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 $\delta$-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

\[ H(p, q) = h(p) + \epsilon P(p, q), \quad (p, q) \in \mathbb{R}^n \times T^n. \]  

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

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

\[ \mathcal{C}_a = \{ \lambda P : P \in \mathcal{R}_a, \lambda \in R_P \}. \]
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 $\delta > 0$, $E \geq \min h$ with $h^{-1}(E) \subset B_D$ and any two points $p^*, p^* \in h^{-1}(E)$, there exists a cusp-residual set $\mathcal{C}_o \subset C^r(B_D \times \mathbb{T}^3, \mathbb{R})$ such that for each $eP \in \mathcal{C}_o$ there exists an orbit $(p(t), q(t))$ of $\Phi^t_H$ which intersects the $\delta$-neighborhood of $p^*$ and of $p^*$, namely, some $t^*, t^* \in \mathbb{R}$ exist such that $\|p(t^*) - p^*\| < \delta$ 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 $\partial^2_{p^*p^*} H$ 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 $\tilde{\pi} : \hat{M} \to \mathbb{T}^n$ be a finite covering of $\mathbb{T}^n$, let $\mathcal{N}(c, \hat{M}), \mathcal{A}(c, \hat{M})$ denote the Mañé set, Aubry set with respect to $\hat{M}$. 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 $\tilde{\pi} : \hat{M} \to M$ such that

1. in time-periodic case: $\tilde{\pi}\mathcal{N}(c, \hat{M})|_{t=0}\setminus(\mathcal{A}(c, \hat{M})|_{t=0} + \delta) \neq \emptyset$ is totally disconnected, where $\mathcal{A}(c, \hat{M})|_{t=0} + \delta = \{ x : \text{dist}(x, \mathcal{A}(c, \hat{M})|_{t=0}) \leq \delta \}$;

2. in autonomous case: $\tilde{\pi}\mathcal{N}(c, \hat{M})|_{\Sigma}\setminus(\mathcal{A}(c, \hat{M}) + \delta) \neq \emptyset$ is totally disconnected, where $\Sigma$ is a section of $\hat{M}$.

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

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 $\Sigma_c$ non-degenerately embedded $(n-1)$-dimensional torus by assuming a smooth injection $\varphi : \mathbb{T}^{n-1} \to \mathbb{T}^n$ such that $\Sigma_c$ is the image of $\varphi$, and the induced map $\varphi_* : H_1(\mathbb{T}^{n-1}, \mathbb{Z}) \to H_1(\mathbb{T}^n, \mathbb{Z})$ 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 $\gamma$ transversally intersects $\Sigma_c$. 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. $\mathbb{V}_c$ 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 $\mathbb{V}^\perp_c = \ker i_{U,*}$.

**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 $\Gamma: [0,1] \rightarrow H^1(\mathbb{T}^n, \mathbb{R})$ such that $\Gamma(0) = c$, $\Gamma(1) = c'$, $\alpha(\Gamma(s))$ keeps constant along $\Gamma$, 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 cohomology classes $c, c' \in H^1(M, \mathbb{R})$ are joined by a generalized transition chain if a continuous curve $\Gamma: [0,1] \rightarrow H^1(M, \mathbb{R})$ 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 $\mathcal{A}(\Gamma(0))$ provided $|s' - s|$ is small;
2. there is $\delta > 0$, for each $s' \in (s - \delta, s + \delta)$, $\Gamma(s')$ 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 $\alpha$-limit and $\omega$-limit set respectively.

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

**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 $\mathcal{L}_H(p(t), q(t), t) = (\dot{q}(t), q(t), t)$ 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), $\|p(t) - \tilde{c}\| \ll 1$ holds along any $\tilde{c}$-semi static orbit for the perturbed system.

Given two points $p^*, p^* \in h^{-1}(E)$ and small $\delta > 0$, $\exists$ points $\tilde{p}^*, \tilde{p}^* \in h^{-1}(E)$ and vectors $k^*, k^* \in \mathbb{Z}^3 \setminus \{0\}$ such that $\|p^* - \tilde{p}^*\| < \frac{\delta}{2}$, $\|p^* - \tilde{p}^*\| < \frac{\delta}{2}$, $\langle k^*, \partial h(\tilde{p}^*) \rangle = 0$ and $\langle k^*, \partial h(\tilde{p}^*) \rangle = 0$. One can choose $(\tilde{p}^*, k^*)$ and $(\tilde{p}^*, k^*)$ such that $k^*$ and $k^*$ are totally irreducible. A vector $k = (k_1, k_2, k_3) \in \mathbb{Z} \setminus \{0\}$ 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 $k \in \mathbb{Z}^3 \setminus \{0\}$, one can choose totally irreducible $k' \in \mathbb{Z}^3 \setminus \{0\}$ 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

$$|p| = \max\{|p_1|, |p_2|, |p_3|\}, \quad \|p\| = \left(\sum_{i=1}^3 p_i^2\right)^{\frac{1}{2}}.$$
$\Gamma_{k^*} = \Gamma_{k^*}$. In both cases, one has a path $\Gamma$ connecting $\bar{p}^*$ to $\bar{p}^*$. If $\Gamma_{k^*}$ intersects $\Gamma_{k^*}$, it starts from the point $\bar{p}^*$, moves along the circle $\Gamma_{k^*}$ until it reaches the intersection point of $\Gamma_{k^*}$ with $\Gamma_{k^*}$, after that, it moves along the circle $\Gamma_{k^*}$ until it arrives at the point $\bar{p}^*$. If $\Gamma_{k^*} = \Gamma_{k^*}$, $\Gamma$ is just a piece of the circle, connecting $\bar{p}^*$ to $\bar{p}^*$, see the figure below.

![Figure](image_url)

Treating $\Gamma$ 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 $\Gamma$.

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 $h \in C^r(\mathcal{D}, \mathbb{R})$ 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 $\mathcal{C}_e \subset C^r(\mathcal{D} \times \mathbb{T}^3, \mathbb{R})$ such that for each $\epsilon \mathcal{P} \in \mathcal{C}_e$ and any two points $p^*, p^* \in h^{-1}(E)$, there is a transition chain that connects the class $\bar{c}$ to the class $\bar{c}'$ which satisfy the condition $|\alpha(\bar{c}) - \alpha(\bar{c}')| = E$, $|p^* - \bar{c}| < \delta$ and $|p^* - \bar{c}'| < \delta$.

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 $\bar{c} = (c, -\alpha(c))$, the section $\Sigma_{\bar{c}}$ is chosen for $\mathbb{T}^{n+1}$, $\mathbb{V}_{\bar{c}}$ and $\mathbb{V}_{\bar{c}}^\perp$ 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 $|\tilde{\Gamma}(s) - \tilde{\Gamma}(s_0)| < \mathbb{V}_{\bar{c}}^\perp$ whenever $s \in [0, 1]$ and $|s - s_0| < \delta$.

3. **Normal form.** Given an irreducible integer vector $k' \in \mathbb{Z}^3 \setminus \{0\}$, 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 $k'' \in \mathbb{Z}^3 \setminus \{0\}$ 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 $k' = (k'_1, k'_2, k'_3) \in \mathbb{Z}^3 \setminus \{0\}$ is totally irreducible. In this case, there exist $k^*, k^* \in \mathbb{Z}^3$ such that the matrix $M_0^r = (k', k^*, k^*)$ is uni-modular. Indeed, if $k'$ contains two non-zero entries, e.g. $k'_1, k'_2 \neq 0$, we set $k^* = (0, 0, 1)$ and $k^* = (k'_1, k'_2, 0)$ such that $k'_1 k'_2 - k'_3 k'_3 = 1$. If $k' = e_1$, we set $k^* = e_2$ and $k^* = e_3$, 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 $\mathfrak{M}_0: (\tilde{p}, \tilde{q}) \to (p, q)$ such that $q = M_0^{-1} \tilde{q}$, $p = M_0 \tilde{p}$, we obtain the Hamiltonian $\tilde{H} = \bar{h} + \epsilon \mathcal{P}$ where $\bar{h} = \mathfrak{M}_0^* h$ and $\mathcal{P} = \mathfrak{M}_0^* \mathcal{P}$. It holds along the path $\tilde{\Gamma}_{k'} = M_0^{-1} \Gamma_{k'}$ that $\partial \bar{h}(\tilde{p}) = (0, \omega_2, \omega_3)$. 


In this case, the first and the second resonant condition are \( M_0^{-1} k' = \tilde{k}' = e_1 \) and \( M_0^{-1} k'' = \tilde{k}'' = (0, \tilde{k}_2'', \tilde{k}_3'') \) respectively. If we introduce the canonical transformation of coordinates \( \mathcal{H}^1(u, v) \rightarrow (\tilde{p}, \tilde{q}) \) further

\[
\tilde{q} = M^{-1} u, \quad \tilde{p} = M' v,
\]

where

\[
M' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{k''}{k_2''} & 1 \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{k''}{k_2''} \\ 0 & 0 & 1 \end{bmatrix}
\]

if \( |\tilde{k}_2''| \leq |\tilde{k}_3''| \) and

\[
M' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{k''}{k_2''} \\ 0 & 0 & 1 \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{k''}{k_3''} & 1 \end{bmatrix}
\]

if \( |\tilde{k}_2''| \geq |\tilde{k}_3''| \). The function \( \mathcal{H} \tilde{H} \) is 2\( \pi \)-periodic in \((u_1, u_2)\), 2\( |\tilde{k}'| \)\( \pi \) in \( u_3 \) if \( |\tilde{k}_2''| \geq |\tilde{k}_3''| \) and it is 2\( \pi \)-periodic in \((u_1, u_3)\), 2\( |\tilde{k}''| \)\( \pi \) in \( u_2 \) if \( |\tilde{k}_3''| \geq |\tilde{k}_2''| \).

By the construction of \( \mathcal{H} \tilde{H} \) respects two symmetries in the variable \( u \).

**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

\[
f(u + 2\pi M e_i) = f(u), \quad \forall u \in \mathbb{R}^n, \ e_i \in \mathbb{Z}^n.
\]

Since \( \mathcal{H} \tilde{H} \) is 2\( \pi \)-periodic in \((u_1, u_2)\), 2\( |\tilde{k}'| \)\( \pi \)-periodic in \( u_3 \) in the case that \( |\tilde{k}_2''| \geq |\tilde{k}_3''| \) and 2\( \pi \)-periodic in \((u_1, u_3)\), 2\( |\tilde{k}''| \)\( \pi \) in \( u_2 \) in the case that \( |\tilde{k}_3''| \geq |\tilde{k}_2''| \),

\[
\mathcal{H} \tilde{H}(u, v) = \sum_{k \in \mathbb{Z}^3} \tilde{H}_k(M'v)e^{i(k, M^{-1}u)},
\]

\( \mathcal{H} \tilde{H} \) respects two symmetries in the variable \( u \), \( M \) and \( \text{diag}\{1, 1, |\tilde{k}'|\} \) for \( |\tilde{k}_2''| \geq |\tilde{k}_3''| \), \( M \) and \( \text{diag}\{1, |\tilde{k}'|, 1\} \) for \( |\tilde{k}_2''| \leq |\tilde{k}_3''| \) respectively.

At a double point \( \tilde{p}'' \), the rotation vector \( \tilde{\omega} = (0, \tilde{\omega}_2, \tilde{\omega}_3) = \partial \tilde{H}(\tilde{p}'') \) is rational, i.e. \( \exists T > 0 \) so that \( T \tilde{\omega} \in \mathbb{Z}^n \setminus \{0\} \). If \( t \omega \notin \mathbb{Z}^n \forall t \in (0, T) \), \( T = T(\omega) \) is called the (minimal) period. Since \( k'' \tilde{\omega}_2 + k'' \tilde{\omega}_3 = 0 \), \( T = |\tilde{k}''| \). We consider those double resonant points \( \{\tilde{p}'' \in \tilde{\Gamma}_k\} \) such that \( T = T(\partial \tilde{H}(\tilde{p}'')) \leq K^*e^{-\frac{1}{\kappa}(1-3\epsilon)} \) with \( \kappa \in (0, \frac{1}{6}) \) and \( K^* > 0 \) is independent of \( \epsilon \). In this case, \( |\tilde{k}''|^{\frac{1}{\kappa}} \rightarrow 0 \) as \( \epsilon \rightarrow 0 \).

**Lemma 3.1.** Assume the second resonant condition at \( \tilde{p}'' \in \tilde{\Gamma}_k \), \( k'' = (0, \tilde{k}_2'', \tilde{k}_3'') \), \( \delta' \in (0, 1/2) \). Then, there exists a small number \( \eta_0 > 0 \) such that for each \( \epsilon \in (0, \epsilon_0] \), restricted on the level set \( \tilde{H}^{-1}(E) \) contained in \( \Sigma_{\epsilon} \times T^3 \) with

\[
\tilde{\Sigma}_{\epsilon} = \{\tilde{p} : |\tilde{p} - \tilde{p}''| \leq K^{-1} \epsilon\},
\]

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

\[
G_{\epsilon}(x, y, \theta) = \frac{1}{2} (By, y) - V(x_1, |\tilde{k}''(x_2)|) + R_{\epsilon}(x, y, \frac{\omega_3}{\sqrt{\epsilon}} \theta),
\]

\(^1\)In [C17a], the matrix \( M_0 M \) is set to be uni-modular. It is not always possible and not necessary.
where the $2 \times 2$ matrix $B$ is positive definite, $V \in C^r$ is $2\pi$-periodic in $(x_1, |\vec{k''}|x_2)$, $R_\epsilon(x, y, \vartheta) \in C^{r-2}(\mathbb{T}^2 \times \Sigma'_e \times |\vec{k''}|\mathbb{T}, \mathbb{R})$, $\vartheta = \omega_3 \sqrt{\epsilon^{-1}}\theta$ and $\Sigma'_e$ satisfies the condition
\[
\{y : |y| \leq (1 - \delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\} \subseteq \Sigma'_e \subseteq \{y : |y| \leq (1 + \delta')K^{-1}\epsilon^{\kappa-\frac{1}{2}}\}.
\]
Restricted in $\mathbb{T}^2 \times \Sigma'_e \times |\vec{k''}|\mathbb{T}$, 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
\[
\|R_\epsilon\|_{C^{r-2}(\mathbb{T}^2 \times \Sigma'_e \times |\vec{k''}|\mathbb{T}, \mathbb{R})} \leq a_0\epsilon^k,
\]
if it is treated as a function in $(x, y, \vartheta)$. Finally, the remainder $R_\epsilon(x, y, \vartheta)$ respects the symmetries $M$ and $\text{diag}\{1, 1, |\vec{k''}|\}$ in $(x, \vartheta)$ and the symplectic transformation is uniformly bounded for any second resonant condition $\vec{k''}$.

Remark. Because $V$ is independent of $\vartheta$, the symmetries $M$ and $\text{diag}\{1, 1, |\vec{k''}|\}$ 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-$2\pi$-map $\Phi_{\epsilon F} = \Phi_{\epsilon F}^t|_{t=2\pi}$ of the Hamiltonian flow generated by the function $\epsilon F(p, q)$. The function $F$ solves the homological equation
\[
\left(\frac{\partial h}{\partial p} (p''), \frac{\partial F}{\partial q}\right) = -\bar{P}(\bar{p}, \bar{q}) + Z(\bar{p}, \bar{q})
\]
where
\[
Z(\bar{p}, \bar{q}) = \frac{1}{T} \int_0^T \bar{P}(\bar{p}, \bar{q} + \bar{\omega}t) dt = \sum_{(\ell_1, \ell_2) \in \mathbb{Z}^2} P_{\ell}(\bar{p}) e^{i(\ell_1(\bar{k'}\bar{q}) + \ell_2(\bar{k''}\bar{q}))}, \tag{3.3}
\]
Clearly, the function
\[
F(p, q) = \frac{1}{T} \int_0^T \bar{P}(\bar{p}, \bar{q} + \bar{\omega}t) dt
\]
solves the homological equation and $\|F\| < T\|\bar{P}\|$.

Under the transformation $\Phi_{\epsilon F}$ we obtain a new Hamiltonian
\[
\Phi_{\epsilon F}^* \bar{H} = \bar{h}(\bar{p}) + \epsilon Z(\bar{p}, \bar{q}) + \epsilon \left(\frac{\partial h}{\partial p}(\bar{p}) - \frac{\partial h}{\partial p}(p''), \frac{\partial F}{\partial q}\right) + \epsilon^2 \frac{1}{2} \int_0^1 (1 - t) \{\{\bar{H}, F\}, F\} \circ \Phi_{\epsilon F}^t dt. \tag{3.4}
\]

To simplify the situation further, we introduce another canonical transformation of coordinates $\mathcal{M} : (u, v) \to (\bar{p}, \bar{q})$ 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 $|\bar{k''}|/|\bar{k''}| \leq 1$ in the first case and $|\bar{k''}|/|\bar{k''}| \leq 1$ in the second case.

We consider the case that $|\bar{k''}| \geq |\bar{k''}|$. 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 $(\bar{k''}, \bar{k''})$ is irreducible, it follows from (3.1) that $\mathcal{M}^* \Phi_{\epsilon F}^* \bar{H}$ is $2\pi$-periodic in $(u_1, u_2)$ and $2|\bar{k''}|\pi$-periodic in $u_3$. 


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

$$\langle k',q \rangle = \langle k', (M M_0)^{-1} u \rangle = u_1, \quad \langle k'', q \rangle = \langle k'', (M M_0)^{-1} u \rangle = \tilde{k}_2'' u_2.$$ 

It follows from Formula (3.3) and the transformation (3.1) that the resonant term has the form of $\mathcal{M}^* \mathcal{M}^* Z(p, (k', q), (k'', q)) = Z''(v, u_1, \tilde{k}_2'' u_2)$.

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

$$H' = h'(v) + \epsilon Z'(v, u_1, \tilde{k}_2'' u_2) + \epsilon \left( \frac{\partial h'}{\partial v} (v) - \frac{\partial h'}{\partial u} \frac{\partial F'}{\partial u} \right) + \frac{\epsilon^2}{2} \mathcal{M}^* \int_0^1 (1 - t) \{ \{ H, F \}, F \} \circ \Phi_{\epsilon,F} dt,$$

where $v'' = (M_0 M)^{-1} p''$. The function $H'$ determines its Hamiltonian equation with the generating function as the following:

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

For this equation we introduce another transformation

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

where we use the notation $\tilde{y} = (y_1, y_2, y_3) = (y_1, y_3)$ 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:

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

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

$$\tilde{R}_{e,1} = Z'(v'' + \sqrt{\epsilon} \tilde{y}, \tilde{x}) - Z'(v'', \tilde{x}),$$

$$\tilde{R}_{e,2} = \left( \frac{\partial h'}{\partial v} (v'' + \sqrt{\epsilon} \tilde{y}) - \frac{\partial h'}{\partial u} \frac{\partial F'}{\partial u} (v''), \frac{\partial F'}{\partial u} \right),$$

$$\tilde{R}_{e,3} = \frac{\epsilon}{2} \mathcal{M}^* \int_0^1 (1 - t) \{ \{ \tilde{H}, F \}, F \} \circ \Phi_{\epsilon,F} dt.$$ 

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

$$\| \tilde{R}_{e,i} \|_{C^{r-2}} \leq a_2 \epsilon^k, \quad i = 1, 2, 3, \quad (3.8)$$

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

$$|\tilde{R}_{e,1}| \leq 2 \| P' \|_{C^1} \sqrt{\epsilon} K^{-1} \epsilon^\kappa \leq 2 \| P' \|_{C^1} K^{-1} \epsilon^k,$$

$$|\tilde{R}_{e,2}| \leq m' \sqrt{\epsilon} K^{-1} \epsilon^k T \| P'' \|_{C^1} \leq m' \| P' \|_{C^1} \epsilon^k,$$

$$| (\mathcal{M}^*)^{-1} \tilde{R}_{e,3} | \leq \frac{\epsilon}{2} \| P' \|_{C^2}^2 \| H' \|_{C^1} T^2 = \frac{1}{2} (K^* \| P' \|_{C^2})^2 \| H' \|_{C^1} \epsilon^{\frac{1}{2} + 2\kappa}.$$
The estimate on the derivatives of the terms can also be done inductively.

We introduce another coordinate rescaling further

$$\theta = \frac{\sqrt{\epsilon}}{\omega_3} x_3, \quad I = \frac{\omega_3}{\sqrt{\epsilon}} y_3,$$  \hfill (3.9)

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

$$\tilde{G}_\epsilon(x, y, I, \theta) = I + \frac{1}{2} \langle \tilde{B}(y, \frac{\sqrt{\epsilon}}{\omega_3} I), (y, \frac{\sqrt{\epsilon}}{\omega_3} I) \rangle - V(x_1, |\tilde{k}''| x_2)$$

$$+ \tilde{R}_h(y, \frac{\sqrt{\epsilon}}{\omega_3} I) + \tilde{R}_\epsilon(x_1, x_2, \frac{\omega_3}{\sqrt{\epsilon}} \theta, p'' + (\sqrt{\epsilon} y, \frac{\sqrt{\epsilon}}{\omega_3} I) \rangle$$  \hfill (3.10)

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

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

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

$$\tilde{B} = \begin{bmatrix} B & B' \\ B'' & B'' \end{bmatrix}$$

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

Obviously, restricted on the domain $\{|y| \leq 2K^{-1}\epsilon^{\kappa-\frac{3}{2}}, |I| \leq 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$, there exists a constant $a_3 = a_3(h, E) > 0$ such that

$$\|\tilde{R}_h\|_{C^{r-2}} \leq a_4 \epsilon^\kappa.$$  \hfill (3.11)

Let $\Omega_\epsilon$ be the image of $\tilde{\Sigma}_\epsilon \times \mathbb{T}^3$ under the maps $\Phi_{\epsilon E'}$, (3.1), (3.6) and (3.9). Since the transformation $\Phi_{\epsilon E'}$ is close to identity, $|k_0^2| > |k_0^3|$ is assumed, each section of $\Omega_\epsilon$ where $(x, \theta)$ keeps constant lies in $\{|y| \leq 2K^{-1}\epsilon^{\kappa-\frac{3}{2}}, |I| \leq 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$. Restricted in $\Omega_\epsilon$, 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 $\tilde{G}_\epsilon(x, y, \theta, -\omega_3 \epsilon, \theta) = 0$.

Indeed, restricted in the domain $\{|y| \leq 2K^{-1}\epsilon^{\kappa-\frac{3}{2}}, |I| \leq 2K^{-1}|\omega_3|^{-1}\epsilon^{\kappa-1}\}$, a constant $a_4 = a_4(h) > 0$ exists such that $|\sqrt{\epsilon} (B', y)| \leq a_4 K^{-1}\epsilon^\kappa$ and $|\epsilon (B y, y)| \leq a_4 K^{-2}\epsilon^{2\kappa}$. It guarantees that the solution $I = -G_0(x, y)$ of the following quadratic equation in $I$

$$\tilde{G}_0 = I + \frac{1}{2} \langle \tilde{B}(y, \frac{\sqrt{\epsilon}}{\omega_3} I), (y, \frac{\sqrt{\epsilon}}{\omega_3} I) \rangle - V(x_1, |\tilde{k}''| x_2)$$

$$= \left(1 + \frac{\sqrt{\epsilon}}{\omega_3} \epsilon B', y) \right) I + \frac{1}{2} \langle By, y \rangle + \frac{B''}{2\omega_3^2} \epsilon I^2 - V(x_1, |\tilde{k}''| x_2) = 0$$

has the form of

$$G_0 = \frac{\omega_3^2 (1 + \frac{\sqrt{\epsilon}}{\omega_3} \epsilon B', y) \rangle}{B'' \epsilon} \left(1 - \left(1 - \frac{2B'' \epsilon (\frac{1}{2} \epsilon B, y) - V(x_1, |\tilde{k}''| x_2))}{\omega_3^2 (1 + \frac{\sqrt{\epsilon}}{\omega_3} \epsilon B', y) \rangle^2 \right)^{\frac{1}{2}} \right)$$

$$= \frac{1}{2} \langle By, y \rangle - V(x_1, |\tilde{k}''| x_2) + R_0(x, y)$$
and some \(a_5(h, E) > 0\) exists such that restricted in the domain \(\mathbb{T}^2 \times \{|y| \leq 2K^{-1}e^{\kappa - \frac{1}{2}}\}\) one has \(\|R_0\|_{C^{r-2}} \leq a_5e^\kappa\), provided \(P \in \mathcal{B}_1\).

According to the estimates (3.8) and (3.11), one has
\[
\|\tilde{G}_e - \tilde{G}_0\|_{C^{r-2}} = \|\tilde{R}_h + \tilde{R}_e\|_{C^{r-2}} \leq (3a_2 + a_4)e^\kappa
\]
when \((x, \vartheta, y, I)\) is restricted in \(\mathbb{T}^2 \times \{|y| \leq 2K^{-1}e^{\kappa - \frac{1}{2}}, |I| \leq 2K^{-1}|\omega_3|e^{\kappa - 1}\}\) where \(\vartheta = \sqrt{\epsilon^{-1}}\omega_3\theta\). It follows from the relations \(\partial_I\tilde{G}_e = 1 + O(e^\kappa), \partial_I\tilde{G}_0 = 1 + O(e^\kappa)\) and the theorem of implicit function that some \(a_6 = a_6(h, \delta) > 0\) exists such that
\[
\|G_e - G_0\|_{C^{r-2}} \leq a_6e^\kappa
\]
holds when \(|y| \leq 2K^{-1}e^{\kappa - \frac{1}{2}}\). Indeed, let \(z = (x, y, \vartheta)\), we get from the equations \(\tilde{G}_0(z, -G_0) = 0\) and \(\tilde{G}_1(z, -G_0) = 0\)
\[
(3.12)
\]
where \(\lambda = \lambda(z, G_e - G_0) \in [0, 1]\). It follows from the relation \(\partial_I\tilde{G}_0 = 1 + O(e^\kappa)\) that \(\max_z |G_e - G_0| \leq 2(3a_2 + a_4)e^\kappa\). For \(\xi = 0, \epsilon\), the derivative of \(G_\xi\) in \(z\) satisfies the equation
\[
\partial_z\tilde{G}_\xi(z, -G_\xi(z)) - \partial_I\tilde{G}_\xi(z, -G_\xi(z))\partial_zG_\xi(z) = 0.
\]
Because \(\partial_z^2\tilde{G}_0 = \omega_3^{-2}B''\epsilon, \partial_z\tilde{G}_e = |\partial_I\tilde{G}_e|^{-1}\partial_z\tilde{G}_e = O(1)\), by taking the difference of these equations we obtain the estimate on first derivative of \(G_e - G_0\)
\[
\|\tilde{G}_e(z, G_e - G_0)\| \leq |\tilde{G}_e(z, -G_0)|\left(|B''\omega_3^{-2}\epsilon(G_\xi - G_0)\partial_zG_e| + \left|\frac{d}{dz}(\tilde{R}_h + \tilde{R}_e)(z, -G_\xi)\right|\right)
\]
\[
\leq 2(3a_2 + a_4)e^\kappa\left(|2\omega_3^{-2}B''\epsilon + 1|\partial_zG_e| + 2|B'|\omega_3^{-1}|\sqrt{\epsilon + 1}\right).
\]
The estimate on higher order derivatives can be done similarly.

From the formula (3.2), we see that \(|G_e(x, y, \theta)| \leq a_4K^{-2}e^{2\kappa - 1}\) if \(\epsilon > 0\) is suitably small such that \(|V| + |R_e| < \frac{1}{2}a_4K^{-2}e^{2\kappa - 1}\). Consequently, the energy level set \(\tilde{G}_e^{-1}(0)\) intersects the domain \(|y| \leq K^{-1}e^{\kappa - \frac{1}{2}}, |I| \leq K^{-1}|\omega_3|e^{\kappa - 1}\) \(\times \mathbb{T}^3\) at the place where \(|I| \leq a_4e^{2\kappa - 1}\).

Under the composition of \(\Phi_{\epsilon,F}\), (3.1), (3.6) and (3.9), \(\tilde{\Sigma}_e \times \mathbb{T}^3\) is mapped onto \(\Omega_e\). If \(\Phi_{\epsilon,F}\) is an identity map, each section of \(\Omega_e\) where \((x, \theta)\) keeps constant contains the disk \(|y| \leq K^{-1}e^{\kappa - \frac{1}{2}}, I = 0\), because \(|k''_e| \geq |k''_0|\) is assumed in (3.1). Since \(\Phi_{\epsilon,F}\) approaches to identity and \(\frac{a_4e^{2\kappa - 1}}{K^{-1}|\omega_3|e^{\kappa - 1}} = a_4K\omega_3\epsilon^\kappa \to 0\) as \(\epsilon \to 0\), some \(\epsilon_0(\delta') > 0\) exists such that, for any \(\epsilon \in (0, \epsilon_0]\), the set \(|y| \leq (1 - \delta')K^{-1}e^{\kappa - \frac{1}{2}}, |I| \leq a_4e^{2\kappa - 1}\) is contained in each section of \(\Omega_e\) where \((x, \theta)\) is fixed, i.e. \(\Sigma_e \supset \{|y| \leq (1 - \delta')K^{-1}e^{\kappa - \frac{1}{2}}\}\). Similarly, we can show that \(\Sigma_e \subset \{|y| \leq (1 + \delta')K^{-1}e^{\kappa - \frac{1}{2}}\}\).

The transformation \(\mathcal{M}\) of (3.1) is uniformly bounded for all resonant conditions along \(\Gamma_{\kappa'}\) and all constants \(a_7\) with \(2 \leq \ell \leq 7\) 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 \(|k''_0| \geq |k''_1|\).

If \(|k''_0| \leq |k''_1|\), in the new coordinates \((u, v)\), the frequency at the double resonance has the form of \(M''\partial h(p'') = (0, \omega_3, 0)\). From (3.1) one obtains
\[
\langle k', q \rangle = \langle k', (MM_0)^{-1}u \rangle = u_1, \quad \langle k'', q \rangle = \langle k'', (MM_0)^{-1}u \rangle = k''_3u_3.
\]
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. \( \square \)

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 \( G_\epsilon \) 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 \( \tilde{M}_H(c) \) 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

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

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 \( M_0 \) 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 \( T_i \leq K^*\epsilon^{-\frac{1}{3} (1 - 3\kappa)} \) where \( \kappa \in (0, \frac{1}{6}) \) and \( K^* \) is independent of \( \epsilon \);
2. \( \Gamma_{k'} \) is covered by the union of the disks \( \{ \| p - p''_i \| \leq T_i^{-1} \epsilon^\kappa \} \).

**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

\[
\|x\|_Z = \min\{\{x\}, 1 - \{x\}\} = \text{dist}(x, \mathbb{Z}).
\]

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

\[
\| k\omega \|_Z \leq K^{-1}.
\]

For \( \omega = (\omega_2, \omega_3) \in \mathbb{R}^2 \), let \( |\omega| = \max\{\|\omega_2\|, |\omega_3|\} \). Let \( S^1 = \{ \omega \in \mathbb{R}^2 : |\omega| = 1 \} \) be the boundary of unit square, it has four sides. By applying Dirichlet’s approximation theorem, for any \( \omega \in S^1 \) and any integer \( K > 0 \) there exists \( 1 \leq k < K \) such that
\[ \|k \min\{\omega_2, \omega_3\}\|_2 \leq K^{-1}. \] In other words, given any \( \omega \in S^1 \), some rational vector \( \omega^* \) on the same side exists such that \( T\omega^* \in \mathbb{Z}^2 \) with \( T \leq K \) and

\[ \text{dist}(T\omega, T\omega^*) = \|T\omega - T\omega^*\| \leq K^{-1}. \tag{4.1} \]

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 \( M_0 \) 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 \( \lambda_\omega(\omega_2, \omega_3) \in S^1 \). 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

\[ \|\omega - \omega^*\| \leq d \|\lambda_\omega \omega - \lambda_\omega^* \omega^*\|. \]

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

The map \( \partial h \) 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

\[ \|\omega - \omega^*\| \leq d \|\lambda_\omega \omega - \lambda_\omega \omega^*\| \leq \frac{d\lambda_{\omega^*}}{KT}, \]

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

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

\[ m\|v\|^2 \leq \langle \partial^2 h(y)v, v \rangle \leq m'\|v\|^2, \quad \forall \ y \in B_D, \ v \in \mathbb{R}^3. \]

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 \( \|y - y^*\| \leq m^{-1}\|\omega - \omega^*\| \). Because \( \lambda_\omega^{-1} T \leq K \), the covering property \( \|y - y^*\| \leq T^{-1} \epsilon \) is guaranteed if we choose

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

Again, because \( \lambda_\omega^{-1} T \leq K \), one has \( T \leq K^* \epsilon^{-\frac{1}{3}(1-3\kappa)} \) if \( K = \frac{d\lambda_{\omega^*}}{m} \epsilon^{-\kappa} \leq \frac{K^*}{\Lambda} \epsilon^{-\frac{1}{3}(1-3\kappa)} \).

It holds for each \( \epsilon \in (0, \epsilon_0] \) if \( \epsilon_0 \) satisfies the condition

\[ \frac{1-3\kappa}{\epsilon_0} \leq \frac{mK^*}{d\Lambda^2}, \]

where \( \Lambda = \max_{\omega \in \Gamma_{\omega, k'}} \lambda_\omega \). For \( \kappa < \frac{1}{6} \), such \( \epsilon_0 > 0 \) exists. \( \square \)

5. The finiteness of strong double resonance. Because of Theorem 4.1, the path of resonance \( \Gamma_{k'} \) is covered by the discs \( \{ \|p - p_i''\| < T_i^{-1} \epsilon \} \), where \( \kappa < \frac{1}{6} \), \( T_i \leq K^* \epsilon^{-\frac{1}{3}(1-3\kappa)} \) is the period of the double resonance at \( p_i'' \), \( K^* \) is independent of \( \epsilon \). 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 $\epsilon$ in generic case.

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

$$Z(p, q) = Z_{k'}(p, (k', q)) + Z_{k', k''}(p, (k', q), (k'', q))$$

(5.1)

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

$$Z_{k'} = \sum_{j \in \mathbb{Z} \setminus \{0\}} P_{jk'}(p) e^{j(k', q)i},$$

$$Z_{k', k''} = \sum_{(j,l) \in \mathbb{Z}^2 \setminus \{0\}} P_{jk'+lk''}(p) e^{j(k', q)+l(k'', q)i}. (5.2)$$

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

Treated as a function of $x = (k', q)$, we consider $Z_{k'}(p, x)$ as a family of functions $Z_{k'}(p, \cdot) : \mathbb{T} \to \mathbb{R}$, 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 $F_\lambda$ is Lipschitz in the parameter $\lambda \in [0, 1]$. Then, there exists an open-dense set $\mathcal{V} \subset C^r(\mathbb{T}, \mathbb{R})$ ($r \geq 4$) such that for each $V \in \mathcal{V}$ 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

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

where $Z_{k'}$ is defined in (5.2) consisting of Fourier modes of $P$ in span$_\mathbb{Z}\{k'\}$, and $P' = P - Z_{k'} \in C^r(B_D \times \mathbb{T} \setminus \{0\}, \mathbb{R})/C^r(B_D \times \mathbb{T}, \mathbb{R})$.

Therefore, there exists an open-dense set $\mathcal{V} \subset C^r(B_D \times \mathbb{T}, \mathbb{R})$ and consequently an open-dense set $\mathcal{W} = \mathcal{V} \oplus C^r(B_D \times \mathbb{T}^3, \mathbb{R})/C^r(B_D \times \mathbb{T}, \mathbb{R})$ such that for all $P \in \mathcal{W}$, 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 $B y_0 = (0, 1)^T$. A non-degenerate maximal point $x_{1,i}$ of $Z_{k'}(p'', \cdot)$ corresponds to a normally hyperbolic invariant cylinder (NHIC)

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

of the Hamiltonian system $\frac{1}{2} \langle By, y \rangle + Z_{k'}(p'', x_1)$. In the $(p, q)$-coordinates, the cylinder passes through a neighborhood of the double resonant point $p''$. 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 \( \frac{1}{2} \langle By, y \rangle + Z_{k'} + Z_{k',k''} \), there exists some weakly invariant cylinder \( \Pi_i' \) lying in a small neighborhood of \( \Pi_i \cap \{ ||y|| \leq K_i^{-1}\epsilon^{\kappa-\frac{3}{2}} - 1 \} \) provided \( |k''_i| \) is large enough.

Indeed, if we introduce a cut-off \( C^\infty \)-function \( \chi: [0, \infty) \to \mathbb{R} \) satisfying \( \chi(\nu) = 0 \) for \( \nu \geq K_i^{-1}\epsilon^{\kappa-\frac{3}{2}} - 1 \) and \( \chi(\nu) = 1 \) for \( \chi \leq K_i^{-1}\epsilon^{\kappa-\frac{3}{2}} - 1 \). Then by applying the theorem of NHIM to the Hamiltonian \( \frac{1}{2} \langle By, y \rangle + Z_{k'}(p''_i, x_1) + \chi(||y||)Z_{k',k''}(p''_i, x) \) we see that the cylinder \( \Pi_i \) survives the small perturbation \( \chi(||y||)Z_{k',k''}(p''_i, x) \) if \( |k''_i| \) is sufficiently large. Restricted on the region \( \{ ||y|| \leq K_i^{-1}\epsilon^{\kappa-\frac{3}{2}} - 1 \} \) 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 \( Z_{k',k''} \) is so small that can be treated as a small perturbation

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

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 \( \{ ||p - p''_i|| < T_i^{-1}\epsilon \} \) depends on \( \epsilon \), 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 \( \epsilon \).

**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 \( \mathfrak{P}' \subset C^\infty(B_D \times \mathbb{T}^3, \mathbb{R}) \), for each \( P \in \mathfrak{P}' \) there exists some \( d_P > 0 \) such that the second derivative \( \partial^2_{x'}Z_{k'}(p, x_p) < -d_P \) holds for all \( p \in \Gamma_{k'} \). Since \( \lambda P \in \mathfrak{P}' \) for any \( \lambda \neq 0 \) if \( P \in \mathfrak{P}' \), the restriction of \( \mathfrak{P}' \) to \( \mathfrak{S}_1 \) is clearly open-dense. Since the double resonant term \( Z_{k',k''} \) will be sufficiently small provided \( ||k''_i|| \) is sufficiently large, we complete the proof. \( \square \)

**Notation 5.1.** Let \( \Lambda_\epsilon \) denote the set of the subscripts \( i \) such that the resonant circles \( \Gamma_{k'} \cup \Gamma_{k''} \) is covered by the disks

\[
\Gamma_{k'} \cup \Gamma_{k''} \subset \bigcup_{i \in \Lambda_\epsilon} \{ p \in \mathbb{R}^3 : ||p - p''_i|| < T_i^{-1}\epsilon \}
\]

where the period \( T_i \) of the double resonant point \( p''_i \) is not larger than \( K_i^{\kappa}\epsilon^{-(1-3\kappa)/3} \) 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 \( k''_i = M_0^{-1}k''_i = (0, k''_i^2, k''_i^3) \). 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)

\[
G_{\epsilon,i}(x, y, \vartheta) = \tilde{G}_i(x, y) + R_{\epsilon,i}(x, y, \vartheta(\vartheta)), \quad (6.1)
\]
where $G_i = \frac{1}{2} \langle B_i y, y \rangle - V_i(x_1, |k''_i|x_2)$, $V_i \in C^r$ is 2π-periodic in $(x_1, |k''_i|x_2)$ such that \( \max V_i = 0 \), \( \theta = \frac{\omega_i}{\sqrt{\epsilon}} \), $R_{\epsilon,i} \in C^{r-2}(\mathbb{T}^2 \times \Sigma_{\epsilon,i} \times |k''_i|\mathbb{T}, \mathbb{R})$ with

\[
\{|y| \leq (1 - \delta')(\eta T_i)^{-1} \epsilon^{\frac{\epsilon}{2}} \} \subset \Sigma'_{\epsilon,i} \subset \{|y| \leq (1 + \delta')(\eta T_i)^{-1} \epsilon^{\frac{\epsilon}{2}} \},
\]

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

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

\[
G_{\epsilon,i}(x_1, x_2 + \tilde{k''_i}/\tilde{k''_{i,2}}, \theta + 1) = G_{\epsilon,i}(x_1, x_2, \theta)
\]

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

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

\[
c \in \mathcal{L}_\alpha(\omega) \iff \alpha(c) + \beta(\omega) = \langle c, \omega \rangle.
\]

By the definition, the $\beta$-function is the Fenchel-Legendre dual of the $\alpha$-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 $\mathfrak{N}_i \subset C^r(\mathbb{T}^2, \mathbb{R})$ with $r \geq 2$. Each $V_i \in \mathfrak{N}_i$ is associated with some positive numbers $\Delta V_i, \epsilon_i$ such that for any $\xi \in (0, \Delta V_i)$, $\epsilon \in [0, \epsilon_i]$ the circle $\alpha_{\epsilon,i}^{-1}(\xi)$ establishes a transition chain (of cohomology equivalence). These circles make up an annulus $\mathfrak{A}_i$ surrounding the set $\mathfrak{F}_i = \{c : \alpha_{\epsilon,i}(c) = 0\}$.

**Proof.** If we are satisfied with the property that $V \in \mathfrak{N}_i$ 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 $\mathfrak{N}_i$ is residual in the space of $C^r$-functions which are 2π-periodic in $(x_1, |k''_i|x_2)$.

For the purpose, we introduce a canonical transformation $\mathcal{T}: (x_1, |k''_i|x_2) \to (\phi_1, \phi_2)$, $(y_1, |k''_i|^{-1}y_2) \to (I_1, I_2)$ and get the Hamiltonian from (6.1)

\[
\mathcal{G}'_i = \mathcal{T}_* \mathcal{G}'_i = \frac{1}{2} \langle B'_i I, I \rangle - V_i(\phi), \quad (\phi, I) \in \mathbb{T}^2 \times \mathbb{R}^2,
\]

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

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

Given an irreducible class $g \in H_1(\mathbb{T}^2, \mathbb{Z}) \setminus \{0\}$, let $\mathfrak{C}_{i,g} = \cup_{\nu > 0} \mathcal{L}_{\alpha_{\epsilon,i}}(\nu g)$. To consider the Aubry set of $G_{\epsilon,i}$ for $c \in \mathfrak{C}_{i,g}$, we consider the truncated Hamiltonian $\mathfrak{G}'_i$ of (6.3) and let $\alpha_i$ be the $\alpha$-function for $\mathfrak{G}'_i$. For $c \in \mathcal{L}_{\alpha_i}(\lambda g)$, each $c$-minimal orbit is periodic. To stress its topological information, we also call it $\lambda g$-minimal orbit. The parameter $\lambda_i$ is called bifurcation point, if there are two or more $\lambda_i$-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 $g' \in H_1(\mathbb{T}^2, \mathbb{Z}) \setminus \{0\}$ and $\lambda_0 > 0$, there exists an open-dense set $\mathfrak{N}_i \subset C^r(\mathbb{T}^2, \mathbb{R})$ with $r \geq 5$ such that for each $V_i \in \mathfrak{N}_i$,
it holds simultaneously for all \( \lambda \in [\lambda_0, \infty) \) that the Mather set of \( G'_i \) for each \( c \in \mathcal{L}_{\bar{\alpha}}(\lambda g') \) consists of hyperbolic periodic orbits. Indeed, except for finitely many \( \{\lambda_i\} \) 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 \( \epsilon \). At each bifurcation point \( \lambda_g \), there exist exactly two \( \lambda'_g \)-minimal orbits of \( G'_i \). So, for \( \lambda'_i \geq \lambda_g \), all \( \lambda'_g \)-minimal orbits make up \( m + 1 \) pieces of NHIC. Let \( \tilde{\Pi}_{i,\ell} \) denote the cylinder made up of \( \lambda'_g \)-minimal orbits of \( \bar{G}_i \) for all \( \lambda' \in [\lambda'_g, \lambda'_{\ell + 1}] \) for \( \ell < m \), let \( \tilde{\Pi}_{i,m} \) be the cylinder made up of \( \lambda'_g \)-minimal orbits of \( \bar{G}_i \) for all \( \lambda' \in [\lambda'_m, \lambda'_g] \) where \( \lambda'_i \) is chosen such that max\( \rho((I_1, k''_i | I_2)(\theta)) = (1 - \delta')(\eta T_i)^{-1}e^{k - \frac{1}{2}} \) holds for the \( \lambda'_g \)-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 \( \mathcal{F}^* \bar{G}'_i \) in the way
\[
(x_1, x_2, y_1, y_2) = \left( \phi_1, \frac{1}{|k''_i|}(\phi_2 + 2\ell\pi), I_1, |k''_i| I_2 \right)
\]
for \( \ell = 0, 1, \cdots, |k''_i| - 1 \). Consequently, each cylinder \( \tilde{\Pi}_{i, \ell} \) of \( \bar{G}_i \), modulo a shift \((x, y) \rightarrow (x + (0, 2\pi/|k''_i|), y)\). Each cylinder is made up of \( \lambda g \)-minimal orbits of \( G_i \) with \( \lambda g = \frac{\lambda'_g}{|k''_i|}(k''_i g_1, g_2) \) if \( \lambda' g' = \lambda'(g_1, g_2) \).

Let \( E_\ell \) 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 \( \tilde{\Pi}_{i, \ell} \) can be extended by the hyperbolic orbits lying in the level set \( \bar{G}_i^{-1}(E) \) with \( E \in [E_{\ell + 1}, E_{\ell + 1} + 2d] \cup [E_{\ell} - 2d, E_{\ell} + 1] \). 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 \( \Phi_{G_{\epsilon,i}}^\theta \). Notice that \( \theta = \omega_{i,3} \sqrt{\epsilon^{-1}} \theta \) and \( G_{\epsilon,i} \) is 2\pi in \( x \), 2\( |k''_i| \pi \) in \( \theta \) and symmetric for \( M_i \), see (6.2), we introduce a shift
\[
\sigma_i : (x, y, \theta) \rightarrow (x + \left(0, 2\pi \frac{k''_i}{k''_{i,2}}, y, \theta + \frac{\sqrt{\epsilon}}{\omega_{i,3}} \right), (6.4)
\]
then \( \sigma_i^* G_{\epsilon,i} = G_{\epsilon,i} \). Let \( \tilde{\Pi}_{i,E_{\ell - d},E_{\ell + 1} + d} = (\tilde{\Pi}_{i,\ell} \times \frac{|k''_i|}{\omega_{i,3}} \mathbb{T}) \cap \{ G_i \in [E_{\ell} - d, E_{\ell + 1} + d] \} \).

**Proposition 6.2.** For sufficiently small \( \epsilon \), there is a cylinder \( \tilde{\Pi}_{i,E_{\ell - d},E_{\ell + 1} + d} \) modulo the shift \( \sigma_i \), which is weakly invariant and normally hyperbolic for the flow \( \Phi_{G_{\epsilon,i}}^\theta \) 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 \( \rho \) be a \( C^2 \)-function such that \( \rho(\mu) = 1 \) for \( \mu \geq 1 \) and \( \rho(\mu) = 0 \) for \( \mu \leq 0 \). 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
\[
G'_{\epsilon,i} = \begin{cases} 
\bar{G}_i + \rho_1 R_{\epsilon,i}, & \text{if } \bar{G}_i(x, y) \in [E_{\ell} - 2d, E_{\ell} - d], \\
\bar{G}_i + \rho_2 R_{\epsilon,i}, & \text{if } \bar{G}_i(x, y) \in [E_{\ell} + 1 + d, E_{\ell} + 1 + 2d], \\
G_{\epsilon,i}, & \text{elsewhere}. 
\end{cases} 
\]
(6.5)

Because \( \|G'_{\epsilon,i} - \bar{G}_i\|_{C2} \ll 1 \) if \( |y| \leq O(\epsilon^{-\frac{1}{2}}) \) and \( \epsilon \ll 1 \), it follows that the NHIC \( \tilde{\Pi}_{i,\ell} \) survives the perturbation \( \bar{G}_i \rightarrow G'_{\epsilon,i} \), denoted by \( \tilde{\Pi}_{i,\ell} \).
Let $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d} = \tilde{\Pi}^{\epsilon}_{i}\cap \{G_i \in [E_\ell - d, E_{\ell+1} + d]\}$, which is a weakly invariant for $\Phi^{\epsilon}_{G_{\epsilon,i}}$ because $G_{\epsilon,i} = G_{\epsilon,i}'$ when they are restricted in the region $\{(x,y) : \tilde{G}_i(x,y) \in [E_\ell - d, E_{\ell+1} + d]\} \times \frac{[k'_\ell, \sqrt{\epsilon}]}{G_i, \tilde{G}_i}$. The normal hyperbolicity is obvious.

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

Notice that $\lambda_\epsilon > 0$ is set such that $\max_{\theta} |y(\theta)| = (1 - \delta')(\eta T)_{\epsilon}^{-1} e^{\kappa - \frac{\epsilon}{2}}$ holds for the $\lambda, g$-minimal orbits $(y(\theta), x(\theta))$. By applying Theorem 1.2 of [C17b], we have

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

1) there are finitely many NHWCs for the flow $G_{\epsilon,i}$: $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ ($\ell = 0, \cdots, m$) and $\tilde{\Pi}^{\epsilon}_{i,E_{\ell-d},E_{\ell+1}+d}$ modulo the shift $\sigma_i$, 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 $\epsilon$, $E_\ell = \tilde{\alpha}_i(\lambda_\epsilon, g) = O(\epsilon^{2^* - 1})$ 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 $\alpha_{\epsilon,i}(c) \in (E_{\ell}, E_{\ell+1}, E_{\ell+1}, E_{\ell+1}+d)$, the Aubry set lies on $\tilde{\Pi}^{\epsilon}_{i,E_\ell-d,E_{\ell+1}+d}$ modulo $\sigma_i$;

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

- $(3)$ if $\alpha_{G_{\epsilon,i}}(c) \in (E_{m+1}, E_\ell)$, the Aubry set lies on $\tilde{\Pi}^{\epsilon}_{E_{m+1}, E_{\ell}}$ modulo $\sigma_i$.

**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 push them back to the space of $(p,q)$.

At a strong double resonant point $p''_i$, $i \in \Lambda_\epsilon$, we consider two classes $g^-, g^+$ 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 $\Gamma_k \cup \Gamma_k^*$:

- $(1)$ the point $p''_i$ is on the path $\Gamma_k$, where $\Gamma_k$ does not intersect with $\Gamma_k^*$;

- $(2)$ the point $p''_i$ is on the path $\Gamma_k^*$, where $\Gamma_k^*$ does not intersect with $\Gamma_k$;

- $(3)$ $p''_i \in \Gamma_k \cap \Gamma_k^*$.

After the linear coordinate transformations $\mathcal{M}_0$ and $\mathcal{M}_i$ 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 $\Gamma_k$, $k = k^*, k^*$ in the first two cases. In the third case, $\partial h(p)$ is proportional to $(0,1)$ along $\Gamma_k^*$ and to $(1,0)$ along $\Gamma_k$.

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 $g_i^+, g_i^- \in H_1(T^2, \mathbb{Z})$ 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 $\mathcal{U}_i \subset C^r(\mathbb{T}^2, \mathbb{R})$ ($r \geq 5$) such that for each $V_i \in \mathcal{U}_i$, there exist $0 < E_{i,0} < \Delta_{V_i}$, $\Delta_i' > 0$ and $\epsilon_i > 0$ such that for all $\epsilon \in (0, \epsilon_i]$ the following holds.

1) there is an annulus $\mathcal{A}_i$ around a double resonant point $p_i''$, made up of the circles 
\{ $c \in \alpha_{\epsilon,i}^{-1}(E) : E \in (0, \Delta_{V_i})$ \}, where $\lambda_i$ is the following holds.

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

$$C_i^\pm = \cup_{\lambda \in [\lambda_i^\pm, \lambda_i^\|]} \mathcal{L}_{\alpha_{\epsilon,i}}(\lambda g_i^\pm), \quad g_i^\pm \in H_1(\mathbb{T}^2, \mathbb{Z}).$$

where $\lambda_i^\pm$, $\lambda_i^\|$ satisfy $\alpha_{\epsilon,i}(\mathcal{L}_{\alpha_{\epsilon,i}}(\lambda_i^\pm g_i^\pm)) = E_{i,0} < \Delta_{V_i}$ and

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

$$(1 - \delta')(\eta T_i)^{-3\epsilon - \frac{1}{2}} - \Delta_i' \leq |y(\theta)| \leq (1 - \delta')(\eta T_i)^{-\epsilon} - \Delta_i';$$

(2) for each $c \in C_i^\pm$ except for finitely many classes \{ $c_j$ \}, the Aubry set $\tilde{A}(c)$ lies in certain cylinder $\Pi'_1, E_{i-1}, E_{i+1}$ modulo the shift $\sigma_i$, while $\tilde{A}(c_j)$ lies in two of the cylinders. For each orbit \{(x(\theta), y(\theta)), \theta \in \mathbb{R}\} in the Aubry set one has

$$\text{Osc}|y(\theta)| = \max \{ |y(\theta) - y(\theta')| \} \leq \Delta_i'. $$

\textbf{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 $G_i$ first. We claim that some constant $\Delta_i' > 0$ exists such that it holds along any $\lambda g$-minimal orbit $(\tilde{y}(\theta), \tilde{x}(\theta))$ of $G_i$ that

$$|\tilde{y}(\theta) - \tilde{y}(\theta')| \leq \frac{1}{2} \Delta_i'. \quad (6.6)$$

Indeed, each $\lambda g$-minimal orbit is entirely contained in certain energy level set $\tilde{G}_i^{-1}(E)$. The higher the energy increases, the shorter the period becomes. Let $E_1$ be the energy that the set $\tilde{G}_i^{-1}(E_1)$ contains $\lambda g$-minimal orbit with period 1. One obtains from the equation $\dot{y} = \partial V_i(x)$ that $|\tilde{y}(\theta) - \tilde{y}(\theta')| \leq \max_{x \in \mathbb{T}^2} |\partial_x V_i(\tilde{x}(\theta))|, |\tilde{y}'(\theta)| |\partial_{xy} V_i(\tilde{x}(\theta))|$. For $(x(\theta), y(\theta)) \in \tilde{G}_i^{-1}(E)$ 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 $(x, y) \in \tilde{G}_i(E)$. Therefore, the estimate (6.6) holds if we set

$$\Delta_i' = \max \left\{ 2 \max_{x \in \mathbb{T}^2} |\partial_x V_i(\tilde{x}(\theta))|, |\tilde{y}'(\theta)| |\partial_{xy} V_i(\tilde{x}(\theta))| \right\}, \frac{4}{m_i} \sqrt{E_1 + \max_x V_i(x)}. \quad (6.7)$$

Next, we consider $G_{\epsilon,i}$ which is a $O(\epsilon^\kappa)$-perturbation of $G_i$. Each $\theta$-section of NHIC of $G_{\epsilon,i}$ is located in $O(\epsilon^\kappa)$-neighborhood of NHIC of $G_i$. By Proposition 5.2 of [C17a], the Aubry set $\tilde{A}(c)$ does not hit the level set $G_{\epsilon_i}^{-1}(E + \epsilon^\frac{1}{2}\kappa)$ if $c \in \alpha_{\epsilon_i}^{-1}(E) \cap \cup_{\lambda} \mathcal{L}_{\alpha_{\epsilon,i}}(\lambda g)$ and $E - \min \alpha_{\epsilon,i} \geq O(1)$. Because the Aubry set $\tilde{A}(c)$ for $G_{\epsilon,i}$ stays in the cylinder, it falls into $O(\epsilon^\frac{1}{2}\kappa)$-neighborhood of the Aubry set for $G_i$. For any $y(\theta)$ there exists $\tilde{y}(\theta')$ such that $|y(\theta) - \tilde{y}(\theta')| < \frac{1}{2} \Delta_i'$ provided $\epsilon > 0$ is suitably small and it holds for any orbit $(x(\theta), y(\theta)) \in \tilde{A}(c)$ with $c \in C_i^\pm$ that

$$\text{Osc}|y(\theta)| = \max_{\theta, \theta'} |y(\theta) - y(\theta')| \leq \max_{\theta, \theta'} |\tilde{y}(\theta) - \tilde{y}(\theta')| + 2O(\epsilon^\frac{1}{2}\kappa) \leq \frac{3}{4} \Delta_i'. $$
If we set $\tilde{\lambda}^\pm_i$ such that \( \max_\theta |\tilde{y}(\theta)| = (1 - \delta')(\eta T_i)^{-\frac{1}{2}} - \frac{1}{2} \Delta_i' \) holds along $\tilde{\lambda}^\pm_i g_i^\pm$-minimal orbit of $\tilde{G}$, one obtains that

\[
(1 - \delta')(\eta T_i)^{-\frac{1}{2}} - \frac{1}{2} \Delta_i' \leq |y(\theta)| \leq (1 - \delta')(\eta T_i)^{-\frac{1}{2}}
\]

if $(x(\theta), y(\theta))$ lies in $\tilde{A}(c)$ with $c \in \mathcal{L}_{\alpha,i}(\tilde{\lambda}^\pm_i g_i^\pm)$.

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

\[
\tilde{G}_i = \frac{1}{2} \langle B_i y, y \rangle - V'_i(x_1) - V''_i(x_1, |\tilde{k}''_i|x_2),
\]

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

**Theorem 6.4.** There exists an open-dense set $\mathfrak{U}_k \subset C^r(B_d \times \mathbb{T}, \mathbb{R})$ $(r \geq 5)$, for each $Z_{k'} \in \mathfrak{U}_k$, 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 $i \in \Lambda_\epsilon \setminus \Lambda_a$ there is a channel

\[
\mathcal{C}_i^w = \bigcup_{\lambda \in [-\bar{\lambda}_i, \bar{\lambda}_i]} \mathcal{L}_{\alpha,i}(\lambda g) \subset H^1(\mathbb{T}^2, \mathbb{R}), \quad g = (0, 1)
\]

with the properties that

1. for each $c \in \mathcal{C}_i^w$, the Aubry set $\tilde{A}(c)$ lies on some NHWIC entirely contained in the region $\Sigma_{\epsilon,i}$, modulo the shift $\sigma_1$. Along each orbit \(\{ (x(\theta), y(\theta)), \theta \in \mathbb{R} \}\) in the Aubry set one has

\[
\text{Osc} |y(\theta)| = \max_{\theta, \theta'} |y(\theta) - y(\theta')| \leq \Delta'_i,
\]

2. the numbers $\lambda_i$ and $\bar{\lambda}_i > 0$ are chosen such that for $c \in \mathcal{L}_{\alpha,i}(\tilde{\lambda}_i g) \cup \mathcal{L}_{\alpha,i}(-\bar{\lambda}_i g)$ it holds for $c$-minimal orbit \(\{(x(\theta), y(\theta)), \theta \in \mathbb{R}\}\) that

\[
(1 - \delta')(\eta T_i)^{-\frac{1}{2}} - \frac{1}{2} \Delta'_i \leq |y(\theta)| \leq (1 - \delta')(\eta T_i)^{-\frac{1}{2}}.
\]

**Proof.** According to Theorem 5.1, there exists an open-dense set $\mathfrak{U}_k \subset C^r(B_d \times \mathbb{T}, \mathbb{R})$, for each $Z_{k'} \in \mathfrak{U}_k$, 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 $\frac{1}{2} \langle B y, y \rangle - V'_i(x_1)$ admits a normally hyperbolic invariant cylinder made up of the minimal periodic orbits of type $(0, 1)$. As the second resonant term $V''_i(x_1, |\tilde{k}''_i|x_2)$ and the remainder $R_{e,i}$ are small, the weakly invariant cylinder survives the perturbation $\frac{1}{2} \langle B y, y \rangle - V'_i \to \frac{1}{2} \langle B y, y \rangle - V'_i - V''_i + R_{e,i}$.

To study the oscillation of $y(t)$, we also consider the truncation $G_i = \frac{1}{2} \langle B y, y \rangle - V'_i(x_1) - V''_i(x_1, |\tilde{k}''_i|x_2)$ first. Let $(\tilde{x}(\theta), \tilde{y}(\theta))$ be the $\lambda g$-minimal orbit, then the second component of $\tilde{y}(\theta)$ satisfies the following relation

\[
\tilde{y}_2(\theta') - \tilde{y}_2(\theta) = \int_\theta^{\theta'} |\tilde{k}''_i| \partial_2 V''_i(x_1(\theta), |\tilde{k}''_i|x_2(\theta))d\theta.
\]

Although the number $|\tilde{k}''_i|$ approaches infinity if the second resonant condition becomes weaker, it does not make trouble to control the oscillation of $\tilde{y}(\theta)$. Indeed, the term $V''_i(x_1, |\tilde{k}''_i|x_2) = Z_{k''_i}(p_i, x_1, |\tilde{k}''_i|x_2)$, which is defined in (5.2). Because $|P_k|$ decrease fast as $|k|$ increases: $|P_k| \leq O(|k|^{-1})$, one has $|\tilde{k}''_i| |\partial_2 V''_i| \to 0$ as $|\tilde{k}''_i| \to \infty$.

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 $\epsilon$-minimal orbits is clear. Each $\epsilon$-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])

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

(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 $\alpha$-function for $H$ and $h$, $\tilde{\beta}_H$ and $\tilde{\beta}_h$ denote the $\beta$-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 $\mathcal{L}_{\tilde{\alpha}_H}(\omega)$ falls into $C_s\sqrt{\epsilon}$-neighborhood of $p$ with $C_s \leq 2\sqrt{\|P\|/m}$, namely

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

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

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

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 $h(\partial h^{-1}(\omega)) = \tilde{\alpha}_H(\mathcal{L}_{\tilde{\alpha}_H}(\nu\omega))$, then some constant $C_r > 0$ exists such that $|\nu - 1| \leq C_r\sqrt{\epsilon}$.

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

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

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

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

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 $\omega$, i.e. $\mathbb{F}_\omega = \mathcal{L}_{\hat{\alpha}_H}(\omega)$. The following lemma establishes the location of $\hat{C}_k$, it lies in $O(\sqrt{\epsilon})$-neighborhood of $\Gamma_k$. 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 $\mathbb{F}_\nu \omega \subset \hat{\alpha}_H^{-1}(E)$ 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 $\mathbb{F}_\nu \omega \subset B_{C_s \epsilon} \sqrt{\epsilon}(p)$. Since $\|p - p_\nu\| \leq |1 - \nu| \|\omega\| m^{-1}$, 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 $\Phi^t_1 H$ can not be $\hat{c}$-minimal if $\|p(t) - \hat{c}\| > D_H \sqrt{\epsilon}$ for all $t \in \mathbb{R}$.

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

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

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

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

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

$$\langle p(t) - \hat{c}, \dot{q}(t) \rangle = \langle p(t) - \hat{c}, \partial h(p) \rangle + \langle p(t) - \hat{c}, \epsilon \partial_P P \rangle$$

$$\geq \frac{m}{2} \|\hat{c} - p(t)\|^2 - 2\epsilon \|P\| - \epsilon \|\partial_P P\|$$

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 $\|p(t) - \hat{c}\| > D_H \sqrt{\epsilon} \forall \, t \in \mathbb{R}$, then $L(\hat{q}(t), q(t)) - \langle \hat{c}, \dot{q}(t) \rangle + \alpha_H(\hat{c}) > 0$ 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 $\hat{c}^* = p^*$ to a small neighborhood of $\hat{c} = p^*$.

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

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

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

then one has $\hat{c} = (c, -\lambda \alpha_c(c)) \in \hat{\alpha}_c^{-1}(E)$. 


It was proved in [C17b] (Theorem 3.3 there). Let $\tilde{x} = (x, \lambda^{-1} t)$, $\tilde{y} = (y, -\lambda G_\epsilon)$ and $\tau$ denote the time of $\tilde{G}_\epsilon$, the theorem follows from the identity
\[
\int \left( \left\langle \frac{dx}{dt} \cdot y - c \right\rangle - G_\epsilon + \alpha_\epsilon(c) \right) dt = \int \left( \left\langle \frac{d\tilde{x}}{d\tau} \cdot \tilde{y} - \tilde{c} \right\rangle - \tilde{G}_\epsilon + E \right) d\tau.
\]

For the application of the theorem in the paper, one has $n = 3$. If we regard the graph of $\alpha_{\epsilon,i}$ over $\tilde{A}_i \cup \tilde{C}^-_i \cup \tilde{C}^+_i$ as a set in $\mathbb{R}^3$,
\[
\{(c_i, -\alpha_{\epsilon,i}(c)) : c \in \tilde{A}_i \cup \tilde{C}^-_i \cup \tilde{C}^+_i \}
\]
it precisely lies in the surface $\tilde{\alpha}_{\epsilon,i}^{-1}(0)$. The graph of $\tilde{A}_i$, $\tilde{C}^-_i$ and $\tilde{C}^+_i$ are denoted by $\tilde{A}_i$, $\tilde{C}^-_i$ and $\tilde{C}^+_i$ 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
\[
\Psi_i^\epsilon : \tilde{c} \to M_i \tilde{c}. \quad (8.1)
\]

Let $\tilde{\alpha}_{\Phi^*_\epsilon F_i} H$ be the $\alpha$-function for the Hamiltonian $\Phi^*_\epsilon F_i H$, where $\epsilon F_i$ is the generating function for the KAM iteration. Because the rescaling (3.9) induces
\[
\int \tilde{p} d\tilde{x} - \Phi^*_\epsilon F_i H dt = \sqrt{\epsilon} \left( \int \tilde{y} d\tilde{x} - \tilde{G}_{\epsilon,i} ds \right),
\]
one obtains the rescaling of the first cohomology class, from $\tilde{\alpha}^{-1}_{\epsilon,i}(0)$ to $\tilde{\alpha}_{\Phi^*_\epsilon F_i}^{-1}(0)$
\[
\Psi_i^\epsilon : c - c_i \to \sqrt{\epsilon} (c - c_i), \quad c_3 - c_{i,3} \to \frac{\epsilon}{\omega_{3,i}} (c_3 - c_{i,3}), \quad (8.2)
\]
where $\tilde{c}_i = (c_1, c_i, 3) = (c_{i,1}, c_{i,2}, c_{i,3}) = p''_i$ if we treat both as the points in $\mathbb{R}^3$.

Because $\Phi^*_\epsilon F_i H$ is a Hamiltonian, it does not change the Mather set, Aubry set and Mañe 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 $\Gamma_k$, and $\Gamma_k^-$ constructed in Section 3. For generic perturbation $P$, due to Proposition 5.1, the number of strong double resonant points is independent of $\epsilon$. For each $i \in \Lambda_\alpha$, the composition of $\Psi_i^\epsilon$ and $\Psi_i'$ maps $\tilde{A}_i$ to an annulus of cohomology equivalence $\tilde{A}_i \subset \tilde{\alpha}_H^{-2}(E)$ where $\tilde{\alpha}_H$ denotes the $\alpha$-function of $H$. It also maps $\tilde{C}_i^\epsilon$ for all $i \in \Lambda_\alpha$ to a local channel $\tilde{C}_i \subset \tilde{\alpha}_H^{-2}(E)$ ($i = \pm, w$), namely
\[
\tilde{A}_i = \Psi_i^\epsilon \Psi_i'(A_i), \quad \tilde{C}_i = \Psi_i^\epsilon \Psi_i'(C_i), \quad i = \pm, w.
\]

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
\[
\tilde{C}_k = \{ \tilde{c} \in \tilde{\alpha}_H^{-1}(E) : \langle k, \omega \rangle = 0 \ \forall \omega \in \mathcal{L}^{\alpha_H}(\tilde{c}) \},
\]
Under generic perturbation $\epsilon P$, 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 $\omega \in \mathcal{L}^{\alpha_H}(\tilde{C}_k)$ is resonant, $\mathcal{F}_\omega = \mathcal{L}^{\alpha_H}(\omega)$ is a flat of dimension one or two.
The set $\tilde{\mathcal{C}}_{k^*} \cup \tilde{\mathcal{C}}_{k^*}$ obviously contains the flats $\mathcal{F}_i \ (i \in \Lambda_k)$ for strong double resonance we are concerned about

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

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

$$\tilde{\mathcal{C}} = (\tilde{\mathcal{C}}_{k^*} \cup \tilde{\mathcal{C}}_{k^*} \setminus \{(\cup_{i \in \Lambda_k} \mathcal{F}_i + d_i \sqrt{\epsilon}) \cup \left( \cup_{i \in \Lambda_k} \tilde{\mathcal{A}}_i \right)$$

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

For any $\delta > 0$, there exists $\epsilon_0 > 0$ such that $\forall \epsilon \leq \epsilon_0$ there exist totally irreducible $k^*, k^* \in \mathbb{Z}^3 \setminus \{0\}$ such that $\tilde{\mathcal{C}}_{k^*} \cup \tilde{\mathcal{C}}_{k^*}$ contains two points $\tilde{c}^*, \tilde{c}^* \in \tilde{\mathcal{C}}$ with $\|\tilde{c}^* - p^*\| < \frac{\delta}{2}$ and $\|\tilde{c}^* - p^*\| < \frac{\delta}{2}$. Let $\Gamma_c = \Gamma_c(\epsilon P)$ be a path lying inside $\tilde{\mathcal{C}}$, connecting $\tilde{c}^*$ to $\tilde{c}^*$. It keeps away from the boundary of $\tilde{\mathcal{C}}$ 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 $\tilde{c} \in \tilde{\mathcal{C}}$. We are going to show that $\Gamma_c$ 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 $\cup_{i \in \Lambda_k} \tilde{\Sigma}_{\epsilon,i}$ where

$$\tilde{\Sigma}_{\epsilon,i} = \{ p : |p - M_{0^{-1}} p''_i| < (\eta T_i)^{-1} \epsilon^\xi, \ \eta \in (0, 1)) \}.$$  

We assume the subscripts $\{ i \in \Lambda_k \}$ are well-ordered such that $\tilde{\Sigma}_{\epsilon,i}$ is adjacent to $\tilde{\Sigma}_{\epsilon,i+1}$. Restricted on the energy level set $H^{-1}(E)$ contained in $\tilde{\Sigma}_{\epsilon,i} \times \mathbb{T}^3$, the Hamiltonian $\tilde{H} = \tilde{h} + \epsilon P$ is reduced to the normal form $G_{\epsilon,i}$ defined on $\tilde{\Sigma}_{\epsilon,i}$ as shown in Lemma 3.1. The corresponding channels $\tilde{C}^\pm_i$ and $\tilde{C}^w_i$ are described in Theorem 6.3 and 6.4.

Since $T_i^{-1} \epsilon^\xi \geq \epsilon^\frac{1}{2}$, by replacing $\eta$ 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 $\{ |p - p'| \leq \xi \sqrt{\epsilon} \}$ is contained in $\tilde{\Sigma}_{\epsilon,i} \cap \tilde{\Sigma}_{\epsilon,i+1}$ and is mapped into $\tilde{\Sigma}'_{\epsilon,i}$ and into $\tilde{\Sigma}'_{\epsilon,i+1}$ 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 $\Gamma_k$ is smooth and $\epsilon$ is small, $\Gamma_{k,i}$ is o(\epsilon)-close to a straight line. For each disk $\tilde{\Sigma}_{\epsilon,i}$ we have local channel $\tilde{\mathcal{C}}_i^i$ made up of the flats $\mathcal{L}_{\tilde{\alpha}_H}(\nu_p \partial h(p))$ for $p \in \Gamma_{k,i}$, where $\nu_p$ is close to 1 such that $\tilde{\alpha}_H(\mathcal{L}_{\tilde{\alpha}_H}(\nu_p \partial h(p))) = h(p)$.

If $\tilde{\mathcal{C}}_i^i$ and $\tilde{\mathcal{C}}_{i+1}^i$ overlap, each flat $\mathcal{L}_{\tilde{\alpha}_H}(\nu_p \partial h(p)) \subset \tilde{\mathcal{C}}_i^i \cup \tilde{\mathcal{C}}_{i+1}^i$ either entirely lies in the intersection $\tilde{\mathcal{C}}_i^i \cap \tilde{\mathcal{C}}_{i+1}^i$, 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 $\tilde{\mathcal{C}}_i^i, \tilde{\mathcal{C}}_{i+1}^i$ overlap over a channel

$$\tilde{\mathcal{C}}_i^i \cap \tilde{\mathcal{C}}_{i+1}^i \supset \cup_{p \in \Gamma_{k,i} \cap \Gamma_{k,i+1}} \mathcal{L}_{\tilde{\alpha}_H}(\nu_p \partial h(p)),$$
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 $\Delta_{j}'$ if $(x(\theta),y(\theta))$ is an orbit lying in $\bar{A}(c)$ with $c \in C^w_{\bar{\epsilon}}$. Since $|k_i''||\partial_2V_{r''_i}| \to 0$ as $|k_i''| \to \infty$, we find from (6.7) that $\Delta_{j}'$ 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 $\text{Osc}||p(t)|| \leq C_V \sqrt{\epsilon}$ if $(p(t),q(t))$ is an orbit in $\bar{A}(c)$ for $c \in C^w_{\bar{\epsilon}}$.

Let us investigate the location of $p(t)$ if $(p(t),q(t))$ is a $\bar{c}$-minimal orbit. If both $\bar{c}$ 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 $\tilde{C}_k (k = k^*,k^*)$ is entirely contained in $C_{H_1} \sqrt{\epsilon}$-neighborhood of $\Gamma_k$. It follows from Lemma 7.4 that $(\partial h_t)$ would not be $\bar{c}$-minimal orbit of $\Phi^t_{\epsilon,F}$ if $p(t)$ entirely keeps away from $D_h \sqrt{\epsilon}$-neighborhood of $\tilde{c}$. Since $\text{Osc}||p(t)|| \leq C_V \Delta_{j}' \sqrt{\epsilon}$ if $(p(t),q(t)) \subset \bar{A}(c)$ for $\bar{c} = \Psi_k^t\Phi_k^t(c)$, each Aubry set $\bar{A}(\tilde{c})$ is entirely contained in $(C_V + D_h + C_H) \sqrt{\epsilon}$-neighborhood of $\tilde{c} \in \tilde{C}_k$ if we treat both as the points in $\mathbb{R}^3$.

The parameters $\xi$ and $\epsilon_0$ in the lemma are set such that

$$\xi \geq 4(C_V + D_h + C_H), \quad \xi \sqrt{\epsilon_0} \leq \frac{1}{2} \epsilon_0^\frac{3}{4}. $$

Then, there exists a segment $\Gamma'_{k,i}$ of $\Gamma_k$ 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 $\{||p - p'|| < \frac{\xi}{4} \sqrt{\epsilon}\} \subseteq \Sigma_{\epsilon,i} \cup \Sigma_{\epsilon,i+1}$. Indeed, $p' \in \Gamma'_{k,i}$ implies that the distance between $p'$ and the boundary of $\Sigma_{\epsilon,j}$ is not smaller than $\frac{\xi}{4} \sqrt{\epsilon}$ for $j = i,i + 1$.

For each $p' \in \Gamma'_k$, let $c' \in L_{\alpha_h}(\nu_{p'} \partial h(p'))$. Then, the Aubry set $\bar{A}(c')$ 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 $\Sigma_{\epsilon,j} \to \Sigma'_{\epsilon,j}$ 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 $(\Psi_{\epsilon,F}^t)^{-1}c'$ for $G_{\epsilon,j}$ is contained in the domain $T^2 \times \Sigma_{\epsilon,j} \times |X_j''| T$. It implies that $c' \in \tilde{C}^w_{\epsilon}$ for $j = i,i + 1$. □

Since $\tilde{C}^+_i$ and $\tilde{C}^-_i$ are joined by $\Psi_{\epsilon,F}^t(A_i)$, the whole path $\Gamma_\epsilon$ is covered by

$$\Gamma_\epsilon \subseteq \bigcup_{i \in A_\epsilon} \left( \Psi_{\epsilon,F}^t(A_i) \cup \tilde{C}^-_i \cup \tilde{C}^+_i \right) \cup \bigcup_{i \in A_\epsilon \setminus A_s} \tilde{C}^w_{\epsilon}. $$

(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 $C_{\epsilon_0}$ cusp-residual in $\mathcal{B}_{\epsilon_0} \subset C^\tau(B_D \times \mathbb{T}^3,\mathbb{R})$ with $\tau \geq 6$ such that for each $c \in C_{\epsilon_0}$, the path $\Gamma_\epsilon$ is a transition chain.

**Proof.** To check if $\Gamma_\epsilon$ is a transition chain under generic perturbation $\epsilon P$, one only needs to check, for generic perturbation $\epsilon P$, the condition of transition chain for each $c \in A_i \cup C^-_i \cup C^+_i$ if $i \in A_s$ and for each $c \in C^w_\epsilon$ if $A_\epsilon \setminus A_s$.

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

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

**CT:** $i \in A_\epsilon \setminus A_s$. The Aubry set $\bar{A}(c)$ lies on certain NHIC for each $c \in C^w_i$. 


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

**Lemma 8.2.** Let \( \mathfrak{B}_i \) be the residual set used in Theorem 6.1, Proposition 6.1, Theorem 6.2 and Theorem 6.3. Then, the following set

\[
\mathfrak{B}_i = \{ V_i + P''_i : V_i \in \mathfrak{B}_i, (\text{CT} : i \in \Lambda_s) \text{ holds for } h + \epsilon(V_i + P''_i) \}
\]

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 \( T^3 \times \{(p-p''_i) \leq (\eta T_i)^{-1}\epsilon\} \), the neighborhood of double resonant point \( p''_i \). In such a normal form, the main part is independent of \( \epsilon \) and the remainder \( R_{\epsilon,i} \) is uniformly bounded in the sense that \( \| R_{\epsilon,i} \|_{C^{r-2}} \leq a_0 \epsilon^a \) for all \( P \in \mathfrak{B}_1 \) if we consider it as the function of \((x, y, \vartheta)\) where \( \vartheta = \omega_3 |k''_i|^{-1/2} \theta \).

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 \( \mathfrak{B}_i \in C''(\mathbb{T}^2, \mathbb{R}) \), each \( V_i \in \mathfrak{B}_i \) is associated some small \( \epsilon_{V_i} > 0 \) such that the condition \((\text{CT} : i \in \Lambda_s) \) holds for each \( R_{\epsilon,i} \in \mathfrak{B}_{\epsilon,\vartheta} \subset C''(\Sigma_{\epsilon,\vartheta} \times T^3, \mathbb{R}) \). Because \( \| R_{\epsilon,i} \|_{C^{r-2}} \leq a_0 \epsilon^a \), for \( P = V_i + P''_i \in \mathfrak{B}_1 \) with \( V_i \in \mathfrak{B}_i \), the condition \((\text{CT} : i \in \Lambda_s) \) holds for each \( \epsilon \in (0, (a_0^{-1} \epsilon_{V_i})^{1/\kappa}) \).

Next, we consider NHICs away from strong double resonance. The path \( \Gamma_{k'} \) induces a decomposition \( C''(B_D \times \mathbb{T}^3, \mathbb{R}) = C''(B_D \times \mathbb{T}^1, \mathbb{R}) \oplus C''(B_D \times \mathbb{T}^3, \mathbb{R})/C''(B_D \times \mathbb{T}^1, \mathbb{R}) \) via \( P(p, q) = Z_{k'}(p, (k', q)) + P'(p, q) \) \((k' = k^*, k'')\) where \( Z_{k'} \) is defined in (5.2) consisting of Fourier modes of \( P \) in span \{k'\}, and \( P' = P - Z_{k'} \in C''(B_D \times \mathbb{T}^1, \mathbb{R})/C''(B_D \times \mathbb{T}^1, \mathbb{R}) \). Treating \((k', q)\) as a scalar variable \( x \) defined on \( T \), there exists an open-dense set \( \mathfrak{3}_{k'}^c \) of \( C''(\Gamma_{k'} \times T, \mathbb{R}) \) such that for each \( Z_{k'} \in \mathfrak{3}_{k'} \) 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 \( \epsilon \).

Given \( Z_{k'} \in \mathfrak{3}_{k'} \), let \{\( p''_{i_k} \in \Gamma_{k'} \)\} denote the set of bifurcation points, i.e. \( Z_{k'}(p''_{i_k}, x) \) has two maximal points in \( x \). Let \( N_p \) denote the cardinality of \{\( p''_{i_k} \in \Gamma_{k'} \)\}, it is finite. Let \( H_1 = h(p) + \epsilon Z_{k'}(p, (k', q)) \), then \( \Phi^t_{H_1} \) admits \( N_w + 1 \) pieces of NHIC consisting of minimal periodic orbits of \( \Phi^t_{H} \). 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 \( N \leq N_s + N_w + 2 \), where \( N_s = \#(\Lambda_s) \).

Indeed,

**Lemma 8.3.** In \( \mathfrak{3}_{k'} \) the following set

\[
\mathfrak{3}_{k'} = \cup \{ Z_{k'} + P' : Z_{k'} \in \mathfrak{3}_{k'}, (\text{CT} : i \in \Lambda \setminus \Lambda_s) \text{ holds for } h + \epsilon(Z_{k'} + P') \}
\]

is cusp-residual.

**Proof.** The normal form around weak resonant point \( p''_{i_k} \) 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 \( \frac{1}{2}(B_i, y, y) - V_i(x_1) \) survives the perturbations \( V''_i \) and \( R_{\epsilon,i} \) if the second derivative of \( V''_i \) at its minimal point is positive and some \( \epsilon \) is sufficiently small.

We notice \( V''_i(x) = -Z_{k'}(p''_{i_k}, x) \). Although the number \( \#(\Lambda \setminus \Lambda_s) \) increases to infinity as \( \epsilon \downarrow 0 \), it does not cause trouble. There is an open-dense set \( \mathfrak{3}_{k'} \subset C''(B_d \times \mathbb{T}^3, \mathbb{R}) \) with small \( \epsilon_{V_i} > 0 \) such that the condition \((\text{CT} : i \in \Lambda_s) \) holds for each \( \epsilon \in (0, (a_0^{-1} \epsilon_{V_i})^{1/\kappa}) \).
T, R), for each Zk′(p, x) ∈ Zk′, the second derivative of Zk′(p, x) in x at its maximal point is uniformly upper bounded below zero. So, the normal hyperbolicity of all NHICs of Gε,i is uniformly bounded from below as ε decreases to 0. Consequently, some εZk′ > 0 exists such that for each i ∈ Λs \ Λε the NHIC for Gε,i persists provided ∥Rε,i∥C2 ≤ εZk′.

By Lemma 3.1 again, given P = Zk′ + P′ ∈ B1 with Zk′ ∈ Zk′, the condition (CT: i ∈ Λs \ Λε) holds for each ε ∈ (0, (a−1 εZk′)1/κ].

For the coordinate transformation (3.1) the matrix Mi is set according to whether |k′′i,2| ≥ |k′′i,3| or not. The path ∂h(Γk′) admits a partition of four arcs. The condition |k′′i,2| > |k′′i,3| holds on two arcs and |k′′i,2| < |k′′i,3| holds on other two arcs. □

Let Ϩk′ = ∪ν∈R νP : P ∈ Ϩk′} ∩ B1 and ϩi = ∪ν∈R νP : P ∈ ϩi} ∩ B1, then the set Ϩk′ ∩ ϩk′ ∩ (∩ν ϩi) is residual in B1. Applying the Kuratowski-Ulam theorem (categorical analogue of the Fubini theorem, c.f. Chapter 15 of [Ox]), one obtains a residual set Ρ ∈ Ω1, each P ∈ Ρ is associated with a set IP residual in [0, 1] such that ∪ε∈Ω1 λP ∈ ϩk′ ∩ ϩk′ ∩ (∩ν ϩi).

Each P ∈ Ρ determines finitely many strong double resonant points {P′′i, i ∈ Λs(P)}. For each i ∈ Λs(P), one obtains the potential V = V = Vi(P) which determines a number εV > 0. Let aP = min{εV, εK, εV, i ∈ Λs(P)} where εV, εK > 0 are determined by the first resonant conditions k* and k* respectively (see Lemma 8.3). For any ε ∈ (0, aP), the flow ΦH with H = h + εP admits the conditions (CT: i ∈ Λs) and (CT: i ∈ Λs \ Λε).

Therefore, a cusp-residual set C′′o exists such that for each εP ∈ C′′o the conditions (CT: i ∈ Λs) and (CT: i ∈ Λs \ Λε) hold.

Some cohomology equivalence exists around each class c lying in the channels if 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 A(c) is a 2-dimensional torus. For each Hamiltonian of normal form Gε,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 Cε−2-topology instead of Cε-topology. One step KAM iteration for the construction of Φε,P makes Gε,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 εP ∈ C′′o is associated with a set ϨεP residual in a ball Bδ(εP) ⊂ C′′(B1 × T3, R) with small radius δ(εP) > 0, such that for each εP′ ∈ ϨεP the Hamiltonian h + εP + P′ possesses the property: the condition (HA) holds for each first cohomology classes c ∈ Ck∗ ∪ Ck∗ \ (∪i∈Λs Fε + di √ε) if the Aubry set A(c) consists of a two-dimensional torus.

Obviously, εP ∈ C′′o(εP + Bδ(εP)) ⊃ C′′o. Let Ρ′ = ∪εP ∈ C′′o(εP + ϩεP), because of the Kuratowski-Ulam theorem, it contains a cusp-residual set: there is a set Ρ′ residual in Ω1, each P ∈ Ω1 is associated with a set Rp residual in (0, aP) such that εP ∈ Ρ′ holds for all ε ∈ Rp.

For P ∈ Ρ ∩ Ρ′, Λs(P) is a finite set. Therefore, a set Rp residual in (0, aP) exists such that for each ε ∈ Rp, the conditions (CT: i ∈ Λs), (CT: i ∈ Λs \ Λε) and (HA) hold, namely, the flow ΦH with H = h + εP admits a transition chain. It verifies the cusp-residual property of the transition chain Γε. The proof of Theorem 8.2 is completed. □

Proof of Lemma 8.4. For εP ∈ C′′o, the Hamiltonian h + εP behaves like an a priori unstable system when it is restricted in neighborhood of c-minimal orbits.
with \( c \in \hat{C}_k \cup \hat{C}_k \setminus (\cup_{i \in \Lambda} F_i + d_i \sqrt{\epsilon}) \). Although \( \epsilon \) is small, it is treated as a fixed number since \( \epsilon P \) is fixed. There exists a small number \( \delta = \delta(\epsilon P) \) such that for any \( \epsilon P' \in \mathcal{B}_{\delta(\epsilon P)} \), the condition \( (CT; i \in \Lambda) \) and \( (CT; i \in \Lambda \setminus \Lambda \delta) \) holds for \( h + \epsilon (P + P') \).

Due to the covering property (8.3), we only need to check the condition \( (HA) \) for \( \Gamma_c \cap \hat{C}_k^+ \) and \( \Gamma_c \cap \hat{C}_k^- \). Restricted in \( \Sigma_{c,i} \cap \{ H^{-1} (E) \} \), 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 \( \Gamma_c \cap \hat{C}_k^+ \), the proof for the cases \( \Gamma_c \cap \hat{C}_k^- \) and \( \Gamma_c \cap \hat{C}_k^- \) is the same. Restricted on \( (\Sigma_{c,i} \times T^3) \cap \{ H^{-1} (E) \} \), the Hamiltonian \( \tilde{H} = \tilde{h}(\tilde{p}) + \epsilon P(\tilde{p}, \tilde{q}) \) is restored to the form \( G_{c,i} \) of (6.1), due to Lemma 3.1. Because of Theorem 6.2, there exist finitely many normally hyperbolic weakly invariant cylinders, denoted by \( \Pi_{c,i} \), such that for any \( c \in C_i^+ = (\Psi_i^+ \Psi_i^-)^{-1} \hat{C}_k^+ \), the Aubry set lies in these cylinders. Denoted by \( \Pi_{c,i} \) the time-0-section of \( \Pi_{c,i} \), i.e. \( \Pi_{c,i} = \tilde{H}_{\ell} \mid_{\theta = 0} \).

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

\[
\psi^* \omega = D\psi dx_1 \wedge dy_1
\]

where \( D\psi \) is the Jacobian of \( \psi \). Since the second de Rham cohomology group of \( \Pi \) is trivial, it follows from Moser’s argument on the isotopy of symplectic forms [Mo] that there exists a diffeomorphism \( \psi_1 \) on \( \Pi \) such that

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

Let \( \theta_1 = \frac{2k_j \sqrt{\epsilon}}{\omega_{2,i}} \). Since \( \Pi_{c,i} \) is invariant for the time-periodic map \( \Phi_{G_{c,i}} = \Phi_{G_{c,i}}^\theta_1 \) and \( \Phi_{G_{c,i}}^* \omega = \omega \), one has

\[
((\psi \circ \psi_1)^{-1} \circ \Phi_{G_{c,i}} \circ (\psi \circ \psi_1))^* dx_1 \wedge dy_1 = dx_1 \wedge dy_1
\]

(8.4)

i.e. \( (\psi \circ \psi_1)^{-1} \circ \Phi_{G_{c,i}} \circ (\psi \circ \psi_1) \) preserves the standard area. Each invariant circle \( \Gamma_{\sigma} \subset \Pi_{c,i} \) is pulled back to the standard cylinder, denoted by \( \Gamma_{\sigma}^* \) which is Lipschitz. The parameter \( \sigma \) 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 \( \psi, \psi_0 \) 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 \( \Pi_{c,i} \) may be crumpled 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, \( \alpha_{c,i} (L_{\alpha_{c,i}} (\lambda_{c,i}^{+} g_{c,i}^{+})) = d_i > 0 \), the cylinders are moderately crumpled. The constant \( b_1 \) is therefore uniformly bounded for \( \sigma \) if we are restricted on the cylinder \( \Pi_{c,i} \).

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 \( \mathcal{R}_i \)

\[
\mathcal{R}_i : \begin{cases} 
& x_j = u_j, \quad y_j = \sqrt{\epsilon}^{-1} (v_j - v''_{i,j}), \quad j = 1, 2, \\
& \theta = \sqrt{\epsilon}^{-1} x_3, \quad I = \sqrt{\epsilon}^{3} (v_3 - v''_{i,3}).
\end{cases}
\]

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

\[
G_{c,i}^*(u_1, u_2, u_3, v_1, v_2) = \epsilon G_{c,i}(x(u_1, u_2), \theta(u_3), y(v_1, v_2))
\]

(8.5)
where \( u_3 \) plays the role of time. Because of (8.5), the function \( G^*_{\epsilon,i} \) solves the equation 
\[ \mathcal{M}_i^* \Phi_{\epsilon,F} \mathcal{M}_0^* H(u, v, u_2, -G^*_{\epsilon,i}) = E, \]  
see Formula (3.10).

Obviously, \( \hat{\Pi}_\epsilon^* = \mathcal{R}_{\epsilon} \Pi_\epsilon \) is weakly invariant for the flow of \( G^*_{\epsilon,i} \). Let \( \Pi_\epsilon^* = \hat{\Pi}_\epsilon^* |_{u_3=0} \). Compared with \( \Pi_\epsilon \), 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 \( \Gamma^*_{i,\sigma} \) denote the invariant circle in \( \Pi_\epsilon^* \), one has \( \| \Gamma^*_\sigma - \Gamma^*_{\sigma'} \|_{C^0} \leq \frac{b_1}{\sqrt{\epsilon}} \sqrt{|\sigma - \sigma'|} \). Let 
\[ \hat{\Gamma}^*_\sigma = \cup_{u_3 \in [0, |k^*_\sigma|]} \Phi_{u_3}^{k^*_3} \Gamma^*_{i,\sigma} \], one has
\[ \| \hat{\Gamma}^*_\sigma - \hat{\Gamma}^*_{\sigma'} \|_{C^0} \leq \frac{b_1}{\sqrt{\epsilon}} \max_{u_3 \in [0, |k^*_\sigma|]} \| D\Phi_{u_3}^{k^*_3} \| \sqrt{|\sigma - \sigma'|} \]
Let \( H^* = \mathcal{M}_i^* \Phi_{\epsilon,F} \mathcal{M}_0^* H \). Corresponding to the cylinder \( \hat{\Pi}_\epsilon^* \), there is a cylinder \( \hat{\Pi}_\epsilon^* \) modulo the shift \( \sigma_i \), it shrinks in the energy level set \( \{ H^*-1(E) \} \)
\[ \hat{\Pi}_\epsilon^* = \{(u, v) \in \mathbb{T}^3 \times \mathbb{R}^3 : (u, v_1, v_2) \in \hat{\Pi}_\epsilon^*, \ v_3 = G^*_{\epsilon,i}(u, v_1, v_2) \} \].

Those invariant 2-tori \( \{ \hat{\Gamma}^*_\sigma \} \) lying in the cylinder. Under the inverse of \( \mathcal{M}_i \) and \( \Phi_{\epsilon,F} \), the tori \( \{ \hat{\Gamma}^*_\sigma \} \) and the cylinder \( \hat{\Pi}_\epsilon^* \) are mapped to the invariant tori and weakly invariant cylinder for the flow of the Hamiltonian \( \mathcal{M}_0^* H \):
\[ \hat{\Pi}_\epsilon = \Phi_{\epsilon,F}^{-1} \mathcal{M}_i^{-1} \hat{\Pi}_\epsilon^*, \quad \hat{\Gamma}^*_\sigma = \Phi_{\epsilon,F}^{-1} \mathcal{M}_i^{-1} \hat{\Gamma}^*_\sigma. \]

Since \( \Phi_{\epsilon,F} \) is a diffeomorphism close to identity and \( \mathcal{M}_i \) is linear, there exists a number \( b_2 > 0 \) such that
\[ \| \hat{\Gamma}^*_\sigma - \hat{\Gamma}^*_{\sigma'} \|_{C^0} \leq \frac{b_2}{\sqrt{\epsilon}} \sqrt{|\sigma - \sigma'|}. \]  
(8.6)

Because \( \partial \mathcal{M}_0^* h(\bar{p}^\prime) = \bar{\omega} = (0, \bar{\omega}_2, \bar{\omega}_3) \) satisfies the resonant condition \( k^\prime_{i,2,3} \), we construct another canonical transformation \( \mathcal{M}_i : \phi = \tilde{M}_i^{-1} \bar{q} \) and \( I = \tilde{M}_i \bar{p} \), where
\[ \tilde{M}_i = \text{diag}(1, \tilde{M}_{i,2}), \quad \text{with} \quad \tilde{M}_{i,2} = \begin{bmatrix} k^\prime_{i,2} & j_2 \\ k^\prime_{i,3} & j_3 \end{bmatrix} \]
the integers \( j_2, j_3 \) are chosen such that \( k^\prime_{i,2,j_3} - k^\prime_{i,3,j_2} = 1 \). Clearly, \( \tilde{M}_i \) is uni-modular and the first derivative of \( \mathcal{M}_i^* \mathcal{M}_0^* h \) in \( I_3 \) is not equal to zero at the point \( I^\prime \) \( = \tilde{M}_i \bar{p}^\prime \). Therefore, there exists a \( C^r \)-function \( G^*_{\epsilon,i} \) which solves the equation 
\[ \mathcal{M}_i^* \mathcal{M}_0^* H(\phi, I_1, I_2, -G^*_{\epsilon,i}(\phi, I_1, I_2)) = E \]  
(8.7)
when it is restricted in a neighborhood of \( I^\prime \) which covers the domain \( \tilde{\Sigma}_{\epsilon,i} \times \mathbb{T}^3 \) under the map \( \mathcal{M}_i \). The function \( G^*_{\epsilon,i} \) defines a Hamiltonian system with two and half degrees of freedom where \( \phi_3 \) plays the role of time, there is a normally hyperbolic and weakly invariant cylinder \( \Pi_\epsilon^* \) for \( G^*_{\epsilon,i} \) such that
\[ \hat{\Pi}_\epsilon = \mathcal{M}_i^{-1}((\phi, I) \in \mathbb{T}^3 \times \mathbb{R}^3 : (\phi, I_1, I_2) \in \hat{\Pi}_\epsilon^*, \ I_3 = G^*_{\epsilon,i}(\phi, I_1, I_2)) \].

Let \( \hat{\Gamma}^*_\sigma = \mathcal{M}_0^{-1} \hat{\Gamma}^*_\sigma \), they lie in the cylinder \( \hat{\Pi}_\epsilon^* \) for which the modulus of continuity of (8.6) still holds, probably with a larger coefficient \( b_2^* \geq b_2 \).
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

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

is not larger than $\delta_i$ (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 $G_{\epsilon,i}^* \to G_{\epsilon,i}^* + G_\delta$ can be achieved by the perturbation $\mathcal{M}_i^* \mathcal{M}_0^* H(I_1, I_2, I_3, \phi) \to \mathcal{M}_i^* \mathcal{M}_0^* H(I_1, I_2, I_3 + G_\delta, \phi)$. Therefore, a set $\mathcal{R}_{i,\ell}$ residual in $\mathcal{B}_\delta$ is correspondingly obtained for $H$ such that for each $\epsilon P^\ell \in \mathcal{B}_\delta$ the condition (HA) holds for cylinder $\hat{P}_\ell$. Taking the intersection $\cap \mathcal{R}_{i,\ell}$ which is still residual in $\mathcal{B}_\delta$. $\square$

**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 $\delta$-neighborhood of some resonant path $\Gamma_k$ 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 $\Gamma: [0, 1] \to H^1(T^3, \mathbb{R})$ joining $c$ to $c'$. Then, there exists an orbit of the Lagrange flow $\Phi_L^{d\gamma}: \mathbb{R} \to T^3$ that connects the two Aubry sets: $\alpha(d\gamma) \subset \hat{A}(c)$ and $\omega(d\gamma) \subset \hat{A}(c')$.

**Proof of Theorem 1.1.** If both $p$ and $\bar{c}$ are treated as point in $\mathbb{R}^3$, $p(t) \equiv \bar{c}$ holds along each $\bar{c}$-minimal orbit $(p(t), q(t))$ of $\Phi_h^{d\gamma}$. 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 $\hat{A}(\bar{c})$ one has $\|p(t) - \bar{c}\| < \frac{\delta}{2}$ if $\epsilon > 0$ is suitably small. So, any orbit connecting $\hat{A}(\bar{c})$ to $\hat{A}(c')$ with $\|\bar{c} - p\| < \frac{\delta}{2}$ and $\|\bar{c} - p\| < \frac{\delta}{2}$ satisfies the condition: some $t^*, t^* \in \mathbb{R}$ exist such that $\|p(t^*) - p^*\| < \delta$ and $\|p(t^*) - p^*\| < \delta$.

According to Theorem 5.1 of [C17b], for each $\epsilon P \in C_{t,,0}$, there is a transition chain that connects the first cohomology classes $\bar{c}^*, c^* \in H^1(T^3, \mathbb{R})$ satisfying the condition $\alpha(\bar{c}^*) = \alpha(c^*) = E$, $\|p^* - \bar{c}^*\| < \frac{\delta}{2}$ and $\|p^* - \bar{c}^*\| < \frac{\delta}{2}$. In this case, one obtains from Theorem 3.1 of [LC] an orbit connecting the Aubry set $\hat{A}(\bar{c}^*)$ to $\hat{A}(c^*)$. 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 $\delta$-dense of the diffusion orbits, slightly stronger than Theorem 1.1. A diffusion orbit is said to be $\delta$-dense if it passes through $\delta$-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 Tonelli $C^r$-Hamiltionian $H(p, q, t)$ where $(p, q, t) \in \mathbb{R}^2 \times T^3$, $r \geq 3$, we have a Tonelli Lagrangian $L(q, q, t) = \max_p \langle \dot{q}, p \rangle - H(p, q, t)$. Let $\Phi_H^{t,\epsilon}$ denote the
Hamiltonian flow of $H$, it maps the initial value at the time-$t$-section to the time-$t'$-section. For $\Phi^t_{H}$ 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 $\Gamma_c : [0, 1] \to H^1(\mathbb{T}^2, \mathbb{R})$ 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 $\tilde{A}(c)$ is an invariant 2-dimensional torus $\tilde{\Upsilon}_c$ lying in $\tilde{\Pi}$, it is a deformation of the torus $\{\tilde{p}, \tilde{q}, \tilde{t} \in \mathbb{R}^2 \times \mathbb{T}^3 : (p_1, q_1) = 0, p_2 = \text{const.}\}$

Let $\pi : \mathcal{M} = \{q : q_1 \mod 4\pi, q_2 \mod 2\pi \} \to \mathbb{T}^2$ be a finite covering space of $\mathbb{T}^2$. The lift of $\tilde{\Pi}$ to $\mathbb{T}^3 \times \mathbb{T}^2$ consists of two copies, denoted by $\tilde{\Pi}_{0,1}$ and $\tilde{\Pi}_{1,0}$. For $c \in \Gamma_c$, if the Aubry set $\tilde{A}(c)$ is an invariant torus $\tilde{\Upsilon}_c \subset \tilde{\Pi}$, its lift also consists of two components, $\tilde{\Upsilon}_{c,t} \subset \tilde{\Pi}_{0,t}$ and $\tilde{\Upsilon}_{c,r} \subset \tilde{\Pi}_{1,r}$. Let $\tilde{\Pi}_0, \tilde{\Pi}_{0,0}, \tilde{\Pi}_{1,0}, \tilde{\Upsilon}_{0,0}$ and $\tilde{\Upsilon}_{c,0}$ denote the time-0-section of $\tilde{\Pi}$, $\tilde{\Pi}_{0,t}$, $\tilde{\Pi}_{1,r}$, $\tilde{\Upsilon}_{c,t}$ and $\tilde{\Upsilon}_{c,r}$ respectively. Denote by $\pi$ the projection such that $\pi(p, q, t) = (q, t)$, let $\Upsilon = \pi\tilde{\Upsilon}$. Let $\Gamma^*_c \subset \Gamma_c$ such that

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

Let $B_D \subset \mathbb{R}^2$ denote 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 $\mathcal{B}_c \subset C''(B_D \times \mathbb{T}^3, \mathbb{R})$ denote a ball about the origin of radius $\epsilon > 0$.

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

$$N(c, M)|_{t=0}\setminus(A(c, M) + \delta)|_{t=0} \neq \emptyset$$

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

$$\int^\tau_t L(\dot{\gamma}^-_{c,x}(s), \gamma^-_{c,x}(s), s) - \langle c, \gamma^-_{c,x}(s) \rangle ds$$

$$\leq \int^{\tau'}_{t'} L(\dot{\xi}(s), \xi(s), s) - \langle c, \xi(s) \rangle ds + (t - t')\alpha(c)$$

holds for any absolutely continuous function $\xi : [t', \tau] \to \mathbb{T}$ with $\xi(\tau) = \gamma^-_{c,x}(\tau)$, $\xi(t') = \gamma^-_{c,x}(t)$ with $t' = t \mod 2\pi$. It produces an orbit $\langle \gamma^-_{c,x}(t), \gamma^-_{c,x}(t) \rangle$ 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 $\langle \gamma^+_{c,x}(t), \gamma^+_{c,x}(t) \rangle$ 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 $u^+_c$ such that $\gamma^+_{c,x}(\tau) = \partial_p H(\partial_q u^+_c(q, t, c, q, \tau))$.

Given an Aubry class for $c \in \Gamma_c$ we can define its elementary weak KAM solution. In the covering space $\mathcal{M}$, there are two Aubry classes for $c \in \Gamma_c$, $\tilde{\Upsilon}_{c,t}$ and $\tilde{\Upsilon}_{c,r}$. To define the elementary weak KAM solution $u^+_{c,t}$ for $\tilde{\Upsilon}_{c,t}$, we construct a perturbation $L(q, q, t) \to L(q, q, t) + V_\ell(q, t)$, where $V_\ell \geq 0$ and $\Upsilon_{c,r} \subset \text{supp}V_\ell \subset (\Upsilon_{c,r} + \delta) =$
\{(q,t) : \text{dist}((q,t), \Upsilon_{c,r}) \leq \delta\}. For \( V_t \neq 0 \), there exists a unique weak KAM solution \( u_{c,V_t}^{\pm} \) modulo constant. For almost every point \((q,t) \in M \times T\), the function \( u_{c,V_t}^{\pm} \) determines a unique forward and backward minimal orbit \((\gamma_{c,x}^{\pm}(t)), \tau_{c,x}^{\pm}(t))\) such that \( \dot{\gamma}_{c,x}^{\pm}(\tau) = \partial_{\tau}H(\partial_q u_{c,x}^{\pm}(q,\tau) + c,q,\tau) \) and the orbits approaches \( \Upsilon_{c,r} \) as \( t \to \pm \infty \) respectively. Let \( V_{c,\ell} \downarrow 0 \), the function \( u_{c,V_{c,\ell}}^{\pm} \) converges to a function \( u_{c,\ell}^{\pm} \) which is obviously a weak KAM solution for \( H \), it is called elementary weak KAM solution for \( \Upsilon_{c,\ell} \). The elementary weak KAM solution for \( \Upsilon_{c,\ell} \) is defined in the same way, denoted by \( u_{c,\ell}^{\pm} \).

For almost every point \((q,t) \in \tilde{M} \times T \setminus \Upsilon_{c,\ell} \) the initial condition \((\partial_{\tau}u_{c,\ell}^{\pm}(q,t) + c,q,t)\) determines a forward (backward) c-minimal orbit that approaches \( \Upsilon_{c,\ell} \) as \( t \to \pm \infty \).

For points \((q,t) \in \tilde{M} \times T \), \( u_{c,\ell}^{\pm} \) determines c-minimal orbit approaching \( \Upsilon_{c,\ell} \).

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

\[
B_c^\ell(q,t) = u_{c,\ell}^-(q,t) - u_{c,\ell}^+(q,t), \quad B_c^r(q,t) = u_{c,r}^-(q,t) - u_{c,\ell}^+(q,t).
\]

In the following, we only study \( B_c^\ell \). The arguments for \( B_c^r \) 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 \( B_c^\ell \), 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 \( u(x + x') - u(x) \leq \langle p, x' \rangle \) 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 \( B_c^\ell \) is a sum of two semi-concave functions, its sup-derivative is the sum of the sup-derivatives of \( u_{c,\ell}^- \) and \( -u_{c,\ell}^+ \). Therefore, \( D^+B_c^\ell \) is a single point iff both \( D^+u_{c,\ell}^- \) and \( D^+(-u_{c,\ell}^+) \) are singleton [CaC].

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

**Proof.** By the definition, \( \partial u_{c,\ell}^- = \partial u_{c,\ell}^+ \) holds at a global minimal point of \( B_c^\ell \), 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 \( \tilde{M} \).

For a class \( c \in \Gamma^*_c \), the covering space \( \tilde{M} \times T \) is divided into two annuli \( A_{c,x} \) and \( \hat{A}_{c,\ell} \), bounded by \( \Upsilon_{c,\ell} \) and \( \Upsilon_{c,r} \). Clearly, one has \( \hat{\pi} A_{c,r} = \hat{\pi} A_{c,\ell} \). The set \( N(c,M) \setminus A(c,M) \) 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 \( \Upsilon_c \).

**Lemma A.3.** There is a finite partition of \( \Gamma_c^* : \Gamma_c = \bigcup I_k \), each \( I_k \) is a segment of \( \Gamma^*_c \). For each \( I_k \), there is an annulus \( N_k \subset \hat{A}_{c,r} |_{t=0} \) such that for each \( c \in I_k \cap \Gamma^*_c \)

1. dist\((N_k, \Upsilon_{c,\ell} \cup \Upsilon_{c,r}) \geq \delta\);
2. each curve \( (\gamma(t),t) \) lying in \( (N(c,M) \setminus A(c,M)) \cap A_{c,x} \) passes through \( N_k \);
3. for each backward (forward) c-minimal curve \( \gamma \), let \( \{q_i = \gamma(2i\pi) \in N_k\} \), then \( |q_i - q_j| \geq \delta \) if \( i \neq j \).
Proof. Because $\Gamma_\epsilon$ is compact, the speed of each $c$-minimal orbit is uniformly upper bounded for all $c \in \Gamma_\epsilon^*$. 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 $N_c = A_{c,\epsilon,\alpha,\beta} \setminus (\Upsilon_{c,\ell} \cup \Upsilon_{c,r} + \delta_c)$ 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 $|c - c'| \leq \delta'_c$ and $c', c \in \Gamma_\epsilon^*$. The first two items are then proved if we notice $\Gamma_\epsilon^*$ is compact.

For the third one, we notice that the condition $\gamma(2i\pi) = \gamma(2j\pi)$ for $i \neq j$ implies that $\gamma$ 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.

By the definition, the Aubry set $\tilde{\mathcal{A}}(c)$ is an invariant torus if $c \in \Gamma_\epsilon^*$. 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 cylinders. 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 $\max_{q} |\hat{\Upsilon}_{\sigma,0}(q) - \hat{\Upsilon}_{\sigma',0}(q)|$ and the length of the horizontal diagonal is not shorter than $\frac{1}{C_L} \max_{q} |\hat{\Upsilon}_{\sigma,0}(q) - \hat{\Upsilon}_{\sigma',0}(q)|$. So, one has

$$\max_{q} |\hat{\Upsilon}_{\sigma,0}(q) - \hat{\Upsilon}_{\sigma',0}(q)| \leq \sqrt{2C_L}|\sigma - \sigma'|.$$ 

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 $c \in \Gamma_\epsilon$ such that the Aubry set is the circle. The parameter $\sigma$ is usually defined on a Cantor set, denoted by $\Sigma$. 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

$$|B_{c(\sigma)}(q, 0) - B_{c(\sigma')}(q, 0)| \leq C(\sqrt{|\sigma - \sigma'|} + |c - c'|).$$

**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 $\tilde{M}$, we have introduced the elementary weak KAM solutions $u_{c,\ell}^\pm$ and $u_{c,r}^\pm$. Since the projection $\pi$ is an injection when it is restricted in the neighborhood $\Upsilon_{c,i} + \delta$ for $i = \ell, r$ respectively, for $(q, t) \in \Upsilon_c + \delta$ one has

$$u_{c,\ell}^\pm(\pi^{-1}(q, t) \cap (\Upsilon_{c,\ell} + \delta)) = u_{c,r}^\pm(\pi^{-1}(q, t) \cap (\Upsilon_{c,r} + \delta)) = u_c^\pm(q, t). \quad (A.1)$$

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

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

1

1
which becomes an equality when $\gamma$ 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

$$u^-_{c, \ell}(q, 0) - u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) = \int_{-2K\pi}^{0} (L(\dot{\gamma}_{c, q}(s), \gamma_{c, q}(s), s) - \langle c, \dot{\gamma}_{c, q}(s) \rangle) ds$$

$$+ 2K\pi a(c),$$

$$u^-_{c, \ell}(q, 0) - u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) \leq \int_{-2K\pi}^{0} (L(\dot{\gamma}_{c, q}(s), \gamma_{c, q}(s), s) - \langle c, \dot{\gamma}_{c, q}(s) \rangle) ds$$

$$+ 2K\pi a(c').$$

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 $\Gamma_c$, the $\alpha$ function is convex and finite everywhere, there is some constant $C_1$ such that $|\alpha(c') - \alpha(c)| \leq C_1 |c - c'|$. Let $\gamma_{c, q}$ be the lift of $\gamma_{c, q}$ to the universal covering space, one has $|\gamma_{c, q}(0) - \gamma_{c, q}(-2K\pi)| \leq 2C_2 K\pi$.

$$u^-_{c, \ell}(q, 0) - u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) - (u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) - u^-_{\ell}(\gamma_{c, q}(-2K\pi), 0))$$

$$\leq 2K\pi(C_1 + C_2)|c - c'|.$$

In the same way one can also obtain

$$u^-_{c, \ell}(q, 0) - u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) - (u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0) - u^-_{c, \ell}(\gamma_{c, q}(-2K\pi), 0))$$

$$\leq 2K\pi(C_1 + C_2)|c - c'|.$$
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

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

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 $\max\{|q - q'| : q, q' \in \text{supp} S_\delta\} \leq d$, 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

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

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

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

We introduce a smooth function $\kappa$ such that $\kappa(q', q) = 1$ if $|q' - q| \leq K$ and $\kappa(q', q) = 0$ if $|q' - q| \geq K + 1$. Let $\Phi'$ 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 $S_\delta$ is $C^r$-small. We choose a smooth function $\rho(s)$ with $\rho(-\tau) = 0$, $\rho(0) = 1$ and let $\Phi_s$ be the symplectic map produced by $S_{-\tau} + \rho(s)\kappa S_\delta$ and let $\Psi_s = \Phi_s' \circ \Phi^{-1}$. Clearly, $\Psi_s$ defines a symplectic isotopy between the identity map and $\Psi$. 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 Hamiltonorphism is a subgroup of symplectic diffeomorphism, there is a function $H'$, sufficiently close to $H$, such that $\Phi_{H'}^{0,0} = \Phi_{H_1}^{0,0} \circ \Phi_{H,0}^{-1}$.

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

$$ u_{c,i}^{\pm, S_\delta}(q, 0) = \min_{|q' - q| \leq 2D \tau} (S_{-\tau}(q', q) + S_\delta(q) - \langle c, q - q' \rangle + u_{c,i}^-(q', -\tau)) $$

$$ = S_\delta(q) + \min_{|q' - q| \leq 2D \tau} (S_{-\tau}(q', q) - \langle c, q - q' \rangle + u_{c,i}^-(q', -\tau)) $$

$$ = S_\delta(q) + u_{c,i}^-(q, 0). $$
Obviously, one has $u_{e,3}^+(S\delta, q, 0) = u_{e,1}^+(q, 0)$. The lemma is proved because the barrier function is the difference of the two functions. □

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

$$\text{Argmin}(S_{d_1}(q^*), B) = \{q \in S_{d_1}(q^*): B(q) = \min B\}.$$ 

Let $\pi_i$ 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 $S_{d_1}(q^*)$ if $\pi_i V \cap S_{d_1}(q^*) = \pi_i S_{d_1}(q^*)$ holds for $i = 1$ or $2$. Otherwise, it is said to be trivial for $S_{d_1}(q^*)$. Let $B_{c, \delta}$ 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 $\mathcal{D}$ open-dense in $\mathcal{B}_c$ such that for each $H_\delta \in \mathcal{D}$, it holds simultaneously for all $c \in I_k \cap \Gamma_c^*$ that the set $\text{Argmin}(S_{d_1}(q^*), B_{c, \delta})$ is trivial for $S_{d_1}(q^*)$ provided $S_{d_1}(q^*) \subset N_k$ and $d_1 < d/3$ is suitably small.

**Proof.** The openness is obvious. To show the density, we construct the perturbations $H_\delta \in \mathcal{B}_c$ 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 $q \in \text{supp}S_\delta$. Because of Lemma A.5, it works.

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

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

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

$$\pi_1 \text{Argmin}(S_{d_1}(q^*), B_{c, \delta}) \subset (q_1^* - d_1, q_1^* + d_1].$$

(A.2)

Let $\mathcal{F}_c = \{B_{c, \delta}(q, 0): c \in \Gamma_c^*\}$ be the set of barrier functions. For $i = 1, 2$ we set

$$\mathcal{F}_i = \{B \in C^0(S_{d_1}(q^*), \mathbb{R}): \pi_i \text{Argmin}(S_{d_1}(q^*), B) = [q_i^* - d_1, q_i^* + d_1]\},$$

where $q_i^* = (q_{i1}^*, q_{i2}^*)$. If the density does not exist, there would be small $\epsilon > 0$, for each $S_\delta \in \mathcal{B}_{1, \epsilon}$, some $c \in \Gamma_c^*$ exists such that $B_{c, \delta} + S_\delta \in \mathcal{F}_1$. Let $\mathcal{F}_{1, \epsilon}$ be the intersection of $\mathcal{B}_{1, \epsilon}$ with a $k$-dimensional subspace. The box-dimension of $\mathcal{F}_{1, \epsilon}$ in $C^0$-topology will not be smaller than $k$.

For any $B_{c, \delta} \in \mathcal{F}_c$ there is only one $S_\delta \in \mathcal{B}_{1, \epsilon}$ such that $B_{c, \delta} + S_\delta \in \mathcal{F}_1$. Otherwise, there would be $S'_\delta \neq S_\delta$ such that $B_{c, \delta} + S'_\delta \in \mathcal{F}_1$ also. As we have $B_{c, \delta} + S'_\delta = B_{c, \delta} + S_\delta + S'_\delta - S_\delta$, where $B_{c, \delta} + S_\delta \in \mathcal{F}_1$ and $S'_\delta \sim S_\delta$, which contradicts the definition of $\mathcal{G}_1$. For $S_\delta \in \mathcal{B}_{1, \epsilon}$, let $\mathcal{G}_{S_{\delta}} = \{B_{c, \delta} \in \mathcal{F}_c: B_{c, \delta} + S_\delta \in \mathcal{F}_1\}$. If the density does not exist, $\mathcal{G}_{S_{\delta}}$ is non-empty. For any $S'_\delta, S'' \in \mathcal{B}_{1, \epsilon}$, each $B_{c, \delta} \in \mathcal{G}_{S_{\delta}}$ and each $B_{c, \delta} \in \mathcal{G}_{S''}$
one has
\[ d(B_c^\epsilon, B_{c'}^\epsilon) = \max_{q \in S_{d_1}(q^*)} |B_c^\epsilon(q, 0) - B_{c'}^\epsilon(q, 0)| \]
\[ \geq \max_{|q_1 - q'_1| \leq d_1} \left| \min_{|q_2 - q_2'| \leq d_1} B_c^\epsilon(q, 0) - \min_{|q_2 - q_2'| \leq d_1} B_{c'}^\epsilon(q, 0) \right| \tag{A.3} \]
\[ = \max_{|q_1 - q'_1| \leq d_1} |S_\delta(q) - S_\delta'(q)| = d(S_\delta, S_\delta') \]

where \( q = (q_1, q_2) \) and \( d(\cdot, \cdot) \) denotes the \( C^0 \)-metric. It implies that the box-dimension of the set \( S_c \) is not smaller than the box-dimension of \( B_1^\epsilon, \epsilon \) in \( C^0 \)-topology. Guaranteed by the modulus continuity of Lemma A.4, the box dimension of the set \( S_c \) 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 \( \Omega_{2, \epsilon} \) open-dense in \( B_{2, \epsilon} \) such that for each \( S_\delta \in \Omega_{2, \epsilon} \) it holds simultaneously for all \( c \in I_k \cap \Gamma_c^* \) that
\[ \pi_2 \text{Argmin}(S_{d_1}(q^*), B_c^\epsilon + S_\delta) \subseteq [q_2^* - d_1, q_2^* + d_1]. \tag{A.4} \]

Therefore, \( \exists \) arbitrarily small \( S_{1, \delta} \in B_{1, \epsilon} \) such that \( \pi_1 \text{Argmin}(S_{d_1}(q^*), B_c^\epsilon + S_{1, \delta} + S_{2, \delta}) \) is trivial for \( S_{d_1}(q^*) \) and for all \( c \in I_k \cap \Gamma_c^* \). Due to Lemma A.5 we obtain the density. \( \square \)

To finish the proof of Theorem A.1, we split the annulus \( N_k \) equally into squares \( \{ S_j = |q - q_j| \leq d_j \} \). For each \( S_j \), there exists an open-dense set \( \Omega_{k,j} \subset B_c \), for each \( H_\delta \in \Omega_{k,j} \) it holds simultaneously for all \( c \in I_k \cap \Gamma_c^* \) that the set \( \text{Argmin}(S_j, B_c^\epsilon) \) is trivial for \( S_j \). The intersection \( \cap \Omega_{k,j} \) is still open-dense in \( B_c \). For each \( H_\delta \in \cap_{k,j} \Omega_{k,j} \), it holds simultaneously for all \( c \in \Gamma_c^* \) that the diameter of each connected component of the Mañé set is not larger than \( \frac{2}{5} d_1 \) if it keeps away from the Aubry set. \( \square \)

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., 26/8 (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.
ARNOLD DIFFUSION IN NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

[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. Delshams and G. Huguet, Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems, Nonlinearity, 22 (2009), pp. 1997–2077.
[DH2] A. Delshams 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 $A^3$, 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.