ASIAN J. MATH.
Vol. 23, No. 3, pp. 401–438, June 2019

003

THE GENERICITY OF ARNOLD DIFFUSION IN NEARLY INTEGRABLE HAMILTONIAN SYSTEMS*

CHONG-QING CHENG†

Abstract. In this paper, we prove that the net of transition chain is δ-dense for nearly integrable positive definite Hamiltonian systems with 3 degrees of freedom in the cusp-residual generic sense in C^r-topology, $r \geq 6$. The main ingredients of the proof existed in [CZ, C17a, C17b]. As an immediate consequence, Arnold diffusion exists among this class of Hamiltonian systems. The question of [C17c] is answered in Section 9 of the paper.

Key words. Dynamical instability, Arnold diffusion.

Mathematics Subject Classification. 37J40, 37D05.

1. Introduction. After he constructed the celebrated example of *a priori* unstable systems in [A64], Arnold raised the conjecture in [A66] on the dynamical instability of nearly integrable Hamiltonian

$$H(p,q) = h(p) + \epsilon P(p,q), \qquad (p,q) \in \mathbb{R}^n \times \mathbb{T}^n. \tag{1.1}$$

CONJECTURE. *The "general case" for a Hamiltonian system (1.1) with $n \geq 3$ is represented by the situation that for an arbitrary pair of neighborhoods of toruses $p = p'$, $p = p''$ in one component of the level set $h(p') = h(p'')$, there exists, for sufficiently small ϵ, an orbit intersecting both neighborhoods.*

The research on the conjecture has two stages: *a priori* unstable and *a priori* stable cases. After the problem in *a priori* unstable case was solved, one has to study how to cross double resonance. It was pointed out by Arnold in [A66] that *in order to take the final step in the proof of the above conjecture, it is necessary to examine the transition from single to double resonance.* Indeed, one was able to establish the existence of global transition chain (Theorem 5.1 in [C17b]) after the double resonance problem was solved there. A positive answer to the conjecture for smooth and positive definite Hamiltonian with $n = 3$ is an immediate consequence, see Section 9.

The main part of the paper is to prove the existence of the δ-dense transition chain (Theorem 2.1), a slightly stronger form of Theorem 5.1 in Section 5 of [C17b]. The main ingredients of the proof are included in [CZ, C17a, C17b].

To study the problem, one needs to specify what is the genericity. Mather used the cusp-residual genericity [M04], we follow him.

DEFINITION 1.1. Let $B_D = \{p \in \mathbb{R}^3 : \|p\| \leq D\}$. Let $\mathfrak{S}_a, \mathfrak{B}_a \subset C^r(B_D \times \mathbb{T}^3, \mathbb{R})$ denote the sphere and the ball about the origin of radius $a > 0$ respectively: $F \in \mathfrak{S}_a$ if and only $\|F\|_{C^r} = a$ and $F \in \mathfrak{B}_a$ if and only $\|F\|_{C^r} \leq a$. They inherit the topology from $C^r(B_D \times \mathbb{T}^3, \mathbb{R})$.

Let $\mathfrak{R}_a$ be a set residual in $\mathfrak{S}_a$, each $P \in \mathfrak{R}_a$ is associated with a set R_P residual in the interval $[0, a_P]$ with $0 < a_P \leq a$. A set $\mathfrak{C}_a$ is said to be *cusp-residual* in $\mathfrak{B}_a$ if

$$\mathfrak{C}_a = \{\lambda P : P \in \mathfrak{R}_a, \lambda \in R_P\}.$$

*Received September 1, 2017; accepted for publication December 8, 2017.
†Department of Mathematics, Nanjing Univerisity, Nanjing 210093, China (chengcq@nju.edu.cn).

A function h is called positive definite if its Hessian matrix $\partial^2 h$ is positive definite.

THEOREM 1.1. *Assume $h \in C^r(B_D, \mathbb{R})$ is positive definite, $r \geq 6$. For any small $\delta > 0$, $E > \min h$ with $h^{-1}(E) \subset B_D$ and any two points $p^\star, p^* \in h^{-1}(E)$, there exists a cusp-residual set $\mathfrak{C}_{\epsilon_0} \subset C^r(B_D \times \mathbb{T}^3, \mathbb{R})$ such that for each $\epsilon P \in \mathfrak{C}_{\epsilon_0}$ there exists an orbit $(p(t), q(t))$ of Φ_H^t which intersects the δ-neighborhood of $p^\star$ and of p^*, namely, some $t^\star, t^* \in \mathbb{R}$ exist such that $\|p(t^\star) - p^\star\| < \delta$ and $\|p(t^*) - p^*\| < \delta$.*

This theorem proves the conjecture for positive definite Hamiltonian systems with three degrees of freedom in the cusp-residual generic sense in C^r-topology with $r \geq 6$. We apply the variational method for the proof, it is based on Mather's theory and the weak KAM theory [M91, M93, Fa]. Since we study the dynamics on the energy level set $\{H^{-1}(E)\}$, we can modify h outside of a neighborhood of $\{h^{-1}(E)\}$ such that h is Tonelli. A Hamiltonian $H(p, q, t)$ is called Tonelli if it satisfies the conditions:

(1) its Hessian matrix $\partial^2_{p_i p_j} H$ in p is positive definite everywhere;
(2) for each (q, t) it holds that $H(p, q, t)/\|p\| \to \infty$ as $\|p\| \to \infty$;
(3) each solution of the Hamilton's equation has all of $\mathbb{R}$ as its domain of definition.

For autonomous system, the third condition is automatically satisfied since each orbit lie on compact energy level set.

The definition of transition chain is recalled and the theorem of global transition chain is stated in Section 2 (Theorem 2.1). In Section 3, we derive the normal form of H when it is restricted in a neighborhood of double resonant point. In Section 4, we show that as a path, the candidate of transition chain is covered by discs around double resonances with controlled periods. In each disc, one Hamiltonian normal form holds. In Section 5, we distinguish strong from weak double resonances and prove that there are only finitely many strong double resonances. The weak double resonance can be reduced to *a priori* unstable case such the problem is reduced to the finite number of strong double resonances. In Section 6, we construct transition chain crossing strong double resonances by applying the main results of [CZ, C17a, C17b]. By preparing some technical estimates for the nearly integrable system including the deviation of the rotation vectors, the location of the flat and the estimate of orbits in the Aubry sets in Section 7, we prove Theorem 2.1 in Section 8. As an immediate consequence, Theorem 1.1 is proved in Section 9.

A lot of works have been contributed to the topic since the conjecture was raised half a century ago. Normally hyperbolic invariant cylinder is assumed by the *a priori* unstable condition, along which the diffusion is well understood, by variational method and geometric methods, cf. [B08, CY1, CY2, DLS1, LC, Tr, Zh1]. There are also many works for the problem, for instance, see [Bs, BCV, DH1, DH2, DLS2, FM, GL, GR1, GR2, KL1, KL2, X].

Nearly integrable Hamiltonian is also called *a priori* stable system. Unlike *a priori* unstable system, multiple resonant points destruct the cylinder into many small pieces. Away from the multiple resonant points, some piece of invariant cylinder was found in [B10] and the method for *a priori* unstable system was applied in [BKZ] to obtain local diffusion. Restricted in a neighborhood of multiple resonant point p'', $\{\|p - p''\| \leq K\sqrt{\epsilon}\}$ with $K \gg 1$, the normal form is non-integrable. So, it is a challenge to construct cylinder in such a disc. The condition $n = 3$ allows us to apply a variational method to construct cylinder which extends $o(\sqrt{\epsilon})$-close to double resonant point, see [CZ, C17a]. Because of the result and by a new cohomology equivalence, we found a way in [C17b] to pass through the small neighborhood by turning around the strong double resonant point and joining two cylinders.

Earlier than us, Mather suggested a way to cross the double resonance [M04], which is based on an observation that the periodic orbits of the averaged system (6.3) may approach two homoclinics simultaneously. He suggested to move the first cohomology class in the channel determined by the prescribed homology class and switch it to the channel determined by one of the homoclinics when it is getting close to the double resonance. From geometric point of view, one expects to construct diffusion orbit that moves along the cylinder with the prescribed homology class, jumps to the cylinder with hole and passes through the neighborhood of double resonance in a way similar to *a priori* unstable case, as it was announced in [KZ2], see [KZ1, Mar] also. For this approach one needs to prove that the cylinder is C^1-differentiable.

2. The definition of the transition chain. The terminology (*generalized*) *transition chain* used in the paper is defined in [CY1, CY2, LC], borrowed from [A64] where it is defined by geometrical language. Definition in our setting is in a variational language.

For the definition, let $\check{\pi}:\check{M}\to\mathbb{T}^n$ be a finite covering of $\mathbb{T}^n$, let $\mathcal{N}(c,\check{M})$, $\mathcal{A}(c,\check{M})$ denote the Mañé set, Aubry set with respect to $\check{M}$. The condition (**HA**) (*hypothesis of Arnold*) is a variational version of Arnold's condition, the stable manifold of a circle intersects its unstable manifold transversally. Such intersection points lie in the Mañé set, but not in the Aubry set.

(**HA**): there exists a finite covering $\check{\pi}:\check{M}\to M$ such that
(1) in time-periodic case: $\check{\pi}\mathcal{N}(c,\check{M})|_{t=0}\backslash(\mathcal{A}(c,\check{M})|_{t=0}+\delta)\neq\varnothing$ is totally disconnected, where $\mathcal{A}(c,\check{M})|_{t=0}+\delta=\{x:\mathrm{dist}(x,\mathcal{A}(c,\check{M})|_{t=0})\le\delta\}$;
(2) in autonomous case: $\check{\pi}\mathcal{N}(c,\check{M})|_{\Sigma}\backslash(\mathcal{A}(c,\check{M})+\delta)\neq\varnothing$ is totally disconnected, where Σ is a section of $\check{M}$.

It is not necessary to work always in nontrivial finite covering space. If the Aubry set contains more than one class, one can choose $\check{M}=\mathbb{T}^n$.

To state the definition of transition chain, we also need the concept of *cohomology equivalence*. The first version was introduced in [M93], however, it does not apply to interesting problem in autonomous systems (cf. [B02]). A new version of cohomology equivalence was introduced for autonomous system in [LC]. For a Tonelli Lagrangian defined on $T\mathbb{T}^n$, it is defined not with respect to the whole $\mathbb{T}^n$ as in [M93], but to a section. For n-torus $\mathbb{T}^n$, the section is chosen as a non-degenerately embedded section $(n-1)$-dimensional torus. We call Σ_c non-degenerately embedded $(n-1)$-dimensional torus by assuming a smooth injection φ: $\mathbb{T}^{n-1}\to\mathbb{T}^n$ such that Σ_c is the image of φ, and the induced map φ_*: $H_1(\mathbb{T}^{n-1},\mathbb{Z})\to H_1(\mathbb{T}^n,\mathbb{Z})$ is an injection.

For a first cohomology class c, we assume that there is a non-degenerate embedded $(n-1)$-dimensional torus $\Sigma_c\subset\mathbb{T}^n$ such that each c-semi static curve γ transversally intersects Σ_c. Let

$$\mathbb{V}_c=\bigcap_U\{i_{U*}H_1(U,\mathbb{R}):U\,\text{is a neighborhood of}\,\mathcal{N}(c)\cap\Sigma_c\},$$

here i_U: $U\to M$ denotes inclusion map. $\mathbb{V}_c^{\perp}$ is defined to be the annihilator of $\mathbb{V}_c$, i.e. if $c'\in H^1(\mathbb{T}^n,\mathbb{R})$, then $c'\in\mathbb{V}_c^{\perp}$ if and only if $\langle c',h\rangle=0$ for all $h\in\mathbb{V}_c$. Clearly,

$$\mathbb{V}_c^{\perp}=\bigcup_U\{\ker i_U^*:U\,\text{is a neighborhood of}\,\mathcal{N}(c)\cap\Sigma_c\}.$$

There is a neighborhood U of $\mathcal{N}(c)\cap\Sigma_c$ such that $\mathbb{V}_c=i_{U*}H_1(U,\mathbb{R})$ and $\mathbb{V}_c^{\perp}=\ker i_U^*$.

Definition 2.1. In autonomous case, $c,c'\in H^1(\mathbb{T}^n,\mathbb{R})$ are said to be cohomologically equivalent if there is a continuous curve Γ: $[0,1]\to H^1(\mathbb{T}^n,\mathbb{R})$ such that $\Gamma(0)=c$, $\Gamma(1)=c'$, $\alpha(\Gamma(s))$ keeps constant along Γ, and for each $s_0\in[0,1]$ there exists $\delta>0$ such that $\Gamma(s)-\Gamma(s_0)\in\mathbb{V}^{\perp}_{\Gamma(s_0)}$ whenever $s\in[0,1]$ and $|s-s_0|<\delta$.

With the terminologies introduced as above, we are able to state the definition of transition chain for autonomous system.

Definition 2.2. Two cohomolgy classes $c,c'\in H^1(M,\mathbb{R})$ are joined by a generalized transition chain if a continuous curve Γ: $[0,1]\to H^1(M,\mathbb{R})$ exists such that $\alpha(\Gamma(s))$ keeps constant and for each $s\in[0,1]$ at least one of the following cases takes place:

(1) the condition (**HA**) holds for $\Gamma(s)$, $\mathcal{A}(\Gamma(s'))$ lies in a small neighborhood of $\mathcal{A}(\Gamma(s))$ provided $|s'-s|$ is small;
(2) there is $\delta_s>0$, for each $s'\in(s-\delta_s,s+\delta_s)$, $\Gamma(s')$ is cohomologically equivalent to $\Gamma(s)$.

It is proved in Theorem 3.1 of [LC] for autonomous case that $\tilde{\mathcal{A}}(\Gamma(0))$ is dynamically connected to $\tilde{\mathcal{A}}(\Gamma(1))$, namely, there is an orbit of the system which takes $\tilde{\mathcal{A}}(\Gamma(0))$ and $\tilde{\mathcal{A}}(\Gamma(1))$ as its α-limit and ω-limit set respectively.

A Tonelli Lagrangian L is uniquely related to a Tonelli Hamiltonian H through Legendre transformation $L(q,\dot q,t)=\max_p\langle\dot q,p\rangle-H(p,q,t)$, which determines a map $\mathscr{L}_H$: $T^*\mathbb{T}^n\times\mathbb{T}\to T\mathbb{T}^n\times\mathbb{T}$: $(p,q,t)\to(\dot q,q,t)$ with $\dot q=\partial_pH(p,q,t)$.

Definition 2.3. An orbit $(p(t),q(t),t)\subset T^*\mathbb{T}^n\times\mathbb{T}$ is said to be *$\tilde c$-semi-static (static)* if $\mathscr{L}_H(p(t),q(t),t)=(\dot q(t),q(t),t)$ is $\tilde c$-semi-static (static). If the system is autonomous, we skip the component of t (see [Man, M93]).

Since $H^1(\mathbb{T}^3,\mathbb{R})=\mathbb{R}^3$, we treat the first cohomology class $\tilde c\in H^1(\mathbb{T}^3,\mathbb{R})$ as a point $\tilde c\in\mathbb{R}^3$. In this case, the choice of diffusion path relies on the observation as follows. It holds along each $\tilde c$-semi-static orbit of the integrable Hamiltonian $h(p)$ that $p(t)\equiv\tilde c$. Since Mañé set is upper semi-continuous with respect to the perturbation (see Lemma 2.3 of [CY2] and follow the proof there), $\|p(t)-\tilde c\|\ll1$ holds along any $\tilde c$-semi static orbit for the perturbed system.

Given two points $p^\star,p^*\in h^{-1}(E)$ and small $\delta>0$, $\exists$ points $\bar p^\star,\bar p^*\in h^{-1}(E)$ and vectors $k^\star,k^*\in\mathbb{Z}^3\backslash\{0\}$ such that $\|p^\star-\bar p^\star\|<\frac{\delta}{2}$, $\|p^*-\bar p^*\|<\frac{\delta}{2}$, $\langle k^\star,\partial h(\bar p^\star)\rangle=0$ and $\langle k^*,\partial h(\bar p^*)\rangle=0$. One can choose $(\bar p^\star,k^\star)$ and $(\bar p^*,k^*)$ such that $k^\star$ and k^* are totally irreducible. A vector $k=(k_1,k_2,k_3)\in\mathbb{Z}\backslash\{0\}$ is said to be *totally irreducible* if the greatest common divisor of k_i and k_j is equal to 1 for any $i\neq j$ and $i,j=1,2,3$. It is based on the observation that, for any $k\in\mathbb{Z}^3\backslash\{0\}$, one can choose totally irreducible $k'\in\mathbb{Z}^3\backslash\{0\}$ such that $\langle k,k'\rangle/\|k\|\|k'\|$ is close to 1. For a point $p=(p_1,p_2,p_3)\in\mathbb{R}^3$, we adopt the following notation for its maximum and Euclidean norm respectively

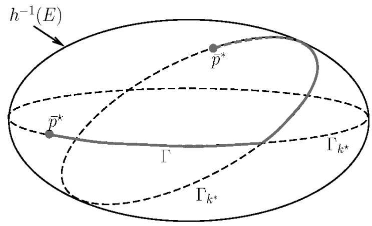
$$|p|=\max\{|p_1|,|p_2|,|p_3|\},\qquad \|p\|=\Big(\sum_{i=1}^3p_i^2\Big)^{\frac12}.$$

An integer $k\in\mathbb{Z}^3\backslash\{0\}$ determines a path of single resonance (a circle on a sphere)

$$\Gamma_k=\{p\in\mathbb{R}^3:h(p)=E>\min h;\langle k,\partial h(p)\rangle=0\}.$$

As h is positive definite, ∂h maps $\{p\in\mathbb{R}^3:h(p)\le E\}$ to a ball containing the origin. The circles $\Gamma_{k^\star}$ and Γ_{k^*} intersect at two points if $k^\star$ is independent of k^*, otherwise

$\Gamma_{k^*} = \Gamma_{k^\star}$. In both cases, one has a path Γ connecting $\bar{p}^\star$ to $\bar{p}^*$. If $\Gamma_{k^\star}$ intersects Γ_{k^*}, it starts from the point $\bar{p}^\star$, moves along the circle $\Gamma_{k^\star}$ until it reaches the intersection point of $\Gamma_{k^\star}$ with Γ_{k^*}, after that, it moves along the circle Γ_{k^*} until it arrives at the point $\bar{p}^*$. If $\Gamma_{k^\star} = \Gamma_{k^*}$, Γ is just a piece of the circle, connecting $\bar{p}^\star$ to $\bar{p}^*$, see the figure below.



Treating Γ as a path in $H^1(\mathbb{T}^3,\mathbb{R})$, we have a *candidate of the transition chain.* We shall show that the transition chain lies in a small neighborhood of Γ.

The following theorem is a slightly stronger version of Theorem 5.1 of [C17b], the main part of this paper is for the proof of this theorem.

THEOREM 2.1. *Assume $h \in C^r(B_D,\mathbb{R})$ is positive definite with $r \geq 6$. For any small $\delta > 0$, $E > \min h$ with $h^{-1}(E) \subset B_D$, there is a cusp-residual set $\mathfrak{C}_{\epsilon_0} \subset C^r(B_D \times \mathbb{T}^3,\mathbb{R})$ such that for each $\epsilon P \in \mathfrak{C}_{\epsilon_0}$ and any two points $p^\star, p^* \in h^{-1}(E)$, there is a transition chain that connects the class $\tilde{c}$ to the class $\tilde{c}'$ which satisfy the condition $\alpha(\tilde{c}) = \alpha(\tilde{c}') = E$, $|p^\star - \tilde{c}| < \delta$ and $|p^* - \tilde{c}'| < \delta$.*

The definition of cohomology equivalence can be further extended to more general version if *we treat the time t as an angle variable and choose a section in the extended configuration space $\mathbb{T}^{n+1}$ where the extra dimension is for t.* If we write the cohomology class in coordinates $\tilde{c} = (c, -\alpha(c))$, the section $\Sigma_{\tilde{c}}$ is chosen for $\mathbb{T}^{n+1}$, $\mathbb{V}_{\tilde{c}}$ and $\mathbb{V}_{\tilde{c}}^{\perp}$ are defined in $H_1(\mathbb{T}^{n+1},\mathbb{R})$ and $H^1(\mathbb{T}^{n+1},\mathbb{R})$ respectively.

DEFINITION 2.4. In time-periodic case, $c, c' \in H^1(\mathbb{T}^n,\mathbb{R})$ are said to be cohomologically equivalent if there exists a continuous curve $\tilde{\Gamma}$: $[0,1] \to H^1(\mathbb{T}^{n+1},\mathbb{R})$ such that $\tilde{\Gamma}(0) = (c, -\alpha(c))$, $\tilde{\Gamma}(1) = (c', -\alpha(c'))$, and for each $s_0 \in [0,1]$ there exists $\delta > 0$ such that $\tilde{\Gamma}(s) - \tilde{\Gamma}(s_0) \in \mathbb{V}^{\perp}_{\tilde{\Gamma}(s_0)}$ whenever $s \in [0,1]$ and $|s - s_0| < \delta$.

3. Normal form. Given an irreducible integer vector $k' \in \mathbb{Z}^3\backslash\{0\}$, one has a path of single resonance $\Gamma_{k'} = \{p \in H^{-1}(E) : \langle k', \partial h(p)\rangle = 0\}$. A point $p'' \in \Gamma_{k'}$ is said to be *double resonant* if there is an additional vector $k'' \in \mathbb{Z}^3\backslash\{0\}$ independent of k' such that $\langle k'', \partial h(p'')\rangle = 0$. Along each resonant path $\Gamma_{k'}$, there are many double resonant points.

We assume that $k' = (k'_1, k'_2, k'_3) \in \mathbb{Z}^3\backslash\{0\}$ is totally irreducible. In this case, there exist $k^*, k^\star \in \mathbb{Z}^3$ such that the matrix $M_0^t = (k', k^*, k^\star)$ is uni-modular. Indeed, if k' contains two non-zero entries, e.g. $k'_1, k'_2 \neq 0$, we set $k^\star = (0,0,1)$ and $k^* = (k^*_1, k^*_2, 0)$ such that $k'_1 k^*_2 - k'_2 k^*_1 = 1$. If $k' = e_1$, we set $k^* = e_2$ and $k^\star = e_3$, where we use the notation that all other entries of e_i are equal to zero except for the i-th entry, which is equal to 1. Other cases can be handled similarly.

Under a linear canonical transformation $\mathfrak{M}_0$: $(\bar{p}, \bar{q}) \to (p, q)$ such that $q = M_0^{-1}\bar{q}$, $p = M_0^t \bar{p}$, we obtain the Hamiltonian $\bar{H} = \bar{h} + \epsilon \bar{P}$ where $\bar{h} = \mathscr{M}_0^* h$ and $\bar{P} = \mathscr{M}_0^* P$. It holds along the path $\bar{\Gamma}_{k'} = M_0^{-t}\Gamma_{k'}$ that $\partial \bar{h}(\bar{p}) = (0, \omega_2, \omega_3)$.

In this case, the first and the second resonant condition are $M_0^{-t}k'=\bar k'=e_1$ and $M_0^{-t}k''=\bar k''=(0,\bar k''_2,\bar k''_3)$ respectively. If we introduce the canonical transformation of coordinates $\mathscr{M}^1$ $(u,v)\to(\bar p,\bar q)$ further

$$\bar q=M^{-1}u,\qquad \bar p=M^tv, \tag{3.1}$$

where

$$M^t=\begin{bmatrix}1&0&0\\0&1&0\\0&\frac{\bar k''_3}{\bar k''_2}&1\end{bmatrix},\qquad M^{-1}=\begin{bmatrix}1&0&0\\0&1&\frac{-\bar k''_3}{\bar k''_2}\\0&0&1\end{bmatrix}$$

if $|\bar k''_2|\ge|\bar k''_3|$ and

$$M^t=\begin{bmatrix}1&0&0\\0&1&\frac{\bar k''_2}{\bar k''_3}\\0&0&1\end{bmatrix},\qquad M^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&\frac{-\bar k''_2}{\bar k''_3}&1\end{bmatrix}$$

if $|\bar k''_2|\le|\bar k''_3|$. The function $\mathscr{M}^*\bar H$ is 2π-periodic in (u_1,u_2), $2|\bar k''|\pi$ in u_3 if $|\bar k''_2|\ge|\bar k''_3|$ and it is 2π-periodic in (u_1,u_3), $2|\bar k''|\pi$ in u_2 if $|\bar k''_3|\ge|\bar k''_2|$.

By the construction of M which may not be uni-modular, we see that the function $\mathscr{M}^*H$ respects two symmetries in u.

Definition 3.1. Let M be a non-degenerate matrix. A function $f(u)\in C^r(\mathbb{R}^n,\mathbb{R})$ is said to respect the symmetry M if

$$f(u+2\pi Me_i)=f(u),\qquad \forall\ u\in\mathbb{R}^n,\ e_i\in\mathbb{Z}^n.$$

Since $\mathscr{M}^*\bar H$ is 2π-periodic in (u_1,u_2), $2|\bar k''|\pi$-periodic in u_3 in the case that $|\bar k''_2|\ge|\bar k''_3|$ and 2π-periodic in (u_1,u_3), $2|\bar k''|\pi$ in u_2 in the case that $|\bar k''_3|\ge|\bar k''_2|$

$$\mathscr{M}^*\bar H(u,v)=\sum_{k\in\mathbb{Z}^3}\bar H_k(M^tv)e^{i\langle k,M^{-1}u\rangle},$$

$\mathscr{M}^*\bar H$ respects two symmetries in the variable u, M and $\mathrm{diag}\{1,1,|\bar k''|\}$ for $|\bar k''_2|\ge|\bar k''_3|$, M and $\mathrm{diag}\{1,|\bar k''|,1\}$ for $|\bar k''_2|\le|\bar k''_3|$ respectively.

At a double point $\bar p''$, the rotation vector $\bar\omega=(0,\bar\omega_2,\bar\omega_3)=\partial\bar h(\bar p'')$ is rational, i.e. $\exists\ T>0$ so that $T\omega\in\mathbb{Z}^n\backslash\{0\}$. If $t\omega\notin\mathbb{Z}^n\ \forall\ t\in(0,T)$, $T=T(\omega)$ is called *the (minimal) period*. Since $\bar k''_2\bar\omega_2+\bar k''_3\bar\omega_3=0$, $T=|\bar k''|$. We consider those double resonant points $\{\bar p''\in\bar\Gamma_{k'}\}$ such that $T=T(\partial\bar h(\bar p''))\le K^*\epsilon^{-\frac13(1-3\kappa)}$ with $\kappa\in(0,\frac16)$ and $K^*>0$ is independent of ϵ. In this case, $|\bar k''|\sqrt{\epsilon}\to0$ as $\epsilon\to0$.

Lemma 3.1. *Assume the second resonant condition at $\bar p''\in\bar\Gamma_{k'}$ is $\bar k''=(0,\bar k''_2,\bar k''_3)$, $\delta'\in(0,1/2)$. Then, there exists a small number $\epsilon_0>0$ such that for each $\epsilon\in(0,\epsilon_0]$, restricted on the level set $\bar H^{-1}(E)$ contained in $\tilde\Sigma_\epsilon\times\mathbb{T}^3$ with*

$$\tilde\Sigma_\epsilon=\{\bar p:|\bar p-\bar p''|\le K^{-1}\epsilon^\kappa\},$$

where $K=\eta T$, $\eta\in(0,1]$ is independent of ϵ, the Hamiltonian $\bar H$ is reduced, by a symplectic transformation and an energetic reduction, to a time-periodic perturbation of mechanical system with two degrees of freedom

$$G_\epsilon(x,y,\theta)=\frac12\langle By,y\rangle-V(x_1,|\bar k''|x_2)+R_\epsilon\Big(x,y,\frac{\omega_3}{\sqrt{\epsilon}}\theta\Big), \tag{3.2}$$

[1] In [C17a], the matrix M_0M is set to be uni-modular. It is not always possible and not necessary.

where the 2×2 matrix B is positive definite, $V\in C^r$ is 2π-periodic in $(x_1,|\bar{k}''|x_2)$, $R_\epsilon(x,y,\vartheta)\in C^{r-2}(\mathbb{T}^2\times\Sigma'_\epsilon\times|\bar{k}''|\mathbb{T},\mathbb{R})$, $\vartheta=\omega_3\sqrt{\epsilon}^{-1}\theta$ and Σ'_ϵ satisfies the condition

$$\{y:|y|\le(1-\delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}\subseteq\Sigma'_\epsilon\subseteq\{y:|y|\le(1+\delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}.$$

Restricted in $\mathbb{T}^2\times\Sigma'_\epsilon\times|\bar{k}''|\mathbb{T}$, some number $a_0=a_0(h,E,k')>0$ exists, independent of T and P, such that for each $P\in\mathfrak{B}_1$ one has

$$\|R_\epsilon\|_{C^{r-2}(\mathbb{T}^2\times\Sigma'_\epsilon\times|\bar{k}''|\mathbb{T},\mathbb{R})}\le a_0\epsilon^\kappa,$$

if it is treated as a function in (x,y,ϑ). Finally, the remainder $R_\epsilon(x,y,\vartheta)$ respects the symmetries M and $\mathrm{diag}\{1,1,|\bar{k}''|\}$ *in (x,ϑ) and the symplectic transformation is uniformly bounded for any second resonant condition $\bar{k}''$.*

Remark. Because V is independent of ϑ, the symmetries M and $\mathrm{diag}\{1,1,|\bar{k}''|\}$ for V are the same as the identity.

Proof of Lemma 3.1. To get the normal form, we introduce a coordinate transformation $\Phi_{\epsilon F}$ which is defined as the time-2π-map $\Phi_{\epsilon F}=\Phi^t_{\epsilon F}|_{t=2\pi}$ of the Hamiltonian flow generated by the function $\epsilon F(p,q)$. The function F solves the homological equation

$$\Big\langle\frac{\partial\bar{h}}{\partial\bar{p}}(\bar{p}''),\frac{\partial F}{\partial\bar{q}}\Big\rangle=-\bar{P}(\bar{p},\bar{q})+Z(\bar{p},\bar{q})$$

where

$$Z(\bar{p},\bar{q})=\frac{1}{T}\int_0^T\bar{P}(\bar{p},\bar{q}+\bar{\omega}t)dt=\sum_{(\ell_1,\ell_2)\in\mathbb{Z}^2}\bar{P}_\ell(\bar{p})e^{i(\ell_1\langle\bar{k}',\bar{q}\rangle+\ell_2\langle\bar{k}'',\bar{q}\rangle)},\tag{3.3}$$

Clearly, the function

$$F(p,q)=\frac{1}{T}\int_0^T\bar{P}(\bar{p},\bar{q}+\bar{\omega}t)tdt$$

solves the homological equation and $\|F\|<T\|\bar{P}\|$.

Under the transformation $\Phi_{\epsilon F}$ we obtain a new Hamiltonian

$$\begin{aligned}\Phi^*_{\epsilon F}\bar{H}=&\bar{h}(\bar{p})+\epsilon Z(\bar{p},\bar{q})+\epsilon\Big\langle\frac{\partial\bar{h}}{\partial\bar{p}}(\bar{p})-\frac{\partial\bar{h}}{\partial\bar{p}}(\bar{p}''),\frac{\partial F}{\partial\bar{q}}\Big\rangle\\&+\frac{\epsilon^2}{2}\int_0^1(1-t)\{\{\bar{H},F\},F\}\circ\Phi^t_{\epsilon F}dt.\end{aligned}\tag{3.4}$$

To simplify the situation further, we introduce another canonical transformation of coordinates $\mathscr{M}$: $(u,v)\to(\bar{p},\bar{q})$ defined in (3.1). Although the norm $\|\bar{k}''\|$ will be large if the second resonant condition is weak, the linear coordinate transformation (3.1) is uniformly bounded in the second resonant condition since $|\bar{k}''_3/\bar{k}''_2|\le 1$ in the first case and $|\bar{k}''_2/\bar{k}''_3|\le 1$ in the second case.

We consider the case that $|\bar{k}''_2|\ge|\bar{k}''_3|$. In the new coordinates (u,v), the rotation vector takes the form of $(0,0,\omega_3)$. Clearly, $|\omega_3|$ is uniformly lower bounded above zero for all double resonant points on $\Gamma_{k'}$. Because $(\bar{k}''_2,\bar{k}''_3)$ is irreducible, it follows from (3.1) that $\mathscr{M}^*\Phi^*_{\epsilon F}\bar{H}$ is 2π-periodic in (u_1,u_2) and $2|\bar{k}''|\pi$-periodic in u_3.

By the construction, the function $\mathscr{M}^*\Phi^*_{\epsilon F}\bar{H}$ possesses the symmetries of M and $\mathrm{diag}\{1,1,|\bar{k}''|\}$. We need to be careful when a perturbation is added, it should respect the symmetries as well. By the relation (3.1) one has

$$\langle k',q\rangle = \langle k',(MM_0)^{-1}u\rangle = u_1, \qquad \langle k'',q\rangle = \langle k'',(MM_0)^{-1}u\rangle = \bar{k}''_2u_2.$$

It follows from Formula (3.3) and the transformation (3.1) that the resonant term has the form of $\mathscr{M}^*\mathscr{M}_0^*Z(p,\langle k',q\rangle,\langle k'',q\rangle) = Z'(v,u_1,\bar{k}''_2u_2)$.

Let $h' = \mathscr{M}^*\bar{h}$, $F' = \mathscr{M}^*F$ and $H' = \mathscr{M}^*\Phi^*_{\epsilon F}\bar{H}$. Since the transformations (3.1) is canonical, it preserves the Poison bracket. We obtain from Formula (3.4) that

$$\begin{aligned} H' =& h'(v) + \epsilon Z'(v,u_1,|\bar{k}''|u_2) + \epsilon\Big\langle \frac{\partial h'}{\partial v}(v) - \frac{\partial h'}{\partial v}(v''), \frac{\partial F'}{\partial u}\Big\rangle \\ &+ \frac{\epsilon^2}{2}\mathscr{M}^*\int_0^1(1-t)\{\{H,F\},F\}\circ\Phi^t_{\epsilon F}dt, \end{aligned}$$

where $v'' = (M_0M)^{-1}p''$. The function H' determines its Hamiltonian equation

$$\frac{du}{dt} = \frac{\partial}{\partial v}H'(u,v), \qquad \frac{dv}{dt} = -\frac{\partial}{\partial u}H'(u,v). \tag{3.5}$$

For this equation we introduce another transformation

$$\tilde{G}_\epsilon = \frac{1}{\epsilon}H', \qquad \tilde{y} = \frac{1}{\sqrt{\epsilon}}\Big(v - v''\Big), \qquad \tilde{x} = u, \qquad s = \sqrt{\epsilon}t, \tag{3.6}$$

where we use the notation $\tilde{y} = (y_1,y_2,y_3) = (y,y_3)$ and $\tilde{x} = (x_1,x_2,x_3) = (x,x_3)$. In the new canonical variables $(\tilde{x},\tilde{y})$ and the new time s, Equation (3.5) turns out to be the Hamiltonian equation with the generating function as the following:

$$\tilde{G}_\epsilon = \frac{1}{\epsilon}\Big(h'(v''+\sqrt{\epsilon}\tilde{y}) - h'(v'')\Big) - V(x_1,|\bar{k}''|x_2) + \tilde{R}_\epsilon(\tilde{x},\tilde{y}), \tag{3.7}$$

where $V(x_1,|\bar{k}''|x_2) = -Z'(v'',x_1,|\bar{k}''|x_2)$ and $\tilde{R}_\epsilon = \tilde{R}_{\epsilon,1} + \tilde{R}_{\epsilon,2} + \tilde{R}_{\epsilon,3}$ with

$$\begin{aligned} \tilde{R}_{\epsilon,1} &= Z'(v''+\sqrt{\epsilon}\tilde{y},\tilde{x}) - Z'(v'',\tilde{x}), \\ \tilde{R}_{\epsilon,2} &= \Big\langle \frac{\partial h'}{\partial v}\Big(v''+\sqrt{\epsilon}\tilde{y}\Big) - \frac{\partial h'}{\partial v}(v''), \frac{\partial F'}{\partial u}\Big\rangle, \\ \tilde{R}_{\epsilon,3} &= \frac{\epsilon}{2}\mathscr{M}^*\int_0^1(1-t)\{\{\bar{H},F\},F\}\circ\Phi^t_{\epsilon F}dt. \end{aligned}$$

One step of KAM iteration makes the remainder $\tilde{R}_\epsilon$ lose two times of differentiability. Since the Hamiltonian $\bar{H}$ is defined in $\tilde{\Sigma}_\epsilon\times\mathbb{T}^2\times|\bar{k}''|\mathbb{T}$, we see from (3.1), (3.6) and $\Phi_{\epsilon F}$ that $\tilde{G}$ is defined on the domain that is contained in $\{|\tilde{y}|\le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}\times\mathbb{T}^2\times|\bar{k}''|\mathbb{T}$. Restricted in the domain, we claim that there exists a number $a_2>0$, depending on h,E,k' only, such that

$$\|\tilde{R}_{\epsilon,i}\|_{C^{r-2}} \le a_2\epsilon^\kappa, \qquad i=1,2,3, \tag{3.8}$$

for small $\epsilon>0$. Indeed, let m' be the upper bound of the largest eigenvalue of $\partial^2h(p)$ for all $p\in h^{-1}(E)$. Since $T\le K^*\epsilon^{-\frac{1}{3}(1-3\kappa)}$, one has

$$\begin{aligned} |\tilde{R}_{\epsilon,1}| &\le 2\|P'\|_{C^1}\sqrt{\epsilon}K^{-1}\epsilon^{\kappa-\frac{1}{2}} \le 2\|P'\|_{C^1}K^{-1}\epsilon^\kappa, \\ |\tilde{R}_{\epsilon,2}| &\le m'\sqrt{\epsilon}K^{-1}\epsilon^{\kappa-\frac{1}{2}}T\|P'\|_{C^1} \le m'\|P'\|_{C^1}\epsilon^\kappa, \\ |(\mathscr{M}^*)^{-1}\tilde{R}_{\epsilon,3}| &\le \frac{\epsilon}{2}\|P'\|^2_{C^2}\|H'\|_{C^1}T^2 = \frac{1}{2}(K^*\|P'\|_{C^2})^2\|H'\|_{C^1}\epsilon^{\frac{1}{3}+2\kappa}. \end{aligned}$$

The estimate on the derivatives of the terms can also be done inductively.

We introduce another coordinate rescaling further

$$\theta=\frac{\sqrt{\epsilon}}{\omega_3}x_3,\qquad I=\frac{\omega_3}{\sqrt{\epsilon}}y_3, \tag{3.9}$$

By expanding $\tilde{G}_\epsilon$ in $O(K^{-1}\epsilon^\kappa)$ neighborhood of v'' in Taylor formula, we obtain

$$\begin{aligned}\tilde{G}_\epsilon(x,y,I,\theta)=&I+\frac{1}{2}\Big\langle\tilde{B}\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big),\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big)\Big\rangle-V(x_1,|\bar{k}''|x_2)\\&+\tilde{R}_h\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big)+\tilde{R}_\epsilon\Big(x_1,x_2,\frac{\omega_3}{\sqrt{\epsilon}}\theta,p''+\Big(\sqrt{\epsilon}y,\frac{\epsilon}{\omega_3}I\Big)\Big)\end{aligned} \tag{3.10}$$

where $\tilde{B}=\frac{\partial^2h'}{\partial v^2}(v'')$ and term $\tilde{R}_h$ represents the following

$$\frac{1}{\epsilon}\left[h'\Big(v''+\Big(\sqrt{\epsilon}y,\frac{\epsilon}{\omega_3}I\Big)\Big)-\Big[h'(v'')+\epsilon I+\frac{\epsilon}{2}\Big\langle\tilde{B}\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big),\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big)\Big\rangle\Big]\right].$$

By the construction, all entries of $\tilde{B}$ are of order $O(1)$, independent of the period T of the rotation vector $\partial h(p'')$. We write

$$\tilde{B}=\begin{bmatrix}B & B'\\ B'^t & B''\end{bmatrix}$$

where B is a 2×2 matrix, B' is a vector with two entries and $B''>0$.

Obviously, restricted on the domain $\{|y|\le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}},|I|\le 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$, there exists a constant $a_3=a_3(h,E)>0$ such that

$$\|\tilde{R}_h\|_{C^{r-2}}\le a_4\epsilon^\kappa. \tag{3.11}$$

Let Ω_ϵ be the image of $\tilde{\Sigma}_\epsilon\times\mathbb{T}^3$ under the maps $\Phi_{\epsilon F'}$, (3.1), (3.6) and (3.9). Since the transformation $\Phi_{\epsilon F'}$ is close to identity, $|k''_2|\ge|k''_3|$ is assumed, each section of Ω_ϵ where (x,θ) keeps constant lies in $\{|y|\le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}},|I|\le 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$. Restricted in Ω_ϵ, we find by direct calculation that $\partial_I\tilde{G}_\epsilon=1+O(\epsilon^\kappa)$. Therefore, there exists a function $G_\epsilon(x,y,\theta)$ solves the equation $\tilde{G}_\epsilon(x,y,-G_\epsilon,\theta)=0$.

Indeed, restricted in the domain $\{|y|\le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}},|I|\le 2K^{-1}\omega_3^{-1}\epsilon^{\kappa-1}\}$, a constant $a_4=a_4(h)>0$ exists such that $|\sqrt{\epsilon}\langle B',y\rangle|\le a_4K^{-1}\epsilon^\kappa$ and $|\epsilon\langle By,y\rangle|\le a_4K^{-2}\epsilon^{2\kappa}$. It guarantees that the solution $I=-G_0(x,y)$ of the following quadratic equation in I

$$\begin{aligned}\tilde{G}_0=&I+\frac{1}{2}\Big\langle\tilde{B}\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big),\Big(y,\frac{\sqrt{\epsilon}}{\omega_3}I\Big)\Big\rangle-V(x_1,|\bar{k}''|x_2)\\=&\Big(1+\frac{\sqrt{\epsilon}}{\omega_3}\langle B',y\rangle\Big)I+\frac{1}{2}\langle By,y\rangle+\frac{B''}{2\omega_3^2}\epsilon I^2-V(x_1,|\bar{k}''|x_2)=0\end{aligned}$$

has the form of

$$\begin{aligned}G_0=&\frac{\omega_3^2(1+\frac{\sqrt{\epsilon}}{\omega_3}\langle B',y\rangle)}{B''\epsilon}\Big(1-\Big(1-\frac{2B''\epsilon(\frac{1}{2}\langle By,y\rangle-V(x_1,|\bar{k}''|x_2))}{\omega_3^2(1+\frac{\sqrt{\epsilon}}{\omega_3}\langle B',y\rangle)^2}\Big)^{\frac{1}{2}}\Big)\\=&\frac{1}{2}\langle By,y\rangle-V(x_1,|\bar{k}''|x_2)+R_0(x,y)\end{aligned}$$

and some $a_5(h,E) > 0$ exists such that restricted in the domain $\mathbb{T}^2 \times \{|y| \le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}$ one has $\|R_0\|_{C^{r-2}} \le a_5\epsilon^\kappa$, provided $P \in \mathfrak{B}_1$.

According to the estimates (3.8) and (3.11), one has

$$\|\tilde{G}_\epsilon - \tilde{G}_0\|_{C^{r-2}} = \|\tilde{R}_h + \tilde{R}_\epsilon\|_{C^{r-2}} \le (3a_2 + a_4)\epsilon^\kappa$$

when (x,ϑ,y,I) is restricted in $\mathbb{T}^2 \times |\bar{k}''|\mathbb{T} \times \{|y| \le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}}, |I| \le 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$ where $\vartheta = \sqrt{\epsilon}^{-1}\omega_3\theta$. It follows from the relations $\partial_I\tilde{G}_\epsilon = 1+O(\epsilon^\kappa)$, $\partial_I\tilde{G}_0 = 1+O(\epsilon^\kappa)$ and the theorem of implicit function that some $a_6 = a_6(h,d) > 0$ exists such that $\|G_\epsilon - G_0\|_{C^{r-2}} \le a_6\epsilon^\kappa$ holds when $|y| \le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}}$. Indeed, let $z = (x,y,\vartheta)$, we get from the equations $\tilde{G}_0(z,-G_0) = 0$ and $\tilde{G}_\epsilon(z,-G_\epsilon) = 0$ that

$$\partial_I\tilde{G}_0(z,-G_0+\lambda(G_\epsilon - G_0))(G_\epsilon - G_0) - (\tilde{R}_h + \tilde{R}_\epsilon)(z,G_\epsilon) = 0, \tag{3.12}$$

where $\lambda = \lambda(z, G_\epsilon - G_0) \in [0,1]$. It follows from the relation $\partial_I G_0 = 1 + O(\epsilon^\kappa)$ that $\max_z |G_\epsilon - G_0| \le 2(3a_2 + a_4)\epsilon^\kappa$. For $\xi = 0, \epsilon$, the derivative of G_ξ in z satisfies the equation

$$\partial_z\tilde{G}_\xi(z,-G_\xi(z)) - \partial_I\tilde{G}_\xi(z,-G_\xi(z))\partial_z G_\xi(z) = 0.$$

Because $\partial_I^2\tilde{G}_0 = \omega_3^{-2}B''\epsilon$, $|\partial_z G_\epsilon| = |\partial_I\tilde{G}_\epsilon|^{-1}|\partial_z\tilde{G}_\epsilon| = O(1)$, by taking the difference of these equations we obtain the estimate on first derivative of $G_\epsilon - G_0$

$$\begin{aligned}|\partial_z(G_\epsilon - G_0)(z)| \le& |\partial_I\tilde{G}_0(z,-G_0)^{-1}|\Big(|B''\omega_3^{-2}\epsilon(G_\epsilon - G_0)\partial_z G_\epsilon| \\ &+ |\partial_z\tilde{G}_0(z,-G_\epsilon) - \partial_z\tilde{G}_0(z,-G_0)| + \Big|\frac{d}{dz}(\tilde{R}_h + \tilde{R}_\epsilon)(z,-G_\epsilon)\Big|\Big) \\ \le& 2(3a_2 + a_4)\epsilon^\kappa\Big((2|\omega_3^{-2}B''|\epsilon + 1)|\partial_z G_\epsilon| + 2|B'||\omega_3^{-1}|\sqrt{\epsilon} + 1\Big).\end{aligned}$$

The estimate on higher order derivatives can be done similarly.

From the formula (3.2), we see that $|G_\epsilon(x,y,\theta)| \le a_4K^{-2}\epsilon^{2\kappa-1}$ for $|y| \le 2K^{-1}\epsilon^{\kappa-\frac{1}{2}}$ if $\epsilon > 0$ is suitably small such that $|V| + |R_\epsilon| < \frac{1}{2}a_4K^{-2}\epsilon^{2\kappa-1}$. Consequently, the energy level set $\tilde{G}_\epsilon^{-1}(0)$ intersects the domain $\{|y| \le K^{-1}\epsilon^{\kappa-\frac{1}{2}}, |I| \le K^{-1}\omega_3^{-1}\epsilon^{\kappa-1}\} \times \mathbb{T}^3$ at the place where $|I| \le a_4\epsilon^{2\kappa-1}$.

Under the composition of $\Phi_{\epsilon F'}$, (3.1), (3.6) and (3.9), $\tilde{\Sigma}_\epsilon \times \mathbb{T}^3$ is mapped onto Ω_ϵ. If $\Phi_{\epsilon F}$ is an identity map, each section of Ω_ϵ where (x,θ) keeps constant contains the disk $\{|y| \le K^{-1}\epsilon^{\kappa-\frac{1}{2}}, I = 0\}$, because $|k_2''| \ge |k_3''|$ is assumed in (3.1). Since $\Phi_{\epsilon F'}$ approaches to identity and $\frac{a_4\epsilon^{2\kappa-1}}{K^{-1}\omega_3^{-1}\epsilon^{\kappa-1}} = a_4K\omega_3\epsilon^\kappa \to 0$ as $\epsilon \to 0$, some $\epsilon_0(\delta') > 0$ exists such that, for any $\epsilon \in (0,\epsilon_0]$, the set $\{|y| \le (1-\delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}, |I| \le a_4\epsilon^{2\kappa-1}\}$ is contained in each section of Ω_ϵ where (x,θ) is fixed, i.e. $\Sigma_\epsilon \supset \{|y| \le (1-\delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}$. Similarly, we can show that $\Sigma_\epsilon \subset \{|y| \le (1+\delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}$.

The transformation $\mathscr{M}$ of (3.1) is uniformly bounded for all resonant conditions along $\Gamma_{k'}$ and all constants a_ℓ with $2 \le \ell \le 7$ can be set to be independent of the second resonant condition. Let $a_0 = a_6 + a_7$, we then complete the proof of the lemma in the case that $|\bar{k}_2''| \ge |\bar{k}_3''|$.

If $|\bar{k}_2''| \le |\bar{k}_3''|$, in the new coordinates (u,v), the frequency at the double resonance has the form of $M'^t\partial h(p'') = (0,\omega_3,0)$. From (3.1) one obtains

$$\langle k', q\rangle = \langle k', (MM_0)^{-1}u\rangle = u_1, \qquad \langle k'', q\rangle = \langle k'', (MM_0)^{-1}u\rangle = \bar{k}_3''u_3.$$

By introducing the permutation $u_2 \leftrightarrow u_3$ and $v_2 \leftrightarrow v_3$, we are again in the situation we have handled. The rest of the proof is the same as above. □

Let $A = B^{-1}$. By the Legendre transformation,

$$L_\epsilon(\dot{x}, x, \theta) = \max_y \{\langle \dot{x}, y\rangle - G_\epsilon(x, y, \theta)\}.$$

one obtains from the Hamiltonian G_ϵ the Lagrangian (1.1) defined in [C17b].

Treated as the set in T^*M, it is shown in [B07] that Mather set, Aubry set and Mañé set are symplectic invariants. Denote by $\tilde{\mathcal{M}}_H(c)$ the Mather set of the Tonelli Hamiltonian $H : T^*M \to \mathbb{R}$ in the cohomology class $c \in H^1(M, \mathbb{R})$.

Theorem 3.1 ([B07]). *Let $\Phi : T^*M \to T^*M$ be a Hamiltonian diffeomorphism. Then one has*

$$\Phi\tilde{\mathcal{M}}_H(c) = \tilde{\mathcal{M}}_{H\circ\Phi}(\Phi^* c), \quad c \in H^1(M, \mathbb{R}).$$

Similarly for Aubry set and Mañé set.

Applying the theorem, we find that the map $\Phi_{\epsilon F}$ does not induce the change of the structure of the Mañé sets and the Aubry sets. The transformations $\mathscr{M}_0$ and (3.1) are linear, (3.6) and (3.9) are rescaling. Therefore, the conditions (1) and (2) in Definition 2.2 remain unchanged under the coordinate transformations.

4. The covering property. We are going to show that the whole resonant path $\Gamma_{k'}$ can be covered by the disks where one obtains the normal form of (3.2).

Theorem 4.1 (Covering property). *Some $\epsilon_0 > 0$ exists such that for each $\epsilon \in (0, \epsilon_0]$ there exists a finite set of double resonant points $\{p_i'' \in \Gamma_{k'}\}$ with the properties*

1. *the period T_i of the frequency $\partial h(p_i'')$ is not large than $T_i \le K^*\epsilon^{-\frac{1}{3}(1-3\kappa)}$ where $\kappa \in (0, \frac{1}{6})$ and K^* is independent of ϵ;*
2. *$\Gamma_{k'}$ is covered by the union of the disks $\{\|p - p_i''\| \le T_i^{-1}\epsilon^\kappa\}$.*

Proof. If $\Gamma_{k'}$ is a set in $\mathbb{R}^n$ the theorem is proved in Chapter 3 of [Lo] with the condition $\kappa < (3n+3)^{-1}$. We use their idea to prove the covering property under the condition $\kappa < \frac{1}{6}$. To do that, we use Dirichlet's approximation theorem.

For real x, let $[x] \in \mathbb{Z}$ denote the integer part and $\{x\} \in (0, 1)$ denote the decimal part. So, one has

$$x = [x] + \{x\},$$

We use notation

$$\|x\|_{\mathbb{Z}} = \min\{\{x\}, 1 - \{x\}\} = \mathrm{dist}(x, \mathbb{Z}).$$

Proposition 4.1 (Dirichlet). *Let $\omega \in \mathbb{R}$ and $K > 1$ be a real number. There exists an integer k, $1 \le k < K$, such that*

$$\|k\omega\|_{\mathbb{Z}} \le K^{-1}.$$

For $\omega = (\omega_2, \omega_3) \in \mathbb{R}^2$, let $|\omega| = \max\{|\omega_2|, |\omega_3|\}$. Let $\mathbb{S}^1 = \{\omega \in \mathbb{R}^2 : |\omega| = 1\}$ be the boundary of unit square, it has four sides. By applying Dirichlet's approximation theorem, for any $\omega \in \mathbb{S}^1$ and any integer $K > 0$ there exists $1 \le k < K$ such that

$\|k\min\{\omega_2,\omega_3\}\|_{\mathbb{Z}} \le K^{-1}$. In other words, given any $\omega\in\mathbb{S}^1$, some rational vector ω^* on the same side exists such that $T\omega^*\in\mathbb{Z}^2$ with $T\le K$ and

$$\mathrm{dist}(T\omega,T\omega^*)=\|T\omega-T\omega^*\|\le K^{-1}. \tag{4.1}$$

To apply the inequality, we notice that $\partial h(\Gamma_{k'})$ is a circle lying on certain 2-dimensional plane.

By the definition, $\Gamma_{k'}$ is a smooth circle. Under the transformation $\mathscr{M}_0$ introduced in the last section, $M_0^{-1}\partial h$ maps the circle $\Gamma_{k'}$ to a smooth circle $\Gamma_{\omega,k'}=M_0^{-1}\partial h(\Gamma_{k'})$ restricted on the plane $\{(\omega_1,\omega_2,\omega_3):\omega_1=0\}$.

Since h is positive definite and $h(\Gamma_{k'})\equiv E>\min h$, each line $\{\lambda(0,\omega_2,\omega_3):\lambda\in\mathbb{R}\}$ intersects the circle $\Gamma_{\omega,k'}$ transversally and each $\omega\in\Gamma_{\omega,k'}$ determines a unique $\lambda_\omega>0$ such that $\lambda_\omega(\omega_2,\omega_3)\in\mathbb{S}^1$. Therefore, some number $d>0$ exists so that the distance between any two points $\omega=(0,\omega_2,\omega_3)$, $\omega^*=(0,\omega_2^*,\omega_3^*)\in\Gamma_{\omega,k'}$ is upper bounded by

$$\|\omega-\omega^*\|\le d\|\lambda_\omega\omega-\lambda_{\omega^*}\omega^*\|.$$

If ω^* is a rational vector with period T, the period of $\lambda_{\omega^*}\omega^*$ will be $\lambda_{\omega^*}^{-1}T$.

The map ∂h establishes a diffeomorphism between $\Gamma_{k'}$ and $\Gamma_{\omega,k'}$. Given an integer $K>0$, it follows from (4.1) that, for any rotation vector $\omega\in\Gamma_{\omega,k'}$, there exists some rational rotation vector $\omega^*\in\Gamma_{\omega,k'}$ such that $\lambda_\omega\omega$, $\lambda_{\omega^*}\omega^*$ lie on the same side, such that

$$\|\omega-\omega^*\|\le d\|\lambda_\omega\omega-\lambda_{\omega^*}\omega^*\|\le\frac{d\lambda_{\omega^*}}{KT},$$

where $T>0$ is the period of ω^* such that $\lambda_{\omega^*}^{-1}T<K$.

For the ball $B_D\subset\mathbb{R}^3$, there are positive numbers $m'=m'(D)\ge m=m(D)>0$ such that

$$m\|v\|^2\le\langle\partial^2h(y)v,v\rangle\le m'\|v\|^2,\qquad\forall\, y\in B_D,\ v\in\mathbb{R}^3.$$

Because h is assumed strictly convex, there exist exactly two points $y,y^*\in B_D$ such that $\partial h(y)=\omega$, $\partial h(y^*)=\omega^*$ and $\|y-y^*\|\le m^{-1}\|\omega-\omega^*\|$. Because $\lambda_{\omega^*}^{-1}T\le K$, the covering property $\|y-y^*\|\le T^{-1}\epsilon^\kappa$ is guaranteed if we choose

$$K=\frac{d\lambda_{\omega^*}}{m}\epsilon^{-\kappa}.$$

Again, because $\lambda_{\omega^*}^{-1}T\le K$, one has $T\le K^*\epsilon^{-\frac13(1-3\kappa)}$ if $K=\frac{d\lambda_{\omega^*}}{m}\epsilon^{-\kappa}\le\frac{K^*}{\Lambda}\epsilon^{-\frac13(1-3\kappa)}$. It holds for each $\epsilon\in(0,\epsilon_0]$ if ϵ_0 satisfies the condition

$$\epsilon_0^{\frac{1-3\kappa}{3}-\kappa}\le\frac{mK^*}{d\Lambda^2},$$

where $\Lambda=\max_{\omega\in\Gamma_{\omega,k}}\lambda_\omega$. For $\kappa<\frac16$, such $\epsilon_0>0$ exists. □

5. The finiteness of strong double resonance. Because of Theorem 4.1, the path of resonance $\Gamma_{k'}$ is covered by the discs $\{\|p-p''_i\|<T_i^{-1}\epsilon^\kappa\}$, where $\kappa<\frac16$, $T_i\le K^*\epsilon^{-\frac13(1-3\kappa)}$ is the period of the double resonance at p''_i, K^* is independent of ϵ. Therefore, the size of each disk is between $O(\epsilon^{1/3})$ and $O(\epsilon^{1/7})$. When the number $\epsilon>0$ decreases, the number of the disks increases. We are going to distinguish strong double resonant points from weak resonant points by the second resonant relation, and we will see that the number of strong double resonant points is finite, independent of ϵ in generic case.

We consider the resonant term (3.3) which takes the form

$$Z(p,q) = Z_{k'}(p,\langle k',q\rangle) + Z_{k',k''_i}(p,\langle k',q\rangle,\langle k''_i,q\rangle) \tag{5.1}$$

where $k' = k^\star$ or k^*, k''_i is the additional resonant condition

$$\begin{aligned} Z_{k'} &= \sum_{j\in\mathbb{Z}\setminus\{0\}} P_{jk'}(p)e^{j\langle k',q\rangle i}, \\ Z_{k',k''_i} &= \sum_{(j,l)\in\mathbb{Z}^2,l\neq 0} P_{jk'+lk''_i}(p)e^{(j\langle k',q\rangle+l\langle k''_i,q\rangle)i}. \end{aligned} \tag{5.2}$$

Because $|P_k|$ decrease fast as $\|k\|$ increases $|P_k|\le O(\|k\|^{-r})$, the term Z_{k',k''_i} is treated as a small perturbation to $Z_{k'}$ provided $\|k''_i\|$ is large.

Treated as a function of $x=\langle k',q\rangle$, we consider $Z_{k'}(p,x)$ as a family of functions $Z_{k'}(p,\cdot)\colon \mathbb{T}\to\mathbb{R}$, where $p\in\Gamma_{k'}$ is treated as a parameter. Such an observation allows us to apply the result of [Zh2].

Theorem 5.1 (Theorem 1.1 of [Zh2]). *Let $F_\lambda:\mathbb{T}\to\mathbb{R}$ be a family of C^4-smooth functions so that F_λ is Lipschitz in the parameter $\lambda\in[0,1]$. Then, there exists an open-dense set $\mathfrak{V}\subset C^r(\mathbb{T},\mathbb{R})$ $(r\ge 4)$ such that for each $V\in\mathfrak{V}$ and each $\lambda\in[0,1]$, every global minimal point of $F_\lambda - V$ is non-degenerate.*

To apply the theorem here, we notice that the path $\Gamma_{k'}$ induces a decomposition

$$C^r(B_D\times\mathbb{T}^3,\mathbb{R}) = C^r(B_D\times\mathbb{T},\mathbb{R})\oplus C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(B_D\times\mathbb{T},\mathbb{R})$$

via

$$P(p,q) = Z_{k'}(p,\langle k',q\rangle) + P'(p,q), \quad (k'=k^\star,k^*),$$

where $Z_{k'}$ is defined in (5.2) consisting of Fourier modes of P in $\mathrm{span}_{\mathbb{Z}}\{k'\}$, and $P' = P - Z_{k'}\in C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(B_D\times\mathbb{T},\mathbb{R})$.

Therefore, there exists an open-dense set $\mathfrak{V}\subset C^r(B_D\times\mathbb{T},\mathbb{R})$ and consequently an open-dense set $\mathfrak{P} = \mathfrak{V}\oplus C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(B_D\times\mathbb{T},\mathbb{R})\subset C^r(B_D\times\mathbb{T}^3)$ such that for all $P\in\mathfrak{P}$, it holds simultaneously for each $p\in\Gamma_{k'}$ that the resonant term $Z_{k'}(p,\cdot)$, treated as a function of x, is non-degenerate at its maximal point, namely, the second derivative $\partial_x^2 Z_{k'}$ at its maximum is uniformly upper bounded below 0.

Let y_0 be a vector such that $By_0=(0,1)^t$. A non-degenerate maximal point $x_{1,i}$ of $Z_{k'}(p''_i,\cdot)$ corresponds to a normally hyperbolic invariant cylinder (NHIC)

$$\Pi_i = \{(x,y)\in\mathbb{T}^2\times\{\|y\|\le K_i^{-1}\epsilon^{\kappa-\frac12}\} : x_1 = x_{1,i}, y=\lambda y_0, \lambda\in\mathbb{R}\}$$

of the Hamiltonian system $\frac12\langle By,y\rangle + Z_{k'}(p''_i,x_1)$. In the (p,q)-coordinates, the cylinder passes through a neighborhood of the double resonant point p''_i. Such a phenomenon allows us to distinguish strong double resonant points from weak ones by the existence of weakly invariant cylinder.

Definition 5.1 (cf. [B10]). An open manifold is said to be weakly invariant for a flow if its vector field is tangent to the manifold.

Applying the theorem of normally hyperbolic invariant manifold (NHIM), we see that, for the Hamiltonian $\frac{1}{2}\langle By,y\rangle+Z_{k'}+Z_{k',k''_i}$, there exists some weakly invariant cylinder Π'_i lying in a small neighborhood of $\Pi_i\cap\{\|y\|\le K_i^{-1}\epsilon^{\kappa-\frac{1}{2}}-1\}$ provided $|k''_i|$ is large enough.

Indeed, if we introduce a cut-off C^∞-function χ: $[0,\infty)\to\mathbb{R}$ satisfying $\chi(\nu)=0$ for $\nu\ge K_i^{-1}\epsilon^{\kappa-\frac{1}{2}}$ and $\chi(\nu)=1$ for $\chi\le K_i^{-1}\epsilon^{\kappa-\frac{1}{2}}-1$. Then by applying the theorem of NHIM to the Hamiltonian $\frac{1}{2}\langle By,y\rangle+Z_{k'}(p''_i,x_1)+\chi(\|y\|)Z_{k',k''_i}(p''_i,x)$ we see that the cylinder Π_i survives the small perturbation $\chi(\|y\|)Z_{k',k''_i}(p''_i,x)$ if $|k''_i|$ is sufficiently large. Restricted on the region $\{\|y\|\le K_i^{-1}\epsilon^{\kappa-\frac{1}{2}}-1\}$ the survived cylinder is obviously weakly invariant. Since the normally hyperbolic splitting still exists on the survived cylinder, we call it normally hyperbolic weakly invariant cylinder, or NHWIC for short.

Definition 5.2. A double resonance is said to be *weak* if the double resonant term Z_{k',k''_i} is so small that can be treated as a small perturbation

$$\frac{1}{2}\langle By,y\rangle+Z_{k'}\to\frac{1}{2}\langle By,y\rangle+Z_{k'}+Z_{k',k''_i}$$

such that one can apply the theorem of NHIM. In this case, the NHWIC survives the perturbation. Otherwise, the double resonance is said to be *strong*.

Although the number of the disks $\{\|p-p''_i\|<T_i^{-1}\epsilon^\kappa\}$ depends on ϵ, we have

Proposition 5.1. *There exists a set $\mathfrak{P}$ open-dense in $\mathfrak{S}_1$ such that for each $P\in\mathfrak{P}$, the number of strong double resonances along $\Gamma_{k'}$ is finite and independent of ϵ.*

Proof. For perturbation P, let x_p be the maximal point of the single resonant term $Z_{k'}(p,\cdot)$ with respect to the variable x. By Theorem 5.1, there exists an open-dense set $\mathfrak{P}'\subset C^r(B_D\times\mathbb{T}^3,\mathbb{R})$, for each $P\in\mathfrak{P}'$ there exists some $d_P>0$ such that the second derivative $\partial_x^2Z_{k'}(p,x_p)<-d_P$ holds for all $p\in\Gamma_{k'}$. Since $\lambda P\in\mathfrak{P}'$ for any $\lambda\ne0$ if $P\in\mathfrak{P}'$, the restriction of $\mathfrak{P}'$ to $\mathfrak{S}_1$ is clearly open-dense. Since the double resonant term Z_{k',k''_i} will be sufficiently small provided $\|k''_i\|$ is sufficiently large, we complete the proof. □

Notation 5.1. *Let Λ_ϵ denote the set of the subscripts i such that the resonant circles $\Gamma_{k^\star}\cup\Gamma_{k^*}$ is covered by the disks*

$$\Gamma_{k^\star}\cup\Gamma_{k^*}\subset\cup_{i\in\Lambda_\epsilon}\{p\in\mathbb{R}^3:\|p-p''_i\|<T_i^{-1}\epsilon^\kappa\}$$

where the period T_i of the double resonant point p''_i is not larger than $K^*\epsilon^{-(1-3\kappa)/3}$ with $\kappa\in(0,\frac{1}{6})$. Let $\Lambda_s\subset\Lambda_\epsilon$ be the subset so that the double resonance at p''_i is strong if and only if $i\in\Lambda_s$.

6. Dynamics around double resonance. Let us apply Lemma 3.1 to the double resonant point p''_i, where the second resonant condition is denoted by k''_i, let $\bar{k}''_i=M_0^{-t}k''_i=(0,\bar{k}''_{i,2},\bar{k}''_{i,3})$. We denote by $G_{\epsilon,i}$ the normal form reduced from H in a neighborhood of p''_i, which takes the form of (3.2)

$$G_{\epsilon,i}(x,y,\theta)=\bar{G}_i(x,y)+R_{\epsilon,i}(x,y,\vartheta(\theta)),\tag{6.1}$$

where $\bar{G}_i = \frac{1}{2}\langle B_i y, y\rangle - V_i(x_1, |\bar{k}_i''|x_2)$, $V_i \in C^r$ is 2π-periodic in $(x_1, |\bar{k}_i''|x_2)$ such that $\max V_i = 0$, $\vartheta = \frac{\omega_{i,3}}{\sqrt{\epsilon}}\theta$, $R_{\epsilon,i} \in C^{r-2}(\mathbb{T}^2 \times \Sigma'_{\epsilon,i} \times |\bar{k}_i''|\mathbb{T}, \mathbb{R})$ with

$$\{|y| \le (1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac{1}{2}}\} \subset \Sigma'_{\epsilon,i} \subset \{|y| \le (1+\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac{1}{2}}\},$$

$\eta \in (0,1)$ is independent of the period T_i, the remainder $R_{\epsilon,i}$ is bounded by $a_0\epsilon^\kappa$ in C^{r-2}-topology if it is considered as a function of (x, y, ϑ).

Because the normal form (6.1) is obtained by the transformation (3.1) where the matrix is denoted by M_i, the function $R_{\epsilon,i}$ respects the symmetry M_i. So we have

$$G_{\epsilon,i}(x_1, x_2 + \bar{k}''_{i,3}/\bar{k}''_{i,2}, \vartheta + 1) = G_{\epsilon,i}(x_1, x_2, \vartheta) \tag{6.2}$$

Pull back to the original space of (p, q), we have $V_i(x_1, |\bar{k}_i''|x_2) = V_i(\langle k', q\rangle, \langle k_i'', q\rangle)$.

Let $\alpha_{\epsilon,i}$ and $\beta_{\epsilon,i}$ denote the α- and β-function for $G_{\epsilon,i}$ respectively, with which one defines the Fenchel-Legendre transformation $\mathscr{L}_\alpha$: $H_1(M,\mathbb{R}) \to H^1(M,\mathbb{R})$ by

$$c \in \mathscr{L}_\alpha(\omega) \quad \Longleftrightarrow \quad \alpha(c) + \beta(\omega) = \langle c, \omega\rangle.$$

By the definition, the β-function is the Fenchel-Legendre dual of the α-function. By adding a constant to $G_{\epsilon,i}$, we can assume $\min \alpha_{\epsilon,i} = 0$.

To understand the dynamics around strong double resonances, we apply the results obtained in [C17a, C17b, CZ].

Theorem 6.1. *There exists a residual set $\mathfrak{V}_i \subset C^r(\mathbb{T}^2, \mathbb{R})$ with $r \ge 2$. Each $V_i \in \mathfrak{V}_i$ is associated with some positive numbers Δ_{V_i}, ϵ_i such that for any $\xi \in (0, \Delta_{V_i})$, $\epsilon \in [0, \epsilon_i]$ the circle $\alpha^{-1}_{\epsilon,i}(\xi)$ establishes a transition chain (of cohomology equivalence). These circles make up an annulus $\mathbb{A}_i$ surrounding the set $\mathbb{F}_i = \{c : \alpha_{\epsilon,i}(c) = 0\}$.*

Proof. If we are satisfied with the property that $V \in \mathfrak{V}_i$ is 2π-periodic both in x_1 and in x_2, it is just an application of Theorem 1.1 of [C17b]. We are going to show that the set $\mathfrak{V}_i$ is residual in the space of C^r-functions which are 2π-periodic in $(x_1, |\bar{k}_i''|x_2)$. For the purpose, we introduce a canonical transformation $\mathscr{T}$: $(x_1, |\bar{k}_i''|x_2) \to (\phi_1, \phi_2)$, $(y_1, |\bar{k}_i''|^{-1}y_2) \to (I_1, I_2)$ and get the Hamiltonian from (6.1)

$$\bar{G}'_i = \mathscr{T}_*\bar{G}'_i = \frac{1}{2}\langle B'_i I, I\rangle - V_i(\phi), \qquad (\phi, I) \in \mathbb{T}^2 \times \mathbb{R}^2, \tag{6.3}$$

where $B'_i = \mathrm{diag}(1, |\bar{k}_i''|)B_i\mathrm{diag}(1, |\bar{k}_i''|)$. The theorem holds for $\bar{G}'_i$ under the condition that $\mathfrak{V}_i$ is residual in $C^r(\mathbb{T}^2, \mathbb{R})$.

Pull $\bar{G}'_i$ back to the space of (x, y), we see that the theorem holds because $G_{\epsilon,i}$ is a small perutrbation of $\mathscr{T}^*\bar{G}'_i$: $G_{\epsilon,i} = \mathscr{T}^*\bar{G}'_i + R_{\epsilon,i}$ and Mañé set is upper semi-continuous with respect to small perturbation. □

Given an irreducible class $g \in H_1(\mathbb{T}^2, \mathbb{Z})\backslash\{0\}$, let $\mathbb{C}_{i,g} = \cup_{\nu>0}\mathscr{L}_{\alpha_{\epsilon,i}}(\nu g)$. To consider the Aubry set of $G_{\epsilon,i}$ for $c \in \mathbb{C}_{i,g}$, we consider the truncated Hamiltonian $\bar{G}'_i$ of (6.3) and let $\bar{\alpha}_i$ be the α-function for $\bar{G}'_i$. For $c \in \mathscr{L}_{\bar{\alpha}_i}(\lambda g)$, each c-minimal orbit is periodic. To stress its topological information, we also call it λg-minimal orbit. The parameter λ_ℓ is called bifurcation point, if there are two or more $\lambda_\ell g$-minimal orbits. Applying Theorem 2.1 of [CZ], Theorem 3.1 and the argument of Section 4 of [C17a] we have

Proposition 6.1. *Given an irreducible class $g' \in H_1(\mathbb{T}^2, \mathbb{Z})\backslash\{0\}$ and $\lambda_0 > 0$, there exists an open-dense set $\mathfrak{V}_i \subset C^r(\mathbb{T}^2, \mathbb{R})$ with $r \ge 5$ such that for each $V_i \in \mathfrak{V}_i$,* *it holds simultaneously for all* $\lambda\in[\lambda_0,\infty)$ *that the Mather set of* $\bar G_i'$ *for each* $c\in\mathscr{L}_{\bar\alpha_i}(\lambda g')$ *consists of hyperbolic periodic orbits. Indeed, except for finitely many* $\{\lambda_\ell\}$ *the Mather set is made up two hyperbolic periodic orbits, for all other* $\lambda\in[\lambda_0,\infty)$ *the Mather set contains exactly one hyperbolic periodic orbit.*

Therefore, for each $V_i\in\mathfrak{V}_i$, there are finitely many bifurcation points $\lambda_0<\lambda_1'<\cdots<\lambda_m'$, the number m is independent of ϵ. At each bifurcation point λ_ℓ', there exist exactly two $\lambda_\ell' g'$-minimal orbits of $\bar G_i'$. So, for $\lambda'\ge\lambda_0$, all $\lambda' g'$-minimal orbits make up $m+1$ pieces of NHIC. Let $\bar\Pi_{i,\ell}'$ denote the cylinder made up of $\lambda' g'$-minimal orbits of $\bar G_i$ for all $\lambda'\in[\lambda_\ell',\lambda_{\ell+1}']$ for $\ell<m$, let $\bar\Pi_{i,m}'$ be the cylinder made up of $\lambda' g'$-minimal orbits of $\bar G_i$ for all $\lambda'\in[\lambda_m',\lambda_\epsilon']$ where λ_ϵ' is chosen such that $\max_\theta|(I_1,|\bar k_i''|I_2)(\theta)|=(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}$ holds for the $\lambda_\epsilon' g'$-minimal orbits $(I(\theta),\phi(\theta))$.

Return back to the coordinates (x,y), each orbit of $\bar G_i'$ is pulled back to the orbit of $\mathscr{T}^*\bar G_i'$ in the way

$$(x_1,x_2,y_1,y_2)=\Big(\phi_1,\frac{1}{|\bar k_i''|}(\phi_2+2\ell\pi),I_1,|\bar k_i''|I_2\Big)$$

for $\ell=0,1,\cdots,|\bar k_i''|-1$. Consequently, each cylinder $\bar\Pi_{i,\ell}'$ is pulled back to a cylinder $\bar\Pi_{i,\ell}$ of $\bar G_i$, modulo a shift $(x,y)\to(x+(0,2\pi/|\bar k_i''|),y)$. Each cylinder is made up of λg-minimal orbits of $\bar G_i$ with $\lambda g=\frac{\lambda'}{|\bar k_i''|}(|\bar k_i''|g_1,g_2)$ if $\lambda' g'=\lambda'(g_1,g_2)$.

Let E_ℓ be the energy such that $\lambda_\ell g$-minimal orbit lies in the energy level set $\bar G_i^{-1}(E_\ell)$. Since each $\lambda_\ell g$-minimal orbit is hyperbolic, each cylinder $\bar\Pi_{i,\ell}$ can be extended by the hyperbolic orbits lying in the level set $\bar G_i^{-1}(E)$ with $E\in[E_{\ell+1},E_{\ell+1}+2d]\cup[E_\ell-2d,E_\ell]$. Since there are finitely many bifurcation points, such a number $d>0$ exists.

Under the time-periodic perturbation $G_{\epsilon,i}=\bar G_i+R_{\epsilon,i}$, major part of these cylinders survives, *weakly invariant* for the Hamiltonian flow $\Phi^\theta_{G_{\epsilon,i}}$. Notice that $\vartheta=\omega_{i,3}\sqrt{\epsilon}^{-1}\theta$ and $G_{\epsilon,i}$ is 2π in x, $2|\bar k_i''|\pi$ in ϑ and symmetric for M_i, see (6.2), we introduce a shift

$$\sigma_i:\ (x,y,\theta)\to\Big(x+\Big(0,2\pi\frac{\bar k_{i,3}''}{\bar k_{i,2}''}\Big),y,\theta+\frac{\sqrt{\epsilon}}{\omega_{i,3}}\Big),\tag{6.4}$$

then $\sigma_i^*G_{\epsilon,i}=G_{\epsilon,i}$. Let $\tilde\Pi_{i,E_\ell-d,E_{\ell+1}+d}=(\bar\Pi_{i,\ell}\times\frac{|\bar k_i''|\sqrt{\epsilon}}{\omega_{i,3}}\mathbb{T})\cap\{\bar G_i\in[E_\ell-d,E_{\ell+1}+d]\}$.

Proposition 6.2. *For sufficiently small* ϵ, *there is a cylinder* $\tilde\Pi^\epsilon_{i,E_\ell-d,E_{\ell+1}+d}$ *modulo the shift* σ_i, *which is weakly invariant and normally hyperbolic for the flow* $\Phi^\theta_{G_{\epsilon,i}}$. *The cylinder lies in a small neighborhood of* $\tilde\Pi_{i,E_\ell-d,E_{\ell+1}+d}$.

Proof. We modify the Hamiltonian $G_{\epsilon,i}$. Let ρ be a C^2-function such that $\rho(\mu)=1$ for $\mu\ge1$ and $\rho(\mu)=0$ for $\mu\le0$. By defining $\rho_1(x,y)=\rho((\bar G_i(x,y)-E_\ell^-+2d)/d)$, $\rho_2(x,y)=1-\rho((\bar G_i(x,y)-E_{\ell+1}-d)/d)$ we introduce

$$G_{\epsilon,i}'=\begin{cases}\bar G_i+\rho_1R_{\epsilon,i}, & \text{if } \bar G_i(x,y)\in[E_\ell^--2d,E_\ell-d],\\ \bar G_i+\rho_2R_{\epsilon,i}, & \text{if } \bar G_i(x,y)\in[E_{\ell+1}+d,E_{\ell+1}+2d],\\ G_{\epsilon,i}, & \text{elsewhere.}\end{cases}\tag{6.5}$$

Because $\|G_{\epsilon,i}'-\bar G_i\|_{C^2}\ll1$ if $|y|\le O(\epsilon^{\kappa-\frac12})$ and $\epsilon\ll1$, it follows that the NHIC $\tilde\Pi_{i,\ell}$ survives the perturbation $\bar G_i\to G_{\epsilon,i}'$, denoted by $\tilde\Pi^\epsilon_{i,\ell}$.

Let $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}=\tilde{\Pi}^{\epsilon}_{i,\ell}\cap\{\bar{G}_i\in[E_\ell-d,E_{\ell+1}+d]\}$, which is a weakly invariant for $\Phi^{\theta}_{G_{\epsilon,i}}$ because $G_{\epsilon,i}=G'_{\epsilon,i}$ when they are restricted in the region $\{(x,y):\bar{G}_i(x,y)\in[E_\ell-d,E_{\ell+1}+d]\}\times\frac{|\bar{k}''_i|\sqrt{\epsilon}}{\omega_{i,3}}\mathbb{T}$. The normal hyperbolicity is obvious.

Due to the symmetry (6.2), the Hamiltonian vector field of $\Phi^{\theta}_{G_{\epsilon,i}}$ is invariant under the shift σ_i. It guarantees the invariance of $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ under the shift σ_i. □

Notice that $\lambda_\epsilon>0$ is set such that $\max_\theta|y(\theta)|=(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac{1}{2}}$ holds for the $\lambda_\epsilon g$-minimal orbits $(y(\theta),x(\theta))$. By applying Theorem 1.2 of [C17b], we have

THEOREM 6.2. *Given a class $g\in H_1(\mathbb{T}^2,\mathbb{R})$ and small $E_0>0$, there is an open-dense set $\mathfrak{V}_i\subset C^r(\mathbb{T}^2,\mathbb{R})$ $(r\geq 5)$. For each $V_i\in\mathfrak{V}_i$, there exists $\epsilon_0>0$ such that for each $\epsilon\in(0,\epsilon_0)$*

1) *there are finitely many NHWICs for the flow $G_{\epsilon,i}$: $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ $(\ell=0,\cdots,m)$ and $\tilde{\Pi}^{\epsilon}_{i,E_m-d,E_\epsilon}$, modulo the shift σ_i, where the integer m, the numbers $E_1<E_2<\cdots<E_m$, $d>0$ and the normal hyperbolicity of each cylinder are all independent of ϵ, $E_\epsilon=\bar{\alpha}_i(\lambda_\epsilon g)=O(\epsilon^{2\kappa-1})$ for small $\epsilon>0$;*

2) *some $E_{\ell,\epsilon}\to E_\ell$ as $\epsilon\to 0$ exists such that for each $c\in\mathbb{C}_{i,g}$*

(1) *if $\alpha_{\epsilon,i}(c)\in(E_{\ell,\epsilon},E_{\ell+1,\epsilon})$, the Aubry set lies on $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ modulo σ_i;*

(2) *if $\alpha_{\epsilon,i}(c)=E_{\ell,\epsilon}$, the Aubry set contains at least two connected components, one is on $\tilde{\Pi}^{\epsilon}_{i,E_{\ell-1}-d,E_\ell+d}$ and the other one is on $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ modulo σ_i;*

(3) *if $\alpha_{G_\epsilon}(c)\in(E_{m,\epsilon},E_\epsilon)$, the Aubry set lies on $\tilde{\Pi}_{E_m-d,E_\epsilon}$ modulo σ_i.*

REMARK. Because of the symmetry M_i, all shifts of the cylinder $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ are projected to the same cylinder, if we pull them back to the space of (p,q).

At a strong double resonant point p''_i, $i\in\Lambda_s$, we consider two classes g_i^-,g_i^+ which determines two channels $\mathbb{C}_i^{\pm}=\mathbb{C}_{i,g^{\pm}}$. The choice of g^-,g^+ depends on where p''_i is. There are three possible locations of a strong double resonant point p''_i along $\Gamma_{k^*}\cup\Gamma_{k^\star}$:

(1) the point p''_i is on the path $\Gamma_{k^\star}$ where $\Gamma_{k^\star}$ does not intersect with Γ_{k^*};
(2) the point p''_i is on the path Γ_{k^*} where Γ_{k^*} does not intersect with $\Gamma_{k^\star}$;
(3) $p''_i\in\Gamma_{k^\star}\cap\Gamma_{k^*}$.

After the linear coordinate transformations $\mathscr{M}_0$ and $\mathscr{M}_i$ defined in (3.1), we have that the first two components of the frequency $\partial h(p)$, i.e. the frequency of the system $G_{\epsilon,i}$, is proportional to $(0,1)$ along Γ_k, $k=k^*,k^\star$ in the first two cases. In the third case, $\partial h(p)$ is proportional to $(0,1)$ along $\Gamma_{k^\star}$ and to $(1,0)$ along Γ_{k^*}.

At a double resonance p''_i, the first two components of the frequency $\partial h(p''_i)$ vanish after the linear transforms. So by *crossing a strong double resonance*, we mean that there exists an orbit of the system $G_{\epsilon,i}$, such that along the orbit, its “frequency” changes from being in the set $\{(0,\nu),\ \nu>0\}$ to the set $\{(0,\nu),\ \nu<0\}$ in the first two cases, and changes from being in the set $\{(0,\nu),\ \nu>0\}$ to the set $\{(\nu,0),\ \nu>0\}$ in the third case. Since $G_{\epsilon,i}$ is not nearly-integrable, here the role of “frequency” of the system $G_{\epsilon,i}$ will be played by the rotation vector of the Mather sets shadowed by the orbit.

We introduce the homology classes $g_i^{\pm}\in H_1(\mathbb{T}^2,\mathbb{Z})$ as follows:

(1) $g_i^+=(0,1)$ and $g_i^-=(0,-1)$ in the first two cases,
(2) $g_i^+=(0,1)$ and $g_i^-=(1,0)$ in the third case.

By adding a constant to $G_{\epsilon,i}$ we assume $\min\alpha_{\epsilon,i}=0$. Applying Theorem 6.1 and 6.2 we obtain the following result for the dynamics around the strong double resonance.

THEOREM 6.3. *There exists an open-dense set* $\mathfrak{V}_i \subset C^r(\mathbb{T}^2,\mathbb{R})$ $(r \geq 5)$ *such that for each* $V_i \in \mathfrak{V}_i$*, there exist* $0 < E_{i,0} < \Delta_{V_i}$, $\Delta'_i > 0$ *and* $\epsilon_i > 0$ *such that* $\forall\ \epsilon \in (0, \epsilon_i]$ *the following holds.*

1) *there is an annulus* $\mathbb{A}_i$ *around a double resonant point* p''_i*, made up of the circles* $\{c \in \alpha_{\epsilon,i}^{-1}(E) : E \in (0, \Delta_{V_i})\}$*, each of which is a path of cohomology equivalence;*

2) *the following two channels are connected by the annulus* $\mathbb{A}_i$

$$\mathbb{C}_i^{\pm} = \cup_{\lambda \in [\lambda_i^{\pm}, \bar{\lambda}_i^{\pm}]} \mathscr{L}_{\alpha_{\epsilon,i}}(\lambda g_i^{\pm}), \qquad g_i^{\pm} \in H_1(\mathbb{T}^2, \mathbb{Z}),$$

where $\bar{\lambda}_i^{\pm}, \lambda_i^{\pm} > 0$ *satisfy* $\alpha_{\epsilon,i}(\mathscr{L}_{\alpha_{\epsilon,i}}(\lambda_i^{\pm} g_i^{\pm})) = E_{i,0} < \Delta_{V_i}$ *and*

(1) *for* $c \in \mathscr{L}_{\alpha_{\epsilon,i}}(\bar{\lambda}_i^{\pm} g_i^{\pm})$ *it holds for every orbit* $\{(x(\theta), y(\theta)),\ \theta \in \mathbb{R}\}$ *in* $\tilde{\mathcal{A}}(c)$ *that*

$$(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac{1}{2}} - \Delta'_i \leq |y(\theta)| \leq (1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac{1}{2}};$$

(2) *for each* $c \in \mathbb{C}_i^{\pm}$ *except for finitely many classes* $\{c_j\}$*, the Aubry set* $\tilde{\mathcal{A}}(c)$ *lies in certain cylinder* $\tilde{\Pi}^{\epsilon}_{i,E_\ell - d, E_{\ell+1}+d}$*, modulo the shift* σ_i*, while* $\tilde{\mathcal{A}}(c_j)$ *lies in two of the cylinders. For each orbit* $\{(x(\theta), y(\theta)),\ \theta \in \mathbb{R}\}$ *in the Aubry set one has*

$$\mathrm{Osc}|y(\theta)| = \max |y(\theta) - y(\theta')| \leq \Delta'_i.$$

Proof. We only need to verify the estimate on the oscillation of $y(\theta)$ if $(x(\theta), y(\theta))$ is an orbit in the Aubry sets. Let us consider the problem for $\bar{G}_i$ first.

We claim that some constant $\Delta'_i > 0$ exists such that it holds along any λg-minimal orbit $(\bar{y}(\theta), \bar{x}(\theta))$ of $\bar{G}_i$ that

$$|\bar{y}(\theta) - \bar{y}(\theta')| \leq \frac{1}{2}\Delta'_i. \tag{6.6}$$

Indeed, each λg-minimal orbit is entirely contained in certain energy level set $\bar{G}_i^{-1}(E)$. The higher the energy increases, the shorter the period becomes. Let E_1 be the energy that the set $\bar{G}_i^{-1}(E_1)$ contains λg-minimal orbit with period 1. One obtains from the equation $\dot{y} = \partial V_i(x)$ that $|\bar{y}(\theta) - \bar{y}(\theta')| \leq \max_{x \in \mathbb{T}^2}\{|\partial_{x_1} V_i(\bar{x}(\theta))|, |\bar{k}''_i||\partial_{x_2} V_i(\bar{x}(\theta))|\}$ if $(x(\theta), y(\theta)) \in \bar{G}_i^{-1}(E)$ with $E \geq E_1$. Let m_i be the smallest eigenvalue of B_i, one has $\|y\| \leq \frac{1}{m_i}\sqrt{E + V_i(x)}$ if $(x, y) \in \bar{G}_i(E)$. Therefore, the estimate (6.6) holds if we set

$$\Delta'_i = \max\Big\{2\max_{x \in \mathbb{T}^2}\{|\partial_{x_1} V_i(\bar{x}(\theta))|, |\bar{k}''_i||\partial_{x_2} V_i(\bar{x}(\theta))|\}, \frac{4}{m_i}\sqrt{E_1 + \max_x V_i(x)}\Big\}. \tag{6.7}$$

Next, we consider $G_{\epsilon,i}$ which is a $O(\epsilon^{\kappa})$-perturbation of $\bar{G}_i$. Each θ-section of NHIC of $G_{\epsilon,i}$ is located in $O(\epsilon^{\kappa})$-neighborhood of NHIC of $\bar{G}_i$. By Proposition 5.2 of [C17a], the Aubry set $\tilde{\mathcal{A}}(c)$ does not hit the level set $G_{\epsilon,i}^{-1}(E \pm \epsilon^{\frac{1}{3}\kappa})$ if $c \in \alpha_{\epsilon,i}^{-1}(E) \cap \cup_{\lambda}\mathscr{L}_{\alpha_{\epsilon,i}}(\lambda g)$ and $E - \min \alpha_{\epsilon,i} \geq O(1)$. Because the Aubry set $\tilde{\mathcal{A}}(c)$ for $G_{\epsilon,i}$ stays in the cylinder, it falls into $O(\epsilon^{\frac{1}{3}\kappa})$-neighborhood of the Aubry set for $\bar{G}_i$. For any $y(\theta)$ there exists $\bar{y}(\theta^*)$ such that $|y(\theta) - \bar{y}(\theta^*)| < \frac{1}{4}\Delta'_i$ provided $\epsilon > 0$ is suitably small and it holds for any orbit $(x(\theta), y(\theta)) \in \tilde{\mathcal{A}}(c)$ with $c \in \mathbb{C}_i^{\pm}$ that

$$\mathrm{Osc}|y(\theta)| = \max_{\theta,\theta'} |y(\theta) - y(\theta')| \leq \max_{\theta,\theta'} |\bar{y}(\theta) - \bar{y}(\theta')| + 2O(\epsilon^{\frac{1}{3}\kappa}) \leq \frac{3}{4}\Delta'_i.$$

If we set $\bar\lambda_i^{\pm}$ such that $\max_\theta|\bar y(\theta)|=(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}-\frac14\Delta'_i$ holds along $\bar\lambda_i^{\pm}g_i^{\pm}$-minimal orbit of $\bar G_i$, one obtains that

$$(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}-\Delta'_i\le|y(\theta)|\le(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}$$

if $(x(\theta),y(\theta))$ lies in $\tilde{\mathcal A}(c)$ with $c\in\mathscr L_{\alpha_{\epsilon,i}}(\bar\lambda_i^{\pm}g_i^{\pm})$. ◻

For weak double resonant point p''_i, the truncated part of $G_{\epsilon,i}$ is

$$\bar G_i=\frac12\langle B_iy,y\rangle-V'_i(x_1)-V''_i(x_1,|\bar k''_i|x_2), \tag{6.8}$$

where $V'_i(x_1)=-Z_{k'}(p''_i,x_1)$, $V''_i(x_1,|\bar k''_i|x_2)=-Z_{k',k''_i}(p''_i,x_1,|\bar k''_i|x_2)$ and the term V''_i is treated as a small perturbation.

Theorem 6.4. *There exists an open-dense set $\mathfrak V_{k'}\subset C^r(B_d\times\mathbb T,\mathbb R)$ $(r\ge5)$, for each $Z_{k'}\in\mathfrak V_{k'}$, there exist $\Delta'_i>0$ and $\epsilon_{k'}>0$ which are independent of k''_i such that for each $\epsilon\in(0,\epsilon_{k'}]$ and each $i\in\Lambda_\epsilon\backslash\Lambda_s$ there is a channel*

$$\mathbb C_i^w=\cup_{\lambda\in[-\lambda_i,\bar\lambda_i]}\mathscr L_{\alpha_{\epsilon,i}}(\lambda g)\subset H^1(\mathbb T^2,\mathbb R),\qquad g=(0,1) \tag{6.9}$$

with the properties that

(1) *for each $c\in\mathbb C_i^w$, the Aubry set $\tilde{\mathcal A}(c)$ lies on some NHWIC entirely contained in the region $\Sigma'_{\epsilon,i}$, modulo the shift σ_i. Along each orbit $\{(x(\theta),y(\theta)),\ \theta\in\mathbb R\}$ in the Aubry set one has*

$$\mathrm{Osc}|y(\theta)|=\max_{\theta,\theta'}|y(\theta)-y(\theta')|\le\Delta'_i,$$

(2) *the numbers λ_i and $\bar\lambda_i>0$ are chosen such that for $c\in\mathscr L_{\alpha_{\epsilon,i}}(\bar\lambda_ig)\cup\mathscr L_{\alpha_{\epsilon,i}}(-\lambda_ig)$ it holds for c-minimal orbit $\{(x(\theta),y(\theta)),\ \theta\in\mathbb R\}$ that*

$$(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}-\Delta'_i\le|y(\theta)|\le(1-\delta')(\eta T_i)^{-1}\epsilon^{\kappa-\frac12}.$$

Proof. According to Theorem 5.1, there exists an open-dense set $\mathfrak V_{k'}\subset C^r(B_d\times\mathbb T,\mathbb R)$, for each $Z_{k'}\in\mathfrak V_{k'}$ it holds simultaneously for all $p\in\Gamma_{k'}$ that the maximal point of $Z_{k'}(p,\cdot)$ in x is non-degenerate, namely, the second derivative of $Z_{k'}(p,\cdot)$ in x at the maximal point is uniformly upper-bounded below zero for all $p\in\Gamma_{k'}$.

In this case, the Hamiltonian system $\frac12\langle By,y\rangle-V'_i(x_1)$ admits a normally hyperbolic invariant cylinder made up of the minimal periodic orbits of type $(0,1)$. As the second resonant term $V''_i(x_1,|\bar k''_i|x_2)$ and the remainder $R_{\epsilon,i}$ are small, the weakly invariant cylinder survives the perturbation $\frac12\langle By,y\rangle-V'_i\to\frac12\langle By,y\rangle-V'_i-V''_i+R_{\epsilon,i}$.

To study the oscillation of $y(t)$, we also consider the truncation $\bar G_i=\frac12\langle By,y\rangle-V'_i(x_1)-V''_i(x_1,|\bar k''_i|x_2)$ first. Let $(\bar x(\theta),\bar y(\theta))$ be the λg-minimal orbit, then the second component of $\bar y(\theta)$ satisfies the following relation

$$\bar y_2(\theta')-\bar y_2(\theta)=\int_\theta^{\theta'}|\bar k''_i|\partial_2V''_i(x_1(\theta),|\bar k''_i|x_2(\theta))d\theta.$$

Although the number $|\bar k''_i|$ approaches infinity if the second resonant condition becomes weaker, it does not make trouble to control the oscillation of $\bar y(\theta)$. Indeed, the term $V''_i(x_1,|\bar k''_i|x_2)=Z_{k'_i,k''_i}(p_i,x_1,|\bar k''_i|x_2)$, which is defined in (5.2). Because $|P_k|$ decrease fast as $\|k\|$ increases: $|P_k|\le O(\|k\|^{-r})$, one has $|\bar k''_i||\partial_2V''_i|\to0$ as $\|\bar k''_i\|\to\infty$.

The rest of the proof is similar to the proof of Theorem 6.3. ◻

7. The estimate of the deviation of Aubry set. For integrable Hamiltonian h, the location of its c-minimal orbits is clear. Each c-minimal orbit is nothing else but the orbit $(p(t) = c, q(t) = \partial h(c)t + q_0)$. We want to know the deviation of Aubry set when h is under small perturbation $h \to H = p + \epsilon P$.

For nearly integrable Hamiltonian $H = h + \epsilon P$ with convex h, one has (see Formulae (4.3) and (4.4) of [C11])

$$|\tilde{\alpha}_H(\tilde{c}) - \tilde{\alpha}_h(\tilde{c})| < \epsilon\|P\|, \qquad |\tilde{\beta}_H(\rho) - \tilde{\beta}_h(\rho)| < \epsilon\|P\| \tag{7.1}$$

where $\|P\| = \max_{(p,q)\in B_D\times\mathbb{T}^3} |P(p,q)|$ denote the C^0-norm of P, $\tilde{\alpha}_H$ and $\tilde{\alpha}_h$ denote the α-function for H and h, $\tilde{\beta}_H$ and $\tilde{\beta}_h$ denote the β-function for H and h respectively.

Since $H^1(\mathbb{T}^3, \mathbb{R}) = \mathbb{R}^3$, we treat $\tilde{c} \in H^1(\mathbb{T}^3, \mathbb{R})$ as a point in $\mathbb{R}^3$. Since h is positive definite, there exists $m > 0$ such that

$$h(p') - h(p) \geq \langle \partial h(p), p' - p\rangle + \frac{m}{2}\|p' - p\|^2.$$

Lemma 7.1. *Let $p = \partial h(\omega) \in B_D$, then the set $\mathscr{L}_{\tilde{\alpha}_H}(\omega)$ falls into $C_s\sqrt{\epsilon}$-neighborhood of p with $C_s \leq 2\sqrt{\|P\|/m}$, namely*

$$\mathrm{dist}(\mathscr{L}_{\tilde{\alpha}_H}(\omega), p) \leq C_s\sqrt{\epsilon}.$$

Proof. Assume $\tilde{c} + p \in \mathscr{L}_{\tilde{\alpha}_H}(\omega)$, then by the definition one has

$$\begin{aligned}
\langle \omega, \tilde{c} + p\rangle &= \tilde{\beta}_H(\omega) + \tilde{\alpha}_H(\tilde{c} + p) \geq \tilde{\beta}_h(\omega) + \tilde{\alpha}_h(\tilde{c} + p) - 2\|P\|\epsilon \\
&\geq \tilde{\beta}_h(\omega) + \tilde{\alpha}_h(p) + \langle \omega, \tilde{c}\rangle + \frac{m}{2}\|\tilde{c}\|^2 - 2\|P\|\epsilon \\
&= \langle \omega, \tilde{c} + p\rangle + \frac{m}{2}\|\tilde{c}\|^2 - 2\|P\|\epsilon
\end{aligned}$$

from which one obtains $\|\tilde{c}\| \leq C_s\sqrt{\epsilon}$. ◻

Small perturbation may induce small rescaling of the rotation vector when both h and H are restricted on the level set with the same energy.

Lemma 7.2. *Given a rotation vector $\omega \neq 0$, let $h(\partial h^{-1}(\omega)) = \tilde{\alpha}_H(\mathscr{L}_{\tilde{\alpha}_H}(\nu\omega))$, then some constant $C_r > 0$ exists such that $|\nu - 1| \leq C_r\sqrt{\epsilon}$.*

Proof. Let ν be a number close to 1. For each rotation vector ω, $\exists$ unique p, p_ν such that $\omega = \partial h(p)$ and $\nu\omega = \partial h(p_\nu)$. It follows that

$$p - p_\nu = (1 - \nu)B_\nu^{-1}\omega, \qquad \|p - p_\nu\| \leq |1 - \nu|\|\omega\|m^{-1},$$

where $B_\nu = \partial^2 h(\lambda_\nu p + (1 - \lambda_\nu)p_\nu)$ is positive definite, $\lambda_\nu \in [0, 1]$. For $p_\nu + \tilde{c} \in \mathscr{L}_{\tilde{\alpha}_H}(\nu\omega)$, one obtains from the relation $|\tilde{\alpha}_H - \tilde{\alpha}_h| \leq \epsilon\|P\|$ and $\|\tilde{c}\| \leq C_s\sqrt{\epsilon}$ that

$$\begin{aligned}
\epsilon\|P\| &\geq |\tilde{\alpha}_H(p_\nu + \tilde{c}) - \tilde{\alpha}_h(p_\nu + \tilde{c})| = |h(p) - h(p_\nu + \tilde{c})| \\
&\geq |\langle \omega, p_\nu - p + \tilde{c}\rangle| - \frac{m'}{2}\|p - p_\nu - \tilde{c}\|^2 \\
&\geq |1 - \nu|\langle B_\nu^{-1}\omega, \omega\rangle - C_s\|\omega\|\sqrt{\epsilon} - \frac{m'}{2}\|p - p_\nu - \tilde{c}\|^2.
\end{aligned}$$

Therefore, some number $C_r = C_r(\omega, C_s) > 0$ exists such that $|1 - \nu| \leq C_r\sqrt{\epsilon}$. ◻

Let $\mathbb{F}_\omega$ denote the Fenchel-Legendre dual of a rotation vector ω, i.e. $\mathbb{F}_\omega = \mathscr{L}_{\tilde{\alpha}_H}(\omega)$. The following lemma establishes the location of $\tilde{\mathbb{C}}_k$, it lies in $O(\sqrt{\epsilon})$-neighborhood of Γ_k. The rescaling $\omega \to \nu\omega$ is bounded by Lemma 7.2.

LEMMA 7.3. *For $E > \min h$, there is a constant $C_H > 0$ such that $\forall\ p \in h^{-1}(E)$, $\omega = \partial h^{-1}(p)$, the set $\mathbb{F}_{\nu\omega} \subset \tilde{\alpha}_H^{-1}(E)$ lies in $C_H\sqrt{\epsilon}$-neighborhood of p, i.e. $\mathbb{F}_{\nu\omega} \subset B_{C_H\sqrt{\epsilon}}(p)$.*

Proof. By Lemma 7.1, one has $\mathbb{F}_{\nu\omega} \subset B_{C_s\sqrt{\epsilon}}(p_v)$. Since $\|p-p_\nu\| \le |1-\nu|\|\omega\|m^{-1}$, by Lemma 7.2 and setting $C_H = C_s + C_r \max_{p\in\Gamma_k} \|\omega(p)\|m^{-1}$, we finish the proof. □

LEMMA 7.4. *Some number $D_H > 0$ exists such that each orbit $(p(t), q(t))$ of Φ_H^t can not be $\tilde{c}$-minimal if $\|p(t) - \tilde{c}\| > D_H\sqrt{\epsilon}$ for all $t \in \mathbb{R}$.*

Proof. Let $L(\dot{q}, q)$ be the Lagrangian related to the Hamiltonian $H = h + \epsilon P$ through the Legendre transformation, then $L(\dot{q}, q) = \langle p, \dot{q}\rangle - H(p, q)$ where $\dot{q} = \partial_p H(p, q)$. If $(p(t), q(t))$ is $\tilde{c}$-minimal for $\tilde{c} \in \tilde{\alpha}_H^{-1}(E)$, one obtains from the identity $H(p(t), q(t)) \equiv \tilde{\alpha}_H(\tilde{c})$

$$L(\dot{q}(t), q(t)) - \langle \tilde{c}, \dot{q}(t)\rangle + \alpha_H(\tilde{c}) = \langle p(t) - \tilde{c}, \dot{q}(t)\rangle.$$

One obtains from Taylor's formula that for certain $\lambda \in [0, 1]$ the following holds

$$h(\tilde{c}) - h(p(t)) = \langle \tilde{c} - p, \partial h(p)\rangle + \frac{1}{2}\Big\langle \partial^2 h(\lambda\tilde{c} + (1-\lambda)(\tilde{c} - p))(\tilde{c} - p), (\tilde{c} - p)\Big\rangle.$$

Since h is positive definite and $\dot{q} = \partial_p h + \epsilon\partial P$, we get from the formula as above that

$$\begin{aligned}\langle p(t) - \tilde{c}, \dot{q}(t)\rangle &= \langle p(t) - \tilde{c}, \partial h(p)\rangle + \langle p(t) - \tilde{c}, \epsilon\partial_p P\rangle \\ &\ge \frac{m}{2}\|\tilde{c} - p(t)\|^2 - 2\epsilon\|P\| - \epsilon\|\partial_p P\|\end{aligned}$$

where $m > 0$ is the lower bound of the eigenvalues of $\partial^2 h$ for $p \in h^{-1}(E)$. We set

$$D_H > \sqrt{2m^{-1}(2\|P\| + \|\partial_p P\|)},$$

if $\|p(t) - \tilde{c}\| > D_H\sqrt{\epsilon}\ \forall\ t \in \mathbb{R}$, then $L(\dot{q}(t), q(t)) - \langle \tilde{c}, \dot{q}(t)\rangle + \tilde{\alpha}_H(\tilde{c}) > 0$ holds along the whole orbit $(p(t), q(t))$. It contradicts the minimality of the orbit. □

8. The construction of global transition chain. We come to the stage to construct a global transition chain that connects a small neighborhood of $\tilde{c}^\star = p^\star$ to a small neighborhood of $\tilde{c}^* = p^*$.

8.1. Invariance of the α-function. Let $\tilde{\alpha}_{\epsilon,i}$ be the α-function for $\tilde{G}_{\epsilon,i}$ with the form of (3.7). The isoenergetic reduction from systems with three degrees of freedom to two and half establishes the relation between $\tilde{\alpha}_{\epsilon,i}^{-1}(0)$ and the graph of $\alpha_{\epsilon,i}$:

THEOREM 8.1. *For the Hamiltonian $\tilde{G}_\epsilon(x, x_n, y, y_n)$, we assume that $\partial_{y_n}\tilde{G}_\epsilon \neq 0$ holds on $(\mathbb{T}^n \times B) \cap \{\tilde{G}_\epsilon^{-1}(E)\}$ where $E > \min\tilde{\alpha}_\epsilon$, $B \subset \mathbb{R}^n$ is a ball. Let $y_n = -\lambda G_\epsilon(x, y, t)$ be the solution of $\tilde{G}_\epsilon(x, \frac{1}{\lambda}t, y, -\lambda G_\epsilon) = E$ where $\lambda > 0$ is a real number. For a class $c \in H^1(\mathbb{T}^{n-1}, \mathbb{R})$, if the c-minimal curve $x(t)$ of G_ϵ satisfies the condition*

$$(x(t), \lambda^{-1}t, y(t), -\lambda G_\epsilon(x(t), y(t), t)) \in \mathbb{T}^n \times B, \qquad \forall\ t \in \mathbb{R}$$

then one has $\tilde{c} = (c, -\lambda\alpha_\epsilon(c)) \in \tilde{\alpha}_\epsilon^{-1}(E)$.

It was proved in [C17b] (Theorem 3.3 there). Let $\tilde{x}=(x,\lambda^{-1}t)$, $\tilde{y}=(y,-\lambda G_\epsilon)$ and τ denote the time of $\tilde{G}_\epsilon$, the theorem follows from the identity

$$\int\Big(\Big\langle\frac{dx}{dt},y-c\Big\rangle-G_\epsilon+\alpha_\epsilon(c)\Big)dt=\int\Big(\Big\langle\frac{d\tilde{x}}{d\tau},\tilde{y}-\tilde{c}\Big\rangle-\tilde{G}_\epsilon+E\Big)d\tau.$$

For the application of the theorem in the paper, one has $n=3$. If we regard the graph of $\alpha_{\epsilon,i}$ over $\mathbb{A}_i\cup\mathbb{C}_i^-\cup\mathbb{C}_i^+$ as a set in $\mathbb{R}^3$,

$$\{(c,-\alpha_{\epsilon,i}(c)):c\in\mathbb{A}_i\cup\mathbb{C}_i^-\cup\mathbb{C}_i^+\}$$

it precisely lies in the surface $\tilde{\alpha}_{\epsilon,i}^{-1}(0)$. The graph of $\mathbb{A}_i$, $\mathbb{C}_i^-$ and $\mathbb{C}_i^+$ are denoted by $\tilde{\mathbb{A}}_i$, $\tilde{\mathbb{C}}_i^-$ and $\tilde{\mathbb{C}}_i^+$ respectively. Formula (3.1) induces a linear transformation in $H^1(\mathbb{T}^3,\mathbb{R})$ under which the sphere $\tilde{\alpha}_\epsilon$ undergoes a linear transformation

$$\Psi_i^\ell:\ \tilde{c}\to M_i\tilde{c}. \tag{8.1}$$

Let $\tilde{\alpha}_{\Phi^*_{\epsilon F_i}H}$ be the α-function for the Hamiltonian $\Phi^*_{\epsilon F_i}H$, where ϵF_i is the generating function for the KAM iteration. Because the rescaling (3.9) induces

$$\int\tilde{p}d\tilde{x}-\Phi^*_{\epsilon F_i}Hdt=\sqrt{\epsilon}\Big(\int\tilde{y}d\tilde{x}-\tilde{G}_{\epsilon,i}ds\Big),$$

one obtains the rescaling of the first cohomology class, from $\tilde{\alpha}_{\epsilon,i}^{-1}(0)$ to $\tilde{\alpha}^{-1}_{\Phi^*_{\epsilon F_i}H}(0)$

$$\Psi_i^r:\ c-c_i\to\sqrt{\epsilon}(c-c_i),\qquad c_3-c_{i,3}\to\frac{\epsilon}{\omega_{3,i}}(c_3-c_{i,3}), \tag{8.2}$$

where $\tilde{c}_i=(c_i,c_{i,3})=(c_{i,1},c_{i,2},c_{i,3})=p_i''$ if we treat both as the points in $\mathbb{R}^3$.

Because $\Phi_{\epsilon F_i}$ is a Hamiltomorphism, it does not change the Mather set, Aubry set and Mañé set, due to Theorem 3.1.

8.2. Construction of the global transition chain . In this section, we show how to construct a global transition chain from the local transition chains.

Recall the circle $\Gamma_{k^\star}$ and Γ_{k^*} constructed in Section 3. For generic perturbation P, due to Proposition 5.1, the number of strong double resonant points is independent of ϵ. For each $i\in\Lambda_s$, the composition of Ψ_i^ℓ and Ψ_i^r maps $\mathbb{A}_i$ to an annulus of cohomology equivalence $\tilde{\mathbb{A}}_i\subset\tilde{\alpha}_H^{-1}(E)$ where $\tilde{\alpha}_H$ denotes the α-function of H. It also maps $\mathbb{C}_i^\imath$ for all $i\in\Lambda_\epsilon$ to a local channel $\tilde{\mathbb{C}}_i^\imath\subset\tilde{\alpha}_H^{-1}(E)$ ($\imath=\pm,w$), namely

$$\tilde{\mathbb{A}}_i=\Psi_i^\ell\Psi_i^r(\mathbb{A}_i),\qquad\tilde{\mathbb{C}}_i^\imath=\Psi_i^\ell\Psi_i^r(\mathbb{C}_i^\imath),\qquad\imath=\pm,w.$$

Since H is a small perturbation of h, the sphere $\tilde{\alpha}_H^{-1}(E)$ lies in $O(\epsilon)$-neighborhood of $\tilde{\alpha}_h^{-1}(E)=h^{-1}(E)$ if we treat both as the set in $\mathbb{R}^3$. Let

$$\tilde{\mathbb{C}}_k=\{\tilde{c}\in\tilde{\alpha}_H^{-1}(E):\langle k,\omega\rangle=0\ \forall\,\omega\in\mathscr{L}^{-1}_{\tilde{\alpha}_H}(\tilde{c})\},$$

Under generic perturbation ϵP, it looks like a channel made up of flats. A subset is said to be a flat of $\tilde{\alpha}_H$ if $\tilde{\alpha}_H$ is affine when it is restricted on the set, no longer affine on any set properly containing the set. Since H is autonomous, $E>\min\alpha_H$ and each $\omega\in\mathscr{L}^{-1}_{\tilde{\alpha}_H}(\tilde{\mathbb{C}}_k)$ is resonant, $\mathbb{F}_\omega=\mathscr{L}_{\tilde{\alpha}_H}(\omega)$ is a flat of dimension one or two.

The set $\tilde{\mathbb{C}}_{k^\star} \cup \tilde{\mathbb{C}}_{k^*}$ obviously contains the flats $\mathbb{F}_i$ ($i \in \Lambda_s$) for strong double resonance we are concerned about

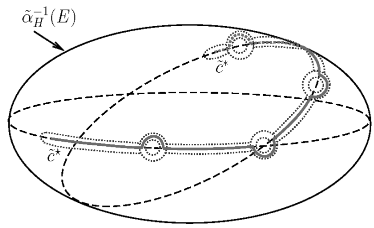
$$\mathbb{F}_i = \{\tilde{c} \in \tilde{\alpha}_H^{-1}(E) : \langle k_i', \omega\rangle = \langle k_i'', \omega\rangle = 0,\ \forall \omega \in \mathscr{L}_{\tilde{\alpha}_H}^{-1}(\tilde{c}),\ k_i' = k^\star \text{ or } k^*\}.$$

Let $\mathbb{F}_i + d_i\sqrt{\epsilon} = \{\tilde{c} \in \tilde{\alpha}_H^{-1}(E) : \mathrm{dist}(\tilde{c}, \mathbb{F}_i) \le d_i\sqrt{\epsilon}\}$ with $d_i < \Delta_i$, the set

$$\tilde{\mathbb{C}} = \Big(\tilde{\mathbb{C}}_{k^\star} \cup \tilde{\mathbb{C}}_{k^*}\backslash(\cup_{i\in\Lambda_s}\mathbb{F}_i + d_i\sqrt{\epsilon})\Big) \cup \Big(\ \cup_{i\in\Lambda_s}\tilde{\mathbb{A}}_i\Big)$$

is path-connected (Theorem 6.3 and Lemma 8.1 below). According to Lemma 7.3, $\tilde{\mathbb{C}}_k$ lies in $C_H\sqrt{\epsilon}$-neighborhood of Γ_k. The rescaling of the corresponding frequencies is bounded by Lemma 7.2.

For any $\delta > 0$, there exists $\epsilon_0 > 0$ such that $\forall\ \epsilon \le \epsilon_0$ there exist totally irreducible $k^\star, k^* \in \mathbb{Z}^3\backslash\{0\}$ such that $\tilde{\mathbb{C}}_{k^\star} \cup \tilde{\mathbb{C}}_{k^*}$ contains two points $\tilde{c}^\star, \tilde{c}^* \in \tilde{\mathbb{C}}$ with $\|\tilde{c}^\star - p^\star\| < \frac{\delta}{2}$ and $\|\tilde{c}^* - p^*\| < \frac{\delta}{2}$. Let $\Gamma_c = \Gamma_c(\epsilon P)$ be a path lying inside $\tilde{\mathbb{C}}$, connecting $\tilde{c}^\star$ to $\tilde{c}^*$. It



keeps away from the boundary of $\tilde{\mathbb{C}}$ and moves along a circle of cohomology equivalence when it turns around the strong double resonance. By definition, $\tilde{\alpha}_H(\tilde{c}) \equiv E > \min h$ for all $\tilde{c} \in \tilde{\mathbb{C}}$. We are going to show that Γ_c is a global transition chain.

8.3. The covering property. According to Theorem 4.1, the path $M_0^{-1}\Gamma_k$ is covered by the disks $\cup_{i\in\Lambda_\epsilon}\tilde{\Sigma}_{\epsilon,i}$ where

$$\tilde{\Sigma}_{\epsilon,i} = \{p : |p - M_0^{-1}p_i''| < (\eta T_i)^{-1}\epsilon^\kappa,\ \eta \in (0,1]\}.$$

We assume the subscripts $\{i \in \Lambda_\epsilon\}$ are well-ordered such that $\tilde{\Sigma}_{\epsilon,i}$ is adjacent to $\tilde{\Sigma}_{\epsilon,i\pm1}$. Restricted on the energy level set $\bar{H}^{-1}(E)$ contained in $\tilde{\Sigma}_{\epsilon,i} \times \mathbb{T}^3$, the Hamiltonian $\bar{H} = \bar{h} + \epsilon\bar{P}$ is reduced to the normal form $G_{\epsilon,i}$ defined on $\Sigma'_{\epsilon,i}$ as shown in Lemma 3.1. The corresponding channels $\mathbb{C}_i^\pm$ and $\mathbb{C}_i^w$ are described in Theorem 6.3 and 6.4.

Since $T_i^{-1}\epsilon^\kappa \ge \epsilon^{\frac{1}{3}}$, by replacing η with $\frac{\eta}{2}$, we can assume some $p' \in M_0^{-1}\Gamma_k$ exists, between p_i'' and p_{i+1}'', such that the disc $\{\|p - p'\| \le \xi\sqrt{\epsilon}\}$ is contained in $\tilde{\Sigma}_{\epsilon,i} \cap \tilde{\Sigma}_{\epsilon,i+1}$ and is mapped into $\Sigma'_{\epsilon,i}$ and into $\Sigma'_{\epsilon,i+1}$ by the transformations to get the normal form (6.1), as shown in Section 3. The number $\xi > 0$ can be large if $\epsilon > 0$ is small.

Let $\Gamma_{k,i} = M_0^{-1}\Gamma_k \cap \tilde{\Sigma}_{\epsilon,i}$. As Γ_k is smooth and ϵ is small, $\Gamma_{k,i}$ is $o(\epsilon)$-close to a straight line. For each disk $\tilde{\Sigma}_{\epsilon,i}$ we have local channel $\tilde{\mathbb{C}}_i^\imath$ made up of the flats $\mathscr{L}_{\tilde{\alpha}_H}(\nu_p\partial h(p))$ for $p \in \Gamma_{k,i}$, where ν_p is close to 1 such that $\tilde{\alpha}_H(\mathscr{L}_{\tilde{\alpha}_H}(\nu_p\partial h(p))) = h(p)$. If $\tilde{\mathbb{C}}_i^\imath$ and $\tilde{\mathbb{C}}_{i+1}^\imath$ overlap, each flat $\mathscr{L}_{\tilde{\alpha}_H}(\nu_p\partial h(p)) \subset \tilde{\mathbb{C}}_i^\imath \cup \tilde{\mathbb{C}}_{i+1}^\imath$ either entirely lies in the intersection $\tilde{\mathbb{C}}_i^\imath \cap \tilde{\mathbb{C}}_{i+1}^\imath$, or completely stays outside.

LEMMA 8.1. *There exist two numbers: a suitably small $\epsilon_0 > 0$ and a suitably large $\xi > 0$, such that for each $\epsilon \in (0, \epsilon_0]$, the adjacent channels $\tilde{\mathbb{C}}_i^\imath, \tilde{\mathbb{C}}_{i+1}^\imath$ overlap over a channel*

$$\tilde{\mathbb{C}}_i^\imath \cap \tilde{\mathbb{C}}_{i+1}^\imath \supseteq \cup_{p\in\Gamma_{k,i}\cap\Gamma_{k,i+1}}\mathscr{L}_{\tilde{\alpha}_H}(\nu_p\partial h(p)),$$

the length of the segment $\Gamma_{k,i}\cap\Gamma_{k,i+1}$ *is not shorter than* $\frac{\xi}{2}\sqrt{\epsilon}$.

Proof. We study the case that the double resonance at both adjacent points p''_i, p''_{i+1} are weak. Other cases can be treated similarly.

Because of Lemma 6.4, the oscillation of $y(\theta)$ is bounded by Δ'_j if $(x(\theta),y(\theta))$ is an orbit lying in $\tilde{\mathcal{A}}(c)$ with $c\in\mathbb{C}^w_j$. Since $|\bar{k}''_i||\partial_2 V''_i|\to 0$ as $|\bar{k}''_i|\to\infty$, we find from (6.7) that Δ'_j is uniformly bounded for all j. In the original coordinate, because of the three steps of coordinate changes, $\Phi_{\epsilon F}$, (3.1) and (3.6), there exists a constant $C_V>0$ such that $\mathrm{Osc}\|p(t)\|\le C_V\sqrt{\epsilon}$ if $(p(t),q(t))$ is an orbit in $\tilde{\mathcal{A}}(\tilde{c})$ for $\tilde{c}\in\tilde{\mathbb{C}}^w_j$.

Let us investigate the location of $p(t)$ if $(p(t),q(t))$ is a $\tilde{c}$-minimal orbit. If both $\tilde{c}$ and p are treated as points in $\mathbb{R}^3$, one has the relation $\tilde{\alpha}_h=h$. It follows from Lemma 7.3 that the channel $\tilde{\mathbb{C}}_k$ $(k=k^*,k^\star)$ is entirely contained in $C_H\sqrt{\epsilon}$-neighborhood of Γ_k. It follows from Lemma 7.4 that $(p(t),q(t))$ would not be $\tilde{c}$-minimal orbit of Φ^t_H if $p(t)$ entirely keeps away from $D_h\sqrt{\epsilon}$-neighborhood of $\tilde{c}$. Since $\mathrm{Osc}|p(t)|\le C_V\Delta'_j\sqrt{\epsilon}$ if $(p(t),q(t))\subset\tilde{\mathcal{A}}(\tilde{c})$ for $\tilde{c}=\Psi^\ell_j\Psi^r_j(c)$, each Aubry set $\tilde{\mathcal{A}}(\tilde{c})$ is entirely contained in $(C_V+D_h+C_H)\sqrt{\epsilon}$-neighborhood of $\tilde{c}\in\Gamma_k$ if we treat both as the points in $\mathbb{R}^3$.

The parameters ξ and ϵ_0 in the lemma are set such that

$$\xi\ge 4(C_V+D_h+C_H),\qquad \xi\sqrt{\epsilon_0}\le\frac{1}{2}\epsilon_0^{\frac{1}{3}}.$$

Then, there exists a segment $\Gamma'_{k,i}$ of Γ_k between p''_i and p''_{i+1} with the length $\frac{\xi}{2}\sqrt{\epsilon}$ such that for each $p'\in\Gamma'_{k,i}$ one has $\{\|p-p'\|<\frac{\xi}{4}\sqrt{\epsilon}\}\subseteq\tilde{\Sigma}_{\epsilon,i}\cup\tilde{\Sigma}_{\epsilon,i+1}$. Indeed, $p'\in\Gamma'_{k,i}$ implies that the distance between p' and the boundary of $\tilde{\Sigma}_{\epsilon,j}$ is not smaller than $\frac{\xi}{4}\sqrt{\epsilon}$ for $j=i,i+1$.

For each $p'\in\Gamma'_{k,i}$, let $\tilde{c}'\in\mathscr{L}_{\tilde{\alpha}_H}(\nu_{p'}\partial h(p'))$. Then, the Aubry set $\tilde{\mathcal{A}}(\tilde{c}')$ for H lies in the disc $\{\|p-p'\|<\frac{\xi}{4}\sqrt{\epsilon}\}$. Under the composition of the maps $\Phi_{\epsilon F}$, (3.1), (3.6) and (3.9), for $\tilde{\Sigma}_{\epsilon,j}\to\Sigma'_{\epsilon,j}$ for $j=i,i+1$ the domain $\{\|p-p'\|<\frac{\xi}{4}\sqrt{\epsilon}\}$ is mapped onto a domain entirely contained $\Sigma'_{\epsilon,j}$. The Aubry set of $(\Psi^\ell_j\Psi^r_j)^{-1}\tilde{c}'$ for $G_{\epsilon,j}$ is contained in the domain $\mathbb{T}^2\times\Sigma'_{\epsilon,j}\times|\bar{k}''_j|\mathbb{T}$. It implies that $\tilde{c}'\in\tilde{\mathbb{C}}^w_j$ for $j=i,i+1$. □

Since $\tilde{\mathbb{C}}^+_i$ and $\tilde{\mathbb{C}}^-_i$ are joined by $\Psi^\ell_i\Psi^r_i(\mathbb{A}_i)$, the whole path Γ_c is covered by

$$\Gamma_c\subset\bigcup_{i\in\Lambda_s}\Big(\Psi^\ell_i\Psi^r_i(\mathbb{A}_i)\cup\tilde{\mathbb{C}}^-_i\cup\tilde{\mathbb{C}}^+_i\Big)\cup\bigcup_{i\in\Lambda_\epsilon\backslash\Lambda_s}\tilde{\mathbb{C}}^w_i.\tag{8.3}$$

8.4. The genericity. Before the proof of Theorem 2.1, we focus on a prescribed path. The following theorem is Theorem 5.1 of [C17b].

THEOREM 8.2. *There exists a set* $\mathfrak{C}_{\epsilon_0}$ *cusp-residual in* $\mathfrak{B}_{\epsilon_0}\subset C^r(B_D\times\mathbb{T}^3,\mathbb{R})$ *with* $r\ge 6$ *such that for each* $\epsilon P\in\mathfrak{C}_{\epsilon_0}$, *the path* Γ_c *is a transition chain.*

Proof. To check if Γ_c is a transition chain under generic perturbation ϵP, one only needs to check, for generic perturbation ϵP, the condition of transition chain for each $c\in\mathbb{A}_i\cup\mathbb{C}^-_i\cup\mathbb{C}^+_i$ if $i\in\Lambda_s$ and for each $c\in\mathbb{C}^w_i$ if $\Lambda_\epsilon\backslash\Lambda_s$.

As the first step, we show the cusp-residual property that, for every $c\in\mathbb{C}^-_i\cup\mathbb{C}^+_i$ with $i\in\Lambda_s$ and for $c\in\mathbb{C}^w_i$ with $i\in\Lambda_\epsilon\backslash\Lambda_s$, the Aubry set $\tilde{\mathcal{A}}(c)$ lies on some NHIC (candidate of transition chain):

(**CT**: $i\in\Lambda_s$). *There is an annulus of cohomology equivalence* $\mathbb{A}_i$ *connecting channel* $\mathbb{C}^-_i$ *to* $\mathbb{C}^+_i$, *and* $\tilde{\mathcal{A}}(c)$ *lies on some NHIC for* $c\in\mathbb{C}^-_i\cup\mathbb{C}^+_i$.

(**CT**: $i\in\Lambda_\epsilon\backslash\Lambda_s$). *The Aubry set* $\tilde{\mathcal{A}}(c)$ *lies on certain NHIC for each* $c\in\mathbb{C}^w_i$.

At strong resonant point p_i'', one has a decomposition $C^r(B_D\times\mathbb{T}^3,\mathbb{R}) = C^r(\mathbb{T}^2,\mathbb{R})\oplus C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(\mathbb{T}^2,\mathbb{R})$ via $P(p,q) = V_i(\langle k',q\rangle,\langle k_i'',q\rangle) + P''(p,q)$ where the resonant term $V_i(\langle k',q\rangle,\langle k_i'',q\rangle) = Z(p_i'',\langle k',q\rangle,\langle k_i'',q\rangle)$ is defined in (5.1), $P_i'' = P - V_i \in C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(\mathbb{T}^2,\mathbb{R})$ is the non-resonant term and $k' = k^*, k^\star$.

LEMMA 8.2. *Let $\mathfrak{V}_i$ be the residual set used in Theorem 6.1, Proposition 6.1, Theorem 6.2 and Theorem 6.3. Then, the following set*

$$\mathfrak{P}_i = \{V_i + P_i'' : V_i \in \mathfrak{V}_i,\ (\mathbf{CT}: i\in\Lambda_s) \text{ holds for } h+\epsilon(V_i+P_i'')\}$$

is cusp-residual in $\mathfrak{B}_1$.

Proof. According to Lemma 3.1, the Hamiltonian H is reduced to the normal form $G_{\epsilon,i}(x,y,\theta)$ of Formula (6.1) when it is restricted on the energy level set $H^{-1}(E)$ lying in $\mathbb{T}^3\times\{|p-p_i''|\le(\eta T_i)^{-1}\epsilon^\kappa\}$, the neighborhood of double resonant point p_i''. In such a normal form, the main part is independent of ϵ and the remainder $R_{\epsilon,i}$ is uniformly bounded in the sense that $\|R_{\epsilon,i}\|_{C^{r-2}}\le a_0\epsilon^\kappa$ holds for all $P\in\mathfrak{B}_1$ if we consider it as the function of (x,y,ϑ) where $\vartheta = \omega_3|k_i''|^{-1}\sqrt{\epsilon}^{-1}\theta$.

By applying Theorem 1.1 and 1.2 of [C17b] to the Lagrangian obtained from $G_{\epsilon,i}$, we find that there exists an open-dense set $\mathfrak{V}_i$ in $C^r(\mathbb{T}^2,\mathbb{R})$, each $V_i\in\mathfrak{V}_i$ is associated some small $\epsilon_{V_i}>0$ such that the condition ($\mathbf{CT}$: $i\in\Lambda_s$) holds for each $R_{\epsilon,i}\in\mathfrak{B}_{\epsilon_{V_i}}\subset C^{r-2}(\Sigma_{\epsilon,i}'\times\mathbb{T}^3,\mathbb{R})$. Because $\|R_{\epsilon,i}\|_{C^{r-2}}\le a_0\epsilon^\kappa$, for $P = V_i + P_i''\in\mathfrak{B}_1$ with $V_i\in\mathfrak{V}_i$, the condition ($\mathbf{CT}$: $i\in\Lambda_s$) holds for each $\epsilon\in(0,(a_0^{-1}\epsilon_{V_i})^{1/\kappa}]$. $\square$

Next, we consider NHICs away from strong double resonance. The path $\Gamma_{k'}$ induces a decomposition $C^r(B_d\times\mathbb{T}^3,\mathbb{R}) = C^r(B_D\times\mathbb{T}^1,\mathbb{R})\oplus C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(B_D\times\mathbb{T}^1,\mathbb{R})$ via $P(p,q) = Z_{k'}(p,\langle k',q\rangle) + P'(p,q)$ ($k' = k^\star, k^*$) where $Z_{k'}$ is defined in (5.2) consisting of Fourier modes of P in $\mathrm{span}_{\mathbb{Z}}\{k'\}$, and $P' = P - Z_{k'}\in C^r(B_D\times\mathbb{T}^3,\mathbb{R})/C^r(B_D\times\mathbb{T}^1,\mathbb{R})$. Treating $\langle k',q\rangle$ as a scalar variable x defined on $\mathbb{T}$, there exists an open-dense set $\mathfrak{Z}_{k'}$ of $C^r(\Gamma_{k'}\times\mathbb{T},\mathbb{R})$ such that for each $Z_{k'}\in\mathfrak{Z}_{k'}$ it holds simultaneously for all $p\in\Gamma_{k'}$ that the second derivative of $Z_{k'}$ in x at its maximal point is uniformly upper bounded below zero, because of Theorem 1.1 of [Zh2]. In this case, the number of strong double resonant points is independent of ϵ.

Given $Z_{k'}\in\mathfrak{Z}_{k'}$, let $\{p_i'\in\Gamma_{k'}\}$ denote the set of bifurcation points, i.e. $Z_{k'}(p_i',x)$ has two maximal points in x. Let N_w denote the cardinality of $\{p_i'\in\Gamma_{k'}\}$, it is finite. Let $H_1 = h(p)+\epsilon Z_{k'}(p,\langle k,q\rangle)$, then $\Phi_{H_1}^t$ admits N_w+1 pieces of NHIC consisting of minimal periodic orbits of $\Phi_{\bar H}^t$. Because of the presence of strong double resonances, the NHICs may break into more pieces of NHICs if we take the second resonant term into account. However, outside of the neighborhoods of strong double resonances, the number of NHICs will be not more than $N\le N_s+N_w+2$, where $N_s = \#(\Lambda_s)$. Indeed,

LEMMA 8.3. *In $\mathfrak{B}_1$ the following set*

$$\mathfrak{P}_{k'} = \cup\{Z_{k'} + P' : Z_{k'}'\in\mathfrak{Z}_{k'},\ (\mathbf{CT}: i\in\Lambda_\epsilon\backslash\Lambda_s) \text{ holds for } h+\epsilon(Z_{k'}+P')\}$$

is cusp-residual.

Proof. The normal form around weak resonant point p_i'' takes the form of (6.1) where $V_i(x) = V_i'(x_1) + V_i''(x_1,|\bar k''|x_2)$ with small V_i'', the NHIC for the Hamiltonian flow of $\frac{1}{2}\langle B_iy,y\rangle - V_i'(x_1)$ survives the perturbations V_i'' and $R_{\epsilon,i}$ if the second derivative of V_i' at its minimal point is positive and some ϵ is sufficiently small.

We notice $V_i'(x) = -Z_{k'}'(p_i'',x)$. Although the number $\#(\Lambda_\epsilon\backslash\Lambda_s)$ increases to infinity as $\epsilon\downarrow 0$, it does not cause trouble. There is an open-dense set $\mathfrak{Z}_{k'}\subset C^r(B_d\times$

$\mathbb{T},\mathbb{R})$, for each $Z_{k'}(p,x)\in\mathfrak{Z}_{k'}$, the second derivative of $Z_{k'}(p,x)$ in x at its maximal point is uniformly upper bounded below zero. So, the normal hyperbolicity of all NHICs of $G_{\epsilon,i}$ is uniformly bounded from below as ϵ decreases to 0. Consequently, some $\epsilon_{Z_{k'}}>0$ exists such that for each $i\in\Lambda_\epsilon\backslash\Lambda_s$ the NHIC for $G_{\epsilon,i}$ persists provided $\|R_{\epsilon,i}\|_{C^2}\le\epsilon_{Z_{k'}}$.

By Lemma 3.1 again, given $P=Z'_{k'}+P'_{k'}\in\mathfrak{B}_1$ with $Z'_{k'}\in\mathfrak{Z}_{k'}$, the condition (**CT**: $i\in\Lambda_\epsilon\backslash\Lambda_s$) holds for each $\epsilon\in(0,(a_0^{-1}\epsilon_{Z_{k'}})^{1/\kappa}]$.

For the coordinate transformation (3.1) the matrix M_i is set according to whether $|\bar{k}''_{i,2}|\ge|\bar{k}''_{i,3}|$ or not. The path $\partial h(\Gamma_{k'})$ admits a partition of four arcs. The condition $|\bar{k}''_{i,2}|>|\bar{k}''_{i,3}|$ holds on two arcs and $|\bar{k}''_{i,2}|<|\bar{k}''_{i,3}|$ holds on other two arcs. □

Let $\bar{\mathfrak{P}}_{k'}=\cup_{\nu\in\mathbb{R}}\{\nu P:P\in\mathfrak{P}_{k'}\}\cap\mathfrak{B}_1$ and $\bar{\mathfrak{P}}_i=\cup_{\nu\in\mathbb{R}}\{\nu P:P\in\mathfrak{P}_i\}\cap\mathfrak{B}_1$, then the set $\bar{\mathfrak{P}}_{k^\star}\cap\bar{\mathfrak{P}}_{k^*}\cap(\cap_i\bar{\mathfrak{P}}_i)$ is residual in $\mathfrak{B}_1$. Applying the Kuratowski-Ulam theorem (categorical analogue of the Fubini theorem, c.f. Chapter 15 of [Ox]), one obtains a residual set $\mathfrak{R}\subset\mathfrak{S}_1$, each $P\in\mathfrak{R}$ is associated with a set I_P residual in $[0,1]$ such that $\cup_{\lambda\in I_P}\lambda P\subset\bar{\mathfrak{P}}_{k^\star}\cap\bar{\mathfrak{P}}_{k^*}\cap(\cap_i\bar{\mathfrak{P}}_i)$.

Each $P\in\mathfrak{R}$ determines finitely many strong double resonant points $\{p''_i,i\in\Lambda_s(P)\}$. For each $i\in\Lambda_s(P)$, one obtains the potential $V_i=V_i(P)$ which determines a number $\epsilon_{V_i}>0$. Let $a_P=\min\{\epsilon_{k^\star},\epsilon_{k^*},\epsilon_{V_i},i\in\Lambda_s(P)\}$ where $\epsilon_{k^\star},\epsilon_{k^*}>0$ are determined by the first resonant conditions $k^\star$ and k^* respectively (see Lemma 8.3). For any $\epsilon\in(0,a_P)$, the flow Φ^t_H with $H=h+\epsilon P$ admits the conditions (**CT**:$i\in\Lambda_s$) and (**CT**:$i\in\Lambda_\epsilon\backslash\Lambda_s$). Therefore, *a cusp-residual set* $\mathfrak{C}'_{\epsilon_0}$ *exists such that for each* $\epsilon P\in\mathfrak{C}'_{\epsilon_0}$ *the conditions* (**CT**: $i\in\Lambda_s$) *and* (**CT**: $i\in\Lambda_\epsilon\backslash\Lambda_s$) *hold.*

Some cohomology equivalence exists around each class c lying in the channels if $\tilde{\mathcal{A}}(c)$ is not a 2-dimensional torus, as it was shown in [CY1]. To finish the proof of Theorem 8.2, we need to verify the condition (**HA**) for each class c lying in the channels if $\tilde{\mathcal{A}}(c)$ is a 2-dimensional torus. For each Hamiltonian of normal form $G_{\epsilon,i}$, it has been proved in Theorem 1.2 of [C17b]. However, we shall not apply that result since the residual set obtained there is for C^{r-2}-topology instead of C^r-topology. One step KAM iteration for the construction of $\Phi_{\epsilon F}$ makes $\tilde{G}_{\epsilon,i}$ lose two times of differentiability.

We apply the following lemma to finish the proof of Theorem 8.2. The proof of the lemma will be presented afterward.

Lemma 8.4. *Each perturbation* $\epsilon P\in\mathfrak{C}'_{\epsilon_0}$ *is associated with a set* $\mathfrak{P}'_{\epsilon P}$ *residual in a ball* $\mathfrak{B}_{\delta(\epsilon P)}\subset C^r(B_D\times\mathbb{T}^3,\mathbb{R})$ *with small radius* $\delta(\epsilon P)>0$*, such that for each* $\epsilon P'\in\mathfrak{P}'_{\epsilon P}$ *the Hamiltonian* $h+\epsilon(P+P')$ *possesses the property: the condition* (**HA**) *holds for each first cohomology classes* $\tilde{c}\in\tilde{\mathbb{C}}_{k^\star}\cup\tilde{\mathbb{C}}_{k^*}\backslash(\cup_{i\in\Lambda_s}\mathbb{F}_i+d_i\sqrt{\epsilon})$ *if the Aubry set* $\tilde{\mathcal{A}}(\tilde{c})$ *consists of a two-dimensional torus.*

Obviously, $\cup_{\epsilon P\in\mathfrak{C}'_{\epsilon_0}}(\epsilon P+\mathfrak{B}_{\delta(\epsilon P)})\supseteq\mathfrak{C}'_{\epsilon_0}$. Let $\mathfrak{R}'=\cup_{\epsilon P\in\mathfrak{C}'_{\epsilon_0}}(\epsilon P+\mathfrak{P}'_{\epsilon P})$, because of the Kuratowski-Ulam theorem, it contains a cusp-residual set: there is a set $\bar{\mathfrak{R}}'$ residual in $\mathfrak{S}_1$, each $P\in\mathfrak{S}_1$ is associated with a set R_p residual in $(0,a_P)$ such that $\epsilon P\in\mathfrak{R}'$ holds for all $\epsilon\in R_p$.

For $P\in\bar{\mathfrak{R}}\cap\bar{\mathfrak{R}}'$, $\Lambda_s(P)$ is a finite set. Therefore, a set R_P residual in $(0,a_P)$ exists such that for each $\epsilon\in R_P$, the conditions (**CT**:$i\in\Lambda_s$), (**CT**:$i\in\Lambda_\epsilon\backslash\Lambda_s$) and (**HA**) hold, namely, the flow Φ^t_H with $H=h+\epsilon P$ admits a transition chain. It verifies the cusp-residual property of the transition chain Γ_c. The proof of Theorem 8.2 is completed. □

Proof of Lemma 8.4. For $\epsilon P\in\mathfrak{C}'_{\epsilon_0}$, the Hamiltonian $h+\epsilon P$ behaves like an *a priori* unstable system when it is restricted in neighborhood of $\tilde{c}$-minimal orbits with $\tilde{c} \in \tilde{\mathbb{C}}_{k^\star} \cup \tilde{\mathbb{C}}_{k^*} \backslash (\cup_{i \in \Lambda_s} \mathbb{F}_i + d_i \sqrt{\epsilon})$. Although ϵ is small, it is treated as a fixed number since ϵP is fixed. There exists a small number $\delta = \delta(\epsilon P)$ such that for any $\epsilon P' \in \mathfrak{B}_{\delta(\epsilon P)}$ the condition (**CT**: $i \in \Lambda_s$) and (**CT**: $i \in \Lambda_\epsilon \backslash \Lambda_s$) holds for $h + \epsilon(P + P')$.

Due to the covering property (8.3), we only need to check the condition (**HA**) for $\Gamma_c \cap \tilde{\mathbb{C}}_i^{\pm}$ and $\Gamma_c \cap \tilde{\mathbb{C}}_i^w$. Restricted in $\tilde{\Sigma}_{\epsilon,i} \cap \{H^{-1}(E)\}$, we reduce $H(p,q)$ to a system with two and half degrees of freedom so that we can apply the result of ([CY1, CY2]).

We consider the case $\Gamma_c \cap \tilde{\mathbb{C}}_i^+$, the proof for the cases $\Gamma_c \cap \tilde{\mathbb{C}}_i^-$ and $\Gamma_c \cap \tilde{\mathbb{C}}_i^w$ is the same. Restricted on $(\tilde{\Sigma}_{\epsilon,i} \times \mathbb{T}^3) \cap \{\bar{H}^{-1}(E)\}$, the Hamiltonian $\bar{H} = \bar{h}(\bar{p}) + \epsilon \bar{P}(\bar{p}, \bar{q})$ is reduced to the normal form $G_{\epsilon,i}$ of (6.1), due to Lemma 3.1. Because of Theorem 6.2, there exist finitely many normally hyperbolic weakly invariant cylinders, denoted by $\tilde{\Pi}_\ell = \tilde{\Pi}^\epsilon_{i,E_\ell - d, E_{\ell+1}+d}$, modulo the shift σ_i. For each $c \in \mathbb{C}_i^+ = (\Psi_i^\ell \Psi_i^r)^{-1} \tilde{\mathbb{C}}_i^+$, the Aubry set lies in these cylinders. Denoted by Π_ℓ the time-0-section of $\tilde{\Pi}_\ell$, i.e. $\Pi_\ell = \tilde{\Pi}_\ell|_{\theta=0}$.

Similar to the argument in Section 4.1 of [C17b], the cylinder Π_ℓ can be treated as a part of the image of a standard cylinder $\Pi = \{(x,y) : (x_2, y_2) = 0, x_1 \in \mathbb{T}, y_1 \in [0,1]\}$ under a map ψ: $\Pi \to \Pi_\ell$. This map induces a 2-form $\psi^*\omega$ on Π

$$\psi^*\omega = D\psi dx_1 \wedge dy_1$$

where $D\psi$ is the Jacobian of ψ. Since the second de Rham cohomology group of Π is trivial, it follows from Moser's argument on the isotopy of symplectic forms [Mo] that there exists a diffeomorphism ψ_1 on Π such that

$$(\psi \circ \psi_1)^*\omega = dx_1 \wedge dy_1.$$

Let $\theta_1 = \frac{2|k_i''|\sqrt{\epsilon}\pi}{\omega_{3,i}}$. Since Π_ℓ is invariant for the time-periodic map $\Phi_{G_{\epsilon,i}} = \Phi^{\theta_1}_{G_{\epsilon,i}}$ and $\Phi^*_{G_{\epsilon,i}}\omega = \omega$, one has

$$((\psi \circ \psi_1)^{-1} \circ \Phi_{G_\epsilon} \circ (\psi \circ \psi_1))^* dx_1 \wedge dy_1 = dx_1 \wedge dy_1 \tag{8.4}$$

i.e. $(\psi \circ \psi_1)^{-1} \circ \Phi_{G_\epsilon} \circ (\psi \circ \psi_1)$ preserves the standard area. Each invariant circle $\Gamma_\sigma \subset \Pi_\ell$ is pulled back to the standard cylinder, denoted by Γ^*_σ which is Lipschitz. The parameter σ is set to be the algebraic area bounded by the circle and a prescribed one, $\|\Gamma^*_\sigma - \Gamma^*_{\sigma'}\|_{C^0} \leq b_0\sqrt{|\sigma - \sigma'|}$ (see [CY1]). Since the maps ψ, ψ_0 are smooth, back to the current coordinates one has $\|\Gamma_\sigma - \Gamma_{\sigma'}\|_{C^0} \leq b_1\sqrt{|\sigma - \sigma'|}$. We notice that the cylinder Π_ℓ may be crumple and slanted, the constant b_1 might approach infinity if the crumpled cylinder extends to the homoclinics. However, since the cylinder is kept away from the double resonance for certain distance (see Lemma 6.3, $\alpha_{\epsilon,i}(\mathscr{L}_{\alpha_{\epsilon,i}}(\lambda_i^\pm g_i^\pm)) = d_i > 0$), the cylinders are moderately crumpled. The constant b_1 is therefore uniformly bounded for σ if we are restricted on the cylinder Π_ℓ.

Recall the rescaling (3.9), (3.6) and let $v_i'' = (v''_{i,1}, v''_{i,2}, v''_{i,3})$ be the double resonant point, we have a transformation $\mathscr{R}_i$

$$\mathscr{R}_i : \begin{cases} x_j = u_j, \qquad y_j = \sqrt{\epsilon}^{-1}(v_j - v''_{i,j}), \qquad j = 1, 2, \\ \theta = \frac{\sqrt{\epsilon}}{\omega_{i,3}} u_3, \qquad I = \frac{\omega_{i,3}}{\epsilon}(v_3 - v''_{i,3}). \end{cases}$$

Replacing (x, y, θ) in the normal form $G_{\epsilon,i}$ of (6.1) by (u, v_1, v_2) defined as above, we obtain a Hamiltonian with two and half degrees of freedom

$$G^\star_{\epsilon,i}(u_1, u_2, u_3, v_1, v_2) = \epsilon G_{\epsilon,i}(x(u_1, u_2), \theta(u_3), y(v_1, v_2)) \tag{8.5}$$

where u_3 plays the role of time. Because of (8.5), the function $G^\star_{\epsilon,i}$ solves the equation $\mathscr{M}_i^*\Phi^*_{\epsilon F}\mathscr{M}_0^* H(u,v_1,v_2,-G^\star_{\epsilon,i})=E$, see Formula (3.10).

Obviously, $\tilde\Pi^\star_\ell=\mathscr{R}_i\tilde\Pi_\ell$ is weakly invariant for the flow of $G^\star_{\epsilon,i}$. Let $\Pi^\star_\ell=\tilde\Pi^\star_\ell|_{u_3=0}$. Compared with Π_ℓ, it shrinks in the coordinate v by the scale $\sqrt{\epsilon}$. The cylinder may become more crumpled, but it is controlled by the factor $1/\sqrt{\epsilon}$. Let $\Gamma^\star_{i,\sigma}$ denote the invariant circle in $\Pi^\star_\ell$, one has $\|\Gamma^\star_\sigma-\Gamma^\star_{\sigma'}\|_{C^0}\le\frac{b_1}{\sqrt{\epsilon}}\sqrt{|\sigma-\sigma'|}$. Let $\tilde\Gamma^\star_\sigma=\cup_{u_3\in[0,|\bar k''_i|]}\Phi^{u_3}_{G^\star_{\epsilon,i}}\Gamma^\star_\sigma$, one has

$$\|\tilde\Gamma^\star_\sigma-\tilde\Gamma^\star_{\sigma'}\|_{C^0}\le\frac{b_1}{\sqrt{\epsilon}}\max_{u_3\in[0,|\bar k''_i|]}\|D\Phi^{u_3}_{G^\star_{\epsilon,i}}\|\sqrt{|\sigma-\sigma'|}$$

Let $H^\star=\mathscr{M}_i^*\Phi^*_{\epsilon F}\mathscr{M}_0^* H$. Corresponding to the cylinder $\tilde\Pi^\star_\ell$, there is a cylinder $\hat\Pi^\star_\ell$ modulo the shift σ_i, lying in the energy level set $\{H^{\star-1}(E)\}$

$$\hat\Pi^\star_\ell=\{(u,v)\in\mathbb{T}^3\times\mathbb{R}^3:(u,v_1,v_2)\in\tilde\Pi^\star_\ell,\ v_3=G^\star_{\epsilon,i}(u,v_1,v_2)\}.$$

Those invariant 2-tori $\{\tilde\Gamma^\star_\sigma\}$ lying in the cylinder. Under the inverse of $\mathscr{M}_i$ and $\Phi_{\epsilon F}$, the tori $\{\tilde\Gamma^\star_\sigma\}$ and the cylinder $\hat\Pi^\star_\ell$ are mapped to the invariant tori and weakly invariant cylinder for the flow of the Hamiltonian $\mathscr{M}_0^*H$:

$$\hat\Pi_\ell=\Phi^{-1}_{\epsilon F}\mathscr{M}_i^{-1}\hat\Pi^\star_\ell,\qquad\hat\Gamma_\sigma=\Phi^{-1}_{\epsilon F}\mathscr{M}_i^{-1}\hat\Gamma^\star_\sigma.$$

Since $\Phi_{\epsilon F}$ is a diffeomorphism close to identity and $\mathscr{M}_i$ is linear, there exists a number $b_2>0$ such that

$$\|\hat\Gamma_\sigma-\hat\Gamma_{\sigma'}\|_{C^0}\le\frac{b_2}{\sqrt{\epsilon}}\sqrt{|\sigma-\sigma'|}. \tag{8.6}$$

Because $\partial\mathscr{M}_0^*h(\bar p''_i)=\bar\omega=(0,\bar\omega_2,\bar\omega_3)$ satisfies the resonant condition $\bar k''_i=(0,\bar k''_{i,2},\bar k''_{i,3})$, we construct another canonical transformation $\bar{\mathscr{M}}_i$: $\phi=\bar M_i^{-t}\bar q$ and $I=\bar M_i\bar p$, where

$$\bar M_i=\mathrm{diag}(1,\bar M_{i,2}),\qquad\text{with }\ \bar M_{i,2}=\begin{bmatrix}\bar k''_{i,2} & j_2\\ \bar k''_{i,3} & j_3\end{bmatrix}$$

the integers j_2,j_3 are chosen such that $\bar k''_{i,2}j_3-\bar k''_{i,3}j_2=1$. Clearly, $\bar M_i$ is uni-modular and the first derivative of $\bar{\mathscr{M}}_i^*\mathscr{M}_0^*h$ in I_3 is not equal to zero at the point $I''_i=\bar M_i\bar p''_i$. Therefore, there exists a C^r-function $G^*_{\epsilon,i}$ which solves the equation

$$\bar{\mathscr{M}}_i^*\mathscr{M}_0^*H(\phi,I_1,I_2,-G^*_{\epsilon,i}(\phi,I_1,I_2))=E \tag{8.7}$$

when it is restricted in a neighborhood of I''_i which covers the domain $\tilde\Sigma_{\epsilon,i}\times\mathbb{T}^3$ under the map $\bar{\mathscr{M}}_i$. The function $G^*_{\epsilon,i}$ defines a Hamiltonian system with two and half degrees of freedom where ϕ_3 plays the role of time, there is a normally hyperbolic and weakly invariant cylinder $\hat\Pi^*_\ell$ for $G^*_{\epsilon,i}$ such that

$$\hat\Pi_\ell=\bar{\mathscr{M}}_i^{-1}\{(\phi,I)\in\mathbb{T}^3\times\mathbb{R}^3:(\phi,I_1,I_2)\in\hat\Pi^*_\ell,\ I_3=G^*_{\epsilon,i}(\phi,I_1,I_2)\}.$$

Let $\hat\Gamma^*_\sigma=\bar{\mathscr{M}}_i^{-1}\hat\Gamma_\sigma$, they lie in the cylinder $\hat\Pi^*_\ell$ for which the modulus of continuity of (8.6) still holds, probably with a larger coefficient $b^*_2\ge b_2$.

Let $\delta_i \downarrow 0$ be a countable sequence. Because of the modulus of continuity of (8.6), it is proved in [CY1, CY2] that, for any small $\delta > 0$, it is an open-dense condition for $G^*_{\epsilon,i}$ in C^r-topology that the diameter of each connected component of the set

$$\mathcal{N}(c(\sigma), \check{M})|_{\phi_3=0} \backslash (\mathcal{A}(c(\sigma), \check{M}) + \delta)|_{\phi_3=0} \tag{8.8}$$

is not larger than δ_i (For the convenience of reader, we shall present the proof of this property in the appendix). The residual set is obtained by taking the intersection of the countably many open-dense sets. Since $G^*_{\epsilon,i}$ solves Equation (8.7), the perturbation $G^*_{\epsilon,i} \to G^*_{\epsilon,i} + G_\delta$ can be achieved by the perturbation $\bar{\mathscr{M}}^*_i \mathscr{M}^*_0 H(I_1, I_2, I_3, \phi) \to \bar{\mathscr{M}}^*_i \mathscr{M}^*_0 H(I_1, I_2, I_3 + G_\delta, \phi)$. Therefore, a set $\mathfrak{R}_{i,\ell}$ residual in $\mathfrak{B}_\delta$ is correspondingly obtained for H such that for each $\epsilon P' \in \mathfrak{B}_\delta$ the condition (**HA**) holds for cylinder $\tilde{\Pi}_\ell$. Taking the intersection $\cap \mathfrak{R}_{i,\ell}$ which is still residual in $\mathfrak{B}_\delta$. $\square$

Proof of Theorem 2.1. Given $\delta > 0$ there exists an integer $K = K(\delta)$ such that any point $p \in h^{-1}(E)$ falls into δ-neighborhood of some resonant path Γ_k with $|k| \leq K$. Since there are finitely resonant circles, we take the intersection of finitely many cusp-residual sets obtained in Theorem 8.2 which is still a cusp-residual set. $\square$

9. Proof of Theorem 1.1. The proof of Theorem 1.1 is simple once one has Theorem 5.1 of [C17b] (Theorem 8.2 here) and Theorem 3.1 of [LC] as follows.

THEOREM 3.1 OF [LC]. Suppose that there is a generalized transition chain Γ: $[0,1] \to H^1(\mathbb{T}^n, \mathbb{R})$ joining c to c'. Then, there exists an orbit of the Lagrange flow Φ^t_L $d\gamma$: $\mathbb{R} \to T\mathbb{T}^n$ that connects the two Aubry sets: $\alpha(d\gamma) \subset \tilde{\mathcal{A}}(c)$ and $\omega(d\gamma) \subset \tilde{\mathcal{A}}(c')$.

Proof of Theorem 1.1. If both p and $\tilde{c}$ are treated as point in $\mathbb{R}^3$, $p(t) \equiv \tilde{c}$ holds along each $\tilde{c}$-minimal orbit $(p(t), q(t))$ of Φ^t_h. By Lemma 2.1 of [CY2], the set of minimal orbits is upper-semi continuous with respect to perturbation. Therefore, along each orbit $(p(t), q(t))$ in the Aubry set $\tilde{\mathcal{A}}(\tilde{c})$ one has $\|p(t) - \tilde{c}\| < \frac{\delta}{2}$ if $\epsilon > 0$ is suitably small. So, an orbit connecting $\tilde{\mathcal{A}}(\tilde{c}^\star)$ to $\tilde{\mathcal{A}}(\tilde{c}^*)$ with $\|\tilde{c}^\star - p^\star\| < \frac{\delta}{2}$ and $\|\tilde{c}^* - p^*\| < \frac{\delta}{2}$ satisfies the condition: some $t^\star, t^* \in \mathbb{R}$ exist such that $\|p(t^\star) - p^\star\| < \delta$ and $\|p(t^*) - p^*\| < \delta$.

According to Theorem 5.1 of [C17b], for each $\epsilon P \in \mathfrak{C}_{\epsilon_0}$, there is a transition chain that connects the first cohomology classes $\tilde{c}^\star, \tilde{c}^* \in H^1(\mathbb{T}^3, \mathbb{R})$ satisfying the condition $\alpha(\tilde{c}^\star) = \alpha(\tilde{c}^*) = E$, $\|p^\star - \tilde{c}^\star\| < \frac{\delta}{2}$ and $\|p^* - \tilde{c}^*\| < \frac{\delta}{2}$. In this case, one obtains from Theorem 3.1 of [LC] an orbit connecting the Aubry set $\tilde{\mathcal{A}}(\tilde{c}^\star)$ to $\tilde{\mathcal{A}}(\tilde{c}^*)$. It completes the proof of the theorem. $\square$

Theorem 1.1 is the elaboration and justification of the sentence in the end of Section 5 of [C17b]: *the conjecture of Arnold diffusion for positive definite Hamiltonian turns out to be a theorem for* $n = 3$. It is an immediate consequence of Theorem 5.1 of [C17b].

Theorem 2.1 leads to the existence of certain δ-dense of the diffusion orbits, slightly stronger than Theorem 1.1. A diffusion orbit is said to be δ-dense if it passes through δ-neighborhood of any point $p \in H^{-1}(E)$.

Appendix A. The proof of genericity. For the convenience of reader, we present a proof of the property (8.8) by applying the ideas and the techniques of [CY1, CY2]. Another version appeared in [BKZ].

Given a Tonell C^r-Hamiltonian $H(p,q,t)$ where $(p,q,t) \in \mathbb{R}^2 \times \mathbb{T}^3$, $r \geq 3$, we have a Tonelli Lagrangian $L(\dot{q}, q, t) = \max_p \langle \dot{q}, p \rangle - H(p,q,t)$. Let $\Phi^{t,t'}_H$ denote the

Hamiltonian flow of H, it maps the initial value at the time-t-section to the time-t'-section. For $\Phi_H^{t,t'}$ we assume that

(1) there exists a normally hyperbolic and weakly invariant cylinder $\tilde{\Pi}$, which is a deformation of a standard cylinder $\{(p,q,t)\in\mathbb{R}^2\times\mathbb{T}^2\times\mathbb{T}:\ (p_1,q_1)=0.\}$;

(2) there is a continuous path Γ_c: $[0,1]\to H^1(\mathbb{T}^2,\mathbb{R})$ such that for any $c\in\Gamma_c$, the Aubry set entirely lies in the cylinder $\tilde{\Pi}$;

(3) for $c\in\Gamma_c$ if $\tilde{\mathcal{A}}(c)$ is an invariant 2-dimensional torus $\tilde{\Upsilon}_c$ lying in $\tilde{\Pi}$, it is a deformation of the torus $\{(p,q,t)\in\mathbb{R}^2\times\mathbb{T}^3:\ (p_1,q_1)=0,\ p_2=\text{const.}\}$

Let $\check{\pi}$: $\check{M}=\{q:\ q_1 \bmod 4\pi, q_2 \bmod 2\pi\}\to\mathbb{T}^2$ be a finite covering space of $\mathbb{T}^2$. The lift of $\tilde{\Pi}$ to $T^*\check{M}\times\mathbb{T}$ consists two copies, denoted by $\tilde{\Pi}_\ell$ and $\tilde{\Pi}_r$. For $c\in\Gamma_c$, if the Aubry set $\tilde{\mathcal{A}}(c)$ is an invariant torus $\tilde{\Upsilon}_c\subset\tilde{\Pi}$, its lift also consists of two components, $\tilde{\Upsilon}_{c,\ell}\subset\tilde{\Pi}_\ell$ and $\tilde{\Upsilon}_{c,r}\subset\tilde{\Pi}_r$. Let $\tilde{\Pi}_0$, $\tilde{\Pi}_{\ell,0}$, $\tilde{\Pi}_{r,0}$, $\tilde{\Upsilon}_{c,0}$, $\tilde{\Upsilon}_{c,\ell,0}$ and $\tilde{\Upsilon}_{c,r,0}$ denote the time-0-section of $\tilde{\Pi}$, $\tilde{\Pi}_\ell$, $\tilde{\Pi}_r$, $\tilde{\Upsilon}_c$, $\tilde{\Upsilon}_{c,\ell}$ and $\tilde{\Upsilon}_{c,r}$ respectively. Denote by π the projection such that $\pi(p,q,t)=(q,t)$, let $\Upsilon=\pi\tilde{\Upsilon}$. Let $\Gamma_c^*\subset\Gamma_c$ such that

$$\Gamma_c^*=\{c\in\Gamma_c:\ \tilde{\mathcal{A}}(c)\,\text{is an invariant torus}\}.$$

Let $B_D\in\mathbb{R}^2$ denotes a ball about the origin of radius D. We assume that $D>0$ is suitably large, such that for all $c\in\Gamma_c$ the c-minimal orbits of H entirely stay in $B_D\times\mathbb{T}^3$. Let $\mathfrak{B}_\epsilon\subset C^r(B_D\times\mathbb{T}^3,\mathbb{R})$ denote a ball about the origin of radius $\epsilon>0$.

Theorem A.1. *For any small $d_1>0$, there exists a set $\mathfrak{O}$ open-dense in $\mathfrak{B}_\epsilon$ such that for each $H_\delta\in\mathfrak{O}$, it holds for $H+H_\delta$ and simultaneously for all $c\in\Gamma_c^*$ that the diameter of each connected component of the set*

$$\mathcal{N}(c,\check{M})|_{t=0}\backslash(\mathcal{A}(c,\check{M})+\delta)|_{t=0}\neq\varnothing$$

is not larger than d_1.

Before the proof, we review some properties of the barrier functions. Starting from every point $x=(q,\tau)\in\mathbb{T}^3$ there exists at least one backward minimal curve $\gamma_{c,x}^-$: $(-\infty,\tau]\to\mathbb{T}^2$, namely $\gamma_{c,x}^-(\tau)=q$ and

$$\begin{aligned}&\int_t^\tau L(\dot\gamma_{c,x}^-(s),\gamma_{c,x}^-(s),s)-\langle c,\dot\gamma_{c,x}^-(s)\rangle ds\\ &\le\int_{t'}^\tau L(\dot\xi(s),\xi(s),s)-\langle c,\dot\xi(s)\rangle ds+(t-t')\alpha(c)\end{aligned}$$

holds for any absolutely continuous curve ξ: $[t',\tau]\to\mathbb{T}$ with $\xi(\tau)=\gamma_{c,x}^-(\tau)$, $\xi(t')=\gamma_{c,x}^-(t)$ with $t'=t \mod 2\pi$. It produces an orbit $(\dot\gamma_{c,x}^-(t)),\gamma_{c,x}^-(t))$ which approaches the Aubry set for c as $t\to-\infty$. Similarly, starting from every point $x=(q,\tau)\in\mathbb{T}^3$ there exists one forward minimal curve $\gamma_{c,x}^+$: $[\tau,\infty)\to\mathbb{T}^3$, the orbit $(\dot\gamma_{c,x}^+(t)),\gamma_{c,x}^+(t))$ approaches the Aubry set for c as $t\to\infty$.

According to the weak KAM theory, for almost every point $(q,t)\in\mathbb{T}^3$, the forward and backward orbit is uniquely determined by the forward and backward weak KAM solution respectively. The initial condition of the orbit is determined by the solution $u_c^\pm$ such that $\dot\gamma_{c,x}^\pm(\tau)=\partial_pH(\partial_qu_c^\pm(q,\tau)+c,q,\tau)$.

Given an Aubry class for $c\in\Gamma_c$ we can define its elementary weak KAM solution. In the covering space $\check{M}$, there are two Aubry classes for $c\in\Gamma_c$, $\tilde{\Upsilon}_{c,\ell}$ and $\tilde{\Upsilon}_{c,r}$. To define the elementary weak KAM solution $u_{c,\ell}^\pm$ for $\tilde{\Upsilon}_{c,\ell}$, we construct a perturbation $L(\dot q,q,t)\to L(\dot q,q,t)+V_\ell(q,t)$, where $V_\ell\ge 0$ and $\Upsilon_{c,r}\subset \text{supp}V_\ell\subset(\Upsilon_{c,r}+\delta)=$

$\{(q,t) : \mathrm{dist}((q,t),\Upsilon_{c,r}) \le \delta\}$. For $V_\ell \neq 0$, there exists a unique weak KAM solution $u^{\pm}_{c,V_\ell}$ modulo constant. For almost every point $(q,t) \in \check{M} \times \mathbb{T}$, the function $u^{\pm}_{c,V_\ell}$ determines a unique forward and backward minimal orbit $(\dot\gamma^{\pm}_{c,x}(t)), \gamma^{\pm}_{c,x}(t))$ such that $\dot\gamma^{\pm}_{c,x}(\tau) = \partial_p H(\partial_q u^{\pm}_{c,V_\ell}(q,\tau) + c, q, \tau)$ and the orbits approaches $\tilde\Upsilon_{c,\ell}$ as $t \to \pm\infty$ respectively. Let $V_{c,\ell} \downarrow 0$, the function $u^{\pm}_{c,V_\ell}$ converges to a function $u^{\pm}_{c,\ell}$ which is obviously a weak KAM solution for H, it is called elementary weak KAM solution for $\tilde\Upsilon_{c,\ell}$. The elementary weak KAM solution for $\tilde\Upsilon_{c,r}$ is defined in the same way, denoted by $u^{\pm}_{c,r}$.

For almost every point $(q,t) \in \check{M}\times\mathbb{T}\backslash\Upsilon_{c,\ell}$ the initial condition $(\partial_p u^{\pm}_{c,r}(q,t)+c, q, t)$ determines a forward (backward) c-minimal orbit that approaches $\tilde\Upsilon_{c,r}$ as $t \to \pm\infty$. For points $(q,t) \in \check{M} \times \mathbb{T}\backslash\Upsilon_{c,r}$, $u^{\pm}_{c,\ell}$ determines c-minimal orbit approaching $\tilde\Upsilon_{c,\ell}$.

Definition A.1. The barrier functions for $c \in \Gamma_c$ are defined as follows

$$B^{\ell}_c(q,t) = u^-_{c,\ell}(q,t) - u^+_{c,r}(q,t), \qquad B^r_c(q,t) = u^-_{c,r}(q,t) - u^+_{c,\ell}(q,t).$$

In the following, we only study B^{ℓ}_c. The arguments for B^r_c are the same. Since the backward weak KAM is semi-concave and the forward weak KAM is semi-convex, the barrier function is semi-concave. Therefore,

Lemma A.1. *At each minimal point of B^{ℓ}_c, both $u^-_{c,r}$ and $u^+_{c,\ell}$ are differentiable.*

Proof. By the definition, semi-concave function admits a local decomposition as the sum of a smooth function and a concave function. For a concave function u, one can define its sup-derivative $D^+u(x)$ at a point x such that $u(x+x') - u(x) \le \langle p, x'\rangle$ holds for any $p \in D^+u(x)$ which is a convex set. The function u is differentiable at x iff $D^+u(x)$ is a singleton.

Since B^{ℓ}_c is a sum of two semi-concave functions, its sup-derivative is the sum of the sup-derivatives of $u^-_{c,\ell}$ and $-u^+_{c,r}$. Therefore, $D^+B^{\ell}_c$ is a single point iff both $D^+u^-_{c,\ell}$ and $D^+(-u^+_{c,r})$ are singleton [CaC]. ◻

Lemma A.2. *If $(q,t) \in \check{M}\times\mathbb{T}\backslash((\Upsilon_{c,\ell}\cup\Upsilon_{c,r})+\delta)$ is a global minimal point of B^{ℓ}_c, then $(q,t) \subset \mathcal{N}(c,\check{M})$, namely, passing through the point (q,t) there is a c-semi-static curve in the covering space $\check{M} \times \mathbb{T}$.*

Proof. By the definition, $\partial u^-_{c,\ell} = \partial u^+_{c,r}$ holds at a global minimal point of B^{ℓ}_c, denoted by $x = (q,t)$. Therefore, the backward minimal curve $\gamma^-_{c,x}$ is joined smoothly to the forward minimal curve $\gamma^+_{c,x}$. They make up a c-semi-static curve for $\check{M}$. ◻

For a class $c \in \Gamma^*_c$, the covering space $\check{M} \times \mathbb{T}$ is divided into two annuli $\mathbb{A}_{c,r}$ and $\mathbb{A}_{c,\ell}$, bounded by $\Upsilon_{c,\ell}$ and $\Upsilon_{c,r}$. Clearly, one has $\check\pi\mathbb{A}_{c,r} = \check\pi\mathbb{A}_{c,\ell}$. The set $\mathcal{N}(c,\check{M})\backslash\mathcal{A}(c,\check{M})$ contains c-minimal curves which cross the annulus from one side to another side or vice versa. Each of the curves produces a homoclinic orbit to the torus $\tilde\Upsilon_c$.

Lemma A.3. *There is a finite partition of Γ_c: $\Gamma_c = \cup I_k$, each I_k is a segment of Γ_c. For each I_k there is an annulus $N_k \subset \mathbb{A}_{c,r}|_{t=0}$, two numbers $\delta > 0$ and $d > 0$ such that for each $c \in I_k \cap \Gamma^*_c$*

1. *$\mathrm{dist}(N_k, \Upsilon_{c,\ell} \cup \Upsilon_{c,r}) \ge \delta$;*
2. *each curve $(\gamma(t), t)$ lying in $(\mathcal{N}(c,\check{M})\backslash\mathcal{A}(c,\check{M})) \cap \mathbb{A}_{c,r}$ passes through N_k;*
3. *for each backward (forward) c-minimal curve γ, let $\{q_i = \gamma(2i\pi) \in N_k\}$, then $|q_i - q_j| \ge d$ if $i \neq j$.*

Proof. Because Γ_c is compact, the speed of each c-minimal orbit is uniformly upper bounded for all $c \in \Gamma_c^*$. Given an integer $m > 0$, there will be small $\delta_c > 0$ such that the period for each c-minimal curve to cross the annulus $N_c = \mathbb{A}_{c,r} \setminus ((\Upsilon_{c,\ell} \cup \Upsilon_{c,r}) + \delta_c)$ is not shorter than $4m\pi$. Because of the upper semi-continuity of Mañé set in c, there exists some $\delta_c' > 0$ such that $\Upsilon_{c',\ell} \cup \Upsilon_{c',r}$ does not touch N_c and the period for each c'-minimal curve to cross the annulus N_c is not shorter than $2m\pi$ provided $|c - c'| \le \delta_c'$ and $c' \in \Gamma_c^*$. The first two items are then proved if we notice Γ_c^* is compact.

For the third one, we notice that the condition $\gamma(2i\pi) = \gamma(2j\pi)$ for $i \neq j$ implies that γ is a curve in the Aubry set. It contradicts the assumption. Since both N_k and I_k are compact, such a constant $d > 0$ exists. $\square$

By the definition, the Aubry set $\tilde{\mathcal{A}}(c)$ is an invariant torus if $c \in \Gamma_c^*$. Its time-2π-section is an invariant circle lying in the cylinder. Fix one of the circles, we are able to parameterize other circle by the algebraic area bounded by the circles. Let us consider the twist map on the standard cylinder first. It is well-known that all invariant circles are Lipschitz with the constant C_L which depends on the twist condition only. Treating each circle as the graph of some periodic function and fixing one as $\tilde{\Upsilon}_{0,0}$ one can parameterize another circle by the algebraic area bounded by these two circles. The annulus bounded by the circle $\tilde{\Upsilon}_{\sigma,0}$ and $\tilde{\Upsilon}_{\sigma',0}$ contains a diamond, the height of the vertical diagonals is $\max_q |\tilde{\Upsilon}_{\sigma,0}(q) - \tilde{\Upsilon}_{\sigma',0}(q)|$ and the length of the horizontal diagonal is not shorter than $\frac{1}{C_L} \max_q |\tilde{\Upsilon}_{\sigma,0}(q) - \tilde{\Upsilon}_{\sigma',0}(q)|$. So, one has

$$\max_q |\tilde{\Upsilon}_{\sigma,0}(q) - \tilde{\Upsilon}_{\sigma',0}(q)| \le \sqrt{2C_L|\sigma - \sigma'|}.$$

A non-standard cylinder can be regarded as the image of the standard cylinder under some diffeomorphism. So, the $\frac{1}{2}$-Hölder continuity still holds, refer to the argument for Formula (8.4).

Each invariant circle corresponds to a unique $c \in \Gamma_c$ such that the Aubry set is the circle. The parameter σ is usually defined on a Cantor set, denoted by Σ. Because of the normal hyperbolicity of the cylinder, we have

LEMMA A.4. *For $\sigma, \sigma' \in \Sigma$, let $c = c(\sigma)$, $c' = c(\sigma')$. If $c, c' \in I_k$ and $q \in N_k$, then*

$$|B^\ell_{c(\sigma)}(q,0) - B^\ell_{c(\sigma')}(q,0)| \le C(\sqrt{|\sigma - \sigma'|} + |c - c'|).$$

Proof. For $c = c(\sigma)$ with $\sigma \in \Sigma$, the minimal measure is uniquely ergodic. There is only one pair of weak KAM solutions $u_c^\pm$ for the configuration space $\mathbb{T}^2$. With respect to the covering space $\check{M}$, we have introduced the elementary weak KAM solutions $u^\pm_{c,\ell}$ and $u^\pm_{c,r}$. Since the projection $\check{\pi}$ is an injection when it is restricted in the neighborhood $\Upsilon_{c,\imath} + \delta$ for $\imath = \ell, r$ respectively, for $(q,t) \in \Upsilon_c + \delta$ one has

$$u^\pm_{c,\ell}(\check{\pi}^{-1}(q,t) \cap (\Upsilon_{c,\ell} + \delta)) = u^\pm_{c,r}(\check{\pi}^{-1}(q,t) \cap (\Upsilon_{c,r} + \delta)) = u^\pm_c(q,t). \tag{A.1}$$

By the definition of weak KAM solutions, for any $t' < t$ one has

$$u^-_{c,\ell}(\gamma(t),t) - u^-_{c,\ell}(\gamma(t'),t') \le \int_{t'}^t (L(\dot\gamma(s),\gamma(s),s) - \langle c, \dot\gamma(s)\rangle) ds + (t - t')\alpha(c)$$

which becomes an equality when γ is a backward c-semi static curve. Assume $\gamma^-_{c,q}$ is a backward c-minimal curve such that $\gamma^-_{c,q}(0)=q$, we have

$$\begin{aligned}u^-_{c,\ell}(q,0)-u^-_{c,\ell}(\gamma^-_{c,q}(-2K\pi),0)=&\int_{-2K\pi}^0(L(\dot\gamma^-_{c,q}(s),\gamma^-_{c,q}(s),s)-\langle c,\dot\gamma^-_{c,q}(s)\rangle)ds\\&+2K\pi\alpha(c),\\u^-_{c',\ell}(q,0)-u^-_{c',\ell}(\gamma^-_{c,q}(-2K\pi),0)\le&\int_{-2K\pi}^0(L(\dot\gamma^-_{c,q}(s),\gamma^-_{c,q}(s),s)-\langle c',\dot\gamma^-_{c,q}(s)\rangle)ds\\&+2K\pi\alpha(c').\end{aligned}$$

Since N_k keeps away from $\Upsilon_{c,\ell}$, some $K>0$ exists such that for each $q\in N_k$, $c\in I_k$ and $q\in N_k$ one has $\gamma^-_{c,q}(-2K\pi)\in(\Upsilon_{c,\ell}+\delta)$. Since c and c' are located in a compact set Γ_c, the α function is convex and finite everywhere, there is some constant C_1 such that $|\alpha(c')-\alpha(c)|\le C_1|c-c'|$. Let $\bar\gamma^-_{c,q}$ be the lift of $\gamma^-_{c,q}$ to the universal covering space, one has $|\bar\gamma^-_{c,q}(0)-\bar\gamma^-_{c,q}(-2K\pi)|\le 2C_2K\pi$.

$$\begin{aligned}&u^-_{c',\ell}(q,0)-u^-_{c,\ell}(q,0)-(u^-_{c',\ell}(\gamma^-_{c,q}(-2K\pi),0)-u^-_{c,\ell}(\gamma^-_{c,q}(-2K\pi),0))\\&\le 2K\pi(C_1+C_2)|c-c'|.\end{aligned}$$

In the same way one can also obtain

$$\begin{aligned}&u^-_{c,\ell}(q,0)-u^-_{c',\ell}(q,0)-(u^-_{c,\ell}(\gamma^-_{c',q}(-2K\pi),0)-u^-_{c',\ell}(\gamma^-_{c',q}(-2K\pi),0))\\&\le 2K\pi(C_1+C_2)|c-c'|.\end{aligned}$$

For $u^+_{c,r}$, $u^+_{c'r}$ we also have similar inequalities. Therefore, it follows from (A.1) that some points $(q_\ell,0),(q_r,0)\in\Upsilon_c+\delta$ exist such that

$$\begin{aligned}|B_c(q,0)-B_{c'}(q,0)|\le{}&4K\pi(C_1+C_2)|c-c'|\\&+|u^-_c(q_\ell,0)-u^-_{c'}(q_\ell,0)-u^+_c(q_r,0)+u^+_c(q_r,0)|.\end{aligned}$$

By the assumption, both u^-_c and u^+_c are $C^{1,1}$ when they are restricted in $\Upsilon_c+\delta$. Due to the normal hyperbolic property, each $(p,q)\in\tilde\Pi_0$ has its stable and unstable fiber which is C^{r-1}-smoothly depends on the point (p,q). The fibers are defined by $\partial_q u^\pm_c+c$ and one has that

$$|\partial_q u^\pm_c-\partial_q u^\pm_{c'}+c-c'|\le C_3\sqrt{|\sigma-\sigma'|}$$

holds for some constant $C_3>0$, independent of c,c'. Combining above two inequalities, one finishes the proof of the lemma. □

We consider the c-minimal curves for $c\in I_k$. Because I_k is compact, there exists a constant $D>0$ such that $|\dot\gamma(t)|\le D$ holds for any c-minimal curve with $c\in I_k$. Let $\Omega_\tau=\{(q',q)\in\mathbb{R}^2\times\mathbb{R}^2:|q'-q|\le 2D\tau \text{ with } \tau>0\}$. We consider the action

$$S_{-\tau}(q',q)=\min_{\substack{\xi(-\tau)=q'\\ \xi(0)=q}}\int_{-t}^0 L(\dot\xi(s),\xi(s),s)ds.$$

For suitably small $\tau>0$, there exists a unique minimal curve if $(q',q)\in\Omega_\tau$. Indeed, because L is Tonelli, the second derivative of any solution $q(t)$ of the Euler-Lagrange equation is bounded by $|\ddot q|\le|\partial_{\dot q\dot q}L^{-1}(\partial_q L-\partial^2_{\dot qq}L\dot\gamma-\partial_{\dot qt}L)|$. Recall the Taylor formula

$$q(t')=q(t)+\dot q(t)(t'-t)+\frac12\ddot q(\lambda t+(1-\lambda)t')(t'-t)^2$$

holds for small $|t'-t|$, where both entries of $\lambda \in \mathbb{R}^2$ takes value in $[0,1]$. Therefore, for small $|t'-t|$, there is an one to one correspondence the initial speed $\dot\gamma(t)$ and the end point $\gamma(t')$. In this case, $S_{-\tau}(q',q)$ is C^r-differentiable in both q' and q. By the definition of weak KAM, for $c \in I_k$ one has

$$u_c^-(q,0) = \min_{q'\in\mathbb{T}^2,\, |q'-q|\le 2D\tau} (S_{-\tau}(q',q) - \langle c, q-q'\rangle + u_c^-(q',-\tau)).$$

We extend $S_{-\tau}$ smoothly to the whole $\mathbb{R}^2\times\mathbb{R}^2$ such that it satisfies the twist condition. Recall the quantities defined in Lemma A.3 such as the annulus N_k and the number $d>0$.

LEMMA A.5. *Let $S_\delta(q)$ be a C^r-function such that* $\max\{|q-q'| : q,q' \in \mathrm{supp} S_\delta\} \le d$, $\mathrm{supp} S_\delta \subset N_k$ *and* $\|S_\delta\|_{C^r}$ *is sufficiently small. Then, restricted on I_k, there exists a perturbation $H \to H' = H + H_\delta$ such that $\|H_\delta\|_{C^r}$ is small and the barrier function is subject to a translation*

$$B_c(q,0) \to B_c(q,0) + S_\delta(q) \qquad \forall\ c\in I_k,\ q\in \mathrm{supp} S_\delta.$$

Proof. The function $S_{-\tau}(q',q)$ induces a symplectic map between the time $-\tau$ section and the time-0-section Φ: $(p',q')\to(p,q)$

$$p = \frac{\partial S_{-\tau}}{\partial q}(q',q) \qquad p' = -\frac{\partial S_{-\tau}}{\partial q'}(q',q).$$

We introduce a smooth function κ such that $\kappa(q',q)=1$ if $|q'-q|\le K$ and $\kappa(q',q)=0$ if $|q'-q|\ge K+1$. Let Φ' be the map determined by the generating function $S_{-\tau}+\kappa S_\delta$, the symplectic diffeomorphism $\Psi = \Phi'\circ\Phi^{-1}$ is close to identity if S_δ is C^r-small. We choose a smooth function $\rho(s)$ with $\rho(-\tau)=0$, $\rho(0)=1$ and let Φ'_s be the symplectic map produced by $S_{-\tau}+\rho(s)\kappa S_\delta$ and let $\Psi_s = \Phi'_s\circ\Phi^{-1}$. Clearly, Ψ_s defines a symplectic isotopy between the identity map and Ψ. Thus, there is a unique family of symplectic vector fields X_s: $T^*\mathbb{T}^2 \to TT^*\mathbb{T}^2$ such that

$$\frac{d}{ds}\Psi_s = X_s\circ\Psi_s.$$

By the choice of perturbation, there is a simply connected and compact domain D such that $\Psi_s|_{T^*\mathbb{T}^2\setminus D} = id$. It follows that there exists a Hamiltonian $H_1(p,q,s)$ such that $X_s = J\nabla H_1(p,q,s)$. Re-parametrizing s by t, we can make X_s smoothly depend on t and smoothly connected to the zero vector field at $t=-\tau,0$. To show the smallness of dH' we apply a theorem of Weinstein [W]. A neighborhood of the identity in the symplectic diffeomorphism group of a compact symplectic manifold can be identified with a neighborhood of the zero in the vector space of closed 1-forms on the manifold. Since Hamiltomorphism is a subgroup of symplectic diffeomorphism, there is a function H', sufficiently close to H, such that $\Phi_{H'}^{-\tau,0} = \Phi_{H_1}^{-\tau,0}\circ\Phi_H^{-\tau,0}$.

For all $c\in\Gamma_c$, by the assumption, any backward (forward) c-minimal curve will not return back to $\mathrm{supp} S_{-\tau}$ if its initial point falls into the support. Let $u_{c,\imath}^{\pm,S_\delta}$ denotes the elementary weak KAM solution for the perturbed Hamiltonian

$$\begin{aligned} u_{c,\imath}^{-,S_\delta}(q,0) &= \min_{|q'-q|\le 2D\tau}(S_{-\tau}(q',q) + S_\delta(q) - \langle c,q-q'\rangle + u_{c,\imath}^-(q',-\tau)) \\ &= S_\delta(q) + \min_{|q'-q|\le 2D\tau}(S_{-\tau}(q',q) - \langle c,q-q'\rangle + u_{c,\imath}^-(q',-\tau)) \\ &= S_\delta(q) + u_{c,\imath}^-(q,0). \end{aligned}$$

Obviously, one has $u^{+,S_\delta}_{c,\imath}(q,0)=u^+_{c,\imath}(q,0)$. The lemma is proved because the barrier function is the difference of the two functions. $\square$

Proof of Theorem A.1. Given $q^*\in\mathbb{T}^2$, let $\mathbb{S}_{d_1}(q^*)=\{|q-q^*|\le d_1\}$ denote a square. Given a function $B\in C^0(\mathbb{S}_{d_1}(q^*),\mathbb{R})$, let

$$\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*),B)=\{q\in\mathbb{S}_{d_1}(q^*):B(q)=\min B\}.$$

Let π_i be the projection so that $\pi_i(q_1,q_2)=q_i$ $(i=1,2)$. A connected set V is said to be non-trivial for $\mathbb{S}_{d_1}(q^*)$ if $\pi_iV\cap\mathbb{S}_{d_1}(q^*)=\pi_i\mathbb{S}_{d_1}(q^*)$ holds for $i=1$ or 2. Otherwise, it is said to be trivial for $\mathbb{S}_{d_1}(q^*)$. Let $B^\ell_{c,\delta}$ be the barrier function for the Hamiltonian $H+H_\delta$ and the class c, we have

LEMMA A.6. *For any small $\epsilon>0$, there is a set $\mathfrak{O}$ open-dense in $\mathfrak{B}_\epsilon$ such that for each $H_\delta\in\mathfrak{O}$, it holds simultaneously for all $c\in I_k\cap\Gamma^*_c$ that the set* $\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*),B^\ell_{c,\delta})$ *is trivial for $\mathbb{S}_{d_1}(q^*)$ provided $\mathbb{S}_{d_1}(q^*)\subset N_k$ and $d_1<d/3$ is suitably small.*

Proof. The openness is obvious. To show the density, we construct the perturbations $H_\delta\in\mathfrak{B}_\epsilon$ such that the barrier function is under a translation $B_c(q,0)\to B_c(q,0)+S_\delta(q)$ for all $c\in I_k\cap\Gamma^*_c$ and $q\in\mathrm{supp}S_\delta$. Because of Lemma A.5, it works.

Recall the number $d>0$ defined in Lemma A.3. Given a square $\mathbb{S}_{d_1}(q^*)\subset N_k$ with $3d_1<d$, we consider the space of C^r-functions $\mathfrak{S}_1$, a function $S\in\mathfrak{S}_1$ if it satisfies the conditions that $\mathrm{supp}S\subset B_{d/2}(q^*)$ and S is constant in q_2 when it is restricted in $\mathbb{S}_{d_1}(q^*)$. Similarly, we can define $\mathfrak{S}_2$ such that $S\in\mathfrak{S}_2$ implies that $\mathrm{supp}S\subset B_{d/2}(q^*)$ and it is constant in q_1 when it is restricted in $\mathbb{S}_{d_1}(q^*)$.

In $\mathfrak{S}_i$ we define an equivalent relation $\sim$, two functions $S_1\sim S_2$ implies $S_1-S_2=$ constant when they are restricted on $\mathbb{S}_{d_1}(q^*)$. Obviously, $\mathfrak{S}_i/\sim$ is a linear space with infinite dimensions. For $S_1,S_2\in\mathfrak{S}_i/\sim$, $\|S_1-S_2\|_r$ measures the C^r-distance if they are regarded as the functions defined on $\mathbb{S}_{d_1}(q^*)$. We also use $\mathfrak{B}_{i,\epsilon}$ to denote a ball in $\mathfrak{S}_i/\sim$, about the origin of radius ϵ in the sense of the C^r-topology.

We claim that there exists a set $\mathfrak{O}_{1,\epsilon}$ open-dense in $\mathfrak{B}_{1,\epsilon}$ such that for each $S_\delta\in\mathfrak{O}_{1,\epsilon}$ it holds simultaneously for all $c\in I_k\cap\Gamma^*_c$ that

$$\pi_1\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*),B^\ell_c+S_\delta)\subsetneqq[q^*_1-d_1,q^*_1+d_1]. \tag{A.2}$$

Let $\mathfrak{F}_c=\{B^\ell_c(q,0):c\in\Gamma^*_c\}$ be the set of barrier functions. For $i=1,2$ we set

$$\mathfrak{Z}_i=\{B\in C^0(\mathbb{S}_{d_1}(q^*),\mathbb{R}):\pi_i\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*),B)=[q^*_i-d_1,q^*_i+d_1]\},$$

where $q^*=(q^*_1,q^*_2)$. If the density does not exist, there would be small $\epsilon>0$, for each $S_\delta\in\mathfrak{B}_{1,\epsilon}$, some $c\in\Gamma^*_c$ exists such that $B^\ell_c+S_\delta\in\mathfrak{Z}_1$. Let $\mathfrak{B}^k_{1,\epsilon}$ be the intersection of $\mathfrak{B}_{1,\epsilon}$ with a k-dimensional subspace. The box-dimension of $\mathfrak{B}^k_{1,\epsilon}$ in C^0-topology will not be smaller than k.

For any $B^\ell_c\in\mathfrak{F}_c$ there is only one $S_\delta\in\mathfrak{B}_{1,\epsilon}$ such that $B^\ell_c+S_\delta\in\mathfrak{Z}_1$. Otherwise, there would be $S'_\delta\neq S_\delta$ such that $B^\ell_c+S'_\delta\in\mathfrak{Z}_1$ also. As we have $B^\ell_c+S'_\delta=B^\ell_c+S_\delta+S'_\delta-S_\delta$ where $B^\ell_c+S_\delta\in\mathfrak{Z}_1$ and $S'_\delta\sim S_\delta$, which contradicts the definition of $\mathfrak{S}_1$. For $S_\delta\in\mathfrak{B}_{1,\epsilon}$, let $\mathfrak{S}_{S_\delta}=\{B^\ell_c\in\mathfrak{F}_c:B^\ell_c+S_\delta\in\mathfrak{Z}_1\}$. If the density does not exist, $\mathfrak{S}_{S_\delta}$ is non-empty. For any $S_\delta,S'_\delta\in\mathfrak{B}^k_{1,\epsilon}$, each $B^\ell_c\in\mathfrak{S}_{S_\delta}$ and each $B^\ell_{c'}\in\mathfrak{S}_{S'_\delta}$

one has

$$
\begin{aligned}
d(B^\ell_c, B^\ell_{c'}) &= \max_{q\in\mathbb{S}_{d_1}(q^*)} |B^\ell_c(q,0) - B^\ell_{c'}(q,0)| \\
&\geq \max_{|q_1-q_1^*|\leq d_1} \Big| \min_{|q_2-q_2^*|\leq d_1} B^\ell_c(q,0) - \min_{|q_2-q_2^*|\leq d_1} B^\ell_{c'}(q,0)\Big| \\
&= \max_{|q_1-q_1^*|\leq d_1} |S_\delta(q) - S'_\delta(q)| = d(S_\delta, S'_\delta)
\end{aligned}
\tag{A.3}
$$

where $q=(q_1,q_2)$ and $d(\cdot,\cdot)$ denotes the C^0-metric. It implies that the box-dimension of the set $\mathfrak{F}_c$ is not smaller than the box-dimension of $\mathfrak{B}^k_{1,\epsilon}$ in C^0-topology. Guaranteed by the modulus continuity of Lemma A.4, the box dimension of the set $\mathfrak{F}_c$ is not larger than 3. Therefore, we will obtain an absurdity if we choose $k\geq 4$.

In the same way, we can show that there exists a set $\mathfrak{O}_{2,\epsilon}$ open-dense in $\mathfrak{B}_{2,\epsilon}$ such that for each $S_\delta\in\mathfrak{O}_{2,\epsilon}$ it holds simultaneously for all $c\in I_k\cap\Gamma^*_c$ that

$$
\pi_2\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*), B^\ell_c + S_\delta) \subsetneq [q_2^* - d_1, q_2^* + d_1]. \tag{A.4}
$$

Therefore, $\exists$ arbitrarily small $S_{i,\delta}\in\mathfrak{B}_{i,\epsilon}$ such that $\pi_i\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*), B^\ell_c + S_{1,\delta}+S_{2,\delta})$ is trivial for $\mathbb{S}_{d_1}(q^*)$ and for all $c\in I_k\cap\Gamma^*_c$. Due to Lemma A.5 we obtain the density. □

To finish the proof of Theorem A.1, we split the annulus N_k equally into squares $\{\mathbb{S}_j = |q-q_j|\leq \frac{d_1}{5}\}$. For each $\mathbb{S}_j$, there exists an open-dense set $\mathfrak{O}_{k,j}\subset\mathfrak{B}_\epsilon$, for each $H_\delta\in\mathfrak{O}_{k,j}$ it holds simultaneously for all $c\in I_k\cap\Gamma^*_c$ that the set $\mathrm{Argmin}(\mathbb{S}_j, B^\ell_{c,\epsilon})$ is trivial for $\mathbb{S}_j$. The intersection $\cap\mathfrak{O}_{k,j}$ is still open-dense in $\mathfrak{B}_\epsilon$. For each $H_\delta\in\cap_{k,j}\mathfrak{O}_{k,j}$, it holds simultaneously for all $c\in\Gamma^*_c$ that the diameter of each connected component of the Mañé set is not larger than $\frac{4}{5}d_1$ if it keeps away from the Aubry set. □

Acknowledgement. The author is grateful to the referees for their comments which help a lot in the revision. He also got benefits from the comments of the referees for the papers [C17a, C17b, CZ]. The work is supported by NNSF of China (No.11790272 and No.11631006) and a program PAPD of Jiangsu Province, China.

REFERENCES

[A64] V. I. Arnold, *Instability of dynamical systems with several degrees of freedom,* Sov. Math., Dokl., 5 (1964), pp. 581–585; translation from Dokl. Akad. Nauk SSSR, 156 (1964), pp. 9–12.

[A66] V. I. Arnold, *The stability problem and ergodic properties for classical dynamical systems.* Proceedings of International Congress of Mathematicians, Moscow (1966) pp. 387–392, in V. I. Arnold-Collected Works. Springer Berlin Heidelberg (2014) pp. 107–113.

[B02] P. Bernard, *Connecting orbits of time dependent Lagrangian systems*, Ann. Inst. Fourier, Grenoble, 52 (2002), pp. 1533–1568.

[B07] P. Bernard, *Symplectic aspects of Mather theory*, Duke Math. J., 136 (2007), pp. 401–420.

[B08] P. Bernard, *The dynamics of pseudographs in convex Hamiltonian systems*, J. Amer. Math. Soc., 21 (2008), pp. 615–669.

[B10] P. Bernard, *Large normally hyperbolic cylinders in a priori stable Hamiltonian systems*, Annales H. Poincare, 11 (2010) pp. 929–942.

[BKZ] P. Bernard, V. Kaloshin and K. Zhang, *Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders*, Acta Math., 217 (2016), pp. 1–79.

[Bs] U. Bessi, *An approach to Arnold's diffusion through the calculus of variations,* Nonlinear Anal., 26:6 (1996), pp. 1115–1135.

[BCV] U. Bessi, L. Chierchia and E. Valdinoci, *Upper bounds on Arnold diffusion times via Mather theory,* J. Math. Pures Appl., 80:1 (2001), pp. 105–129.

[CaC] P. Cannarsa and W. Cheng, *Generalized characteristics and Lax-Oleinik operators: global theory*, Calc. Var. Partial Differential Equations, 56:5 (2017) Art. 125.

[C11] C.-Q. Cheng, *Non-existence of KAM torus*, Acta Math. Sinica, 27 (2011), pp. 397–404.
[C13] C.-Q. Cheng, *Arnold diffusion in nearly integrable Hamiltonian systems*, arXiv:1207.4016v2 (2013).
[C17a] C.-Q. Cheng, *Uniform hyperbolicity of invariant cylinder*, J. Diff. Geometry, 106 (2017), pp. 1–43.
[C17b] C.-Q. Cheng, *Dynamics around the double resonance*, Cambridge J. Mathematics, 5:2 (2017), pp. 153–228.
[C17c] C.-Q. Cheng, *Addendum to: "Dynamics around the double resonance"*, Cambridge J. Mathematics, 5:4 (2017), p. 571.
[CY1] C.-Q. Cheng and J. Yan, *Existence of diffusion orbits in a priori unstable Hamiltonian systems,* J. Differential Geometry, 67 (2004), pp. 457–517.
[CY2] C.-Q. Cheng and J. Yan, *Arnold diffusion in Hamiltonian Systems: a priori Unstable Case,* J. Differential Geometry, 82 (2009), pp. 229–277.
[CZ] C.-Q. Cheng and M. Zhou, *Global normally hyperbolic cylinders in Lagrangian systems*, Math. Res. Lett., 23 (2016) , pp. 685–705.
[DLS1] A. Delshams, R. de la Llave and T. M. Seara, *Geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristic and rigorous verification of a model*, Memoirs Amer. Math. Soc., 179:844 (2006).
[DLS2] A. Delshams, R. de la Llave and T. M. Seara, *Instability of high dimensional Hamiltonian systems: multiple resonance do not impede diffusion*, Advance Math., 294 (2016), pp. 689–755.
[DH1] A. Delshames and G. Huguet, *Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems*, Nonlineairty, 22 (2009), pp. 1997–2077.
[DH2] A. Delshames A. and G. Huguet, *A geometric mechanism of diffusion: rigorous verification in a priori unstable Hamiltonian systems*, J. Diff. Eqns., 250 (2011), pp. 2601–2623.
[FM] E. Fontich and P. Martin, *Arnold diffusion in perturbations of analytic integrable Hamiltonian systems*, Discrete Contin. Dyn. Syst., 7 (2001), pp. 61–84.
[Fa] A. Fathi, *Weak KAM theorem in Lagrangian dynamics*, preprint (7-th preliminary version), (2005).
[GL] M. Gidea and R. de la Llave, *Topological methods in the instability problem of Hamiltonian systems*, Discrete and Continuous Dynamical Systems, 14 (2006), pp. 294–328.
[GR1] M. Gidea and C. Robinson, *Shadowing orbits for transitive chains of invariant tori alternating with Birkhoff zone of instability* Nonlinearity, 20 (2007), pp. 1115–1143.
[GR2] M. Gidea and C. Robinson, *Obstruction argument for transition chains of tori interspersed with gaps*, Discrete and Continuous Dynamical Systems series S, 2 (2009), pp. 393–416.
[KL1] V. Kaloshin and M. Levi, *An example of Arnold diffusion for near-integrable Hamiltonians*, Bulletin AMS, 45 (2008), pp. 409–427.
[KL2] V. Kaloshin and M. Levi, *Geometry of Arnold diffusion*, SIAM Review, 50 (2008), pp. 702–720.
[KZ1] V. Kaloshin and K. Zhang, *A strong form of Arnold diffusion for two and a half degrees of freedom,* arXiv:1212.1150v2 (2013).
[KZ2] V. Kaloshin and K. Zhang, *Arnold diffusion for smooth convex systems of two and a half degrees of freedom*, Nonlinearity, 28 (2015), pp. 2699–2720.
[LC] X. Li and C.-Q. Cheng, *Connecting orbits of autonomous Lagrangian systems*, Nonlinearity, 23 (2009), pp. 119–141.
[Lo] P. Lochak, *Canonical perturbation theory via simultaneous approximation*, Russian Math. Surveys, 47 (1992), pp. 57–133.
[Man] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity, 9 (1996), pp. 273–310.
[Mar] J. P. Marco, *Arnold diffusion in cusp-residual nearly integrable systems on* $\mathbb{A}^3$, preprint (2015).
[M91] J. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems,* Math. Z., 207:2 (1991), pp. 169–207.
[M93] J. Mather, *Variational construction of connecting orbits,* Ann. Inst. Fourier (Grenoble), 43:5 (1993), pp. 1349–1386.
[M04] J. Mather, *Arnold diffusion, I: Announcement of results,* J. Mathematical Sciences, 124:5 (2004), pp. 52759–5289. (Russian translation in Sovrem. Mat. Fundam. Napravl, 2 (2003), pp. 116–130).
[Mo] J. K. Moser, *On the volume elements on a manifold*, Trans. Amer. Math. Soc., 120 (1965), pp. 286–294.

[Ox] J. C. Oxtoby, *Measure and category: A survey of the analogies between topological and measure spaces,* Vol. 2. Springer Science & Business Media, 2013.

[Tr] D. V. Treschev, *Evolution of slow variables in a priori unstable Hamiltonian systems*, Nonlinearity, 17 (2004), pp. 1803–1841.

[X] Z. Xia, *Arnold diffusion: a variational construction*, Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998), Doc. Math. 1998, Extra Vol. II, pp. 867–877.

[W] A. Weinstein, *Symplectic manifolds and their Lagrangian submanifolds*, Advances in Math., 6 (1971), pp. 329–346.

[Zh1] M. Zhou, *Hölder regularity of barrier functions in a priori unstable case*, Math. Res. Lett., 18 (2011), pp. 77–94.

[Zh2] M. Zhou, *Non-degeneracy of extremal points*, Chin. Ann. Math., 35B:6 (2014), pp. 1–6.