Università degli Studi di Padova

DIPARTIMENTO DI INGEGNERIA CIVILE, EDILE E AMBIENTALE

Corso di Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

KAM theory for properly–degenerate systems: remarks on the measure of the invariant set

Master Candidate
Liu Xiang
Student Number: 1216948

Supervisor
Prof. Gabriella Pinzari

Academic Year: 2019-2020

Abstract

The thesis extends the properly-degenerate KAM theory proposed by L. Chierchia and G. Pinzari to a very general setting. The original theorem assumes that the averaged Hamiltonian function possesses the first order Birkhoff invariants which are non resonant up to order 4. And this thesis generalizes this assumption to an arbitary order. Following the original proof by L. Chierchia and G. Pinzari, the measure of the set of invariant tori in the new setting is also generalized.

Acknowledgements

I would like to thank my advisor Prof. Gabriella Pinzari for giving me this opportunity to pursue my thesis. And I believe that this would be a good start to my future study. I would like to express my sincere and deep gratitude for her unconditional guidance and support. I felt very lucky that all my questions, no matter how stupid they were, could be answered and clearly explained.

A special thanks goes to the Padova University, and the students and professors I met here. It's my honor to study with them, to talk with them and to enjoy the time with them.

Contents

Chapter 1

Introduction

Is the solar system stable? Since Newton's era this has been a key problem of celestial dynamics. To answer this question, we are required to study the integrability of perturbed Hamiltonian. In 1954, Kolmogorov reported his breakthrough result [10, 11] on the integrability of the perturbed Hamiltonian system at the International Congress of Mathematicians, which marked the beginning of the KAM theory.

After that, J. Moser and V.I. Arnold enriched the result [12, 1], which states that a small perturbation added to an integrable Hamiltonian system will not break the major part of the unperturbed motions, provided that the "non-degeneracy" conditions hold for the integrable part of the perturbed Hamiltonian system.

Unfortunately, the result cannot be directly applied to the solar system, or more generally, the planetary $(n+1)$-body problem. The planetary problem is equivalent to the integrable n uncoupled two-body problem with a small perturbation, which is due to the interactions among the planets. However, its associate integrable part does not depend fully on the action variables, which, called proper degeneracy, violates the non-degeneracy condition. This problem was solved in 1963 with the "Fundamental Theorem" proposed by Arnold, and the following result was obtained for the planar, nearly-circular three-body problem

"For the majority of initial conditions under which the instantaneous orbits of the planets are close to circles lying in a single plane, perturbation of the planets on one another produces, in the course of an infinite interval of time, little change on these orbits provided the masses of the planets are sufficiently samll.

In particular, ... in the n-body problem there exists a set of initial conditions having a positive Lebesgue measure and such that, if the initial positions and velocities of the bodies belong to this set, the distances of the bodies from each other will remain perpetually bounded. "

It was until the 1995 that the result in the general spatial three-body problem was proved. And in 2004, the n-body case was first completely proved with a different KAM method. And in [3], which is the main reference of the thesis, the hypotheses of the Fundamental theorem were weakened and the measure estimates on the Kolmogorov set was improved. And the new theorem was successfully applied to the planetary n-body problem [8].

A properly–degenerate system is a conservative system whose Hamiltonian function is of the form [3, 8]

$$H(I, \varphi, p, q; \mu) := H_0(I) + \mu f(I, \varphi, p, q; \mu)$$

where $(I, \varphi) \in V \times \mathbb{T}^{n_1} \subset \mathbb{R}^{n_1} \times \mathbb{T}^{n_1}$ and $(p, q) \in B^{2n_2} \subset \mathbb{R}^{2n_2}$ are standard symplectic variables, here V is an open, connected, bounded subset of $\mathbb{R}^{n_1}$, B^{2n_2} is a ball around the origin, and, finally, μ is a small and positive parameter. The phase space is

$$\mathcal{P} := V \times \mathbb{T}^{n_1} \times B^{2n_2}$$

We have the following assumptions:

- **(a)** The "frequency map" $I \to \omega_0(I) := \partial_I H_0(I)$ is a diffeomorphism in a complex open neighborhood of V;
- **(b)** The "averaged perturbation"

  $$f_{av}(p, q; I, \mu) := \frac{1}{(2\pi)^{n_1}} \int_{\mathbb{T}^{n_1}} f(I, \varphi, p, q; \mu) d\varphi$$

  is "in Birkhoff normal form of order 2s", i.e.:

  $$f_{av}(p, q; I, \mu) = f_0(I) + \Omega(I) \cdot r + \frac{1}{2} r \cdot B(I) r + \cdots + \sum_{i_1, \cdots, i_s \in \{1, \cdots, n_2\}} b_i(I) r_{i_1} \cdots r_{i_s} + O(|(p, q)|^{2s+1})$$

  where $r = (r_1, \cdots, r_{n_2})$ and $r_i := \frac{p_i^2 + q_i^2}{2}$;

- **(c)** The "first order Birkhoff invariants are non resonant up to order $2s$", i.e., the vector–valued function $I \to \Omega(I) = (\Omega_1(I), \cdots, \Omega_{n_2}(I))$ verifies the following inequality

$$|\Omega(I) \cdot k| = |\sum_{j=1}^{n_2} \Omega_j(I) k_j| \geq const > 0, \ \forall I \in V, \ \forall 0 < |k| \leq 2s, \ k \in \mathbb{Z}^{n_2}$$

- **(d)** The "matrix of the second order Birkhoff invariants is not singular", i.e.,

$$|\det(B(I))| \geq \text{const} > 0, \ \forall I \in V$$

Properly–degenerate systems have been introduced by V.I.Arnold in [1] as a mathematical tool for the study of the stability of planetary systems. Such systems were next reconsidered in the paper [3], in turn based on [8]. The case treated in this thesis is a particularization of the statements in [8, 3], where the integer number s in item **(b)** above was chosen to be 2 (in [1] it was taken to be 3). The generalization is here introduced in order to obtain a finer estimate of the so–called "resonant set" in the case of high order non–resonance of the first order Birkhoff invariants. We argue some problems of Celestial Mechanics would benefit of such a finer estimate. As an example, it can be applied to the planetary problem, as the first order Birkhof invariants of the planetary Hamiltonian written in a suitable set of coordinates are non–resonant at any order, as proved in [4].
Techniques used here strictly follow [8, 3].

Denote by $B_\epsilon = B_\epsilon^{2n_2} = \{y \in \mathbb{R}^{2n_2} : |y| < \epsilon\}$ the $2n_2-$ball of radius ϵ and let

$$\mathcal{P}_\epsilon := V \times \mathbb{T}^{n_1} \times B_\epsilon$$

We prove the following theorem:

Theorem 1 ([8, 3])**.** *Let H be real-analytic on $\mathcal{P}$ and assume **(a)**, **(b)**, **(c)** and **(d)**, and let n_1, n_2 be positive integers, $\tau > n_1$, and $\tau_* > n := n_1 + n_2$. Then, there exist positive numbers $\epsilon_* < 1$, $C_*, \gamma_* > 1$ such that, if the following conditions hold*

$$0 < \epsilon < \epsilon_*, \quad \mu < C_* \epsilon^2 \tag{1.1}$$

and if $\bar{\gamma}, \gamma_1, \hat{\gamma}_2$ are taken so as to satisfy $\mu \hat{\gamma}_2 \leq \gamma_1$ and

$$\begin{cases} \gamma_* \max\{\sqrt{\mu}(\log(\epsilon^{-2s-1}))^{\tau_*+1}, \epsilon(\log(\epsilon^{-2s-1}))^{\tau_*+1}\} < \bar{\gamma} < \gamma_* \\ \gamma_* \epsilon^{(2s+1)/2}(\log(8(12)^{2(\tau+1)}))^{\tau_*+1} < \gamma_1 < \gamma_* \\ \gamma_* \epsilon^{(2s+1)/2}(\log(8(12)^{2(\tau+1)}))^{\tau_*+1} < \hat{\gamma}_2 < \gamma_* \epsilon^2 \end{cases} \tag{1.2}$$

there exists a set $\mathcal{K} \subset \mathcal{P}$ formed by the union of $H-$invariant $n-$dimensional tori, on which the $H-$motion is analytically conjugated to linear Diophantine quasi-periodic motions. The set $\mathcal{K}$ is ok positive measure and satifies

$$\begin{aligned} meas(\mathcal{P}_\epsilon) > meas(\mathcal{K}) > \Big[&(1 - \bar{C}\bar{\gamma}) \left(1 - C_1 \epsilon^{(2s-3)n_2/2}\right) \\ &- C_2 \left(\gamma_1 + \frac{\hat{\gamma}_2}{\epsilon^2} + \bar{\gamma} + \epsilon^{(2s-3)n_2/2} \right) \Big] meas(\mathcal{P}_\epsilon) \end{aligned} \tag{1.3}$$

Chapter 2

Notations and Useful Tools

First we introduce some basic notations and definitions.

- in $\mathbb{R}^{n_1}$ we fix the 1-norm: $|I| := |I|_1 := \sum_{i=1}^{n_1} |I_i|$;
- in $\mathbb{T}^{n_1}$ we fix the sup-metric: $|\varphi| := |\varphi|_\infty := \max_{1\le i\le n_1} |\varphi_i| \pmod{2\pi}$;
- in $\mathbb{R}^{n_2}$ we fix the sup-norm: $|p| := |p|_\infty := \max_{1\le i\le n_2} |p_i|, |q| := |q|_\infty := \max_{1\le i\le n_2} |q_i|$;
- if $A \subset \mathbb{R}^{n_i}$, or $A \subset \mathbb{T}^{n_1}$, and $r > 0$, with the specific norms defined above, the complex $r-$neighborhood of A is denoted by

  $$A_r := \bigcup_{x\in A} \{z \in \mathbb{C}^{n_1} : |z - x| < r\}$$

- let $U \subset \mathbb{R}^d$, a real-analytic function $f : U \times \mathbb{T}^m \to \mathbb{R}$ is identified with its analytic extension $\tilde{f} : U_r \times \mathbb{T}^m_s \to \mathbb{R}$ over a $(r, s)-$ neighborhood of its domain. The sup-Fourier norm $\|f\|_{U_r\times\mathbb{T}^m_s} := \|f\|_{r,s}$ is defined by

  $$\|f\|_{r,s} := \sum_{k\in\mathbb{Z}^m} \sup_{I\in U_r} |f_k(I)|e^{|k|s}$$

  where $|k| := |(k_1, \cdots, k_m)| := \sum_{i=1}^m |k_i|$, and $f_k(I)$ is the k-th Fourier coefficient of f:

  $$f_k(I) := \frac{1}{(2\pi)^m} \int_{\mathbb{T}^m} f(I, \varphi)e^{-ik\cdot\varphi} d\varphi$$

note that the 0th Fourier coefficient $f_0(I)$ is the the averaged term of f with respect to φ:

$$\bar{f}(I) := f_0(I) = \frac{1}{(2\pi)^m} \int_{\mathbb{T}^m} f(I,\varphi) d\varphi$$

and we also have that

$$f(I,\varphi) = \sum_{k \in \mathbb{Z}^m} f_k(I) e^{ik\cdot\varphi}$$

If $K > 0$ and Λ is a sub-lattice of $\mathbb{Z}^m$, $T_K f$ and $\Pi_\Lambda f$ denote the K-truncation and the Λ-projection of f

$$T_K f := \sum_{|k|\le K} f_k(I) e^{ik\cdot\varphi}, \quad \Pi_\Lambda f := \sum_{k\in\Lambda} f_k(I) e^{ik\cdot\varphi}$$

and if $K = 0$ or $\Lambda = \{0\}$ we have $T_0 f = \Pi_{\{0\}} f = f_0(I) = \bar{f}(I)$.

- given specific norms, $f : A \subset \mathbb{R}^d \to \mathbb{R}^n$ is Lipschitz if

$$\mathcal{L}(f) := \sup_{I \neq I' \in A} \frac{|f(I) - f(I')|}{|I - I'|}$$

exists and $\mathcal{L}(f)$ is called Lipschitz constant of f.

A function f is called bi-Lipschitz if f is a injective Lipschitz function with a Lipschitz inverse function. For bi-Lipschitz functions we have two Lipschitz constants $0 < \mathcal{L}_-(f) \le \mathcal{L}_+(f)$ such that for all $I, I' \in A$

$$\mathcal{L}_-(f)|I - I'| \le |f(I) - f(I')| \le \mathcal{L}_+(f)|I - I'|$$

where

$$\mathcal{L}_+(f) = \mathcal{L}(f), \quad \mathcal{L}_-(f) = \frac{1}{\mathcal{L}(f^{-1})}$$

- We denote by $\mathcal{D}_{\gamma,\tau} \subset \mathbb{R}^n$ the Diophantine (γ,τ)-numbers

$$\mathcal{D}_{\gamma,\tau} = \{\omega \in \mathbb{R}^n \| \omega \cdot k| \ge \frac{\gamma}{|k|^\tau}, \ \forall k \in \mathbb{Z}^n \backslash \{0\}\}$$

- $\mathcal{D}_{\gamma_1,\gamma_2,\tau} \subset \mathbb{R}^{n_1+n_2}$ denotes the "two–scale Diophantine set", defined as the set of vectors $\omega = (\omega_1, \omega_2) \in \mathbb{R}^{n_1+n_2}$ satisfying, $\forall k = (k_1, k_2) \in (\mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2}) \backslash \{0\}$,

$$|\omega \cdot k| = |\omega_1 \cdot k_1 + \omega_2 \cdot k_2| \ge \begin{cases} \frac{\gamma_1}{|k|^\tau}, & \text{if } k_1 \neq 0 \\ \frac{\gamma_2}{|k|^\tau}, & \text{if } k_1 = 0, k_2 \neq 0 \end{cases}$$

Note that, when $\gamma_1 = \gamma_2 := \gamma$, then $\mathcal{D}_{\gamma,\gamma,\tau} = \mathcal{D}_{\gamma,\tau}$, with $n = n_1 + n_2$. Here we take $\gamma_1 > \gamma_2$.

2.1 Averaging theory

The proof of Theorem 1 strictly follows the proof of Proposition 3 in [3]. As in such papers, we use the "normal form theorem" due to [6, 7], in the form revisited in [9] and adapted to "two–scale systems", as in [8, 3].

Proposition 1 ([9, 8, 3])**.** *(Averaging theory) Let $\bar K$, $\bar s$ and s be positive numbers such that $\bar K s \geq 6$ and let $\alpha_1 \geq \alpha_2 > 0$; let $A \times B \times B' \subset (\mathbb{R}^{l_1} \times \mathbb{R}^{l_2}) \times \mathbb{R}^m \times \mathbb{R}^m$, and $v = (r, r_p, r_q)$ a triple of positive numbers. Let $H := h(I) + f(I, \varphi, p, q)$ be a real-analytic Hamiltonian on $W_{v,\bar s+s} := A_r \times B_{r_p} \times B'_{r_q} \times \mathbb{T}^{l_1+l_2}_{\bar s+s}$. Finally, let Λ be a (possibly trivial) sub-lattice of $\mathbb{Z}^{l_1+l_2}$ and let $\omega = (\omega_1, \omega_2)$ denote the gradient $(\partial_{I_1} h, \partial_{I_2} h) \in \mathbb{R}^{l_1+l_2}$. Let $k = (k_1, k_2) \in \mathbb{Z}^{l_1} \times \mathbb{Z}^{l_2}$ and assume that*

$$|\omega \cdot k| \geq \begin{cases} \alpha_1, & \text{if } k_1 \neq 0 \\ \alpha_2, & \text{if } k_1 = 0 \end{cases} \forall I \in A_r, \forall k = (k_1, k_2) \notin \Lambda, |k| \leq \bar K \tag{2.1}$$

$$E := \|f\|_{v,\bar s+s} < \frac{\alpha_2 d}{2^7 c_m \bar K s}, \text{ where } d := \min\{rs, r_p r_q\}, c_m := \frac{e(1+em)}{2} \tag{2.2}$$

Then, there exists a real-analytic, symplectic transformation

$$\Psi : (I', \varphi', p', q') \in W_{v/2,\bar s+s/6} \to (I, \varphi, p, q) \in W_{v,\bar s+s}$$

such that

$$H_* := H \circ \Psi = h + g + f_* \tag{2.3}$$

with g in normal form and f_* small:

$$g = \sum_{k \in \Lambda} g_k(I', p', q') e^{ik\cdot\varphi'} \tag{2.4}$$

with

$$\begin{aligned} \|g - \Pi_\Lambda T_{\bar K} f\|_{v/2,\bar s+s/6} &\leq \frac{12}{11} \frac{2^7 c_m E^2}{\alpha_2 d} \leq \frac{E}{4}, \\ \|f_*\|_{v/2,\bar s+s/6} &\leq e^{-\bar K s/6} \frac{2^9 c_m E^2}{\alpha_2 d} \leq e^{-\bar K s/6} E \end{aligned} \tag{2.5}$$

Moreover, denoting by $z = z(I', \varphi', p', q')$, the projection of $\Psi(I', \varphi', p', q')$ onto the $z-$variables ($z = I_1, I_2, \varphi, p$ or q) one has

$$\max\{\alpha_1 s|I_1 - I'_1|, \alpha_2 s|I_2 - I'_2|, \alpha_2 r|\varphi - \varphi'|, \alpha_2 r_q|p - p'|, \alpha_2 r_p|q - q'|\} \leq 9E \tag{2.6}$$

Note by setting the constant $\bar{K}$ to be very large (with respect to s), we may obtain a transformed Hamiltonian H_* with a very small perturbed term f_* (see equation (2.5)). And since the choice of m is arbitary, we can set $m = 0$ to eliminate the dependence of the Hamiltonian on the variables (p, q). Furthermore, if we set $\Lambda := \{0\}$ we find that the term g is only dependent on the variable I (see equation (2.4)), according to which we may apply the average theory infinitely many times. For the proof of this proposition the reader may refer to [2, 3].

2.2 Birkhoff normal form

Proposition 2. *(Birkhoff normal form [3, 5]) Let $\alpha > 0, s \geq 3$, and $\Omega = (\Omega_1, \cdots, \Omega_m) \in \mathbb{R}^m$ be non-resonant of order s, i.e.,*

$$|\Omega \cdot k| \geq \alpha > 0, \ \forall k \in \mathbb{Z}^m \text{ with } 0 < |k| \leq s \tag{2.7}$$

and let $z = (p, q) \in B^{2m}_{\epsilon_0} = \{z : |z| < \epsilon_0\} \subset \mathbb{R}^{2m} \to H(z)$ be a real-analytic function of the form

$$H(z) = \sum_{i=1}^{m} \Omega_i r_i + O(|z|^3), \ \text{where } r_i := \frac{p_i^2 + q_i^2}{2} \tag{2.8}$$

Then, there exists $0 < \tilde{\epsilon} \leq \epsilon_0$ and a real-analytic and symplectic transformation

$$\phi : \tilde{z} = (\tilde{p}, \tilde{q}) \in B^{2m}_{\tilde{\epsilon}} \to \tilde{z} + \hat{z}(\tilde{z}) \in B^{2m}_{\epsilon_0} \tag{2.9}$$

which puts H into Birkhoff normal form up to order s, i.e.,

$$\tilde{H} := H \circ \phi = \sum_{i=1}^{m} \Omega_i \tilde{r}_i + \sum_{j=2}^{[s/2]} Q_j(\tilde{r}) + O(|\tilde{z}|^{s+1}) \tag{2.10}$$

where, for $2 \leq j \leq [s/2]$, the Q_j's are homogeneous polynomials of degree j in $\tilde{r} = (\tilde{r}_1, \cdots, \tilde{r}_m)$ with $\tilde{r}_j := \frac{\tilde{p}_j^2 + \tilde{q}_j^2}{2}$. The polynomials Q_j do not depend on ϕ.

2.3 Some Useful Theorems and Formulae

Theorem 2. *(Quantitative Implicit Function Theorem [8]) Let $F = f + g : C^1(D^n_R(0), \mathbb{C}^n)$, where:*

(1) f is a diffeomorphism of $D_R^n(0)$ such that $f(0) = 0$ and Jacobian matrix ∂f non degenerate on $D_R^n(0)$;

(2) $\sup_{D_R(0)} \|\partial g\| \sum_{D_R(0)} \|(\partial f)^{-1}\| \leq \frac{1}{2}$*;*

(3) $\frac{\sup_{D_R(0)} |g| \sum_{D_R(0)} \|(\partial f)^{-1}\|}{r} \leq \frac{1}{2}$*, where $0 < r < R$;*

Then, there exists a unique $z_0 \in B_r^n(0)$ such that $F(z_0) = 0$.

Lemma 1. *Suppose A and B are two $n \times n$ matrices, and A is invertible, and $\|A^{-1}\|\|B\| \leq \frac{1}{2}$ then $A + B$ is invertible and*

$$\|(A+B)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|B\|} \tag{2.11}$$

Proof. Since $\|A^{-1}\|\|B\| \leq \frac{1}{2}$, with quantitative implicit function theorem we obtain that $I + A^{-1}B$ is invertible. let $Q := A^{-1}B$, and

$$\tilde{Q}_n = \sum_{k=0}^{n} (-1)^k Q^k$$

we show that $\tilde{Q}_n$ converges as $n \to \infty$. Indeed, we have the following norm estimate

$$\begin{aligned}\|\tilde{Q}_n\| &= \|\sum_{k=0}^{n} (-1)^k Q^k\| \\ &\leq \sum_{k=1}^{n} \|Q^k\| \leq \sum_{k=1}^{n} \|Q\|^k\end{aligned}$$

since $\|Q\| \leq \frac{1}{2}$, we have $\|\tilde{Q}_n\|$ converges as $n \to \infty$. And let $n, m \to \infty$ and $n \geq m$, we obtain

$$\begin{aligned}\|\tilde{Q}_n - \tilde{Q}_m\| &= \|\sum_{k=m+1}^{n} (-1)^k Q^k\| \leq \sum_{k=m+1}^{n} \|Q\|^k \\ &\leq \frac{\|Q\|^{m+1}}{1 - \|Q\|}\end{aligned}$$

which converges to zero as $n, m \to \infty$. The above computation indicates $\tilde{Q}_n$ converges (Cauchy sequence), and we denote $\tilde{Q} := \lim_{n\to\infty} \tilde{Q}_n$.

Now we prove that $\tilde{Q} = (I+Q)^{-1}$. Indeed we can compute

$$\begin{aligned}\|\tilde{Q}_n(I+Q) - I\| &= \|\sum_{k=0}^{n}(-1)^k Q^k + \sum_{k=0}^{n}(-1)^k Q^{k+1} - I\| \\ &= \|I + (-1)^n Q^{n+1} - I\| \\ &= \|(-1)^n Q^{n+1}\|\end{aligned}$$

which converges to zero, hence we have

$$\begin{aligned}\lim_{n\to\infty} \tilde{Q}_n(I+Q) &= I \\ \Longrightarrow \tilde{Q}(I+Q) &= I\end{aligned}$$

Therefore we have

$$\|(I+Q)^{-1}\| = \|\tilde{Q}\| = \|\sum_{k=0}^{\infty}(-1)^k Q^k\| \leq \frac{1}{1-\|Q\|}$$

which gives us

$$\begin{aligned}\|(A+B)^{-1}\| &= \|(I+A^{-1}B)^{-1}A^{-1}\| \leq \|A^{-1}\|\|(I+A^{-1}B)^{-1}\| \\ &\leq \frac{\|A^{-1}\|}{1-\|A^{-1}B\|} \\ &\leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|B\|}\end{aligned}$$

□

Lemma 2 ([3])**.** *Let $n_1, n_2 \in \mathbb{N}, \tau > n := n_1 + n_2, \gamma_1, \gamma_2 > 0, 0 < \hat{r} < 1, \bar{D}$ be a compact set. Let*

$$\omega = (\omega_1, \omega_2) : \bar{D} \times \bar{B}^{n_2}_{\hat{r}} \to \Omega \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$$

be a function which can be extended to a diffeomorphism on an open neighborhood of $\bar{D} \times \bar{B}^{n_2}_{\hat{r}}$, with ω_2 of the form

$$\omega_2(I_1, I_2) = \omega_{02}(I_1) + \beta(I_1)I_2$$

where $I_1 \to \beta(I_1)$ is a $(n_2 \times n_2)$-matrix, non singular on $\bar{D}$. Let

$$R_1 > \max_{\bar{D}\times\bar{B}^{n_2}_{\hat{r}}} |\omega_1|, \quad a > \max_{\bar{D}} \|\beta\|, \quad c(n,\tau) := \sum_{0\neq k\in\mathbb{Z}^n} \frac{1}{|k|^\tau}$$

and denote

$$\mathcal{R}_{\gamma_1,\gamma_2,\tau} := \left\{ I = (I_1, I_2) \in \bar{D} \times \bar{B}^{n_2}_{\hat{r}} : \omega(I) \notin \mathcal{D}_{\gamma_1,\gamma_2,\tau} \right\}$$

Then,

$$meas\,(\mathcal{R}_{\gamma_1,\gamma_2,\tau}) \leq \left(c_1\gamma_1 + c_2\frac{\gamma_2}{\hat{r}} \right) meas(\bar{D} \times B^{n_2}_{\hat{r}})$$

where

$$\begin{cases} c_1 & := \max_{\bar{D}\times\bar{B}^{n_2}_{\hat{r}}} \|(\partial\omega)^{-1}\|^n \frac{R_1^{n_1-1}}{meas\bar{D}} a^{n_2} c(n,\tau) p \\ c_2 & := \max_{\bar{D}} \|\beta^{-1}\|^{n_2} a^{n_2-1} c(n_2,\tau) \end{cases} \tag{2.12}$$

for a suitable integer p *depending on* $\bar{D}$ *and* ω_1.

Chapter 3

Proof of Theorem 1

The proof of Theorem 1 closely follows the proof of the Theorem 1.4 in [3, 8]. As in [3, 8], we divide it in 6 steps. Steps 1–5 are precisely as in [3, 8], apart for taking into account (in an almost obvious way) the number s of item **(b)** in the Introduction, which in [3, 8] was absent. More careful estimates appear in Step 6, which deals with the measure of the invariant set.

As in [3, 8], we assume that H has an analytic extension to a domain $\mathcal{P}_{\rho_0,\epsilon_0,s_0} := V_{\rho_0} \times \mathbb{T}^{n_1}_{s_0} \times B^{2n_2}_{\epsilon_0}$ with $s_0 < 1$, and the frequency map $\omega_0 := \partial_I H_0$ is a diffeomorphism of V_{ρ_0}.

Step 1 Averaging over the fast angles φ's

Let $0 < \epsilon < e^{-1}$. First we want to use averaging theory, with $l_1 = n_1, l_2 = 0, m = n_2, h = H_0, f = \mu f, B = B' = \{0\}, r_p = r_q = \epsilon_0, s = s_0, \bar{s} = 0, \Lambda = \{0\}$, and $\bar{K}$ such that

$$e^{-\bar{K}s_0/6} := \epsilon^{2s+1}, \quad i.e., \quad \bar{K} = \frac{6}{s_0} \log\left(\epsilon^{-(2s+1)}\right) \tag{3.1}$$

Let $\tau > n_1$, $\tilde{M} := \max_{i,j} \sup_{V_{\rho_0}} |\partial^2_{ij} H_0(I)|$, and $\bar{\gamma}$ be a suitable constant which will be given later. Then, let

$$\bar{D} := \omega_0^{-1}(\mathcal{D}_{\bar{\gamma},\tau}) \cap V \tag{3.2}$$

where $\mathcal{D}_{\bar{\gamma},\tau}$ is the $(\bar{\gamma}, \tau)$-Diophantine set. And we have the following measure estimate

$$\text{meas}(V \backslash \bar{D}) \leq C\bar{\gamma}\text{meas}(V) \tag{3.3}$$

To apply averaging theorem, we need to check condition (2.1). Let $I \in \bar{D}_{\bar{\rho}}$, we obtain

$$|\omega_0(I)\cdot k| \geq \frac{\bar{\gamma}}{\bar{K}^\tau} - \bar{\rho}\bar{K}\tilde{M}$$

and let the right-hand side be larger than or equal to $\frac{\bar{\gamma}}{2\bar{K}^\tau} = \alpha_1 = \alpha_2$, we get

$$\bar{\rho} \leq \frac{\bar{\gamma}}{2\tilde{M}\bar{K}^{\tau+1}}$$

hence we can choose $\bar{\rho}$ as follows (C is a suitable constant):

$$\bar{\rho} := \min\{\frac{\bar{\gamma}}{2\tilde{M}\bar{K}^{\tau+1}}, \rho_0\} \sim \frac{\bar{\gamma}}{C2\tilde{M}\bar{K}^{\tau+1}} \tag{3.4}$$

Now we also need to check condition (2.2), and assume that $\bar{\rho}s_0$ is smaller than ϵ_0^2 (this can accomplished by choosing a very large C), we have $d = \bar{\rho}s_0$, hence

$$\begin{aligned}\mu||f||\frac{2^7 c_{n_2}\bar{K}s_0}{\alpha_2 d} &= \mu||f||\frac{2^8 c_{n_2}\bar{K}^{\tau+1}}{\bar{\gamma}\bar{\rho}}\\ &= \mu||f||\frac{C2^9 c_{n_2}\tilde{M}\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2}\end{aligned}$$

Since from the condition (2.2) we want this to be smaller than 1, $\bar{\gamma}$ need to satisfy the following:

$$\bar{\gamma} > \sqrt{C\mu||f||2^9 c_{n_2}\tilde{M}\bar{K}^{2(\tau+1)}} \tag{3.5}$$

indeed, we can let $\gamma_* >> 2^5(6/s_0)^{\tau+1}\sqrt{C\|f\|c_{n_2}\tilde{M}}$ be a very large constant, and make sure that $\bar{\gamma} \geq \gamma_*\sqrt{\mu}(\log(\epsilon^{-2s-1}))^{\tau+1}$, in the end we also need $\bar{\gamma} \geq \gamma_*\epsilon(\log(\epsilon^{-2s-1}))^{\tau+1}$.

Now conditions (2.1) and (2.2) are verified, and $\bar{K}s_0 \geq 6$ is trivial. Let $A = \bar{D}, r = \bar{\rho}$, then according to the averaging theorem, we have a real-analytic symplectomorphism

$$\phi_1 : (I_1, \varphi_1, p_1, q_1) \in W_{\bar{\rho}/2,\epsilon_0/2,s_0/6} \to (I, \varphi, p, q) \in W_{\bar{\rho},\epsilon_0,s_0} \tag{3.6}$$

where $W_{\bar{\rho},\epsilon_0,s_0} := \bar{D}_{\bar{\rho}} \times \mathbb{T}^{n_1}_{s_0} \times B^{2n_2}_{\epsilon_0}$. And H is transformed into

$$\begin{aligned}H_1 &:= H \circ \phi_1\\ &= H_0(I_1) + \mu g_1(I_1, p_1, q_1) + \mu f_1(I_1, \varphi_1, p_1, q_1)\end{aligned} \tag{3.7}$$

and, according to (2.5), we have

$$\begin{aligned}
\|\mu g_1 - \mu f_{av}\|_{\bar{\rho}/2,\epsilon_0/2,s_0/6} &\le \frac{12}{11}\frac{2^7 c_{n_2}(\mu\|f\|)^2}{\alpha_2 d} = \frac{12}{11}\frac{2^7 c_{n_2}(\mu\|f\|)^2}{\frac{\bar{\gamma}}{2K^\tau}\bar{\rho}s_0} \\
&\le \frac{12}{11}\frac{2^7 c_{n_2}(\mu\|f\|)^2}{\frac{\bar{\gamma}}{2K^\tau}s_0\frac{\bar{\gamma}}{2C\tilde{M}\bar{K}^{\tau+1}}} \\
&\le \frac{12}{11}\frac{2^9 C c_{n_2}(\mu\|f\|)^2\tilde{M}\bar{K}^{2\tau+1}}{\bar{\gamma}^2 s_0} \\
&\le \frac{12}{11}\frac{2^9 C c_{n_2}(\mu\|f\|)^2\tilde{M}\bar{K}^{2\tau+1}}{C\mu\|f\|2^9 c_{n_2}\tilde{M}\bar{K}^{2(\tau+1)}s_0} \\
&\le \frac{12}{11}\frac{\mu\|f\|}{\bar{K}s_0} \\
&\le \frac{2}{11}\frac{\mu\|f\|}{\log(\epsilon^{-2s-1})}
\end{aligned}$$

note $\mu f_{av} = \Pi_\Lambda T_{\bar{K}} f$ since $\Lambda = \{0\}$. Hence

$$\|g_1 - f_{av}\|_{\bar{\rho}/2,\epsilon_0/2,s_0/6} \sim O(\frac{1}{\log(\epsilon^{-2s-1})})$$

And by equation (2.6), we have $\alpha_1 s_0 |I_1 - I| \le 9\mu\|f\|$, i.e.

$$\begin{aligned}
|I_1 - I| &\le \frac{9\mu\|f\|}{\frac{\bar{\gamma}}{2\bar{K}^\tau}s_0} = \frac{18\mu\|f\|\bar{K}^\tau}{\bar{\gamma}s_0} \\
&\le \frac{18\mu\|f\|\bar{K}^\tau}{\sqrt{C\mu\|f\|2^9 c_{n_2}\tilde{M}}\bar{K}^{\tau+1}s_0} \\
&\le \frac{\sqrt{\mu\|f\|}}{\sqrt{Cc_{n_2}\tilde{M}}\log(\epsilon^{-2s-1})} \\
&\sim O(\frac{\sqrt{\mu}}{\log(\epsilon^{-2s-1})})
\end{aligned}$$

Similarly, we have that

$$|p_1 - p|, |q_1 - q| \sim O(\frac{\sqrt{\mu}}{\log(\epsilon^{-2s-1})})$$

and

$$
\begin{aligned}
|\varphi_1 - \varphi| &\leq \frac{9\mu\|f\|}{\alpha_2\bar{\rho}} \leq \frac{9\mu\|f\|}{\frac{\bar{\gamma}}{2\bar{K}^{\tau}}\frac{\bar{\gamma}}{C2\tilde{M}\bar{K}^{\tau+1}}} \\
&\leq \frac{36\mu\|f\|C\tilde{M}\bar{K}^{2\tau+1}}{\bar{\gamma}^2} \\
&\sim O(\frac{1}{\log(\epsilon^{-2s-1})})
\end{aligned}
$$

And by (2.5), we have

$$
\begin{aligned}
\|\mu f_1\|_{\bar{\rho}/2,\epsilon_0/2,s_0/6} &\leq e^{\bar{K}s_0/6}\frac{2^9 c_{n_2}(\mu\|f\|)^2}{\alpha_2 d} \\
&\leq e^{\bar{K}s_0/6}\frac{2^9 c_{n_2}(\mu\|f\|)^2}{\frac{\bar{\gamma}}{2\bar{K}^{\tau+1}}\frac{\bar{\gamma}}{C2\bar{M}\bar{K}^{\tau+1}}6} \\
&\leq e^{\bar{K}s_0/6}\frac{2^9 c_{n_2}(\mu\|f\|)^2}{\frac{C\mu\|f\|2^9 c_{n_2}\bar{M}\bar{K}^{2(\tau+1)}}{C4\bar{M}\bar{K}^{2(\tau+1)}}6} \\
&= e^{\bar{K}s_0/6}\frac{2}{3}\mu\|f\| \\
&\leq \epsilon^{2s+1}\mu\|f\| \\
&\sim O(\mu\epsilon^{2s+1})
\end{aligned}
$$

Step 2 Determination of the elliptic equilibrium for the secular system

Since we assume that f_{av} has a non-degenerate elliptic equilibrium point at the origin and g_1 is close to f_{av} (of order $O(\frac{1}{\log(\epsilon^{-2s-1})})$) , we have that, for any $I_1 \in \bar{D}_{\bar{\rho}/4}$, g_1 also has a elliptic equilibrium point close to the origin of order $O(\frac{1}{\log(\epsilon^{-2s-1})})$, which we call $(p_e(I_1), q_e(I_1))$, and assume that $|(p_e(I_1), q_e(I_1))| < \epsilon_0/4$. Now let $0 < \epsilon < \epsilon_0/4$, (note ϵ is defined before) we have a transformation

$$
\phi_2 : (I_2, \varphi_2, p_2, q_2) \in W_{\bar{\rho}/4,\epsilon,s_0/12} \to (I_1, \varphi_1, p_1, q_1) \in W_{\bar{\rho}/2,\epsilon_0/2,s_0/6} \tag{3.8}
$$

whose generating function is

$$
S(I_2, p_2, \varphi_1, q_1) = I_2 \cdot \varphi_1 + (p_2 + p_e(I_2)) \cdot (q_1 - q_e(I_2))
$$

and that is equivalent to

$$
\begin{aligned}
I_1 &= I_2 \\
\varphi_1 &= \varphi_2 - \partial_{I_2}(p_2 + p_e(I_2)) \cdot (q_1 - q_e(I_2)) \\
p_1 &= p_2 + p_e(I_2) \\
q_1 &= q_2 + q_e(I_2)
\end{aligned}
$$

from which we obtain

$$|p_2 - p_1|, |q_2 - q_1| \leq C_1 \frac{1}{\log(\epsilon^{-2s-1})}$$

where C_1 is a suitable constant, and by Cauchty estimate

$$\begin{aligned}|\varphi_2 - \varphi_1| &= |\partial_{I_2}(p_2 + p_e(I_2)) \cdot (q_1 - q_e(I_2))| \\ &\leq \frac{|(p_2 + p_e(I_2)) \cdot (q_1 - q_e(I_2))|}{\bar{\rho}/4} \\ &\leq \frac{|(p_2 + p_e(I_2)) \cdot q_2|}{\bar{\rho}/4} \\ &\leq 4\frac{(\epsilon + C_1/\log(\epsilon^{-2s-1})) \cdot \epsilon}{\frac{\bar{\gamma}}{C2\tilde{M}\bar{K}^{\tau+1}}} \\ &\leq 4\frac{C2\tilde{M}\bar{K}^{\tau+1}(\epsilon + C_1/\log(\epsilon^{-2s-1})) \cdot \epsilon}{\gamma_* \epsilon (\log \epsilon^{-2s-1})^{\tau+1}} \\ &= \frac{8C\tilde{M}(6/s_0)^{\tau+1}}{\gamma_*}(\epsilon + \frac{C_1}{\log(\epsilon^{-2s-1})}) \\ &\sim O(\epsilon + \frac{C_1}{\log(\epsilon^{-2s-1})})\end{aligned}$$

this term can be very small due to the propriate choice of γ_*.

The Hamiltonian will be transformed into

$$\begin{aligned}H_2 &= H_1 \circ \phi_2 \\ &= H_0(I_2) + \mu g_2(I_2, p_2, q_2) + \mu\epsilon^{2s+1} f_2(I_2, \varphi_2, p_2, q_2)\end{aligned} \tag{3.9}$$

where $g_2 = g_1 \circ \phi_2, \mu\epsilon^{2s+1} f_2 = \mu f_1 \circ \phi_2$. And note that in such case g_2 has a $2s$-non resonant and non-degenerate elliptic equilibrium point in the origin of (p_2, q_2)-coordinate.

Step 3 Symplectic diagonalization of the secular system

Following the [3], we have a symplectic map

$$\phi_3 : (I_3, \varphi_3, p_3, q_3) \in W_{\bar{\rho}/8, \epsilon/2, s_0/24} \to (I_2, \varphi_2, p_2, q_2) \in W_{\bar{\rho}/4, \epsilon, s_0/12} \tag{3.10}$$

which acts as the identity on the I_3-variables, and is close to the identity on other variables in the sense

$$|p_2 - p_3|, |q_2 - q_3| \leq C_2 \frac{\epsilon}{\log(\epsilon^{-2s-1})}$$

$$|\varphi_2 - \varphi_3| \leq C_3(\epsilon^2 + \frac{C_1\epsilon}{\log(\epsilon^{-2s-1})})$$

And we obtain that

$$
\begin{aligned}
H_3 &= H_2 \circ \phi_3 \\
&= H_0(I_3) + \mu g_3(I_3, p_3, q_3) + \mu\epsilon^{2s+1} f_3(I_3, \varphi_3, p_3, q_3)
\end{aligned}
\tag{3.11}
$$

where $f_3 := f_2 \circ \phi_3$, and

$$
g_3(I_3, p_3, q_3) := g_2 \circ \phi_3(I_3, p_3, q_3) = \hat{f}_0(I_3) + \hat{\Omega}(I_3) \cdot r_3 + \hat{R}
$$

where $r_3 = (r_{31}, \cdots, r_{3n_2}), r_{3i} = \frac{p_{3i}^2 + q_{3i}^2}{2}, i = 1, 2, \cdots, n_2$ and

$$
|\hat{\Omega} - \Omega|, |\hat{R}| < C_4 \frac{\epsilon}{\log(\epsilon^{-2s-1})}
$$

where $\hat{R}$ has a zero of order 3 for $(p_3, q_3) = 0$.

Step 4 Birkhoff normal form of the secular part

Given that $\hat{\Omega}$ is non resonant up to order $(2s)$, Proposition2 shows that there exists a Birkhoff transformation

$$
\phi_4 : (I_4, \varphi_4, p_4, q_4) \in W_{\bar{\rho}/16, \epsilon/4, s_0/48} \to (I_3, \varphi_3, p_3, q_3) \in W_{\bar{\rho}/8, \epsilon/2, s_0/24} \tag{3.12}
$$

which puts g_3 into Birkhoff normal form up to order $2s$, i.e., the Hamiltonian is transformed into the form

$$
\begin{aligned}
H_4 &= H_3 \circ \phi_4 \\
&= H_0(I_4) \quad + \mu\Big(\hat{f}_0(I_4) + \hat{\Omega}(I_4) \cdot r_4 + \frac{1}{2} r_4 \cdot \hat{B}(I_4) r_4 + \cdots \\
&\qquad\qquad + \sum_{i_1, \cdots, i_s \in \{1, \cdots, n_2\}} \hat{b}_i(I_4) r_{4i_1} r_{4i_2} \cdots r_{4i_s} + O(|(p_4, q_4)|^{2s+1})\Big) + \\
&\qquad\qquad + \mu\epsilon^{2s+1} f_3 \circ \phi_4(I_4, \varphi_4, p_4, q_4) \\
&=: H_0(I_4) \quad + \mu g_4(I_4, p_4, q_4) + \mu\epsilon^{2s+1} f_4(I_4, \varphi_4, p_4, q_4)
\end{aligned}
\tag{3.13}
$$

where $r_{4i} = \frac{p_{4i}^2 + q_{4i}^2}{2}, i = 1, 2, \cdots, n_2$ and $\mu\epsilon^{2s+1} f_4 = \mu O(|(p_4, q_4)|^{2s+1}) + \mu\epsilon^{2s+1} f_3 \circ \phi_4$. And we have that

$$
|p_3 - p_4|, |q_3 - q_4| \leq C(n_2) \frac{\epsilon^2}{\log(\epsilon^{-2s-1})}
$$

$$
|\varphi_3 - \varphi_4| \leq C(n_2) \frac{\epsilon^3}{\sqrt{\mu} \log(\epsilon^{-2s-1})}
$$

Step 5 Global action-angle variables for the full system

Fix the real n_2-dimensional annulus

$$\mathcal{A}(\epsilon) := \left\{ J \in \mathbb{R}^{n_2} : c_1 \epsilon^{(2s+1)/2} < J_i < c_2 \epsilon^2, \quad 1 < i < n_2 \right\} \tag{3.14}$$

and let

$$\mathcal{D} := \bar{D} \times \mathcal{A}(\epsilon), \quad \breve{\rho} := \min\{c_1 \epsilon^{(2s+1)/2}, \bar{\rho}/16\}, \quad \breve{s} := s_0/48 \tag{3.15}$$

where $\bar{D}$ is defined in equation (3.2). On $\mathcal{D}_{\breve{\rho}} \times \mathbb{T}^n_{\breve{s}}$, we define a transformation

$$\phi_5 : (J, \psi) = ((J_1, J_2), (\psi_1, \psi_2)) \to (I_4, \varphi_4, p_4, q_4) \tag{3.16}$$

with

$$\begin{aligned} I_4 &= J_1, \\ \varphi_4 &= \psi_1, \\ p_{4i} &= \sqrt{2J_{2i}} \cos \psi_{2i}, \\ q_{4i} &= \sqrt{2J_{2i}} \sin \psi_{2i} \end{aligned}$$

where $i = 1, \cdots, n_2$. The transformation ϕ_5 transform H_4 into the form

$$\begin{aligned} H_5(J, \psi) &:= H_4 \circ \phi_5 \\ &= H_0(J_1) \quad + \mu g_5(J_1, J_2) + \mu \epsilon^{2s+1} f_5(J, \psi) \\ &= H_0(J_1) \quad + \mu \Big(\hat{f}_0(J_1) + \hat{\Omega}(J_1) \cdot J_2 + \frac{1}{2} J_2 \cdot \hat{B}(J_1) J_2 + \cdots \\ &\qquad + \sum_{i_1, \cdots, i_s \in \{1, \cdots, n_2\}} \hat{b}_i(J_1) J_{2i_1} J_{2i_2} \cdots J_{2i_s} \Big) \\ &\qquad + \mu \epsilon^{2s+1} f_4 \circ \phi_5 \end{aligned} \tag{3.17}$$

From the above steps (3.6,3.8,3.10,3.12,3.16), we have the following transformation

$$\phi := \phi_1 \circ \phi_2 \circ \phi_3 \circ \phi_4 \circ \phi_5 : (J, \psi) = (J_1, J_2, \psi_1, \psi_2) \to (I, \varphi, p, q) \tag{3.18}$$

which is well defined, and the following inequalities hold

$$\begin{aligned} |I - J_1| &= |I - I_1 + I_1 - I_2 + I_2 - I_3 + I_3 - I_4 + I_4 - J_1| \\ &= |I - I_1| \le C_I \frac{\sqrt{\mu}}{\log(\epsilon^{-2s-1})} \\ |\varphi - \psi_1| &= |\varphi - \varphi_1 + \varphi_1 - \varphi_2 + \varphi_2 - \varphi_3 + \varphi_3 - \varphi_4 + \varphi_4 - \psi_1| \\ &\le |\varphi - \varphi_1| + |\varphi_1 - \varphi_2| + |\varphi_2 - \varphi_3| + |\varphi_3 - \varphi_4| + |\varphi_4 - \psi_1| \\ &\le \frac{1}{\log(\epsilon^{-2s-1})} + (\epsilon + \frac{C_1}{\log(\epsilon^{-2s-1})}) + (\epsilon^2 + \frac{C_1 \epsilon}{\log(\epsilon^{-2s-1})}) + C(n_2) \frac{\epsilon^3}{\sqrt{\mu} \log(\epsilon^{-2s-1})} \\ &\le C_\varphi(\epsilon + \frac{1}{\log(\epsilon^{-2s-1})}) \end{aligned}$$

and

$$|p_i - p_{e,i} - \sqrt{2J_{2i}}\cos\psi_{2i}| = |p_i - p_{1i} + p_{1i} - p_{2i} + p_{2i} - p_{3i} + p_{3i} - p_{4i} + p_{4i} - \sqrt{2J_{2i}}\cos\psi_{2i} - p_{e,i}|$$
$$\le \frac{\sqrt{\mu}}{\log(\epsilon^{-2s-1})} + \frac{\epsilon}{\log(\epsilon^{-2s-1})} + C(n_2)\frac{\epsilon^2}{\log(\epsilon^{-2s-1})}$$
$$\le C_p\frac{\max\{\sqrt{\mu},\epsilon\}}{\log(\epsilon^{-2s-1})}$$
$$|q_i - q_{e,i} - \sqrt{2J_{2i}}\sin\psi_{2i}| = |q_i - q_{1i} + q_{1i} - q_{2i} + q_{2i} - q_{3i} + q_{3i} - q_{4i} + q_{4i} - \sqrt{2J_{2i}}\sin\psi_{2i} - q_{e,i}|$$
$$\le \frac{\sqrt{\mu}}{\log(\epsilon^{-2s-1})} + \frac{\epsilon}{\log(\epsilon^{-2s-1})} + C(n_2)\frac{\epsilon^2}{\log(\epsilon^{-2s-1})}$$
$$\le C_q\frac{\max\{\sqrt{\mu},\epsilon\}}{\log(\epsilon^{-2s-1})}$$

Step 6 Construction of the Kolmogorov set and estimate of its measure

Fix γ_1 and $\gamma_2 = \mu\hat{\gamma_2}$. We apply the two-scale KAM theorem (following [3]), with

$$H = H_5, \quad h = H_0 + \mu g_5, \quad f = \mu\epsilon^{2s+1} f_5, \quad D = \mathcal{D}, \quad \rho = \breve{\rho}$$

and $s = \breve{s}/5, \bar{s} = 4\breve{s}/5$.

We need to check that the frequency map $\omega_\mu : \partial(H_0 + \mu g_5)$ is a diffeomorphism of $\mathcal{D}_{\breve{\rho}}$ and its Hessian matrix $\partial^2(H_0 + \mu g_5)$ is non-singular.

Claim: The frequency map $\omega_\mu := \partial(H_0 + \mu g_5)$ is a diffeomorphism of $\mathcal{D}_{\breve{\rho}}$.

Proof: Assume that ∂_1 represents taking derivative with respect to J_1, and ∂_2 w.r.t. J_2. We have

$$\omega_\mu = \partial(H_0 + \mu g_5)$$
$$= \begin{pmatrix} \partial_1 H_0 + \mu\partial_1\hat{f}_0 + \mu\partial_1\hat{\Omega}\cdot J_2 + \frac{\mu}{2}J_2\cdot(\partial_1\hat{B})J_2 + \cdots + \mu\sum_i(\partial_1\hat{b}_i)J_{2i_1}J_{2i_2}\cdots J_{2i_s} \\ \mu\hat{\Omega} + \mu\hat{B}J_2 + \mu\sum_{ijk}\hat{b}_{ijk}\partial_2(J_{2i}J_{2j}J_{2k}) + \cdots + \mu\sum_i\hat{b}_i\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s}) \end{pmatrix}$$

Since the analyticity of the frequency map is assumed, we only need to prove ist injectivity. That is, we want to prove, for any frequency $\nu = (\nu_1; \mu\nu_2)$, the following equation has only one solution on $(J_1, J_2) \in \mathcal{D}_{\breve{\rho}}$

$$\partial_1 H_0 + \mu\partial_1\hat{f}_0 + \mu\partial_1\hat{\Omega}\cdot J_2 + \frac{\mu}{2}J_2\cdot(\partial_1\hat{B})J_2 + \cdots + \mu\sum_i(\partial_1\hat{b}_i)J_{2i_1}J_{2i_2}\cdots J_{2i_s} = \nu_1$$
$$\mu\hat{\Omega} + \mu\hat{B}J_2 + \mu\sum_{ijk}\hat{b}_{ijk}\partial_2(J_{2i}J_{2j}J_{2k}) + \cdots + \mu\sum_i\hat{b}_i\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s}) = \mu\nu_2$$

First we consider the second equation: for any fixed $J_1 \in \bar{D}$, we have

$$\hat{\Omega} + \hat{B}J_2 + \sum_{ijk} \hat{b}_{ijk}\partial_2(J_{2i}J_{2j}J_{2k}) + \cdots + \sum_i \hat{b}_i\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s}) = \nu_2$$

Since $\hat{B}$ is non-singular (according to our assumption), we find that $J_2 = \hat{B}^{-1}(\nu_2 - \hat{\Omega})$ solves the following equation (the major part of the second equation)

$$\hat{\Omega} + \hat{B}J_2 = \nu_2$$

Given that the remaining part is of order $O(\epsilon^4)$, we could assume that the expected solution is of the form

$$J_2 = \hat{B}^{-1}(\nu_2 - \hat{\Omega}) + \delta$$

with δ is of order $O(\epsilon^2)$, substituting it into the second equation we obtain

$$\hat{\Omega} + \hat{B}(\hat{B}^{-1}(\nu_2 - \hat{\Omega}) + \delta) + \sum_{ijk} \hat{b}_{ijk}\partial_2(J_{2i}J_{2j}J_{2k})|_{J_2} + \cdots + \sum_i \hat{b}_i\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s})|_{J_2} = \nu_2$$

$$\Rightarrow \hat{B}\delta + \sum_{ijk} \hat{b}_{ijk}\partial_2(J_{2i}J_{2j}J_{2k})|_{J_2} + \cdots + \sum_i \hat{b}_i\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s})|_{J_2} = 0$$

which is denoted by $f(\delta) + g(\delta) = 0$, where $f = \hat{B}\delta$ and g is the remaining terms.

Note $\partial_\delta f = \hat{B}$ is non-degenerate and $\|(\partial_\delta f)^{-1}\| = \|\hat{B}^{-1}\|$ is bounded by a suitable constant N, i.e. $\sup_{\bar{D}}\|\hat{B}^{-1}\| \leq N$. And by Cauchy estimates,

$$\begin{aligned}
\sup_{\mathcal{A}(\epsilon)_{\bar{\rho}/16}} \|\partial_\delta g\| &\leq C\frac{\|g\|}{\bar{\rho}} \leq C\frac{\epsilon^4}{\bar{\rho}} \\
&\leq C\frac{\epsilon^4}{\frac{\bar{\gamma}}{2\tilde{M}\bar{K}^{\tau+1}}} \\
&\leq C\frac{\epsilon^4}{\sqrt{\mu}} \leq \frac{1}{2N}
\end{aligned}$$

provided that μ and ϵ satisfy some conditions ($\mu > C\epsilon^8$). Hence by quantitative implicit function theorem there exists a unique δ_0 in $\mathcal{A}(\epsilon)_{\bar{\rho}/16}$ such that $f(\delta_0) + g(\delta_0) = 0$, as a consequence of which, $J_2 = \hat{B}^{-1}(\nu_2 - \hat{\Omega}) + \delta_0$ is unique.

Second we consider the first equations

$$\partial_1 H_0 + \mu\partial_1\hat{f}_0 + \mu\partial_1\hat{\Omega}\cdot J_2 + \frac{\mu}{2}J_2\cdot(\partial_1\hat{B})J_2 + \cdots + \mu\sum_i(\partial_1\hat{b}_i)J_{2i_1}J_{2i_2}\cdots J_{2i_s} = \nu_1$$

Note $\omega_0 := \partial_1 H_0$ is well defined in $\mathcal{D}_{\breve{\rho}}$, and we assume that $\|(\partial_1\omega_0)^{-1}\|$ is bounded by M_0 in $\mathcal{D}_{\breve{\rho}}$. Let the remaining terms be denoted by ω_1, by Cauchy estimates,

$$\begin{aligned}
\sup_{\mathcal{D}_{\bar{\rho}/16}} \|\partial_1\omega_1\| &\leq C\frac{\|\omega_1\|}{\bar{\rho}} \leq C\frac{\mu}{\bar{\rho}^2} \\
&\leq C\frac{\mu}{\frac{\bar{\gamma}^2}{4\tilde{M}^2\bar{K}^{2(\tau+1)}}} \\
&\leq C\mu\frac{4\tilde{M}^2\bar{K}^{2(\tau+1)}}{\gamma_*^2\epsilon^2(\log\epsilon^{-2s-1})^{2(\tau+1)}} \\
&\leq C\frac{\mu}{\epsilon^2} \\
&< \frac{1}{2M_0}
\end{aligned}$$

provided that μ and ϵ satisfy some conditions ($\mu < C\epsilon^2$). Similarly by the quantitative implicit function theorem, we have J_1 is also uniquely determined. ■

Claim: The frequency map ω_μ has a non-singular Jacobian matrix, i.e. $\partial\omega_\mu$, on $\mathcal{D}_{\breve{\rho}}$.

Proof: Denote $W := \partial\omega_\mu$, then we have

$$\begin{aligned}
W_{11} &= \partial_1^2 H_0 + \mu\partial_1^2\hat{f}_0 + \mu\partial_1^2\hat{\Omega}\cdot J_2 + \frac{\mu}{2}J_2\cdot(\partial_1^2\hat{B})J_2 + \cdots + \mu\sum_i(\partial_1^2\hat{b}_i)J_{2i_1}J_{2i_2}\cdots J_{2i_s} \\
W_{12} &= \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 + \mu\sum_{ijk}(\partial_1\hat{b}_{ijk})\partial_2(J_{2i}J_{2j}J_{2k}) + \cdots + \mu\sum_i(\partial_1\hat{b}_i)\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s}) \\
W_{21} &= \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 + \mu\sum_{ijk}(\partial_1\hat{b}_{ijk})\partial_2(J_{2i}J_{2j}J_{2k}) + \cdots + \mu\sum_i(\partial_1\hat{b}_i)\partial_2(J_{2i_1}J_{2i_2}\cdots J_{2i_s}) \\
W_{22} &= \mu\hat{B} + \mu\sum_{ijk}\hat{b}_{ijk}\partial_2^2(J_{2i}J_{2j}J_{2k}) + \cdots + \mu\sum_i\hat{b}_i\partial_2^2(J_{2i_1}J_{2i_2}\cdots J_{2i_s})
\end{aligned}$$

note here W_{12} and W_{21} are symmetric. Indeed we can rewrite the Jacobian matrix in the following form

$$W = W^0 + W^1$$

where

$$W^0 := \begin{pmatrix} \partial_1^2 H_0 + \mu\partial_1^2\hat{f}_0 & \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 \\ \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 & \mu\hat{B} \end{pmatrix}$$

and

$$W^1 := \begin{pmatrix} \mu\partial_1^2\hat{\Omega}\cdot J_2 + \frac{\mu}{2}J_2\cdot(\partial_1^2\hat{B})J_2 + O(\mu\epsilon^6) & \mu\sum_{ijk}(\partial_1\hat{b}_{ijk})\partial_2(J_{2i}J_{2j}J_{2k}) + O(\mu\epsilon^6) \\ \mu\sum_{ijk}(\partial_1\hat{b}_{ijk})\partial_2(J_{2i}J_{2j}J_{2k}) + O(\mu\epsilon^6) & \mu\sum_{ijk}\hat{b}_{ijk}\partial_2^2(J_{2i}J_{2j}J_{2k}) + O(\mu\epsilon^6) \end{pmatrix}$$

Due to the smallness of μ and ϵ, W^1 is a very small term. We first consider W^0 and prove its invertibility

$$\begin{aligned} W^0 &= \begin{pmatrix} \partial_1^2 H_0 & \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 \\ \mu\partial_1\hat{\Omega} + \mu(\partial_1\hat{B})J_2 & \mu\hat{B} \end{pmatrix} + \begin{pmatrix} \mu\partial_1^2\hat{f}_0 & 0 \\ 0 & 0 \end{pmatrix} \\ &=: \begin{pmatrix} \partial_1^2 H_0 & \mu\hat{A} \\ \mu\hat{A}^T & \mu\hat{B} \end{pmatrix} + \begin{pmatrix} \mu\hat{C} & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

note the first part of W^0 is invertible and its inverse is

$$\begin{pmatrix} (\partial_1^2 H_0 - \mu\hat{A}(\mu\hat{B})^{-1}\mu\hat{A}^T)^{-1} & -(\partial_1^2 H_0 - \mu\hat{A}(\mu\hat{B})^{-1}\mu\hat{A}^T)^{-1}\mu\hat{A}(\mu\hat{B})^{-1} \\ -(\mu\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\mu\hat{A})^{-1}\mu\hat{A}^T(\partial_1^2 H_0)^{-1} & (\mu\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\mu\hat{A})^{-1} \end{pmatrix}$$

Hence we have

$$\begin{aligned} W^0 &= \begin{pmatrix} \partial_1^2 H_0 & \mu\hat{A} \\ \mu\hat{A}^T & \mu\hat{B} \end{pmatrix} \left(id + \begin{pmatrix} \partial_1^2 H_0 & \mu\hat{A} \\ \mu\hat{A}^T & \mu\hat{B} \end{pmatrix}^{-1} \begin{pmatrix} \mu\hat{C} & 0 \\ 0 & 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} \partial_1^2 H_0 & \mu\hat{A} \\ \mu\hat{A}^T & \mu\hat{B} \end{pmatrix} \left(id + \begin{pmatrix} \mu\hat{C}(\partial_1^2 H_0 - \mu\hat{A}(\mu\hat{B})^{-1}\mu\hat{A}^T)^{-1} & 0 \\ -\mu\hat{C}(\mu\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\mu\hat{A})^{-1}\mu\hat{A}^T(\partial_1^2 H_0)^{-1} & 0 \end{pmatrix} \right) \\ &=: \begin{pmatrix} \partial_1^2 H_0 & \mu\hat{A} \\ \mu\hat{A}^T & \mu\hat{B} \end{pmatrix} (id + \tilde{W^0}) \end{aligned}$$

note in the big parentheses we have an identity matrix added by a small term, $\tilde{W^0}$, of which the norm is bounded by

$$\begin{aligned} \|\tilde{W^0}\| \le &\max\{\mu\|\hat{C}\|\|(\partial_1^2 H_0 - \mu\hat{A}(\mu\hat{B})^{-1}\mu\hat{A}^T)^{-1}\|, \\ &\qquad \mu\|\hat{C}\|\|(\mu\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\mu\hat{A})^{-1}\|\|\mu\hat{A}^T\|\|(\partial_1^2 H_0)^{-1}\|\} \\ = &\max\{\mu\|\hat{C}\|\|(\partial_1^2 H_0 - \mu\hat{A}\hat{B}^{-1}\hat{A}^T)^{-1}\|, \\ &\qquad \mu\|\hat{C}\|\|(\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\hat{A})^{-1}\|\|\hat{A}^T\|\|(\partial_1^2 H_0)^{-1}\|\} \end{aligned}$$

where

$$\mu\|\hat{C}\| = \mu\|\partial_1^2 \hat{f}_0\| \leq \frac{C\mu}{\bar{\rho}^2} \leq \mu \frac{C}{\frac{\bar{\gamma}^2}{4\tilde{M}^2\bar{K}^{2(\tau+1)}}} \leq C\frac{\mu}{\epsilon^2}$$

$$\begin{aligned}
\|(\partial_1^2 H_0 - \mu\hat{A}\hat{B}^{-1}\hat{A}^T)^{-1}\| &= \|(id - \mu(\partial_1^2 H_0)^{-1}\hat{A}\hat{B}^{-1}\hat{A}^T)^{-1}(\partial_1^2 H_0)^{-1}\| \\
&\leq \|(\partial_1^2 H_0)^{-1}\|\|(id - \mu(\partial_1^2 H_0)^{-1}\hat{A}\hat{B}^{-1}\hat{A}^T)^{-1}\| \\
&\leq M_0\|(id - \mu(\partial_1^2 H_0)^{-1}\hat{A}\hat{B}^{-1}\hat{A}^T)^{-1}\| \\
&\leq M_0\sum_{k=0}^{\infty}\|\mu(\partial_1^2 H_0)^{-1}\hat{A}\hat{B}^{-1}\hat{A}^T\|^k \\
&\leq M_0\sum_{k=0}^{\infty}(\mu M_0)^k\|\hat{A}\|^k\|\hat{B}^{-1}\|^k\|\hat{A}^T\|^k \\
&\leq M_0\sum_{k=0}^{\infty}(\mu M_0)^k(\frac{C}{\bar{\rho}^2})^k \\
&\leq M_0\sum_{k=0}^{\infty}\left(\frac{C\mu M_0}{\bar{\rho}^2}\right)^k \leq \frac{M_0}{1-\frac{C\mu M_0}{\bar{\rho}^2}}
\end{aligned}$$

and

$$\|\hat{A}^T\| \leq \frac{C}{\bar{\rho}}$$

$$\begin{aligned}
\|(\hat{B} - \mu\hat{A}^T(\partial_1^2 H_0)^{-1}\hat{A})^{-1}\| &= \|(id - \mu\hat{B}^{-1}\hat{A}^T(\partial_1^2 H_0)^{-1}\hat{A})^{-1}\hat{B}^{-1}\| \\
&\leq \|\hat{B}^{-1}\|\|(id - \mu\hat{B}^{-1}\hat{A}^T(\partial_1^2 H_0)^{-1}\hat{A})^{-1}\| \\
&\leq \frac{C}{1-\frac{C\mu M_0}{\bar{\rho}^2}}
\end{aligned}$$

Hence we have that

$$\|\tilde{W^0}\| \leq \max\{C\frac{\mu}{\epsilon^2}\frac{M_0}{1-\frac{C\mu M_0}{\bar{\rho}^2}}, CM_0\frac{\mu}{\epsilon^2}\frac{C}{\bar{\rho}}\frac{C}{1-\frac{C\mu M_0}{\bar{\rho}^2}}\}$$

and $\|\tilde{W^0}\| < \frac{1}{2}$ is possible. Hence hence the matrix W^0 is invertible, and

$$\|(W^0)^{-1}\| = \frac{C}{\mu}$$

Now we consider the Jacobian matrix W,

$$W = W^0 + W^1 = W^0(id + (W^0)^{-1}W^1)$$

and the matrix $(W^0)^{-1}W^1$ has the norm bounded by

$$\begin{aligned}\|(W^0)^{-1}W^1\| &\le \|(W^0)^{-1}\|\|W^1\| \\ &\le \frac{C}{\mu}C\mu\epsilon^2 \\ &\le \frac{1}{2}\end{aligned}$$

This proves that the Jacobian matrix W is also invertible. ■

With a suitable constant C_+, we have that

$$\|\partial\omega_\mu\| \le M := C_+, \quad \hat{M} := C_+\mu, \quad \bar{M} := C_+\mu^{-1}, \quad E := \mu\epsilon^{2s+1}C_+,$$
$$\bar{M}_1 = C_+, \quad \bar{M}_2 = C_+\mu^{-1}$$

and we also have

$$\begin{aligned}K &= \frac{6}{s}\log\left(8(12)^{2(\tau+1)}\right) \\ \hat{\rho} &= \min\{\frac{\gamma_1}{2MK^{\tau+1}}, \frac{\gamma_2}{2\hat{M}K^{\tau+1}}, \rho\} \\ &= \min\{\frac{\gamma_1}{2C_+K^{\tau+1}}, \frac{\mu\hat{\gamma}_2}{2C_+\mu K^{\tau+1}}, c_1\epsilon^{(2s+1)/2}, \rho_0/16, \frac{\bar{\gamma}}{32\tilde{M}\bar{K}^{\tau+1}}\} \\ &= \min\{\frac{\gamma_1}{2C_+K^{\tau+1}}, \frac{\hat{\gamma}_2}{2C_+K^{\tau+1}}, c_1\epsilon^{(2s+1)/2}, \rho_0/16, \frac{\bar{\gamma}}{32\tilde{M}\bar{K}^{\tau+1}}\}\end{aligned} \tag{3.19}$$

We need to check that the following conditions hold

$$\hat{c}\frac{E}{M\hat{\rho}^2} < 1, \quad \hat{c}\frac{E}{\hat{M}\hat{\rho}^2} < 1, \quad \hat{c}\frac{E\bar{M}}{\hat{\rho}^2} < 1 \tag{3.20}$$

Claim: The smallness conditions hold when $\gamma_1, \hat{\gamma}_2 >> \gamma_*\epsilon^{(2s+1)/2}(\log(8(12)^{2(\tau+1)}))^{\tau+1}$.

Proof: Substituting $\hat{\rho}$ into the smallness conditions, we check the first term of equation (3.20), i.e.

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{C_+}\max\{\frac{4C_+^2K^{2(\tau+1)}}{\gamma_1^2}, \frac{4C_+^2K^{2(\tau+1)}}{\hat{\gamma}_2^2}, \frac{1}{c_1^2\epsilon^{2s+1}}, \frac{16^2}{\rho_0^2}, \frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2}\} < 1$$

since $\bar{\gamma}^2 \geq C\mu\|f\|2^9 c_{n_2}\tilde{M}\bar{K}^{2(\tau+1)}$, clearly we have

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{C_+}\frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2} \leq \hat{c}\frac{\mu\epsilon^{2s+1}C_+}{C_+}\frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{C\mu\|f\|2^9 c_{n_2}\tilde{M}\bar{K}^{2(\tau+1)}}$$
$$\leq \hat{c}\frac{\epsilon^{2s+1}C_+}{C_+}\frac{\tilde{M}}{C\|f\|c_{n_2}} << 1$$

and

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{C_+}\frac{1}{c_1^2\epsilon^{2s+1}} = \hat{c}\frac{\mu C_+}{c_1^2 C_+} << 1$$

we also need

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{C_+}\max\{\frac{4C_+^2K^{2(\tau+1)}}{\gamma_1^2}, \frac{4C_+^2K^{2(\tau+1)}}{\hat{\gamma}_2^2}\} << 1$$

namely

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+4C_+^2(6/s)^{2(\tau+1)}(\log(8(12)^{2(\tau+1)}))^{2(\tau+1)}}{C_+}\max\{\frac{1}{\gamma_1^2}, \frac{1}{\hat{\gamma}_2^2}\} << 1$$

Hence we need that

$$\gamma_1, \hat{\gamma}_2 >> \gamma_*\sqrt{\mu}\epsilon^{(2s+1)/2}(\log(8(12)^{2(\tau+1)}))^{(\tau+1)}$$

Then the second term of equation (3.20) can also be verified

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{\mu C_+}\max\{\frac{4C_+^2K^{2(\tau+1)}}{\gamma_1^2}, \frac{4C_+^2K^{2(\tau+1)}}{\hat{\gamma}_2^2}, \frac{1}{c_1^2\epsilon^{2s+1}}, \frac{16^2}{\rho_0^2}, \frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2}\} < 1$$

with similar computations, we have that

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{\mu C_+}\frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{\bar{\gamma}^2} < \hat{c}\frac{\mu\epsilon^{2s+1}C_+}{\mu C_+}\frac{32^2\tilde{M}^2\bar{K}^{2(\tau+1)}}{\gamma_*^2\epsilon^2(\log(\epsilon^{-2s-1}))^{2(\tau+1)}} << 1$$

and

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{\mu C_+}\frac{1}{c_1^2\epsilon^{2s+1}} < \frac{\hat{c}}{c_1^2} < 1$$

with a suitable c_1. And

$$\hat{c}\frac{\mu\epsilon^{2s+1}C_+}{\mu C_+}\max\{\frac{4C_+^2K^{2(\tau+1)}}{\gamma_1^2}, \frac{4C_+^2K^{2(\tau+1)}}{\hat{\gamma}_2^2}\} << 1$$

gives us

$$\gamma_1, \hat{\gamma}_2 >> \gamma_* \epsilon^{(2s+1)/2} (\log(8(12)^{2(\tau+1)}))^{(\tau+1)}$$

Finally, the third smallness condition is similar with the second one, hence it is obviously satisfied. ■

Hence, according to the two-scale KAM theorem 3, for any $\omega \in \Omega_* := \omega_\mu(\mathcal{D}) \cap \mathcal{D}_{\gamma_1,\gamma_2,\tau}$, one can find a unique real-analytic embedding

$$\phi_\omega : \theta \in \mathbb{T}^n \to ((v_1(\theta;\omega), v_2(\theta;\omega)), \theta + u(\theta;\omega)) \in Re(D_r) \times \mathbb{T}^n$$

with $r = 2\frac{\mu\epsilon^{2s+1}C_+}{\hat{\rho}}(\frac{1}{C_+} + \frac{1}{C_+\mu} + 4nC_+\mu^{-1})$, which can be proved to be very small. Then, the H-flow is analytically conjugated to $\theta \to \theta + \omega t$. We have $T_\omega := \phi_\omega(\mathbb{T}^n)$ and set $\mathcal{T}_\omega := \phi(T_\omega)$, where ϕ is the symplectic transformation defined before.

Now we have the measure of the Kolmogorov set to estimates

$$\mathcal{K} := \phi(K) = \cup_{\omega \in \Omega_*} \mathcal{T}_\omega \tag{3.21}$$

where $K := \cup_{\omega\in\Omega_*} T_\omega$. Let $\mathcal{D}^*_{\gamma_1,\gamma_2,\tau} := \omega_\mu^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap \mathcal{D}$. Then since the map ϕ (see equation (3.18)) is volume preserving, we have, from the two-scale KAM theorem

$$\begin{aligned} \text{meas}\mathcal{K} &= \text{meas}K \\ &\geq \text{meas}(Re(\mathcal{D}_r) \times \mathbb{T}^n) - c_n(\text{meas}(\mathcal{D} \backslash \mathcal{D}^*_{\gamma_1,\gamma_2,\tau} \times \mathbb{T}^n) + \text{meas}((Re(\mathcal{D}_r) \backslash \mathcal{D}) \times \mathbb{T}^n)) \end{aligned}$$

Let $\mathcal{V} := V \times B^{n_2}_{c_2\epsilon^2}$, where $B^{n_2}_{c_2\epsilon^2} := \{|J_i| < c_2\epsilon^2\}$. Indeed, $\bar{D} \subset V$ and $\mathcal{A}(\epsilon) \subset B^{n_2}_{c_2\epsilon^2}$ indicates that $\mathcal{D} \subset \mathcal{V}$. Define $\mathcal{P}_\epsilon := V \times \mathbb{T}^{n_1} \times \{p_i^2 + q_i^2 < \epsilon^2\}$. We have that

$$\begin{aligned} \text{meas}(\mathcal{D}_r \times \mathbb{T}^n) &\geq \text{meas}(\mathcal{D} \times \mathbb{T}^n) \\ &= \text{meas}(\bar{D})\text{meas}(\mathcal{A}(\epsilon))\text{meas}(\mathbb{T}^n) \end{aligned} \tag{3.22}$$

from equation (3.3) we know that

$$\begin{aligned} \text{meas}(V \backslash \bar{D}) &= \text{V} - \text{meas}(\bar{D}) \\ &\leq \bar{C}\bar{\gamma}\text{meas}(V) \end{aligned}$$

which indicates

$$\text{meas}(\bar{D}) \geq (1 - \bar{C}\bar{\gamma})\text{meas}(V) \tag{3.23}$$

and due to the definition of $\mathcal{A}(\epsilon)$, we get

$$\begin{aligned}
\text{meas}(\mathcal{A}(\epsilon)) &= \text{meas}(B^{n_2}_{c_2\epsilon^2}) - \text{meas}(B^{n_2}_{c_1\epsilon^{(2s+1)/2}}) \\
&= \sigma_{n_2}(c_2\epsilon^2)^{n_2} - \sigma_{n_2}(c_1\epsilon^{(2s+1)/2})^{n_2} \\
&= \sigma_{n_2}(c_2\epsilon^2)^{n_2}\left[1 - \left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right] \qquad (3.24)\\
&= \text{meas}(B^{n_2}_{c_2\epsilon^2})\left[1 - \left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]
\end{aligned}$$

Hence from equation (3.22,3.23,3.24) we have

$$\begin{aligned}
\text{meas}(\mathcal{D}_r \times \mathbb{T}^n) &\geq (1-\bar{C}\bar{\gamma})\left[1 - \left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]\text{meas}(V)\text{meas}(B^{n_2}_{c_2\epsilon^2})\text{meas}(\mathbb{T}^n) \\
&\geq (1-\bar{C}\bar{\gamma})\left[1 - \left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]\text{meas}(\mathcal{V})\text{meas}(\mathbb{T}^n) \\
&= (1-\bar{C}\bar{\gamma})\left[1 - \left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}})
\end{aligned} \qquad (3.25)$$

indeed, we have $(J_1, J_2, \psi_1, \psi_2) \in \mathcal{V} \times \mathbb{T}^n = V \times B^{n_2}_{c_2\epsilon^2} \times \mathbb{T}^n$, and $(I, \varphi, p, q) \in \mathcal{P}_{\sqrt{2c_2\epsilon^2}} = V \times \mathbb{T}^n \times \{p_i^2 + q_i^2 < 2c_2\epsilon^2\}$, and we also have a map ϕ, which is symplectic, hence volume preserving, mapping from $\mathcal{V} \times \mathbb{T}^n$ to $\mathcal{P}_{\sqrt{2c_2\epsilon^2}}$. Hence we can conclude that $\text{meas}(\mathcal{V})\text{meas}(\mathbb{T}^n) = \text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}})$.

Claim: $\frac{r}{\epsilon^2} < 1$

Proof: We have that, for some suitable constant C

$$r = 2\frac{\mu\epsilon^{2s+1}C_+}{\hat{\rho}}\left(\frac{1}{C_+} + \frac{1}{C_+\mu} + 4nC_+\mu^{-1}\right) = C\frac{\epsilon^{2s+1}}{\hat{\rho}}$$

The expression of $\hat{\rho}$ is already known, we can substitute each term (see equation 3.19) to prove our claim:

$$\begin{aligned}
\frac{r}{\epsilon^2} &= C\frac{\epsilon^{2s+1}}{\hat{\rho}\epsilon^2} = C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{2C_+K^{\tau+1}}{\gamma_1} \\
&\leq C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{2C_+}{\gamma_*\epsilon^{(2s+1)/2}} \\
&\leq \frac{2CC_+}{\gamma_*}\epsilon^{s+\frac{1}{2}-2} << 1
\end{aligned}$$

and similar we obtain

$$\begin{aligned}\frac{r}{\epsilon^2} &= C\frac{\epsilon^{2s+1}}{\hat\rho\epsilon^2} = C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{2C_+K^{\tau+1}}{\hat\gamma_2} \\ &\le \frac{2CC_+}{\gamma_*}\epsilon^{s+\frac{1}{2}-2} << 1\end{aligned}$$

and

$$\begin{aligned}\frac{r}{\epsilon^2} &= C\frac{\epsilon^{2s+1}}{\hat\rho\epsilon^2} = C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{1}{c_1\epsilon^{(2s+1)/2}} \\ &\le \frac{C}{c_1}\epsilon^{s+\frac{1}{2}-2} << 1\end{aligned}$$

and also

$$\begin{aligned}\frac{r}{\epsilon^2} &= C\frac{\epsilon^{2s+1}}{\hat\rho\epsilon^2} = C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{32\tilde M\bar K^{\tau+1}}{\bar\gamma} \\ &\le C\frac{\epsilon^{2s+1}}{\epsilon^2}\frac{1}{\sqrt{\mu}} << 1\end{aligned}$$

given that μ is of order $O\left((\log\epsilon^{-1})^{-2b}\right)$. ■

Denoting $B := B^{n_2}_{c_2\epsilon^2}$, one has that

$$\text{meas}(Re(V_r)\backslash V) \le C_V r\text{meas}(V)$$

hence

$$\text{meas}(V_r) \le (1 + C_V r)\text{meas}(V) \tag{3.26}$$

and

$$\begin{aligned}\text{meas}(B_r\backslash B) &= \text{meas}(B_r) - \text{meas}(B) \\ &= \sigma_{n_2}[(c_2\epsilon^2 + r)^{n_2} - (c_2\epsilon^2)^{n_2}] \\ &= \sigma_{n_2}(c_2\epsilon^2)^{n_2}[\left(1 + \frac{r}{c_2\epsilon^2}\right)^{n_2} - 1] \\ &= \text{meas}(B)[\left(1 + \frac{r}{c_2\epsilon^2}\right)^{n_2} - 1] \\ &\cong \text{meas}(B)\frac{n_2 r}{c_2\epsilon^2}\end{aligned} \tag{3.27}$$

$$\begin{aligned}\text{meas}(\mathcal{V}\backslash\mathcal{D}) &= \text{meas}(V \times B\backslash\bar D \times \mathcal{A}(\epsilon)) \\ &\le \text{meas}(V\backslash\bar D \times B) + \text{meas}(V \times B\backslash\mathcal{A}(\epsilon)) \\ &\le \bar C\bar\gamma\text{meas}(V \times B) + C'\epsilon^{(2s-3)n_2/2}\text{meas}(V \times B) \\ &\le \hat C(\bar\gamma + \epsilon^{(2s-3)n_2/2})\text{meas}(\mathcal{V})\end{aligned} \tag{3.28}$$

and from equation (3.26,3.27)

$$\begin{aligned}
\text{meas}(\mathcal{V}_r\backslash\mathcal{V}) &\leq \text{meas}((V_r\times B_r)\backslash(V\times B))\\
&\leq \text{meas}(V_r\backslash V\times B_r)+\text{meas}(V_r\times B_r\backslash B)\\
&\leq C_V r\text{meas}(V\times B_r)+\left[\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2}-1\right]\text{meas}(V_r\times B)\\
&\leq C_V r\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2}\text{meas}(V\times B)\\
&\quad+\left[\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2}-1\right](1+C_V r)\text{meas}(V\times B)\\
&\leq C_V r\left(1+\frac{n_2 r}{c_2\epsilon^2}\right)\text{meas}(V\times B)+\left(\frac{n_2 r}{c_2\epsilon^2}\right)(1+C_V r)\text{meas}(V\times B)\\
&\leq\left(C_V r+2C_V r\frac{n_2 r}{c_2\epsilon^2}+\frac{n_2 r}{c_2\epsilon^2}\right)\text{meas}(\mathcal{V})
\end{aligned}\tag{3.29}$$

Thus from equation (3.28,3.29)

$$\begin{aligned}
&\text{meas}(Re(\mathcal{D}_r)\backslash\mathcal{D}\times\mathbb{T}^n)\leq\text{meas}(Re(\mathcal{V}_r)\backslash\mathcal{D}\times\mathbb{T}^n)\\
&\leq\text{meas}(Re(\mathcal{V}_r)\backslash\mathcal{V}\times\mathbb{T}^n)+\text{meas}(\mathcal{V}\backslash\mathcal{D}\times\mathbb{T}^n)\\
&\leq\left(2C_V r\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2}+\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2}-C_V r-1\right)\text{meas}(\mathcal{V}\times\mathbb{T}^n)\\
&+\hat{C}(\bar{\gamma}+\epsilon^{(2s-3)n_2/2})\text{meas}(\mathcal{V}\times\mathbb{T}^n)\\
&\leq\left(C_V r+2C_V r\frac{n_2 r}{c_2\epsilon^2}+\frac{n_2 r}{c_2\epsilon^2}+\hat{C}(\bar{\gamma}+\epsilon^{(2s-3)n_2/2})\right)\text{meas}(\mathcal{V}\times\mathbb{T}^n)
\end{aligned}\tag{3.30}$$

And we can also prove that (with some suitable constant C)

$$\begin{aligned}
\text{meas}(\mathcal{D}\backslash\mathcal{D}^*_{\gamma_1,\mu\hat{\gamma}_2,\tau}\times\mathbb{T}^n)&\leq\text{meas}(\mathcal{R}_{\gamma_1,\hat{\gamma}_2,\tau}\times\mathbb{T}^n)\\
&\leq C(\gamma_1+\frac{\hat{\gamma}_2}{\epsilon^2})\text{meas}(\bar{D}\times B_{\check{c}_2\epsilon^2}\times\mathbb{T}^n)\\
&\leq C(\gamma_1+\frac{\hat{\gamma}_2}{\epsilon^2})\text{meas}(\mathcal{P}_{\sqrt{2\check{c}_2\epsilon^2}})
\end{aligned}\tag{3.31}$$

Hence from equations (3.25,3.30,3.31) we have that

$$
\begin{aligned}
\text{meas}\mathcal{K} &= \text{meas}K \\
&\geq \text{meas}(Re(\mathcal{D}_r)\times\mathbb{T}^n) - c_n\left[\text{meas}(\mathcal{D}\backslash\mathcal{D}^*_{\gamma_1,\gamma_2,\tau}\times\mathbb{T}^n) + \text{meas}((Re(\mathcal{D}_r)\backslash\mathcal{D})\times\mathbb{T}^n)\right] \\
&\geq (1-\bar{C}\bar{\gamma})\left[1-\left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) - c_nC\left(\gamma_1+\frac{\hat{\gamma}_2}{\epsilon^2}\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) \\
&\quad - c_n\left(2C_Vr\left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2} + \left(1+\frac{r}{c_2\epsilon^2}\right)^{n_2} - C_Vr - 1 + \hat{C}(\bar{\gamma}+\epsilon^{(2s-3)n_2/2})\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) \\
&\geq (1-\bar{C}\bar{\gamma})\left[1-\left(\frac{c_1}{c_2}\right)^{n_2}\epsilon^{(2s-3)n_2/2}\right]\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) - c_nC\left(\gamma_1+\frac{\hat{\gamma}_2}{\epsilon^2}\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) \\
&\quad - c_n\left(2C_Vr\frac{n_2r}{c_2\epsilon^2} + \frac{n_2r}{c_2\epsilon^2} + C_Vr + \hat{C}(\bar{\gamma}+\epsilon^{(2s-3)n_2/2})\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) \\
&\geq (1-\bar{C}\bar{\gamma})\left[1-C_1\epsilon^{(2s-3)n_2/2}\right]\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) - C_2\left(\gamma_1+\frac{\hat{\gamma}_2}{\epsilon^2}\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}}) \\
&\quad - C_2\left(\frac{\hat{\gamma}_2}{\epsilon^2}+\bar{\gamma}+\epsilon^{(2s-3)n_2/2}\right)\text{meas}(\mathcal{P}_{\sqrt{2c_2\epsilon^2}})
\end{aligned}
$$

by replacing $\sqrt{2c_2\epsilon^2}$ with ϵ, we obtain equation (1.3), hence the Theorem 1 is proved.

3.1 Two–scale KAM theorem [3]

Theorem 3 ([3], Proposition 3)**.** *Let* $n_1, n_2 \in \mathbb{N}, n := n_1+n_2, \tau > n, \gamma_1 > \gamma_2 > 0, s > 0, \bar{s} > 0, \rho > 0, D \subset \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}, A := D_\rho$, *and let*

$$H(I,\psi) = h(I) + f(I,\psi)$$

be real-analytic on $A\times\mathbb{T}^n_{\bar{s}+s}$*. Assume that* $\omega_0 := \partial h$ *is a diffeomorphism of* A *with non singular Hessian matrix* $U := \partial\omega_0 = \partial^2 h$ *and let* $\hat{U}$ *denote the* $n\times n_1$ *submatrix of* U*. Let*

$$M \geq \sup_A\|U\|, \quad \hat{M} \geq \sup_A\|\hat{U}\|, \quad \bar{M} \geq \sup_A\|U^{-1}\|, \quad E \geq \|f\|_{\rho,\bar{s}+s}$$

define

$$
\begin{aligned}
\hat{c} &:= \max\{2^8 n, 12^\tau\} \\
\hat{\rho} &:= \min\{\frac{\gamma_1}{2MK^{\tau+1}}, \frac{\gamma_2}{2\hat{M}K^{\tau+1}}, \rho\} \\
K &:= \frac{6}{s}\log\left(8(12)^{2(\tau+1)}\right)
\end{aligned}
$$

and let $\bar{M}^{(1)}, \bar{M}^{(2)}$ be the upper bounds on the norms of the sub-matrices $n_1 \times n, n_2 \times n$ of U^{-1} of the first n_1, last n_2 rows. Assume the following conditions hold

$$\hat{c}\frac{E}{M\hat{\rho}^2} < 1, \hat{c}\frac{E}{\hat{M}\hat{\rho}^2} < 1, \hat{c}\frac{E\bar{M}}{\hat{\rho}^2} < 1 \tag{3.32}$$

Then, for any $\omega \in \Omega_* := \omega_0(D) \cap \mathcal{D}_{\gamma_1,\gamma_2,\tau}$, one can find a unique real-analytic embedding

$$\phi_\omega : \theta \in \mathbb{T}^n \to ((v_1(\theta;\omega), v_2(\theta;\omega)), \theta + u(\theta;\omega)) \in Re(D_r) \times \mathbb{T}^n \tag{3.33}$$

where $r := 2\frac{E}{\hat{\rho}}(\frac{1}{M}+\frac{1}{\hat{M}}+4n\bar{M})$ such that $T_\omega := \phi_\omega(\mathbb{T}^n)$ is a real-analytic $n-$dimensional H-invariant torus, on which the H-flow is analytically conjugated to $\theta \to \theta + \omega t$. Furthermore, the map $(\theta;\omega) \to \phi_\omega(\theta)$ is Lipschitz and one-to-one and the invariant set $\mathcal{K} := \bigcup_{\omega\in\Omega_*} T_\omega$ satisfies the following measure estimate

$$meas\,(Re(D_r) \times \mathbb{T}^n \backslash \mathcal{K}) \leq c_n\,(meas(D\backslash D_0 \times \mathbb{T}^n) + meas((Re(D_r)\backslash D) \times \mathbb{T}^n)) \tag{3.34}$$

where $D_0 := D_{\gamma_1,\gamma_2,\tau} := \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap D$ and $c_n := 1 + (1 + 32n\frac{E}{\hat{\rho}^2}\max\{\frac{1}{M}, \frac{1}{\hat{M}}\})^{2n}$. Finally, on $\mathbb{T} \times \Omega_*$, the following uniform estimates hold

$$\begin{aligned} |v_1(\cdot;\omega) - I_1^0(\omega)| &\leq 2\frac{E}{M\hat{\rho}} + 5n\frac{\bar{M}^{(1)}E}{\hat{\rho}} \\ |v_2(\cdot;\omega) - I_2^0(\omega)| &\leq 2\frac{E}{\hat{M}\hat{\rho}} + 5n\frac{\bar{M}^{(2)}E}{\hat{\rho}} \\ |u(\cdot;\omega)| &\leq 2\frac{Es}{M\hat{\rho}^2} \end{aligned} \tag{3.35}$$

where v_i denotes the projection of $v \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ over $\mathbb{R}^{n_i}$ and $I^0(\omega) = (I_1^0(\omega), I_2^0(\omega)) \in D$ is the ω_0-pre-image of $\omega \in \Omega_*$.

Theorem 3 is proven in [8, 3]. We report its proof here for sake of completeness. As in the tradition of KAM theory in the real–analytic class, an iterative lemma is used.

3.2 Iterative Lemma [3]

Lemma 3. *[Lemma B.1, [3]] Let $n_1, n_2 \in \mathbb{N}, n := n_1 + n_2, \tau > n, \gamma_1 > \gamma_2 > 0, s > 0, \bar{s} > 0, \rho > 0, D \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, A := D_\rho$, and let*

$$H(I,\psi) = h(I) + f(I,\psi)$$

for any $1 \leq j \in \mathbb{N}$, there exists a domain $D_j \subset \mathbb{R}^n$, two positive numbers ρ_j, s_j and a real-analytic and symplectic transformation Φ_j on $W_j := (D_j)_{\rho_j} \times \mathbb{T}^n_{\bar{s}+s_j}$ which conjugates $H_0 := H$ to

$$H_j = H_0 \circ \Phi_j = h_j + f_j$$

Letting, for $j = 0$, $D_0 := \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap D, s_0 = s, \rho_0 := \rho, M_0 := M, \hat{M}_0 := \hat{M}, \bar{M}_0 := \bar{M}, \bar{M}_0^{(1)} := \bar{M}^{(1)}, \bar{M}_0^{(2)} := \bar{M}^{(2)}, E_0 := E, K_0 := K, \hat{\rho}_0 := \hat{\rho}$,

and, for $j = 1$, $\rho_1 = \hat{\rho}_0/16, s_1 = s_0/12$, $M_1 = M_0 + \frac{1}{2}M_0, \hat{M}_1 = \hat{M}_0 + \frac{1}{2}\hat{M}_0, \bar{M}_1 = 2\bar{M}_0, \bar{M}_1^{(1)} = 2\bar{M}_0^{(1)}, \bar{M}_1^{(2)} = 2\bar{M}_0^{(2)}$

$$\begin{aligned} K_1 &= \frac{6}{s_1} \log\left(8(12)^{2(\tau+1)} \frac{M_0 \hat{\rho}_0^2}{E_0}\right), \\ E_1 &= \frac{1}{8(12)^{2(\tau+1)}} E_0, \\ \hat{\rho}_1 &= \min\left\{\frac{\gamma_1}{2M_1 K_1^{\tau+1}}, \frac{\gamma_2}{2\hat{M}_1 K_1^{\tau+1}}, \rho_1\right\} \end{aligned} \tag{3.36}$$

and for $j > 1$, $\rho_j = \hat{\rho}_{j-1}/16, s_j = s_{j-1}/12, M_j = M_{j-1} + \frac{1}{2^{j-1}} M_{j-1}, \hat{M}_j = \hat{M}_{j-1} + \frac{1}{2^{j-1}} \hat{M}_{j-1}, \bar{M}_j = 2\bar{M}_{j-1}, \bar{M}_j^{(1)} = 2\bar{M}_{j-1}^{(1)}, \bar{M}_j^{(2)} = 2\bar{M}_{j-1}^{(2)}$

$$\begin{aligned} K_j &= \frac{6}{s_j} \log\left(8(12)^{2(\tau+1)} \frac{M_0 \hat{\rho}_0^2}{E_0}\right), \\ E_j &= \frac{1}{8(12)^{2(\tau+1)}} \frac{E_0}{M_0 \hat{\rho}_0^2} E_{j-1}, \\ \hat{\rho}_j &= \min\left\{\frac{\gamma_1}{2M_j K_j^{\tau+1}}, \frac{\gamma_2}{2\hat{M}_j K_j^{\tau+1}}, \rho_j\right\} \end{aligned} \tag{3.37}$$

then

(1) the map

$$\hat{\iota}_j = (\hat{\iota}_j^{(1)}, \hat{\iota}_j^{(2)}) : I \in D_{j-1} \to \omega_j^{-1} \circ \omega_{j-1}(I) \in D_j \tag{3.38}$$

verifies

$$\begin{aligned} \sup_{D_{j-1}} |\hat{\iota}_j^{(1)} - id| &\leq \frac{52n}{11} \frac{\bar{M}_{j-1}^{(1)} E_{j-1}}{\hat{\rho}_{j-1}} \\ \sup_{D_{j-1}} |\hat{\iota}_j^{(2)} - id| &\leq \frac{52n}{11} \frac{\bar{M}_{j-1}^{(2)} E_{j-1}}{\hat{\rho}_{j-1}} \end{aligned} \tag{3.39}$$

and

$$\mathcal{L}(\hat{\imath}_j - id) \le 2^6 n \frac{\bar{M}_{j-1} E_{j-1}}{\hat{\rho}_{j-1}^2} \tag{3.40}$$

(2) the perturbation f_j has sup-Fourier norm on W_j

$$\|f_j\|_{W_j} \le E_j \tag{3.41}$$

(3) the real-analytic symplectomorphism Φ_j is defined as $\Phi_j = \Psi_1 \circ \cdots \circ \Psi_j$, where

$$\Psi_k : (I_k, \psi_k) \in W_k \to (I_{k-1}, \psi_{k-1}) \in W_{k-1} \tag{3.42}$$

verifies

$$\begin{aligned} \sup_{(D_k)_{\rho_k} \times \mathbb{T}^n_{\bar{s}+s_k}} |I^{(1)}_{k-1}(I_k, \psi_k) - I^{(1)}_k| &\le \frac{3E_{k-1}}{2M_{k-1}\hat{\rho}_{k-1}} \\ \sup_{(D_k)_{\rho_k} \times \mathbb{T}^n_{\bar{s}+s_k}} |I^{(2)}_{k-1}(I_k, \psi_k) - I^{(2)}_k| &\le \frac{3E_{k-1}}{2\hat{M}_{k-1}\hat{\rho}_{k-1}} \\ \sup_{(D_k)_{\rho_k} \times \mathbb{T}^n_{\bar{s}+s_k}} |\psi_{k-1}(I_k, \psi_k) - \psi_k| &\le \frac{3E_{k-1}s_{k-1}}{2\hat{M}_{k-1}\hat{\rho}^2_{k-1}} \end{aligned} \tag{3.43}$$

and the rescaled dimensionless map $\check{\Psi}_k - id := 1_{\hat{\rho},s}\Psi_k \circ 1^{-1}_{\hat{\rho},s} - id$ has Lipschitz constant on $(\check{D}_k)_{\rho_k/\rho} \times \check{\mathbb{T}}^n_{(\bar{s}+s_k)/s}$

$$\mathcal{L}(\check{\Psi}_k - id) \le \begin{cases} 8n \max\{\frac{E_{k-1}}{M_{k-1}\hat{\rho}_{k-1}}, \frac{E_{k-1}}{\hat{M}_{k-1}\hat{\rho}_{k-1}}\} \left(\frac{12^{\tau+1}}{6}\right)^{k-1}, & k > 1 \\ 36n \max\{\frac{E_0}{M_0\hat{\rho}_0}, \frac{E_0}{\hat{M}_0\hat{\rho}_0}\}, & k = 1 \end{cases} \tag{3.44}$$

where id denotes the identity map, 1_d denotes the $d \times d$ identity matrix, $1_{\rho,s} := (\rho^{-1}1_n, s^{-1}1_n)$, $\check{D}_k := \hat{\rho}^{-1}D_k, \check{\mathbb{T}} := \mathbb{R}/(2\pi/s)\mathbb{Z}$.

Chapter 4

Proofs of the Iterative Lemma and Theorem 3

4.1 Proof of the Iterative Lemma

Proof. We apply the averaging theorem (Prop 1), with $m = 0, l_1 = n_1, l_2 = n_2, B = B' = \{0\}$. We have that $D_0 = \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap D$.

First we consider the 1st iteration. Our Hamiltonian is

$$H_0(I, \psi) = H(I, \psi) = h(I) + f(I, \psi) =: h_0(I) + f_0(I, \psi) \tag{4.1}$$

and $\omega_0 := \partial h$, clearly $\omega_0(D_0) \subset \mathcal{D}_{\gamma_1,\gamma_2,\tau}$.

To apply the averaging theorem, the non-resonance condition (2.1) needs to be verified. For $0 < |k| \leq K_0, I \in (D_0)_{\hat{\rho}_0}$, and the choice of $\hat{\rho}_0$,we have that

$$|\omega_0(I) \cdot k| \geq \begin{cases} \frac{\gamma_1}{K_0^\tau} - M_0\hat{\rho}_0 K_0 \geq \frac{\gamma_1}{2K_0^\tau} =: \alpha_1, & k_1 \neq 0 \\ \frac{\gamma_2}{K_0^\tau} - \hat{M}_0\hat{\rho}_0 K_0 \geq \frac{\gamma_2}{2K_0^\tau} =: \alpha_2, & k_1 = 0 \end{cases} \tag{4.2}$$

indeed, if $I_0 \in D_0$, we have

$$|\omega_0(I_0) \cdot k| \geq \begin{cases} \frac{\gamma_1}{k^\tau} \geq \frac{\gamma_1}{K_0^\tau}, & k_1 \neq 0 \\ \frac{\gamma_2}{k^\tau} \geq \frac{\gamma_2}{K_0^\tau}, & k_1 = 0 \end{cases}$$

and $\forall I \in (D_0)_{\hat{\rho}_0}, \exists I_0 \in D_0$ be such that

$$|I - I_0| \leq \hat{\rho}_0$$

hence

$$\begin{aligned}|\omega_0(I)\cdot k| &= |\omega_0(I)\cdot k - \omega_0(I_0)\cdot k + \omega_0(I_0)\cdot k|\\ &\geq |\omega_0(I_0)\cdot k| - |\omega_0(I)\cdot k - \omega_0(I_0)\cdot k|\\ &\geq |\omega_0(I_0)\cdot k| - |\omega_0(I) - \omega_0(I_0)|K_0\\ &\geq |\omega_0(I_0)\cdot k| - \|\omega_0\||I - I_0|K_0\\ &\geq \begin{cases}\frac{\gamma_1}{K_0^\tau} & -M_0\hat{\rho}_0 K_0, k_1 \neq 0\\ \frac{\gamma_2}{K_0^\tau} & -\hat{M}_0\hat{\rho}_0 K_0, k_1 = 0\end{cases}\end{aligned}$$

since $K_0 = \frac{6}{s_0}\log\left(8(12)^{2(\tau+1)}\right)$, clearly we have $K_0 s_0 \geq 6$. Now we need to verify boundedness condition (2.2) with $d_0 := \hat{\rho}_0 s_0$ and $c_0 = e/2$

$$\begin{aligned}E_0\frac{2^7 c_0 K_0 s_0}{\alpha_2 d_0} &\leq E_0\frac{2^7 c_0 K_0}{\alpha_2\hat{\rho}_0}\\ &\leq E_0\frac{2^7 c_0 K_0}{\frac{\gamma_2}{2K_0^\tau}\hat{\rho}_0}\\ &\leq \frac{2^7 c_0 E_0}{M_0\hat{\rho}_0^2}\\ &< 1\end{aligned} \tag{4.3}$$

Hence, according to the average theorem we have a transformation

$$\Psi_1 : (I_1, \psi_1) \in (D_0)_{\hat{\rho}_0/2} \times \mathbb{T}^n_{\bar{s}+s_0/6} \to (I_0, \psi_0) \in W_0 := (D_0)_{\hat{\rho}_0} \times \mathbb{T}^n_{\bar{s}+s_0} \tag{4.4}$$

and the Hamiltonian is transformed into the form

$$\begin{aligned}H_1 &:= H_0 \circ \Psi_1\\ &= h_0(I_1) + g_1(I_1) + f_1(I_1, \psi_1)\\ &=: h_1(I_1) + f_1(I_1, \psi_1)\end{aligned} \tag{4.5}$$

From averaging theorem we have that

$$\begin{aligned}\|g_1 - \bar{f}_0\|_{\hat{\rho}_0/2, s_0/6} &\leq \frac{12}{11}\frac{2^7 c_0 E_0^2}{\alpha_2 d_0} = \frac{12}{11}\frac{2^7 c_0 E_0^2}{\frac{\gamma_2}{2K_0^\tau}\hat{\rho}_0 s_0}\\ &\leq \frac{2}{11}\frac{2^7 c_0 E_0^2}{\frac{\hat{M}_0\gamma_2}{2\hat{M}_0 K_0^{\tau+1}}\hat{\rho}_0} \leq \frac{2}{11}\frac{2^7 c_0 E_0^2}{\hat{M}_0\hat{\rho}_0^2}\\ &\leq \frac{2}{11}E_0\end{aligned} \tag{4.6}$$

and

$$\begin{aligned} \|f_1\|_{\hat{\rho}_0/2, s_0/6} &\leq e^{-K_0 s_0/6} \frac{2^9 c_0 E_0^2}{\alpha_2 d_0} \\ &\leq \frac{1}{8(12)^{2(\tau+1)}} \frac{2}{3} E_0 \leq E_1 \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} |I_0^{(1)}(I_1, \psi_1) - I_1^{(1)}| &\leq \frac{9E_0}{\alpha_1 s_0} \leq \frac{9E_0}{\frac{\gamma_1}{2K_0^\tau} s_0} \\ &\leq \frac{3E_0}{\frac{\gamma_1}{K_0^{\tau+1}}} \\ &\leq \frac{3E_0}{2M_0 \hat{\rho}_0} \end{aligned} \tag{4.8}$$

similarly we obtain

$$\begin{aligned} |I_0^{(2)}(I_1, \psi_1) - I_1^{(2)}| &\leq \frac{9E_0}{\alpha_2 s_0} \leq \frac{3E_0}{2\hat{M}_0 \hat{\rho}_0} \\ |\psi_0(I_1, \psi_1) - \psi_1| &\leq \frac{9E_0}{\alpha_2 \hat{\rho}_0} \leq \frac{3E_0 s_0}{2\hat{M}_0 \hat{\rho}_0^2} \end{aligned} \tag{4.9}$$

We also estimate the strength of g_1

$$\begin{aligned} \sup_{(D_0)_{\hat{\rho}_0/2}} |g_1| &\leq \sup_{(D_0)_{\hat{\rho}_0/2}} |g_1 - \bar{f}_0| + \sup_{(D_0)_{\hat{\rho}_0/2}} |\bar{f}_0| \\ &\leq \frac{13}{11} E_0 \end{aligned} \tag{4.10}$$

and with the Cauchy estimate we obtain

$$\begin{aligned} \sup_{(D_0)_{\hat{\rho}_0/4}} \|\partial^2 g_1\| = \sup_{(D_0)_{\hat{\rho}_0/4}} \|\partial^2 g_1\|_\infty &\leq \frac{\sup_{(D_0)_{3\hat{\rho}_0/8}} |\partial g_1|_\infty}{\hat{\rho}_0/8} \\ &\leq \frac{\sup_{(D_0)_{\hat{\rho}_0/2}} |g_1|}{(\hat{\rho}_0/8)^2} \\ &\leq 2^6 \frac{13}{11} \frac{E_0}{\hat{\rho}_0^2} \end{aligned} \tag{4.11}$$

from which we have, for $\omega_1 := \partial h_1 = \partial h_0 + \partial g_1$,

$$\begin{aligned} \sup_{(D_0)_{\hat{\rho}_0/4}} \|\omega_1\| = \sup_{(D_0)_{\hat{\rho}_0/4}} \|\partial^2 h_0 + \partial^2 g_1\| \\ &\leq M_0 + 2^6 \frac{13}{11} \frac{E_0}{\hat{\rho}_0^2} \\ &\leq M_0 + \frac{1}{2} M_0 =: M_1 \end{aligned} \tag{4.12}$$

and for $\|\hat{\omega}_1\|$ we have

$$\sup_{(D_0)_{\hat{\rho}_0/4}} \|\hat{\omega}_1\| = \sup_{(D_0)_{\hat{\rho}_0/4}} \|\widehat{\partial^2 h_0} + \widehat{\partial^2 g_1}\| \tag{4.13}$$
$$\leq \hat{M}_0 + 2^6 \frac{13}{11} \frac{E_0}{\hat{\rho}_0^2}$$
$$\leq \hat{M}_0 + \frac{1}{2} \hat{M}_0 =: \hat{M}_1$$

To estimate $\|(\partial\omega_1)^{-1}\|$, we first estimate the following

$$\sup_{(D_0)_{\hat{\rho}_0/4}} \|\partial^2 g_1 (\partial^2 h_0)^{-1}\| \leq \sup\|\partial^2 g_1\| \sup\|(\partial^2 h_0)^{-1}\| \tag{4.14}$$
$$\leq 2^6 \frac{13}{11} \frac{E_0}{\hat{\rho}_0^2} \bar{M}_0 \leq \frac{1}{2}$$

and, by Woodbury formula we have

$$\begin{aligned}\sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0 + \partial^2 g_1)^{-1}\| &= \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1} - (\partial^2 h_0)^{-1} \partial^2 g_1 (\partial^2 h_0 + \partial^2 g_1)^{-1}\| \\ &\leq \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1}\| + \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1} \partial^2 g_1 (\partial^2 h_0 + \partial^2 g_1)^{-1}\| \\ &\leq \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1}\| + \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1} \partial^2 g_1\| \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0 + \partial^2 g_1)^{-1}\|\end{aligned} \tag{4.15}$$

from which we have

$$\begin{aligned}\sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial\omega_1)^{-1}\| = \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0 + \partial^2 g_1)^{-1}\| &\leq \frac{\sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1}\|}{1 - \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1} \partial^2 g_1\|} \\ &\leq 2 \sup_{(D_0)_{\hat{\rho}_0/4}} \|(\partial^2 h_0)^{-1}\| \\ &\leq 2\bar{M}_0 =: \bar{M}_1\end{aligned} \tag{4.16}$$

Now we define a map

$$\hat{\iota}_1 = (\hat{\iota}_1^{(1)}, \hat{\iota}_1^{(2)}) := \omega_1^{-1} \circ \omega_0|_{D_0} \tag{4.17}$$

let $I_0 \in (D_0)_{\hat{\rho}_0/16}$ be such that $|I - I_0| = r < \frac{\hat{\rho}_0}{16}$, then

$$\begin{aligned}|\omega_1(I) - \omega_1(I_0)| &= |\omega_0(I) + \partial g_1(I) - \omega_0(I_0) - \partial g_1(I_0)| \\ &= |\omega_0(I) - \omega_0(I_0) + \partial g_1 \circ \omega_0^{-1}(\omega_0(I)) - \partial g_1 \circ \omega_0^{-1}(\omega_0(I_0))| \\ &\geq (1 - \|\partial^2 g_1 (\partial^2 h_0)^{-1}\|)|\omega_0(I) - \omega_0(I_0)| \\ &\geq \frac{1}{2} \frac{1}{\bar{M}_0} |I - I_0| = \frac{r}{2\bar{M}_0}\end{aligned}$$

this indicates that

$$\omega_1\left((D_0)_{\hat\rho_0/16}\right) \supseteq (\omega_1(D_0))_{\hat\rho_0/(32\bar M_0)} \tag{4.18}$$

since we also have

$$\begin{aligned}\sup_{D_0}|\omega_0-\omega_1|_1 = \sup_{D_0}|\partial g_1|_1 &\le n\frac{13/11E_0}{\hat\rho_0/2}\\ &\le n\frac{26}{11}\frac{\hat\rho_0}{\bar M_0}\frac{\bar M_0E_0}{\hat\rho_0^2}\\ &\le n\frac{26}{11}\frac{\hat\rho_0}{\bar M_0}\frac{1}{2^8n} \le \frac{\hat\rho_0}{32\bar M_0}\end{aligned} \tag{4.19}$$

hence

$$\omega_0(D_0) \subset (\omega_1(D_0))_{\hat\rho_0/(32\bar M_0)} \subseteq \omega_1\left((D_0)_{\hat\rho_0/16}\right) \tag{4.20}$$

By our definition, $\hat\iota_1$ maps from D_0 to D_1, and we have

$$\begin{aligned}\sup_{D_0}|\hat\iota_1^{(1)}(I_0)-I_0^{(1)}| &= \sup_{D_0}|\hat\iota_1^{(1)}(I_0)-I_0^{(1)}|_1\\ &= \sup_{D_0}|(\omega_1^{-1})^{(1)}\circ\omega_0(I_0)-(\omega_1^{-1})^{(1)}\circ\omega_1(I_0)|_1\\ &\le \bar M_1^{(1)}\sup_{D_0}|\omega_0(I_0)-\omega_1(I_0)|\\ &= \frac{26n}{11}\frac{\bar M_1^{(1)}E_0}{\hat\rho_0}\\ &= \frac{52n}{11}\frac{\bar M_0^{(1)}E_0}{\hat\rho_0}\end{aligned} \tag{4.21}$$

And similarly we could get

$$\sup_{D_0}|\hat\iota_1^{(2)}(I_0)-I_0^{(2)}| = \sup_{D_0}|\hat\iota_1^{(2)}(I_0)-I_0^{(2)}|_1 \le \frac{52n}{11}\frac{\bar M_0^{(2)}E_0}{\hat\rho_0} \tag{4.22}$$

and

$$\begin{aligned}\sup_{(D_0)_{\hat\rho_0/8}}|\hat\iota_1(I)-I|_\infty \le \bar M_1 \sup_{(D_0)_{\hat\rho_0/8}}|\partial g_1|_\infty &\le 2\bar M_0\frac{\sup_{(D_0)_{\hat\rho_0/2}}|g_1|}{3\hat\rho_0/8}\\ &\le 2\bar M_0\frac{13/11E_0}{3\hat\rho_0/8}\\ &\le \frac{208}{33}\frac{\bar M_0E_0}{\hat\rho_0}\end{aligned} \tag{4.23}$$

And the Lipschitz constant of $\hat{\imath}_1 - id$ is bounded by

$$\begin{aligned} \mathcal{L}(\hat{\imath}_1 - id) &\leq n \inf_{0<r<\hat{\rho}_0/8} \sup_{(D_0)_r} \|D(\hat{\imath}_1 - id)\| \\ &\leq n \inf_{(0,\hat{\rho}_0/8)} \frac{\sup_{(D_0)_{\hat{\rho}_0/8}} |\hat{\imath}_1(I) - I|_\infty}{\hat{\rho}_0/8 - r} \\ &\leq \frac{8n}{\hat{\rho}_0} \frac{208}{33} \frac{\bar{M}_0 E_0}{\hat{\rho}_0} \leq 2^6 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2} \end{aligned} \tag{4.24}$$

Since $\omega_0(D_0) = \omega_1(D_1)$ and $\omega_0(D_0) \subseteq \omega_1((D_0)_{\hat{\rho}_0/16})$, this indicates that

$$D_1 \subseteq (D_0)_{\hat{\rho}_0/16} \tag{4.25}$$

hence

$$\begin{aligned} &(D_1)_{\hat{\rho}_1} \subset (D_0)_{\hat{\rho}_0/8} \\ \Rightarrow &\left(\frac{D_1}{\hat{\rho}_0}\right)_{\hat{\rho}_1/\hat{\rho}_0} \subseteq \left(\frac{D_0}{\hat{\rho}_0}\right)_{\hat{\rho}_0/(8\hat{\rho}_0)} \subseteq \left(\frac{D_0}{\hat{\rho}_0}\right)_{\hat{\rho}_0/(2\hat{\rho}_0)} \end{aligned}$$

We define a scaled map, on $(\frac{D_1}{\hat{\rho}_0})_{\hat{\rho}_1/\hat{\rho}_0} \times (\frac{\mathbb{T}^n}{s_0})_{(\bar{s}+s_1)/s_0}$,

$$\check{\Psi}_1 - id : 1_{\hat{\rho}_0,s_0} \Psi_1 \circ 1_{\hat{\rho}_0,s_0}^{-1} - id$$

and with Cauchy estimate

$$\begin{aligned} \|D(\check{\Psi}_1 - id)\|_{\left(\frac{D_1}{\hat{\rho}_0}\right)_{\hat{\rho}_1/\hat{\rho}_0} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_1)/s_0}} &\leq \|D(\check{\Psi}_1 - id)\|_{\left(\frac{D_0}{\hat{\rho}_0}\right)_{\hat{\rho}_0/(8\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_0/12)/s_0}} \\ &\leq \frac{1}{\min\{\frac{3\hat{\rho}_0}{8\hat{\rho}_0}, \frac{s_0}{12 s_0}\}} \|(\check{\Psi}_1 - id)\|_{\left(\frac{D_0}{\hat{\rho}_0}\right)_{\hat{\rho}_0/(2\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_0/6)/s_0}} \end{aligned}$$

and this can be evaluated (let $A := \left(\frac{D_0}{\hat{\rho}_0}\right)_{\hat{\rho}_0/(2\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_0/6)/s_0}$) because

$$\begin{aligned} \|(\check{\Psi}_1 - id)\|_A &\leq \max\{\frac{1}{\hat{\rho}_0}\|I_1 - I(I_1, \varphi_1)\|_A, \frac{1}{s_0}\|\varphi_1 - \varphi(I_1, \varphi_1)\|_A\} \\ &\leq \max\{\frac{1}{\hat{\rho}_0}\frac{3E_0}{2M_0\hat{\rho}_0}, \frac{1}{\hat{\rho}_0}\frac{3E_0}{2\hat{M}_0\hat{\rho}_0}, \frac{1}{s_0}\frac{3E_0 s_0}{2\hat{M}_0\hat{\rho}_0^2}\} \\ &\leq \max\{\frac{3E_0}{2M_0\hat{\rho}_0^2}, \frac{3E_0}{2\hat{M}_0\hat{\rho}_0^2}\} \end{aligned}$$

hence

$$\|D(\check{\Psi}_1 - id)\|_{\left(\frac{D_1}{\hat{\rho}_0}\right)_{\hat{\rho}_1/\hat{\rho}_0}\times\left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_1)/s_0}} \leq \frac{\max\{\frac{3E_0}{2M_0\hat{\rho}_0^2}, \frac{3E_0}{2\hat{M}_0\hat{\rho}_0^2}\}}{\min\{\frac{3\hat{\rho}_0}{8\hat{\rho}_0}, \frac{s_0}{12s_0}\}}$$
$$\leq 18\max\{\frac{E_0}{M_0\hat{\rho}_0^2}, \frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}$$

Now for the $(j+1)-$th iteration, $j \geq 1$, we have Hamiltonian of the form

$$H_j = h_j(I_j) + f_j(I_j, \psi_j) \tag{4.26}$$

and $\omega_j := \partial h_j$, note we have that $\omega_j(D_j) \in \mathcal{D}_{\gamma_1,\gamma_2,\tau}$. We need to check conditions (2.1) and (2.2) before applying averaging theorem. For $0 < |k| < K_j, I_j \in (D_j)_{\hat{\rho}_j}$, and due to the definition of $\hat{\rho}_j$, we have

$$|\omega_j(I_j)\cdot k| \geq \begin{cases} \frac{\gamma_1}{K_j^\tau} - M_j\hat{\rho}_jK_j = \frac{\gamma_1}{2K_j^\tau} =: \alpha_1, & k_1 \neq 0 \\ \frac{\gamma_2}{K_j^\tau} - \hat{M}_j\hat{\rho}_jK_j = \frac{\gamma_2}{2K_j^\tau} =: \alpha_2, & k_1 = 0 \end{cases} \tag{4.27}$$

and clearly we have $K_js_j \geq 6$ given the definition of K_j, and

$$\begin{aligned} E_j\frac{2^7c_0K_js_j}{\alpha_2 d_j} &= E_j\frac{2^7c_0K_j}{\alpha_2\hat{\rho}_j} = E_j\frac{2^7c_0K_j}{\frac{\gamma_2}{2K_j^\tau}\hat{\rho}_j} \\ &\leq 2^7c_0\frac{E_j}{\hat{M}_j\hat{\rho}_j^2} \end{aligned} \tag{4.28}$$

From the averaging theorem, we have a transformation

$$\Psi_{j+1} : (I_{j+1}, \psi_{j+1}) \in (D_j)_{\hat{\rho}_j/2} \times \mathbb{T}^n_{\bar{s}+s_j/6} \to (I_j, \psi_j) \in W_j := (D_j)_{\hat{\rho}_j} \times \mathbb{T}^n_{\bar{s}+s_j} \tag{4.29}$$

and the new Hamiltonian is

$$\begin{aligned} H_{j+1} &:= H_j \circ \Psi_{j+1} \\ &= h_j(I_{j+1}) + g_{j+1}(I_{j+1}) + f_{j+1}(I_{j+1}, \psi_{j+1}) \\ &=: h_{j+1}(I_{j+1}) + f_{j+1}(I_{j+1}, \psi_{j+1}) \end{aligned} \tag{4.30}$$

and we have the following estimates

$$\begin{aligned} \|g_{j+1} - \bar{f}_j\|_{\hat{\rho}_j/2, s_j/6} &\leq \frac{12}{11}\frac{2^7c_0E_j^2}{\alpha_2 d_j} \leq \frac{2}{11}\frac{2^7c_0E_j^2}{\hat{M}_j\hat{\rho}_j^2} \\ &\leq \frac{2}{11}E_j \end{aligned} \tag{4.31}$$

and

$$\begin{aligned}
\|f_{j+1}\|_{\hat{\rho}_j/2,s_j/6} &\leq e^{-K_js_j/6}\frac{2^9c_0E_j^2}{\alpha_2d_j} \leq e^{-K_js_j/6}\frac{2}{3}E_j \\
&\leq \frac{2}{3}\frac{1}{8(12)^{2(\tau+1)}}\frac{E_0}{M_0\hat{\rho}_0^2}E_j \\
&=: \frac{2}{3}E_{j+1}
\end{aligned}$$

and also

$$\begin{aligned}
|I_j^{(1)}(I_{j+1},\psi_{j+1}) - I_{j+1}^{(1)}| &\leq \frac{3E_j}{2M_j\hat{\rho}_j} \\
|I_j^{(2)}(I_{j+1},\psi_{j+1}) - I_{j+1}^{(2)}| &\leq \frac{3E_j}{2\hat{M}_j\hat{\rho}_j} \\
|\psi_j(I_{j+1},\psi_{j+1}) - \psi_{j+1}| &\leq \frac{3E_js_j}{2\hat{M}_j\hat{\rho}_j^2}
\end{aligned} \tag{4.32}$$

And the strength of g_{j+1} can be estimates

$$\begin{aligned}
\sup_{(D_j)_{\hat{\rho}_j/2}}|g_{j+1}| &\leq \sup_{(D_j)_{\hat{\rho}_j/2}}|g_{j+1}-\bar{f}_j| + \sup_{(D_j)_{\hat{\rho}_j/2}}|\bar{f}_j| \\
&\leq \frac{2}{11}E_j + E_j \leq \frac{13}{11}E_j
\end{aligned} \tag{4.33}$$

and from the Cauchy estimate we have

$$\begin{aligned}
\sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial^2 g_{j+1}\| = \sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial^2 g_{j+1}\|_\infty &\leq \frac{\sup_{(D_j)_{3\hat{\rho}_j/8}}|\partial g_{j+1}|_\infty}{\hat{\rho}_j/8} \\
&\leq \frac{\sup_{(D_j)_{\hat{\rho}_j/2}}|g_{j+1}|}{(\hat{\rho}_j/8)^2} \\
&\leq 2^6\frac{13}{11}\frac{E_j}{\hat{\rho}_j^2}
\end{aligned} \tag{4.34}$$

from which we have, for $\omega_{j+1} := \partial h_{j+1} = \partial h_j + \partial g_{j+1}$,

$$\begin{aligned}
\sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial\omega_{j+1}\| = \sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial^2 h_j + \partial^2 g_{j+1}\| &\leq \sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial^2 h_j\| + \sup_{(D_j)_{\hat{\rho}_j/4}}\|\partial^2 g_{j+1}\| \\
&\leq M_j + 2^6\frac{13}{11}\frac{E_j}{\hat{\rho}_j^2} = M_j + 2^6\frac{13}{11}\frac{E_j}{M_j\hat{\rho}_j^2}M_j
\end{aligned}$$

And now since $M_j < M_{j+1} < 2M_j$, $\hat{M}_j < \hat{M}_{j+1} < 2\hat{M}_j$ and, by definition, $K_{j+1} = 12K_j$, we obtain

$$\begin{aligned}
\hat{\rho}_{j+1} &= \min\{\frac{\gamma_1}{2M_{j+1}K_{j+1}^{\tau+1}}, \frac{\gamma_2}{2\hat{M}_{j+1}K_{j+1}^{\tau+1}}, \rho_{j+1} = \frac{\hat{\rho}_j}{16}\} \\
&\geq \min\{\frac{\gamma_1}{2\cdot 2M_j 12^{\tau+1}K_j^{\tau+1}}, \frac{\gamma_2}{2\cdot 2\hat{M}_j 12^{\tau+1}K_j^{\tau+1}}, \rho_{j+1} = \frac{\hat{\rho}_j}{16}\} \\
&\geq \frac{1}{2(12)^{\tau+1}} \min\{\frac{\gamma_1}{2M_jK_j^{\tau+1}}, \frac{\gamma_2}{2\hat{M}_jK_j^{\tau+1}}, \rho_j\} = \frac{\hat{\rho}_j}{2(12)^{\tau+1}}
\end{aligned}$$

therefore, according to the definition of E_j

$$\begin{aligned}
\frac{E_j}{M_j\hat{\rho}_j^2} &\leq \frac{E_0}{M_j\hat{\rho}_j^2}\left(\frac{E_0}{M_0\hat{\rho}_0^2}\frac{1}{8(12)^{2(\tau+1)}}\right)^{j-1}\frac{1}{8(12)^{2(\tau+1)}} \\
&\leq \frac{E_0}{M_0\hat{\rho}_j^2}\left(\frac{E_0}{M_0\hat{\rho}_0^2}\frac{1}{8(12)^{2(\tau+1)}}\right)^{j-1}\frac{1}{8(12)^{2(\tau+1)}} \\
&\leq \frac{E_0}{M_0\hat{\rho}_0^2}\left(\frac{E_0}{M_0\hat{\rho}_0^2}\right)^{j-1}\left(\frac{4(12)^{2(\tau+1)}}{8(12)^{2(\tau+1)}}\right)^{j} \\
&\leq \left(\frac{E_0}{M_0\hat{\rho}_0^2}\right)^{j}\frac{1}{2^j}
\end{aligned}$$

hence

$$\sup_{(D_j)_{\hat{\rho}_j/4}} \|\partial\omega_{j+1}\| \leq M_j + \frac{1}{2^j}M_j =: M_{j+1} \tag{4.35}$$

And similarly for $\|\hat{\omega}_{j+1}\|$, we have

$$\begin{aligned}
\sup_{(D_j)_{\hat{\rho}_j/4}} \|\partial\hat{\omega}_{j+1}\| &= \sup_{(D_j)_{\hat{\rho}_j/4}} \|\partial^2\hat{h}_j + \partial^2\hat{g}_{j+1}\| \\
&\leq \hat{M}_j + 2^6\frac{13}{11}\frac{E_j}{\hat{\rho}_j^2} \\
&\leq \hat{M}_j + \frac{\hat{M}_j}{2^j} =: \hat{M}_{j+1}
\end{aligned} \tag{4.36}$$

To estimate $\|(\partial\omega_{j+1})^{-1}\|$, first we compute

$$\begin{aligned}\sup_{(D_j)_{\hat\rho_j/4}} \|\partial^2 g_{j+1}(\partial^2 h_j)^{-1}\| &\le \sup\|\partial^2 g_{j+1}\| \sup\|(\partial^2 h_j)^{-1}\| \\ &\le 2^6 \frac{13}{11}\frac{E_j}{\hat\rho_j^2}\bar M_j \\ &\le 2^6 \frac{13}{11}\frac{E_0}{\hat\rho_0^2}\frac{\bar M_j}{2^j}\end{aligned}$$

if we have $\bar M_j = 2\bar M_{j-1}$, we get

$$\sup_{(D_j)_{\hat\rho_j/4}} \|\partial^2 g_{j+1}(\partial^2 h_j)^{-1}\| \le \frac{1}{2}$$

and this gives us

$$\sup_{(D_j)_{\hat\rho_j/4}} \|(\partial^2 h_j + \partial^2 g_{j+1})^{-1}\| \le 2\bar M_j =: \bar M_{j+1} \tag{4.37}$$

Now we define a new map

$$\hat\iota_{j+1} = (\hat\iota_{j+1}^{(1)}, \hat\iota_{j+1}^{(2)}) := \omega_{j+1}^{-1} \circ \omega_j|_{D_j} \tag{4.38}$$

let $I' \in D_j, I \in (D_j)_{\hat\rho_j/16}$ be such that $|I - I'| < \frac{\hat\rho_j}{16}$, then

$$\begin{aligned}|\omega_{j+1}(I) - \omega_{j+1}(I')| &= |\omega_j(I) + \partial g_{j+1}(I) - \omega_j(I') - \partial g_{j+1}(I')| \\ &= |\omega_j(I) - \omega_j(I') + \partial g_{j+1} \circ \omega_j^{-1}(\omega_j(I)) - \partial g_{j+1} \circ \omega^{-1}(\omega_j(I'))| \\ &= |\omega_j(I) - \omega_j(I') + \partial g_{j+1} \circ (\partial h_j)^{-1}(\omega_j(I)) - \partial g_{j+1} \circ (\partial h_j)^{-1}(\omega_j(I'))| \\ &\ge (1 - \|\partial^2 g_{j+1}(\partial^2 h_j)^{-1}\|)|\omega_j(I) - \omega_j(I')| \\ &\ge \frac{1}{2}\frac{1}{\bar M_j}|I - I'| = \frac{r}{2\bar M_j}\end{aligned}$$

this implies that

$$\omega_{j+1}\left((D_j)_{\hat\rho_j/16}\right) \supseteq (\omega_{j+1}(D_j))_{\hat\rho_j/(32\bar M_j)} \tag{4.39}$$

and we also have

$$\begin{aligned}\sup_{D_j} |\omega_j - \omega_{j+1}|_1 = \sup_{D_j} |\partial g_{j+1}|_1 &\le n\frac{13/11 E_j}{\hat\rho_j/2} \\ &\le n\frac{26}{11}\frac{E_j \bar M_j}{\hat\rho_j^2}\frac{\hat\rho_j}{\bar M_j} \\ &\le n\frac{26}{11}\frac{E_0 \bar M_0}{\hat\rho_0^2}\frac{\hat\rho_j}{\bar M_j} \\ &\le \frac{1}{32}\frac{\hat\rho_j}{\bar M_j}\end{aligned}$$

hence

$$\omega_j(D_j) \subset (\omega_{j+1}(D_j))_{\hat{\rho}_j/(32\bar{M}_j)} \subseteq \omega_{j+1}\left((D_j)_{\hat{\rho}_j/16}\right) \tag{4.40}$$

We explore the properties of the map $\hat{\iota}_{j+1}$, first

$$\begin{aligned}
\sup_{D_j} |\hat{\iota}^{(1)}_{j+1}(I_j) - I^{(1)}_j| &= \sup_{D_j} |\hat{\iota}^{(1)}_{j+1}(I_j) - I^{(1)}_j|_1 \\
&= \sup_{D_j} |(\omega^{-1}_{j+1})^{(1)} \circ \omega_j(I_j) - (\omega^{-1}_{j+1})^{(1)} \circ \omega_{j+1}(I_j)|_1 \\
&\leq \bar{M}^{(1)}_{j+1} \sup_{D_j} |\omega_j(I_j) - \omega_{j+1}(I_j)| = \bar{M}^{(1)}_{j+1} \sup_{D_j} |\partial g_{j+1}|_1 \\
&\leq \bar{M}^{(1)}_{j+1} n \frac{13/11}{\hat{\rho}_j/2} E_j = \frac{26n}{11} \frac{\bar{M}^{(1)}_{j+1} E_j}{\hat{\rho}_j} \\
&= \frac{52n}{11} \frac{\bar{M}^{(1)}_j E_j}{\hat{\rho}_j}
\end{aligned} \tag{4.41}$$

similarly

$$\sup_{D_j} |\hat{\iota}^{(2)}_{j+1}(I_j) - I^{(2)}_j| = \sup_{D_j} |\hat{\iota}^{(1)}_{j+1}(I_j) - I^{(2)}_j|_1 \leq \frac{52n}{11} \frac{\bar{M}^{(2)}_j E_j}{\hat{\rho}_j} \tag{4.42}$$

and

$$\begin{aligned}
\sup_{(D_j)_{\hat{\rho}_j/8}} |\hat{\iota}_{j+1}(I) - I|_\infty &\leq \bar{M}_{j+1} \sup_{(D_j)_{\hat{\rho}_j/8}} |\partial g_{j+1}|_\infty \\
&\leq 2\bar{M}_j \frac{\sup_{(D_j)_{\hat{\rho}_j/2}} |g_{j+1}|}{3\hat{\rho}_j/8} \\
&\leq 2\bar{M}_j \frac{13/11 E_j}{3\hat{\rho}_j/8} \\
&\leq \frac{208}{33} \frac{\bar{M}_j E_j}{\hat{\rho}_j}
\end{aligned} \tag{4.43}$$

With this and Cauchy estimate we have

$$\begin{aligned}
\mathcal{L}(\hat{\iota}_{j+1} - id) &\leq n \inf_{0<r<\hat{\rho}_j/8} \sup_{(D_j)_r} \|D(\hat{\iota}_{j+1} - id)\| \\
&\leq n \inf_{(0,\hat{\rho}_j/8)} \frac{\sup_{(D_j)_{\hat{\rho}_j/8}} |\hat{\iota}_{j+1}(I) - I|_\infty}{\hat{\rho}_j/8 - r} \\
&\leq \frac{8n}{\hat{\rho}_j} \frac{208}{33} \frac{\bar{M}_j E_j}{\hat{\rho}_j} \leq \frac{1664n}{33} \frac{\bar{M}_j E_j}{\hat{\rho}^2_j} \\
&\leq 2^6 n \frac{\bar{M}_j E_j}{\hat{\rho}^2_j}
\end{aligned} \tag{4.44}$$

Note by our definition, $\hat{\iota}_{j+1}(D_j) = D_{j+1}$, i.e., $\omega_j(D_j) = \omega_{j+1}(D_{j+1})$, this indicates

$$D_{j+1} \subseteq (D_j)_{\hat{\rho}_j/16} \tag{4.45}$$

and moreover

$$\begin{aligned} &(D_{j+1})_{\hat{\rho}_{j+1}} \subset (D_j)_{\hat{\rho}_j/8} \\ \Rightarrow &\left(\frac{D_{j+1}}{\hat{\rho}_0}\right)_{\hat{\rho}_{j+1}/\hat{\rho}_0} \subseteq \left(\frac{D_j}{\hat{\rho}_0}\right)_{\hat{\rho}_j/(8\hat{\rho}_0)} \subseteq \left(\frac{D_j}{\hat{\rho}_0}\right)_{\hat{\rho}_j/(2\hat{\rho}_0)} \end{aligned} \tag{4.46}$$

On $(\frac{D_{j+1}}{\hat{\rho}_0})_{\hat{\rho}_{j+1}/\hat{\rho}_0} \times (\frac{\mathbb{T}^n}{s_0})_{(\bar{s}+s_{j+1})/s_0}$, we define a scaled map

$$\check{\Psi}_{j+1} - id := 1_{\hat{\rho}_0,s_0}\Psi_j \circ 1^{-1}_{\hat{\rho}_0,s_0} - id \tag{4.47}$$

and by applying the Cauchy estimate,

$$\begin{aligned} \|D(\check{\Psi}_{j+1} - id)\|_{\left(\frac{D_{j+1}}{\hat{\rho}_0}\right)_{\hat{\rho}_{j+1}/\hat{\rho}_0} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_{j+1})/s_0}} &\leq \|D(\check{\Psi}_{j+1} - id)\|_{\left(\frac{D_j}{\hat{\rho}_0}\right)_{\hat{\rho}_j/(8\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_j/12)/s_0}} \\ &\leq \frac{1}{\min\{\frac{3\hat{\rho}_j}{8\hat{\rho}_0}, \frac{s_j}{12s_0}\}} \|(\check{\Psi}_{j+1} - id)\|_{\left(\frac{D_j}{\hat{\rho}_0}\right)_{\hat{\rho}_j/(2\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_j/6)/s_0}} \end{aligned}$$

let $A := \left(\frac{D_j}{\hat{\rho}_0}\right)_{\hat{\rho}_j/(2\hat{\rho}_0)} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_j/6)/s_0}$, this can be evaluated because

$$\begin{aligned} \|(\check{\Psi}_{j+1} - id)\|_A &\leq \max\{\frac{1}{\hat{\rho}_0}\|I_{j+1} - I(I_{j+1}, \varphi_{j+1})\|_A, \frac{1}{s_0}\|\varphi_{j+1} - \varphi(I_{j+1}, \varphi_{j+1})\|_A\} \\ &\leq \max\{\frac{1}{\hat{\rho}_0}\frac{3E_j}{2M_j\hat{\rho}_j}, \frac{1}{\hat{\rho}_0}\frac{3E_j}{2\hat{M}_j\hat{\rho}_j}, \frac{1}{s_0}\frac{3E_js_j}{2\hat{M}_j\hat{\rho}_j^2}\} \end{aligned}$$

and then

$$\begin{aligned} \|D(\check{\Psi}_{j+1} - id)\|_{\left(\frac{D_{j+1}}{\hat{\rho}_0}\right)_{\hat{\rho}_{j+1}/\hat{\rho}_0} \times \left(\frac{\mathbb{T}^n}{s_0}\right)_{(\bar{s}+s_{j+1})/s_0}} &\leq \frac{\max\{\frac{1}{\hat{\rho}_0}\frac{3E_j}{2M_j\hat{\rho}_j}, \frac{1}{\hat{\rho}_0}\frac{3E_j}{2\hat{M}_j\hat{\rho}_j}, \frac{1}{s_0}\frac{3E_js_j}{2\hat{M}_j\hat{\rho}_j^2}\}}{\min\{\frac{3\hat{\rho}_j}{8\hat{\rho}_0}, \frac{s_j}{12s_0}\}} \\ \leq \frac{\max\{\frac{1}{\hat{\rho}_0}\frac{3E_j}{2M_j\hat{\rho}_j}, \frac{1}{\hat{\rho}_0}\frac{3E_j}{2\hat{M}_j\hat{\rho}_j}, \frac{1}{s_0}\frac{3E_js_j}{2\hat{M}_j\hat{\rho}_j^2}\}}{\min\{\frac{3}{8}(\frac{1}{2(12)^{\tau+1}})^j, \frac{1}{12^{j+1}}\}} &\leq \frac{\max\{\frac{\hat{\rho}_j}{\hat{\rho}_0}\frac{3E_j}{2M_j\hat{\rho}_j^2}, \frac{\hat{\rho}_j}{\hat{\rho}_0}\frac{3E_j}{2\hat{M}_j\hat{\rho}_j^2}, \frac{s_j}{s_0}\frac{3E_j}{2\hat{M}_j\hat{\rho}_j^2}\}}{\min\{\frac{3}{8}(\frac{1}{2(12)^{\tau+1}})^j, \frac{1}{12^{j+1}}\}} \\ \leq \frac{3}{2}\frac{\max\{\frac{E_j}{M_j\hat{\rho}_j^2}, \frac{E_j}{\hat{M}_j\hat{\rho}_j^2}, \frac{E_j}{\hat{M}_j\hat{\rho}_j^2}\} \cdot \max\{\frac{1}{16^j}, \frac{1}{12^j}\}}{\min\{\frac{3}{8}(\frac{1}{2(12)^{\tau+1}})^j, \frac{1}{12^{j+1}}\}} & \\ \leq 4\max\{\frac{E_j}{M_j\hat{\rho}_j^2}, \frac{E_j}{\hat{M}_j\hat{\rho}_j^2}\}\left(\frac{12^{\tau+1}}{6}\right)^j & \end{aligned}$$

And we have the following Lipschitz estimate

$$\begin{aligned}\mathcal{L}(\check{\Psi}_{j+1} - id) &\leq 2n \sup \|D(\check{\Psi}_{j+1} - id)\|_\infty \\ &\leq 8n \max\{\frac{E_j}{M_j \hat{\rho}_j^2}, \frac{E_j}{\hat{M}_j \hat{\rho}_j^2}\} \left(\frac{12^{\tau+1}}{6}\right)^j \end{aligned} \tag{4.48}$$

□

Now we prove theorem 1

4.2 Proof of Theorem 3

Proof. **Step 1.** Construction of the "limit actions"

Recall our definition of $\hat{\iota}_j$, we have that, for $j \geq 0$,

$$\hat{\iota}_{j+1} = (\hat{\iota}^{(1)}_{j+1}, \hat{\iota}^{(2)}_{j+1}) := \omega^{-1}_{j+1} \circ \omega_j|_{D_j} : D_j \to D_{j+1} \tag{4.49}$$

and on $D_0 = \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap D$, we define

$$\check{\iota}_0 := id, \quad \check{\iota}_j := \hat{\iota}_j \circ \hat{\iota}_{j-1} \circ \cdots \circ \hat{\iota}_1, \quad j \geq 1 \tag{4.50}$$

We first prove the uniform convergence of $\check{\iota}$, let $i > j$, we have

$$\begin{aligned}
\sup_{D_0} |\check{\iota}_i - \check{\iota}_j|_1 &\leq \sum_{l=j}^{i-1} \sup_{D_0} |\check{\iota}_{l+1} - \check{\iota}_l|_1 \leq \sum_{l=j}^{i-1} \sup_{D_0} |\hat{\iota}_{l+1} \circ \check{\iota}_l - id \circ \check{\iota}_l|_1 \leq \sum_{l=j}^{i-1} \sup_{D_0} |(\hat{\iota}_{l+1} - id) \circ \check{\iota}_l|_1 \\
&\leq \sum_{l=j}^{i-1} \sup_{D_l} |\hat{\iota}_{l+1} - id|_1 \leq \sum_{l=j}^{i-1} \sup_{(D_l)_{\hat{\rho}_l/8}} |\hat{\iota}_{l+1} - id|_1 \leq \sum_{l=j}^{i-1} n \sup_{(D_l)_{\hat{\rho}_l/8}} |\hat{\iota}_{l+1} - id|_\infty \\
&\leq \sum_{l=j}^{i-1} n \frac{208}{33} \frac{\bar{M}_l E_l}{\hat{\rho}_l}
\end{aligned}$$

since the right-hand side can be very small, the uniform convergence is verified. If we let $i \to \infty$, the following estimates will be verified

$$\begin{aligned}
\sup_{D_0} |\check{\iota} - \check{\iota}_j| &\leq \sum_{l=j}^{\infty} n \frac{208}{33} \frac{\bar{M}_l E_l}{\hat{\rho}_l} \\
&\leq \sum_{l=j}^{\infty} n \frac{208}{33} \bar{M}_j 2^{l-j} \frac{1}{(8(12)^{2(\tau+1)})^{l-j}} (\frac{E_0}{M_0 \hat{\rho}_0^2})^{l-j} E_j \frac{(2(12)^{\tau+1})^{l-j}}{\hat{\rho}_j} \\
&\leq \sum_{l=j}^{\infty} n \frac{208}{33} \frac{1}{(2(12)^{(\tau+1)})^{l-j}} (\frac{1}{2^8 n})^{l-j} \frac{E_j \bar{M}_j}{\hat{\rho}_j} \\
&\leq \sum_{l=j}^{\infty} n \frac{208}{33} \frac{1}{(2(12)^{(\tau+1)})^{l-j}} \frac{1}{(2^8 n)^{l-j}} \frac{1}{2^8 n} \hat{\rho}_j \\
&\leq \frac{\hat{\rho}_j}{2^5} \sum_{l=j}^{\infty} \frac{1}{(2(12)^{(\tau+1)})^{l-j}} \frac{1}{(2^8 n)^{l-j}} \leq \frac{\hat{\rho}_j}{2^4} \leq \rho_j
\end{aligned}$$

We define

$$\check{\iota} := \lim_j \check{\iota}_j, \quad D_* := \check{\iota}(D_0) \tag{4.51}$$

and the following is alread proved

$$D_* \subseteq \bigcap_j (\check{\iota}_j(D_0))_{\rho_j} = \bigcap_j (D_j)_{\rho_j} \tag{4.52}$$

and if we let j to be equal to zero, we have that

$$\begin{aligned}\sup_{D_0} |\check{\iota} - id|_1 &\leq n\frac{208}{33}\frac{E_0\bar{M}_0}{\hat{\rho}_0}\sum_{l=0}^{\infty}\frac{1}{(2(12)^{(\tau+1)})^l}\frac{1}{(2^8 n)^l} \\ &\leq 8n\frac{E_0\bar{M}_0}{\hat{\rho}_0}\end{aligned} \tag{4.53}$$

which implies

$$D_* \subset (D_0)_{8nE_0\bar{M}_0/\hat{\rho}_0} \tag{4.54}$$

With the similar method and equation (3.39) the following estimates can also be verified

$$\begin{aligned}\sup_{D_0} |\check{\iota}_i^{(1)} - \check{\iota}_j^{(1)}|_1 &\leq \sum_{l=j}^{i-1} \sup_{D_l} |\hat{\iota}_{l+1}^{(1)} - id^{(1)}| \\ &\leq \sum_{l=j}^{i-1} \frac{52n}{11}\frac{\bar{M}_l^{(1)}E_l}{\hat{\rho}_l}\end{aligned}$$

by letting $i \to \infty$, and $j = 0$, we get

$$\sup_{D_0} |\check{\iota}^{(1)} - id^{(1)}|_1 \leq 5n\frac{\bar{M}_0^{(1)}E_0}{\hat{\rho}_0} \tag{4.55}$$

and we can also prove

$$\sup_{D_0} |\check{\iota}^{(2)} - id^{(2)}|_1 \leq 5n\frac{\bar{M}_0^{(2)}E_0}{\hat{\rho}_0} \tag{4.56}$$

Let $\mu_j = 2^6 n \frac{\bar{M}_j E_j}{\hat{\rho}_j^2}$, we have

$$\begin{aligned}
\mathcal{L}(\breve{\imath}_{j+1} - id) &\leq \prod_{j=1}^{\infty}(1+\mu_{j-1}) - 1 \\
&= \exp\left[\sum_{j=0}^{\infty} \log(1+\mu_j)\right] - 1 \\
&\leq \exp\left[\sum_{j=0}^{\infty} \mu_j\right] - 1
\end{aligned}$$

where

$$\sum_{j=0}^{\infty} \mu_j = \sum_{j=0}^{\infty} 2^6 n \frac{\bar{M}_j E_j}{\hat{\rho}_j^2} \leq 2^7 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2}$$

hence

$$\mathcal{L}(\breve{\imath}_{j+1} - id) \leq \exp\left[2^7 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2}\right] - 1 \leq 2^8 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2} \tag{4.57}$$

and let $j \to \infty$, we have

$$\mathcal{L}(\breve{\imath} - id) \leq \limsup_{j} \mathcal{L}(\breve{\imath}_{j+1} - id) \leq 2^8 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2} \tag{4.58}$$

hence on D_0, the lower and upper Lipschitz constants are

$$\mathcal{L}_-(\breve{\imath}) \geq 1 - 2^8 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2}, \quad \mathcal{L}_+(\breve{\imath}) \leq 1 + 2^8 n \frac{\bar{M}_0 E_0}{\hat{\rho}_0^2} \tag{4.59}$$

Step 2.<u>Construction of ϕ_ω</u>

On $W_{j-1} = (D_{j-1})_{\rho_{j-1}} \times \mathbb{T}^n_{\bar{s}+s_{j-1}}$,

$$\Phi_j := \Psi_1 \circ \cdots \circ \Psi_j \tag{4.60}$$

It is clear that $\{\Phi\}$ converges uniformly on $W_* := D_* \times \mathbb{T}^n \subset \cap_j W_j$, and we call its limit Φ, i.e. $\Phi := \lim_{j\to\infty} \Phi_j$ and define

$$\phi_\omega(\theta) = ((v_1(\theta;\omega), v_2(\theta;\omega)), \theta + u(\theta;\omega)) := \Phi(\breve{\imath}(\omega_0^{-1}(\omega)), \theta) \tag{4.61}$$

note here we must make sure $\omega \in \mathcal{D}_{\gamma_1,\gamma_2,\tau} \cap \omega_0(D)$. And on D_*

$$
\begin{aligned}
|\Pi_1\Phi - id|_1 &\leq \lim_j \sum_{1\leq k\leq j} \sup_{W_k} |\Psi_k^{(1)} - id|_1 \leq \lim_j \sum_{1\leq k\leq j} \frac{3}{2}\frac{E_{k-1}}{M_{k-1}\hat{\rho}_{k-1}} \\
&\leq \frac{3}{2}\sum_{k=0}^{\infty}\frac{E_k}{M_k\hat{\rho}_k} \leq \frac{3}{2}\sum_{k=0}^{\infty}\frac{1}{M_0}\frac{(2(12)^{\tau+1})^k}{(8(12)^{2(\tau+1)})^k}\left(\frac{E_0}{M_0\hat{\rho}_0^2}\right)^k\frac{E_0}{\hat{\rho}_0} \\
&\leq 2\frac{E_0}{M_0\hat{\rho}_0}
\end{aligned}
\tag{4.62}
$$

and, similarly,

$$
\begin{aligned}
|\Pi_2\Phi - id|_1 &\leq \lim_j \sum_{1\leq k\leq j} \sup_{W_k} |\Psi_k^{(2)} - id|_1 \\
&\leq \lim_j \sum_{1\leq k\leq j} \frac{3}{2}\frac{E_{k-1}}{\hat{M}_{k-1}\hat{\rho}_{k-1}} \\
&\leq 2\frac{E_0}{\hat{M}_0\hat{\rho}_0}
\end{aligned}
\tag{4.63}
$$

and

$$
\begin{aligned}
|\Pi_\psi\Phi - id|_1 &\leq \lim_j \sum_{1\leq k\leq j} \sup_{W_k} |\Psi_k^{(\psi)} - id|_1 \leq \lim_j \sum_{1\leq k\leq j} \frac{3}{2}\frac{E_{k-1}s_{j-1}}{\hat{M}_{k-1}\hat{\rho}_{k-1}^2} \\
&\leq \frac{3}{2}\sum_{k=0}^{\infty}\frac{s_0}{\hat{M}_0}\frac{(4(12)^{2(\tau+1)})^k}{(8(12)^{2(\tau+1)})^k}\left(\frac{E_0}{M_0\hat{\rho}_0^2}\right)^k\frac{1}{12^k}\frac{E_0}{\hat{\rho}_0^2} \\
&\leq 2\frac{E_0 s_0}{M_0\hat{\rho}_0^2}
\end{aligned}
\tag{4.64}
$$

Let $\Omega_* := \omega_0(D) \cap \mathcal{D}_{\gamma_1,\gamma_2,\tau}$, $\omega \in \Omega_*$, and $I^0(\omega) = \omega_0^{-1}(\omega) = (I_1^0(\omega), I_2^0(\omega))$, then we have

$$
\begin{aligned}
|v_1(\cdot,\omega) - I_1^0(\omega)| &= \sup_{\Omega_*} |\Pi_1\Phi(\breve{\imath}(\omega_0^{-1}(\omega)),\theta) - I_1^0(\omega)| \\
&= \sup_{\Omega_*} |\Pi_1\Phi(\breve{\imath}(\omega_0^{-1}(\omega)),\theta) - \Pi_1\breve{\imath}(I^0(\omega)) + \Pi_1\breve{\imath}(I^0(\omega)) - I_1^0(\omega)| \\
&\leq \sup_{\Omega_*} |\Pi_1\Phi(\breve{\imath}(\omega_0^{-1}(\omega)),\theta) - \Pi_1\breve{\imath}(I^0(\omega))| + \sup_{\Omega_*} |\Pi_1\breve{\imath}(I^0(\omega)) - I_1^0(\omega)| \\
&\leq \sup_{D_*} |\Pi_1\Phi - id| + \sup_{D_0} |\Pi_1\breve{\imath} - id| \\
&\leq 2\frac{E_0}{M_0\hat{\rho}_0} + 5n\frac{\bar{M}_0^{(1)}E_0}{\hat{\rho}_0}
\end{aligned}
\tag{4.65}
$$

with the same computation we have

$$|v_2(\cdot,\omega) - I_2^0(\omega)| \le 2\frac{E_0}{\hat{M}_0\hat{\rho}_0} + 5n\frac{\bar{M}_0^{(2)}E_0}{\hat{\rho}_0} \tag{4.66}$$

$$|u(\cdot,\omega)| = |\theta + u(\theta;\omega) - \theta| = \sup_{W_*} |\Pi_\psi \Phi(\check{\imath}(I^0(\omega)),\theta) - \theta| \le 2\frac{E_0 s_0}{M_0\hat{\rho}_0^2} \tag{4.67}$$

From the above estimates, we have

$$T_\omega = \phi_\omega(\mathbb{T}^n) \subset (D_*)_{2\frac{E_0}{M_0\hat{\rho}_0}+2\frac{E_0}{\hat{M}_0\hat{\rho}_0}} \times \mathbb{T}^n \subset D_r \times \mathbb{T}^n$$

where $r = 2\frac{E_0}{M_0\hat{\rho}_0} + 2\frac{E_0}{\hat{M}_0\hat{\rho}_0} + 8n\frac{\bar{M}_0E_0}{\hat{\rho}_0}$.

Similar with $\check{\imath}$, we need to verify the Lipschitz constant of the rescaled map of Φ. First we define the following rescaled domain:

$$\check{D}_* \times \check{\mathbb{T}}^n_{\bar{s}/s_0} := \hat{\rho}_0^{-1} D_* \times (\mathbb{R}/(2\pi/s_0)\mathbb{Z})_{\bar{s}/s_0} \tag{4.68}$$

and the rescaled map is defined as

$$\begin{aligned} \check{\Phi} &= 1_{\hat{\rho}_0,s_0}\Phi \circ 1^{-1}_{\hat{\rho}_0,s_0} \\ &= \lim_j 1_{\hat{\rho}_0,s_0} \circ \Phi_j \circ 1^{-1}_{\hat{\rho}_0,s_0} \\ &= \lim_j \check{\Psi}_1 \circ \cdots \circ \check{\Psi}_j \end{aligned} \tag{4.69}$$

To estimate its Lipschitz constant, we define

$$\begin{aligned} i_j &:= \check{\Phi}_j - id = \check{\Psi}_1 \circ \cdots \circ \check{\Psi}_j - id \\ &= \check{\Phi}_{j-1} \circ \check{\Psi}_j - id - \check{\Psi}_j + \check{\Psi}_j \\ &= (\check{\Phi}_{j-1} - id) \circ \check{\Psi}_j + \check{\Psi}_j - id \\ &= i_{j-1} \circ \check{\Psi}_j + \check{\Psi}_j - id \end{aligned}$$

and its Lipschitz constant satisfies

$$\begin{aligned} \mathcal{L}(i_j) &= \mathcal{L}(i_{j-1})\mathcal{L}(\check{\Psi}_j) + \mathcal{L}(\check{\Psi}_j - id) \\ &= \mathcal{L}(i_{j-1})\mathcal{L}(\check{\Psi}_j - id + id) + \mathcal{L}(\check{\Psi}_j - id) \\ &\le \mathcal{L}(i_{j-1})\big(\mathcal{L}(\check{\Psi}_j - id) + 1\big) + \mathcal{L}(\check{\Psi}_j - id) \\ &\le \mathcal{L}(i_{j-2})\big(\mathcal{L}(\check{\Psi}_{j-1} - id) + 1\big)\big(\mathcal{L}(\check{\Psi}_j - id) + 1\big) + \mathcal{L}(\check{\Psi}_{j-1} - id)\big(\mathcal{L}(\check{\Psi}_j - id) + 1\big) + \mathcal{L}(\check{\Psi}_j - id) \\ &\ddots \\ &\le \prod_{k=1}^{j}\big(\mathcal{L}(\check{\Psi}_k - id) + 1\big) - 1 \end{aligned}$$

therefore

$$\begin{aligned}
\mathcal{L}(\check{\Phi}-id) &= \limsup_j \mathcal{L}(\check{\Phi}_j - id) \le \lim_j \prod_{k=1}^{j}\left(\mathcal{L}(\check{\Psi}_k - id)+1\right)-1 \\
&= \prod_{k=1}^{\infty}\left(\mathcal{L}(\check{\Psi}_k - id)+1\right)-1 \\
&= \exp\left[\sum_{k=1}^{\infty}\log\left(\mathcal{L}(\check{\Psi}_k - id)+1\right)\right]-1 \le \exp\left[\sum_{k=1}^{\infty}\mathcal{L}(\check{\Psi}_k - id)\right]-1 \\
&\le \exp\left[\sum_{k=1}^{\infty} 8n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}\left(\frac{E_0}{M_0\hat{\rho}_0^2}12^{\tau}\right)^{k-1}\right]-1
\end{aligned}$$

note the exponent part is

$$\begin{aligned}
\sum_{k=1}^{\infty} 8n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}\left(\frac{E_0}{M_0\hat{\rho}_0^2}12^{\tau}\right)^{k-1} &\le 8n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}\frac{1}{1-\frac{E_0}{M_0\hat{\rho}_0^2}12^{\tau}} \\
&\le 16n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}
\end{aligned}$$

from which we have

$$\begin{aligned}
\mathcal{L}(\check{\Phi}-id) &\le \exp\left[16n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\}\right]-1 \\
&\le 32n\max\{\frac{E_0}{M_0\hat{\rho}_0^2},\frac{E_0}{\hat{M}_0\hat{\rho}_0^2}\} \qquad (4.70)
\end{aligned}$$

Hence $\check{\Phi}, \Phi$, and the map $(\theta,\omega)\to\phi_\omega(\theta)$ are bi-Lipschitz and injective.

Step 3 For any $\omega\in\mathcal{D}_{\gamma_1,\gamma_2,\tau}\cap\omega_0(D), T_\omega := \phi_\omega(\mathbb{T}^n)$ is a Lagrangian H-invariant torus with frequency ω

Let $I_*\in D_*$ be an independent variable, and $\phi_t(I_0,\psi_0)$ denote the H-flow starting at (I_0,ψ_0), we want to prove

$$\phi_t(\Phi(I_*,\theta)) = \Phi(I_*,\theta+\omega_*(I_*)t) \qquad (4.71)$$

where $\omega_*(I_*) := \omega_0(\check{\iota}^{-1}(I_*))$.

We try to estimate the bound of the following

$$\begin{aligned}|\phi_t(\Phi(I_*,\theta)) - \Phi(I_*,\theta+\omega_*(I_*)t)| &:= |\phi_t(\Phi(I_*,\theta)) - \Phi(I_*,\theta+\omega_*(I_*)t)|_\infty \\ &\le |\phi_t(\Phi(I_*,\theta)) - \phi_t(\Phi_j(I_*,\theta))| + \\ &+ |\phi_t(\Phi_j(I_*,\theta)) - \Phi_j(I_*,\theta+\omega_*(I_*)t)| + \\ &+ |\Phi_j(I_*,\theta+\omega_*(I_*)t) - \Phi(I_*,\theta+\omega_*(I_*)t)|\end{aligned}$$

by letting $j \to \infty$, due to the uniform convergence of Φ_j to Φ on W_* , we have that as $j \to \infty$

$$\begin{aligned}|\phi_t(\Phi(I_*,\theta)) - \phi_t(\Phi_j(I_*,\theta))| \to 0 \\ |\Phi_j(I_*,\theta+\omega_*(I_*)t) - \Phi(I_*,\theta+\omega_*(I_*)t)| \to 0\end{aligned}$$

Now we have to prove that as $j \to 0$

$$|\phi_t(\Phi_j(I_*,\theta)) - \Phi_j(I_*,\theta+\omega_*(I_*)t)| \to 0$$

since the transformation (composition of canonical transformations) Φ_j is canonical on W_j, let $\phi_t^j(I_*,\theta)$ be the flow of $H_j = H \circ \Phi_j = h_j + f_j$ starting from (I_*,θ), we have

$$\phi_t(\Phi_j(I_*,\theta)) = \Phi_j(\phi_t^j(I_*,\theta))$$

hence we are reduced to prove that as $j \to \infty$,

$$|\Phi_j(\phi_t^j(I_*,\theta)) - \Phi_j(I_*,\theta+\omega_*(I_*)t)| \to 0$$

that is

$$\phi_t^j(I_*,\theta) - (I_*,\theta+\omega_*(I_*)t) \to 0$$

Indeed, let $(I_j(t),\psi_j(t)) = \phi_t^j(I_*,\theta)$ be the H_j-evolution of (I_*,θ), then from the Hamiltonian equation

$$\begin{cases}\dot{I}_j(t) &= -\partial_\psi H_j = -\partial_\psi f_j(I_j(t),\psi_j(t)) \\ \dot{\psi}_j(t) &= \partial_I H_j = \partial_I h_j(I_j(t)) + \partial_I f_j(I_j(t),\psi_j(t))\end{cases}$$

we obtain

$$\begin{aligned}I_j(t) &= I_* - \int_0^t \partial_\psi f_j(I_j(\tau),\psi_j(\tau))d\tau \\ \psi_j(t) &= \theta + \int_0^t \partial_I h_j(I_j(\tau)) + \partial_I f_j(I_j(\tau),\psi_j(\tau))d\tau\end{aligned}$$

Therefore

$$
\begin{aligned}
|I_j(t) - I_*| &= \left| \int_0^t \partial_\psi f_j(I_j(\tau), \psi_j(\tau)) d\tau \right| \\
&\leq \int_0^t |\partial_\psi f_j(I_j(\tau), \psi_j(\tau))| \, d\tau \\
&\leq t \|f_j\| \frac{1}{s_{j-1}/12} \\
&\leq t \frac{1}{8(12)^{2(\tau+1)}} \frac{E_0}{M_0 \hat{\rho}_0^2} E_{j-1} \frac{1}{s_{j-1}/12} \\
&\leq t \left(\frac{1}{8(12)^{2(\tau+1)}} \right)^j \left(\frac{E_0}{M_0 \hat{\rho}_0^2} \right)^j E_0 \frac{1}{s_0/12^{j+1}} \\
&\leq \left(\frac{12}{8(12)^{2(\tau+1)}} \right)^j \left(\frac{E_0}{M_0 \hat{\rho}_0^2} \right)^j \frac{12tE_0}{s_0} \to 0 \text{ as } j \to \infty
\end{aligned}
$$

which shows that $|I_j(t) - I_*| \to 0$, i.e., $I_j(t) \to I_*$. And

$$
\begin{aligned}
|\psi_j(t) - \theta - \omega_*(I_*)t| &= \left| \int_0^t \partial_I h_j(I_j(\tau)) + \partial_I f_j(I_j(\tau), \psi_j(\tau)) - \omega_*(I_*) d\tau \right| \\
&\leq \int_0^t |\partial_I h_j(I_j(\tau)) + \partial_I f_j(I_j(\tau), \psi_j(\tau)) - \omega_*(I_*)| d\tau \\
&\leq \int_0^t |\omega_j(I_j(\tau)) - \omega_j(I_*)| d\tau + \int_0^t |\omega_j(I_*) - \omega_*(I_*)| d\tau + \\
&+ \int_0^t |\partial_I f_j(I_j(\tau), \psi_j(\tau))| d\tau
\end{aligned}
$$

where we have

$$
\begin{aligned}
\int_0^t |\omega_j(I_j(\tau)) - \omega_j(I_*)| d\tau &\leq t |\omega_j(I_j(t)) - \omega_j(I_*)| \\
&\leq t \sup_{D_j} |\partial \omega_j| \sup |I_j(\tau) - I_*| \\
&\leq t M_j \sup |I_j(\tau) - I_*| \to 0 \text{ as } j \to \infty
\end{aligned}
$$

and

$$
\int_0^t |\omega_j(I_*) - \omega_*(I_*)| d\tau \leq t \sup |\omega_j(I_*) - \omega_*(I_*)| \to 0 \text{ as } j \to \infty
$$

since

$$\begin{aligned}\omega_*(I_*) &= \omega_0(\check{\iota}^{-1}(I_*)) \\ &= \lim_j \omega_0(\check{\iota}_j^{-1}(I_*)) \\ &= \lim_j \omega_0 \circ \omega_0^{-1} \circ \omega_j(I_*) \\ &= \lim_j \omega_j(I_*)\end{aligned}$$

and also as $j \to \infty$

$$\int_0^t |\partial_I f_j(I_j(\tau), \psi_j(\tau))| d\tau \le t|\partial_I f_j| \le \frac{t}{7\hat{\rho}_{j-1}/16}|f_j| \to 0$$

hence $|\psi_j(t) - \theta - \omega_*(I_*)t| \to 0$, i.e., $\psi_j(t) \to \theta + \omega_*(I_*)t$.

Step 4 Measure estimates

We want to prove

$$\text{meas}\,(Re(D_r) \times \mathbb{T}^n \backslash K) \le c_n\,(\text{meas}(D\backslash D_{\gamma_1,\gamma_2,\tau} \times \mathbb{T}^n) + \text{meas}((Re(D_r)\backslash D) \times \mathbb{T}^n)) \tag{4.72}$$

where $D_{\gamma_1,\gamma_2,\tau} = \omega_0^{-1}(\mathcal{D}_{\gamma_1,\gamma_2,\tau}) \cap D, c_n = 1 + \left(1 + 32n\frac{E}{\hat{\rho}^2}\max\{\frac{1}{M}, \frac{1}{\bar{M}}\}\right)^{2n}$ and $K = \bigcup_{\omega\in\Omega_*} T_\omega$ First we decompose the left-hand side

$$\text{meas}\,(Re(D_r) \times \mathbb{T}^n \backslash K) \le \text{meas}((Re(D_r)\backslash D) \times \mathbb{T}^n) + \text{meas}(D \times \mathbb{T}^n \backslash K)$$

Note the map

$$\check{\iota} - id : D_0 \to D_*$$

is Lipschitz on D_0 and verifies $\mathcal{L}(\check{\iota} - id) \le 2^8 n\frac{\bar{M}E}{\hat{\rho}^2}$. And we could extend this map to a Lipschitz function with the same Lipschitz constant on $(D_0)_{\rho_1}$, with $\rho_1 := 8nE\bar{M}/\hat{\rho}/(1 - 2^8 n\bar{M}E/\hat{\rho}^2)$. We denote this new map as $\check{\iota}_e - id$, clearly we have $\mathcal{L}_-(\check{\iota}_e) \ge 1 - 2^8 n\frac{\bar{M}E}{\hat{\rho}^2}$.

From the Lipschitz constant of the extended map $\check{\iota}_e$, we could conclude that $\check{\iota}_e$ sends a ball with radius ρ_1 centered at $I_0 \in D_0$ over a ball with radius $(1 - 2^8 n\frac{\bar{M}E}{\hat{\rho}^2})\rho_1 = 8n\frac{E\bar{M}}{\hat{\rho}}$, note $8n\frac{\bar{M}E}{\hat{\rho}} \ge \sup_{D_0} |\check{\iota} - id|_1$. Hence

$$D_0 \subset \check{\iota}_e((D_0)_{\rho_1})$$

Then

$$\begin{aligned}
\operatorname{meas}(D_0\backslash D_*) &= \operatorname{meas}(D_0\backslash \breve{\imath}(D_0)) \\
&= \operatorname{meas}(D_0\backslash \breve{\imath}_e(D_0)) \\
&\leq \operatorname{meas}\left(\breve{\imath}_e(Re(D_0)_{\rho_1})\backslash \breve{\imath}_e(D_0)\right) \\
&\leq \operatorname{meas}(\breve{\imath}_e(Re(D_0)_{\rho_1}\backslash D_0)) \\
&\leq \mathcal{L}(\breve{\imath}_e)^n \operatorname{meas}(Re(D_0)_{\rho_1}\backslash D_0) \\
&\leq \mathcal{L}(\breve{\imath})^n \operatorname{meas}(Re(D_{\rho_1})\backslash D_0) \\
&\leq \mathcal{L}(\breve{\imath})^n \left(\operatorname{meas}(Re(D_{\rho_1})\backslash D) + \operatorname{meas}(D\backslash D_0)\right)
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{meas}(D\backslash D_*) &\leq \operatorname{meas}(D\backslash D_0) + \operatorname{meas}(D_0\backslash D_*) \\
&\leq (1+\mathcal{L}(\breve{\imath})^n)\operatorname{meas}(D\backslash D_0) + \mathcal{L}(\breve{\imath})^n \operatorname{meas}(D_{\rho_1}\backslash D)
\end{aligned}$$

Note the previous discussion is related to the map $\breve{\imath}$, more precisely, related to

$$\breve{\imath},\ D_*,\ D_0,\ n,\ \mathcal{L}(\breve{\imath}),\ D,\ \rho_1$$

we can repeat the above argument with

$$\check{\Phi},\ \check{K} = \check{\Phi}(\check{D}_* \times \check{\mathbb{T}}^n),\ \check{D}_* \times \check{\mathbb{T}}^n,\ 2n,\ \mathcal{L}(\check{\Phi}),\ \check{D} \times \check{\mathbb{T}}^n,\ \check{\rho}_2$$

where

$$\check{\rho}_2 = \frac{2\frac{E}{\hat{\rho}^2}(\frac{1}{M}+\frac{1}{\hat{M}})}{1-32n\frac{E}{\hat{\rho}^2}\max\{\frac{1}{M},\frac{1}{\hat{M}}\}} \tag{4.73}$$

then

$$\begin{aligned}
\operatorname{meas}(\check{D}\times\check{\mathbb{T}}^n\backslash\check{K}) &\leq (1+\mathcal{L}(\check{\Phi})^{2n})\operatorname{meas}(\check{D}\times\check{\mathbb{T}}^n\backslash\check{D}_*\times\check{\mathbb{T}}^n) + \mathcal{L}(\check{\Phi})^{2n}\operatorname{meas}((\check{D}\times\check{\mathbb{T}}^n)_{\check{\rho}_2}\backslash\check{D}\times\check{\mathbb{T}}^n) \\
&\leq (1+\mathcal{L}(\check{\Phi})^{2n})\operatorname{meas}(\check{D}\backslash\check{D}_*\times\check{\mathbb{T}}^n) + \mathcal{L}(\check{\Phi})^{2n}\operatorname{meas}((\check{D})_{\check{\rho}_2}\backslash\check{D}\times\check{\mathbb{T}}^n) \\
&\leq (1+\mathcal{L}(\check{\Phi})^{2n})\operatorname{meas}(\check{D}\backslash\check{D}_0\times\check{\mathbb{T}}^n) + \mathcal{L}(\check{\Phi})^{2n}\operatorname{meas}((\check{D})_{\check{\rho}_2}\backslash\check{D}\times\check{\mathbb{T}}^n)
\end{aligned}$$

after rescaling all the sets

$$\begin{aligned}
\operatorname{meas}(D\times\mathbb{T}^n\backslash K) &\leq (1+\mathcal{L}(\check{\Phi})^{2n})\operatorname{meas}(D\backslash D_0\times\mathbb{T}^n) + \mathcal{L}(\check{\Phi})^{2n}\operatorname{meas}((D)_{\hat{\rho}_0\check{\rho}_2}\backslash D\times\mathbb{T}^n) \\
&\leq (1+\mathcal{L}(\check{\Phi})^{2n})\operatorname{meas}(D\backslash D_0\times\mathbb{T}^n) + \mathcal{L}(\check{\Phi})^{2n}\operatorname{meas}(D_r\backslash D\times\mathbb{T}^n)
\end{aligned}$$

note here we should choose $\check{\rho}_2$ to be such that $\hat{\rho}\check{\rho}_2 \leq r = 2\frac{E}{M\hat{\rho}} + 2\frac{E}{\hat{M}\hat{\rho}} + 8n\frac{\bar{M}E}{\hat{\rho}}$. And finally

$$\begin{aligned}\text{meas}\,(Re(D_r) \times \mathbb{T}^n \backslash K) &\leq \text{meas}((Re(D_r)\backslash D) \times \mathbb{T}^n) + \text{meas}(D \times \mathbb{T}^n \backslash K) \\ &\leq (1 + \mathcal{L}(\check{\Phi})^{2n})\,(\text{meas}(D\backslash D_0 \times \mathbb{T}^n) + \text{meas}((Re(D_r)\backslash D) \times \mathbb{T}^n))\end{aligned}$$

□

Bibliography

[1] V.I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. Russian Math. Surveys, 18(6):85–191, 1963.

[2] Luca Biasco, Luigi Chierchia, and Enrico Valdinoci. Elliptic two-dimensional invariant tori for the planetary three-body problem. Archive for rational mechanics and analysis, 170(2):91–135, 2003.

[3] Luigi Chierchia and Gabriella Pinzari. Properly-degenerate kam theory (following v. i. arnold). Discrete and Continuous Dynamical Systems - S, 3(4):545–578, 2010.

[4] L. Chierchia and G. Pinzari. Planetary Birkhoff normal forms. J. Mod. Dyn., 5(4):623–664, 2011.

[5] Helmut Hofer and Eduard Zehnder. Symplectic invariants and Hamiltonian dynamics. Birkhäuser, 2012.

[6] N. N. Nehorošev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. Uspehi Mat. Nauk, 32(6(198)):5–66, 287, 1977.

[7] N. N. Nehorošev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. II. Trudy Sem. Petrovsk., (5):5–50, 1979. In Topics in Modern Mathematics, Petrovskii Seminar, No. 5, edited by O. A. Oleinik (Consultant Bureau, New York, 1985).

[8] Gabriella Pinzari. On the Kolmogorov set for many-body problems. arXiv preprint arXiv:1309.7028, 2013.

[9] Jürgen Pöschel. Nekhoroshev estimates for quasi-convex hamiltonian systems. Math. Z, 213(2):187–216, 1993.

[10] A.N. Kolmogorov. On dynamical systems with an integral invariant on the torus. Dokl. Akad. Nauk SSSR (NS) 93: 763–766, 1953.

[11] A.N. Kolmogorov. On the preservation of conditionally periodic motions under a small change in Hamilton's function. Dokl. Akad. Nauk SSSR 98: 527–530, 1954.

[12] J.K. Moser. On invariant curves of area-preserving mappings of an annulus. Matematika 6(5): 51–68, 1962.