^{3+\tau}} \geq \rho_{n}^{1+3\tau} > \rho_{n+2}$$

Therefore $\text{dist}(\Gamma_{k_{n}}, \omega_{m}^{*}) > \rho_{m}$ for any $m \geq n + 2$. Since a Dirichlet resonant segment $S(\Gamma_{k})$ associated to a frequency $\omega$ of the generation $m \geq n + 2$ satisfies $S(\Gamma_{k_{n}}) \subset B_{\rho_{n}}(\omega)$, it cannot intersect $S(\Gamma_{k_{n}})$. 

32
Recall that, by statement 6, we know that $S(\Gamma_{k_n}) \subset B_{\rho_n}(\omega_{n-1}^\star)$. Reasoning analogously, the resonant lines associated to frequencies in $D^m_{n,\tau}$ with $m \leq n - 3$ cannot intersect $B_{\rho_n}(\omega_{n-1}^\star)$ and therefore, they cannot intersect $S(\Gamma_{k_n})$ either. This completes the proof of Theorem 1.

Analyzing the resonances belong to a generation $n \geq 1$, we can see that in suitable neighborhoods of the intersections of resonances, there cannot exist other intersections. This fact will be crucial then, in the analysis of the resonances. It is stated in the next two lemmas.

**Lemma 7.** Consider $S(\Gamma_{k'})$ and $S(\Gamma_{k''})$, two of the resonant segments obtained in Theorem 1 which intersect at a point $\omega' = S(\Gamma_{k'}) \cap S(\Gamma_{k''})$. Assume $k$ belongs to the $n$-th generation and $k'$ belongs to the $n + 1$ generation. Then, any other resonant segment $S(\Gamma_k)$ of those obtained in Theorem 1 intersecting the ball $B_{\rho_n^2}(\omega')$ must also contain $\omega'$.

**Lemma 8.** Fix $\theta > 1$, $n \in \mathbb{N}$ and consider $S(\Gamma_{k'})$ one of the resonant segments of the $n$ generation obtained in Theorem 1. Assume that a resonant line $\Gamma_k'$ with $|k''| \leq R_n^\theta$ intersects $S(\Gamma_{k'})$ and call $\omega'$ to the intersection point. Then,

- Any other resonant line $\Gamma_k$ with $|k| \leq R_n^\theta$ intersecting the ball $B_{\rho_n^2(1+2\theta)/3-5\tau}(\omega')$ must also contain $\omega'$.
- Any other resonant segment $S(\Gamma_k)$ of those obtained in Theorem 1 intersecting the ball $B_{\rho_n^2(1+2\theta)/3-5\tau}(\omega')$ must also contain $\omega'$.

We devote the rest of the sections to prove Lemmas 7 and 8.

**Proof of Lemma 7.** Let $\omega' \in S(\Gamma_{k'}) \cap S(\Gamma_{k''})$ be an essential double resonance. Let $|\omega'| \geq \nu$.

By definition $\omega' \cdot k' = \omega' \cdot k'' = 0$. Therefore, $\omega \parallel k' \times k''$, where $\times$ is the wedge product. Let $k^\star$ be a resonance such that $\Gamma_{k^\star}$ passes through the $\rho_n^2$-neighborhood of $\omega'$ but does not contain $\omega'$. Then, $k^\star$ satisfies $|\omega \cdot k^\star| \leq \rho_n^2|k^\star|$. On the other hand, since $k^\star$ is not a linear combination of $k'$ and $k''$ (otherwise $\Gamma_{k^\star}$ would contain $\omega'$, we have that $(k' \times k'') \cdot k^\star \neq 0$. Since they are integer vectors, this implies that $|(k' \times k'') \cdot k^\star| \geq 1$. Then,

$$|\omega' \cdot k^\star| \geq \frac{\nu|k' \times k''| \cdot k^\star|}{|k' \times k''|} \geq \frac{\nu}{|k' \times k''|} \geq \nu R_n^{-(2+2\tau)}.$$ 

Thus, recalling that $|\omega \cdot k^\star| \leq \rho_n^2|k^\star|$ and the definition of $\rho_n$, we obtain

$$\nu R_n^{-(2+2\tau)} \leq \rho_n^2|k^\star| \leq R_n^{-(6-10\tau)}|k^\star|,$$

which implies that $|k^\star| \geq \nu R_n^{4-12\tau}$ and therefore, it cannot belong to the generations $n$ or $n + 1$. Assume that it belongs to a generation $m \geq n + 2$ and we reach a contradiction.

If $k^\star$ belongs to the $m$ generation, we have that $R_m \geq \nu R_n^{4-12\tau}$. Then, using the relation between $R_m$ and $\rho_m$, we can ensure that there is $\omega_m \in D^m_{n,\tau}$ which satisfies $\text{dist}(\Gamma_{k'}, \omega_m) \leq R_m$. 

33
\( \rho_m \leq \rho_n^{4 - 13\tau} \). This implies that \( \omega' \in \mathcal{S}(\Gamma_{k'}) \cap \mathcal{S}(\Gamma_{k''}) \) satisfies \( \text{dist}(\omega', \omega_m) \leq 2\rho_n^2 \). Therefore, \( \text{dist}(\mathcal{S}(\Gamma_{k'}), \omega) \leq 2\rho_n^2 \) and \( \text{dist}(\mathcal{S}(\Gamma_{k''}), \omega) \leq 2\rho_n^2 \). Nevertheless, this violates the Diophantine condition satisfied by \( \omega_m \) and therefore we have reached a contradiction.

**Proof of Lemma 8.** It follows the same lines as the proof of Lemma 7. We take \( \nu \) such that \( |\omega'| \geq \nu \) and \( \Gamma_{k^*} \) which passes through the \( \rho_1 + \theta \)-neighborhood of \( \omega' \) but does not contain \( \omega' \). Then, proceeding as in the proof of Lemma 7, one can see that \( k^* \) must satisfy

\[
\text{dist}(\omega', \Gamma_{k^*}) \geq \frac{\nu}{R_1^{1+\theta}|k^*|} \geq \nu \rho_n^{(1+2\theta)/(3-5\tau)},
\]

This proves the first statement of the lemma. The second one can be proved as in Lemma 7.

**4 The deformation procedure**

We devote this section to explain the different deformations that we need to perform to the Hamiltonian \( H_1 \) in (2). These deformations have been explained in Section 2.

**4.1 The mollification of the Hamiltonian: proof of Lemma 2**

We mollify \( H_1 \) away from KAM tori and we prove Lemma 2. Recall that we have defined the sequence

\[
\rho_{n+1} = \rho_n^{1+2\tau}, \ n \geq 1.
\]

and the union of KAM tori \( \text{KAM}^U_{\eta, \tau} \) in (5). For a positive \( \lambda > 0 \) denote \( U_\lambda(\text{KAM}_{\eta, \tau}) \) the \( \lambda \)-neighborhood of \( \text{KAM}_{\eta, \tau} \) in \( U_\eta' \times \mathbb{T}^2 \times \mathbb{T} \). Consider the annuli

\[
U_{\rho_n}(\text{KAM}_{\eta, \tau}) \setminus U_{2\rho_n^{1+2\tau}}(\text{KAM}_{\eta, \tau}).
\]

For a function \( f \in C^l(\mathbb{R}^n) \), in Lemma 33, we define the family of convolution operators

\[
S_r f(x) = r^{-n} \int_{\mathbb{R}^n} K_r(r^{-1}(x - y)) f(y) dy, \quad 0 < r \leq 1,
\]

where \( K_r \) is an explicitly defined function supported in the ball of radius \( 2\sqrt{r} \) around the origin. Notice that if \( f \) is periodic in some variables, then so is \( S_r f \).

Take \( r = \rho_1^{1+4\tau} \) and

\[
H^1_\varepsilon = H_0 + \varepsilon H^1_1,
\]

where

\[
H^1_1(I, \varphi, t) := \begin{cases} 
H_1(I, \varphi, t) & \text{if } (I, \varphi, t) \notin U_{2\rho_1}(\text{KAM}_{\eta, \tau}) \setminus U_{\rho_1^{1+2\tau}}(\text{KAM}_{\eta, \tau}) \\
S_{\rho_1^{1+4\tau}} H_1(I, \varphi, t) & \text{if } (I, \varphi, t) \in U_{\rho_1}(\text{KAM}_{\eta, \tau}) \setminus U_{2\rho_1^{1+2\tau}}(\text{KAM}_{\eta, \tau}) \\
\chi_{\rho_1}(I, \varphi, t) H_1(I, \varphi, t) + (1 - \chi_{\rho_1}(I, \varphi, t)) S_{\rho_1^{1+4\tau}} H_1(I, \varphi, t) & \text{otherwise}
\end{cases}
\]
with a $C^\infty$ function

$$
\chi_{\rho_1} = \begin{cases} 
1 & \text{in } U_{\rho_1}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_1^{1+2\tau}}(\text{KAM}_{\eta,\tau}) \\
0 & \text{in } U_{2\rho_1}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_1^{1+2\tau}}(\text{KAM}_{\eta,\tau})
\end{cases}
$$

smoothly interpolating in between.

By Lemma 33 we have

$$
\|S_{\rho_1^{1+\tau}} H_1 - H_1\|_{c^s} \leq c\|H_1\|_{c^r}(1+4\tau)(r-s) \text{ for any } 0 \leq s \leq r.
$$

Rewrite

$$
\|S_{\rho_1^2} H_1\|_{c^s} \leq c\|H_1\|_{c^r}(1+4\tau)(s-r) \text{ for any } r \leq s.
$$

We have

$$
\|(1 - \chi_{\rho_1})(S_{\rho_1^{1+\tau}} H_1 - H_1)\|_{c^r} \\
\|(1 - \chi_{\rho_1})\|_{c^0} \|S_{\rho_1^2} H_1 - H_1\|_{c^r} + \|(1 - \chi_{\rho_1})\|_{c^r} \|S_{\rho_1^2} H_1 - H_1\|_{c^0} \leq 
\leq c(\rho_1^2)\|H_1\|_{c^r} + \rho_1^r \rho_1^{1+4\tau}\|H_1\|_{c^r} \leq (c(\rho_1^2) + \rho_1^{4\tau})\|H_1\|_{c^r},
$$

where $d(\rho)$ is given by

$$
\|f(x) - S_{\rho} f(x)\|_{c^r} \leq d(\rho)\|f\|_{c^r}.
$$

Since any continuous function on a compact set is uniformly continuous, for any $d > 0$ there is $\rho = \rho(d)$ such that $\max_{|x-y| \leq \rho} |f(x) - f(y)| \leq d$. Denote the inverse function $d(\rho)$.

Now define inductively in $n$: let $r_n = \rho_n^{(1+4\tau)}$ and

$$
H^n_\varepsilon = H_0 + \varepsilon H^n_1,
$$

where

$$
H^n_1(I, \varphi, t) := \begin{cases} 
H^{n-1}_1(I, \varphi, t) & \text{if } (I, \varphi, t) \not\in U_{\rho_n}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_n^{1+2\tau}}(\text{KAM}_{\eta,\tau}) \\
S_{\rho_n^{1+4\tau}} H_1^{n-1}(I, \varphi, t) & \text{if } (I, \varphi, t) \in U_{\rho_n}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_n^{1+2\tau}}(\text{KAM}_{\eta,\tau}) \\
\chi_{\rho_n}(I, \varphi, t) H_1^{n-1}(I, \varphi, t) + (1 - \chi_{\rho_n}(I, \varphi, t)) S_{\rho_n^{1+4\tau}} H_1^{n-1}(I, \varphi, t) & \text{otherwise}
\end{cases}
$$

with the function $\chi_{\rho}$ defined in the same way as before. By definition supports of only two steps of deformation overlap this procedure defines

$$
H_\varepsilon^{\text{mol}} := H_0 + \varepsilon H_1^{\text{mol}} := H_0 + \varepsilon \lim_{n \to \infty} H^n_1
$$

which is a $C^r$ small perturbation of $H_\varepsilon$. Thus, using monotonicity of $d(\rho)$ we can choose $\rho_1$ small enough so that $d(\rho_1^{1+4\tau}) \leq \gamma/2$. This completes the proof of Lemma 2.
4.2 Perturbation along resonant segments

As explained in Section 2, the perturbation along resonances has two different terms. One supported in the neighborhoods of single resonances and the other one supported in the neighborhood of double resonances. In fact, this perturbation is constructed as two infinite sums:

$$\Delta H_{\text{sr}} + \Delta H_{\text{dr}} = \sum_{n=1}^{\infty} \Delta H_{\text{sr}}^n + \Delta H_{\text{dr}}^n. \quad (15)$$

The support of $\Delta H_{\text{sr}}^0$ is localized along the single resonances in $S_k, k \in \mathcal{F}_0$ (see Definition 4). The support of $\Delta H_{\text{dr}}^n$ is localized in the neighborhoods of the double resonances which appear along the resonances in $S_0$. The supports of $\Delta H_{\text{sr}}^n, n \geq 1$ are localized in action space in a neighborhood of the Dirichlet resonances $S_{\omega k-n}$ of order $n$ selected in Key Theorem 1. The supports of $\Delta H_{\text{dr}}^n$ are localized in action space in the neighborhoods of $\theta$-strong double resonances. By construction supports of $\Delta H_{\text{sr}}^n$ (resp. $\Delta H_{\text{dr}}^n$) and $\Delta H_{\text{sr}}^m$ (resp. $\Delta H_{\text{dr}}^m$) are disjoint if $|n-m| > 1$.

Recall that we are dealing with the $C^\infty$ Hamiltonian

$$H'_\varepsilon = H_0 + \varepsilon H_1 + \varepsilon \Delta H_{\text{mod}},$$

given by Lemma 2. While the $C^r$-norm of $H_1 + \Delta H_{\text{mod}}$ is bounded, its $C^{r+\kappa}$-norms with $\kappa > 0$ are bounded by inverse powers of the mollification parameter $\rho^{-\kappa}$ (see Lemma 2).

4.2.1 The perturbation along the resonance of $S_0$

The analysis of the step 0 resonances is done in [BKZ11] and [KZ12]. The step 0 perturbation $\Delta H^{(0)} = \Delta H_{\text{sr}}^0 + \Delta H_{\text{dr}}^0$ is designed to construct diffusion along the zero generation of resonant segments $\{S_k\}_{k \in \mathcal{F}_0}$. Let $\{S_k\}_{k \in \mathcal{F}_0}$ be all single resonant segments of the zero generation. Then

$$\Delta H_{0}^{\text{sr}} = \sum_{k \in \mathcal{F}_0} \Delta H_0^k \circ \Phi_k^{-1}, \quad (16)$$

where $\Phi_k$ is a canonical change of variables which is obtained in [BKZ11] and is defined in $B_{\varepsilon/4}(\mathcal{S}) \times \mathbb{T}^2 \times \mathbb{T}$. As explained in [BKZ11], this change is chosen such that

$$H'_\varepsilon \circ \Phi_k = H_0 + \varepsilon Z_k + \delta \varepsilon R_k$$

for some small $\delta > 0$ and $\|Z_k\|_{C^2} \leq C$, $\|R_k\|_{C^2} \leq C$, for some $C$ independent of $\rho$ and $\varepsilon$.

Let $\omega_{ij} = S_k \cap S_{k_i}, k_i, k_j \in \mathcal{F}_0$ be all double resonances given by non-empty intersections of resonant segments of the first generation with other first generation resonant segments, or other resonant segments creating double strong resonances, as defined in [BKZ11] [KZ12]. Let $E = E(H_0, H_1)$ be a large number. Then

$$\Delta H_{0}^{\text{dr}} = \sum_{k_i, k_j \in \mathcal{F}_0} \Delta H_{0}^{ij} \circ \Phi_{ij}^{-1}, \quad (17)$$
where $\Phi_{ij}$ is obtained in [KZ12]. It is the composition of a canonical change of variables and a rescaling and is defined in $B_{E_{\sqrt{\varepsilon}}(\omega^{ij})} \times \mathbb{T}^2 \times \mathbb{T}$. As explained in [KZ12], is constructed so that

$$H'_{\varepsilon} \circ \Phi_{ij} = K_{ij} + \varepsilon Z_{ij} + \varepsilon^{3/2} R_{ij},$$

where $K_{ij}$ is a positive definite quadratic form, $Z_{ij}$ is a potential depending on a 2-dimensional slow angle $\varphi_{ij}^s$, and $\|R_{ij}\|_{L^2} \leq C$.

After this step 0 deformation, we have the Hamiltonian

$$H^{(0)}_{\varepsilon} := H'_{\varepsilon} + \Delta H_{0}^{dr} + \Delta H_{0}^{sr}. \quad (18)$$

### 4.2.2 The perturbation along resonances of $S_n$ for $n \geq 1$

The rest of the perturbations

$$\Delta H^{(n)} = \Delta H^{sr}_n + \Delta H^{dr}_n \quad (19)$$

are designed to construct diffusion along the generations of resonant segments $\{S_{k_n}^n\}_{k_n \in F_n}$, $n \geq 1$. Each such segment is a Dirichlet resonance of the $n$ generation. We construct these deformations inductively. Assume that one has already modified the Hamiltonian along the Dirichlet resonant segments belonging to the generations 1, 2, ..., $n-1$. We denote by $H^{(n-1)}_{\varepsilon}$ the Hamiltonian after adding the generation 1, 2, ..., $n-1$ deformations. Now we show how to construct the generation $n$ deformation.

Since the Voronoi cells constructed in Lemma 4 are pairwise disjoint, we restrict our discussion to one Voronoi cell $\text{Vor}_n(\omega^*)$ with $\omega^* \in \mathcal{D}_{\eta,\tau}^n$. Denote by $\mathcal{F}_n(\omega^*)$ the set of resonant vectors of the $n$ generation such that the associated resonant segments intersect the Voronoi cell $\text{Vor}_n(\omega^*)$. That is,

$$\mathcal{F}_n(\omega^*) = \left\{ k_n \in \mathcal{F}_n : S_{k_n}^n \cap \text{Vor}_n(\omega^*) \neq \emptyset \right\}.$$

Let $\theta > 1$ be specified later. Then we define

$$\Delta H^{sr,\omega^*}_n = \sum_{k_n \in \mathcal{F}_n(\omega^*)} \Delta H_{k_n}^n \circ \Phi^{-1}_{k_n}, \quad (20)$$

where $\Phi_{k_n}$ is a change of variables defined for $(I, \varphi) \in B_{\rho_n}(S_{k_n}^n) \times \mathbb{T}^2 \times \mathbb{T})$. The functions $\Delta H_{k_n}^n$ are obtained in Section 9 and the change $\Phi_{k_n}$ is obtained through a normal form procedure in Section 5.2. Now we define

$$H^{(n)}_{\varepsilon} := H^{(n-1)}_{\varepsilon} + \sum_{\omega^* \in \mathcal{D}^n_{\eta,\tau}} \Delta H^{sr,\omega^*}_n. \quad (21)$$

Let $W^n_{\omega^*} = \{\omega^j\}_{j \in J_n}$ be the set of all $\theta$-strong double resonances inside the Voronoi cell $\text{Vor}_n(\omega^*)$ except the ones involving Dirichlet resonant segments belonging to the previous generations. Define

$$\Delta H^{dr,\omega^*}_n = \sum_{j \in J_n} \Delta H_j^n \circ \Phi^{-1}_j \quad (21)$$

37
where the function $\Delta H^\delta_n$ is obtained in Appendix B. The change of variables $\Phi_j$ is defined for $(I, \varphi) \in B_{\rho_n}(\omega^j) \times T^2 \times T)$. This change of coordinates is a proper mollification of Bounemoura Normal Form and is obtained in Section 5.3 (see Theorem 12 which is based on [Bou10]). As explained in that section, the change $\Phi_j$ is constructed such that

$$\tilde{H}^{(n)} \circ \Phi_j = K_j + Z_j + R_j,$$

where $K_j$ is Hamiltonian which only depends on the actions, $Z_j$ is function which only depends on the actions and the 2-dimensional slow angle $\varphi^j_s$, and $\|R_j\|_{C^2} \leq C\rho_n^{q(r+1)}$ (see Section 5.3).

Thus, now we consider the Hamiltonian

$$H^{(n)}_{\delta} := \tilde{H}^{(n)} + \sum_{\omega^* \in D^\eta_\tau} \Delta H^{sr,n}_{\omega^*} + \sum_{\omega^* \in D^\eta_\tau} \Delta H^{dr,n}_{\omega^*},$$

(22)

which is the Hamiltonian after the generation $n$ deformation.

5 Different regimes and normal forms while approaching Diophantine frequencies

In this section we want to analyze the Hamiltonian (8) when we approach the Diophantine frequencies through the resonant segments obtained in Theorem 1. To this end, we fix some $\omega^* \in D^\eta_\tau$ for some $n \geq 1$ and we analyze a small neighborhood of it. Recall that the analysis of the generation 0 resonances is done in [BKZ11] and [KZ12].

As we have explained in Section 2 (see also Section 3), we approach the Diophantine frequencies through Dirichlet resonant segments, which intersect at double resonances. To analyze them, we cover the resonant segments with balls centered at strong double resonances (see Definition 7). Then, in each of these balls we will perform certain normal forms.

First in Section 5.1 we consider one of the resonant vectors $k_n$ obtained in Theorem 1. Then, we show that the associated resonant segment $I_{\omega^* k_n}$ in action space can be covered by neighborhoods of double resonances of sufficiently low order. We divide these neighborhoods into two parts where we perform different normal forms.

In Section 5.2 we analyze the so-called transition zones, which are annulus around the double resonances not containing them. That is, they contain points which are not very close to the double resonances. We proceed analogously as the analysis of the transition zones in [KZ12]. Nevertheless, here we need more accurate normal forms. In Section 5.3 we analyze what we call the core of the double resonances, which are a small ball around the double resonances.

Thanks to these normal form procedure, we obtain good first orders of the Hamiltonian system in the neighborhoods of those double resonances. Later, in Section 6 we use these normal forms and new first orders to prove the existence of a normally hyperbolic invariant cylinders along the resonances in the different regimes.
5.1 Covering the Dirichlet segments by neighborhoods of strong double resonances

We need to analyze the Hamiltonian (8) in a neighborhood of the Dirichlet resonant segments $\mathcal{I}_{\omega_n}^k$ obtained in Theorem 1 (recall that $\mathcal{I}_{\omega_n}^k$ is a resonant segment in action space and $\mathcal{I}_{\omega_n}^k$ is the same resonant segment in frequency space). To analyze this segment we follow the approach in [KZ12]. Nevertheless, instead of considering single and double resonance regimes, here we cover $\mathcal{I}_{\omega_n}^k$ by a sequence of balls centered at $\theta$-strong double resonances. This spares us the analysis of the single resonance regime and allows us to deal only with normal forms around double resonances.

In this section, we fix $k = k_n$ and consider the associated resonant segment $\mathcal{I}_k = \mathcal{I}_{\omega_n}^k$. Recall $\mathcal{S}(\Gamma_k)$ belongs to the ball $B_{3\rho_{n-1}}(\omega_{n1})$ centered at a frequency of the previous generation $\mathcal{D}_{n-1}^{n-1}$. To simplify notation, we define $\rho = \rho_{n-1}$.

First we need to show that double resonances of low order are dense enough in $\mathcal{I}_k$.

**Lemma 9.** Consider the resonant segment $\mathcal{I}_k$ and consider the double resonant points

$$P_{k'} = \mathcal{I}_k \cap \{k' \cdot \omega = 0\}$$

for $|k'| \leq K$. Then, for each such point $P_{k'}$ there exists another point $P_{k''}$ such that

$$|P_{k'} - P_{k''}| \leq CK^{-2}|k|$$

for some constant $C$ independent of $k$ and $K$.

**Proof.** To prove this lemma we make first a change of variables in the frequency space to put $\Gamma_k$ as a horizontal line. We define the new frequencies $v_1 = k \cdot \omega$, $v_2 = \omega_2$ and $v_3 = \omega_3 = 1$. Now, the new frequency $v = (v_1, v_2, 1) \in \Gamma_k$ if $v_1 = 0$. A double resonance corresponds to $v_2 \in \mathbb{Q}$. Now we can apply Dirichlet Theorem to ensure that for any $v_2 \in \mathcal{S}(\Gamma_k)$ there exist integers $q$ and $p$ with $|q| \leq K$ such that

$$\left| v_2 - \frac{p}{q} \right| \leq \frac{1}{K^2}.$$

Thus, in the $v$ coordinates, we obtain that the double resonances are at a distance at most $2K^{-2}$. Undoing the change of coordinates and measuring the distance of the double resonances in the original frequency space we obtain the wanted estimates. Recall that the frequency map $\Omega$ has estimates independent of $\varepsilon$ and $\rho$ and therefore, the estimates in frequency and action space are equivalent.

We use this lemma to cover the resonant segment $\mathcal{I}_k$ by neighborhoods centered at double resonances. We denote balls centered at the double resonance points $P_{k'}$ and radius $\mu$ as

$$B_\mu(P_{k'}) = \{|I - P_{k'}| \leq \mu\}.$$
Now, according to the definition of strong double resonance, we take
\[ K = R^\theta_n = R^\theta_{n-1} = \rho^{-\theta \frac{1+2r}{3+5r}}. \]

Then, Lemma 9 ensures us that \( I_k \) satisfies
\[ I_k \subset \bigcup_{|k'| \leq R_n^\theta} B_\mu(P_{k'}) \]
with
\[ \mu = C^3 \rho^{(2\theta - 1) \frac{1+2r}{3+5r}}. \] (23)

for some constant \( C_3 > 0 \).

We fix a Dirichlet resonant segment \( I_k = I_{k_n}^n \) of the \( n \) generation. As explained in Section 2.1, we fix \( \theta > 0 \) and we consider all the \( \theta \)-strong double resonances (see Definition 7). Then, following Lemma 8, we consider balls of radius \( \rho_n^\theta \) centered at such double resonances. Then, we know that such disk does not contain other \( \theta \)-strong double resonances.

As explained in Section 2, we divide these disks into two different parts: the core of the double resonance and the transition zones. The first one is the \( B_{C\rho^{m/2}}(P_{k'}) \) (see (9)). This region is studied in Section 5.3. The second is what we call transistion zone, the annulus \( A(k_n, k') = \{ I \in B_\mu(P_{k'}) : c \rho^{m/2} \leq |I - P_{k'}| \leq \mu \} \) (see (10) and recall that \( m = r/10 \) has been defined in (9)). The constant \( c \) is chosen so that \( c > C \) to have overlap with (9). In Section 5.2 we study neighborhoods along single resonance which belong to the transition zones of two consecutive double resonaces.

5.2 Normal form in the transition zones

We study the Hamiltonian (8) in certain open sets in the transition zones (10) following [KZ12]. Here we perform several steps of normal form instead of just one, since we need an additional smallness of the remainder.

First we add a varialbe \( E \) conjugate to time to Hamiltonian (8) to have an autonomous Hamiltonian system, as done in [BKZ11]. We define the Hamiltonian
\[ H(I, E, \varphi, t) = N'(I, \varphi, t) + E \]
In this section we abuse notation and we take \( I = (I_1, I_2, E) \) and \( \varphi = (\varphi_1, \varphi_2, t) \). Now this Hamiltonian can be split as \( H = H_0 + H_1 \) where
\[ H_0(I) = H'_0(I_1, I_2) + E \]
\[ H_1(I, \varphi) = R(I_1, I_2, \varphi_1, \varphi_2, t) + \Delta H_1^{\mathrm{mol}} \circ \Phi_{\text{Pöss}}(I_1, I_2, \varphi_1, \varphi_2, t). \]

We perform a translation to assume that the double resonance is located at \( I = 0 \) and we perform the following rescaling to the action variables
\[ I = \rho^{m/2} J \]
where $m = r/10$ has been defined in (9). This change is conformally symplectic. Then, after a time rescaling, Hamiltonian (9) becomes

$$\mathcal{H}(J, \varphi) = \mathcal{H}_0(\rho^{m/2}J) + \mathcal{H}_1(\rho^{m/2}J, \varphi).$$

(24)

Denote by $\tilde{A}(k_n, k')$ the rescaled annulus. We perform several steps of normal form to this Hamiltonian as in [KZ12]. Consider two consecutive annulus $\tilde{A}(k_n, k')$ and $\tilde{A}(k_n, k'')$ which, by construction, overlap. We do not perform the normal form in the union of these annulus but in a strip along the single resonance,$$
\mathcal{D}_{C_0}(I_{k_n}) = \{ |\partial_J(\mathcal{H}_0(\rho^{m/2}J)) \cdot k_n | \leq C_0 \} \cap (\tilde{A}(k_n, k') \cup \tilde{A}(k_n, k'')). $$

Note that in the original coordinates, this strip would be of width of order $C_0 \rho^{m/2}$. One can easily see that in such set, the Hamiltonian (24) satisfies

$$\|\mathcal{H}_0\|_{C^{r-6-2\tau}} \leq C, \quad \|\mathcal{H}_1\|_{C^{r-6-2\tau-m}} \leq C\rho^m$$

and

$$\|\mathcal{H}_0\|_{C^{r-6-2\tau+\kappa}} \leq C\rho^{-(1+4\tau)\kappa} \quad 0 < \kappa \leq 6 + 2\tau$$
$$\|\mathcal{H}_1\|_{C^{r-6-2\tau-m+\kappa}} \leq C\rho^{m-(1+4\tau)\kappa} \quad 0 < \kappa \leq 6 + 2\tau + m$$

(25)

**Theorem 11.** Consider the Hamiltonian (24). There exists a change of coordinates $\Phi : \mathcal{D}_C(I_{k_n}) \times T^3 \rightarrow \mathbb{R}^2 \times T^3$ such that $\mathcal{H} \circ \Phi$ is of the form

$$\mathcal{H} \circ \Phi(J, \varphi) = \tilde{\mathcal{H}}_0(J) + \mathcal{Z}(J, k_n \cdot \varphi) + \mathcal{R}(J, \varphi)$$

(26)

where $\tilde{\mathcal{H}}_0(J) = \mathcal{H}_0(\rho^{m/2}J)$ and the other terms satisfy

$$\|\mathcal{Z}\|_{C^{r-6-2\tau-3\kappa}} \leq C\rho^{\kappa-3m-(1+4\tau)\kappa}, \quad \|\mathcal{R}\|_{C^{r-6-2\tau-3m+\kappa}} \leq C\rho^{(r+1)-(1+4\tau)\kappa},$$

for some constant $C$ independent of $\rho$ and for any $\kappa$ such that $0 \leq \kappa < r - (r - 6 - 2\tau - 3m) = 6 + 2\tau + 3m$.

Moreover,

$$\|\Phi - \text{Id}\|_{C^{r-6-2\tau-3m}} \leq C\rho^{3-3m-(1+4\tau)\kappa},$$

for any $\kappa$ such that $0 \leq \kappa < r - (r - 6 - 2\tau - 3m) = 6 + 2\tau + 3m$.

This theorem gives the normal form in Key Theorem 2. It only suffices to see in this proof of Theorem 11 that the normal form procedure always keeps the form $H(J, \varphi, t) + E$ and therefore it can be only be taken back to a non-autonomous Hamiltonian.
5.2.1 Proof of Theorem 11

The proof of this proposition is obtained performing several steps of normal form. We do them through an iterative process. We modify the strategy in [KZ12]. Here we divide the steps into two parts. In the first parts will reduce the size of the low harmonics (to be defined properly later) and then, in the second part we will reduce the size of the high harmonics. Each part requires several steps of normal form.

We define as low harmonics, those which satisfy $|\tilde{k}| \leq \frac{R^4}{5n}$. First, we show that, taking $\rho \ll \eta$, in the original coordinates (before rescaling) the low harmonic resonances are 100$\rho$ away from the ball $B_{3\rho}(\omega_n^*)$. Indeed, for such $\tilde{k}$,

$$
\text{dist}(\Gamma_{\tilde{k}}, \omega_n^*) = \frac{|\tilde{k} \cdot \omega_n^*|}{|\tilde{k}|} \geq \frac{\eta}{3^4 + \tau} \geq \eta R_n^{-4(3^4 + \tau)/5}.
$$

Then, recalling the relation between $\rho = \rho_n$ and $R_n$ given in (6) and taking $\rho$ small enough with respect to $\eta$, we have that

$$
\text{dist}(\Gamma_{\tilde{k}}, \omega_n^*) \geq 100\rho.
$$

Then, $\Gamma_{\tilde{k}}$ is 50$\rho$ away from the annulus where we are doing the normal form.

We define the projections associated to low and high harmonics ($\pi \leq f$)($J, \varphi$) = $\sum_{k \in \mathbb{Z}^4, 0 < |k| \leq R^4_n} f^{[k]}(J)e^{2\pi ik \cdot \varphi}$ and $\pi > f = f - \pi \leq f$.

First we reduce the size of the low harmonics by performing some steps of averaging. We perform the averaging procedure only to the low harmonics. That is, we do not remove any high harmonic term. This will be done later on. We see that at step $j$ we have a Hamiltonian $H^j$ of the form

$$
H^j(J, \varphi) = H^j_0(J) + R^j_\leq(I, \varphi) + R^j_> (I, \varphi)
$$

where $R^j_\leq$ only contains low harmonics (that is $\pi \leq R^j_\leq = R^j_\leq$) and $R^j_>$ only contains high harmonics (that is $\pi > R^j_\geq = R^j_\geq$). Note that all the $k_n$-resonant harmonics belong to $R^j_>$. In the first step, we take $R^0_\leq = \pi \leq H_1$, $R^0_\geq = \pi > H_1$ and $H^0_0(J) = H_0(\rho^{m/2}J)$.

We prove inductively that at the $j$ step, the terms in the Hamiltonian satisfy

$$
\|H^j_0 - H^0_0\|_{C^{r-6-2r-2j}} \leq C\rho^{j^m} \tag{27}
$$

$$
\|R^j_\geq\|_{C^{r-6-2r-2j}} \leq C\rho^m \tag{28}
$$

$$
\|R^j_\leq\|_{C^{r-6-2r-2j}} \leq C\rho^{m + \frac{1}{2}r}jm. \tag{29}
$$

In the step 0, these estimates are trivially satisfied.
To reduce the size of $\mathcal{R}_\xi^j$, we perform a change of coordinates defined by the time one map of the flow associated to the Hamiltonian

$$\Gamma^j(J, \varphi) = \sum_{k \in \mathbb{Z}^3, 0 < |k| \leq R^{5/4}_n} \frac{\mathcal{R}_\xi^j(J)}{\partial J \mathcal{H}_0^j(J) \cdot k} e^{ik \varphi}.$$ 

We bound this Hamiltonian. To bound the small divisors, we do it first for the zero-iteration Hamiltonian $\mathcal{H}_0^0(J) = \mathcal{H}_0(\rho^{m/2} J)$. By Theorem \[\Box\] we have that for a resonance $\Gamma_k$ with $|k| \leq R^{5/4}_n$, in the original coordinates $I \in B_{\xi \rho}(\omega^*_n)$,

$$\frac{|\partial I \mathcal{H}_0^0(I) \cdot k|}{|k|} = \text{dist}(\Gamma_k, B_{\xi \rho}(\omega^*_n)) \geq 50 \rho,$$

which implies $|\partial I \mathcal{H}_0^0(I) \cdot k| \geq 50 \rho |k|$. Thus, after rescaling, we obtain

$$\left| \left( \partial J \left( \mathcal{H}_0(\rho^{m/2} J) \right) \cdot k \right)^{-1} \right| \leq 50 \rho^{-\frac{m}{2}} |k|^{-1}.$$ 

Taking derivatives, one can easily see that for $\ell \in \mathbb{Z}^2$ with $|\ell| = 0, \ldots, |r - 7 - 2\tau - 2j|$,

$$\left| \partial J \left( \mathcal{H}_0^j(J) \cdot k \right)^{-1} \right| \leq 50 \rho^{-\frac{m}{2}} |k|^{-1}.$$ 

Then, using the inductive hypothesis \[\[\Box\]\], one has that for $\ell \in \mathbb{Z}^2$ with $|\ell| = 0, \ldots, r - 6 - 2\tau - 2j - 1$,

$$\left| \partial J \left( \mathcal{H}_0^j(J) \cdot k \right)^{-1} \right| \leq 50 \rho^{-\frac{m}{2}} |k|^{-1}.$$ 

Thus, we can conclude

$$\left\| \left( \partial J \mathcal{H}_0^j(J) \cdot k \right)^{-1} \right\|_{C^r-6-2\tau-2j-1} \leq C \rho^{-\frac{m}{2}} \leq C \rho^{-\frac{2m}{\tau}} |k|^{-1}.$$ 

We use these estimates, to bound $\Gamma^j$.

$$\left\| \Gamma^j \right\|_{C^r-6-2\tau-2j-1} \leq C \rho^{-\frac{2m}{\tau}} \|\mathcal{R}_\xi^j\|_{C^m-2j} \leq C \rho^{\frac{1}{2}(j+1)m}.$$ 

We define $\Phi^j$ the flow associated to the Hamiltonian $\Gamma^j$ and the time one map $\Phi^j = \Phi^j_1$. Using the Faa di Bruno formula, one can see that

$$\left\| \Phi^j - \text{Id} \right\|_{C^r-6-2\tau-2(j+1)} \leq C \left\| \Gamma^j \right\|_{C^r-6-2\tau-2j-1} \leq C \rho^{\frac{1}{2}(j+1)m}.$$ 

Now, we analyze the new Hamiltonian $\mathcal{H}^{j+1} = \mathcal{H}^j \circ \Phi^j$. To this end, we define $F^j_t = \mathcal{R}^j_\xi + t \mathcal{R}^j_\leq$. Then, $\mathcal{H}^j$ can be written as

$$\mathcal{H}^{j+1}(J, \varphi) = \mathcal{H}_0^j + \mathcal{R}^j_\xi + \int_0^1 \{F^j_t, \Gamma^j\} \circ \Phi^j_t dt.$$ 

43
We define then,
\[ H^{j+1}_0 = H^j_0 + \left\langle \int_0^1 \{ F^j_t, \Gamma^j \} \circ \Phi^j_t dt \right\rangle \]
\[ R^{j+1}_\leq = \pi \int_0^1 \{ F^j_t, \Gamma^j \} \circ \Phi^j_t dt - H^{j+1}_0 \]
\[ R^{j+1}_> = \pi > + \int_0^1 \{ F^j_t, \Gamma^j \} \circ \Phi^j_t dt. \]

We use the estimates for \( R_j \leq, R_j > \) and \( \Gamma_j \) to bound the Poisson Bracket appearing in these terms by
\[ \| \{ F^j_t, \Gamma^j \} \|_{C^{r-6-2\tau-2(j+1)}} \leq C \rho^{m+\frac{1}{2}(j+1)m}. \]

Then, applying the Faa di Bruno formula, one obtains analogous bounds for \( \| \{ F^j_t, \Gamma^j \} \circ \Phi^j_t \|_{C^{r-6-2\tau-2(j-1)}} \).

Using them, one can easily deduce the desired estimates for \( H^{j+1}_0, R^{j+1}_\leq \) and \( R^{j+1}_> \).

We perform \( m/2 \) steps of normal form. With our choice of parameters, we have that \( m^2/6 > qr \). Therefore, we obtain a Hamiltonian
\[ H^{m/2}(J, \varphi) = H^{m/2}_0(J) + R^{m/2}_\leq(J, \varphi) + R^{m/2}_>(J, \varphi) \]
satisfying
\[ \| H^{m/2}_0 - H^0_0 \|_{C^{r-6-2\tau-m}} \leq C \rho^{4m} \]
\[ \| R^{m/2}_\leq \|_{C^{r-6-2\tau-m}} \leq C \rho^{q(r+1)} \]
\[ \| R^{m/2}_> \|_{C^{r-6-2\tau-m}} \leq C \rho^m. \]

Moreover, using the high regularity estimates of \( H_0 \) and \( H_1 \) given in (25), proceeding inductively, one can easily see that
\[ \| H^{m/2}_0 - H^0_0 \|_{C^{r-6-2\tau-m+\kappa}} \leq C \rho^{4m-\kappa(1+4\tau)} \]
\[ \| R^{m/2}_\leq \|_{C^{r-6-2\tau-m+\kappa}} \leq C \rho^{q(r+1)-\kappa(1+4\tau)} \]
\[ \| R^{m/2}_> \|_{C^{r-6-2\tau-m+\kappa}} \leq C \rho^{m-\kappa(1+4\tau)}. \]

for \( \kappa = 1 \ldots 6 + 2\tau + m \).

Now, we perform additional steps of normal form to \( H^{m/2} \) to reduce the size of certain terms involving high harmonics. To this end, we split it in more terms. We define slightly different projections. We take
\[ \tilde{\pi}_k f(J, \varphi) = \sum_{j \in \mathbb{Z}} f^{jk_n}(J) e^{2\pi j k_n \cdot \varphi} \]
\[ \tilde{\pi}_> f(J, \varphi) = \pi> f(J, \varphi) - \tilde{\pi}_k f(J, \varphi) \]
and \( \tilde{\pi}_\leq = \pi_\leq \). Then, we rewrite \( \tilde{\mathcal{H}} = \mathcal{H}^{m/2} \) as

\[
\tilde{\mathcal{H}}(J, \varphi) = \tilde{\mathcal{H}}_0(J) + \tilde{Z}(J, k_n \cdot \varphi) + \tilde{R}_\leq(J, \varphi) + \tilde{R}_>(J, \varphi)
\]

where

\[
\tilde{\mathcal{H}}_0(J) = \mathcal{H}_0^{m/2}(J) \\
\tilde{R}_\leq(J, \varphi) = \mathcal{R}_\leq^{m/2}(J, \varphi) \\
\tilde{Z}(J, k_n \cdot \varphi) = \tilde{\pi}_k \mathcal{R}_\leq^{m/2}(J, \varphi) \\
\tilde{R}_>(J, \varphi) = \tilde{\pi}_\mathcal{R}_\leq^{m/2}(J, \varphi) = \mathcal{R}_\leq^{m/2}(J, \varphi) - \tilde{Z}(J, k_n \cdot \varphi).
\]

We do some more steps of normal form to reduce the size of \( \tilde{R}_> \). Nevertheless, before doing them we take advantage of the fact that \( \tilde{Z} \) and \( \tilde{R}_> \) only contain high harmonics and therefore, decreasing the regularity, it has better estimates. Indeed, we have that

\[
\| \tilde{R}_\leq \|_{C^{r-6-2r-m}} \leq C\rho^m \\
\| \tilde{Z} \|_{C^{r-6-2r-m}} \leq C\rho^m \\
\| \tilde{R}_> \|_{C^{r-6-2r-m}} \leq C\rho^m.
\]

implies

\[
\| \tilde{R}_\leq \|_{C^{r-6-2r-2m}} \leq C\rho^m \\
\| \tilde{Z} \|_{C^{r-6-2r-2m}} \leq C\rho^{m+\frac{1}{5}r} \leq C\rho^{m+\frac{1}{5}m} \\
\| \tilde{R}_> \|_{C^{r-6-2r-2m}} \leq C\rho^{m+\frac{1}{5}r} \leq C\rho^{m+\frac{1}{5}m}.
\]

Note that decreasing regularity, \( \tilde{R}_\leq \) does not reduce its norm, since it contains low harmonics. Using the better estimates for the high harmonics we perform more steps of normal form to reduce the size of \( \mathcal{R}_> \). Now at the step \( j \) we will have a Hamiltonian of the form

\[
\tilde{\mathcal{H}}^j(J, \varphi) = \tilde{\mathcal{H}}_0(J) + \tilde{Z}(J, k_n \cdot \varphi) + \tilde{R}_\leq(J, \varphi) + \tilde{R}_>(J, \varphi)
\]

which is \( C^{r-6-2r-m-2j} \) and satisfies

\[
\| \tilde{\mathcal{H}}^j - \mathcal{H}_0 \|_{C^{r-6-2r-2m-2j}} \leq C\rho^{\frac{1}{5}m} \quad (30) \\
\| \tilde{Z}^j \|_{C^{r-6-2r-2m-2j}} \leq C\rho^{m+\frac{1}{5}m} \quad (31) \\
\| \tilde{R}_\leq^j \|_{C^{r-6-2r-2m-2j}} \leq C\rho^{m+\frac{1}{5}m} \quad (32) \\
\| \tilde{R}_>^j \|_{C^{r-6-2r-2m-2j}} \leq C\rho^{\rho r} \quad (33)
\]

We proceed as before composing \( m/2 \) changes of coordinates \( \tilde{\Phi}^j \) defined as the time one map associated to the flow \( \tilde{\Phi}^j_{\mathcal{H}_0} \) of Hamiltonians \( \tilde{\mathcal{H}}^j \). We define at each step

\[
\Gamma^j(J, \varphi) = \sum_{k \in \mathbb{Z}^3 \setminus k_n \mathbb{Z}} \frac{\mathcal{R}^{j,k}_{\mathcal{H}_0}(J)}{\partial J \mathcal{H}_0^j(J) \cdot k} e^{ik \cdot \varphi}.
\]

45
Now, to estimate the small divisors, we need to take into account that in the original coordinates the puncture are at a distance $\rho^{m/2}$ of the annulus $A_{k'}$ where we are doing the normal form (and the same for the nearby annulus $A_{k''}$ and the corresponding puncture). Then, proceeding as before and using the inductive hypothesis (30), one can see that
\[
\left\| \left( \partial_J \tilde{H}_0^j(J) \cdot k \right)^{-1} \right\|_{C^{r-6-2r-2m-2j-1}} \leq C\rho^{-m}.
\]

Then, by (32), we have that
\[
\|\tilde{\Gamma}^j\|_{C^{r-6-2r-2m-2j-1}} \leq C\rho^{-m} \|\mathcal{R}^j\|_{C^{r-6-2r-2m-2j-1}} \leq C\rho^{\frac{1}{2}jm}.
\]

and
\[
\|\tilde{\Phi}^j - \text{Id}\|_{C^{r-6-2r-2m-2(j+1)}} \leq C\|\Gamma^j\|_{C^{r-6-2r-2m-2j-1}} \leq C\rho^{\frac{1}{2}jm}.
\]

Now, taking $\tilde{F}_t^j = \tilde{Z}^j + \tilde{R}^j_\leq + t\mathcal{R}^j_\geq$, the transformed Hamiltonian $\tilde{H}^{j+1} = \tilde{H}^j \circ \tilde{\Phi}^j$ is given by
\[
H^{j+1}(J, \varphi) = H_0^j + \tilde{Z}^j + \tilde{R}^j_\leq + \int_0^1 \{\tilde{F}_t^j, \tilde{\Gamma}^j\} \circ \tilde{\Phi}_t^j dt.
\]

We define then,
\[
\tilde{H}_0^{j+1} = \tilde{H}_0^j + \left< \int_0^1 \{F_t^j, \Gamma^j\} \circ \Phi_t^j dt \right>
\]
\[
\tilde{Z}^j_\leq = \pi_{k_n} \left( \int_0^1 \{F_t^j, \Gamma^j\} \circ \Phi_t^j dt \right) - \tilde{H}_0^{j+1}
\]
\[
\tilde{R}^j_\leq = \pi_\leq \left( \int_0^1 \{F_t^j, \Gamma^j\} \circ \Phi_t^j dt \right)
\]
\[
\tilde{R}^j_\geq = \mathcal{R}_\geq + \pi_\geq \left( \int_0^1 \{F_t^j, \Gamma^j\} \circ \Phi_t^j dt \right) - \tilde{H}_0^{j+1}.
\]

Now proceeding as before, one can easily see that the new terms are $C^{r-6-2r-2m-2(j+1)}$ and satisfy the estimates (30), (33), (32) and (31).

Thus, taking $\mathcal{Z} = Z^{m/2}$ and $\mathcal{R} = R^{m/2}_\leq + R^{m/2}_\geq$, we obtain the Hamiltonian stated in Theorem 11. One can easily keep track of the high regularity norms to obtain the needed estimates.

### 5.3 Bounemoura Normal Form near in the core of the double resonances

Consider $\mu$-neighborhoods of a double resonance of period $T$ such that $T\mu < 1$. We derive a normal form describing behavior near a double resonance, proposed by Bounemoura [Bou10]. Denote by
\[
\mathcal{D}_R = B_R(I^*) \times \mathbb{T}^2 \times \mathbb{T}.
\]
Call a frequency $\omega^\# \in \mathbb{R}^2$ rational if

$$T = \inf \{ t > 0 : \ t(\omega^\#, 1) \in \mathbb{Z}^3 \setminus 0 \} < \infty$$

is well defined. By definition $T$ is the period of the corresponding periodic orbit of the unperturbed flow. Recall that $m = r/10$ has been defined in (9). Below $p$ stands for the number of steps of averaging.

**Theorem 12.** Let $p, m$ be integers with $2 \leq p < r - 6 - 2\tau$ and $\tau > 0$ satisfies $4\tau(p - 1) < 1$. For any $0 < \sigma, \rho < 1$ and $0 < \mu \leq C\rho^{m/2}$ consider a $C^\infty$ smooth Hamiltonian of the form

$$\mathcal{H}(I, \varphi, t) = \mathcal{H}_0(I) + \mathcal{H}_1(I, \varphi, t)$$

for any $0 \leq \kappa < 6 + 2\tau + m$ satisfying

$$\|\mathcal{H}_0\|_{C^{r - 6 - 2\tau - m + \kappa}(\mathcal{D}_\mu(I^*)))} \leq \sigma^{-\kappa} \quad \text{and} \quad \|\mathcal{H}_1\|_{C^{r - 6 - 2\tau - m + \kappa}(\mathcal{D}_\mu(I^*)))} \leq \rho^m \sigma^{-\kappa}.$$

Suppose $I^*$ is such that the frequency vector $\omega^\# = \nabla\mathcal{H}_0(I^*)$ is rational of some period $T$ such that $T\rho^m < C$ and $|\omega^\#| < C$. Then there exists a $C^{r - 6 - 2\tau - p}$ symplectic transformation

$$\Phi_B^p : \mathcal{D}_{\mu/2}(I^*) \rightarrow \mathcal{D}_\mu(I^*)$$

with $\|\Pi_1 \Phi_B^p - \text{Id}\|_{C^0(\mathcal{D}_{\mu/2}(I^*))} \leq C (T\rho^m)$ such that

$$\mathcal{H} \circ \Phi_B^p = \mathcal{H}_0 + \mathcal{Z} + \rho^m \mathcal{R}$$

with $\{\mathcal{Z}, \omega^\# I\} = 0$, and for any $0 \leq \kappa < (4 + 2\tau)m$ we have

$$\|\mathcal{Z}\|_{C^{r - 6 - 2\tau - p + \kappa}(\mathcal{D}_{\mu/2}(I^*))} \leq C \rho^m \sigma^{-\kappa},$$

$$\|\mathcal{R}\|_{C^{r - 6 - 2\tau - p + \kappa}(\mathcal{D}_{\mu/2}(I^*))} \leq C (T\rho^m)^p \sigma^{-\kappa},$$

and

$$\|\Phi_B^p - \text{Id}\|_{C^{r - 6 - 2\tau - p + \kappa}(\mathcal{D}_{\mu/2}(I^*))} \leq C (T\rho^m).$$

Notice that $\{\omega^\# I, \mathcal{Z}\} = 0$ implies that, expanding $\mathcal{Z}$ in a Fourier series, it contains only resonant harmonics, namely, integers $k \in (\mathbb{Z}^2 \setminus \{0\}) \times \mathbb{Z}$ given by $k \cdot (\omega^\#, 1) = 0$. By Key Theorem II any vector $k$ such that $\omega^\# \cdot k = 0$ satisfies $|k| \geq R_n^{1 - \tau}$. Then, since $\|\mathcal{H}_1\|_{C^{r - 6 - 2\tau - m}} \leq C\rho^m$, $\rho = \rho_n = R_n^{-(3 - 5\tau)}$ we have that its Fourier coefficients $h_k$ decay as $|h_k| \leq c, \rho^m |k|^{-(r - 6 - 2\tau - m)}$. Therefore,

$$\|\mathcal{Z}\|_{C^0} \leq c \rho^m |k|^{-(r - 6 - 2\tau - m)} \leq c \rho^m R_n^{1 - (r - 6 - 2\tau - m)} = c \rho^m R_n^{1 - (r - 6 - 2\tau - m)} = c \rho^{m + 1, r - 6 - 2\tau - m}.$$ 

**Warning:** The resonant term $\mathcal{Z}$, in general, is not determined by resonant Fourier coefficients of $\mathcal{H}_1$. Indeed, after one step of averaging we have remainders of order $\rho^{2m}$, while resonant Fourier coefficients of $\mathcal{H}_1$ are bounded by $c \rho^m R_n^{1 + (r - 6 - 2\tau - m)}$. For our range of parameters the former dominates the latter.
Corollary 5. In notations of Theorem 12 let $\omega^\# \in \mathcal{S}(\Gamma_k) \cap \mathcal{S}(\Gamma_k')$ be $\theta$-strong double resonance. Let $\mu = \rho^{m/2}$. Then there exists a $C^{r-6-2r-m-p}$ symplectic transformation

$$
\Phi_B^p : \mathcal{D}_{\mu/2}(I^*) \longrightarrow \mathcal{D}_{\mu}(I^*)
$$

with $\|\Pi_I \Phi_B^p - \text{Id}\|_{C^0(\mathcal{D}_{\mu}(I^*))} \leq C \rho^{m/2-(1+\theta)(1+2r)/3}$ such that

$$
\mathcal{H} \circ \Phi_B^p = \mathcal{H}_0 + \mathcal{Z} + \rho^m \mathcal{R}
$$

with $\{\mathcal{Z}, \omega^\#I\} = 0$ we have

$$
\|\mathcal{Z}\|_{C^{r-6-2r-p}(\mathcal{D}_{\mu}(I^*))} \leq C \rho^{m+(1+2r)(r-6-2r-p)/3},
$$

and for any $0 \leq \kappa < (3 + \tau)(m + 1)$

$$
\|\mathcal{R}\|_{C^{r-6-2r-p+\kappa}(\mathcal{D}_{\mu}(I^*))} \leq C \rho^{p(m-(1+\theta)(1+2r)/3)-\kappa(1+4r)}.
$$

Finally,

$$
\|\Phi_B^p - \text{Id}\|_{C^{r-6-2r-p+\kappa}(\mathcal{D}_{\mu}(I^*))} \leq C \rho^{m-(1+\theta)(1+2r)/3-\kappa(1+4r)}.
$$

In particular, for $\kappa = 6 + 2\tau + p$, we have

$$
\|\mathcal{R}\|_{C^r(\mathcal{D}_{\mu}(I^*))} \leq C \rho^{p(m-(1+\theta)(1+2r)/3-((3+\tau)(m+1)+p)(1+4r))}.
$$

(34)

Proof of Corollary: Since the double resonance is $\theta$-strong, we have that $\rho^{-(1+2r)/3} < \rho^{-(1+2r)/3}$. Notice that a periodic orbit of the unperturbed system has period

$$
T \leq |k| \cdot |k'| < R^{1+\theta} \leq \rho^{-(1+\theta)(1+2r)/3}.
$$

By Theorem 12 $\|\mathcal{Z}\|_{C^{r-6-2r-p}} \leq C \rho^m$ and by construction $\mathcal{Z}$ has only resonant Fourier coefficients whose indices given by integer combinations of $k$ and $k'$. This implies that

$$
\|\mathcal{Z}\|_{C^{r-6-2r-m}} \leq C \rho^{(1+2r)(r-6-2r-m)/3+m}.
$$

This implies the claim of the Corollary. \hfill \Box

5.3.1 Application of Corollary 5

We would like to deform

$$
\rho^m \mathcal{R} \circ (\Phi_B^p)^{-1}
$$

to zero outside of the $\rho^\theta$-neighborhood of $I^*$. In order to keep perturbation $C^r$ small it suffices to have

$$
\|\rho^m \mathcal{R} \circ (\Phi_B^p)^{-1}\|_{C^0} \leq \|\rho^m \mathcal{R} \circ (\Phi_B^p)^{-1}\|_{C^r} \leq C \rho^{\theta(r+1)}.
$$
Choose \( p \) so that the following inequality holds
\[
p(m - (1 + \theta)(1 + 2\tau)/3 - ((3 + \tau)(m + 1) + p)(1 + 4\tau) = m(p - 3\tau) - \frac{(4 + \theta)\tau}{3} p \geq \vartheta(r + 1).
\]
\[
\text{(35)}
\]

Denote
\[
z = m - (1 + \theta)(1 + 2\tau)/3 + (1 + 4\tau)(s - p) = r 1_r r - 4_r m - \frac{1 + \theta}{3} 1_r - 3_r - 1_r p,
\]
where \( =_r \) means that we make error \( c_r \) with \( c \) independent of \( \tau \). We prove the following

**Lemma 10.** In the above setting we have
\[
\| \rho^m \mathcal{R} \circ (\Phi_B^p)^{-1} \|_{C^r} \leq 2c\rho^{m+(1+p)z-(r+1)(1+4\tau)}.
\]

In the case
\[
m + (1 + p)z - (r + 1)(1 + 4\tau) \geq \vartheta(r + 1) \iff
\]
\[
\begin{align*}
\frac{p-\vartheta}{p+1} r & \geq \frac{\vartheta + 1_r}{p+1} + \left(4_r - \frac{1}{p+1} \right) m + \frac{1 + \theta}{3} 1_r + 3_r + 1_r p
\end{align*}
\]
we have that \( \rho^m \mathcal{R} \circ (\Phi_B^p)^{-1} \) can be deformed to zero with a \( C^r \) small perturbation localized in \( \rho^\vartheta \)-neighborhood. We shall choose \( m \geq p \). If \( p > 10\vartheta \), then \( r > 6m \) and large enough suffices.

**Proof:** By Lemma \[35\] we have
\[
\| \rho^m \mathcal{R} \circ (\Phi_B^p - \text{Id})^{-1} \|_{C^r} \leq \| \rho^m \mathcal{R} \|_{C^r} + \| \rho^m \mathcal{R} \circ (\Phi_B^p - \text{Id})^{-1} \|_{C^r}.
\]
\[
\leq \rho^m \| \mathcal{R} \|_{C^r} + c \rho^m \left( \| \mathcal{R} \|_{C^r} \| (\Phi_B^p - \text{Id})^{-1} \|_{C^r} + \sum_{g=2}^{r} \| \mathcal{R} \|_{C^r} \sum_{j_1 + \cdots + j_g = r} \| (\Phi_B^p - \text{Id})^{-1} \|_{C^{j_1}} \cdots \| (\Phi_B^p - \text{Id})^{-1} \|_{C^{j_g}} \right).
\]
By Corollary \[5\] we have
\[
\| \Phi_B^p - \text{Id} \|_{C^{r-(3+\tau)(m+1)-p+\kappa}(\mathcal{D}_\mu(I^*))} \leq C \rho^{m-(1+\theta)(1+2\tau)/3-\kappa(1+4\tau)}.
\]
We also have
\[
\| \mathcal{R} \|_{C^{r-(3+\tau)(m+1)-p+\kappa}(\mathcal{D}_\mu(I^*))} \leq C \rho^{m-(1+\theta)(1+2\tau)/3-\kappa(1+4\tau)}.
\]
This implies that
\[
\| \Phi_B^p - \text{Id} \|_{C^{\kappa}(\mathcal{D}_\mu(I^*)))} \leq C \rho^{z-(1+4\tau)}.
\]
and \[ \| \mathcal{R} \|_{C^j(D_{\mu}(I^*))} \leq C \rho^{p^2 - j(1 + 4\tau)}. \]

By Lemma 37 we have that \[ \| (\Phi_B^p - \text{Id})^{-1} \|_{C^j} \leq C \rho^{z - j(1 + 4\tau)}. \]

Notice that for \( j > z/(1 + 4\tau) \) exponents is negative and estimate blows up. Now substitute these estimates into \[ \| R \|_{C^r} + c(\| R \|_{C^1} \| (\Phi_B^p - \text{Id})^{-1} \|_{C^r} + \sum_{g=2}^r \| R \|_{C^g} \sum_{j_1 + \ldots + j_a = r} \| (\Phi_B^p - \text{Id})^{-1} \|_{C^{j_1}} \ldots \| (\Phi_B^p - \text{Id})^{-1} \|_{C^{j_a}}). \]

We have the following upper bound
\[
c \left( \rho^{(1+p)z-(r+1)(1+4\tau)} + \sum_{g=1}^r \rho^{pz+gz-(r+1)(1+4\tau)} \right) \leq c \rho^{(1+p)z-(r+1)(1+4\tau)} (1 + \sum_{g=1}^r \rho^{pz}) \leq 2c \rho^{(1+p)z-(r+1)(1+4\tau)}. \]

5.3.2 Proof of Theorem 12

We follow the same strategy as in the proof of Proposition 3.2 [Bou10]. To simplify the exposition we define the parameter \( s = r - 6 - 2\tau - m \).

Before we proceed with averaging we need to write the Hamiltonian in the form: the linear Hamiltonian plus a remainder (see (**), sect. 3.2 [Bou10]). We have
\[ H_0(I) + H_1(I, \varphi, t) = H_0(I^*) + \omega^# \cdot I + Q_2(I - I^*) + H_1(I, \varphi, t), \]
where \( Q_2(I - I^*) = H_0(I) - H_0(I^*) + \omega^# \cdot I \) is the quadratic remainder. It can be bounded by \( C\mu^2 \leq C\rho^m \) for \( \mu \leq \rho^m/2 \).

Denote by \( \Phi_{H_0}^t \) the time \( t \) map of the new integrable Hamiltonian \( \widetilde{H}_0(I) = \omega^# \cdot I \). The change of coordinates will be given by a composition of \( p \) averaging transformations. Define \( \Delta \mu = \rho^m/2p \) and let \( \mu_j = \rho^m - j\Delta \mu \). We show that for any \( j \in \{0, 1, \ldots, p\} \) there exists a canonical transformation \( \Phi_j : D_{\mu_j}(I^*) \to D_{\rho^m}(I^*) \) with
\[
\| \Phi_j - \text{Id} \|_{C^{r-6-2\tau-m+k}(D_{\mu_j}(I^*))} \leq C(T\rho^m) \sigma^{-\kappa} \quad \text{for any } 0 \leq \kappa \leq 6 + 2\tau + m + j
\]
and such that
\[ \mathcal{H} \circ \Phi_j = \mathcal{H}_0 + \mathcal{Z}_j + \mathcal{R}_j, \]
where \( Z_j \) has \( \{ Z_j, \omega^\# \cdot I \} = 0 \) and along with \( R_j \) satisfies the estimates
\[
\| R \|_{C^{r-6-2r-\kappa}(D_\mu(I^*))} \leq C(T \rho^m)^j \rho^m \sigma^{-\kappa},
\]
\[
\| Z \|_{C^{r-\kappa}(D_\mu(I^*))} \leq C \rho^m \sigma^{-\kappa}.
\]
Theorem follows now taking \( j = p - 1 \). We shall prove now the claim by induction in \( j = \{ 0, 1, \ldots, p - 1 \} \). For \( j = 0 \) there is nothing to prove. We simple write
\[
Z_0 = 0, R_0 = H^1.
\]
Now assume that the claim is true for some \( j_0 \in \{ 0, 1, \ldots, p - 1 \} \), and consider
\[
H \circ \Phi_j = H_0 + Z_j + R_j.
\]
Define the average
\[
[R_j] = \frac{1}{T} \int_0^T R_j \circ \Phi_{H_0}^T dt
\]
and
\[
\chi_j = \frac{1}{T} \int_0^T t(R_j - [R_j]) \circ \Phi_{H_0}^T dt.
\]
For any \( 0 \leq \kappa \leq 6 + 2\tau + m + j \) we have
\[
\| [R_j] \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq \| R_j \circ \Phi_{H_0}^T \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))}.
\]
Therefore, we have
\[
\| [R_j] \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq \| R_j \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))}
\]
and by inductive hypothesis gives
\[
\| [R_j] \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq C \| R_j \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq C(T \rho^m)^j \rho^m \sigma^{-\kappa}.
\]
Similarly,
\[
\| [\chi_j] \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq C(T \rho^m)^{j+1} \sigma^{-\kappa}.
\]
Since \( T \rho^m \leq \rho^{m-(1+\theta)(1+2\tau)/3} \) is small, we have that for any \( 0 \leq \kappa \leq 6 + 2\tau + m + j \)
\[
\| [\chi_j] \|_{C^{r-6-2r+\kappa-j}(D_\mu(I^*))} \leq C \rho^{(j+1)[m-(1+\theta)(1+2\tau)/3]} \sigma^{-\kappa}.
\]
Let \( \Phi^{\chi_j} \) be the time one map of the Hamiltonian vector field given by \( \chi_j \), then we need to show that the map
\[
\Phi_{j+1} = \Phi_{j+1} \circ \Phi^{\chi_j}
\]
satisfies the assumptions. Since \( \rho^{m-(1+\theta)(1+2\tau)/3} \) is small, we can check that \( \Phi^{\chi_j} \) has the \( C^r \)-norms small and, therefore, is a well-defined embedding
\[
\Phi^{\chi_j} : D_{\mu_{j+1}}(I^*) \rightarrow D_{\mu_j}(I^*). \]
Moreover, having
\[ \|X_{\chi_j}\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq C\rho^{m-(1+\theta)(1+2\tau)/3}\sigma^{-\kappa} \]
and using smallness of \( \rho \) long with Lemma 36 for any \( 0 \leq \kappa \leq 6 + 2\tau + m + j \) we get
\[ \|\Phi_{\chi_j} - \text{Id}\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq C\rho^{m-(1+\theta)(1+2\tau)/3}\sigma^{-\kappa}. \]
Now
\[ \|\Phi_{j+1} - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} = \]
\[ \|\Phi_j \circ \Phi_{\chi_j} - \Phi_{\chi_j} + \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq \]
\[ \|\Phi_j - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} + \|\Phi_{\chi_j} - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq \]
\[ \|\Phi_j - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} + \|\Phi_{\chi_j} - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))}. \]
here we used Lemma 35. By inductive hypothesis, this gives
\[ \|\Phi_{j+1} - \text{Id}\|_{C^{r-6-2\tau+\kappa-1-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq C\rho^{m-(1+\theta)(1+2\tau)/3}\sigma^{-\kappa}. \]
Now using Taylor’s formula with integral remainder, we get
\[ H_{j+1} = H \circ \Phi_{j+1} = H_0 + Z_{j+1} + R_{j+1} \]
where
\[ Z_{j+1} = Z_j + [R_j], \quad R_{j+1} = \rho^{-m} \int_0^1 \{Z_j + R^t_j, \chi_j\} \circ \Phi_{t}^{H_0} dt \]
where \( R^t_j = tR_j + (1-t)[R_j] \). Since \( \{H_0, Z_j\} = 0 \) by our hypothesis and \( \{H_0, [R_j]\} = 0 \) by construction, we have \( \{H_0, Z_{j+1}\} = 0 \). For any \( 0 \leq \kappa \leq 6 + 2\tau + m + j \) we also obtain an estimate
\[ \|Z_{j+1}\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j}})} \leq \|Z_j\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j}})} + \rho^m \|R_{j+1}\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j}})} \leq \]
\[ \leq C\rho^{m-(1+\theta)(1+2\tau)/3}\rho^{-\kappa(1+4\tau)} + C\rho^{j-(1+\theta)(1+2\tau)/3}\rho^{-\kappa} \leq \]
\[ \leq 2C\rho^{m-(1+\theta)(1+2\tau)/3}\rho^{-\kappa}. \]
Estimate now the remainder \( R_{j+1} \) we follows
\[ \|R_{j+1}\|_{C^{\kappa-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \leq \|\{Z_j + \rho^m R^t_j, \chi_j\} \circ \Phi_{\chi_j}^t\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j+1}}(I^*)}. \]
By Lemma 35 can upper bound by
\[ c\|\{Z_j + \rho^m R^t_j, \chi_j\}\|_{C^{r}(\mathcal{D}_{\nu_{j+1}}(I^*))} \|\Phi_{\chi_j}^t\|_{C^{r-6-2\tau+\kappa-j}(\mathcal{D}_{\nu_{j+1}}(I^*))} \]
\[ + c \sum_{p=2}^\infty \|\{Z_j + \rho^m R^t_j, \chi_j\}\|_{C^{p}(\mathcal{D}_{\nu_{j+1}}(I^*))} \times \]
\[ 52 \]
\[ \times \sum_{j_1+\cdots+j_p=r-6-2\tau+\kappa-j} \| \Phi_t^{\chi_j} \|_{C^{r_1}((D_{\mu_{j+1}}(I^*)) \cdots \| \Phi_t^{\chi_j} \|_{C^{r_p}(D_{\mu_{j+1}}(I^*))} \] \\
\leq C \| \{ Z_j + \rho^m R_j^t, \chi_j \} \|_{C^{r_6-2r+\kappa-j}(D_{\mu_j}(I^*))} \sum_{p=1}^{r-6-2r+\kappa-j} (T \rho^m)^{p(j+1)} \rho^{(r-6-2r+\kappa-j)(1+4r)} \\
\leq 2C(T \rho^m)^{(j+1)} \rho^{(r-6-2r+\kappa-j)(1+4r)} \| \{ Z_j + R_j^t, \chi_j \} \|_{C^{r_6-2r+\kappa-j}(D_{\mu_j}(I^*))} \\
\leq C \| \{ Z_j, \chi_j \} \|_{C^{r_6-2r+\kappa-j}(D_{\mu_j}(I^*))} + C \{ \mathcal{R}_j^t, \chi_j \} \|_{C^{r_6-2r+\kappa-j}(D_{\mu_j}(I^*))} \\
\leq C \rho^m (T \rho^m)^{j+1} \sigma^{-\kappa} + C (T \rho^m)^{j} (T \rho^m)^{j+1} \sigma^{-\kappa} \\
\leq 2C \rho^m (T \rho^m)^{j+1} \sigma^{-\kappa} \leq 2C \rho^m (\rho^{m-(1+\theta)(1+2r)/3})^{j+1} \sigma^{-\kappa}. \]

5.4 End of the proof of Key Theorem 4

Once we have obtained the normal form in the core of double resonances (Theorem 12) and killed the remainder by the deformation procedure (Corollary 5), last step is to change variables to have o have explicitly in the Hamiltonian the slow and fast variables.

We define

\[ (\varphi_1^a, \varphi_2^a, \varphi_3^a) = (k \cdot \varphi, k' \cdot \varphi, \varphi_3) \]

which can be written in matrix notation as

\[ \begin{pmatrix} \varphi_1^a \\ \varphi_2^a \\ \varphi_3^a \end{pmatrix} = A \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} k \\ k' \\ e_3 \end{pmatrix} \] (37)

where \( e_3 = (0, 0, 1) \). To have a symplectic change of coordinates it is easy to see that we have to perform the change of coordinates

\[ \begin{pmatrix} J_1^a \\ J_2^a \\ J_3^a \end{pmatrix} = A^{-T} \begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} \] (38)

to the conjugate actions. To simplify notation, we call \( J \) to \( J = (J_1^a, J_2^a, J_3^a) \). It can be seen that such matrix and its inverse, satisfy

\[ \| A \|, \| A^{-1} \| \lesssim \rho^{-2(1+\tau+\theta)} \] (39)

and the same bounds are satisfied by their transposed matrices.
Lemma 11. If we apply the symplectic change of coordinates \((37), (38)\) to the Hamiltonian \((26)\), we obtain a Hamiltonian of the form

\[
\hat{\mathcal{H}}(J, \varphi) = \hat{\mathcal{H}}_0(J) + \hat{\mathcal{Z}}(J, \varphi^s) + \hat{\mathcal{R}}(J, \varphi^s, \varphi^f, t),
\]

where \(\hat{\mathcal{Z}}\) only depends on the slow angles and \(\hat{\mathcal{R}}\) depends on all three angles, and they satisfy the following bounds

\[
\left\| \hat{\mathcal{Z}} \right\|_{c^2} \leq \rho^{m+\frac{1}{4}k_0^2(r-2\tau-2m)-4(\frac{1}{4}-\tau-\theta)}
\]

and

\[
\left\| \hat{\mathcal{R}} \right\|_{c^2} \leq \rho^q.
\]

The proof of this lemma is straightforward. This lemma completes the proof of Key Theorem 4.

6 The normally hyperbolic cylinders in the transition zones

In this section we prove Key Theorem 3 by proving the existence of normally hyperbolic cylinders in the transition zones along the Dirichlet resonant segment \(S_{k_0}^{\omega_n}\) given by Theorem 9. We use the normal forms that we have obtained in Theorem 11. We follow the techniques developed in [BKZ11]. Nevertheless, the approach in that paper needs to be modified since now the different parameters involved satisfy different relations.

6.1 Change to slow-fast variables

First we perform a change of variables to Hamiltonian \((26)\) separate the slow and fast angles. This change was also done in [BKZ11]. Nevertheless, now this change is \(\rho_n\)-dependent. Therefore, we need to be more accurate in the estimates. On the other hand, note that in the present setting we only have one slow angle. To simplify notation, in this section we take \(\rho = \rho_n\).

We define

\[
\begin{pmatrix}
\varphi^s \\
\varphi^f \\
t
\end{pmatrix} = \tilde{A} \begin{pmatrix}
\varphi_1 \\
\varphi_2 \\
t
\end{pmatrix}
\]

with \(\tilde{A} = \begin{pmatrix} k^s \\ e_2 \\ e_3 \end{pmatrix}\) \((41)\)

where \(e_i\) are the standard coordinate vectors (if this change is singular, replace some of the coordinate vectors). To have a symplectic change of coordinates, one has to perform the change of coordinates

\[
\begin{pmatrix}
J^s \\
J^f_1 \\
J^f_2 \\
E
\end{pmatrix} = \tilde{A}^{-T} \begin{pmatrix}
J_1 \\
J_2 \\
E
\end{pmatrix}
\]

\((42)\)
to the conjugate actions. \( E \) is the variable conjugate to time. It can be easily seen that \( E \) is not modified when the change (42) is performed. We call \( J \) to \( J = (J^s, J^f) \). It can be seen that the matrix \( \tilde{A} \) and its inverse, satisfy
\[
\|\tilde{A}\|, \|\tilde{A}^{-1}\| \lesssim \rho^{-\frac{1}{3}}
\]
and the same bounds are satisfied by their transposed matrices. Next lemma gives estimates for the transformed Hamiltonian.

**Lemma 12.** If we apply the symplectic change of coordinates (41), (42) to the Hamiltonian (26), we obtain a Hamiltonian of the form
\[
\hat{\mathcal{H}}(J, \varphi) = \hat{\mathcal{H}}_0(J) + \hat{Z}(J, \varphi^s) + \hat{R}(J, \varphi^s, \varphi^f, t),
\]
where \( \hat{Z} \) only depends on the slow angle, \( \hat{R} \) depends on all three angles, and they satisfy the following bounds
\[
\left\| \hat{Z} \right\|_c \leq C_5 \rho^\frac{m-2}{s} \leq C_5 \rho^m
\]
and
\[
\left\| \hat{R} \right\|_c \leq C_5 \rho^{q(r+1)-\frac{2}{s}} \leq C_5 \rho^{q(r+1)-1}.
\]
Moreover, \( \hat{\mathcal{H}}_0 \) satisfies
\[
\frac{1}{2} D^{-1} \rho^{m+1} \text{Id} \leq \frac{1}{2} D^{-1} \rho^{m+\frac{2}{s}+\tau} \text{Id} \leq \partial^2 \mathcal{H}_0(J) \leq 2D \rho^{m-\frac{2}{s}-\tau} \text{Id} \leq 2D \rho^{m-1} \text{Id}
\]
where \( D \) is the constant introduced in (1).

In this system of coordinates the resonant vector has become \( \mathbf{k}_n = (1, 0, 0) \). We prove the existence of a cylinder along this resonance for the Hamiltonian (44). The resonance \( S_{\omega_k^n} \) is now defined by
\[
\partial_{J^s} \hat{\mathcal{H}}_0(J) = 0.
\]
Since \( \partial_{J^s} \hat{\mathcal{H}}_0(J) \neq 0 \) along the resonance. The resonant segment \( S_{\omega_k^n} \) can be parameterized as a graph as \( J^s = J^s(J^f) \) for \( J^f \in [b^-_{k_n}, b^+_{k_n}] \), for some \( b^-_{k_n} < b^+_{k_n} \).

### 6.2 Existence of the cylinders

Consider the averaged potential \( \hat{Z}(J, \varphi^s) \) in (44) for \( J^f \in [b^-_{k_n}, b^+_{k_n}] \). We can look at it along the resonance, that is, we can take \( \hat{Z}(J^s(J^f), J^f, \varphi^s) \). Call a value \( J^f \in [b^-_{k_n}, b^+_{k_n}] \) regular if \( \hat{Z}(J^s(J^f), J^f, \varphi^s) \) has a unique global maximum on \( T^s \supset \varphi^s \) at some \( \varphi^{s*} = \varphi^s(J^f) \). We say the maximum is non-degenerate if
\[
\partial_{\varphi^s, \varphi^{s*}} \hat{Z}(J^s(J^f), J^f, \varphi^s) < 0.
\]
This maximum depends smoothly on $J^f$ and can be extended to a bigger interval $\varphi^s_*(J^f) : [a_- - \lambda, a_+ + \lambda] \to \mathbb{T}^2$, in which is a local maximum but might not be the global one. From the deformation procedure done in \cite{9} we have proved that $\widehat{Z}$ has a regular maximum in the needed interval, and that it satisfies

$$-\partial_{\varphi^s_*, \varphi^s} \widehat{Z}(J^s_*(J^f), J^f)(\varphi^s) \geq \rho^{dr}.$$  

**Lemma 13.**

The key idea, used in \cite{BKZ11}, \cite{KZ12}, is to construct an isolating block around

$$\{\varphi^s(J^f)\} \times \{J^s(J^f)\} \times [a_- - \lambda, a_+ + \lambda] \times \mathbb{T} \ni (\varphi^s, J^s, \varphi^f, J^f, t).$$

We shall at some occasions lift the map $\varphi^s_*$ to a $C^2$ map taking values in $\mathbb{R}$ without changing its name. To simplify notations, we will be using the $O(\cdot)$ notation, where $f = O(g)$ means $|f| \leq C g$ for a constant $C$ independent of $\rho, \lambda, \delta$, and $r$. In particular, we will not be keeping track of the parameter $D$, which is considered fixed throughout the paper.

**Theorem 13.** There exists a $C^1$ map

$$(\Theta^s, P^s)(\varphi^f, J^f) : \mathbb{T}^2 \times \mathcal{S}(\Gamma_k) \to \mathbb{T} \times \mathbb{R}$$

such that the cylinder

$$\mathcal{C} = \{(\varphi^s, J^s) = (\Theta^s, P^s)(\varphi^f, J^f); \ J^s_*(p^f) \in \mathcal{S}(\Gamma_k), \ \varphi^f \in \mathbb{T}^2\}$$

is weakly invariant with respect to the vector field associated to the Hamiltonian \cite{44}, in the sense that the vector field is tangent to $\mathcal{C}$. The cylinder $\mathcal{C}$ is contained in the set

$$V := \{(\varphi, J^f) \in \mathcal{S}(\Gamma_k), \quad \|\varphi^s - \varphi^s_*(J^f)\| \leq M' \rho^{m-1+2dr}, \quad \|J^s - J^s_*(J^f)\| \leq M' \rho^{m-1+2dr}\},$$

for some constant $M'$ independent of $\rho$ and it contains all the full orbits of \cite{14} contained in $V$. We have the estimates

$$\|\Theta^s(\varphi^f, J^f) - \varphi^s_*(J^f)\| \lesssim M' \rho^{m-1+5dr},$$

$$\|P^s(\varphi^f, J^f, t) - J^s_*(J^f)\| \lesssim M' \rho^{m-1+5dr},$$

$$\left\|\frac{\partial \Theta^s}{\partial J^f}\right\| \lesssim M' \rho^{\frac{3}{2}m-\frac{m+1}{2}}, \quad \left\|\frac{\partial \Theta^s}{\partial \varphi^f}\right\| \lesssim \rho^{\frac{3}{2}m-\frac{m+1}{2}}.$$
6.3 Proof of Theorem 13

Theorem 13 is a consequence of Theorem 4.1 in [BKZ11]. Nevertheless we need to slightly modify the system associated to the Hamiltonian to be able to apply that theorem since the different parameters involved in [BKZ11] satisfy different relations. Moreover, recall that the change (42) makes convexity $\rho$-dependent.

We redo the proof in [BKZ11]. We start estimating different quantities and computing its $\rho$ dependence. The Hamiltonian flow admits the following equations of motion

\[
\begin{align*}
\dot{\varphi}^s &= \partial_{J^s} \mathcal{H}_0 + \partial_{J^s} \mathcal{I} + \partial_{J^s} \mathcal{R} \\
\dot{J}^s &= -\partial_{\varphi^s} \mathcal{I} - \partial_{\varphi^s} \mathcal{R} \\
\dot{\varphi}^f &= \partial_{J^f} \mathcal{H}_0 + \partial_{J^f} \mathcal{I} + \partial_{J^f} \mathcal{R} \\
\dot{J}^f &= -\partial_{\varphi^f} \mathcal{R} \\
\dot{t} &= 1,
\end{align*}
\]  

we denote this vector field by $F$.

First we look for a good first order. Since in the present setting is much smaller than the size of the angle dependent part of the Hamiltonian, the first order considered in [BKZ11] cannot be used. Instead, we consider a different one. This new first order will allow us to look for a system of coordinates suitable for the isolating block procedure.

\[
\begin{align*}
\varphi^s &= \partial_{J^s} \mathcal{H}_0 + \partial_{J^s} \mathcal{I} \\
\dot{J}^s &= -\partial_{\varphi^s} \mathcal{I} \\
\dot{\varphi}^f &= \partial_{J^f} \mathcal{H}_0 \\
\dot{J}^f &= 0 \\
\dot{t} &= 1
\end{align*}
\]  

Now, due to Lemma 13 for any fixed $J^f \in [J^f_i - \delta, J^f_{i+1} + \delta]$, the point $(\varphi^s(J^f), J^s(J^f))$ is a hyperbolic critical point for the system

\[
\dot{\varphi}^s = \partial_{J^s} \mathcal{H}_0 + \partial_{J^s} \mathcal{I}, \quad \dot{J}^s = -\partial_{\varphi^s} \mathcal{I}.
\]

Moreover, the differential associated to the the hyperbolic point is given by

\[
M(J^f) = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}
\]

where

\[
\begin{align*}
a &= \partial_{\varphi^s J^s} \mathcal{I}(\varphi^s(J^f), \hat{J}^s(J^f)) \\
b &= \partial_{J^s J^f} \mathcal{H}_0(\hat{J}^s(J^f)) + \partial_{J^s J^s} \mathcal{I}(\varphi^s(J^f), \hat{J}^s(J^f)) \\
c &= -\partial_{\varphi^s \varphi^s} \mathcal{I}(\varphi^s(J^f), \hat{J}^s(J^f))
\end{align*}
\]  

57
Thanks to Lemma 13, these coefficients satisfy
\[ |a| \lesssim \rho^{8m-1} \quad \text{and} \quad 0 < \rho^{d-1} \lesssim b \lesssim \rho^{m+1} \]
\[ 0 < \rho^{dr} \lesssim c \lesssim \rho^{8m-1}. \]  \quad (50)

The eigenvalues of the hyperbolic critical points are given by \( \pm \lambda \) where
\[ \lambda = \sqrt{a^2 + bc} \]  \quad (51)

Therefore, they satisfy
\[ \rho^{d+\frac{m+1}{2}} \lesssim \lambda \lesssim \rho^{10m-2}. \]

We diagonalize the matrix associated to these critical points. To this end we define the matrix
\[ S = \begin{pmatrix} a + \sqrt{a^2 + bc} & -b \\ c & a + \sqrt{a^2 + bc} \end{pmatrix} \]

First, we give some estimates on the critical point \((\varphi^*(J^f), J^s_*(J^f))\) and the matrix \(S\).

**Lemma 14.** The hyperbolic critical point \((\varphi^*(J^f), J^s_*(J^f))\) and the matrix \(S\) satisfy the following estimates
\[ |\partial_{J^f} \varphi^*(J^f)|, |\partial_{J^f} J^s_*(J^f)| \lesssim \rho^{8m-dr-3} \]
and
\[ |S| \lesssim \rho^{m-1}, \quad |S^{-1}| \lesssim \rho^{-dr-1} \]
\[ |\partial J^f S| \lesssim \rho^{3m-\frac{d}{2}dr-7}, \quad |\partial J^f S^{-1}| \lesssim \rho^{3m-\frac{d}{2}dr-2}. \]

**Proof.** The estimates for the critical point can be obtained applying implicit derivation, since one obtains
\[ M \begin{pmatrix} \partial_{J^f} \varphi^* \\ \partial_{J^f} J^s_*(J^f) \end{pmatrix} = \begin{pmatrix} -\partial_{J^f J^s_*} \hat{H}_0 - \partial_{J^f, J^f_*} \hat{Z} \\ \partial_{J^f, \varphi^*} \hat{Z} \end{pmatrix} \]
where the right hand side is evaluated at the critical points. Thus,
\[ \begin{pmatrix} \partial_{J^f \varphi^*} \\ \partial_{J^f J^s_*} \end{pmatrix} = \frac{1}{a^2 + bc} \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \begin{pmatrix} -\partial_{J^f J^s_*} \hat{H}_0 - \partial_{J^f, J^f_*} \hat{Z} \\ \partial_{J^f, \varphi^*} \hat{Z} \end{pmatrix} \]

We have that \( \det M \gtrsim \rho^{m+dr+1} \). Then, using also the estimates for \(a, b\) and \(c\) in (50), one obtains the desired bounds.

The estimate for \(|S|\) is straightforward. For \(S^{-1}\), one just needs to take into account that \(\det S = 2a^2 + 2bc + 2a\sqrt{a^2 + bc}\) satisfies \(\det S \geq 2bc \gtrsim \rho^{m+dr+1}\).

Now, we compute the bounds for the derivatives \(\partial_{J^f} S\) and \(\partial_{J^f} S^{-1}\). To this end, we need to compute the derivative for \(a, b, c\) and \(S\). By the definition of \(a\) in (19), we have that
\[ |\partial_{J^f} a| \leq \left\| \hat{Z} \right\|_{C^3} (1 + |\partial_{J^f} J^s_*| + |\partial_{J^f \varphi^*}|) \lesssim \rho^{16m-dr-5} \]
The computations for $\partial_{Jf}c$ is analogous. For $\partial_{Jf}b$ it is enough to recall also that

$$\left| \partial_{J^*Jf} \hat{H}_0 \right| \lesssim \rho^{3m-2}.$$  

Therefore, we obtain also that $|\partial_{Jf}b| \lesssim \rho^{\frac{3}{10}m-\frac{3}{2}dr-5}$.

Using these estimates and also (50), we obtain the bound for $\partial_{Jf}S$. For $\partial_{Jf}S^{-1}$, we need upper bounds for

$$\partial_{Jf} \det S = \left(2 - \frac{a}{\sqrt{a^2 + bc}}\right) \left(a \partial_{Jf}a + c \partial_{Jf}b + b \partial_{Jf}c + 2 \partial_{Jf}a \sqrt{a^2 + bc}\right)$$

Using the just obtained bounds and (50), one can see that $|\partial_{Jf} \det S| \lesssim \rho^{\frac{3}{10}m-\frac{3}{2}dr-7} \leq \rho^{5m-\frac{3}{2}dr}$. Then, using all these estimates, we obtain that $|\partial_{Jf}S^{-1}| \lesssim \rho^{3m-\frac{3}{2}dr-2}$.

We define the new coordinates

$$\left(\begin{array}{c} x \\ y \end{array} \right) = S^{-1} \left(\begin{array}{c} \varphi^s - \varphi^s(J_f) \\ J_s - J^s(J_f) \end{array} \right)$$

(52)

The inverse change is given by

$$\left(\begin{array}{c} \varphi^s \\ J^s \end{array} \right) = \left(\begin{array}{c} \varphi^s(J_f) \\ J^s(J_f) \end{array} \right) + S \left(\begin{array}{c} x \\ y \end{array} \right)$$

We look for an isolating block in the region $|x| \leq \eta, |y| \leq \eta$ where $\eta$ is a parameter to be determined and depends on $\rho$. We need also to rescale the other variables,

$$I = \nu^{-1} J_f$$

$$\Theta = \gamma \varphi^f$$

(53)

where the parameters $\nu, \gamma \ll 1$ will be determined later.

**Lemma 15.** Assume that $|x| \leq \eta, |y| \leq \eta$. Then, the following estimates are satisfied

$$|\varphi^s - \varphi^s(J_f)| \lesssim \rho^{m-1}\eta,$$

$$|J^s - J^s(J_f)| \lesssim \rho^{m-1}\eta$$

**Lemma 16.** In the new variables, the equation (47) takes the following form

$$\left(\begin{array}{c} \dot{x} \\ \dot{y} \end{array} \right) = \Lambda \left(\begin{array}{c} x \\ y \end{array} \right) + O \left(\rho^{\frac{3}{10}m+(q-2d)r-1} + \rho^{4m+(q-\frac{7}{2}d)r-8}\eta + \rho^{\frac{7}{2}m-\frac{3}{2}dr-7}\eta^2\right)$$

(54)

where $\Lambda = \text{diag}(\lambda, -\lambda)$.  

59
Proof. Applying the change (52), (53), we have
\[
\begin{align*}
(\dot{x}, \dot{y}) = & \partial_{J^f} S^{-1} J^f \left( \varphi^s - \varphi^s_s(J_f) \right) + S^{-1} \left( \dot{\varphi}^s - \dot{\varphi}^s_s \right) + S^{-1} \left( \dot{\varphi}^s - \dot{\varphi}^s_s \right) J^f_s \left( x, y \right) + S^{-1} \left( \partial_{J^f} \varphi^s \right) J^f_s \left( x, y \right) + S^{-1} \left( \partial_{J^f} \varphi^s \right) J^f_s \left( x, y \right) \right).
\end{align*}
\]

We analyze of each three terms. For the first one, we use that \( \dot{J}^f = O(\rho^r) \) and the estimates in Lemmas (14) and (15) to obtain
\[
\left| \partial_{J^f} S^{-1} j^f S \left( \begin{array}{c} x \\ y \end{array} \right) \right| \leq \rho^{4m+(q-2d)r - 8}.\]

Using the same lemmas, for the third term we obtain
\[
\left| S^{-1} \left( \frac{\partial_{J^f} \varphi^s_s}{\partial_{J^f} J^f_s} \right) j^f \right| \leq \rho^{2m+(q-2d)r - 4} \leq \rho^{m+(q-2d)r}.
\]

For the second term, we use equation (17). Indeed, we have that
\[
\begin{align*}
\left( \frac{\dot{\varphi}^s}{\dot{J}^s} \right) = & M \left( \varphi^s - \varphi^s_s \right) + \rho^m \left( \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \right) \left( \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \right) + \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \left( \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \right) + O \left( \rho^r \right) \\
= & M \left( \varphi^s - \varphi^s_s \right) + \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \left( \frac{\partial_{J^f} \varphi^s}{\partial_{J^f} J^f_s} \right) + O \left( \rho^r \right).
\end{align*}
\]

Therefore
\[
S^{-1} \left( \frac{\dot{\varphi}^s}{\dot{J}^s} \right) = \Lambda \left( \begin{array}{c} x \\ y \end{array} \right) + O \left( \rho^r \right)
\]

Using the relation between \( d \) and \( q \), one obtains the desired estimates. \( \square \)

To prove the existence of an isolating block, we need to analyze also the linearized equation. We first analyze the linearization of the change of coordinates (52), (53).

**Lemma 17.** The linearization of the change of coordinates \( (x, y, \Theta, I, t) \to (\varphi^s, J^s, \varphi^f, J^f, t) \) is of the following form
\[
T = \left( \begin{array}{c} S \varphi^s \varphi^f J^f \end{array} \right) = \left( \begin{array}{cccc} S & 0 & 0 & 0 \\ 0 & \nu A & 0 & 0 \\ 0 & 0 & \nu & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)
\]

where
\[
A = \partial_{J^f} \left( \begin{array}{c} \varphi^s_s \\ J^f_s \end{array} \right) + \partial_{J^f} S \left( \begin{array}{c} x \\ y \end{array} \right)
\]
The linearization of the inverse change of coordinates is of the form

\[ T^{-1} = \begin{pmatrix} S^{-1} & 0 & B & 0 \\ 0 & 0 & \gamma & 0 \\ 0 & 0 & \nu^{-1} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \]

where

\[ B = \partial_{J^f} S^{-1} \left( \varphi^s - \varphi^s_0 \right) + S^{-1} \partial_{J^f} \left( \varphi^s_0 \right) \]

Moreover, the norms of \( T \) and \( T^{-1} \) have the following bounds

\[ |T| \lesssim \gamma^{-1} + \rho^{3m-2dr-2} + \rho^{4m-2dr-3} \eta \]

\[ |T^{-1}| \lesssim \nu^{-1} + \rho^{3m-2dr-2} + \rho^{4m-2dr-3} \eta \]

**Proof.** The computation of \( T \) and \( T^{-1} \) is straightforward. To compute their norms we first bound \( A \) and \( B \). By Lemma 14, one can see that \( A \) and \( B \) satisfy

\[ |A| \leq \rho^{3m-2dr-2} + \rho^{4m-2dr-3} \eta \]

\[ |B| \leq \rho^{3m-2dr-2} + \rho^{4m-2dr-3} \eta \]

From these estimates, one can obtain the bounds for \( T \) and \( T^{-1} \), recalling that \( m = r/10 \) (see (9)) and that \( \gamma \ll 1 \) and \( \nu \ll 1 \).

**Lemma 18.** In the coordinate system \((x, y, \Theta, I, t)\), the linearized system is given by the block matrix

\[ L = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} + \mathcal{O} \left( \rho^{m-2\gamma} \right) + \mathcal{O} \left( \rho^{3m-3\gamma} \right) + \mathcal{O} \left( \rho^{4m-3\gamma} \right) + \mathcal{O} \left( \rho^m \right) \]

where

\[ v_1 = \begin{pmatrix} \partial_{J^f} \varphi^s \hat{Z} \\ \partial_{J^f} \varphi^s (\hat{H}_0 + \hat{Z}) \end{pmatrix} \]

\[ v_2 = \begin{pmatrix} \partial_{J^f} \varphi^s \hat{Z} \\ \partial_{J^f} \varphi^s (\hat{H}_0 + \hat{Z}) \end{pmatrix} \]

and \( C = \partial_{J^f} \left( \hat{H}_0 + \hat{Z} \right) \). Therefore, using the estimates (12), we have that

\[ |v_1| \leq \rho^m, \quad |v_2| \leq \rho^m, \quad |C| \leq \rho^m \].

61
The matrix $L$ is given by $L = T^{-1} \cdot DF \cdot T$. Thus, we just need to multiply the matrices. We obtain

$$
L = \begin{pmatrix}
\Lambda + \mathcal{O}(|S^{-1}||S|q^2m\eta) & 0 & S^{-1}((M + \mathcal{O}(q^2m\eta))A + v_2)\nu & 0 \\
\gamma v_1^T S & 0 & (v_1^T A + C)\nu & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} + \mathcal{O}(|T^{-1}||T|q^\nu).
$$

To bound each term, we take advantage of a cancellation which arises in $L$. We define

$$
v_2^* = \left(\frac{\partial_{Jf} J_s}{\partial_{Jf} \varphi_s} + \mathcal{Z}(\varphi_s^*, J_s^*, J_f^*)\right).
$$

Then, we have that

$$
M \left(\frac{\partial_{Jf} \varphi_s^*}{\partial_{Jf} J_s^*}\right) + v_2^* = 0.
$$

Thus, using the definition of $A$ we have that

$$
|S^{-1}((M + \mathcal{O}(q^2m\eta))A + v_2)\nu| \lesssim |S^{-1}||A|q^{2m-2}\eta\nu + |S^{-1}||v_2 - v_2^*|\nu \\
+ |S^{-1}||M||\partial_{Jf} S| \left|\begin{array}{c}
x \\
y
\end{array}\right| \nu.
$$

Then, using the estimates of $v_1$, $v_2$ and $C$, the fact that $\gamma \ll 1$ and $\nu \ll 1$, the estimates of $A$ given in Lemma 17 and the estimates given in Lemma 14, we obtain

$$
L = \begin{pmatrix}
\Lambda & 0 \\
0 & 0
\end{pmatrix} + \mathcal{O}(|T^{-1}||T|q^{\nu}, q^{2m-2}\gamma, q^{\frac{5m-2m-2}{2}\eta}, q^{\frac{3m-2m-2}{2}\nu\gamma}, q^{4m-2\nu\gamma})
$$

Finally it is enough to use the estimates for $|T^{-1}|$ and $|T|$ given in Lemma 17.

Now, it is enough to choose the parameters $\eta$, $\nu$ and $\gamma$ to show that the remainder is smaller than the first order of $\Lambda$. This allows us to show the existence of the isolating block. We make two different choice of parameters. The first choice is to construct the smallest possible isolating block. This gives the sharper estimates for the size of the cylinder and also gives good bounds for its derivatives. The second one is to obtain the largest isolating block where we can prove the existence of the cylinder. This gives the maximal set where we can ensure the existence and uniqueness of the cylinder.

First, we take $\eta = \rho^{q-2d\nu} = \rho^{5d\nu}$, $\nu = 1$ and $\gamma = \rho^{2d\nu}$. Recalling that $q = 7d$, we obtain that the equation obtained in Lemma 16 becomes

$$
\begin{pmatrix}
x \\
y
\end{pmatrix} = \Lambda \begin{pmatrix}
x \\
y
\end{pmatrix} + \mathcal{O}(\eta \rho^{d\nu})
$$

(55)
and the matrix $L$ obtained in Lemma 18

$$L = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} + \mathcal{O} \left( \rho^{2dr} \right)$$

(56)

Now we are ready to apply the isolating block argument as done in [BKZ11]. We apply Proposition A.1 [BKZ11] to the system in coordinates $(x, y, \Theta, I, t)$. More precisely, with the notations of appendix B [BKZ11], we set

$$u = x, s = y, c_1 = (\Theta, t), c_2 = I, \Omega = \mathbb{R}^2 \times \Omega^c = \mathbb{R}^2 \times \left[ \frac{a_- - \lambda^2}{\nu}, \frac{a_+ + \lambda^2}{\nu} \right].$$

We fix $\alpha = \lambda/2$ and we take $B^u = \{u : \|u\| \leq \eta\}$ and $B^s = \{s : \|s\| \leq \eta\}$. Then, by (55), one can easily see that

$$\dot{x} \cdot x \geq \alpha x^2 \quad \text{if} \quad |x| = \eta, |y| \leq \eta, (\Theta, I, t) \in \Omega$$
$$\dot{y} \cdot y \leq -\alpha y^2 \quad \text{if} \quad |x| \leq \eta, |y| = \eta, (\Theta, I, t) \in \Omega$$

Moreover, using (56) and recalling that $\alpha = \lambda/2$, we have that

$$L_{uu} = \lambda + \mathcal{O} \left( \rho^{2dr} \right) \geq \alpha$$
$$L_{ss} = \lambda + \mathcal{O} \left( \rho^{2dr} \right) \geq \alpha$$

on $B^u \times B^s \times \Omega$. Finally, following the notations of Proposition A.1 of [BKZ11], we have that $m \lesssim \rho^{2dr}$. Therefore, $K \leq 1/\sqrt{2}$ since

$$K = \frac{m}{\alpha - 2m} \lesssim \rho^{2dr-m+1} \ll 1$$

Thus, we can apply Proposition A.1 of [BKZ11]. This implies that there exists a $C^1$ map

$$w^c = (w^c_u, w^c_s) : \Omega \to \mathbb{R}^2$$

which satisfies $\|dw^c\| \leq 2K$, is $\gamma^{-1}$-periodic in $\Theta$ and $1$-periodic in $t$, and the graph of which is weakly invariant. Now it only remains to go back to the original variables by defining

$$\begin{pmatrix} w_{\varphi^c}^c \\ w_{j^c}^c \end{pmatrix} = \begin{pmatrix} \varphi^c_s(J^f) \\ J^c_s(J^f) \end{pmatrix} + S \begin{pmatrix} w_u^c(\gamma \varphi^f, \nu^{-1}J^f, t) \\ w_s^c(\gamma \varphi^f, \nu^{-1}J^f, t) \end{pmatrix}$$

Then, we obtain

$$|w_{\varphi^c}^c - \varphi^c_s(J^f)| \leq \rho^{m-1}\eta \leq \rho^{m+5dr}, \quad |w_{j^c}^c - J^c_s(J^f)| \leq \rho^{m-1}\eta \leq \rho^{m+5dr}$$

and for the derivatives

$$\left| \partial_{J^f} w_{\varphi^c}^c \right| \lesssim K \nu^{-1} \lesssim \rho^{2dr-m+1}, \quad \left| \partial_{(\varphi^f, t)} w_{\varphi^c}^c \right| \lesssim K \lesssim \rho^{2dr-m+1}$$

The second choice of parameters is $\eta = \rho^{2dr}, \nu = \rho^{2dr}, \gamma = \rho^{2dr}$. Proceeding analogously one can see that the isolating block argument goes through also with these choice of parameters. This choice of parameters gives the set $V$ given in Theorem 13 where there is no other invariant set except the cylinder.
7 Aubry sets in transition zones: proof of Key Theorem 7

We devote this section to study the properties of the Aubry and Mañé sets corresponding to Dirichlet resonances in the transition zones around double resonances. We follow the approach developed in [BKZ11]. We consider the Hamiltonian (44) and show that the Aubry sets related to certain cohomology classes $c \in H^1(T^2)$ belong to the invariant cylinder along the resonance $S_{k_n}$ obtained in Theorem 13. We also prove the Mather graph property for the Aubry sets.

The proof of Theorem 7 is a consequence of the following two theorems, which give vertical and horizontal estimates for the Aubry set. The first result replicates Theorem 4.1 in [BKZ11] and provides vertical estimates and also a graph property for the Weak KAM solutions analogous to the Mather graph principle. This theorem is proved in Section 7.1.

**Theorem 14.** Consider the Hamiltonian (44). Then, for each cohomology class $c \in \mathbb{R}^2$ and each weak KAM solution $u$ of $H$ at cohomology $c$, the set $\tilde{I}(u,c)$ is contained in a $C^{1,\frac{3dr}{2h}} \times \mathbb{T}^n \times \mathbb{T}$, and in the domain $\|J - c\| \leq C\rho^{\frac{3m}{10} - 1}$, for some constant $C > 0$ independent of $\rho$.

The next theorem also gives estimates for the horizontal localization of the Aubry sets. We adapt the statement of [BKZ11] to our purposes. This theorem is proved in Section 7.2.

**Theorem 15.** Consider the Hamiltonian (44). Then, for $\rho > 0$ small enough, the Mañé set $\tilde{N}(c)$ satisfies

$$s\tilde{N}(c) \subset B_{\rho,\|J - c\|} \times \mathbb{T}^n \times B_{\rho,\|J - c\|} \subset \mathbb{T}^n \times \mathbb{T}$$

for some constant $\kappa > 0$ independent of $\rho$.

With these two theorems we are ready to prove Key Theorem 7.

**Proof of Key Theorem 7** Theorem 15 implies that the suspended Mañé set $s\tilde{N}(c)$ sets belong to the set $V$ introduced in Theorem 13. Since the set $s\tilde{N}(c)$ is invariant, belonging to $V$ implies that it must belong to the cylinder $C_{k_n}$.

Now it only remains to show that the set $\tilde{I}(u,c)$ is a Lipschitz graph over the fast angle $\varphi^f$. Let $(\varphi_i, J_i), i = 1, 2$ be two points in $\tilde{I}(u,c)$. By Theorem 14 we know that

$$\|J_2 - J_1\| \leq C\rho^{\frac{3m}{10} - 1}\|\varphi_2 - \varphi_1\| \leq C\rho^{\frac{3m}{10} - 1}\left(\|\varphi_2^s - \varphi_1^s\| + \|\varphi_2^f - \varphi_1^f\|\right).$$

Now, using the fact that $\tilde{I}(u,c)$ belongs to $C_{k_n}$, we know that

$$\|\varphi_2^s - \varphi_1^s\| \leq C\rho^{\frac{3dr}{2h}-(m+1)/2} \left(\|J_2 - J_1\| + \|\varphi_2^f - \varphi_1^f\|\right).$$

64
Combining these two estimates, we obtain
\[
\|J_2 - J_1\| \leq C \rho^{13m/10 - 1} \|\varphi_2^f - \varphi_1^f\|.
\]

\[\square\]

7.1 Vertical estimates: proof of Theorem 14

The proof of Theorem 14 follows the same lines of the proof of Theorem 4.1 in [BKZ11], taking into account the different choice of parameters. We define \(L(\varphi,v,t)\) the Lagrangian associated to \(\hat{H}\). Then,

\[
\partial_{vv} L(\varphi,v,t) = \left(\partial_{JJ} \hat{H}(\varphi,\partial_v L(\varphi,v,t),t)\right)^{-1}
\]

and

\[
\partial_{\varphi v} L(\varphi,v,t) = -\partial_{\varphi J} \hat{H}(\varphi,\partial_v L(\varphi,v,t)) \partial_{vv} L(\varphi,v,t)
\]

Therefore, we have the estimates

\[
\|\partial_{vv} L\|_{C^0} \lesssim \rho^{m-1}, \quad \|\partial_{\varphi v} L\|_{C^0} \lesssim \rho^{5m-2}, \quad \|\partial_{\varphi \varphi} L\|_{C^0} \lesssim \rho^{5m-1}.
\]

(57)

Recall that a function \(u : \mathbb{T}^n \rightarrow \mathbb{R}\) is called \(K\)-semi-concave if the function

\[
x \mapsto u(x) - K \frac{\|x\|^2}{2}
\]

is concave on \(\mathbb{R}^n\), where \(u\) is seen as a periodic function on \(\mathbb{R}^n\). We state two lemmas. The first is proven in [BKZ11] and is a simple case of Lemma A.10 of [Ber08]. The second one is proven in [Fat11].

**Lemma 19.** If \(u : \mathbb{T}^n \rightarrow \mathbb{R}\) is \(K\)-semi-concave, then it is \((K \sqrt{n})\)-Lipschitz.

**Lemma 20.** Let \(u\) and \(v\) be \(K\)-semi-concave functions, and let \(I \subset \mathbb{T}^n\) be the set of points where the sum \(u + v\) is minimal. Then, the functions \(u\) and \(v\) are differentiable at each point of \(I\), and the differential \(x \mapsto du(x)\) is \(6K\)-Lipschitz on \(I\).

The Weak KAM solutions of cohomology \(c\) are defined as fixed points of the operator \(\mathcal{T}_c : \mathcal{C}(\mathbb{T}^n) \rightarrow \mathcal{C}(\mathbb{T}^n)\) defined by

\[
\mathcal{T}_c(u)(\varphi) = \min_{\gamma} u(\gamma(0)) + \int_0^T (\mathcal{L}(\gamma(t),\dot{\gamma}(t),t) + c \cdot \dot{\gamma}(t)) \, dt
\]

where the minimum is taken on the set of \(C^1\) curves \(\gamma : [0,1] \rightarrow \mathbb{T}^n\) satisfying the final condition \(\gamma(T) = \varphi\).
Proposition 1. For each \( c \in \mathbb{R}^n \), each Weak KAM solution \( u \) of cohomology \( c \) is \( C^\rho \frac{m-1}{2} \)-semi-concave and \( C^\rho \frac{m-1}{2} \)-Lipschitz for some constant \( C > 0 \) independent of \( \rho > 0 \).

Proof. Consider \( \Phi : [0, T] \to T^n \) an optimal curve for \( T_c \). That is a curve satisfying
\[
u(\varphi) = u(\Phi(0)) + \int_0^T \left( L(\Phi(t), \dot{\Phi}(t), t) + c\dot{\Phi}(t) \right) dt.
\]
We lift \( \Phi \) to a curve in \( \mathbb{R}^n \) and consider \( \Phi_x(t) = \Phi(t) + \frac{t}{T} x \). Then, \( \Phi_x(T) = \varphi + x \). Then, we have that
\[
u(\varphi + x) - \nu(\varphi) \leq \int_0^T \left( L(\Phi_x(t), \dot{\Phi}_x(t), t) - L(\Phi(t), \dot{\Phi}(t), t) + \frac{c\dot{x}}{T} \right) dt.
\]
Proceeding as in [BKZ11], one can bound the integral by
\[
u(\varphi + x) - \nu(\varphi) \leq \left( c + \partial_v L(\Phi(T), \dot{\Phi}(T), T) \right) \cdot x + K \left( \rho^{\frac{m}{5}-1} T + \rho^{13m/5-2} + \frac{\rho^{m-1}}{T} \right) |x|^2
\]
Then, taking \( T \in [\rho^{-3m/10}/2, \rho^{-3m/10}] \) (this interval contains an integer since \( \rho \ll 1 \)), we obtain
\[
u(\varphi + x) - \nu(\varphi) \leq \left( c + \partial_v L(\Phi(T), \dot{\Phi}(T), T) \right) \cdot x + K \rho^{13m/10-1} |x|^2.
\]
This estimate gives the semi-concavity constant and therefore, also the Lipschitz constant.

\[
\square
\]

7.2 Horizontal estimates: proof of Theorem 15

Hamiltonian (44) is of the form \( \hat{H}(\varphi, J) = H_1(\varphi^s, J) + \hat{\mathcal{R}}(\varphi^s, J) \) where
\[
H_1(\varphi^s, J) = \tilde{H}_0(J) + \tilde{\mathcal{E}}(\varphi^s, J).
\] (58)

First we perform a change of coordinates which puts the hyperbolic critical point of \( H_1 \) at the origin and removes some terms in its associated linearization. A similar change has been done in Section 6 but now we need it to be symplectic.

Lemma 21. There exists a symplectic change of coordinates \( (\varphi, J) = \Psi(\tilde{\varphi}, \tilde{J}) \) such that, for each fixed \( \tilde{J} \in [\tilde{a}_-, \tilde{a}_+] \),
\[
\mathcal{K}_1 = H_1 \circ \Psi^{-1},
\] (59)

as a function of \( (\tilde{\varphi}^s, \tilde{J}^s) \), has a hyperbolic critical point at \( (\tilde{\varphi}^s, \tilde{J}^s) = (0, 0) \) and moreover it satisfies
\[
D^{-1} \rho^{m+1} \leq \partial_{J^s, J^s} \mathcal{K}_1(0, 0) \leq D \rho^{m-1} \\
\rho^{dr} \leq - \partial_{\tilde{\varphi}^s, \tilde{\varphi}^s} \mathcal{K}_1(0, 0) \leq \rho^{8m/5-1}.
\]

and
\[
\partial_{J^s, \tilde{\varphi}^s} \mathcal{K}_1(0, 0) = 0, \quad \partial_{J^s, \tilde{\varphi}^s} \mathcal{K}_1(0, 0) = 0.
\]
Proof. We perform the change of variables in two steps. First we put the curve of critical points at \((0, 0)\) and then we modify its linearization. The first change is given by

\[
(\hat{\varphi}^s, \hat{J}^s) = (\varphi^s - \varphi^s(\hat{J}^f), J^s - J^s(\hat{J}^f))
\]

\[
(\hat{\varphi}^f, \hat{J}^f) = (\varphi^f - \partial_{J^f} \varphi^s(J^f), J^s - \partial_{J^f} J^s(J^f) \varphi^s, J^f)
\]

This change is symplectic, and sets the hyperbolic critical points at \((\hat{\varphi}^s, \hat{J}^s) = (0, 0)\). Note that \(\mathcal{H}_1\) is independent of \(\varphi^f\). Nevertheless, we need to consider it in the change of coordinates so that we can apply it later on to the full Hamiltonian \(\hat{\mathcal{H}}\) keeping the symplectic structure.

Now we eliminate the crossed derivatives. By construction, the Hamiltonian \(\hat{\mathcal{H}}_1\) after this change of coordinates is of the form,

\[
\mathcal{H}'(\hat{\varphi}^s, \hat{J}^f) = a_0(\hat{J}^f) + a_{20}(\hat{J}^f)(\hat{\varphi}^s)^2 + a_{11}(\hat{J}^f)\varphi^s \hat{J}^s + a_{02}(\hat{J}^f)(\hat{J}^s)^2
\]

+ higher order terms,

where by hypothesis \(a_{20} = \partial_{J^f} \hat{\mathcal{H}}_1(0, 0) > 0\) and \(D^{-1} \rho^{m+1} \leq a_{20} \leq D \rho^{m-1}\), \(a_{02} = \partial_{\varphi^s \varphi^s} \hat{\mathcal{H}}_1(0, 0) < 0\) and \(\rho^{dr} \leq -a_{02} \lesssim \rho^{sm/5-1}\), and \(a_{11} = \partial_{\varphi^f \varphi^s} \hat{\mathcal{H}}_1(0, 0)\) and \(|a_{11}| \lesssim \rho^{sm/5-1}\).

Note that we already have that \(\partial_{\varphi^f \varphi^s} \hat{\mathcal{H}}_1(0, 0) = 0\). To set the crossed derivative \(\partial_{\varphi^f \varphi^s}\) to zero, it is enough to consider the change of coordinates

\[
(\tilde{\varphi}^s, \tilde{J}^s) = \left(\tilde{\varphi}^s, \tilde{J}^s + \frac{a_{11}(\hat{J}^f)}{2a_{20}(\hat{J}^f)} \tilde{\varphi}^s\right)
\]

\[
(\tilde{\varphi}^f, \tilde{J}^f) = \left(\varphi^f - \partial_{J^f} \left[\frac{a_{11}(\hat{J}^f)}{2a_{20}(\hat{J}^f)}\right] \tilde{\varphi}^s, \tilde{J}^f\right)
\]

This change is well defined since \(a_{20} > 0\). For the estimates of the second derivatives with respect to \(\tilde{J}^s\), it is enough to point out that \(\partial_{\tilde{J}^f, \tilde{J}^s} \mathcal{K}_1(0, 0) = a_{20}(\hat{J}^f) = \partial_{J^f, J^s} \hat{\mathcal{H}}_1(0, 0)\) which satisfies the desired estimates. For the other one

\[
\partial_{\tilde{\varphi}^s, \tilde{\varphi}^s} \mathcal{K}_1(0, 0) = a_{02}(\hat{J}^f) - \frac{a_{11}^2(\hat{J}^f)}{2a_{20}(\hat{J}^f)} \leq a_{02}(\hat{J}^f) \leq -\rho^{dr}
\]

and

\[
|\partial_{\tilde{\varphi}^s, \tilde{\varphi}^s} \mathcal{K}_1(0, 0)| \leq |a_{02}(\hat{J}^f)| + \left|\frac{a_{11}^2(\hat{J}^f)}{2a_{20}(\hat{J}^f)}\right| \lesssim \rho^{sm/5}.
\]

\(\square\)

If we apply the change obtained in Lemma 21 to the Hamiltonian \(\hat{\mathcal{H}}\), we obtain a new Hamiltonian \(\mathcal{K}\) which is of the form

\[
\mathcal{K} = \mathcal{K}_1 + \mathcal{K}_2
\]

67
where $\mathcal{K}_2 = \widehat{R} \circ \Psi$ and therefore satisfies $\|\mathcal{K}_2\|_{c^2} \lesssim \rho^{dr}$. We study the horizontal localization of the Aubry sets along a fixed single resonance for the Hamiltonian $\mathcal{K}$. Then we deduce them for (14) and we prove Theorem 15. The results for $\mathcal{K}$ are summarized in the next two propositions. The first one gives localization estimates for the Aubry sets with respect to the slow angle.

**Proposition 2.** Consider a cohomology class $c = (c^s, c^f) = (0, c^f)$ with $c^f \in [\tilde{a}_-, \tilde{a}_+]$. Then, the slow angle coordinate of any point belonging to the Mañé set $\tilde{N}(c)$ associated to the Hamiltonian $\mathcal{K}$ satisfies

$$|\tilde{\varphi}^s| \leq \kappa \rho^{5dr/2}$$

for some constant $\kappa > 0$ independent of $\rho$.

**Proposition 3.** Consider a cohomology class $c = (c^s, c^f) = (0, c^f)$ with $c^f \in [\tilde{a}_-, \tilde{a}_+]$. Then, the slow action coordinate of any point belonging to the Mañé set $\tilde{N}(c)$ associated to the Hamiltonian $\mathcal{K}$ satisfies

$$|\tilde{J}^s| \leq \kappa \rho^{9dr/4}$$

for some constant $\kappa > 0$ independent of $\rho$.

These two propositions are proved respectively in Sections 7.2.1 and 7.2.2. Now, using the symplectic invariance of the Mañé set, we are ready to prove Theorem 15.

**Proof of Theorem 15.**

7.2.1 Localization for the slow angle: proof of Proposition 2

We start by studying the Lagrangian $\mathcal{L}$ and the function $\alpha$ associated to the Hamiltonian $\mathcal{K}$. The following lemma is straightforward.

**Lemma 22.** The Lagrangian $\mathcal{L}$, the Legendre dual of $\mathcal{K}$, can be split as

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$$

where $\mathcal{L}_1$ is the Legendre transform of $\mathcal{K}_1$ and $\mathcal{L}_2$ satisfies $\|\mathcal{L}_2\|_{c^2} \lesssim \rho^{dr}$.

**Lemma 23.** The $\alpha$ function associated to $\mathcal{K}$ satisfies

$$\mathcal{K}_1(0, c) - C\rho^{dr} \leq \alpha(c) \leq \mathcal{K}_1(0, c) + C\rho^{dr}.$$

for some constant $C > 0$ independent of $\rho$.

**Proof.** It follows the same lines as the proof of Lemma 4.4 in [BKZ11]. For the upper bound, it is enough to take into account that

$$\alpha(c) \leq \max_{(\phi, t)} \mathcal{K}(\phi, c, t) \leq \max_{\phi^s} \mathcal{K}_1(\phi^s, c) + \max_{(\phi, t)} \mathcal{K}_2(\phi, c, t).$$
For the lower bound, we consider the Haar measure $\mu$ of the torus $\mathbb{T} \times \{0\} \times \{\omega\} \times \mathbb{T}$, which is not necessarily invariant but is closed. Then

$$\alpha(c) \geq c \cdot \omega - \int L \, d\mu \geq c \cdot \omega - L_1(0, \omega) - C\rho^{qr} \geq K_1(0, c) - C\rho^{qr}.$$ 

We use these lemmas to obtain precise upper and lower bounds for the Lagrangian. These estimates are used later on to obtain good upper and lower bounds for the action along curves belonging to the Aubry sets.

**Lemma 24.** For each $c$ in the resonance,

$$L(\varphi, v, t) - cv + \alpha(c) \leq (\partial_v L_1(\varphi^s, \omega) - c)(v - \partial_J K_1(\varphi^s, c)) - (K_1(\varphi^s, c) - K_1(\varphi^s, c))$$

$$+ \frac{D\rho^{m-1}}{2} \|\partial_v L_1(\varphi^s, \omega) - \partial_v L_1(\varphi^s, \omega)\|^2$$

$$+ \rho^{qr} + \frac{D\rho^{m-1}}{2} \|v - w\|^2$$

$$L(\varphi, v, t) - cv + \alpha(c) \geq (\partial_v L_1(\varphi^s, \omega) - c)(v - \partial_J K_1(\varphi^s, c)) - (K_1(\varphi^s, c) - K_1(\varphi^s, c))$$

$$+ \frac{\rho^{m+1}}{2D} \|\partial_v L_1(\varphi^s, \omega) - \partial_v L_1(\varphi^s, \omega)\|^2$$

$$- \rho^{qr} + \frac{\rho^{m+1}}{2D} \|v - w\|^2$$

To prove this lemma, we first prove the following lemma. Call $\omega = \partial_J K_1(0, c)$. This implies that $c = \partial_v L_1(0, \omega)$.

**Lemma 25.** The Lagrangian $L_1$ satisfies

$$L_1(\varphi^s, \omega) - L_1(0, \omega) \leq (\partial_v L_1(\varphi^s, \omega) - c)(\omega - \partial_J K_1(\varphi^s, c)) - (K_1(\varphi^s, c) - K_1(0, c))$$

$$+ \frac{D\rho^{m-1}}{2} \|\partial_v L_1(\varphi^s, \omega) - \partial_v L_1(0, \omega)\|^2$$

$$L_1(\varphi^s, \omega) - L_1(0, \omega) \geq (\partial_v L_1(\varphi^s, \omega) - c)(\omega - \partial_J K_1(\varphi^s, c)) - (K_1(\varphi^s, c) - K_1(0, c))$$

$$+ \frac{D^{-1}\rho^{m+1}}{2} \|\partial_v L_1(\varphi^s, \omega) - \partial_v L_1(0, \omega)\|^2$$

**Proof.** We know that

$$L_1(0, \omega) = c\omega - K_1(0, c).$$
For the other term, we write $L_1(\varphi^s, \omega) = \partial_\theta L_1(\varphi^s, \omega) - K_1(\varphi^s, \partial_\omega L_1(\varphi^s, \omega))$ and applying Taylor to $H_1$ with respect to the actions

\[ L_1(\varphi^s, \omega) \leq \partial_\theta L_1(\varphi^s, \omega) - K_1(\varphi^s, c) - \partial_j K_1(\varphi^s, c)(\partial_\omega L_1(\varphi^s, \omega) - \partial_0 L_1(0, \omega)) + \frac{D\rho^{m-1}}{2} \|\partial_\omega L_1(\varphi^s, \omega) - \partial_\omega L_1(0, \omega)\|^2 \]

Analogously, one can obtain the lower bounds. \hfill \square

Using Lemma 25, the proof of Lemma 24 follows the same lines of the analogous lemma in [BKZ11].

**Proof of Lemma 24.** To compute the upper bound, we apply Taylor at $v = \omega$ and we use Lemma 22.

\[ L(\varphi, v, t) \leq L(\varphi^s, v) + O(\rho^{m_r}) \]

\[ \leq L(\varphi^s, w) + \partial_\omega L(\varphi^s, w)(v - \omega) + \frac{D\rho^{m-1}}{2} \|v - \omega\|^2 + O(\rho^{m_r}) \]

Then, it is enough to use Lemma 25. One can obtain the lower bound analogously. \hfill \square

Now, from Lemma 24 we derive estimates doing Taylor expansion around $\varphi^s = 0$.

**Lemma 26.** The Lagrangian $L$ satisfies

\[ L(\varphi, v, t) - cv + \alpha(c) \leq \rho^{s_0/5 - 1} |\varphi^s|^2 + \frac{D\rho^{m-1}}{2} \|v - w\|^2 \]

\[ + \rho^{s_0/5 - 1} |\varphi^s|^2 \|v - \omega\| + \rho^{16m/5 - 2} |\varphi^s|^4 + \rho^{m_r} \]

\[ L(\varphi, v, t) - cv + \alpha(c) \geq \rho^{d_r} |\varphi^s|^2 + \frac{D\rho^{m+1}}{2} \|v - w\|^2 \]

\[ - \rho^{s_0/5 - 1} |\varphi^s|^2 \|v - \omega\| - \rho^{16m/5 - 2} |\varphi^s|^4 - \rho^{m_r} \]

**Proof.** We bound each term in Lemma 24. For the first one, we take into account that by (57) and Lemma 21, we have that $\partial_{\varphi^s v^s} L(0, \omega) = 0$ and $\partial_{\varphi^s v^s} L(0, \omega) = 0$. Now, applying Taylor at $\varphi^s = \varphi^*_s$ we have

\[ |\partial_\theta L_1(\varphi^s, \omega) - c| \leq \|L_1\|c \|\varphi^s\|^2 \leq \rho^{s_0/5 - 1} |\varphi^s|^2 \]

\[ |v - \partial_j K_1(\varphi^s, c)| \leq \|v - \omega\| + \|K_1\|c \|\varphi^s\|^2 \leq \|v - \omega\| + \rho^{s_0/5 - 1} |\varphi^s|^2 \]

Therefore, we obtain

\[ |(\partial_\theta L_1(\varphi^s, \omega) - c)(v - \partial_j K_1(\varphi^s, c))| \leq \rho^{s_0/5 - 1} |\varphi^s|^2 \|v - \omega\| + \rho^{16m/5 - 2} |\varphi^s|^4 \]

\[ \leq \rho^{s_0/5 - 1} |\varphi^s|^2 \|v - \omega\| + \rho^{16m/5 - 2} |\varphi^s|^4. \]

For the second term, we use Lemma 21 to obtain $K_1(\varphi^s, c) - K_1(\varphi^*_s, c) \leq \rho^{s_0/5} r^2$ and $K_1(\varphi^s, c) - K_1(\varphi^*_s, c) \geq \rho^{d_r} r^2$. The rest of the terms are straightforward. \hfill \square
This lemma gives very precise positive upper bounds for the Lagrangian but only locally close to \( \varphi^s = 0 \). To obtain the horizontal estimates we also need global positive lower bounds. They are given in the next lemma.

**Lemma 27.** Consider \( c = (c^s, c^f) = (0, c^f) \). Then, the Lagrangian \( \mathcal{L} \) satisfies
\[
\mathcal{L}(\varphi, v, t) - cv + \alpha(c) \geq \rho^{dr}|\varphi^s|^2 - C \rho^{gr}.
\]

**Proof.** By definition
\[
\mathcal{L}_1(\varphi^s, v) = \max \{Jv - \mathcal{K}_1(\varphi^s, J)\} \geq cv - \mathcal{K}_1(\varphi^s, c).
\]
So, we have that \( \mathcal{L}(\varphi, v, t) - cv \geq -\mathcal{K}_1(\varphi^s, c) - C \rho^{gr} \). Now it only remains to use the lower bounds of \( \alpha(c) \) obtained in Lemma \ref{lemma:lowerbounds} which lead to
\[
\mathcal{L}(\varphi, v, t) - cv + \alpha(c) \geq - (\mathcal{K}_1(\varphi^s, c) - \mathcal{K}_1(0, c)) - C \rho^{gr}.
\]
To complete the proof it suffices to use the properties of \( \mathcal{K}_1 \) obtained in Lemma \ref{lemma:lowerbounds} \( \square \)

Now, we obtain upper bounds for the action,

**Lemma 28.** Let \( u(\varphi, t) \) be a weak KAM solution at cohomology \( c \in \Gamma \). Given \( r_1 \geq \rho^{gr/2-8m/5+1} \) and two points \((\varphi_1, t_1)\) and \((\varphi_2, t_2) \in \mathbb{T} \times B(\varphi^s, r_1) \times \mathbb{T} \), we have
\[
u(\varphi_2, t_2) - u(\varphi_1, t_1) \lesssim \rho^m r_1.
\]

**Proof.** We proceed as in \cite{BKZ11} by taking a curve
\[
\varphi(t) = \varphi_1 + (t - \tilde{t}_1)(\tilde{\varphi}_2 - \tilde{\varphi}_1 + (T + \tilde{t}_2 - \tilde{t}_1)\omega)
\]
where \( T \in \mathbb{N} \) is a parameter to be fixed later, and \( \tilde{t}_i \in [0, 1) \) and \( \tilde{\varphi}_i \in [0, 1)^2 \) are representatives of the angular variables \( t_i, \varphi_i \). Then,
\[
u(\varphi_2, t_2) - u(\varphi_1, t_1) \lesssim \int_{\tilde{t}_1}^{\tilde{t}_2 + T} \mathcal{L}(\varphi(t), \dot{\varphi}(t), t) - c \cdot \dot{\varphi}(t) + \alpha(c) \, dt
\]
Now, by Lemma \ref{lemma:lowerbounds}
\[
u(\varphi_2, t_2) - u(\varphi_1, t_1) \lesssim \int_{\tilde{t}_1}^{\tilde{t}_2 + T} (D \rho^{m-1}\|\dot{\varphi}(t) - \omega\|^2 + \rho^{8m/5-1}r_1^2\|\dot{\varphi}(t) - \omega\| + \rho^{8m/5-1}r_1^2 \, dt.
\]
Then,
\[
u(\varphi_2, t_2) - u(\varphi_1, t_1) \lesssim \int_{\tilde{t}_1}^{\tilde{t}_2 + T} \frac{D \rho^{m-1}(T + \tilde{t}_2 - \tilde{t}_1)^2}{(T + \tilde{t}_2 - \tilde{t}_1)^2} + \frac{\rho^{8m/5-1}r_1^2}{T + \tilde{t}_2 - \tilde{t}_1} + \rho^{8m/5-1}r_1^2 \, dt
\]
\[
u(\varphi_2, t_2) - u(\varphi_1, t_1) \lesssim \frac{D \rho^{m-1}}{T + \tilde{t}_2 - \tilde{t}_1} + \rho^{8m/5-1}r_1^2 + \rho^{8m/5-1}r_1^2(T + \tilde{t}_2 - \tilde{t}_1).
\]
Now, taking
\[
\frac{1}{\rho^{3m/10}} \sqrt{\frac{1}{r_1^2}} \leq T + \tilde{t}_2 - \tilde{t}_1 \leq \frac{1}{\rho^{3m/10}} \sqrt{\frac{1}{r_1^2}}
\]
we obtain the estimate given in the lemma. \qed

Now we look for lower bounds of the action so that we can reach a contradiction. We adapt the argument in [BKZ11]. We consider the invariant set \( \tilde{I}(v, c) \subset \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T} \) associated to a suspended weak KAM solution \( v \) at cohomology \( c \), and its projection \( \mathcal{I}(v, c) \subset \mathbb{T}^2 \times \mathbb{T} \). The set \( \tilde{I}(v, c) \) is the suspension of the invariant set \( \tilde{I}(u, c) \) associated to \( u(\varphi) = v(\varphi, 0) \). We consider also a curve \( \varphi(t) : \mathbb{R} \rightarrow \mathbb{T}^2 \) calibrated by \( u \). Namely,
\[
\int_{t_1}^{t_2} L(\varphi(t), \dot{\varphi}(t), t) - c \cdot \dot{\varphi}(t) + \alpha(c) \, dt = u(\varphi(t_2), t_2) - u(\varphi(t_1), t_1)
\]
for each interval \((t_1, t_2)\). Since \( u \) is bounded, this integral is bounded independently of the time interval. Now, if we take \( r_1 = \sqrt{2C \rho^{q-d} r^2} \), by Lemma 23, we know that
\[
L(\varphi, v, t) - cv + \alpha(c) \geq \frac{1}{2} \rho^{dr} r_1^2
\]
if \( \varphi^s \) does not belong to the ball \( B(0, r_1) \). Then, the set of times for which \( \varphi^s \) does not belong to this ball must have finite measure. Call \([t_1, t_2]\) to one of the connected components of this open set of times. Then, \( \varphi^s(t_1) \) and \( \varphi^s(t_2) \) both belong to \( B(0, r_1) \). Therefore, by Lemma 28 we have that
\[
\int_{t_1}^{t_2} L(\varphi(t), \dot{\varphi}(t), t) - c \dot{\varphi}(t) + \alpha(c) \, dt = u(\varphi(t_2), t_2) - u(\varphi(t_1), t_1) \leq \rho^n r_1 \quad (60)
\]
Now, we assume that \( \varphi^s(t) \) is not contained in the ball \( B(0, r_0) \) with \( r_0 = \rho^{5dr/2} \) and we reach a contradiction. Assume that the points out of this ball happen in the time interval \([t_1, t_2]\) (if not change the time connected component). Then, call \( t_4 \) the first time such that \( |\varphi^s(t_4)| = r_0 \) and \( t_3 \) the time the largest time in \([t_1, t_4]\) such that \( |\varphi^s(t_3)| = r_0/2 \). Then,
\[
\int_{t_1}^{t_2} L(\varphi(t), \dot{\varphi}(t), t) - c \dot{\varphi}(t) + \alpha(c) \, dt \geq \int_{t_3}^{t_4} L(\varphi(t), \dot{\varphi}(t), t) - c \dot{\varphi}(t) + \alpha(c) \, dt.
\]
since the function in the integral is always positive.

Now we bound this integral using Lemma 27 and using that in this time interval \(|\varphi^s(t)| \geq r_0/2\).
\[
\int_{t_3}^{t_4} L(\varphi(t), \dot{\varphi}(t), t) - c \dot{\varphi}(t) + \alpha(c) \, dt \geq \int_{t_3}^{t_4} \frac{1}{2} \rho^{dr} (\varphi^s(t))^2 \, dt \geq \frac{1}{8} \rho^{6dr} (t_4 - t_3).
\]
So, joining this estimate with (60) we have that $\rho^{dr}(t_4 - t_3)/8 \leq \rho^{m_1}$ which implies $t_4 - t_3 \leq 8\sqrt{2C} \rho^{m_1 + 5d/2}$. Now, we know that

$$
\frac{r_0}{2} = \varphi^s(t_4) - \varphi^s(t_3) = \int_{t_3}^{t_4} \dot{\varphi}^s(t) \, dt
$$

Using the differential equation we have that $\dot{\varphi}^s = O(\rho^{m_1 - 1} J^s + \rho^{8m/5 - 1}(\varphi^s)^2)$. By the vertical estimates, for the calibrated curve we know that $|J^s| \leq \rho^{23m/10 - 1}$ (since $c^s = 0$) and we are assuming that $|\varphi^s| \leq r_1$. So $|\dot{\varphi}^s(t)| \leq \rho^{23m/10 - 1} \leq \rho^{2m}$. This implies that

$$
\frac{r_0}{2} = \left| \int_{t_3}^{t_4} \dot{\varphi}^s(t) \, dt \right| \leq \rho^{2m}(t_4 - t_3)
$$

So, we obtain that

$$
\frac{1}{2} \rho^{5dr/2} \leq 8\sqrt{2C} \rho^{m_1 + 5d/2}.
$$

So, if we take $q = 18$ we get a contradiction. This implies that all the curve must be inside the ball of radius $r_0 = \rho^{5dr/2}$.

### 7.2.2 Localization for the slow action: proof of Proposition 3

To ensure that the Aubry sets belong to the cylinder obtained in Theorem 13 we need to obtain improved estimates for the slow action $\tilde{J}^s$. As we have done in the previous section we omit the tildes over the variables to simplify notation. We consider the invariant set $\tilde{I}(v, c) \subset \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T}$ associated to a (suspended) weak KAM solution $v$ of the Hamiltonian (59) at a cohomology class $c$, and its projection $I(v, c) \subset \mathbb{T}^3$. The set $\tilde{I}(v, c)$ is the suspension of the invariant set $\tilde{I}(u, c)$ associated to the restriction $u(\varphi) = v(\varphi, 0)$. Consider $\varphi(t) : \mathbb{R} \to \mathbb{T}^2$, a calibrated curve by $u$. That is,

$$
\int_{t_1}^{t_2} \mathcal{L}(\varphi(t), \dot{\varphi}(t), t) - c \cdot \dot{\varphi}(t) + \alpha(c) \, dt = u(\varphi(t_2), t_2) - u(\varphi(t_1), t_1)
$$

for each interval $(t_1, t_2)$. In Section 7.2.1 we have seen that these calibrated curves must satisfy $|\varphi^s| \leq \rho^{5dr/2}$.

First, prove that for a calibrated curve $(\varphi(t), J(t))$, the associated action $J^s(t)$ satisfies $|J^s(t)| \leq \rho^{9dr/4} / 2$ for all time except for a bounded set of times. Indeed, by Lemma 26

$$
\mathcal{L}(\varphi, v, t) - cv + \alpha(c) \geq \rho^{dr} |\varphi^s|^2 + \frac{D \rho^{m_1 + 1}}{2} \|v - w\|^2
$$

$$
- \rho^{8m/5 - 1} |\varphi^s|^2 \|v - \omega\| - \rho^{16m/5 - 2} |\varphi^s|^4 - \rho^{2r}
$$

$$
\geq \rho^{dr} |\varphi^s|^2 + \frac{D \rho^{m_1 + 1}}{2} \left( \|v - w\| - \rho^{3m/5 - 2} \|\varphi^s\|^2 \right)^2
$$

$$
- \frac{\rho^{11m/5 - 3}}{2D} |\varphi^s|^4 - \rho^{16m/5 - 2} |\varphi^s|^4 - \rho^{2r}
$$

73
Now, since $|\varphi^s(t)| \lesssim \rho^{5dr/2}$ for all $t \in \mathbb{R}$, we know that
\[
\mathcal{L}(\varphi(t), \dot{\varphi}(t), t) - cv + \alpha(c) \geq \frac{D\rho^{m+1}}{2} \left( \|\dot{\varphi}(t) - w\| - \frac{\rho^{3m/5-1}}{D} |\varphi^s(t)|^2 \right)^2 - \rho^q.
\]
Consider the corresponding curve in the Hamiltonian setting $(\varphi(t), J(t))$ and assume that $|J^s(t)| \geq \rho^{9dr/4}/2$ for some subset of time $I \subset \mathbb{R}$. Then, since by the Legendre transform $|\varphi^s(t)| \geq |\rho^{m+1}J^s(t)| \geq \rho^{9dr/4+m+1}$. Then, $\mathcal{L}(\varphi, v, t) - cv + \alpha(c) \geq \rho^{9dr/2+3m+3}$. Therefore, the time subset $I \subset \mathbb{R}$ in which $|J^s(t)| \geq \rho^{9dr/4}/2$ must be bounded. Otherwise, the action along the curve would be infinite.

Assume that there exist times $t_1 < t_2$ such that $J^s(t_1) = \rho^{9dr/4}/2$ and $J^s(t_2) = \rho^{9dr/4}$. If they would not exist, one would have that for all time $|J^s(t)| \leq \rho^{9dr/4}$ and the proof of Proposition 3 would be finished. Since the curve is calibrated, we know that, for all $t \in [t_1, t_2]$, $|\varphi^s(t)| \leq \rho^{5dr/2}$.

We obtain upper bounds for $t_2 - t_1$. We have that
\[
\frac{\rho^{9dr/4}}{2} = J^s(t_2) - J^s(t_1) = \int_{t_1}^{t_2} \dot{J}(t) dt
\]
Using the equation associated to Hamiltonian (59), we have that $|\dot{J}(t)| \lesssim \rho^{5dr/2+8m/5-1}$. Therefore, $\rho^{9dr/4}/2 \lesssim (t_2 - t_1)\rho^{5dr/2+8m/5-1}$. This implies that $t_2 - t_1 \gtrsim \rho^{-dr/4-8m/5+1}$. On the other hand we know that $|\varphi^s(t_2) - \varphi^s(t_1)| \lesssim 2\rho^{5dr/2}$ and
\[
\varphi^s(t_2) - \varphi^s(t_1) = \int_{t_1}^{t_2} \dot{\varphi}^s(t) dt
\]
Using the equation associated to Hamiltonian (59), we have that $|\dot{\varphi}^s| \geq D\rho^{m+1}J^s - C(\varphi^s)^2$ for some constant $C > 0$. This implies that $|\dot{\varphi}^s| \gtrsim D\rho^{9dr/4+m+1}/4$. Therefore,
\[
2\rho^{5dr/2} \gtrsim \int_{t_1}^{t_2} |\dot{\varphi}^s(t)| dt \gtrsim \frac{D}{4} \rho^{9dr/4+m+1}(t_2 - t_1),
\]
which implies that $t_2 - t_1 \lesssim \rho^{dr/4-m-1}$. Thus, we have a contradiction and we can deduce that $|J^s(t)| \leq \rho^{9dr/4}$ for all $t \in \mathbb{R}$.

8 Equivalent forcing classes and shadowing

9 Perturbation of single averaged potentials

Fix $n \geq 1$. Let $\omega^* \in \mathcal{D}^n_{n,\tau}$ be a Diophantine frequency and let $\text{Vor}_n(\omega^*)$ be the corresponding Voronoi cell around it. Let $\{S_{k_j}\}_{j \in J_n(\omega^*)}$ be the set of Dirichlet segments which intersec inside $\text{Vor}_n(\omega^*)$. 

74
Consider a $\rho_n^2$-tube neighborhood of the union $\{S_{k_j}^\omega\}_{j \in J_n(\omega^*)}$, denoted by

$$U_n(\omega^*) = U_{\rho_n^2}((\cup_{j \in J_n(\omega^*)} S(\Gamma_{k_j}))) .$$

Select one Dirichlet resonant segment $S_{k^*}^\omega$. Consider its $\rho_n^2$-neighborhood $U_{\rho_n^2}(S_{k^*}^\omega)$. Apply Key Theorem 2 in this tube neighborhood. Now, we have a Hamiltonian

$$H_0(J) + Z(J, \varphi) + R(J, \varphi),$$

where

$$Z(J, \varphi) = \sum_{k \in \mathbb{Z}^3} Z_k(J) \exp 2\pi i (k \cdot \varphi).$$

and $R$ is small (see Key Theorem 2). Namely, for each $J$ only $k$ such that $|k \cdot J| \leq 2\rho^2|k|$ are non-zero for some $J \in U_{\rho_n^2}(S(\Gamma_{k^*}))$. Denote $\varphi_k^* = \varphi \cdot k^*$. Decompose the averaged perturbation $Z$ into the single average and the complement:

$$Z_{k^*}^{SR}(\varphi_k^*, J) = \sum_{p \in \mathbb{Z}} \rho_{pk^*}^{2/m}(J) Z_{pk^*}(J) \exp 2\pi i (p \varphi_k^*).$$

$$\Delta Z_{k^*} = Z - Z_{k^*}^{SR}.$$ 

We have

$$H_0(J) + Z_{k^*}^{SR}(J, \varphi_k^*) + \Delta Z_{k^*}(J, \varphi) + \rho_m^{n(1+4m/\delta)} R(J, \varphi),$$

In the same way as we define single averaged potential $Z_{k^*}^{SR}(\varphi_k^*, J)$ we can define a single averaged potential of any function $\delta Z(J, \varphi_k^*)$ as the sum of subcollection of Fourier coefficients.

Fix one such a segment $S(\Gamma_{k_j})$, $j \in J_n(\omega^*)$ and the Hamiltonian obtained after application of Pöschel normal form $\tilde{H}_0 + \tilde{H}_1$ (see (65)). Using the integrable part of this Hamiltonian we can define Dirichlet resonant segments in the action space associated to Dirichlet resonant lines

$$S^A(\Gamma_{k^*}) \subset \{ J : (\partial \tilde{H}_0, 1) \cdot k^* = 0 \} .$$

Recall that each resonant segment can be smoothly parametrized by $J^f$. Namely, $S^A(\Gamma_{k^*}) = \{ J(J^f) : a_{k^*}^J \leq J^f \leq a_{k^*}^J \}$

$$Z_{k^*}^{SR}(\varphi^*, J^f) = \sum_{m \in \mathbb{Z}} \tilde{Z}_{pk^*}(J(J^f)) \exp 2\pi i (m \varphi^*),$$

where $Z_{pk^*}(J(J^f))$ are the resonant Fourier coefficients of the averaged potential $Z$.

The goal of this section is to prove the following
Theorem 16. Let \( \omega^* \in \mathcal{D}_{n, \tau}^{n} \) be a Diophantine number for some \( n \in \mathbb{Z}_+ \), \( 50\tau < 1 \), let \( \text{Vor}_n(\omega_n) \) be its Voronoi cell. Let \( \{S(\Gamma_{k_j})\}_{j \in J_n(\omega^*)} \) be Dirichlet resonant segments inside \( \text{Vor}_n(\omega^*) \), i.e. \( S(\Gamma_{k_j}) \subset \text{Vor}_n(\omega^*) \). Let \( U_n(\omega^*) \) be the \( \rho_n^2 \)-neighborhood of the union of all these segments. Let

\[
\mathcal{H}_0(J) + \mathcal{Z}(J, \varphi) + \rho^{n(1+4m/81)}_n \mathcal{R}(J, \varphi)
\]

be the normal form. Then there exists a \( \rho_n^{1+4r} \)-small in the \( \mathcal{C}^r \) topology perturbation \( \delta \mathcal{Z}(J, \varphi) \) such that for any \( k_j \in J_n(\omega^*) \) the single averaged potential

\[
\mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{k^*}^s, J^f) = \mathcal{Z}_{k^*}^{\mathcal{SR}}(\varphi_{k^*}^s, J^f) + \delta \mathcal{Z}_{k^*}^{\mathcal{SR}}(\varphi_{k^*}^s, J^f)
\]

satisfies the following conditions for each \( J^f \in [\tilde{a}_-^{k^*}, \tilde{a}_+^{k^*}] \) each global minimum of \( \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{k^*}^s, J^f) \) is \( \rho_n^{d(r+1)} \) nondegenerate, i.e.

\[
\partial_{\varphi_{k^*}^s}^2 \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{k^*}^s, J^f) \geq \rho_n^{d(r+1)}.
\]

There are finitely many \( J^f \in [\tilde{a}_-^{k^*}, \tilde{a}_+^{k^*}] \) such that \( \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{k^*}^s, J^f) \) has exactly two global minima, denoted \( \varphi_{\min}(J) \) and \( \varphi_{\min}^2(J) \) and we have

\[
|\partial_{J^f} \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{\min}^1(J), J^f) - \partial_{J^f} \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{\min}^2(J), J^f)| \geq \rho_n^{7d(r+1)/4},
\]

in other words, speed of change of values wrt \( J^f \) is the minima differs by at least \( \lambda_n^{7/4} \).

There are no \( J^f \in [\tilde{a}_-^{k^*}, \tilde{a}_+^{k^*}] \) such that \( \mathcal{Z}_{k^*}^{\mathcal{SR}, \delta}(\varphi_{k^*}^s, J^f) \) has more than two global minima.

Proof. Here we prove this theorem using Theorem 17 (proven and stated below).

Denote by \( \chi_j^\rho \) a \( \mathcal{C}^\infty \) bump function given by

\[
\chi_j^\rho(J) = \begin{cases} 1 & \text{dist } (J, \mathcal{S}(\Gamma_{k_j})) \leq \rho_n^2, \\ 0 & \text{dist } (J, \mathcal{S}(\Gamma_{k_j})) > 2\rho_n^2. \end{cases}
\]

In the \( \rho_n^2 \) tube neighborhood we would like to perturb by trigonometric polynomials.

\[
\mathcal{Z}_{j}^{\rho}(\varphi_j, J^f) = \mathcal{Z}_{j}(\varphi_j, J^f) + \chi_j^\rho(J(J^f)) (\sigma_1 \cos 2\pi \varphi_j + \sigma_2 \sin 2\pi \varphi_j + \sigma_3 \cos 4\pi \varphi_j + \sigma_4 \cos 4\pi \varphi_j + J^f \sigma_5 \cos 2\pi \varphi_j + J^f \sigma_6 \sin 2\pi \varphi_j).
\]

Since the perturbation is \( \mathcal{C}^r \)-small, it suffices to have

\[
|\sigma_1|, |\sigma_2|, |\sigma_3|, |\sigma_4|, |\sigma_5|, |\sigma_6| \leq \rho_n^{2(1+4r)(r+1)}.
\]

Denote by \( D_\rho \) the 6-dimensional cube of these parameters. Recalling the relation between \( R_n \) and \( \rho_n \) and \( |k_j| < R_n \) we have

\[
\rho_n^{2(1+4r)(r+1)} \leq R_n^{2(3-4r)(1+4r)(r+1)} \leq |k_j|^{-2(1+4r)(3-4r)(r+1)}.
\]
Denote by $\mu_\rho$ the Lebesgue probability measure supported on the product of two disks $D_\nu$ of radii $\nu = \rho_n^{2(1+4r)(r+1)}$.

Apply Theorem 17. Then,

$$
\mu_\rho\{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) \in D_\nu : \exists J^f \in [a_+^k, a_-^k] \text{ such that either } Z^\sigma_j(\varphi_j, J^f) \text{ has a } \lambda_\sigma^-\text{-degenerate minimum} \} \leq \frac{(3C_3 + 7)(1 + |a_+ - a_-|) C_3^2 \sqrt{\lambda_\sigma}}{8 \pi^5 \nu^3},
$$

where $C_3$ is bounded by the $C^1$-norm of $H$.

Since $|J_+ - J_-|$ is bounded by $\rho$, for $\nu = \rho_n^{2(1+4r)(r+1)}$ we have that for any perturbation $Z^\sigma_j(\varphi_j, J^f)$ with $\sigma \in D_\rho^{2(1+4r)(r+1)}$ localized in the $\rho^2$-neighborhood of a Dirichlet resonant segment $\mathcal{I}_{k_n}$ is $C^r$-small. In order to assure that measure of the exceptional set being small it suffices to pick $\lambda_\sigma = \rho_n^{14(r+1)}$. Then the upper bound on measure is $\frac{(3C_3 + 7) C_3^2 \rho^{(1-24r)(r+1)}}{8 \pi^5 \nu^3}$.

\[ \square \]

### 9.1 Perturbation of families of functions on the circle

Let $f_t : T \to \mathbb{T}, t \in [a_-, a_+]$ be a $C^r$ smooth one-parameter family of periodic functions $\theta \in T = \mathbb{R}/\mathbb{Z}$. Consider the following 6-parameter extended family:

$$
F_t(\theta, \sigma) = f_t(\theta) + \sigma_1 \cos 2\pi \theta + \sigma_2 \sin 2\pi \theta + \sigma_3 \cos 4\pi \theta + \sigma_4 \sin 4\pi \theta + t\sigma_5 \cos 2\pi \theta + t\sigma_6 \sin 2\pi \theta.
$$

Fix $\nu > 0$. Denote $D_\nu := \{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4) : \sigma_1^2 + \sigma_2^2 \leq \nu^2, \sigma_3^2 + \sigma_4^2 \leq \nu^2, \sigma_5^2 + \sigma_6^2 \leq \nu^2 \}$ the direct products of three disks. Denote by $\mu_\nu$ the Lebesgue product probability measure on each.

Notice that the set of parameters $D_\nu$ is invariant with respect to rotations. Indeed, using the sine and cosine sum formula for properly chosen $\{ \sigma_i(\theta^*) \}_i$, we have that

$$
\sigma_1 \cos 2\pi(\theta + \theta^*) + \sigma_2 \sin 2\pi(\theta + \theta^*) + \sigma_3 \cos 4\pi(\theta + \theta^*) + \sigma_4 \sin 4\pi(\theta + \theta^*) + t\sigma_5 \cos 2\pi(\theta + \theta^*) + t\sigma_6 \sin 2\pi(\theta + \theta^*)
$$

$$
= \sigma_1(\theta^*) \cos 2\pi \theta + \sigma_2(\theta^*) \sin 2\pi \theta + \sigma_3(\theta^*) \cos 4\pi \theta + \sigma_4(\theta^*) \sin 4\pi \theta + t\sigma_5(\theta^*) \cos 2\pi \theta + t\sigma_6(\theta^*) \sin 2\pi \theta, \tag{61}
$$

where $\sigma_1^2 + \sigma_2^2 = \sigma_1^2(\theta^*) + \sigma_2^2(\theta^*), \sigma_3^2 + \sigma_4^2 = \sigma_3^2(\theta^*) + \sigma_4^2(\theta^*), \sigma_5^2 + \sigma_6^2 = \sigma_5^2(\theta^*) + \sigma_6^2(\theta^*)$.

Fix $\sigma \in D_\nu^1$. For $t \in [a_-, a_+]$ we study global minima of $F_t(\cdot, \sigma)$. Let

$$
\theta_{\text{min}}(t, \sigma) := \{ \theta^* : F_t(\theta^*, \sigma) = \min_\theta F_t(\theta, \sigma) \}
$$
be a global minimum. Let $\lambda(t, \sigma) := \min_{\theta_{\min}(t, \sigma)} \partial_\theta^2 F_t(\theta, \sigma)|_{\theta = \theta_{\min}(t, \sigma)}$.

Call a value $t^* \in [a_-, a_+]$ bifurcation if there are at least two global minima $\theta_{\min}^1(t, \sigma)$ and $\theta_{\min}^2(t, \sigma)$. Denote by $B_\sigma \subseteq [a_-, a_+]$ the set of bifurcation points of $F_t(\cdot, \sigma)$. Introduce the difference between speed of change of values at the global minima with respect to $t$:

$$d_1(t^*, \sigma) = \min_{\theta_{\min}} \{ |\partial_t F_t(\theta, \sigma)|_{\theta = \theta_{\min}^1(t)} - |\partial_t F_t(\theta, \sigma)|_{\theta = \theta_{\min}^2(t)} | \},$$

where minimum is taken over all possible pairs of global minima. For a bifurcation value introduce we have two global minima $\theta_{\min}^1(t)$ and $\theta_{\min}^2(t)$.

$$d_2(t^*, \sigma) = \min_{\theta_{\min}} \{ F(\theta_{\min}^1, \sigma) - F(\theta_{\min}^2(t), \sigma) \},$$

where the minimum is taken over all local minima.

Denote $C_3 = \max_{t \in [a_-, a_+], \sigma \in D_\nu} \{ |\partial_t F_t(\theta, \sigma)|, |\partial_\sigma^2 F_t(\theta, \sigma)|, |\partial_\theta^2 F_t(\theta, \sigma)| \}$, where partial derivates are taken with respect to $\sigma$ an $t$ only.

The goal of this section is to prove the following

**Theorem 17.** Let $f_t(\theta)$ be a $C^4$ smooth one-parameter family of periodic functions $t \in [a_-, a_+]$ with $\| f_t \|_{C_3} \leq C_3$. Then

$$\mu_\nu \{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) \in D_\nu : \lambda(t, \sigma) \leq \lambda_*, \min_{t^* \in B_\sigma} (d_1(t^*, \sigma) + d_2(t^*, \sigma)) \leq \lambda_*^{7/4} \} \leq \frac{(3C_3 + 7)C_3^2}{8\pi^5 \nu^3} \sqrt{\lambda_*}.$$ 

### 9.2 Condition on local minimum

In order to determine the set of almost degenerate minima define the following set:

$$\text{Cr}(\lambda^*) = \{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) \in D_\nu : \exists t \in [a_-, a_+], \partial_\theta F_t(\theta^*, \sigma) = 0, |\partial_\theta^2 F_t(\theta, \sigma)| \leq \lambda_*, \theta^* \text{ is a local minimum} \}.$$ 

We embed this set into a bigger set using the following trick and the discretization method from [KH07]. If $\theta^*$ is a local minimum of a $C^4$ function $g$ on $\mathbb{T}$ and $g''(\theta^*)$ is small, then $g^{(3)}(\theta^*)$ should also be small. Otherwise, it is not a local minimum. Here is the formal claim about relations between first, second, and third derivatives.

Let $\theta^* \in \mathbb{T}$ be a local minimum of $g(\theta)$ for some $\theta^* \in \mathbb{T}$. This implies that $g'(\theta^*) = 0$. Consider the second and third derivative at the critical point $g''(\theta^*)$, $g^{(3)}(\theta^*)$. Since $\theta^*$ is a local minimum, $g''(\theta^*) \geq 0$ and, in case, $g''(\theta^*) = 0$ we have $g''(\theta^*)(\sigma) = 0$.

**Lemma 29.** Let $g \in C^4$, $\| g \|_{C^4} \leq C$ and $\theta^*$ is a local minimum and $0 \leq g''(\theta^*) \leq \lambda$, then $|g^{(3)}| \leq 3(C + 2)\sqrt{\lambda}$.
Proof. Expand $g$ near its local minimum $\theta^*$ using Taylor with the remainder in the Lagrange form. We have

$$g(\theta^* + \delta \theta) = g(\theta^*) + g'(\theta^*)\delta \theta + \frac{1}{2}g''(\theta^*)\delta \theta^2 + \frac{1}{6}g^{(3)}(\theta^* + \xi)\delta \theta^3,$$

where $|\xi| < |\delta \theta|$. By the intermediate value theorem we can rewrite

$$g(\theta^* + \delta \theta) = \frac{1}{2}\delta \theta^2 \left( g''(\theta^*) + \frac{1}{3}g^{(3)}(\theta^* + \xi) \right)$$

Plug in $\delta \theta = -\text{sign}(g^{(3)}(\theta^*))\sqrt{\lambda}$. We have

$$g''(\theta^*) + \frac{1}{3}g^{(3)}(\theta^*)\delta \theta + g^{(4)}(\theta^* + \eta)\xi \delta \theta \leq \lambda - (C + 2)\lambda + C\lambda \leq -\lambda < 0.$$

This is a contradiction. \qed

Let $\lambda^\# = \lambda^*/C_3$ be small positive. Denote by $\mathbb{Z}^2_{\lambda^\#}$ the $\lambda^\#$-grid in $[a_-, a_+] \times \mathbb{T}$. To estimate measure of $\text{Cr}(\lambda^*)$ we use the discretization trick (see e.g. [KH07]).

$$\text{Cr}^d(\lambda^*) = \{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) \in D_\nu : \exists (t^*, \theta^*) \in \mathbb{Z}^2_{\lambda^\#},$$

$$|\partial \theta F_t(\theta^*, \sigma)| \leq 2\lambda^*, \quad |\partial^2 \theta F_t(\theta^*, \sigma)| \leq 2\lambda^*, \quad |\partial^3 \theta F_t(\theta^*, \sigma)| \leq (3C_3 + 7)\lambda^*. \}$$

One of the key element of the proof is the following estimate

**Lemma 30.** With the above notations we have

$$\mu_\nu\{\text{Cr}^d(\lambda^*) \cup B(\lambda_*)\} \leq \frac{(3C_3 + 7)C_3^2(|a_+ - a_-| + 1)\sqrt{\lambda_*}}{16 \pi^5 \nu^3}.$$

This lemma implies Theorem [17] through a simple approximation argument. Suppose $\sigma \in \text{Cr}(\lambda^*)$. Then there exists $t' \in [a_-, a_+]$ and $\theta' \in \mathbb{T}$ such that

$$\partial \theta F_{t'}(\theta', \sigma) = 0, \quad |\partial^2 \theta F_{t'}(\theta', \sigma)| \leq \lambda_*$$

and $\theta'$ is a local minimum. Then by lemma [29] we have

$$|\partial^2 \theta F_{t'}(\theta', \sigma)| \leq 3(C + 2)\sqrt{\lambda_*}.$$

Now we can approximate $(t', \theta', \sigma)$ by a $\mathbb{Z}^2_{\lambda^*/C_3}$-grid point with precision $\lambda_*/C$. Due to derivatives of order up to 4 being bounded by $C_3$ we can transfer the above bounds to a nearby grid point.
Proof. The proof proceeds as follows. We estimate a probability that given \((t^*, \theta^*) \in \mathbb{Z}_N^2\), we have

\[
\mu_\nu\{(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6) \in D_\nu : \\
|\partial_\theta F_{t^*}(\theta^*, \sigma)| \leq 2\lambda_*, \ |\partial_\theta^2 F_{t^*}(\theta^*, \sigma)| \leq 2\lambda_*, \ |\partial_\theta^3 F_{t^*}(\theta^*, \sigma)| \leq (3C_3 + 7)\lambda_* \}
\]
Rewrite the family using (61) as follows

\[
F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma) = f_{t^*}(\theta^* + \delta \theta) + \sigma_1(\theta^*) \cos 2\pi \delta \theta + \sigma_2(\theta^*) \sin 2\pi \delta \theta + \sigma_3(\theta^*) \cos 4\pi \delta \theta + \sigma_4(\theta^*) \sin 4\pi \delta \theta + t \sigma_5(\theta^*) \cos 2\pi \delta \theta + t \sigma_6(\theta^*) \sin 2\pi \delta \theta.
\]
Compute the first derivatives

\[
\partial_\theta F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma) = \partial_\theta f_{t^*}(\theta^* + \delta \theta) + \\
-2\pi \sigma_1(\theta^*) \sin 2\pi \delta \theta + 2\pi \sigma_2(\theta^*) \cos 2\pi \delta \theta + 4\pi \sigma_3(\theta^*) \sin 4\pi \delta \theta + 4\pi \sigma_4(\theta^*) \cos 4\pi \delta \theta - 2\pi t \sigma_5(\theta^*) \sin 2\pi \delta \theta + 2\pi t \sigma_6(\theta^*) \cos 2\pi \delta \theta.
\]
For \(\delta \theta = 0\). Fix any values of \((\sigma_1(\theta^*), \sigma_3(\theta^*), \sigma_4(\theta^*))\) we see that probability

\[
\text{Leb}\{\sigma_1(\theta^*) : |\partial_\theta F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma)| \leq 2\lambda_*\} \leq \frac{4\lambda^*}{2\pi}.
\]
Compute the second derivatives

\[
\partial_\theta^2 F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma) = \partial_\theta^2 f_{t^*}(\theta^* + \delta \theta) + -4\pi^2 \sigma_1(\theta^*) \cos 2\pi \delta \theta + \\
-4\pi^2 \sigma_2(\theta^*) \sin 2\pi \delta \theta - 16\pi^2 \sigma_3(\theta^*) \cos 4\pi \delta \theta + 16\pi^2 \sigma_4(\theta^*) \sin 4\pi \delta \theta - \\
-4\pi^2 t \sigma_5(\theta^*) \cos 2\pi \delta \theta + 4t \pi^2 \sigma_6(\theta^*) \sin 2\pi \delta \theta.
\]
For \(\delta \theta = 0\). Fix any values of \((\sigma_2(\theta^*), \sigma_3(\theta^*), \sigma_4(\theta^*))\) we see that probability

\[
\text{Leb}\{\sigma_2(\theta^*) : |\partial_\theta^2 F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma)| \leq 2\lambda_*\} \leq \frac{4\lambda^*}{4\pi^2}.
\]
Compute the third derivatives

\[
\partial_\theta^3 F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma) = \partial_\theta^3 f_{t^*}(\theta^* + \delta \theta) + 8\pi^3 \sigma_1(\theta^*) \sin 2\pi \delta \theta + \\
-8\pi^3 \sigma_2(\theta^*) \cos 2\pi \delta \theta - 64\pi^3 \sigma_3(\theta^*) \sin 4\pi \delta \theta + 64\pi^3 \sigma_4(\theta^*) \cos 4\pi \delta \theta + \\
+8\pi^3 t \sigma_5(\theta^*) \sin 2\pi \delta \theta - 8\pi^3 t \sigma_6(\theta^*) \cos 2\pi \delta \theta.
\]
For \(\delta \theta = 0\). Fix any values of \((\sigma_1(\theta^*), \sigma_2(\theta^*), \sigma_3(\theta^*))\) we see that probability

\[
\text{Leb}\{\sigma_4(\theta^*) : |\partial_\theta^3 F_{t^*}(\theta^* + \delta \theta, \sigma^* + \delta \sigma)| \leq (3C_3 + 7)\sqrt{\lambda_*} \} \leq \frac{2(3C_3 + 7)\sqrt{\lambda_*}}{64\pi^3}.
\]
The number of grid points is \(|a_+ - a_-|\lambda_y^2 = C_4^2|a_+ - a_-|\lambda_x^2 \). Using the fact that the measure \(\mu_\nu\) is the product of two Lebesgue probability measures of the disks \(\sigma_1^2 + \sigma_2^2 \leq \nu^2\) and \(\sigma_3^2 + \sigma_4^2 \leq \nu^2\) we combine the above estimates and get the required bound. \(\square\)
9.3 Bifurcation of global minima

Suppose \( t^* \in B \) be a bifurcation point, i.e. there are at least two global minima \( \theta_{\min}^1(t, \sigma) \) and \( \theta_{\min}^2(t, \sigma) \). In order to determine the set of almost having two global minima having nearly the same speed of change of values in \( t \) define the following set:

\[
\mathcal{B}^1(\lambda^*) = \{(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \in D_\nu : \exists t \in [a_-, a_+],
|F(\theta_{\min}^1(t, \sigma)) - F(\theta_{\min}^2(t, \sigma))| \leq \lambda_*^{7/4},
|\partial \theta F(\theta_{\min}^1(t, \sigma)) - \partial \theta F(\theta_{\min}^2(t, \sigma))| \leq \lambda_*^{7/4} \}
\]

As before to estimate measure of this set we discretize the set of parameters and of \( \theta \)'s. Consider \( \lambda^* = \lambda_*/C_3^2 \) grid. We can find \( \lambda^* \)-closest points of the grid to each global minima and denote them by \( \bar{\theta}_{\min}^1(t, \sigma) \) and \( \bar{\theta}_{\min}^2(t, \sigma) \). For parameters \( \sigma \not\in \text{Cr}(\lambda_*) \) we have that

\[
|\partial \theta F_t(\bar{\theta}_{\min}^i(t, \sigma), \sigma)| \leq \lambda_* \quad \text{for } i = 1, 2.
\]

Now the discretized versions of the set \( \mathcal{B}(\lambda_*) \) is

\[
\mathcal{B}^{1,d}(\lambda^*) = \{(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \in D_\nu : \exists (\theta_1, \theta_2, t) \in \mathbb{Z}^3_{\lambda^*},
|\partial \theta F(\theta_1, \sigma)|, |\partial \theta F(\theta_2, \sigma)| \leq 2\lambda_*,
|\partial^2 \theta F(\theta_1, \sigma)|, |\partial^2 \theta F(\theta_2, \sigma)| \geq \lambda_*,
|F(\theta_{\min}^1(t, \sigma)) - F(\theta_{\min}^2(t, \sigma))| \leq 2\lambda_*^{7/4},
|\partial \theta F(\theta_{\min}^1(t, \sigma)) - \partial \theta F(\theta_{\min}^2(t, \sigma))| \leq 2\lambda_*^{7/4} \}.
\]

Show that \( \mathcal{B}^1(\lambda^*) \subset \mathcal{B}^{1,d}(\lambda^*) \). We have

\[
|F_t(\bar{\theta}_{\min}^i(t, \sigma), \sigma) - F_t(\bar{\theta}_{\min}^j(t, \sigma), \sigma)| \leq \lambda_*^2/C_3.
\]

Therefore, if

\[
|F_t(\theta_{\min}^1(t, \sigma)) - F_t(\theta_{\min}^2(t, \sigma), \sigma)| \leq \lambda_*^{7/4}
\]

implies

\[
|F_t(\theta_{\min}^1(t, \sigma)) - F_t(\theta_{\min}^2(t, \sigma), \sigma)| \leq \lambda_*^{7/4} + 2\lambda_*^{7/4}/C_3 \leq 2\lambda_*^{7/4}
\]

Also if

\[
|\partial \theta F_t(\theta_{\min}^1(t, \sigma), \sigma) - \partial \theta F_t(\theta_{\min}^2(t, \sigma), \sigma)| \leq \lambda_*^{7/4}
\]

implies

\[
|\partial \theta F_t(\theta_{\min}^1(t, \sigma), \sigma) - \partial \theta F_t(\theta_{\min}^2(t, \sigma), \sigma)| \leq \lambda_*^{7/4} + 2\lambda_*^{7/4}/C_3 \leq 2\lambda_*^{7/4}.
\]

Choose a parameter \( \sigma \) such that

\[
|\partial^2 \theta F_t(\theta_{\min}^1(t, \sigma), \sigma)| \geq \lambda_* \quad \text{for all } t \in [a_-, a_+].
\]

Then

\[
|\theta_{\min}^1(t, \sigma) - \theta_{\min}^2(t, \sigma)| \geq \frac{\lambda_*}{C_3}
\]

81
and

\[ |\bar{\theta}_{\min}^1(t,\sigma) - \bar{\theta}_{\min}^2(t,\sigma)| \geq \frac{\lambda_*}{C_3} - 2\lambda# \leq \frac{\lambda_*}{2C_3}. \]

Therefore,

\[ \text{Leb}\{\sigma_4(\bar{\theta}_{\min}^1(t,\sigma)) : |F_t(\bar{\theta}_{\min}^1(t,\sigma)) - F_t(\bar{\theta}_{\min}^2(t,\sigma))| \leq 2\lambda_*^{7/4}\} \leq 8C_3\lambda_*^{3/4}. \]

and

\[ \text{Leb}\{\sigma_6(\bar{\theta}_{\min}^1(t,\sigma)) : |\partial_t F_t(\bar{\theta}_{\min}^1(t,\sigma)) - \partial_t F_t(\bar{\theta}_{\min}^2(t,\sigma))| \leq 2\lambda_*^{7/4}\} \leq 8C_3\lambda_*^{3/4}. \]

Combining we get

\[ \mu(\nu)\{B_1^{ld}(\lambda_*)\} \leq \frac{C_3^3 \sqrt{\lambda_*}}{2\pi^2 \nu^3}. \]

We now prove the following

**Lemma 31.** With the above notations we have

\[ \mu(\nu)\{B_1^{ld}(\lambda_*)\} \leq \frac{C_3^3 \sqrt{\lambda_*}}{2\pi^2 \nu^3}. \]

In order to study quantitative absence of three global minima we use a very similar strategy. The condition of having three global minima can be written as follows. If a function \( G \) has three global minima, then it has three critical points with the same value:

\[ G(\theta_0) = G(\theta_1) = G(\theta_2), \quad G'(\theta_0) = G'(\theta_1) = G'(\theta_2) = 0. \]

We need a quantitative version of when this condition almost holds. We have

\[ B^2(\lambda_*) = \{(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \in D_\nu : \exists t \in [a_-, a_+], \]

\[ \max_{i=1,2,3} |\partial_{\theta} F(t_{\min}^i(t,\sigma), \sigma)| \leq \lambda_* \]

\[ \max_{i \neq j \in \{1,2,3\}} |F(t_{\min}^i(t,\sigma), \sigma) - F(t_{\min}^j(t,\sigma), \sigma)| \leq \lambda_*^{7/4}. \]

By analogy with we rewrite the family

\[ F_t(\theta, \sigma) = f_t(\theta) + \sigma_1(\theta_1) \cos 2\pi \theta + \sigma_2(\theta_1) \sin 2\pi \theta + \]

\[ + \sigma_3(\theta_1) \cos 4\pi \theta + \sigma_4(\theta_1) \sin 4\pi \theta + t\sigma_5(\theta_1) \cos 2\pi \theta + t\sigma_6(\theta_1) \sin 2\pi \theta =: f_t(\theta) + P_\sigma(\theta; \theta_1). \]

Using independent parameters we evaluate measure of

\[ |\partial_{\theta} F_t(\theta_i, \sigma)| \leq \lambda_*, \quad i = 1, 2, 3 \]
and

\[ |F_t(\theta_1, \sigma) - F_t(\theta_2, \sigma)| \leq \lambda, \quad |F_t(\theta_1, \sigma) - F_t(\theta_3, \sigma)| \leq \lambda, \]

To keep the condition \( \partial_\theta F_t(\theta_1, \sigma) = 0 \) it suffices to have \( \partial_\theta P_\sigma(\theta_1; \theta_1) = 0 \).

To keep the condition \( \partial_\theta F_t(\theta_2, \sigma) = 0 \) it suffices to have \( \partial_\theta P_\sigma(\theta_2; \theta_1) = 0 \).

To keep the condition \( \partial_\theta F_t(\theta_2, \sigma) = 0 \) it suffices to have \( \partial_\theta P_\sigma(\theta_3; \theta_1) = 0 \).

We obtain five conditions and four variables. Linear algebra calculations complete the proof in the same way as before.

### 9.4 Deformation of single averaged potential along all Dirichlet resonant segments

Fix the collection of the Dirichlet resonances \( \{k_n, \{k^j_n\}_{j \in J_n}\} \). Denote this collection by \( J_n \). Each resonance \( k \in \mathbb{Z}^3 \) has the corresponding averaged potential

\[ Z_k(\varphi \cdot k, J) = \sum_{s \in \mathbb{Z}} \tilde{h}_{sk}(J) \exp(isk \cdot \varphi), \]

where \( \tilde{h}_m, m \in \mathbb{Z}^3 \) are the Fourier coefficients of \( \tilde{H}_1(J, \varphi) \). We add a perturbation

\[ \tilde{H}_1(J, \varphi) + \Delta H_1(J, \varphi) \]

so that for each \( k \in \{k_n, \{k^j_n\}_{j \in J_n}\} \) the averaged potential

\[ Z'_k(\varphi \cdot k, J) = Z_k(\varphi \cdot k, J) + \Delta Z_k(\varphi \cdot k, J) \]

has a unique non-degenerate minimum and

\[ \min_{J_f} \partial^2_{\varphi \cdot k} Z'_k(\varphi \cdot k, J^f(J_f), J_f) \geq \rho_n^{5(r+1)}. \]

The perturbation has the following form:

\[ \Delta Z_k(k \cdot \varphi) = \]

\[ = \sum_{k \in J_n} \Delta h_k^1 \cos(k \cdot \varphi) + \Delta h_k^2 \sin(k \cdot \varphi) + \Delta h_k^3 \cos(2k \cdot \varphi) + \Delta h_k^4 \sin(2k \cdot \varphi). \]

In order to have small \( C^r \) norm of the perturbation we have

\[ |\Delta h_k^j| \leq l_n = R_n^{-r-1} \rho_n^{r+1} \sim \rho_n^{4(r+1)/3}. \]

We prove that for the 4-dimensional Lebesgue measure of each resonance
\[
\text{mes \{ } (\Delta h^1_k, \Delta h^2_k, \Delta h^3_k, \Delta h^4_k) : \text{ there is no } J_k \in [J^-_k, J^+_k] \\
\text{such that } Z_k(\varphi \cdot k, J^s(J^f), J^f) + \Delta Z_k(k \cdot \varphi) \text{ has} \\
a \rho_n^{9(r+1)/2}\text{-degenerate minimum} \} \leq C \frac{\rho_n^{9(r+1)/2}}{l^3_n} \leq C \rho_n^{(r+1)/2},
\]

where \( C \) is a universal constant independent of \( Z \). The reason we have \( l^3_n \) in the denominator, because we find three conditions to be almost fullfilled to have a locally degenerate minimum. Each condition, when measures with respect to normalized measure get additional \( l \) in the denominator.

The volume of a 4-dimensional ball of radius \( \delta \) is \( \delta^4 \), so there exists a perturbation of size \( \rho_n^{10(r+1)/2} \). This perturbation provides \( \lambda_n = \rho_n^{4.5(r+1)} \)-nondegenerate minimum. By BKZ theorem ?? cylinder should persist under perturbations of size 

\[ \lambda_n^2 = \rho_n^{9(r+1)} \gg \rho_n^{10(r+1)}. \]

9.5 Hölder norms and approximations

Recall that \( C^\mu(\mathbb{R}^n) \) for \( 0 < \mu < 1 \) denotes the space of bounded Hölder continuous functions \( f : \mathbb{R}^n \to \mathbb{R} \) with the norm

\[
\| f \|_{C^\mu} := \sup_{|x-y|<1} \frac{|f(x) - f(y)|}{|x-y|^{\mu}} + \sup_{x \in \mathbb{R}^n} |f(x)|.
\]

If \( \mu = 0 \), then \( \| f \|_{C^\mu} \) denotes the sup-norm. For \( \ell = k + \mu \) with \( k \in \mathbb{Z}_+ \) and \( 0 \leq \mu < 1 \) we denote by \( C^\ell(\mathbb{R}^n) \) the space of functions \( f : \mathbb{R}^n \to \mathbb{R} \) with Hölder continuous partial derivatives \( \partial^\alpha f \in C^\ell(\mathbb{R}^n) \) for all multi-indices \( \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n \) with \( |\alpha| := \alpha_1 + \cdots + \alpha_n \leq k \). We define the norm

\[
\| f \|_{C^\ell} := \sum_{|\alpha| \leq \ell} |\partial^\alpha f|_{C^\mu}
\]

for \( \mu := \ell - [\ell] < 1 \). We also use the standard abbreviation \( \alpha! := \alpha_1! \cdots \alpha_n! \) and \( x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \) for \( y = (y_1, \ldots, y_n) \in \mathbb{R}^n \).

**Lemma 32.** There is a family of convolution operators

\[ S_r f(x) = r^{-n} \int_{\mathbb{R}^n} K_r(r^{-1}(x-y))f(y)dy, \quad 0 < r \leq 1, \]

from \( C^0(\mathbb{R}^n) \) into the space of entire functions on \( \mathbb{R}^n \) with the following property. For every \( d \geq \ell \geq 0 \), there exists a constant \( c = c(\ell, n, d) > 0 \) such that, for every \( f \in C^\ell(\mathbb{R}^n) \)

we have

\[ \| S_r f - f \|_{C^s} \leq c \| f \|_{C^{r\ell-s}}, \quad s \leq \ell, \]

84
Finally, if $f$ is periodic in some variables, then so are the approximating functions $S_r f$ in the same variables.

We shall also use a $C^\infty$-version of this lemma.

**Lemma 33.** There is a family of convolution operators

$$S_r f(x) = r^{-n} \int_{\mathbb{R}^n} K_r(r^{-1}(x - y)) f(y) dy, \quad 0 < r \leq 1,$$

from $C^0(\mathbb{R}^n)$ into the space of $C^\infty$ functions on $\mathbb{R}^n$ with the following property. For every $d \geq \ell \geq 0$, there exists a constant $c = c(\ell, n, d) > 0$ such that, for every $f \in C^\ell(\mathbb{R}^n)$ we have

$$\|S_r f - f\|_{C^s} \leq c \|f\|_{C^\ell} r^{\ell - s}, \quad s \leq \ell,$$

$$\|S_r f\|_{C^s} \leq c \|f\|_{C^\ell} r^{\ell - s}, \quad \ell \leq s \leq d.$$

Finally, if $f$ is periodic in some variables, then so are the approximating functions $S_r f$ in the same variables.

In the case of convolution operators mapping $C^0(\mathbb{R}^n)$ into the space of entire functions is a classical lemma going back to Jackson, Moser, Zehnder (see e.g. [SZ89], Lemma 3). For the proof of our main result we need to deform a function $f$ on the complement of an open set $U$, keeping the same value in $U$. We slightly adapt the proof from [?], lm. 3.

**Proof**

Define a real analytic function

$$K(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{K}(\xi) \exp(i x \cdot \xi) \, d\xi, \quad x \in \mathbb{R}^n,$$

where its Fourier transform

$$\hat{K}(\xi) = \int_{\mathbb{R}^n} \exp(-i x \cdot \xi) \, d\xi, \quad \xi \in \mathbb{R}^n$$

is a smooth function with a compact support contained in the ball $|\xi| \leq a$, $\hat{K}(\xi) = \hat{K}(-\xi)$, and

$$\partial^\alpha \hat{K}(\xi) = \begin{cases} 1, & \text{if } \alpha = 0 \\ 0, & \text{if } \alpha \neq 0. \end{cases}$$

Denote by $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ a $C^\infty$-bump function

$$\rho(x) = \begin{cases} 1, & \text{if } 0 \leq r < 1 \\ 0, & \text{if } 2 < r, \end{cases}$$

monotonically decreasing in between 1 and 2. Define for each $r > 0$ a truncation of $K(x)$ by $K_r(x) = K(x) \rho(x \sqrt{r})$. This is a compactly supported function whose support belongs
to the ball of radius $2/\sqrt{r}$ around the origin. The standard estimate of an oscillatory integral shows that for any $\alpha \in \mathbb{Z}_+^n$ and any $N \in \mathbb{Z}_+$ there is $C = C(N, \alpha, \hat{K})$ such that

$$|\partial^\alpha K(x)| \leq \frac{1}{(2\pi)^n} \left| \int_{\mathbb{R}^n} \hat{K}(\xi)(i\xi)^\alpha \exp(ix \cdot \xi) \, d\xi \right| \leq C |x|^{-N}.$$ 

Denote

$$S_r^0 f(x) = \frac{1}{r^n} \int_{\mathbb{R}^n} K(r^{-1}(x - y)) \, f(y) \, dy, \quad 0 < r \leq 1.$$ 

and

$$S_r f(x) = \frac{1}{r^n} \int_{\mathbb{R}^n} K(r^{-1}(x - y)) \, f(y) \, dy, \quad 0 < r \leq 1.$$ 

By the aforementioned Lemma 3 \cite{??} we have

$$\|S_r^0 f - f\|_{c^s} \leq c \|f\|_{c^\ell} r^{-s}, \quad s \leq \ell,$$

$$\|S_r^0 f\|_{c^s} \leq c \|f\|_{c^\ell} r^{-s}, \quad \ell \leq s.$$ 

Now we compare $S_r f$ and $S_r^0 f$. Let $\alpha \in \mathbb{Z}_+^n$ and $N > |\alpha| + 1$. Since the operators $S_r$ and $\partial^\alpha$ commute, we have

$$\left| \partial^\alpha S_r f(x) - \partial^\alpha S_r^0 f(x) \right| \leq \frac{1}{r^{n+|\alpha|}} \int_{\mathbb{R}^n} \left| \partial^\alpha K(r^{-1}(x - y)) - \partial^\alpha K(r^{-1}(x - y)) \right| |f(y)| \, dy \leq$$

$$\frac{C}{r^{n+|\alpha|}} \int_{\mathbb{R}^n} \left| \partial^\alpha K(r^{-1}(x - y)) \right| |1 - \rho((x - y)/\sqrt{r})| |f(y)| \, dy \leq$$

$$c(n) C \sup_{x \in \mathbb{R}^n} |f(x)| \int_{\sqrt{r}}^{\infty} r^{-N+|\alpha|} \, dr \leq \frac{c(n)}{N - 1 - |\alpha|} C r^{(N-1-|\alpha|)/2}.$$ 

Choosing $N$ large enough we prove the required estimates. \hfill $\square$

**Lemma 34.** (Lemma 4, \cite{SZL)} Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is the limit of a sequence of real analytic functions $f_k(x)$ defined in the complex strip $|\Re x| < r_k = 2^{-k}$, $x \in \mathbb{C}^n$, with $0 < r_0 \leq 1$ and

$$|f_k(x) - f_{k-1}(x)| \leq A r_k^l, \quad \forall x, \quad |\Re x| \leq r_k.$$

Then $f \in C^s(\mathbb{R}^n)$ for every $s \leq l$ which is not an integer and, moreover,

$$|f(x) - f_0(x)|_{c^s} \leq c A(\theta(1 - \theta)) r_0^l$$

for $0 < \theta = s - [s] < 1$ and a suitable constant $c = c(l, n) > O$. 

86
9.6 Smoothness of compositions and the inverse

In order to estimate smoothness of deformations in the original coordinate system we need some auxiliary claims. Let \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) be coordinate system in \( \mathbb{R}^n \).

**Lemma 35.** (DH09 Appendix C) Let \( F \in C^k(\mathbb{R}^n) \) and \( G \in C^k(\mathbb{R}^n, \mathbb{R}^n) \), \( k \geq 1 \). Then for some \( c = c(n, k) \) we have

\[
\| F \circ G \|_{C^k} \leq c(\| F \|_{C^1} \| G \|_{C^k} + \sum_{p=2}^k \| F \|_{C^p} \sum_{j_1+\cdots+j_p=k} \| G \|_{C^{j_1}} \cdots \| G \|_{C^{j_p}} ).
\]

In particular,

\[
\| F \circ G \|_{C^k} \leq c(\| F \|_{C^1} \| G \|_{C^k} + \| F \|_{C^k} \| G \|_{C^{k-1}} ).
\]

Recall that \( D_R = D_R \times \mathbb{T}^{n+1} \subset \mathbb{R}^n \times \mathbb{T}^{n+1} \ni (I, \varphi, t) \). Let \( H \in C^k(D_R) \) be a \( C^k \) smooth function. Denote by \( X_H \) the Hamiltonian vector field associated to \( H \) given by

\[
X_H = (\partial_\varphi H, -\partial_\varphi H),
\]

where \( \partial_\varphi \) are partial derivatives with respect to coordinate variables. Clearly we have \( X_H \in C^{k-1}(D_R, \mathbb{R}^{2n}) \) and

\[
\| X_H \|_{C^{k-1}(D_R)} \leq \| H \|_{C^k(D_R)}. \]

Assuming that \( |\partial_\varphi H|_{C^0(D_R)} < \delta \) for some \( \delta < R \), then by the mean value theorem the time \( t \) map of the vector field \( \Phi^H_t \) is a well-defined \( C^{k-1} \) map and satisfies

\[
\Phi^H = \Phi^H_1 : D_{R-\delta} \to D_R.
\]

If \( H \) is integrable, then \( \delta \) can be chosen equal to 0.

In what follows we need to estimate the \( C^k \) norm of \( \Phi^H \) in terms of the \( C^k \) norm of the vector field \( X_H \). One can expect to have \( \Phi^H \) being \( C^k \) close to the identity when \( X_H \) is \( C^k \) close to zero. In our case, for smaller \( k \) is will indeed be true, but for higher \( k \) estimates will deteriorate. Recall the standard relation

\[
\Phi^H_t = \text{Id} + \int_0^t X_H \circ \Phi^H_s \, ds.
\]

It follows from the classical formula due to Faà di Bruno (see e.g. R. Abraham-J. Robbin, Transversal mapping and flows, Benjamin, New-York, 1967) that

\[
\| F \circ G \|_{C^k} \leq c \| F \|_{C^k} \| G \|_{C^k}^k
\]

for some \( c = c(n, k) \) and for \( C^k \) vector-valued functions defined on appropriate domains. Recall that at some point of the proof we consider a mollification of a \( C^k \) smooth vector field with a mollifying parameter \( \sigma \). In this setting we have the following
Lemma 36. Let $X_{Hs} \in C^k(D_R, \mathbb{R}^{2n})$ be a one-parameter family of $C^k$ vector field, $\delta, \sigma_0$ be small positive, $0 < m < k$. Assume that

$$\max_{0 \leq \sigma \leq \sigma_0} \|X_{Hs}\|_{C^0(D_R)} < \delta,$$

$$\max_{0 \leq \sigma \leq \sigma_0} \|X_{Hs}\|_{C^k(D_R)} < 1,$$

then

$$\|\Phi^H - \text{Id}\|_{C^k(D_{R-\delta})} \leq c \|\Phi^H_{s}\|_{C^k(D_R)}$$

for some $c = c(n, k, R)$. In addition,

$$\|F \circ \Phi^H - \text{Id}\|_{C^k(D_{R-\delta})} < c \|H\|_{C^k(D_R)} \|\Phi^H\|_{C^k(D_R)}.$$  

This lemma for $m = 0$ is lemma 3.1 [Bou10], (see also similar to lemma 3.15 [DH09]). For $m > 0$ we use estimates from the proof of lemma 3.15 [DH09].

Proof: By the fundamental theorem of calculus we have

$$\Phi^H_{t}(x) = x + \int_{0}^{t} \frac{\partial \Phi^H_{\tau}}{\partial \tau}(x) d\tau = x + \int_{0}^{t} X_{Hs} \circ \Phi^H_{\tau}(x) d\tau, d\tau,$$

where $x = (I, \varphi, t) \in D_R$ and for the canonical matrix $J$ of the sympletic form $\omega = dI \wedge d\varphi + dE \wedge dt$ we have $X_H = J\nabla H$. The extra variable $E$, conjugated to the angle $s$, was introduced to make apparent the sympletic character of the change of variables. Using the last formula of lemma 35 we obtain

$$\|\Phi^H_{t}\|_{C^l} \leq \|\text{Id}\|_{C^l} + \int_{0}^{1} \|\Phi^H_{\tau}\|_{C^l} d\tau \leq \|\text{Id}\|_{C^l} +$$

$$\int_{0}^{1} \left( \|X_{Hs}\|_{C^l} \Phi^H_{\tau}\|_{C^l} + \sum_{p=2}^{l} \|X_{Hs}\|_{C^p} \sum_{j_1 + \ldots + j_p = l} \|\Phi^H_{\tau}\|_{C^{j_1}} \ldots \|\Phi^H_{\tau}\|_{C^{j_p}} \right) d\tau,$$

where $l = 2, \ldots, k + m$ and $c$ is a constant depending on $l$ and $n$. Notice also that in this and the similar sums given below all $\{j_s\}_s$ are strictly positive.

For $l = 1$ we have

$$\|\Phi^H_{t}\|_{C^1} \leq \|\text{Id}\|_{C^1} + \int_{0}^{1} \|X_{Hs}\|_{C^1} \Phi^H_{\tau}\|_{C^1} d\tau.$$

Now by induction define

$$a_l = \max_{0 \leq \sigma \leq \sigma_0} \|\Phi^H_{t}\|_{C^l}.$$  

Denote $d_l = \max_{0 \leq \sigma \leq \sigma_0} \|X_{Hs}\|_{C^l}$. We know that $d_l \leq 1$ for $1 \leq l \leq k$ and $d_{k+l} \leq \sigma^{-l}$ for $1 \leq l \leq m$. Then,

$$a_1 \leq 1 + d_1a_1,$$

88
Lemma 37. Let \( l \) and \( D \) maps of an open set \( \Omega \) for some \( l \leq k \). Then for the inverse \( \Phi^{-1} \) and all other \( j \), moreover, each of its elements there is at most one \( d \) such that \( j > l \). Differentiate \( \Phi^{-1} \). Let \( \Phi^{-1} \sigma(\Omega) \). We have

\[
a_{k+1} \leq 1 + d_1 a_{k+1} + c d_{k+1} a_1^{k+1} + \sum_{p=2}^{k} d_p \sum_{j_1 + \ldots + j_p = l} a_{j_1} \ldots a_{j_p} \leq c + c \sigma^{-1}.
\]

Let \( 1 < l \leq m < k \). We have

\[
a_{k+l} \leq 1 + d_1 a_{k+l} + c d_{k+l} a_1^{k+l} + \sum_{p=2}^{k+l-1} d_p \sum_{j_1 + \ldots + j_p = k+l-1} a_{j_1} \ldots a_{j_p} \leq c + c \sigma^{-l}.
\]

Due to the fact that \( m < k \) and all \( \{j_s\} \) are strictly positive in the second line sum in each of its elements there is at most one \( j_s > k \). In this case \( p < l-1 \leq m-1 < k-1 \) and all other \( j_s < l-1 \leq m-1 < k-1 \). Thus, \( j_s \leq k + l - p \leq k + l - 2 \) and such an element in the sum is bounded by \( c \sigma^{-l+2} \). All others are uniformly bounded. However, \( d_{k+l} \leq \sigma^{-l} \) and gives the leading term.

Lemma 37. Let \( 0 < s < r \). Let \( \Phi_\sigma : \Omega \to \mathbb{R}^d \) be a one-parameter family of \( C^r \) smooth maps of an open set \( \Omega \subset \mathbb{R}^d \) which is \( C^2 \)-close to the identity on the closure and such that for some \( C > 0 \) we have

\[
\|\Phi_\sigma\|_{C^s} \leq C \quad \|\Phi_\sigma\|_{C^{s+j}} \leq C \sigma^{-j} \text{ for } j \leq r - s.
\]

Then for the inverse \( \Phi_\sigma^{-1} \) is \( C^r \)-smooth on \( \Omega'_\sigma = \Phi_\sigma(\Omega) \), namely, \( \|\Phi_\sigma^{-1}(y)\|_{C^r(\Omega')} \) is bounded. Moreover,

\[
\|\Phi_\sigma^{-1}\|_{C^s} \leq C \quad \|\Phi_\sigma^{-1}\|_{C^{s+j}} \leq C \sigma^{-j} \text{ for } j \leq r - s.
\]

Proof: Denote for \( \beta, \alpha \in \mathbb{Z}_+^n \) \( \beta < \alpha \) if \( \beta_i \leq \alpha_i \) for \( i = 1, \ldots, n \).

Differentiate \( \Phi_\sigma^{-1} \circ \Phi_\sigma(x) \) wrt \( x \). In the domain of definition \( x \in \Omega \) we have

\[
D\Phi_\sigma^{-1}(\Phi_\sigma(x))D\Phi_\sigma(x) \equiv Id,
\]

where \( D\Phi_\sigma^{-1}(y) \) and \( D\Phi_\sigma(x) \) are \( n \times n \) linearization matrices.

Let \( \alpha \in \mathbb{Z}_+^n \) with \( |\alpha| < r \). Consider

\[
0 \equiv \partial^\alpha (D\Phi_\sigma^{-1}(\Phi_\sigma(x))D\Phi_\sigma(x)).
\]

89
For any integer $k$ with $0 \leq k \leq r$ denote by $D^k\Phi_\sigma(x)$ the collection of all partial derivatives of order up to $k$. Let $e_j = (0, \ldots, 1_j, \ldots, 0)$. By induction we can prove the following statement:

$$0 \equiv \partial^\alpha(D\Phi_{\sigma}^{-1}(y)\big|_{y=\Phi(x)} D\Phi_\sigma(x)) =$$

$$= \sum_{\beta \prec \alpha} \partial^\beta D\Phi_{\sigma}^{-1}(y)\big|_{y=\Phi(x)} P_{\beta,\alpha}(D\Phi_\sigma(x), \ldots, D^{\alpha-\beta}\Phi_\sigma(x)),$$

where $P_{\beta,\alpha}$ are $n \times n$ matrix polynomials with $\beta \prec \alpha$ satisfying the following inductive formula: for any $j = 1, \ldots, n$

$$P_{\beta,\alpha+e_j}(D\Phi_\sigma(x), \ldots, D^{\alpha-\beta+1}\Phi_\sigma(x)) =$$

$$= \partial x_j P_{\beta,\alpha}(D\Phi_\sigma(x), \ldots, D^{\alpha-\beta}\Phi_\sigma(x)) + \partial x_j \Phi_\sigma(x) P_{\beta-e_j,\alpha}(D\Phi_\sigma(x), \ldots, D^{\alpha-\beta}\Phi_\sigma(x)),$$

and $P_{0,0}(D\Phi_\sigma(x)) = D\Phi_\sigma(x)$. Notice that all partial derivatives of $\Phi_{\sigma}^{-1}$ of order $\leq r$ can be explicitly computed using partial derivatives of $\Phi_\sigma$ of order $\leq r$ and of $\Phi^{-1}$ of order $\leq r - 1$. This implies the claim.

**Proof of Lemma** Consider the sequence of tube neighborhoods of KAM tori, $\text{KAM}_{\eta,\tau}$. For a positive $\lambda > 0$ denote $U_\lambda(\text{KAM}_{\eta,\tau})$ the $\lambda$-tube neighborhood. Let

$$H_\varepsilon^{\text{mol}} = H_0 + \varepsilon H_1 + \varepsilon \sum_{n=1}^\infty \Delta H^n$$

Consider the annuli neighborhood

$$U_{\rho_n}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_n+2r}(\text{KAM}_{\eta,\tau}).$$

Denote by $\chi_n$ a non-negative $C^\infty$ bump function

$$\chi_n(I, \varphi) = \begin{cases} 1 & \text{if } (I, \varphi) \in U_{\rho_n}(\text{KAM}_{\eta,\tau}) \setminus U_{\rho_n+2r}(\text{KAM}_{\eta,\tau}) \\ 0 & \text{if } (I, \varphi) \in U_{\rho_n+2r}(\text{KAM}_{\eta,\tau}) \\ 0 & \text{if } (I, \varphi) \in U_{\rho_n+\rho_n+2r}(\text{KAM}_{\eta,\tau}). \end{cases}$$

Recall that $Srf$ is mollification of a function $f$ from lemma 32. By induction let

$$\Delta H^1(I, \varphi) := \chi_1(I, \varphi) (S_{\rho_1^{1+4r}} H_1(I, \varphi) - H_1(I, \varphi)).$$

$$H_\varepsilon^1(I, \varphi) := H_0(I) + \varepsilon H_1^1(I, \varphi) := H_0(I) + \varepsilon(H_1(I, \varphi) + \Delta H^1(I, \varphi))$$

and

$$\Delta H^{n+1}(I, \varphi) := \chi_n(I, \varphi) (S_{\rho_n^{1+4r}} H^n(I, \varphi) - H^n(I, \varphi))$$

$$H_\varepsilon^{n+1}(I, \varphi) := H_0(I) + \varepsilon H_1^{n+1}(I, \varphi) := H_0(I) + \varepsilon(H_1^n(I, \varphi) + \Delta H^n(I, \varphi)).$$

90
The resulting Hamiltonian

\[ H_{mol}^\varepsilon := H_0 + \varepsilon H_1^{mol} := H_0 + \varepsilon \left( H_1 + \sum_{n=1}^{\infty} \Delta H^n \right). \]

By lemma [32] we have

\[ \| \Delta H^n \|_{C^r} \leq c(\rho_n)^3 \]

for some continuous function \( c(\rho) \to 0 \) as \( \rho \to 0 \). By construction supports of only neighboring \( \Delta H^n \) and \( \Delta H^{n+1} \) overlap. This implies that

\[ \Delta H^{mol} = \sum_{n=1}^{\infty} \Delta H^n \]

(63)

is a \( C^r \)-small perturbation. \( \square \)

A  A time-periodic KAM

Consider a Hamiltonian \( H_0(I) \) in action-angle variables, where angles \( \varphi \in \mathbb{T}^n \) and \( I \) belongs to a bounded open set \( U \subset \mathbb{R}^n \). Consider a smooth time-periodic perturbation

\[ H_\varepsilon(\varphi, I, t) = H_0(I) + \varepsilon H_1(\varphi, I, t). \]

Using the standard procedure associate to it an autonomous Hamiltonian

\[ H^*_\varepsilon(\varphi, I, t, E) = H_0(I) + E + \varepsilon H_1(\varphi, I, t), \]

where \( E \) is conjugate to \( t \). Therefore, orbits of \( H_\varepsilon \) and projected along \( E \) orbits of \( H^*_\varepsilon \) are the same as those of \( H_\varepsilon \) independently of the value of \( E \).

Consider an auxiliary Hamiltonian \( H_0(I, E) = \exp(H_0 + E) \) and its frequency map

\[ \omega(I) = \nabla H_0(I, E) = (\exp(H_0 + E)\nabla H_0, \exp(H_0 + E)). \]

Define

\[ H_\varepsilon(\varphi, I, t, E) = \exp(H_0(I) + E + \varepsilon H_1(\varphi, I, t)). \]

Write this Hamiltonian in the form

\[ H_\varepsilon(\varphi, I, t, E) = H_0 + \varepsilon \mathcal{H}_1(\varphi, I, t, E) \]

Notice that orbits of \( \mathcal{H}_\varepsilon \) coincide with orbits of \( H_\varepsilon \) after constant time change.

Call \( \mathcal{H}_0 \) non-degenerate in \( U \) if for each \( I \in U \) and each \( E \) we have

\[ \det \left( \frac{\partial \omega(I, E)}{\partial(I, E)} \right) \neq 0. \]

Recall that a frequency \( \omega \in \mathcal{D}_{\eta, \tau} \) is called \( (\eta, \tau) \)-Diophantine and it has been defined in [4].

\(^3\text{need to add that as } r \to 0 \text{ the } C^r \text{ norms also converge}\)
Theorem 18. [Pos82] Let $\eta, \tau > 0$ and $r \geq 2n+2\tau$. Let $\mathcal{H}_0$ be a real analytic Hamiltonian on $U$ and $\mathcal{H}_1$ be $C^\infty$ smooth on $U \times \mathbb{R} \times \mathbb{T}^{n+1}$ resp. Suppose $\mathcal{H}_0$ is non-degenerate in $U$. Then there exists $\varepsilon_0 = \varepsilon_0(\mathcal{H}_0, \eta, \tau) > 0$ and $c_0 = c_0(\mathcal{H}_0, r)$ such that for any $\varepsilon$ with $0 < \varepsilon < \varepsilon_0$ and any $\omega \in \mathcal{D}_{\eta, \tau}$ the Hamiltonian $\mathcal{H}_0 + \varepsilon \mathcal{H}_1$ has a $(n+1)$-dimensional (KAM) invariant torus $\mathcal{T}_\omega$ and dynamics restricted to $\mathcal{T}_\omega$ is smoothly conjugate to the constant flow $(\dot{\varphi}, \dot{t}) = \omega$ on $\mathbb{T}^{n+1}$. Moreover, for the union tori

$$\text{Leb}(\cup_{\omega \in \mathcal{D}_{\eta, \tau}} \mathcal{T}_\omega) > (1 - c_0\eta) \cdot \text{Leb}(U \times \mathbb{T}^n \times \mathbb{T}),$$

where $\text{Leb}$ is the Lebesgue measure.

One can improve $\eta$ to be proportional to $\sqrt{\varepsilon}$, but we do not use this modification. Notice that dynamics of $\mathcal{H}_0$ is independent of a choice of $E$. Namely, for any $E \neq E'$ and any initial condition $(\varphi_0, I_0, t_0)$ orbits of $\mathcal{H}_0$ starting $(\varphi_0, I_0, t_0, E)$ and $(\varphi_0, I_0, t_0, E')$ have the same projection along $E$. In particular, if $(\varphi_0, I_0, t_0, E)$ belongs to a KAM torus $\mathcal{T}_\omega$, then $(\varphi_0, I_0, t_0, E')$ also belongs to a KAM torus $\mathcal{T}_{\omega'}$ for $\lambda = \exp(E - E')$. Similarly, if $(\varphi_0, I_0, t_0, E)$ does not belong to a KAM torus, then $(\varphi_0, I_0, t_0, E')$ also does not belong to any KAM torus.

Recall that we have assumed that $H_0(I)$ is called strictly convex (see (1)).

Corollary 6. Let $\eta, \tau > 0$ and $r \geq 2n+2\tau$. Let $H_0$ be a real analytic Hamiltonian on $U$ and $H_1$ be $C^\infty$ smooth on $U \times \mathbb{R} \times \mathbb{T}^n \times \mathbb{T}$ resp. Suppose $H_0$ is strictly convex on $U$. Then there exists $\varepsilon_0 = \varepsilon_0(H_0, \eta, \tau) > 0$ and $c_0 = c_0(H_0, r)$ such that for any $\varepsilon$ with $0 < \varepsilon < \varepsilon_0$ and any $\omega \in \mathcal{D}_{\eta, \tau}$ the Hamiltonian $H_0 + \varepsilon H_1$ has a $(n+1)$-dimensional (KAM) invariant torus $\mathcal{T}_\omega$ and dynamics restricted to $\mathcal{T}_\omega$ is smoothly conjugate to the constant flow $(\dot{\varphi}, \dot{t}) = \lambda \omega$ on $\mathbb{T}^{n+1}$ with $\lambda = \omega^{-1}_{n+1}$. Moreover, for the union tori

$$\mu_E(\cup_{\omega \in \mathcal{D}_{\eta, \tau}} \mathcal{T}_\omega) > (1 - c_0\eta) \cdot \mu_E(U \times \mathbb{T}^n \times \mathbb{T}),$$

where $\mu_E$ is the Lebesgue measure.

To prove this corollary it suffices to show that if $H_0$ is strictly convex in $U$, then $H_0$ is non-degenerate in $U \times \mathbb{R}$.

Lemma 38. If $H_0$ is strictly convex, then the Hessian of $H_0$ is nondegenerate.

Proof. Notice that the Hessian of $H_0$ has the following form

$$H_0 = 
\begin{pmatrix}
\partial^2_{1,1} H_0 + \partial_1 H_0 \partial_1 H_0 & \ldots & \partial^2_{1, n} H_0 + \partial_1 H_0 \partial_n H_0 & \partial_1 H_0 \\
\partial^2_{2,1} H_0 + \partial_2 H_0 \partial_1 H_0 & \ldots & \partial^2_{2, n} H_0 + \partial_2 H_0 \partial_n H_0 & \partial_2 H_0 \\
\vdots & \vdots & \vdots & \vdots \\
\partial^2_{n,1} H_0 + \partial_n H_0 \partial_1 H_0 & \ldots & \partial^2_{n, n} H_0 + \partial_n H_0 \partial_n H_0 & \partial_n H_0 \\
\partial_1 H_0 & \ldots & \partial_n H_0 & 1
\end{pmatrix}.
$$

(64)

92
Denote the last matrix Hess($I, E$). The determinant for this matrix is equal to get
$$H_n^a \det \text{Hess}(I, E).$$

In order to compute $\det \text{Hess}(I, E)$ for each line $j = 1, \ldots, n$ we multiply the last time by $\partial I_j H_0$ and subtract from $j$-th line.

$$H_n^a \det \begin{pmatrix}
\partial^2_{I_1 I_1} H_0 & \cdots & \partial^2_{I_1 I_n} H_0 & 0 \\
\partial^2_{I_2 I_1} H_0 & \cdots & \partial^2_{I_2 I_n} H_0 & 0 \\
\vdots & \cdots & \vdots & \vdots \\
\partial^2_{I_n I_1} H_0 & \cdots & \partial^2_{I_n I_n} H_0 & 0 \\
\partial I_1 H_0 & \cdots & \partial I_n H_0 & 1
\end{pmatrix}.$$

Notice that $H_0$ is always positive and the determinant equals $\det \text{Hess } H_0$. Since $H_0$ is positive definite, determinant is not zero. □

A.1 Regularity of straightening of the union of KAM tori: a simplified version

In this section we state a theorem of Pöschel [Pö82] about simultaneous straightening of all KAM tori within a certain class.

Fix
$$\alpha > 1, \quad \lambda > \tau + n > n,$$

and let $H$ be of class $C^{\alpha, \lambda}$ if it is of class $C^\alpha$ in the $I$-variables, but of class $C^{\lambda}$ in the $\varphi$-variables. The formal definition is given below.

**Theorem 19.** (Pöschel [Pö82]) Let $H_0$ be real analytic and nondegenerate, such that the frequency map $\nabla H_0$ is a diffeomorphism $I \rightarrow \mathbb{R}^n$, and consider a differentiable perturbation $\bar{H} = H_0 + \varepsilon \bar{H}_1$ of class $C^{\lambda+n+1}$ with $\lambda > \tau + n$ and $\alpha > 1$ not in the discrete set $\Lambda = \{i/\lambda + j, \ i, j \geq 0 \text{ integers}\}$. Then, for small $\eta > 0$, one can choose $|\varepsilon|$ proportional to $\eta^2$ so that there exists a diffeomorphism

$$T = \Psi_0 \circ \Phi_\varepsilon : \mathbb{T}^n \times D_{\eta, \tau} \rightarrow \mathbb{T}^n \times \mathbb{R}^n$$

which on $\mathbb{T}^n \times D_{\eta, \tau}$, transforms the Hamiltonian equations of motion into

$$\dot{\theta} = \omega, \quad \dot{\omega} = 0.$$

More precisely, $\Psi_0$ is the real analytic inverse of the frequency map, and $\Phi_\varepsilon$ is a diffeomorphism of class $C^{\alpha, \lambda}$ close to the identity. Its Jacobian determinant uniformly bounded from above and below. In addition, if $H$ is of class $C^{\beta \lambda + \lambda + n - 1 + \tau}$ with $\alpha \leq \beta \leq \infty$, one can modify $\Phi_\varepsilon$ outside $\mathbb{T}^n \times D_{\eta, \tau}$ so that $\Phi_\varepsilon$ is of class $C^{\beta \lambda, \beta}$ for $\beta \notin \Lambda$.

For the proof we need a more technical statement.
A.2 Anisotropic norms

A class of functions with anisotropic differentiability is defined as follows. Let $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ are open sets and set

$$\Omega = \Omega_1 \times \Omega_2.$$ 

Points $x \in \Omega$ are pairs $(x_1, x_2)$ with $x_i \in \Omega_i$, $i = 1, 2$. Similarly we write multi-indices $k$ as pairs $(k_1, k_2)$ of two $n$ vectors with non-negative integer components. For $\rho = (\rho_1, \rho_2)$, $\rho_1, \rho_2 > 0$, we set

$$|x|_\rho = \max\{ |x_1|_{\rho_1}, |x_2|_{\rho_2} \}.$$

$$|k|_\rho = \frac{|k_1|}{\rho_1} + \frac{|k_2|}{\rho_2},$$

where $|x_i|$ denotes the maximum norm, and $|k_i|$ the sum of the moduli of the components of $k_i$.

**Definition 20.** Denote $C^\rho(\Omega) = C^{\rho_1,\rho_2}(\Omega_1, \Omega_2)$ is a class of all functions $u$ on $\Omega$ with partial derivatives $D^k u$ for all $k \in K_\rho = \{|k|_\rho < 1\}$ which satisfy, for some finite $M$,

$$|D^k u| \leq M,$$

$$|D^k u(x) - D^k u(y)| \leq M|x - y|_\rho^\kappa,$$

where $\kappa = 1 - |k|_\rho$ for all $x, y \in \Omega$ and all $k \in K_\rho$, where, however, we require $x_i = y_i$, if $\kappa \rho_i > 1$. Moreover,

$$\|u\|_{\rho, \Omega} = \inf M.$$

is the smallest $M$ for which both inequalities hold.

The $C^\rho(\Omega)$ are Banach spaces with respect to the norm $\| \cdot \|_{\rho, \Omega}$. In addition, we define another norm $\| \cdot \|_{\rho, \Omega; \eta}$ for $\eta > 0$ by

$$\|u\|_{\rho, \Omega; \eta} = \|u \circ \sigma_\eta\|_{\rho, \sigma_\eta^{-1} \Omega},$$

where $\sigma_\eta$ denotes the partial stretching $(x_1, x_2) \to (x_1, \eta x_2)$. This can be viewed as a weighted norm; essentially the $k$-th derivative is multiplied with a weight factor $\eta^{k_2}$. For $\eta \leq 1$ we have

$$\|u\|_{\rho, \eta} \leq \|u\|_{\rho} \leq \eta^{-\rho_2}\|u\|_{\rho, \eta}$$

where we dropped the domains to simplify the notation.
A.3 Regularity of straightening of the union of KAM tori

Let \( \Omega \) be some bounded or unbounded domain in \( \mathbb{R}^n \). We are considering perturbations of an integrable, nondegenerate Hamiltonian system on \( T^n \times \Omega \), described by a real analytic Hamiltonian \( H_0 \) which depends only on the action variable \( p \in \Omega \) and whose frequency map \( \nabla H_0 \) is a diffeomorphism \( \Omega \to \mathcal{D} \). We shall assume, that these properties of the integrable system extend to some complex \( \rho \)-neighborhood

\[
\Omega + \rho = \bigcup_{I \in \Omega} \{ \zeta \in \mathbb{C}^n : |\zeta - I| < \rho \}
\]

of \( I \).

In \( \Omega \) we select the Cantor set of \((\eta, \tau)\)-Diophantine numbers (see 4).

**Theorem 21.** (Pöschel [Pö82]) Let \( \bar{H}_0 \) be real analytic and nondegenerate on \( \Omega + \rho \) for some \( \rho > 0 \) with

\[
|\partial^2_{II} \bar{H}_0|_{\Omega + \rho}, \quad |(\partial^2_{II} \bar{H}_0)^{-1}|_{\Omega + \rho} \leq R.
\]

Furthermore, assume that the frequency map \( \nabla \bar{H}_0 : \Omega \to \mathcal{D} \) is invertible on \( I_{\rho} \). Then, for fixed \( \lambda > \tau + n > n \) and \( \alpha > 1 \), there exists a positive \( \delta = \delta(n, \tau, \lambda, \rho, R) \), but independent of the domain \( \Omega \) or the parameter \( \eta \), such that for a Hamiltonian \( \bar{H} \in C^{\alpha \lambda + \lambda + n - 1 + \tau} (T^n \times \Omega) \) with

\[
\| \bar{H} - \bar{H}_0 \|_{\alpha \lambda + \lambda + n - 1 + \tau; \eta} \leq \eta^2 \delta, \quad 0 < \eta \leq \rho/R,
\]

the following holds.

**Foliation of Invariant Tori:** There exists a diffeomorphism

\[
T = \Psi_0 \circ \Phi_{\varepsilon} : T^n \times \mathcal{D} \to T^n \times \Omega
\]

transforming the Hamiltonian vector field \( X_{\bar{H}} \), on \( T^n \times \Omega \) into the vector field \( T^*X_{\bar{H}} \), on \( T^n \times \mathcal{D}_{\eta, \tau} \) so that

\[
T^*X_{\bar{H}}|_{T^n \times \mathcal{D}_{\eta, \tau}} = \sum \omega_i \frac{\partial}{\partial \varphi_i}
\]

\( \Psi_0 \) is the real analytic inverse of the frequency map leaving the angle variables fixed, and \( \Phi_{\varepsilon} \) is a diffeomorphism on \( T^n \times \mathcal{D} \) of class \( C^{\alpha \lambda + \alpha} \) to be determined in dependence on \( \bar{H} \).

**Solution to Hamilton-Jacobi Equation:** On \( T^n \times I \) there exists a function \( S \) and a nondegenerate Hamiltonian \( \bar{H}_0 \), which is independent of the angle variables, such that

\[
\bar{H}(\varphi, I - S_{\varphi}(\varphi, I))|_{T^n \times \Omega_{\eta, \tau}} = \bar{H}_0(I).
\]

\( \Omega_{\eta, \tau} \) is the inverse image of \( \mathcal{D}_{\eta, \tau} \) under the map \( I \to \bar{H}_0(I) \). The Hamiltonian \( \bar{H}_0 \) is of class \( C^{\alpha + 1} \), and \( S \) is of class \( C^{\alpha \lambda + \alpha} \) with \( \tilde{\alpha} = \alpha - (\lambda - \tau + n + 1)/\lambda \) provided \( \tilde{\alpha} \notin \Lambda \). Denote

\[
\bar{H} \circ \Phi = \bar{H}_0 + \bar{H}_1.
\]

**Additional Smoothness and Estimates:** If \( \bar{H} \in C^{\beta \lambda + \lambda + n - 1 + \tau} (T^n \times I) \) with \( \alpha \leq \beta \) not in \( \Lambda \), then one can modify \( \Phi_{\varepsilon} \) outside \( T^n \times \mathcal{D}_{\eta, \tau} \) so that \( \Phi_{\varepsilon} \in C^{\beta \lambda + \lambda} (T^n \times \Omega) \) with

\[
\| \sigma_{\eta}^{-1}(\Phi - I d) \|_{\beta \lambda, \beta; \eta} \leq c \eta^{-2} \| \bar{H} - \bar{H}_0 \|_{\beta \lambda + \lambda + n - 1 + \tau; \eta}.
\]

95
Corollary 7. In the notations above we have $\partial^a \mathcal{H}_1|_{\text{KAM}} \equiv 0$ for any $|a| \leq \alpha$.

Notice that every point of $\text{KAM}$ is a Lebesgue density point. Therefore, all partial derivatives, that are well defined vanish.

A.4 Application of Pöschel theorem to our setting

In this section we apply Theorem 21 to the situation where $H_0$ is only finitely differentiable

Expand the Hamiltonian $H_0$ near $\xi \in D$ using Taylor formula and write $H_\varepsilon$ as follows:

$$H_\varepsilon(\theta, \xi + \Delta I) = H_0(\xi) + \langle \nabla H_0(\xi), \Delta I \rangle + \frac{1}{2} \langle \partial^2 H_0(\xi) \Delta I, \Delta I \rangle + Q_3(\Delta I, \xi) + \varepsilon H_1(\theta, \xi + \Delta I),$$

where $Q_3(\Delta I, \xi)$ is the cubic remainder. Consider the ball of radius $d = \varepsilon^{1/3}$ around $\xi$. Denote

$$\bar{H}_0(I) = H_0(\xi) + \langle \nabla H_0(\xi), \Delta I \rangle + \frac{1}{2} \langle \partial^2 H_0(\xi) \Delta I, \Delta I \rangle.$$ 

and

$$\varepsilon \bar{H}_1 = \bar{H} - \bar{H}_0 = Q_3(\Delta I, \xi) + \varepsilon H_1(\theta, \xi + \Delta I).$$

Let $r = \alpha \lambda + \lambda + n - 1 + \tau$. Assume that $H_0 \in C^{r+3}$ and $H_1 \in C^r$. Then the remainder $\bar{H}_1 \in C^{r+3}$. The condition

$$|\partial^2_{II} \bar{H}_0|_{\Omega+\rho}, \quad |(\partial^2_{II} \bar{H}_0)^{-1}|_{\Omega+\rho} \leq R$$

are clearly satisfies, because $\partial^2_{II} \bar{H}_0 \equiv \partial^2 H_0(\xi)$.

Application of Theorem 21 gives existence of a diffeomorphism $\Psi_0 \circ \Phi_\varepsilon$ defined in $T^n \times D_{\varepsilon^{1/3}}(\xi)$. Covering $T^n \times D$ by balls of radius $\xi$ and applying Whitney extension theorem (see e.g. [Pöss82]) we obtain a global change of coordinates. Moreover, we have $\mathcal{H}_0 \in C^{\alpha+1}$ and $\mathcal{H}_1 \in C^\alpha$ where $\alpha = (r - n + 1 - \tau)/\lambda$.

B Generic Tonelli Hamiltonians and a Generalized Mapertuis Principle

Consider a $C^r$ smooth function $\mathcal{H} : T^*T^n \to \mathbb{R}$. It is called a Tonelli Hamiltonian if

- $\mathcal{H}(\varphi, I)$ is positive definite, i.e., the Hessian $\partial^2_{II} \mathcal{H}$ is positive definite $\forall (\varphi, I) \in T^n$.
- $\mathcal{H}$ has superlinear growth, i.e. $\mathcal{H}(\varphi, I)/\|I\| \to \infty$ as $\|I\| \to \infty$.

4The authors are grateful to Popov for pointing out this approach
Recall that \( x = (\varphi, I) \) is a critical point if the gradient of \( H \) vanishes. A function \( H \) is called a Morse function if at any critical point, the Hessian \( \partial^2_{xx} H \) is non-singular and critical points have pairwise distinct values. Morse functions form an open dense set in a properly chosen topology (see e.g. \cite{Mi63}).

We say that a property is Mane’s \( C^r \) generic if there is a \( C^r \) generic set of potentials \( U \) on \( \mathbb{T}^2 \) such that \( H_U = H + U \) satisfies this property. We prove that properties in Lemma 4 and Key Theorems 5 and 6 are Mane’s \( C^r \) generic.

We introduce some notations: \( \nabla H(\varphi, I) \) is the gradient, i.e. the set of all partial derivatives, \( \pi: \mathbb{T}^* \mathbb{T} \to \mathbb{T} \) is the natural projection, \( \nabla I H(\varphi, I) \) and \( \nabla \varphi H(\varphi, I) \) are the \( I \) and the \( \varphi \)-gradients, given by the set of partial derivatives with respect to \( I \) and \( \varphi \) respectively.

Lemma 39. The property that \( H \) is a Morse function is Mane’s \( C^r \) generic.

Proof. Due to convexity for each \( \varphi^* \in \mathbb{T}^n \) the map \( \nabla I H(\varphi^*, \cdot): \mathbb{R}^n \to \mathbb{R}^n \) is a diffeomorphism. Therefore, for each \( \varphi^* \) there is at most one point \( I^*(\varphi^*) \) such that \( \nabla I H(\varphi^*, I^*(\varphi^*)) = 0 \). By the implicit function theorem \( I^*(\varphi^*) \) is a smooth function. In particular, outside of some compact set in \( \mathbb{T}^* \mathbb{T}^n \) there are no critical points of \( \nabla H(\varphi, I) = 0 \). Due to superlinearity this compact set is contained in \( \{ H \leq E \} \) for some large \( E \).

Let \( U: \mathbb{T}^n \to \mathbb{R} \) be a potential. Consider a family \( H_U = H + U \) of potential perturbations. Notice that adding a potential does not change the function \( I^*: \mathbb{T}^n \to \mathbb{T}^* \mathbb{T}^n \) and equations for a critical point \( \nabla H_U(\varphi, I) = 0 \) can be rewritten in the form

\[
\nabla \varphi H(\varphi, I)|_{I=I^*(\varphi)} + \nabla U(\varphi) = 0.
\]

Generically in \( U \) this function as a function of \( \varphi \) has isolated solutions to this equation. These solutions lift to critical points of \( H_U \) by means of the function \( I^* \). Once critical points of \( H_U \) are isolated an arbitrary small localized potential perturbation \( H'_U = H_U + U' \) of \( H_U \) we can assure that \( H'_U \) is Morse or, equivalently, all critical points are in \( \{ H'_U \leq E \} \) and nondegenerate. In addition, critical points have pairwise distinct values.

Contreras-Iturriaga-Paternain-Paternain \cite{CIPP98} gave the following definition of \( \alpha(0) \).

Let \( \omega \) be a smooth one form on \( \mathbb{T}^n \). It is a section of the bundle \( \mathbb{T}^n \to \mathbb{T}^* \mathbb{T}^n \). Let \( G_f \subset \mathbb{T}^* \mathbb{T}^n \) be the graph of the differential \( df \) of the smooth function \( f \). It is well known that \( G_f \) is a Lagrangian submanifold of \( \mathbb{T}^* \mathbb{T}^n \). Define

\[
\alpha(0) = \inf_{E \in \mathbb{R}} \mathcal{H}^{-1}(\varphi, E) \text{ contains } G_f \text{ for all } C^\infty \text{ function } f.
\]

This definition is equivalent to the one of Mather as proven in \cite{CIPP98}.

In \cite{CIPP98} it is also shown that if \( E > \alpha(0) \), the dynamics of the Hamiltonian flow of \( H \) restricted to the energy surface \( \{ H = E \} \) is a reparametrization of the geodesic flow on the unit tangent bundle of an appropriately chosen Finsler metric on \( \mathbb{T}^n \). Below we define the corresponding Finsler metric for \( E > \alpha(0) \) and describe this relation.
Lemma 40. Let $\mathcal{H}$ be a Tonelli Hamiltonian and a Morse function. Then $\alpha(0)$ is a critical value of $\mathcal{H}$. Moreover, for $n = 2$ after a possible potential perturbation the corresponding unique critical point $(\varphi^*, I^*)$ is a saddle of the Hamiltonian flow and has real eigenvalues.

Proof. Notice that $E^* = \alpha(0)$ cannot be a regular value of $\mathcal{H}$ as the fact that $\mathcal{H}^{-1}(-\infty, E^*)$ contains the graph $G_f$, implies that $\mathcal{H}^{-1}(-\infty, E^* - \delta)$ also contains $G^{\prime}_f$ for small enough $\delta > 0$ and $f^{\prime}$ close to $f$.

Since critical points have pairwise distinct values, on the critical energy surface $\{\mathcal{H} = E^*\}$ we have exactly one non-degenerate critical point. Show that $\{\mathcal{H} = E^*\}$ has a saddle critical point.

Let $G_f \subset \mathcal{H}^{-1}(-\infty, E^*)$ be a graph of the differential $df$ for some smooth function $f$. The projection $\pi \mathcal{H}^{-1}(-\infty, E^* - \delta)$ cannot contain $\mathbb{T}^n$, otherwise, we can deform $f$ to $f^{\prime}$ and have $G^{\prime}_f \subset \mathcal{H}^{-1}(-\infty, E^* - \delta)$.

Consider a neighborhood $V$ of the only critical point $(\varphi^*, I^*) \in \mathbb{T}^n \mathbb{T}^n$ with $\mathcal{H}(\varphi^*, I^*) = E$. Denote by $\pi_n$ the natural projection of $\mathbb{T}^n \mathbb{T}^n \to \mathbb{T}^n$. By Morse lemma (see e.g. [Mil63]) in local coordinates $(x_1, \ldots, x_{2n}) = \Phi(\varphi, I)$ for some $a$ we have

$$\mathcal{H} \circ \Phi^{-1}(x) = a + \sum_{i=1}^k x_i^2 - \sum_{i=k+1}^{2n} x_i^2.$$ 

Thus, $\pi_n$ is a smooth projection onto a $n$-dimensional surface $\Phi((\mathbb{T}^n \times 0) \cap V)$ given by $\Phi \pi_n \circ \Phi^{-1}$. Denote by $K_L$ the kernel of $\pi_n(0)$. Perturbing $\mathcal{H}$, if necessary, make the subspaces $x_1 \mathbb{R}, \ldots, x_{2n} \mathbb{R}$ transverse to $K_L$ and, therefore, transverse to the kernel of $\pi_n(x)$ for all $x \in U = \Phi(V)$, i.e. near 0. Due to fiber convexity of $\mathcal{H}$ we have that for all $x \in U$ the set $\{\mathcal{H}^{-1}(a') \cap \mathbb{T}^n \mathbb{T}^n\}$ is convex (if nonempty). We know that for small $\delta > 0$

- the projection of $\pi_n(\mathcal{H}^{-1}(a + \delta))$ contains $U$,
- the projection of $\pi_n(\mathcal{H}^{-1}(a - \delta))$ does not contain $U$.

If $k < n$ convexity in the fiber $K_L$ fails. If $k > n$, $\pi_n \mathcal{H}^{-1}(a + \delta)$ does not contain $U$.

In the cases $k = n$ the energy surface $\{\mathcal{H} = a\}$ locally consists of two transverse $n$-dimensional surfaces, that are invariant. If $n = 2$, then linearizing the Hamiltonian vector field at 0 we have two invariant subspaces of dimension two. Consider the eigenvalues of the linearization. Since the vector field is Hamiltonian, each eigenvalue $\lambda$ has a pair $\lambda^{-1}$. The linearization is real, thus, all complex eigenvalues come in pairs: $\lambda \notin \mathbb{R}$ and its conjugate $\bar{\lambda}$. Therefore, if eigenvalues are away from the unit circle, they form quadruples.

By a perturbation move eigenvalues away from 1. If a pair of complex conjugate eigenvalues are on the unit circle, then one pair of conjugate variables is the sum of squares and enters into the Hamiltonian with one sign, while the other pair has the opposite sign to keep $k = 2$. This contradicts fiber convexity. Therefore, eigenvalues should be away from the unit circle. If some nonreal eigenvalues are away from the unit circle, they come in quadruples. In this case there is no pair of invariant 2-dimensional subspaces. Thus, $(\varphi^*, I^*)$ is a saddle with real eigenvalues.  

\[\square\]
Denote by $E_{\text{min}} = \alpha(0)$. Consider a mechanical system $\mathcal{H}(\varphi, I) = K(I) - U(\varphi)$, where $K(I)$ is a positive definite quadratic form and $U(\varphi)$ is a smooth potential on $\mathbb{T}^2$. One can associate $\mathcal{H}$ the Jacobi metric $g_E(\varphi) = 2\sqrt{E + U(\varphi)}K$. Let $E > E_{\text{min}}$. According to the Mapertuis principle the orbits of $\mathcal{H}$ on the energy surface $\{\mathcal{H} = E\}$ are time reparametrization of geodesics of $g_E$. In [KZ12] as a slow system in a double resonance we study mechanical systems of the above form $\mathcal{H} = K - U$ and state all non-degeneracy conditions on the Hamiltonian $\mathcal{H}$ using Mapertuis principle.

In our case we cannot assure that the slow system $\mathcal{H}$ is a mechanical system. However, for a Tonelli Hamiltonian $\mathcal{H}$ there is a generalized Mapertuis principle, described below.

### B.1 A generalized Mapertuis principle

Let $L$ be a Lagrangian given as a Legendre transform of $\mathcal{H}$. Given any $c \in H^1(\mathbb{T}^n, \mathbb{R}) \simeq \mathbb{R}^n$, define

$$\alpha(c) = \inf \int (L - c \cdot I) d\mu(\varphi, I),$$

where the infimum is taken over all probability measures $\mu$ in $T\mathbb{T}^n$. It is called Mather’s $\alpha$-function. $\alpha(c)$ is the average of the Hamiltonian on the support of the $c$-minimal measures.

Denote by $L_c(\varphi, v) = L(\varphi, v) - c \cdot v$ for any $c \in \mathbb{R}^n$. Let $\mathcal{H}_c$ be the Legendre-Fenchel dual of $L_c(\varphi, v)$, then it has the form $\mathcal{H}(\varphi, I + c)$. For any $t > 0$, $x, y \in \mathbb{T}^n$ and $c \in \mathbb{R}^n$ we introduce the following quantity:

$$h^t_c(x, y) = \int_0^t L_c(\xi(s), \dot{\xi}(s)) ds,$$

where the infimum is taken over of the piecewise $C^1$ curve $\xi : [0, t] \rightarrow \mathbb{T}^n$ such that $\xi(0) = x$ and $\xi(t) = y$. It is well know that we can define the Mañé’s critical potential and Peierls barrier respectively as follows

$$\varphi_c(x, y) = \inf_{t > 0} h^t_c(x, y) + \alpha(c) t$$

$$h_c(x, y) = \lim_{t \to \infty} \inf_{t > 0} h^t_c(x, y) + \alpha(c) t.$$

For given $c \in \mathbb{R}^n$, define the projected Aubry set

$$\mathcal{A}_c = \{x \in \mathbb{T}^n : h_c(x, x) = 0\}.$$

In order to define a Finsler metric associate to the flow of $\mathcal{H}$ on the energy surface $\{\mathcal{H} = E\}$ we need the following definition. We use the variant approach in [FS04], where only the case $c = 0$ was considered. For any fixed $x \in \mathbb{T}^n$, define the sublevel set

$$\bar{Z}_c(x) = \{I \in \mathbb{R}^n : \mathcal{H}_c(x, I) \leq \alpha(c)\}, \quad c \in \mathbb{R}^n,$$

and

$$\delta_c(x, v) = \sigma_{\bar{Z}_c(x)}(v), \quad x \in \mathbb{T}^n, \quad v \in \mathbb{R}^n,$$
where \( \sigma_C(v) \) is the support function with respect to the compact convex set \( C \), i.e.,

\[
\sigma_C(v) = \max\{ \langle x, v \rangle : x \in C \}, v \in \mathbb{R}^n.
\]

Denote by \( \sigma_{\alpha(c)}(x, v) = \sigma_{Z_\alpha(x)}(v) \), where \( Z_\alpha(x) = \{ I \in \mathbb{R}^n : \mathcal{H}(x, I) \leq \alpha(c) \} \).

For any \( x, y \in \mathbb{T}^n \), define

\[
S_c(x, y) = \inf \int_0^1 \delta_c(\xi(t), \dot{\xi}(t)) \, dt = \inf \int_0^1 (\sigma_c(\xi(t), \dot{\xi}(t)) - c\dot{\xi}(t)) \, dt,
\]

where the infimum is taken over of the piecewise \( C^1 \) curve \( \xi : [0, t] \to \mathbb{T}^n \) such that \( \xi(0) = x \) and \( \xi(t) = y \).

It is natural to call the quantity \( \delta_c(x, v) \) — the Jacobi–Finsler metric for the generalized Maupertuis principle with the restriction of the energy \( \alpha(c) \). If the kinetic energy function is of the form of Riemannian metric \( g_x(v, v) = \langle v, v \rangle_x \), \( \delta_0(x, v) \) is the usual Jacobi metric \( \sqrt{E - U(x)} \), where \( E = \alpha(c) > \min_c \alpha(c) \). In [Che12] the following useful statement if proven.

**Theorem 22.** The Finsler length of \( \delta_c(x, v) \) and action \( L_c + \alpha(c) \) coincide, i.e.

\[
\varphi_c(x, y) = S_c(x, y).
\]

It follows from Lemma [40] that

**Corollary 8.** Let \( n = 2 \) and \( \mathcal{H}(\varphi, I) \) be a Tonelli Hamiltonian, then the property that \( \mathcal{H} \) is a Morse function and the only critical point \( (\varphi^*, I^*) \in \mathcal{H}^{-1}(\alpha(0)) \) of the Hamiltonian flow associated to \( \mathcal{H} \) is a saddle with real eigenvalues is Mañé generic.

For Tonelli Hamiltonians satisfying this property the associated Jacobi-Finsler metric \( \delta_0 \) has exactly one critical point at \( \varphi^* \).

### B.2 Proof of Lemma [4]

**Proof.** We use a proof from unpublished notes by Kaloshin and Zhang. It is easy to see that all the properties are open. To prove density, we show that by an arbitrarily small perturbation by a potential, the system satisfies condition of this lemma.

Let \( S \subset \mathbb{T}^3 \) be a smooth global section not containing \( m_0 \), and intersecting every curve in homology \( \bar{h} \). Lift \( S \) and the metric \( \delta_0 \) to the universal cover. Denote by \( l_\delta(\gamma) \) \( \delta_0 \)-length of \( \gamma \). Consider the variational problem

\[
\min_{\varphi \in S} l_\delta(\gamma), \quad \gamma(0) = \varphi, \gamma(T) = x + \bar{h}.
\]

Any minimizer \( \gamma_0 \) projects to a shortest geodesic of homology \( \bar{h} \). Perturbing the section if necessary, we may assume at least one of the minimizer intersect \( S \) transversally, i.e. \( \gamma_0 \cap S \) at isolated points. Let \( \varphi_0 \) be one such point, for \( \delta > 0 \), let \( U_0(\varphi) \) be a function satisfying:
• $U(\phi_0) = 0$ and $U(\phi) \geq 0$ for all $\phi \neq \phi_0$.

• $U(\phi) = \text{const}$ for $\phi \notin B_\delta(\phi_0)$ and $\|U\|_{C^r} = o(1)$ as $\delta \to 0$.

Consider the new potential function $U + U_0$, for the new variational problem there exists a unique minimizer intersecting $S$ at $\phi_0$.

Since $\gamma_0$ is a geodesic outside of the critical point $m_0$, it has no self-intersection outside of $m_0$.

If $m_0 \notin \gamma_0$, then $\gamma_0$ must be a simple closed curve of homology $\bar{h}$.

If $m_0 \in \gamma_0$ and $\gamma_0$ is not a simple closed curve, then it must be concatenation of simple closed curves, intersecting at $m_0$. We assume that for each integer homology class $g$ with $l_\delta(g) \leq l_\delta(\bar{h})$ the minimizer is unique. This is open and dense from the earlier argument. Then the components of $\gamma_0$ must have different homology classes.

We claim: for a collection of simple curves with different homologies in $\mathbb{T}^2$, to avoid intersection at more than one point, the only possibility is the following (This is similar to the discussion of Mather in [Mat03]):

• There are two curves $\gamma_1$ and $\gamma_2$ with homologies $h_1$ and $h_2$ generating $\mathbb{Z}^2$, and if there is a third simple curve, it must have homology $\pm(h_1 + h_2)$.

To see this, make a (integer linear) change of basis such that $h_1 = (0, 1)$. It’s easy to see that unless $h_2 = \pm (1, n)$, the curves $\gamma_1, \gamma_2$ intersect at two points. $h_1, h_2$ clearly form a basis. Now $\mathbb{T}^2 \setminus \gamma_1 \cup \gamma_2$ is an open rectangle with $m_0$ at four vertices, and the only room left is the diagonal, i.e. a simple curve of homology $\pm(h_1 + h_2)$.

We now make an additional perturbation, such that

$$l_\delta(h_1) + l_\delta(h_2) < l_\delta(h_1 + h_2),$$

the system now fall into the non-simple case as given in the condition of the Lemma. \qed

References

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

[Arn89] V. I. Arnol’d. Mathematical methods of classical mechanics, volume 60 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1989.

[Ber08] P. Bernard. The dynamics of pseudographs in convex Hamiltonian systems. J. Amer. Math. Soc., 21(3):615–669, 2008.

[BKZ11] P. Bernard, V. Kaloshin, and K. Zhang. Arnold diffusion in arbitrary degrees of freedom and crumpled 3-dimensional normally hyperbolic invariant cylinders. Preprint available at http://arxiv.org/abs/1112.2773, 2011.
A. Bounemoura. Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians. *J. Differential Equations*, 249(11):2905–2920, 2010.

J. W. S. Cassels. *An introduction to Diophantine approximation*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 45. Cambridge University Press, New York, 1957.

L. Chierchia and G. Gallavotti. Drift and diffusion in phase space. *Ann. Inst. H. Poincaré Phys. Théor.*, 60(1):144, 1994.

C. Cheng. Arnold diffusion in nearly integrable hamiltonian systems. 2012.

G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain. Lagrangian graphs, minimizing measures and Mañé’s critical values. *Geom. Funct. Anal.*, 8(5):788–809, 1998.

C.Q. Cheng and J. Yan. Existence of diffusion orbits in a priori unstable Hamiltonian systems. *J. Differential Geom.*, 67(3):457–517, 2004.

C.Q. Cheng and J. Yan. Arnold diffusion in hamiltonian systems: apriori unstable case. *J. Differential Geom.*, 82:229, 2009.

C. Cheng and J. Zhang. Asymptotic trajectories of kam torus. Preprint available at [http://arxiv.org/abs/1312.2102](http://arxiv.org/abs/1312.2102), 2013.

A. Delshams, R. de la Llave, and T.M. Seara. A geometric mechanism for diffusion in hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. *Mem. Amer. Math. Soc.*, 2006.

A. Delshams, R. de la Llave, and T. M. Seara. Geometric properties of the scattering map of a normally hyperbolic invariant manifold. *Adv. Math.*, 217(3):1096–1153, 2008.

A. Delshams, R. de la Llave, and T. M. Seara. Instability of high dimensional hamiltonian systems: Multiple resonances do not impede diffusion. 2013.

A. Delshams and G. Huguet. Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems. *Nonlinearity*, 22(8):1997–2077, 2009.

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

A. Fathi. Weak kam theorem in lagrangian dynamics. Seventh Preliminary Version, 2011.

A. Fathi and A. Siconolfi. Existence of $C^1$ critical subsolutions of the Hamilton-Jacobi equation. *Invent. Math.*, 155(2):363–388, 2004.
[KH07] V. Y. Kaloshin and B. R. Hunt. Stretched exponential estimates on growth of the number of periodic points for prevalent diffeomorphisms. I. *Ann. of Math. (2)*, 165(1):89–170, 2007.

[KMV04] V. Kaloshin, J. Mather, and E. Valdinoci. Instability of totally elliptic points of symplectic maps in dimension 4. *Astérisque*, to appear in 2004.

[KS12] V. Kaloshin and M. Saprykina. An example of a nearly integrable Hamiltonian system with a trajectory dense in a set of maximal Hausdorff dimension. *Comm. Math. Phys.*, 315(3):643–697, 2012.

[KZ12] V. Kaloshin and K. Zhang. A strong form of arnold diffusion for two and a half degrees of freedom. Preprint available at [http://www.terpconnect.umd.edu/~vkaloshi/](http://www.terpconnect.umd.edu/~vkaloshi/), 2012.

[KZZ09] V. Kaloshin, K. Zhang, and Y. Zheng. Almost dense orbit on energy surface. *Proceedings of XVITH International Congress on Mathematical Physics. Published by World Scientific Publishing Co*, pages 314 – 322, 2009.

[LCD83] P Le Calvez and R. Douady. Exemple de point fixe elliptique non topologiquement stable en dimension 4. *C.R.Acad.Sci. Paris*, 296:895–898, 1983.

[Mat03] J.N. Mather. Arnold diffusion II. *Preprint*, 2003.

[Mil63] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

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

[Sul94] J. M. Sullivan. Sphere packings give an explicit bound for the Besicovitch covering theorem. *J. Geom. Anal.*, 4(2):219–231, 1994.

[SZ89] D. Salamon and E. Zehnder. KAM theory in configuration space. *Comment. Math. Helv.*, 64(1):84–132, 1989.

[Tre04] D. Treschev. Evolution of slow variables in a priori unstable hamiltonian systems. *Nonlinearity*, 17(5):1803–1841, 2004.

[Tre12] D. Treschev. Arnold diffusion far from strong resonances in multidimensional a priori unstable hamiltonian systems. 2012.

[Yoc92] J. C. Yoccoz. Travaux de herman sur les tores invariants. *Sminaire N. Bourbaki*, 754:311–344, 1991–1992.