Gevrey genericity of Arnold diffusion in a priori unstable Hamiltonian systems

Qinbo Chen\textsuperscript{1,2,}\textsuperscript{*} and Chong-Qing Cheng\textsuperscript{1}

\textsuperscript{1} Department of Mathematics, Nanjing University, Nanjing 210093, People’s Republic of China
\textsuperscript{2} Morningside Center of Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, People’s Republic of China

E-mail: chenqb@amss.ac.cn and chengcq@nju.edu.cn

Received 28 February 2018, revised 27 July 2020
Accepted for publication 1 September 2020
Published 21 January 2021

Abstract

It is well known that under generic $C^r$ smooth perturbations, the phenomenon of global instability, known as Arnold diffusion, exists in a priori unstable Hamiltonian systems. In this paper, by using variational methods, we will prove that under generic Gevrey smooth perturbations, Arnold diffusion still exists in the a priori unstable Hamiltonian systems of two and a half degrees of freedom.

Keywords: Arnold diffusion, genericity, Gevrey functions, a priori unstable

Mathematics Subject Classification numbers: primary 37J40, 37J50.

(Some figures may appear in colour only in the online journal)

1. Introduction

Throughout this paper, we denote by $T^n \times \mathbb{R}^n$ the cotangent bundle $T^*T^n$ of the torus $T^n$ with $T = \mathbb{R}/\mathbb{Z}$, and endow $T^n \times \mathbb{R}^n$ with its usual coordinates $(q, p)$ where $q = (q_1, \ldots, q_n)$ and $p = (p_1, \ldots, p_n)$. We also endow the phase space with its canonical symplectic form $\Omega = \sum_{i=1}^{n} dq_i \wedge dp_i$. A Hamiltonian system is usually a dynamical system governed by the following Hamilton’s equations

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}, \quad (q, p) \in T^n \times \mathbb{R}^n$$

where $H(q, p, t)$ is a Hamiltonian function and the dependence on the time $t$ is 1-periodic, so $t \in T$.

The goal of this paper is to present global instability in a class of Hamiltonians. The problem of the influence of small perturbations on an integrable Hamiltonian system was considered by

\textsuperscript{*}Author to whom any correspondence should be addressed.
Poincaré to be the fundamental problem of Hamiltonian dynamics. It is customary to consider a nearly-integrable Hamiltonian system of the form $H = H_0(p) + \varepsilon H_1(q, p, t)$. Notice that for $\varepsilon = 0$ such systems do not admit any instability phenomenon. For $0 < \varepsilon \ll 1$, the celebrated KAM theory asserts that a set of nearly full measure in the phase space consists of invariant tori carrying quasi-periodic motions, and the oscillation of the action variables $p$ on each KAM torus is at most of order $\sqrt{\varepsilon}$. For $n \geq 2$, the complement of the set of the union of KAM tori is connected, so the natural question arises whether it is possible to find large evolution of order 1. In the celebrated paper [1], Arnold first proposed an example of a nearly-integrable Hamiltonian system with two and a half degrees of freedom, which admits trajectories whose action variables have large oscillation. Moreover, he also conjectured that such an instability phenomenon occurs in generic nearly-integrable systems. This is known as the Arnold diffusion conjecture, and it has been investigated extensively since then.

The mechanism of Arnold’s original example is based on the existence of a normally hyperbolic invariant cylinder (NHIC) foliated by a family of hyperbolic invariant tori or whiskered tori. The unstable manifold of one torus intersects transversally the stable manifold of another nearby torus. These tori constitute a transition chain along which diffusion takes place. By Nekhoroshev theory, it must be extremely slow. This mechanism has inspired a large number of studies to the Hamiltonians possessing certain hyperbolic geometric structures. In the literature, such a system is referred to as ‘a priori unstable’, to be distinguished from a nearly-integrable system (i.e. ‘a priori stable’). There have been many works devoted to the a priori unstable systems based on Arnold’s geometric mechanism, and most of them have tried to find transition chains in more general cases [8, 9, 29, 52, 60, 79], etc. However, for a general a priori unstable system the transition chain cannot be formed by a continuous family of tori but a Cantorian family, and the size of gaps between the persisting tori could be larger than the size of the intersections of the stable and unstable manifolds. This is known as the large gap problem.

In the last two decades there have been several methods to overcome the large gap problem. Among them there are mainly two methods concerning the genericity of instability: variational methods and geometric methods. The first attempt to study Arnold’s original example in variational viewpoint is by Bessi [10]. Essential progress has also been made by Mather. In the celebrated paper [67], he developed a powerful variational tool to study the global instability in the framework of convex Lagrangian systems. In an unpublished manuscript [68], he further showed the existence of orbits with unbounded energy in perturbations of a geodesic flow on $\mathbb{T}^2$ by a generic time-periodic potential. Based on Mather’s variational mechanism, the authors of [24] constructed diffusing orbits and proved the $C^r$-genericity ($r$ is finite and suitably large) of Arnold diffusion for the a priori unstable systems with two and a half degrees of freedom. On the other hand, several authors have used geometric methods, which also apply to Hamiltonians that are not necessarily convex, to obtain Arnold diffusion. More precisely, the authors in [32–34] defined the so-called scattering map which accounts for the outer dynamics along homoclinic orbits, and overcame the large gap problem by incorporating in the transition chain new invariant objects, like secondary tori and the stable and unstable manifolds of lower dimensional tori; in [73], the author geometrically defined the so-called separatrix map near the NHIC, then he showed in [74] the existence of diffusion by making full use of the dynamics of this map, and even estimated the optimal diffusion speed of order $\varepsilon/|\log \varepsilon|$ (see also [8]). Moreover, for the case of a priori unstable Hamiltonians with higher degrees of freedom, similar results have also been obtained by variational or geometric methods in [3, 25, 36, 37, 46, 59, 75].

The a priori stable case poses a new difficulty: the presence of multiple resonances. In the paper [69] (see also [70]), Mather first made an announcement for systems with two degrees
of freedom in the time-periodic case or with three degrees of freedom in the autonomous case, under a series of cusp-residual conditions. So the diffusion problem in this situation was thought to possess only cusp-residual genericity. The complete proof for the autonomous systems with three degrees of freedom appeared in the preprint [19], and the main ingredients have been published in the recent works [20–22, 26]. Indeed, the main difficulty in this case arises from the dynamics around strong double resonances. It is because away from double resonances, one could apply normal form theory to construct NHICs with a length independent of \( \varepsilon \), along which the local instability can be obtained as in the \textit{a priori} unstable case [4, 7]. To solve the problem of double resonance, the paper [20] presented a new variational mechanism to switch from one resonance to another, which eventually proved the cusp-residual genericity of diffusion in the \( C' \) smooth category [21]. Moreover, we mention that similar results on diffusion have also been obtained, by using variational methods, in the paper [56] and the preprint [53] for systems with 2.5 degrees of freedom. Also, we refer the reader to the preprints [47, 62, 63] for systems with 3 degrees of freedom by using the geometric tools. As for the case of arbitrarily higher degrees of freedom, we refer the reader to the preprint [23] and the announcement [54]. Anyway, there have been many other works related to the problem of Arnold diffusion but we cannot list all of them, see [11, 17, 31, 43, 48, 49, 55, 78], etc.

To the author’s knowledge, the genericity of Arnold diffusion is by now quite well understood in the \( C' \) smooth category, not yet for the analytic category, or the Gevrey smooth category [45]. The present paper is interested in whether the phenomenon of large evolution is generic in the Gevrey smooth Hamiltonians. Given \( \alpha \geq 1 \), a Gevrey-\( \alpha \) function is an ultra-differentiable function whose \( k \)th order partial derivatives are bounded by \( O(M^{-|k|}|k|!^{\alpha}) \). For the case \( \alpha = 1 \), it is exactly a real analytic function. So the Gevrey class is intermediate between the \( C^\infty \) class and the real analytic class. Besides, a key point for the Gevrey class is that it allows the existence of a function with compact support (i.e. bump function). But no analytic function has compact support.

To consider the Arnold diffusion problem in the Gevrey topology, we would adopt the Gevrey norm introduced by Marco and Sauzin in [64] during a collaboration with Herman (see definition 1.1). Apart from the theory of PDE where it has been widely used, the Gevrey class is also studied in the field of dynamical systems. For example, we refer to [12–15, 57, 72], etc for the stability theory, such as KAM theory and Nekhoroshev theory. We also refer the reader to [16, 39, 58, 65, 77], etc for some relevant results on instability. All these studies make us believe that one can also consider the genericity problem of diffusion in the Gevrey case.

So in this paper, we start by considering the \textit{a priori} unstable, Gevrey-\( \alpha \) (\( \alpha > 1 \)) Hamiltonian systems of two and a half degrees of freedom. The case \( \alpha = 1 \) (i.e. the analytic genericity) is more complicated and has not been fully studied. Here we only mention a recent work [46] which proposes a general geometric mechanism that might be useful for analytic genericity. In the same spirit as in [46], the paper [44] gives models where the analytic genericity can be achieved for \textit{a priori} chaotic symplectic maps, provided that the scattering map has no monodromy and is globally defined on the NHIC.

Before stating our main results, we review the concept of Gevrey function and some standard facts.

\textbf{Definition 1.1 (Gevrey function [64]).} Let \( \alpha \geq 1 \), \( L > 0 \) and \( K \) be an \( n \)-dimensional compact domain. A real-valued \( C^\infty \) function \( f(x) \) defined on \( K \) is said to be Gevrey-(\( \alpha \), \( L \)) if

\[
\| f \|_{\alpha, L} := \sum_{k \in \mathbb{N}^n} \frac{L^{|k|\alpha}}{(k!)!!} \| \partial^k f \|_{C^\infty(K)} < +\infty,
\]
with the standard multi-index notation \( k = (k_1, \ldots, k_n) \in \mathbb{N}^n \), \( |k| = k_1 + \cdots + k_n \), \( k! = k_1! \cdots k_n! \) and \( \partial^k = \partial_{k_1} \cdots \partial_{k_n} \).

Let \( G^{\alpha,L}(K) := \{ f \in C^\infty(K) : \| f \|_{\alpha,L} < +\infty \} \). The space \( G^{\alpha,L}(K) \) endowed with the norm \( \| \cdot \|_{\alpha,L} \) is a Banach space. Sometimes we also write \( G^{\alpha}(K) := \bigcup_{L>0} G^{\alpha,L}(K) \). In particular, for \( K \subset \mathbb{R}^d \) and \( \alpha = 1 \), \( G^1(K) \) is exactly the space of real analytic functions on \( K \); any function \( f \in G^1(K) \) is real analytic in \( K \) and admits an analytic extension in the complex domain \( \{ z \in \mathbb{C} : \text{dist}(z, K) < L \} \). Conversely, for any real analytic function \( f \) in \( K \), there exists \( L > 0 \) such that \( f \in G^1(K) \). However, for \( \alpha > 1 \), \( G^{\alpha,L}(K) \) admits non-analytic functions. Therefore, the Gevrey-smooth category is intermediate between the \( C^\infty \) category and the analytic category.

Gevrey class has the following useful properties which have been already proved in [64]:

(G1) The norm \( \| \cdot \|_{\alpha,L} \) is an algebra norm, namely \( \| fg \|_{\alpha,L} \leq \| f \|_{\alpha,L} \| g \|_{\alpha,L} \).

(G2) Suppose \( 0 < \lambda < L \) and \( f \in G^{\alpha,L}(K) \), then all partial derivatives of \( f \) belong to \( G^{\alpha-L}(K) \) and
\[
\| \partial^k f \|_{\alpha-L-\lambda} \leq L^\lambda \lambda^{|k|} \| f \|_{\alpha,L}.
\]

(G3) Let \( f \in G^{\alpha,L}(K_m) \) where \( K_m \) is an \( m \)-dimensional domain and let \( g = (g_1, \ldots, g_m) \) be a mapping whose component \( g_i \in G^{\alpha,L}(K_m) \). If \( g(K_m) \subset K_m \) and \( \| g \|_{\alpha,L} - \| g \|_{\alpha,L} \leq L^\alpha n^{-\alpha} \) for all \( 1 \leq i \leq m \), then \( f \circ g \in G^{\alpha,L}(K_m) \) and \( \| f \circ g \|_{\alpha,L} \leq \| f \|_{\alpha,L} \).

1.1. Setup and main result

The current paper will mainly focus on the convex Hamiltonians of two and a half degrees of freedom. As we will see later, all discussions will be restricted on a compact domain in \( \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T} \), so we fix, once and for all, a constant \( R > 1 \) and a compact set
\[
\mathcal{D}_R = \mathbb{T}^2 \times \tilde{B}_R(0) \times \mathbb{T},
\]
where \( B_R(0) \subset \mathbb{R}^d \) is an open ball of radius \( R \) centered at 0 and \( \tilde{B}_R(0) \) is the closure. By definition 1.1, the space \( G^{\alpha,L}(\mathcal{D}_R) \) consists of all real-valued smooth functions \( f(q,p,t) \) satisfying
\[
\| f \|_{\alpha,L} = \sum_{k \in \mathbb{N}^2} \frac{L^{k \alpha}}{k!} \| \partial^k f \|_{C^0(\mathcal{D}_R)} < +\infty.
\] (1.1)

Let \( C^\alpha(\mathcal{D}_R) \) be the space of all real-valued analytic functions on \( \mathcal{D}_R \), admitting an analytic extension in the complex domain \( \{(q,p,t) \in (\mathbb{C}/\mathbb{Z})^2 \times \mathbb{C}^2 \times (\mathbb{C}/\mathbb{Z}) : |\text{Im} q|_\infty < d, \ \text{dist}(p, \mathbb{B}_R(0)) < d, |\text{Im} t| < d \} \). Set \( C^\alpha(\mathcal{D}_R) = \bigcup_{\alpha \geq 0} C^\alpha(\mathcal{D}_R) \), it is well known that

(a) For \( \alpha > 1, L > 0 \) and any \( d > L^\alpha \),
\[
C^\alpha(\mathcal{D}_R) \subset G^{\alpha,L}(\mathcal{D}_R) \subset C^\infty(\mathcal{D}_R), \quad C^\alpha(\mathcal{D}_R) \subset G^\alpha(\mathcal{D}_R) \subset C^\infty(\mathcal{D}_R).
\]

(b) \( C^\alpha(\mathcal{D}_R) = G^1(\mathcal{D}_R) \).

Now, we introduce the \textit{a priori} unstable Hamiltonian model considered in this paper and state the main assumptions. Let \( q = (q_1, q_2) \in \mathbb{T}^2 \) and \( p = (p_1, p_2) \in \mathbb{R}^2 \). We consider a time-periodic and \( C^r(r > 2) \) smooth Hamiltonian of the form:
\[
H(q, p, t) = H_0(q, p) + H_1(q, p, t), \quad \text{where} \quad H_0(q, p) = h_1(p_1) + h_2(q_2, p_2).
\] (1.2)
Here, the term $H_1$ is a small perturbation which is periodic of period $1$ in $t$. Our main assumptions on $H_0$ are the following:

(H1) **Convexity and superlinearity:** for each $q \in \mathbb{T}^2$, the Hessian $\partial_{pp}H_0(q, p)$ is positive definite, and $\lim_{|p|\to +\infty}H_0(q, p)/\|p\| = +\infty$.

(H2) **A priori hyperbolicity:** the Hamiltonian flow $\Phi_{h_2}$, determined by $h_2$, has a hyperbolic fixed point $(q_2, p_2) = (x^*, y^*)$. Moreover, the function $h_2(q_2, y^*): \mathbb{T} \to \mathbb{R}$ attains its unique maximum at $q_2 = x^*$. Without loss of generality, we can assume $(x^*, y^*) = (0, 0)$.

A prototype example of such a system is the coupling of a rotator and a pendulum

$$H = \frac{p_1^2}{2} + \frac{p_2^2}{2} + (\cos 2\pi q_2 - 1) + H_1(q, p, t),$$

which has been considered many times in the literature. Keep this example in mind will help the reader better understand our result and method. As we will see later, the above assumptions (H1)–(H2) are in the same spirit as in [24] while our main result and approach have some differences.

Let $B_{\varepsilon}^{L,R} = \{H_1 \in C^\infty(\mathcal{D}_R) : \|H_1\|_{0,L} < \varepsilon\} \subset G^{1,L}(\mathcal{D}_R)$ denote the open ball of radius $\varepsilon$ centered at the origin with respect to the norm $\| \cdot \|_{0,L}$.

**Theorem 1.2.** Let $\alpha > 1, R > 1$ and assume that $H_0$ in (1.2) is of class $C^r(r > 2)$, then there exists a positive constant $L_0 = L_0(H_0, \alpha, R)$ such that, for each $L \in (0, L_0]$ and a sequence of open balls $B_1(y_1), \ldots, B_k(y_k) \subset \mathbb{R}^2$, of radius $s$ centered at $y_t \in [-R + 1, R - 1] \times \{0\} \subset \mathbb{R}^2$, $\ell = 1, \ldots, k$, we have:

there exist a positive number $\varepsilon_0 = \varepsilon_0(H_0, \alpha, R, s, L)$ and an open and dense subset $\mathcal{D}_{\varepsilon_0,R} \subset B_{\varepsilon_0,R}$ such that for each perturbation $H_1 \in \mathcal{D}_{\varepsilon_0,R}$, the system $H = H_0 + H_1$ has a trajectory $(q(t), p(t))$ whose action variables $p(t)$ pass through the ball $B_1(y_t)$ at the time $t = t_1$, where $t_1 < t_2 < \cdots < t_k$.

**Remark.** Just as Mather did in [69, 70], the smoothness of the unperturbed Hamiltonian $H_0$ could differ from that of the perturbation term $H_1$. Notice that $L_0$ can not be an arbitrary constant, the reason is that our approach needs to adopt the Gevrey approximation (see theorem 5.3).

**Remark (Autonomous case).** Recall that Mather’s cohomology equivalence is trivial for an autonomous system (cf [2]). The problem is that, unlike the time-periodic case, there is no canonical global transverse section of the flow in an autonomous system. In [59], this difficulty was overcome by taking local transverse sections, which could generalize Mather’s cohomology equivalence. Thus we believe that the Gevrey genericity is still valid for the a priori unstable autonomous Hamiltonians. However, in this paper, we only consider the non-autonomous case.

The perturbation technique used in the current paper can also prove the genericity in the sense of Mañé, which means that the diffusion is still a typical phenomenon when $H_0$ is perturbed by potential functions. More precisely, let $B_{\varepsilon}^{L} \subset G^{1,L}(\mathbb{T}^2 \times \mathbb{T})$ denote the open ball of radius $\varepsilon$ centered at the origin with respect to the norm $\| \cdot \|_{0,L}$, we have

**Theorem 1.3.** Under the same assumptions as in theorem 1.2, there exists $L_0 = L_0(H_0, \alpha, R) > 0$ such that, for each $L \in (0, L_0]$ and a sequence of open balls $B_1(y_1), \ldots, B_k(y_k) \subset \mathbb{R}^2$, of radius $s$ centered at $y_t \in [-R + 1, R - 1] \times \{0\} \subset \mathbb{R}^2$, $\ell = 1, \ldots, k$, we have:
there exist a positive number \( \varepsilon_0 = \varepsilon_0(H_0, \alpha, R, s, L) \) and an open and dense subset \( S_{t_0}^L \subset B_{t_0}^L \) such that for each potential perturbation \( H_1 \in S_{t_0}^L \), the system \( H = H_0 + H_1 \) has a trajectory \((q(t), p(t))\) whose action variables \( p(t) \) pass through the ball \( B_{s(y)} \) at the time \( t = t_0 \), where \( t_1 < t_2 < \cdots < t_k \).

### 1.2. Outline of this paper

This paper mainly adopts variational methods to construct diffusing orbits, so it requires us to transform into Lagrangian formalism. We still denote by \( \tilde{\mathcal{L}} \) an introduction new Lagrangian as a result of hypothesis (H1). Now, we turn to the small perturbation term \( L_1 \). As we will see later, only the information on a compact region outside of that region in our proofs, then it will not affect the study of Arnold diffusion if one modifies the perturbation function \( L_1 \) outside the compact set. For example, one can introduce a new function \( \tilde{L}_1 \) which has compact support, and is identical to \( L_1 \) on a compact set \( \{ \|v\|_q \leq K \} \). In terms of this modification, we then introduce a new Lagrangian \( \tilde{L} := L_0 + L_1 \). Observe that the modified Lagrangian \( \tilde{L} \) satisfies the Tonelli conditions since the perturbation term \( L_1 \) is small enough and has compact support. Also, it is quite clear that \( \tilde{L} \) and \( L \) generate the same Euler–Lagrange flow when restricted on the compact region \( \{ \|v\|_q \leq K \} \). Such a modification is elementary, see for instance [69].

Therefore, in what follows, we can always assume, without loss of generality, that our Lagrangian (1.3) satisfies the Tonelli conditions. Then, through the Legendre transformation

\[
\mathcal{L} : \mathbb{T}^2 \times \mathbb{T} \to \mathbb{T}^2 \times \mathbb{T},
\]

\[ (q, p, t) \mapsto (q, \frac{\partial H}{\partial p}(q, p, t), t), \tag{1.4} \]

we can write

\[ L(q, v, t) = \langle \pi_p \circ \mathcal{L}^{-1}(q, v, t), v \rangle - H \circ \mathcal{L}^{-1}(q, v, t), \]

where \( \pi_p \) denotes the projection \((q, p, t) \mapsto p \). Then, the Hamilton's equations \( \dot{q} = \frac{\partial H}{\partial p}, \dot{p} = -\frac{\partial H}{\partial q} \) is equivalent to the Euler–Lagrange equation

\[ \frac{d}{dt} \left( \frac{\partial L}{\partial v} \right) - \frac{\partial L}{\partial q} = 0. \]
Throughout this paper, we use $\phi^t_L$ to denote the Euler–Lagrange flow determined by $L$ and $\Phi^t_H$ to denote the Hamiltonian flow determined by $H$.

The Fenchel inequality and hypothesis (H2) together give rise to

$$h_2(q_2, 0) + l_2(q_2, v_2) \geq 0, \quad h_2(0, 0) + l_2(0, 0) = 0, \quad (q_2, v_2) \in T^T.$$ 

Since $q_2 = 0 \pmod{1}$ is the unique maximum point of the function $h_2(\cdot, 0) : T \to \mathbb{R}$, one gets

$$l_2(0, 0) = -h_2(0, 0) \leq -h_2(q_2, 0) \leq l_2(q_2, v_2), \quad (q_2, v_2) \in T^T. \quad (1.5)$$

Then the point $(q_2, v_2) = (0, 0)$ is the unique minimum point of the function $l_2$ as a consequence of the strict convexity. Also, $(0, 0)$ is a hyperbolic fixed point for the Euler–Lagrange flow $\phi^t_L$.

Compared with the variational proofs of $C^r$-genericity in [24, 25], the method in this paper contains some new techniques. Indeed, the strategy used in [24, 25], which perturbs the generating functions to create genericity, seems not applicable to the Gevrey genericity. The main difficulty arises from the fact that, when we estimate the Gevrey smoothness of a Hamiltonian flow, we cannot avoid the decrease of Gevrey coefficient $L$ during the switch from a generating function to its corresponding Hamiltonian, or the switch from a Lagrangian to its associated Hamiltonian (see property (G2) above). Thus in this paper, inspired by the ideas in [20], we decide to directly perturb a Hamiltonian by potential functions, one advantage of this approach is that the Lagrangian associated to the perturbed Hamiltonian $H + V(q, t)$ is exactly $L - V(q, t)$. To this end, some quantitative estimations are required, such as the Gevrey approximation and the corresponding inverse function theorem. It is also worth mentioning that one can establish the genericity not only in the usual sense but also in the sense of Mañé. Besides, we also believe that our results could be obtained by geometric tools, such as the scattering maps developed in [30, 34–36], or the separatrix maps in [37, 74, 75].

In our variational proof of genericity, the modulus of continuity of barrier functions is crucial. To implement this argument, the work [24] introduced the following parameterization technique: fixing an invariant curve $\Gamma_0$ on the NHIC, for any other invariant curve $\Gamma_\sigma$ on the NHIC, we parameterize it by $\sigma$, the area between the two curves. Then it can be shown that $\Gamma_\sigma$ is Hölder continuous with respect to $\sigma$ in the $C^0$ topology. However, by taking advantage of the tools in weak KAM theory, now we can show that this ‘area’ parameter $\sigma$ is exactly the cohomology class (see section 6.2). This will help us simplify the proof.

The structure of this paper is as follows. In section 2, we review some standard results, related to Arnold diffusion problem, in Mather theory. Section 3 discusses the elementary weak KAM solutions, and a special ‘barrier function’ whose minimal points correspond to heteroclinic orbits. In section 4, we introduce the concept of generalized transition chain and then give the variational mechanism of constructing diffusing orbits along this chain. In section 5, we present some properties of Gevrey functions which are necessary for our proofs. Section 6 is the main part of this paper, and applies the tools exposed in previous sections to study the Gevrey smooth systems. First, we generalize the genericity of uniquely minimal measure in the Gevrey or analytic topology. Second, we obtain certain regularity of the elementary weak KAM solutions and show how to choose suitable Gevrey space. Finally, by proving total disconnectedness for the minimal sets of barrier functions, we establish the genericity of generalized transition chain along which the global instability occurs, this therefore completes the proofs of theorems 1.2 and 1.3.

We would like to thank the anonymous referees for their insightful comments and valuable suggestions on improving our results.
2. Preliminaries: Mather theory

In this section, we recall some standard results in Mather theory which are necessary for the purpose of our study, the main references are Mather’s original papers [66, 67]. Let $M$ be a connected and compact smooth manifold without boundary, equipped with a smooth Riemannian metric $g$. Let $TM$ denote the tangent bundle, a point of $TM$ will be denoted by $(q, v)$ with $q \in M$ and $v \in T_qM$. We shall denote by $\|\cdot\|_q$ the norm induced by $g$ on the fiber $T_qM$. A time-periodic $C^2$ function $L = L(q, v, t) : TM \times \mathbb{T} \to \mathbb{R}$ is called a Tonelli Lagrangian if it satisfies:

(a) **Convexity:** $L$ is strictly convex in each fiber, i.e., the second partial derivative $\partial^2 L / \partial v^2(q, v, t)$ is positive definite, as a quadratic form, for each $(q, t) \in M \times \mathbb{T}$;

(b) **Superlinear growth:** $L$ is superlinear in each fiber, i.e. for each $(q, t) \in M \times \mathbb{T}$,

$$\lim_{\|v\|_q \to +\infty} \frac{L(q, v, t)}{\|v\|_q} = +\infty.$$ 

(c) **Completeness:** all solutions of the Euler–Lagrange equation are well defined for all $t \in \mathbb{R}$.

Let $I = [a, b]$ be an interval and $\gamma : I \to M$ be any absolutely continuous curve. Given a cohomology class $c \in H^1(M, \mathbb{R})$, we choose and fix a closed 1-form $\eta_c$ with $[\eta_c] = c$. Denote by

$$A_c(\gamma) := \int_a^b L(d\gamma(t), t) - \eta_c(d\gamma(t)) \, dt$$

the action of $L - \eta_c$ along $\gamma$, where $d\gamma(t) = (\gamma(t), \dot{\gamma}(t))$. A curve $\gamma : I \to M$ is called $c$-minimal if

$$A_c(\gamma) = \min_{\xi \in \mathcal{C}(I, M)} \int_a^b L(d\xi(t), t) - \eta_c(d\xi(t)) \, dt,$$

where $\mathcal{C}(I, M)$ denotes the set of absolutely continuous curves. As is known to all, each minimal curve satisfies the Euler–Lagrange equation. A curve $\gamma : \mathbb{R} \to M$ is called globally $c$-minimal if for any $a < b$, the curve $\gamma : [a, b] \to M$ is $c$-minimal. Therefore, we introduce the **globally minimal set**

$$\tilde{G}(c) := \bigcup_{\gamma} \{(d\gamma(t), t) : \gamma : \mathbb{R} \to M \text{ is } c - \text{minimal} \}.$$

Let $\phi^t_L$ be the Euler–Lagrange flow on $TM \times \mathbb{T}$, and $\mathcal{M}$ be the space of all $\phi^t_L$-invariant probability measures on $TM \times \mathbb{T}$. To each $\mu \in \mathcal{M}$, Mather has proved that $\int_{TM \times \mathbb{T}} \lambda \, d\mu = 0$ holds for any exact 1-form $\lambda$, which yields that $\int_{TM \times \mathbb{T}} L - \eta_c \, d\mu = \int_{TM \times \mathbb{T}} L - \eta'_c \, d\mu$ if $\eta_c - \eta'_c$ is exact. This leads us to define Mather’s $\alpha$ function,

$$\alpha(c) := \inf_{\mu \in \mathcal{M}} \int_{TM \times \mathbb{T}} L - \eta_c \, d\mu.$$ 

To some extent, the value $\alpha(c)$ is a minimal average action for $L - \eta_c$. Mather has proved that $\alpha : H^1(M, \mathbb{R}) \to \mathbb{R}$ is finite everywhere, convex and superlinear.
For each $\mu \in \mathcal{M}$, the rotation vector $\rho(\mu)$ associated with $\mu$ is the unique element in $H_1(M, \mathbb{R})$ that satisfies

$$\langle \rho(\mu), [\eta_c] \rangle = \int_{TM \times T} \eta_c \, d\mu,$$

for all closed 1-form $\eta_c$, here $\langle \cdot , \cdot \rangle$ denotes the dual pairing between homology and cohomology classes. Then, we can define Mather’s $\beta$ function as follows:

$$\beta(h) := \inf_{\mu \in \mathcal{M}, \rho(\mu) = h} \int_{TM \times T} L \, d\mu.$$ 

This function $\beta : H_1(M, \mathbb{R}) \to \mathbb{R}$ is also finite everywhere, convex and superlinear. In fact, $\beta$ is the Legendre–Fenchel dual of the function $\alpha$, i.e. $\beta(h) = \max_c \{ \langle h, c \rangle - \alpha(c) \}$.

We define

$$\mathcal{M}_1(L) := \left\{ \mu : \int_{TM \times T} L - \eta_c \, d\mu = -\alpha(c) \right\},$$

$$\mathcal{M}_0(L) := \left\{ \mu : \rho(\mu) = h, \int_{TM \times T} L \, d\mu = \beta(h) \right\}.$$ 

By duality, it can be easily checked that

$$\mathcal{M}_c(L) = \bigcup_{h \in \partial \alpha(c)} \mathcal{M}_0(L),$$

where $\partial \alpha(c)$ is the sub-differential. For one and a half degrees of freedom systems including twist maps, Mather’s $\alpha$ function is of class $C^1$, then

$$\mathcal{M}_c(L) = \mathcal{M}_0(L), \quad d\alpha(c) = h. \quad (2.1)$$

We call each element $\mu \in \mathcal{M}_c(L)$ a c-minimal measure. The Mather set of cohomology class $c$ is then defined by

$$\mathcal{M}(c) := \bigcup_{\mu \in \mathcal{M}_c(L)} \text{supp} \mu.$$ 

To study more dynamical properties, we need to find some ‘larger’ minimal invariant sets and discuss their topological structures. Let $t' > t$, the action function $h_{c}^{t'} : M \times M \to \mathbb{R}$ is defined by

$$h_{c}^{t'}(x, x') := \min_{\gamma(t) = x, \gamma(t') = x'} \int_{t}^{t'} (L - \eta_c)(d\gamma(s), s) \, ds + \alpha(c) \cdot (t' - t).$$

Then we define a real-valued function $\Phi_{c} : (M \times \mathbb{T}) \times (M \times \mathbb{T}) \to \mathbb{R}$ by

$$\Phi_{c}((x, \tau), (x', \tau')) := \inf_{t' > t, \tau' \equiv \tau \text{ mod } 1, t' \equiv t' \text{ mod } 1} h_{c}^{t'}(x, x'),$$

and a real-valued function $h_{c}^{\infty} : (M \times \mathbb{T}) \times (M \times \mathbb{T}) \to \mathbb{R}$ by

$$h_{c}^{\infty}((x, \tau), (x', \tau')) = \lim_{t' \equiv t' \text{ mod } 1, t' \to +\infty} h_{c}^{t'}(x, x'). \quad (2.2)$$
Proposition 2.1. In the literature, $h^\infty_c$ and $\Phi_c$ are called the Peierls barrier function and Mañé’s potential respectively.

A minimal curve $\gamma : \mathbb{R} \to M$ is called $c$-semi static if for any $t < t'$,

$$A_c(\gamma|_{(t,t')}) + \alpha(c) \cdot (t' - t) = \Phi_c \left( (\gamma(t), t \mod 1), (\gamma(t'), t' \mod 1) \right).$$  \hfill (2.3)

A minimal curve $\gamma : \mathbb{R} \to M$ is called $c$-static if for any $t < t'$,

$$A_c(\gamma|_{(t,t')}) + \alpha(c) \cdot (t' - t) = -\Phi_c \left( (\gamma(t'), t' \mod 1), (\gamma(t), t \mod 1) \right).$$ \hfill (2.4)

This gives the so-called Aubry set $\tilde{A}(c)$ and Mañé set $\tilde{N}(c)$ in $TM \times T$:

$$\tilde{A}(c) = \bigcup \{(d\gamma(t), t \mod 1) : \gamma \text{ is } c \text{ - static} \},$$

$$\tilde{N}(c) = \bigcup \{(d\gamma(t), t \mod 1) : \gamma \text{ is } c \text{ - semi static} \}.$$  

The $\alpha$-limit and $\omega$-limit sets of a $c$-minimal curve $(d\gamma(t), t)$ belong to $\tilde{A}(c)$, see for instance [2]. In addition, with the canonical projection $\pi : TM \times T \to M \times T$, one could define the projected Aubry set $\tilde{A}(c) = \pi \tilde{A}(c)$, the projected Mather set $\tilde{M}(c) = \pi \tilde{M}(c)$, the projected Mañé set $\tilde{N}(c) = \pi \tilde{N}(c)$ and the projected globally minimal set $\tilde{G}(c) = \pi \tilde{G}(c)$. Then the following inclusion relations hold (see [2]):

$$\tilde{M}(c) \subset \tilde{A}(c) \subset \tilde{N}(c) \subset \tilde{G}(c), \quad \tilde{M}(c) \subset \tilde{A}(c) \subset \tilde{N}(c) \subset \tilde{G}(c).$$

Next, we present some key properties of the minimal sets above, which will be fully exploited in the construction of diffusing orbits. Property (a) below is a classical result which has been proved by Mather in [66], and the proof of property (b) could be found in [2, 24].

**Proposition 2.1.** For the Tonelli Lagrangian $L$, we have:

(a) (Graph property) let $\pi : TM \times T \to M \times T$ be the canonical projection. Then the restriction of $\pi$ to $\tilde{A}(c)$ is a bi-Lipschitz homeomorphism.

(b) (Upper semi-continuity) the set-valued map $(c,L) \mapsto \tilde{G}(c,L)$ and the set-valued map $(c,L) \mapsto \tilde{N}(c,L)$ are both upper semi-continuous.

For $(x, \tau), (x', \tau') \in M \times T$, we set

$$d_c((x, \tau),(x', \tau')) := h^\infty_c((x, \tau),(x', \tau')) + h^\infty_c((x', \tau'),(x, \tau)).$$

By definition (2.4), it follows that

$$h^\infty_c((x, \tau),(x, \tau)) = 0 \iff (x, \tau) \in \tilde{A}(c),$$

hence $d_c$ is a pseudo-metric on the projected Aubry set $\tilde{A}(c)$. Two points $(x, \tau), (x', \tau') \in \tilde{A}(c)$ are said to be in the same Aubry class if $d_c((x, \tau),(x', \tau')) = 0$. Clearly, each Aubry class is a closed set. If only one $c$-minimal measure exists, then the Aubry class is unique and

$$\tilde{A}(c) = \tilde{N}(c).$$

To characterize the Mañé set from another point of view, we define the following function

$$B'_c(x, \tau) := \min_{(\ell_i, \tau_i) \in \tilde{A}(c)} \left\{ h^\infty_c((x, \tau_1),(x, \tau)) + h^\infty_c((x, \tau),(x_2, \tau_2)) - h^\infty_c((x_1, \tau_1),(x_2, \tau_2)) \right\}. $$
Mather has proved in [69] that \( \min B^*_c = 0 \), and the set of all minimal points is exactly \( N(c) \), i.e.

\[
B^*_c(x, \tau) = 0 \iff (x, \tau) \in N(c).
\]

(2.5)

To prove theorem 6.1 in section 6, it is convenient to adopt the equivalent definition of minimal measures originating from Mañé [61]. In his setting, the minimal measures are obtained through a variational principle not requiring the invariance \textit{a priori}. Let \( C \) be the set of all continuous functions \( f : TM \times T \to \mathbb{R} \) having linear growth at most, i.e.

\[
\|f\|_l := \sup_{(q, v, t)} \frac{|f(q, v, t)|}{1 + \|v\|_q} < +\infty,
\]

and endow \( C \) with the norm \( \|\cdot\|_l \). Let \( C^* \) be the vector space of all continuous linear functionals \( \nu : C \to \mathbb{R} \) provided with the weak-\(*\) topology, namely,

\[
\lim_{k \to +\infty} \nu_k = \nu \iff \lim_{k \to +\infty} \int_{TM \times T} f \, d\nu_k = \int_{TM \times T} f \, d\nu, \quad \forall f \in C.
\]

For each \( N \in \mathbb{Z}^+ \) and each \( N \)-periodic absolutely continuous curve \( \gamma : \mathbb{R} \to M \), one can define a probability measure \( \mu_\gamma \) associated to \( \gamma \) as follows:

\[
\int_{TM \times T} f \, d\mu_\gamma := \frac{1}{N} \int_0^N f(d\gamma(t), t) \, dt, \quad \forall f \in C.
\]

(2.6)

Let

\[
\Gamma := \bigcup_{N \in \mathbb{Z}^+} \{ \mu_\gamma : \gamma \in C^{ac}(\mathbb{R}, M) \text{ is } N \text{-periodic} \} \subset C^*.
\]

and let \( \mathcal{H} \) be the closure of \( \Gamma \) in \( C^* \). It is easily seen that the set \( \mathcal{H} \) is convex.

\( \mu_\gamma \) in \( \Gamma \) has a naturally associated homology class \( \rho(\mu_\gamma) = \frac{\gamma}{N} \in H_1(M, \mathbb{R}) \), where \( [\gamma] \) denotes the homology class of \( \gamma \). The map \( \rho : \Gamma \to H_1(M, \mathbb{R}) \) can extend continuously to \( \rho : \mathcal{H} \to H_1(M, \mathbb{R}) \) which is surjective. Then, Mañé introduced the following minimal measures:

\[
\mathcal{M}^c(L) := \left\{ \mu \in \mathcal{H} \mid \int L - \eta_k \, d\mu = \min_{\nu \in \mathcal{H}} \int L - \eta_k \, d\nu \right\},
\]

\[
\mathcal{M}_h(L) := \left\{ \mu \in \mathcal{H} \mid \rho(\mu) = h, \int L \, d\mu = \min_{\nu \in \mathcal{H}, \rho(\nu) = h} \int L \, d\nu \right\}.
\]

(2.7)

We end this section by the following equivalence property:

**Proposition 2.2 ([61]).** The sets \( \mathcal{M}^c(L) = \mathcal{M}^c(L) \) and \( \mathcal{M}_h(L) = \mathcal{M}_h(L) \).

3. Elementary weak KAM solutions and heteroclinic orbits

3.1. Weak KAM solutions

Weak KAM solution is the basic element in weak KAM theory which builds a link between Mather theory and the theory of viscosity solutions of Hamilton–Jacobi equations. Here we only recall some basic concepts and properties which help us better understand Mather theory.
For more details, we refer the reader to Fathi’s book [38] for time-independent systems, and to [3, 28, 76] for time-periodic systems.

**Definition 3.1.** A continuous function \( u^-_c : M \times \mathbb{T} \rightarrow \mathbb{R} \) is called a **backward weak KAM solution** if

(a) For any absolutely continuous curve \( \gamma : [a, b] \rightarrow M \),

\[
    u^-_c(\gamma(b), b) - u^-_c(\gamma(a), a) \leq \int_a^b (L - \eta_c)(d\gamma(s), s) + \alpha(c) \, ds.
\]

(b) For each \((x, t) \in M \times \mathbb{R}\), there exists a **backward calibrated curve** \( \gamma^- : (-\infty, t] \rightarrow M \) with \( \gamma^-(t) = x \) such that for all \( a < b \leq t \),

\[
    u^-_c(\gamma^-(b), b) - u^-_c(\gamma^-(a), a) = \int_a^b (L - \eta_c)(d\gamma^-(s), s) + \alpha(c) \, ds.
\]

Similarly, a continuous function \( u^+_c : M \times \mathbb{T} \rightarrow \mathbb{R} \) is called a **forward weak KAM solution** if

(a) For any absolutely continuous curve \( \gamma : [a, b] \rightarrow M \),

\[
    u^+_c(\gamma(b), b) - u^+_c(\gamma(a), a) \leq \int_a^b (L - \eta_c)(d\gamma(s), s) + \alpha(c) \, ds.
\]

(b) For each \((x, t) \in M \times \mathbb{R}\), there exists a **forward calibrated curve** \( \gamma^+ : [t, +\infty) \rightarrow M \) with \( \gamma^+(t) = x \) such that for all \( t \leq a < b \),

\[
    u^+_c(\gamma^+(b), b) - u^+_c(\gamma^+(a), a) = \int_a^b (L - \eta_c)(d\gamma^+(s), s) + \alpha(c) \, ds.
\]

For example, it is well known in weak KAM theory that, for each \((x_0, t_0) \in M \times \mathbb{T}\) the barrier function \( h^\infty_\gamma((x_0, t_0), \cdot) : M \times \mathbb{T} \rightarrow \mathbb{R} \) is a backward weak KAM solution and \(-h^\infty_\gamma(\cdot, (x_0, t_0)) : M \times \mathbb{T} \rightarrow \mathbb{R} \) is a forward weak KAM solution. If there is only one Aubry class, then \( h^\infty_\gamma((x_0, t_0), \cdot) \) is the unique backward weak KAM solution up to an additive constant, and \(-h^\infty_\gamma(\cdot, (x_0, t_0)) \) is also the unique forward weak KAM solution up to an additive constant.

It is easily seen that backward (forward) calibrated curves are semi static. The following properties for weak KAM solutions are well known and the proof can be found in [38] or [28]:

**Proposition 3.2.**

(a) \( u^-_c \) is Lipschitz continuous, and is differentiable on \( \mathcal{A}(c) \). If \( u^-_c \) is differentiable at \((x_0, t_0) \in M \times \mathbb{T}\), then

\[
    \partial_1 u^-_c(x_0, t_0) + H(x_0, c + \partial_1 u^-_c(x_0, t_0), t_0) = \alpha(c).
\]

It also determines a unique \( c \)-semi static curve \( \gamma^-_c : (-\infty, t_0] \rightarrow M \) with \( \gamma^-_c(t_0) = x_0 \), and such that \( u^-_c \) is differentiable at each point \((\gamma^-_c(t), t) \) with \( t \leq t_0 \), namely \( c + \partial_1 u^-_c(\gamma^-_c(t), t) = \frac{d}{dt}(d\gamma^-_c(t), t) \).

(b) \( u^+_c \) is Lipschitz continuous, and is differentiable on \( \mathcal{A}(c) \). If \( u^+_c \) is differentiable at \((x_0, t_0) \in M \times \mathbb{T}\), then

\[
    \partial_1 u^+_c(x_0, t_0) + H(x_0, c + \partial_1 u^+_c(x_0, t_0), t_0) = \alpha(c).
\]
It also determines a unique $c$-semi static curve $\gamma_c^+ : [t_0, +\infty) \to M$ with $\gamma_c^+(t_0) = x_0$, and such that $u_i^+$ is differentiable at each point $(\gamma_i^+(t), t)$ with $t \geq t_0$, namely $c + \partial_t u_i^+(\gamma_i^+(t), t) = \frac{\partial}{\partial v} (d\gamma_i^+(t), t)$.

3.2. Elementary weak KAM solutions

It is a generic property that a Lagrangian has finitely many Aubry classes \cite{6} for each cohomology class. Recall that the weak KAM solution is unique (up to an additive constant) if the Aubry class is unique. If two or more Aubry classes exist, there are infinitely many weak KAM solutions, among which we are only interested in the elementary weak KAM solutions. In what follows, we assume that for certain cohomology class $c$ the Aubry classes are $\{A_{ij} : i = 1, 2, \ldots, k\}$, hence the projected Aubry set $A(c) = \bigcup A_{ij}$. The concept of elementary weak KAM solution appeared in the work \cite{3}. However, for the purpose of our applications, we decide to adopt an analogous concept defined in \cite{19}.

**Definition 3.3.** We fix an $i \in \{1, \ldots, k\}$ and perturb the Lagrangian $L \to L + \varepsilon V(x, t)$ where $\varepsilon > 0$ and $V$ is a non-negative $C^\infty$ function satisfying $\text{supp} V \cap A_{ij} = \emptyset$ and $V|_{A_{ij}} > 0$ for each $j \neq i$. Then for the cohomology class $c$, the perturbed Lagrangian has only one Aubry class $A_{ij}$, and its backward weak KAM solution, denoted by $u_{ij}^-$, is unique up to an additive constant. If for a subsequence $\{u_{ij}^{\varepsilon}, \varepsilon > 0\}$, the limit

$$u_{ij}^- := \lim_{\varepsilon \to 0^+} u_{ij}^{\varepsilon}$$

exists, then we call $u_{ij}^-$ a **backward elementary weak KAM solution**. Analogously, one can define a **forward elementary weak KAM solution** $u_{ij}^+$.

In the following theorem, we will prove the existence of elementary weak KAM solutions and give explicit representation formulas as well.

**Theorem 3.4.** For each $i$, the backward (respectively forward) elementary weak KAM solution $u_{ij}^-$ (respectively $u_{ij}^+$) always exists and is unique up to an additive constant. More precisely, let $(x, \tau_i)$ be any point in $A_{ij}$, then there exists a constant $C$ (respectively $C'$) depending on $(x, \tau_i)$, such that

$$u_{ij}^-(x, \tau) = h_c^\infty((x, \tau_i), (x, \tau)) + C$$ (resp. $u_{ij}^+(x, \tau) = -h_c^\infty((x, \tau_i), (x, \tau)) + C'$).

**Proof.** We only give the proof for $u_{ij}^-$ since $u_{ij}^+$ is similar. Denote by $\alpha(\varepsilon)$ and $\alpha_c(\varepsilon)$ the value of Mather’s $\alpha$-function at the cohomology class $c$ for the Lagrangians $L - \eta_\varepsilon$ and $L - \eta_\varepsilon + \varepsilon V$ respectively, and denote by $h_c^\infty((x, \tau_i), (x, \tau))$ and $h_c^\infty((x, \tau_i), (x, \tau))$ the corresponding Peierls barrier functions. We first claim that

$$h_c^\infty((x, \tau_i), (x, \tau)) = \lim_{\varepsilon \to 0} h_c^\varepsilon((x, \tau_i), (x, \tau)).$$

Indeed, as $V \geq 0$ and its support does not intersect with $A_{ij}$, we have $\alpha_c(\varepsilon) = \alpha(\varepsilon)$ and

$$h_c^\infty((x, \tau_i), (x, \tau)) \leq \liminf_{\varepsilon \to 0} h_c^\varepsilon((x, \tau_i), (x, \tau)).$$
Now we turn to the opposite inequality \( \limsup_{k \to 0} h_{\text{Dir}}^\infty((x_i, \tau_i), (x, \tau)) \leq h_{\text{Dir}}^\infty((x_i, \tau_i), (x, \tau)) \). Assume by contradiction that there exist a subsequence \( \{h_{\text{Dir}}^\infty\}_{k \in \mathbb{N}} \) and a point \((x', \tau')\) such that

\[
\lim_{k \to \infty} h_{\text{Dir}}^\infty((x_i, \tau_i), (x', \tau')) > h_{\text{Dir}}^\infty((x_i, \tau_i), (x, \tau))
\]  

(3.3)

For abbreviation, we denote

\[
\phi_{c,i}^-(x, \tau) := h_{c,i}^\infty((x_i, \tau_i), (x, \tau)).
\]  

(3.4)

By definition 3.3, \( \mathcal{A}_{c,i} \) is the only Aubry class for the Lagrangian \( L - \eta_i + \varepsilon_i V \), which gives \( \phi_{c,i}^-(x_i, \tau_i) = 0 \). Moreover, it is not hard to verify that the sequence \( \{\phi_{c,i}^-(x_i, \tau_i)\}_k \) is uniformly Lipschitz. Hence this sequence is also uniformly bounded. So it follows from the Arzelà–Ascoli theorem that, by taking a subsequence if necessary, \( \phi_{c,i}^-(x_i, \tau_i) \) converges uniformly to a Lipschitz function \( \phi_{c,i}^- \).

It is well known in weak KAM theory that \( \phi_{c,i}^- \) is a backward weak KAM solution for \( L - \eta_i \), then

\[
\phi_{c,i}^-(x, \tau) = \phi_{c,i}^-(x, \tau) - \phi_{c,i}^-(x_i, \tau_i) \leq h_{c,i}^\infty((x_i, \tau_i), (x, \tau)),
\]

further, by letting \( k \to \infty \) on both sides of (3.4), we get

\[
\lim_{k \to \infty} h_{c,i}^\infty((x_i, \tau_i), (x, \tau)) = \lim_{k \to \infty} \phi_{c,i}^-(x, \tau) = \phi_{c,i}^-(x, \tau) \leq h_{c,i}^\infty((x_i, \tau_i), (x, \tau)),
\]

which contradicts (3.3). This therefore proves equality (3.2).

Finally, we recall that \( \mathcal{A}_{c,i} \) is the unique Aubry class for \( L - \eta_i + \varepsilon V (\varepsilon > 0) \). Then for any backward weak KAM solution \( u_{c,i} \), one has \( u_{c,i}(\cdot) = h_{c,i}^\infty((x_i, \tau_i), \cdot) + \varepsilon \) with \( \varepsilon \) a constant. Now the theorem is evident from what we have proved.

Remark. By fixing a point \((x_i, \tau_i) \in \mathcal{A}_{c,i} \) for each index \( i \in \{1, \ldots, k\} \), we conclude from theorem 3.4 that the set of all backward elementary weak KAM solutions is exactly \( \{h_{c,i}^\infty((x_i, \tau_i), \cdot) + \varepsilon \} : \varepsilon \in \mathbb{R}, i = 1, \ldots, k \} \), and the set of all forward elementary weak KAM solutions is exactly \( \{-h_{c,i}^\infty((x_i, \tau_i)) + \varepsilon \} : \varepsilon \in \mathbb{R}, i = 1, \ldots, k \} \).

3.3. Heteroclinic orbits between Aubry classes

To study the heteroclinic trajectories from a variational viewpoint, we will use a special type of barrier function. Indeed, let \( u_{c,i}^- \) and \( u_{c,i}^+ \) be a backward and a forward elementary weak KAM solution respectively. Now we define a function

\[
B_{c,i,j}(x, \tau) := u_{c,i}^-(x, \tau) - u_{c,i}^+(x, \tau), \quad \text{for each } (x, \tau) \in M \times T.
\]  

(3.5)

Roughly speaking, it measures the action along curves joining the Aubry class \( \mathcal{A}_{c,i} \) to \( \mathcal{A}_{c,j} \), we refer the reader to \([3, 19, 25]\) for more discussions.

In the sequel, the notation \( \arg \min f \) denotes the minimal set \( \{a | f(a) = \min f \} \), then we have

**Proposition 3.5.** Suppose that the projected Aubry set \( \mathcal{A}(c) = \bigcup_{i=1}^k \mathcal{A}_{c,i} \) consists of \( k(k \geq 2) \) Aubry classes, then the projected Mañé set

\[
\mathcal{N}(c) = \bigcup_{i,j=1}^k \arg \min B_{c,i,j},
\]

468
Proof. We first prove \( \mathcal{N}(c) \supseteq \arg \min B_{c,i,j} \) for each \( i, j \). Taking two points \((x_i, \tau_i) \in A_{c,i}\) and \((x_j, \tau_j) \in A_{c,j}\), theorem 3.4 implies that there exist two constants \( C_i \) and \( C_j \) such that
\[
u_{c,i}(x, \tau) = h_c^\infty((x, \tau), (x, \tau)) + C_i, \quad \nu_{c,j}(x, \tau) = -h_c^\infty((x, \tau), (x, \tau)) + C_j.
\]
So it is easy to compute that
\[
\min B_{c,i,j}(x, \tau) = h_c^\infty((x, \tau), (x, \tau)) + C_i - C_j.
\]
If \((\bar{x}, \bar{\tau}) \in \arg \min B_{c,i,j}\), then
\[
h_c^\infty((x_i, \tau_i), (\bar{x}, \bar{\tau})) + C_i + h_c^\infty((\bar{x}, \bar{\tau}), (x_j, \tau_j)) - C_j
\]
\[
= h_c^\infty((x_i, \tau_i), (x_j, \tau_j)) + C_i - C_j
\]
\[
amely h_c^\infty((x_i, \tau_i), (\bar{x}, \bar{\tau})) + h_c^\infty((\bar{x}, \bar{\tau}), (x_j, \tau_j)) - h_c^\infty((x_j, \tau_j), (x_j, \tau_j)) = 0.
\]
By (2.5) one obtains \((\bar{x}, \bar{\tau}) \in \mathcal{N}(c)\).

Now it remains to show \( \mathcal{N}(c) \subseteq \bigcup_{i=1}^k \arg \min B_{c,i,j}\). For each \((\bar{x}, \bar{\tau}) \in \mathcal{N}(c)\), one deduces from (2.5) that there always exist \( m, n \in \{1, 2, \ldots, k\} \), and two points \((x_m, \tau_m) \in A_{c,m}\), \((x_n, \tau_n) \in A_{c,n}\) such that
\[
h_c^\infty((x_m, \tau_m), (\bar{x}, \bar{\tau})) + h_c^\infty((\bar{x}, \bar{\tau}), (x_n, \tau_n)) = h_c^\infty((x_m, \tau_m), (x_n, \tau_n)).
\]
Combining with theorem 3.4, one gets that for each \((x, \tau) \in M \times \mathbb{T}\),
\[
u_{c,m}(\bar{x}, \bar{\tau}) - \nu_{c,m}(\bar{x}, \bar{\tau}) - \left(\nu_{c,m}(x, \tau) - \nu_{c,m}(x, \tau)\right)
\]
\[
= h_c^\infty((x_m, \tau_m), (\bar{x}, \bar{\tau})) + h_c^\infty((\bar{x}, \bar{\tau}), (x_n, \tau_n))
\]
\[
- \left(\nu_{c,m}(x_m, \tau_m), (x, \tau)) + h_c^\infty((x, \tau), (x, \tau))\right)
\]
\[
= h_c^\infty((x_m, \tau_m), (x, \tau)) - \left(\nu_{c,m}(x_m, \tau_m), (x, \tau)) + h_c^\infty((x, \tau), (x, \tau))\right) \leq 0,
\]
hence \((\bar{x}, \bar{\tau}) \in \arg \min B_{c,m,n}\). This completes the proof.

From now on, we denote by \( \mathcal{N}_{c}(c) \) the set of \( c \)-semi static curves which are negatively asymptotic to \( A_{c,i} \) and positively asymptotic to \( A_{c,j} \), i.e.,
\[
\mathcal{N}_{c}(c) = \left\{(x, \tau) : \exists \text{ a } c \text{-semi static curve } \gamma, \quad \gamma(\tau) = x, \right\},
\]
and
\[
\alpha(\gamma(t), t) \subset A_{c,i}, \quad \omega(\gamma(t), t) \subset A_{c,j} \right\}.
\]
Obviously, \( \mathcal{N}_{c}(c) \subseteq \mathcal{N}(c) \), and each point \((x, \tau) \in \mathcal{N}_{c}(c) \) satisfies
\[
h_c^\infty((x_i, \tau_i), (x, \tau)) + h_c^\infty((x, \tau), (x_j, \tau_j)) = h_c^\infty((x_j, \tau_j), (x_j, \tau_j)),
\]
then \( \mathcal{N}_{c}(c) \subseteq \arg \min B_{c,i,j} \) thanks to theorem 3.4. Moreover, \( A_{c,i} \cup A_{c,j} \cup \mathcal{N}_{c}(c) \subseteq \arg \min B_{c,i,j} \).

Conversely, the equality \( \arg \min B_{c,i,j} \mathcal{A}(c) = \mathcal{N}_{c}(c) \) may not hold in general. For instance, the pendulum Lagrangian \( L = \frac{\dot{\theta}^2}{2} - (\cos 8\pi x - 1) \) has four Aubry classes for the cohomology class \( c = 0 \in H^1(\mathbb{T}, \mathbb{R}) \):
\[
\tilde{A}_1 = (0, 0), \quad \tilde{A}_2 = \left(\frac{1}{4}, 0\right), \quad \tilde{A}_3 = \left(\frac{1}{2}, 0\right), \quad \tilde{A}_4 = \left(\frac{3}{4}, 0\right).
\]
They are all hyperbolic fixed points. By symmetry, it is easy to compute that \( \arg \min B_{c,1,3} = T \) but \( N_{1,3}(c) = \emptyset \).

However, in the case of only two Aubry classes, we can give a precise description.

**Proposition 3.6.** Suppose that the projected Aubry set \( A(c) = A_{c,1} \cup A_{c,2} \) has only two Aubry classes, then

\[
\arg \min B_{c,1,2} = A_{c,1} \cup A_{c,2} \cup N_{1,2}(c)
\]

and

\[
\arg \min B_{c,2,1} = A_{c,1} \cup A_{c,2} \cup N_{2,1}(c).
\]

**Proof.** We only prove \( \arg \min B_{c,1,2} \). The other case is similar. By the analysis above, it only remains for us to verify \( \arg \min B_{c,1,2} \subset A_{c,1} \cup A_{c,2} \cup N_{1,2}(c) \). Indeed, for each point \((x, \tau) \in \arg \min B_{c,1,2} \), we take \( \tau = 0 \) for simplicity, then

\[
B_{c,1,2}(x,0) = u^-_{c,1}(x,0) - u^+_{c,2}(x,0) = \min B_{c,1,2}, \tag{3.7}
\]

and proposition 3.5 implies that there exists a \( c \)-semi static curve \( \gamma: \mathbb{R} \to M \), \( \gamma(0) = x \), to be calibrated by \( u^-_{c,1} \) on \((\infty, 0] \) and be calibrated by \( u^+_{c,2} \) on \([0, +\infty) \).

Next, there exist two points \((\alpha, 0), (\omega, 0) \in A(c) \) and a sequence of positive integers \( \{m_k\}_k, \{n_k\}_k \subset \mathbb{Z}^+ \) such that

\[
\lim_{k \to \infty} \gamma(-m_k) = \alpha \quad \text{and} \quad \lim_{k \to \infty} \gamma(n_k) = \omega.
\]

By the calibration property,

\[
u^-_{c,1}(\gamma(0),0) - u^-_{c,1}(\gamma(-m_k),0) + u^+_{c,2}(\gamma(n_k),0) = \int_{-m_k}^{n_k} L(d\gamma(t), t) - \eta_c(d\gamma(t)) + \alpha(c) \, dt.
\]

Let \( \liminf k \to \infty, \)

\[
B_{c,1,2}(x,0) = u^-_{c,1}(\alpha,0) - u^+_{c,2}(\omega,0) + h^-_c((\alpha,0),(\omega,0)). \tag{3.8}
\]

On the other hand, without loss of generality (see theorem 3.4), we could assume \( u^-_{c,1}(x,0) = h^-_c((x,0),(x,0)) \) with \((x,0) \in A_{c,1} \) and \( u^+_{c,2}(x,0) = -h^-_c((x,0),(x,2,0)) \) with \((x,2) \in A_{c,2} \). Then equalities (3.7) and (3.8) together give rise to

\[
h^-_c((x,0),(x,2,0)) = h^-_c((x,0),(0,0)) + h^-_c((\omega,0),(x,2,0)) + h^-_c((\alpha,0),(\omega,0)),
\]

this could happen only if either \((\alpha,0),(\omega,0) \) belong to the same Aubry class or \((\alpha,0) \in A_{c,1}, (\omega,0) \in A_{c,2} \). This therefore completes the proof. \( \square \)

Proposition 3.6 will be fully exploited in section 6 where we extend the Lagrangian to a double covering space such that the lift of the Aubry set contains two Aubry classes.

### 4. Variational mechanism of diffusing orbits

In this section, we aim to give a master theorem which guarantees the existence of diffusion for Tonelli Lagrangian \( L: TM \times T \to \mathbb{R} \) with \( M = T^n \). Our construction of diffusion is variational,
which requires less information about the geometric structure. The orbits are constructed by shadowing a sequence of local connecting orbits, along each of them the Lagrangian action attains 'local minimum'. Basically, among them there are two types of local connecting orbits, one is based on Mather’s variational mechanism constructing orbits with respect to the cohomology equivalence [67, 68], the other one is based on Arnold’s geometric mechanism [1] whose variational version was first achieved by Bessi [10] for Arnold’s original example, and was later generalized to more general systems [3, 24, 25].

Given a cohomology class $c \in H^1(M, \mathbb{R})$, following Mather, we define

$$
\nabla_c = \bigcap_U \{ i_{U,*} H_1(U, \mathbb{R}) : U \text{ is a neighborhood of } N_0(c) \},
$$

Here, $i_{U,*} : H_1(U, \mathbb{R}) \to H_1(M, \mathbb{R})$ is the mapping induced by the inclusion map $i_U : U \to M$, and $N_0(c)$ denotes the time-0 section of the projected Mañé set $\mathcal{N}(c)$. Let $\nabla_c^\perp \subset H^1(M, \mathbb{R})$ denote the annihilator of $\nabla_c$, i.e. $c' \in \nabla_c^\perp$ if and only if $(c', h) = 0$ for all $h \in \nabla_c$. Clearly,

$$
\nabla_c^\perp = \bigcup_U \{ \ker i_{U,*} : U \text{ is a neighborhood of } N_0(c) \}.
$$

In fact, Mather has proved that there exists a neighborhood $U$ of $N_0(c)$ in $M$ such that $\nabla_c = i_{U,*} H_1(U, \mathbb{R})$ and $\nabla_c^\perp = \ker i_{U,*}$ (see [67]). Then we can introduce the cohomology equivalence (also known as $c$-equivalence).

**Definition 4.1 (Mather’s $c$-equivalence).** We say that $c, c' \in H^1(M, \mathbb{R})$ are $c$-equivalent if there exists a continuous curve $\Gamma : [0, 1] \to H^1(M, \mathbb{R})$ such that $\Gamma(0) = c$, $\Gamma(1) = c'$ and for each $s_0 \in [0, 1]$, $\exists \varepsilon > 0$ such that $\Gamma(s) - \Gamma(s_0) \in \nabla_c^{\perp}$ whenever $|s - s_0| < \varepsilon$ and $s \in [0, 1]$.

By making full use of the cohomology equivalence, Mather obtained a remarkable result on connecting orbits: if $c$ is equivalent to $c'$, the system has an orbit which in the infinite past tends to the Aubry set $\mathcal{A}(c)$ and in the infinite future tends to the Aubry set $\mathcal{A}(c')$ [67].

Next, we recall Arnold’s famous example in [1]: when the stable and unstable manifolds of an invariant circle transversally intersect each other, then the unstable manifold of this circle would also intersect the stable manifold of another invariant circle nearby. To understand this mechanism from a variational viewpoint, we let $\tilde{\pi} : M \to \mathbb{T}^n$ be a finite covering of $\mathbb{T}^n$. Denote by $\tilde{N}(c, M), \tilde{A}(c, M)$ the corresponding Mañé set and Aubry set with respect to $M$. $\tilde{A}(c, M)$ may have several Aubry classes even if $\mathcal{A}(c)$ is unique. Here, we would like to emphasize that $\tilde{\pi}\mathcal{A}(c, M) = \tilde{\mathcal{A}}(c)$. Also, it is not necessary to work always in a nontrivial finite covering space, one can choose $\tilde{M} = M$ if the Aubry set already contains more than one classes. So for Arnold’s famous example, the intersection of the stable and unstable manifolds implies that the set $\tilde{\pi}N(c, M)|_{i=0} \backslash \{ \mathcal{A}(c)|_{i=0} + \delta \}$ is discrete. Here, $\mathcal{A}(c)|_{i=0} + \delta$ stands for a $\delta$-neighborhood of the set $\mathcal{A}(c)|_{i=0}$.

This leads us to introduce the concept of generalized transition chain. This notion could be found in [25, definition 5.1] as a generalization of Arnold’s transition chain [1]. In this paper, we adopt the definition as in [20, definition 4.1] (see also [21, definition 2.2]).

**Definition 4.2 (Generalized transition chain).** Two cohomology classes $c, c' \in H^1(M, \mathbb{R})$ are joined by a generalized transition chain if a continuous path $\Gamma : [0, 1] \to H^1(M, \mathbb{R})$ exists such that $\Gamma(0) = c$, $\Gamma(1) = c'$, and for each $s \in [0, 1]$ at least one of the following cases takes place:

(a) There is $\delta_s > 0$ such that for each $s' \in (s - \delta_s, s + \delta_s) \cap [0, 1]$, $\Gamma(s')$ is $c$-equivalent to $\Gamma(s)$.
Figure 1. A global connecting orbit shadowing the generalized transition chain.

(b) There exist a finite covering $\tilde{\pi} : \tilde{M} \to M$ and a small $\delta_i > 0$ such that the set $\pi N(\Gamma(s), M) \setminus (A(\Gamma(s)))_{|s=0} + \delta_i$ is non-empty and totally disconnected. $A(\Gamma(s'))$ lies in a neighborhood of $A(\Gamma(s))$ provided $|s' - s|$ is small.

We would like to emphasize that, the statement ‘$A(\Gamma(s'))$ lies in a neighborhood of $A(\Gamma(s))$ provided $|s' - s|$ is small’ in condition (b) could be guaranteed by the upper semi-continuity of Aubry sets. In fact, this upper semi-continuity is always true in our model since the number of Aubry classes is only finite (in fact, two at most), see [5]. Also, condition (b) appears weaker than the condition of transversal intersection of stable and unstable manifolds because it still works when the intersection is only topologically transversal. Our condition (b) is usually applied to the case where the Aubry set $A(\Gamma(s))$ is contained in a neighborhood of a lower dimensional torus, while condition (a) is usually applied to the case where the Mañé set $N(\Gamma(s))$ is homologically trivial.

Along a generalized transition chain, one can construct an orbit along which there is a substantial variation:

**Theorem 4.3.** If $c, c' \in H^1(M, \mathbb{R})$ are connected by a generalized transition chain $\Gamma$, then

(a) there exists an orbit $(d\gamma(t), t)$ of the Euler–Lagrangian flow connecting the Aubry set $\tilde{A}(c)$ to $\tilde{A}(c')$, which means the $\alpha$-limit set $\alpha(d\gamma(t), t) \subset \tilde{A}(c)$ and the $\omega$-limit set $\omega(d\gamma(t), t) \subset \tilde{A}(c')$.

(b) for any $c_1, \ldots, c_k \in \Gamma$ and small $\varepsilon > 0$, there exist an orbit $(d\gamma(t), t)$ of the Euler–Lagrangian flow and times $t_1, \ldots, t_k$, such that the orbit $(d\gamma(t), t)$ passes through the $\varepsilon$-neighborhood of $A(c_i)$ at the time $t = t_i$.

The proof of theorem 4.3 is similar to that of [25, section 5] and can also be found in [19, section 7]. This variational mechanism of connecting orbits has already been used in [20, 21]. However, for the reader’s convenience, we provide a proof of the theorem in appendix B. We end this section by a simple illustration of the diffusing orbits in geometry, such orbits constructed in theorems 1.2 and 1.3 would drift near the normally hyperbolic cylinder (see figure 1).

5. Technical estimates on Gevrey functions

In this part, we provide some necessary results for Gevrey functions defined on the torus $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$, which will be useful for our choice of Gevrey space in section 6.3. We present this section in a self-contained way for the reader’s convenience.

The variational proof of the genericity of Arnold diffusion usually depends on the existence of functions with compact support, i.e. bump functions. This technique cannot apply to the problem of analytic genericity since no analytic function has compact support. However, the
bump function does exist in the Gevrey-$\alpha$ category with $\alpha > 1$. Here we give a modified Gevrey bump function which is based on the one constructed in [65].

**Lemma 5.1 (Gevrey bump function).** Let $\alpha > 1$, $L > 0$, $D = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{T}^n$ be an $n$-dimensional cube and $U$ be an open neighborhood of $D$. Then there exists $f \in G^{\alpha,L}(\mathbb{T}^n)$ such that $0 \leq f \leq 1$, supp$f \subset U$, and

$$f(x) = 1 \iff x \in D.$$  

**Proof.** We first claim that for $0 < d < d' < \frac{1}{2}$, there exists a function $g \in G^{\alpha,L}(\mathbb{T})$ such that $0 \leq g \leq 1$ and

$$g(x) = 1 \iff x \in [-d, d], \quad \text{supp} g \subset [-d', d'].$$

Indeed, let $\alpha = 1 + \frac{1}{\sigma}$ ($\sigma > 0$) and define a non-negative function $h \in C^\infty(\mathbb{R})$ as follows: $h(x) = 0$ for $x \leq 0$, $h(x) = \exp(-\frac{x}{\sigma})$ for $x > 0$. Then $h \in G^{\alpha,L}(\mathbb{R})$ if the constant $\lambda > (2L^\sigma / \sin \alpha)^\sigma / \sigma$ with $a = \frac{1}{\sigma} \min\{1, \frac{1}{\sigma}\}$ (cf [64, lemma A.3]). Next, we define

$$\psi(x) = \int_{-d}^{x} h(t + \frac{d' - d}{2}) h\left(-t + \frac{d' - d}{2}\right) \, dt.$$  

It is easy to compute that $\psi \geq 0$ is non-decreasing and

$$\psi(x) = \begin{cases} 0, & x \leq -\frac{d' - d}{2} \\ K, & x \geq \frac{d' - d}{2} \end{cases}$$

where

$$K = \int_{-d}^{\frac{d' - d}{2}} h\left(t + \frac{d' - d}{2}\right) h\left(-t + \frac{d' - d}{2}\right) \, dt > 0.$$  

Then we define the function

$$g(x) = \frac{1}{K^2} \psi\left(x + \frac{d' + d}{2}\right) \psi\left(-x + \frac{d' + d}{2}\right).$$

Obviously, $0 \leq g \leq 1$, supp$g \subset [-d', d']$, and $g(x) = 1 \iff x \in [-d, d]$. It can be viewed as a function defined on $\mathbb{T}$. Hence by property (G1) in section 1, $g \in G^{\alpha,L}(\mathbb{T})$, which proves our claim.

Next, without loss of generality we assume $D = [-d_1, d_1] \times \cdots \times [-d_n, d_n]$ with $0 \leq d_i < \frac{1}{2}$. By assumption, we can find another cube $D' = [-d'_1, d'_1] \times \cdots \times [-d'_n, d'_n]$ such that $D \subset D' \subset U \subset \mathbb{T}^n$. By the claim above, for each $i \in \{1, \ldots, n\}$ there exists a function $f_i \in G^{\alpha,L}(\mathbb{T})$ such that $0 \leq f_i \leq 1$, supp$f_i \subset [-d'_i, d'_i]$, $f_i(x) = 1 \iff x \in [-d_i, d_i]$. Thus we define

$$f(x_1, \ldots, x_n) := \prod_{i=1}^{n} f_i(x_i),$$

which meets our requirements. \[ \square \]

Next, we prove that the inverse of a Gevrey map is still Gevrey smooth. For each high dimensional map $\varphi = (\varphi_1, \ldots, \varphi_n) : V \to \mathbb{R}^n$ where $\varphi_i \in G^{\alpha,L}(V)$, its norm could be defined as follows:

$$\|\varphi\|_{\alpha,L} := \sum_{i=1}^{n} \|\varphi_i\|_{\alpha,L}.$$
In what follows, \((0, 1)^n\) denotes the unit domain \((0, 1) \times \cdots \times (0, 1)\) in \(\mathbb{R}^n\). We also refer the reader to \([51]\) for the inverse function theorem of a general ultra-differentiable mapping.

**Theorem 5.2 (Inverse function theorem of Gevrey class).** Let \(X, Y\) be two open sets in \((0, 1)^n\) and let \(f: X \to Y\) be a Gevrey-(\(\alpha, L\)) map with \(\alpha > 1\). If the Jacobian matrix \(Jf\) is non-degenerate at \(x_0 \in X\), then there exist an open set \(U\) containing \(x_0\), an open set \(V\) containing \(f(x_0)\), a constant \(L_1 < L\), and a unique inverse map \(f^{-1}: V \to U\) such that \(f^{-1} \in G^{\alpha L_1} (V)\).

**Proof.** For simplicity we suppose the Jacobian matrix \(J_{I_n} f = I_n\), where \(I_n = \text{diag}(1, 1, \ldots, 1)\), otherwise we can replace \(f\) by \(f \circ (J_{I_n} f)^{-1}\). We also suppose \(f(x_0) = x_0\), otherwise we can replace \(f\) by \(f + x_0 - f(x_0)\). If we write \(f = id + h\) in a neighborhood of \(x_0\), then \(h(x_0) = 0\), \(J_0 h = 0\). For \(0 < \varepsilon \ll 1\) there exist \(d > 0\) and an open ball \(B_d (x_0) = \{x \in X : ||x - x_0|| < d\}\) such that

\[
||(h||_{C^1 (B_d (x_0)))} \leq \varepsilon. \tag{5.1}
\]

By classical inverse function theorem, there exist two small open sets \(U, V \subset B_{d/2} (x_0)\) containing \(x_0\) and a unique \(C^\infty\) inverse map \(f^{-1}: V \to U\) where \(f^{-1}(x_0) = x_0\). Let \(L_1 = \varepsilon^{\frac{1}{\alpha}}\), next we will prove \(f^{-1} \in G^{\alpha L_1} (V)\) by the contraction mapping principle. We can write \(f^{-1} = id + g\), so \(g \in C^\infty (V)\) and the equality

\[
g(y) = -h(y + \varphi(y)), \quad \forall y \in V
\]

holds. Define the set \(E = \{\varphi = (\varphi_1, \ldots, \varphi_n) : \varphi(x_0) = 0, \varphi \in G^{\alpha L_1} (V), \|\varphi\|_{\alpha L_1} \leq \varepsilon^\frac{1}{\alpha}\}\) with the norm \(\|\cdot\|_{\alpha L_1}\), it is a non-empty, closed and convex set in the space \(G^{\alpha L_1} (V)\). Define the operator

\[
(T\varphi)(y) := -h(y + \varphi(y)), \quad \forall y \in V.
\]

- We first claim that the mapping \(T\varphi \in E, \forall \varphi \in E\). In fact, for each \(\varphi \in E, (T\varphi)(x_0) = 0\). For \(y \in V \subset B_{d/2} (x_0)\), we have \(\|y + \varphi (y) - x_0\| \leq \|y - x_0\| + \|\varphi (y) - \varphi (x_0)\| \leq \frac{d}{2} + \|J \varphi\| y - x_0\| < d\) hence \((id + \varphi)(V) \subset B_d (x_0)\). Moreover, let \(L_2 := L \varepsilon^{\frac{1}{\alpha n}}\) and \(\varepsilon\) be suitably small. For each \(i \in \{1, \ldots, n\}\),

\[
\|\varphi_i\| \leq n L_1^n \left(1 + \frac{\varepsilon^{\frac{1}{\alpha}}}{L_1}\right) + \|\varphi_i\| \leq 2 n \varepsilon^{\frac{1}{\alpha}} + \varepsilon^{\frac{1}{\alpha}} \leq \frac{L_2}{n^{\alpha - 1}}.
\]

where \(\delta_{ij} = 1\) for \(i = j\) and \(\delta_{ij} = 0\) for \(i \neq j\). Hence by property (G3) in section 1, \(\|T \varphi\|_{\alpha L_1} = \|h \circ (id + \varphi)\|_{\alpha L_1} \leq \|h\|_{\alpha L_2, B_d (x_0)}\) since \((id + \varphi)(V) \subset B_d (x_0)\). Now it only remains to verify that

\[
\|h\|_{\alpha L_2, B_d (x_0)} \leq \varepsilon^{\frac{1}{\alpha}}.
\]

Recall that for \(|k| \geq 2\) and \(x \in B_d (x_0)\), \(\partial^k f (x) = \partial^k h (x)\). By using (5.1), we have
\[\|h_i\|_{C^k_{\alpha}B_2(0)} = \|h_i\|_{C^0_{\alpha}B_2(0)} + \sum_{k \in \mathbb{N}, |k| \geq 1} L^2_k \|\partial^k h_i\|_{C^0_{\alpha}B_2(0)} + \sum_{k \in \mathbb{N}, |k| \geq 2} \frac{L^2_k}{k!} \|\partial^k f_i\|_{C^0_{\alpha}B_2(0)} \leq (1 + nL^2)\varepsilon + \sum_{k \in \mathbb{N}, |k| \geq 2} \frac{L^2_k}{k!} \|\partial^k f_i\|_{C^0_{\alpha}B_2(0)} \leq (1 + nL^2)\varepsilon + \varepsilon \|f\|_{C^0_{\alpha}L^1} \leq \frac{\varepsilon^2}{n}, \quad (5.2)\]

which proves the claim.

• On the other hand, for \(\varphi, \tilde{\varphi} \in E\) and \(i \in \{1, \ldots, n\}\), by the Newton–Leibniz formula we have

\[h_i(x + \varphi(x)) - h_i(x + \tilde{\varphi}(x)) = \left( \int_0^1 Jh_i \left( x + s\varphi(x) + (1 - s)\tilde{\varphi}(x) \right) ds \right) (\varphi(x) - \tilde{\varphi}(x)) = F(x) (\varphi(x) - \tilde{\varphi}(x))\]

where \(Jh_i\) is the Jacobian matrix. It follows from property (G2) in section 1 and (5.2) that

\[\|Jh_i\|_{C^0_{\alpha}B_{L_2}(0)} \leq \frac{\|h_i\|_{C^1_{\alpha}B_{L_2}(0)}}{(L_2 - L_2/2)^n} \sim O(\varepsilon^{1/2}) < \frac{1}{2n}\]

provided \(\varepsilon\) is suitably small. By property (G3), \(\|F\|_{C^0_{\alpha}L^1} \leq \|Jh_i\|_{C^0_{\alpha}B_{L_2}(0)} \leq \frac{1}{2n}\).

Finally, we deduce from (G1) that

\[\|h_i \circ (id + \varphi) - h_i \circ (id + \tilde{\varphi})\|_{C^0_{\alpha}L^1} \leq \|F\|_{C^0_{\alpha}L^1} \|\varphi - \tilde{\varphi}\|_{C^0_{\alpha}L^1} \leq \frac{1}{2n}\|\varphi - \tilde{\varphi}\|_{C^0_{\alpha}L^1}.

Hence \(\|h \circ (id + \varphi) - h \circ (id + \tilde{\varphi})\|_{C^0_{\alpha}L^1} \leq \frac{1}{2}\|\varphi - \tilde{\varphi}\|_{C^0_{\alpha}L^1}\), namely

\[\|T\varphi - T\tilde{\varphi}\|_{C^0_{\alpha}L^1} \leq \frac{1}{2}\|\varphi - \tilde{\varphi}\|_{C^0_{\alpha}L^1}.

In conclusion, \(T : E \to E\) is a contraction mapping. By the contraction mapping principle, \(T\) has a unique fixed point, hence the fixed point must be \(g\). Therefore, \(f^{-1} = id + g \in G^{k_1}(V)\). \(\square\)

Sometimes we need to approximate a continuous function by Gevrey smooth ones. Convolution provides us with a systematic technique. More specifically, for any \(\alpha > 1, L > 0\), by lemma 5.1 there exists a non-negative function \(\eta \in G^{1-L}(\mathbb{R}^n)\) such that \(\text{supp}\eta \subset \left[\frac{1}{2}, \frac{3}{2}\right]^n\) and \(\int_{\mathbb{T}^n} \eta(x) \, dx = 1\). Next we set \(\eta_\varepsilon(x) = \frac{1}{\varepsilon} \eta(\frac{x}{\varepsilon})\) \((0 < \varepsilon < 1, x \in \mathbb{R}^n)\) which is called the mollifier. Then we define the convolution of \(\eta_\varepsilon\) and \(f \in C^0(\mathbb{T}^n)\) by

\[\eta_\varepsilon * f(x) = \int_{\mathbb{T}^n} \eta_\varepsilon(x - y)f(y) \, dy, \quad \forall x \in \mathbb{T}^n. \quad (5.3)\]
Theorem 5.3 (Gevrey approximation).

(a) Let $\alpha > 1$, and $U \subset \mathbb{T}^n$, $V \subseteq (0, 1)^n$ be two open sets. If $f : U \to V$ is a continuous map, then there exists a sequence of maps $f^\varepsilon : U \to (0, 1)^n$ such that $f^\varepsilon \in G^{\alpha, \lambda}(U)$. Furthermore, $L_\varepsilon \to 0$ and $\|f^\varepsilon - f\|_c^0 \to 0$ as $\varepsilon$ tends to 0.

(b) Let $\alpha > 1$, $U, V$ be connected open sets satisfying $U, V \subseteq \mathbb{T}^n$ and $f : U \to V$ be a continuous map. Then there exists a sequence of maps $f^\varepsilon : U \to \mathbb{T}^n$ such that $f^\varepsilon \in G^{\alpha, \lambda}(U)$, $L_\varepsilon \to 0$ and $\|f^\varepsilon - f\|_c^0 \to 0$ as $\varepsilon$ tends to 0. Specifically, if $f$ is a diffeomorphism and the determinant $\det(Jf)$ ($Jf$ is the Jacobian matrix) has a uniform positive distance away from zero, then the Gevrey map $f^\varepsilon : U \to V^\varepsilon$ with $V^\varepsilon = f^\varepsilon(U)$ will also be a diffeomorphism provided that $\varepsilon$ is small enough.

Proof. (a) Let $f = (f_1, \ldots, f_n)$ and $f_i$ ($1 \leq i \leq n$) be continuous, we only need to prove that each $f_i$ can be approximated by a Gevrey smooth function. Indeed, let $f_i^\varepsilon = \eta_\varepsilon * f_i$ ($0 < \varepsilon < 1$), where $\eta \in G^{\alpha, \lambda}$. It’s easy to check that $f_i^\varepsilon : U \to (0, 1)$ since $\int_{\mathbb{R}^n} \eta_\varepsilon(x) \, dx = 1$ and $\text{supp} \eta \subseteq \left[\frac{\varepsilon}{2}, \frac{1}{2}\right]^n$. By the classical properties of convolutions, one obtains $f_i^\varepsilon \in C^\infty$ and

$$\|f_i^\varepsilon - f_i\|_c^0 \to 0, \quad \text{as} \ \varepsilon \to 0.$$

It only remains to prove $f_i^\varepsilon$ is Gevrey smooth. In fact, if one sets $L_\varepsilon = L_\varepsilon^\varepsilon$, then

$$\|f_i^\varepsilon\|_{\alpha, L_\varepsilon} \leq \sum_k L_{k\alpha}^{\varepsilon\varepsilon} \|\partial^k \eta\|_c^0 \|f_i\|_c^0$$

$$\leq \sum_k \frac{\|f_i\|_c^0}{\varepsilon^n} \sum_{k \geq k \varepsilon} L_{k\alpha}^{\varepsilon\varepsilon} \|\partial^k \eta\|_c^0$$

$$= \sum_k \frac{\|f_i\|_c^0}{\varepsilon^n} \sum_{k \geq k \varepsilon} L_{k\alpha}^{\varepsilon\varepsilon} \|\partial^k \eta\|_c^0 = \|f_i\|_c^0 \|\eta\|_{\alpha, L}.$$

Obviously, $L_\varepsilon \to 0$ as $\varepsilon \to 0$. This completes the proof of (a).

(b) The first part is not hard to prove by the technique in (a). Furthermore, if $f$ is a diffeomorphism from $U$ to $V$, then by using $\partial^k f^\varepsilon = \eta_\varepsilon * \partial^k f$ with $|k| = 1$, one gets

$$\|f^\varepsilon - f\|_c^1 \to 0, \quad \varepsilon \to 0. \quad (5.4)$$

Since $\det(Jf)$ has a uniform positive distance away from zero, it concludes from (5.4) and theorem 5.2 that $f^\varepsilon : U \to f^\varepsilon(U)$ would also be a diffeomorphism for $\varepsilon$ small enough. \hfill \Box

6. Proof of the main results

This section is the main part of the present paper, which aims to prove theorems 1.2 and 1.3. We will explain how to apply the tools exposed in the previous sections to a priori unstable and Gevrey smooth systems. Before that, we need to do some preparations.
6.1. Genericity of uniquely minimal measure in Gevrey or analytic topology

Let $M = T^n$. We fix an $h \in H_1(M, \mathbb{R})$, it is well known that in the $C^r$ ($2 \leq r \leq \infty$) topology, a generic Lagrangian has only one minimal measure $\mu$ with the rotation vector $\rho(\mu) = h$ (see [61]). Next, we will show that such a property still holds in the Gevrey topology. For this purpose, we shall consider it in a Gevrey space $G^\alpha L(M \times \mathbb{T})$ with $\alpha \geq 1$, $L > 0$. A property is called generic in the sense of Mañé if, for each Lagrangian $L: TM \times \mathbb{T} \to \mathbb{R}$, there exists a residual subset $\mathcal{O} \subset G^\alpha L(M \times \mathbb{T})$ such that the property holds for each Lagrangian $L + \phi$ with $\phi \in \mathcal{O}$.

**Theorem 6.1.** Let $h \in H_1(M, \mathbb{R})$, $\alpha \geq 1$, $L > 0$ and $L: TM \times \mathbb{T} \to \mathbb{R}$ be a Tonelli Lagrangian, then there exists a residual subset $\mathcal{O}(h) \subset G^\alpha L(M \times \mathbb{T})$ such that, for each $\phi \in \mathcal{O}(h)$, the Lagrangian $L + \phi$ has only one minimal measure with the rotation vector $h$.

**Remark.** We shall note that the residual set $\mathcal{O}(h)$ depends on the homology class $h$.

**Proof.** Recall Mañé’s equivalent definition of minimal measure in section 2, we are going to prove this theorem in the following setting based on Mañé’s approach.

(a) Set $E := G^\alpha L(M \times \mathbb{T})$. Obviously, it is a Banach space.

(b) Denote by $F \subset C^r$ the vector space spanned by the set of probability measures $\mu \in \mathcal{H}$ with $\int_{TM \times \mathbb{T}} L d\mu < \infty$, the definitions of the sets $\mathcal{H}$ and $C^r$ are in section 2. Recall that for $\mu_k, \mu \in F$,

$$\lim_{k \to +\infty} \mu_k = \mu \iff \lim_{k \to +\infty} \int_{TM \times \mathbb{T}} f d\mu_k = \int_{TM \times \mathbb{T}} f d\mu, \quad \forall f \in C.$$

(c) Let $L: F \to \mathbb{R}$ be a linear map satisfying $L(\mu) = \int L d\mu$, for every $\mu \in F$.

(d) Let $\varphi: E \to F^*$ be a linear map such that for each $\phi \in E$, $\varphi(\phi) \in F^*$ is defined as follows

$$\langle \varphi(\phi), \mu \rangle := \int_{TM \times \mathbb{T}} \phi d\mu, \quad \mu \in F.$$

(e) $K := \{ \mu \in F | \rho(\mu) = h \}$. It is easy to check that $K$ is a separable metrizable convex subset.

For $\phi \in E$, we denote

$$\text{arg min}(\phi) := \{ \mu \in K \mid L(\mu) + \langle \varphi(\phi), \mu \rangle = \min_{\nu \in K} (L(\nu) + \langle \varphi(\phi), \nu \rangle) \}.$$

It is easy to verify that our setting above satisfies all conditions of that in [61, proposition 3.1], so there exists a residual subset $\mathcal{O}(h) \subset E$ such that each $\phi \in \mathcal{O}(h)$ has the following property:

$$\# \text{ arg min}(\phi) = 1.$$

Since $\text{ arg min}(\phi) = \delta_h(L + \phi)$, see (2.7), it follows from proposition 2.2 that the Lagrangian $L + \phi$ admits only one minimal measure with the rotation vector $h$. \hfill \Box

**Remark.** For $\alpha = 1$, $G^{1L}$ is the space of real analytic functions. This therefore means that the uniqueness of minimal measure is also a generic property in the analytic topology.

**Corollary 6.2.** Let $L > 0$, $\alpha \geq 1$ and $L: T^n \times \mathbb{T}$ be a Tonelli Lagrangian. Then there exists a residual set $\mathcal{O}_1 \subset G^\alpha L(T^n \times \mathbb{T})$ such that for any $V \in \mathcal{O}_1$, the Lagrangian $L + V$ is $\mathcal{O}_1$-generic in the sense of Mañé.

\footnote{A residual subset $X$ of a Baire space is one whose complement is the union of countably many nowhere dense subsets. Every residual set is a dense set.}
has the following property: for each rational \( h = (h_1, \ldots, h_n) \in H_1(\mathbb{T}^n, \mathbb{R}) \) with \( h_i \in \mathbb{Q} \), \( L + V \) has one and only one minimal measure with the rotation vector \( h \).

**Proof.** For each \( h \in H_1(\mathbb{T}^n, \mathbb{R}) \), thanks to theorem 6.1, we obtain a residual subset \( \mathcal{O}(h) \subset G^{\alpha,L}(\mathbb{T}^n \times \mathbb{T}) \) such that for each \( \phi \in \mathcal{O}(h) \), the Lagrangian \( L + \phi \) has only one minimal measure with the rotation vector \( h \). Then we set

\[
\mathcal{O}_1 = \bigcap_{h \in \mathbb{Q}^n} \mathcal{O}(h),
\]

which is the intersection of countably many residual sets. Of course, by the definition of residual set, \( \mathcal{O}_1 \) is non-empty and still residual (also dense) in the Banach space \( G^{\alpha,L}(\mathbb{T}^n \times \mathbb{T}) \). The corollary is now evident from what we have proved. \( \square \)

### 6.2. Hölder regularity of elementary weak KAM solutions

In this part, we will choose a family of elementary weak KAM solutions which can be parameterized so that they are Hölder continuous in the \( C^0 \) topology. Such a property is crucial for our proof of theorem 6.7.

To this end, we need to do study the normally hyperbolic cylinders (refer to appendix A). Let us go back to our two and a half degrees of freedom Hamiltonian model (1.2). Let

\[
\Sigma(0) := \{ (q_1, 0, p_1, 0) : q_1 \in \mathbb{T}, |p_1| \leq R \} \subset \mathbb{T}^2 \times \mathbb{R}^2.
\]

It is a cylinder restricted on the time-0 section, where \( R \) is the constant fixed in section 1. By condition (H2), \( \Sigma(0) \) is a NHIC for the time-1 map of the Hamiltonian flow \( \Phi_{H_0}^1 \). Since the Hamiltonian \( H_0 \) is integrable when restricted in the cylinder \( \Sigma(0) \), the rate \( \mu \) in (A.1) is 1 and \( \log \mu = 0 \), so it follows from theorem A.3 that there exists

\[
\varepsilon_1 = \varepsilon_1(H_0, R) > 0
\]

such that if \( \| H_1 \|_{C^1(\mathbb{T}^n)} \leq \varepsilon_1 \), the time-1 map \( \Phi_{H_1}^1 \) of the Hamiltonian \( H \) still admits a \( C^{r-1} \) NHIC \( \Sigma_{H_1}(0) \), which is a small deformation of \( \Sigma(0) \) and can be considered as the image of the following diffeomorphism (see figure 2)

\[
\psi : \Sigma(0) \to \Sigma_{H_1}(0) \subset \mathbb{T}^2 \times \mathbb{R}^2,
\]

\[
(q_1, 0, p_1, 0) \mapsto (q_1, q_2(q_1, p_1), p_1, p_2(q_1, p_1)).
\]

Here, \( q_2 \) and \( p_2 \) are two \( C^{r-1} \) functions taking values close to zero. Then \( \psi \) induces a 2-form \( \psi^*\Omega \) on the standard cylinder \( \Sigma(0) \) where \( \Omega = \sum_{i=1}^2 dq_i \wedge dp_i \),

\[
\psi^*\Omega = \left( 1 + \frac{\partial q_2}{\partial (q_1, p_1)} \right) dq_1 \wedge dp_1.
\]
Since the second de Rham cohomology group $H^2(\Sigma(0), \mathbb{R}) = \{0\}$, by using Moser’s trick on the isotopy of symplectic forms, one can find a diffeomorphism $\psi_1 : \Sigma(0) \to \Sigma(0)$ such that

$$\psi_1^* \psi_1^* \Omega = dq_1 \wedge dp_1.$$ 

Recall that $\Sigma_{\Omega}(0)$ is invariant under $\Phi^t_{\Omega}$ and $(\Phi^t_{\Omega})^* \Omega = \Omega$, one obtains

$$((\psi \circ \psi_1)^{-1} \circ \Phi^t_{\Omega} \circ (\psi \circ \psi_1))^* dq_1 \wedge dp_1 = dq_1 \wedge dp_1.$$ 

Combining with the fact that $(\psi \circ \psi_1)^{-1} \circ \Phi^t_{\Omega} \circ (\psi \circ \psi_1)$ is a small perturbation of $\Phi^t_{\Omega}$, the map $(\psi \circ \psi_1)^{-1} \circ \Phi^t_{\Omega} \circ (\psi \circ \psi_1)$ is an exact twist map, hence one can apply the classical Aubry–Mather theory to characterize the minimal orbits on $\Sigma(0)$: given a $\rho \in \mathbb{R}$, there exists an Aubry–Mather set with rotation number $\rho$ satisfying

(a) if $\rho \in \mathbb{Q}$, the set consists of periodic orbits;
(b) if $\rho \in \mathbb{R}\setminus\mathbb{Q}$, the set is either an invariant circle or a Denjoy set.

For simplicity, we denote by

$$\Sigma_{\Omega}(s) = \Phi^s_{\Omega}(\Sigma_{\Omega}(0), 0), \quad \Sigma(s) = \Phi^s_{\Omega}(\Sigma(0), 0)$$

the 2-dimensional manifolds, and denote by

$$\bar{\Sigma}_{\Omega} = \bigcup_{s \in \mathbb{T}} \Sigma_{\Omega}(s), \quad \bar{\Sigma} = \bigcup_{s \in \mathbb{T}} \Sigma(s)$$

the 3-dimensional manifolds in $T^*\mathbb{T}^2 \times \mathbb{T}$. By using the Legendre transformation $\mathcal{L}$ (see (1.4)), the set $\mathcal{L}\bar{\Sigma}_{\Omega}$ is $\phi_1^t$-invariant in $T^*\mathbb{T}^2 \times \mathbb{T}$. Given a cohomology class $c = (c_1, 0) \in H^1(\mathbb{T}^2, \mathbb{R})$ with $|c_1| \leq R - 1$, the following lemma shows that the Aubry set $\mathcal{A}(c)$ lies inside the cylinder $\mathcal{L}\bar{\Sigma}_{\Omega}$.

**Lemma 6.3 (Location of the minimal sets).** Let $H$ be the Hamiltonian (1.2) and $L$ be the associated Lagrangian (1.3). There exists $\varepsilon_1 = \varepsilon_1(H_0, R) > 0$ such that if $\|H_1\|_{C^1(\mathbb{R})} \leq \varepsilon_1$, then for each $c = (c_1, 0)$ with $|c_1| \leq R - 1$, the globally minimal set $\mathcal{G}_L(c) \subset \mathcal{L}\bar{\Sigma}_{\Omega}$.

**Proof.** We first consider the autonomous Lagrangian $L_2(q_2, v_2)$. It follows from (1.5) that $(0, 0)$ is the unique minimal point of $L_2$, so the globally minimal set of the Lagrangian $L_2$ is

$$\mathcal{G}_{L_2} = (0, 0) \times \mathbb{T} \subset T\mathbb{T} \times \mathbb{T}.$$ 

Then for all $c = (c_1, 0)$ with $|c_1| \leq R$, the globally minimal set of $L_0 = L_1(v_1) + L_2(q_2, v_2)$ is

$$\mathcal{G}_{L_0}(c) = \{(q_1, 0, Dh_1(c_1), 0, t) : q_1 \in \mathbb{T}, t \in \mathbb{T}\} \quad \text{and} \quad \mathcal{G}_{L_0}(c) \subset \mathcal{L}\bar{\Sigma}.$$ 

Here, the function $h_1$ is given in (1.2).

Next, we take a small neighborhood $U$ of $\mathcal{L}\bar{\Sigma}$ in the space $T^*\mathbb{T}^2 \times \mathbb{T}$ and let $\varepsilon_1 = \varepsilon_1(H_0, R)$ be the constant defined in (6.1). Since $\|H_1\|_{C^1(\mathbb{R})} \leq \varepsilon_1$, by letting $\varepsilon_1$ suitably small, it follows
that $\|L_1\|_{C^1(\mathbb{R}^2)}$ is also sufficiently small. Thus, by the upper semi-continuity in proposition 2.1, $G_L(c) \subset U$ for all $c \in [-R + 1, R - 1] \times \{0\}$ where $L = L_0 + L_1$. Equivalently, $L^{-1}G_L(c) \subset L^{-1}U$.

On the other hand, due to normal hyperbolicity and theorem A.3, $\tilde{\Sigma}_H \subset L^{-1}U$, provided that $\varepsilon_1$ is small enough. Moreover, $\tilde{\Sigma}_H$ is the unique $\phi^t$-invariant set in the neighborhood $L^{-1}U$. This therefore implies $L^{-1}G_L(c) \subset \tilde{\Sigma}_H$ since $L^{-1}G_L(c)$ is $\phi^t$-invariant. $\square$

In the remainder of this section, we will use the following notation for simplicity.

Notation:

(a) In what follows, we use $M$ to denote the manifold $T^2 = \mathbb{R}^2 / \mathbb{Z}^2$. Also, we denote by

$$\hat{M} = T \times 2T = \mathbb{R} / \mathbb{Z} \times \mathbb{R} / 2\mathbb{Z}, \quad \hat{\pi} : M \to M$$

the double covering of $M$. We use such a double covering to distinguish between 0 and 1 in the $q_2$-coordinate, and identify 0 with 2 in the $q_2$-coordinate.

The Hamiltonian $H : T^*M \times T \to \mathbb{R}$ and the Lagrangian $L : TM \times T \to \mathbb{R}$ could extend naturally to $T^*\hat{M}$ and $T\hat{M}$ respectively. So by abuse of notation, we continue to write $H : T^*\hat{M} \times T \to \mathbb{R}$ and $L : T\hat{M} \times T \to \mathbb{R}$ for the new Hamiltonian and Lagrangian respectively. In this setting, the lift of the NHIC $\Sigma_H(0)$ will have two copies $\hat{\Sigma}_H = \hat{\Sigma}_{H,l}(0) \cup \hat{\Sigma}_{H,u}(0)$, where the subscripts $l,u$ are introduced to indicate ‘lower’ and ‘upper’ respectively. Then $\hat{\Sigma}_H = \hat{\Sigma}_{H,l} \cup \hat{\Sigma}_{H,u}$.

(b) For simplicity, we always use $\pi_q$ to denote the natural projection from $T\hat{M}$ (respectively $TM$) to $M$ (respectively $M$) or from $T^*\hat{M}$ (respectively $T^*M$) to $M$ (respectively $M$).

(c) Let $\kappa > 0$ be small, we denote by $U_\kappa = U_{\kappa,l} \cup U_{\kappa,u}$ the disconnected subset in $\hat{M}$ where

$$U_{\kappa,l} = T \times [\kappa, 1 - \kappa], \quad U_{\kappa,u} = T \times [1 + \kappa, 2 - \kappa].$$

Let $N_\kappa = \hat{M} \setminus U_\kappa = N_{\kappa,l} \cup N_{\kappa,u}$ where

$$N_{\kappa,l} = T \times (-\kappa, \kappa), \quad N_{\kappa,u} = T \times (1 - \kappa, 1 + \kappa).$$

The subscripts $l,u$ are also introduced to indicate the ‘lower’ and the ‘upper’ respectively (See figure 3). The number $\kappa$ should be chosen such that

$$\pi_q \circ \Sigma_{H,l}(0) \subset N_{\kappa/2,l}, \quad \pi_q \circ \Sigma_{H,u}(0) \subset N_{\kappa/2,u}.$$ 

Namely, the perturbed cylinder is contained in a $\kappa/2$-neighborhood of the unperturbed one. Also, we let

$$\tilde{\Sigma}_{H,l} \subset N_{\kappa/2,l} \times \mathbb{R}^2 \times T, \quad \tilde{\Sigma}_{H,u} \subset N_{\kappa/2,u} \times \mathbb{R}^2 \times T. \quad (6.4)$$
(d) For \( e = (e_1, 0) \in H^1(M, \mathbb{R}) \), if the Aubry set \( \tilde{\mathcal{A}}_{L}(c, M) \) is an invariant circle, we denote by

\[
\tilde{\mathcal{T}}_e = \mathcal{L}^{-1} \tilde{\mathcal{A}}_L(c, M) |_{t=0} \subset T^*M \times \{ t = 0 \}
\]

the invariant circle in the cotangent space. This leads us to introduce an index set

\[
S := \{ (e_1, 0) \mid |e_1| \leq R - 1, \; \tilde{\mathcal{T}}_e \text{ is an invariant circle lying in } \Sigma_H(0) \}.
\]

(6.5)

(e) Let \( r_0 > 0 \) be small satisfying \( r_0 > \kappa \). Since \( \Sigma_H(0) \) is a NHIM for the time-1 map \( \Phi_t \), we have the associated local stable and unstable manifolds, denoted by \( W^S_{\Sigma_H(0)} \) and \( W^U_{\Sigma_H(0)} \) respectively, in the \( r_0 \)-tubular neighborhood of \( \Sigma_H(0) \). In addition, \( W^S_{\Sigma_H(0)} = \bigcup_{q \in \Sigma_H(0)} \tilde{W}^S_q \) and \( W^U_{\Sigma_H(0)} = \bigcup_{q \in \Sigma_H(0)} \tilde{W}^U_q \).

Now, let us focus on \( c = (e_1, 0) \in S \). For the \( \Phi_t \)-invariant circle \( \tilde{\mathcal{T}}_e \), it has local stable manifold \( \tilde{W}^S_{\tilde{\mathcal{T}}_e} = \bigcup_{q \in \tilde{\mathcal{T}}_e} \tilde{W}^S_q \) and local unstable manifold \( \tilde{W}^U_{\tilde{\mathcal{T}}_e} = \bigcup_{q \in \tilde{\mathcal{T}}_e} \tilde{W}^U_q \). Theorem A.2 tells us that the leaf \( \tilde{W}^S_q \) (respectively \( \tilde{W}^U_q \)) has smooth dependence on the base point \( q \in \Sigma_H(0) \). Consequently, \( \tilde{W}^S_{\tilde{\mathcal{T}}_e} \) (respectively \( \tilde{W}^U_{\tilde{\mathcal{T}}_e} \)) is a Lipschitz manifold since \( \tilde{\mathcal{T}}_e \) is only Lipschitz in general. Besides, the local stable (unstable) manifold can be viewed as a Lipschitz graph over \( \tilde{\pi} \circ N_{r_0} \), namely

\[
\tilde{W}^S_{\tilde{\mathcal{T}}_e} = \{ (q, q_2, \tilde{p}_1^S(q_1, q_2), \tilde{p}_2^S(q_1, q_2)) \in T^*M \times \{ t = 0 \} : (q_1, q_2) \in \tilde{\pi} \circ N_{r_0} \}
\]

\[
\tilde{W}^U_{\tilde{\mathcal{T}}_e} = \{ (q_1, q_2, \tilde{p}_1^U(q_1, q_2), \tilde{p}_2^U(q_1, q_2)) \in T^*M \times \{ t = 0 \} : (q_1, q_2) \in \tilde{\pi} \circ N_{r_0} \}.
\]

Here, \( \tilde{p}_1^S, \tilde{p}_2^U \) are Lipschitz functions on \( \tilde{\pi} \circ N_{r_0} \subset M \), and the domain \( N_{r_0} = N_{r_0,1} \cup N_{r_0,2} \) with

\[
N_{r_0,1} = T \times (-r_0, r_0), \quad N_{r_0,2} = T \times (1 - r_0, 1 + r_0).
\]

Next, in the covering space \( \tilde{M} \), the Aubry set \( \tilde{\mathcal{A}}_{L}(c, \tilde{M}) \) is the union of two disjoint copies of \( \tilde{\mathcal{A}}_{L}(c, M) \) satisfying \( \pi \tilde{\mathcal{A}}_{L}(c, M) = \tilde{\mathcal{A}}_{L}(c, M) \). More precisely, \( \mathcal{L}^{-1} \tilde{\mathcal{A}}_{L}(c, M) |_{t=0} = \tilde{\mathcal{T}}_{c,1} \cup \tilde{\mathcal{T}}_{c,a} \), where \( \tilde{\mathcal{T}}_{c,1} \) lies in \( \Sigma_H(0) \) and its stable and unstable manifolds are

\[
\tilde{W}^S_{\tilde{\mathcal{T}}_{c,1}} = \{ (q, q_2, \tilde{p}_1^S(q_1, q_2), \tilde{p}_2^S(q_1, q_2)) \in T^*\tilde{M} \times \{ t = 0 \} : (q, q_2) \in N_{r_0} \}
\]

\[
\tilde{W}^U_{\tilde{\mathcal{T}}_{c,1}} = \{ (q_1, q_2, \tilde{p}_1^U(q_1, q_2), \tilde{p}_2^U(q_1, q_2)) \in T^*\tilde{M} \times \{ t = 0 \} : (q_1, q_2) \in N_{r_0} \}.
\]
with \( i = 1, u \). Here, by abuse of notation, we have continued to use \( \Pi^u, \Pi^s \) to denote the corresponding Lipschitz functions defined on the lift of \( \pi \circ N_{u,b} \). The lemma below gives the relation between the elementary weak KAM solutions and the local stable and unstable manifolds.

**Lemma 6.4.** There exists \( r_0 > 0 \) such that for each \( c = (c_1, 0) \in S \), we have

(a) for each backward elementary weak KAM solution \( u_{\gamma}^c(q, t) \) with \( i = 1, u \), the function \( u_{\gamma}^c(q, 0) \) is \( C^{1,1} \) in the domain \( N_{r,0} \), and generates the local unstable manifold of \( \gamma_{c,0} \).

\[
W_{\gamma_{c,0}}^{\text{loc}} = \{(q, c + \partial_t u_{\gamma}^c(q, 0)) : q \in N_{r,0}\}, \quad i = 1, u.
\]

(b) for each forward elementary weak KAM solution \( u_{\gamma}^c(q, t) \) with \( i = 1, u \), the function \( u_{\gamma}^c(q, 0) \) is \( C^{1,1} \) in the domain \( N_{r,0} \), and generates the local stable manifold of \( \gamma_{c,0} \), i.e.

\[
W_{\gamma_{c,0}}^{\text{loc}} = \{(q, c + \partial_t u_{\gamma}^c(q, 0)) : q \in N_{r,0}\}, \quad i = 1, u.
\]

**Proof.** We only prove for the case \( u_{\gamma}^c \) since the other cases are similar.

**Step 1.** We first claim that there exists a neighborhood \( V \) of \( \pi_q \circ T_{c,1} \) in \( \tilde{M} \) such that for each \( \xi_0 : (-\infty, 0] \to \tilde{M} \) calibrated by \( u_{\gamma}^c \) with \( \xi_0(0) \in V \), the \( \alpha \)-limit set of the backward minimal configuration \( \{\xi(-t)\}_{t \in \mathbb{Z}^+} \) must be contained in \( V \).

Let \( \gamma_0 : (-\infty, 0] \to \tilde{M} \) be the local stable manifold of \( \gamma_{c,0} \), and \( S_0 = \gamma_0(0) \). Assume by contradiction that there exist a sequence of backward calibrated curves \( \xi_k^c : (-\infty, 0] \to \tilde{M} \) with \( \xi_k^c(0) = x_k \), and a sequence \( \alpha_k \) which belongs to the \( \alpha \)-limit set of the backward minimal configuration \( \{\xi_k^c(-t)\}_{t \in \mathbb{Z}^+} \) satisfying

\[
\lim_{k \to \infty} x_k = x^* \in \pi_q \circ T_{c,1} \quad \text{and} \quad \lim_{k \to \infty} \alpha_k = \alpha^* \notin \pi_q \circ T_{c,1}.
\]

This implies \( \alpha^* \notin \pi_q \circ T_{c,1} \) since the \( \alpha \)-limit set of each minimal curve shall be contained in the Aubry set. By theorem 3.4, each \( \xi_k^c : (-\infty, 0] \to \tilde{M} \) is \( c \)-semi static and calibrated by \( h^c((x^*, 0), -t) \):

\[
h^c((x^*, 0), (\xi_k^c(0), 0)) - h^c((x^*, 0), (\xi_k^c(-t), 0)) = h^c((\xi_k^c(-t), 0), (\xi_k^c(0), 0)), \quad \forall t \in \mathbb{Z}^+.
\]

This further gives \( h^c((x^*, 0), (x_k, 0)) \geq h^c((x^*, 0), (\alpha_k, 0)) \geq h^c((\alpha_k, 0), (x_k, 0)). \) The opposite inequality is obvious, therefore \( h^c((x^*, 0), (x_k, 0)) - h^c((x^*, 0), (\alpha_k, 0)) = h^c((\alpha_k, 0), (x_k, 0)). \) Sending \( k \to \infty \), it follows that

\[
0 = h^c((x^*, 0), (x_k, 0)) - h^c((x^*, 0), (\alpha_k, 0)) + h^c((\alpha_k, 0), (x_k, 0)).
\]

So \((x^*, 0)\) and \((\alpha^*, 0)\) belong to the same Aubry class, which contradicts (6.6).

**Step 2.** By letting the above domain \( V \) be suitably small if necessary, \( W_{\gamma_{c,0}}^{\text{loc}} \) is a Lipschitz graph over \( V \), we will show that there exists a small number \( r_0 > 0 \) such that \( N_{r,0} \subset V \), and each \( u_{\gamma_{c,1}} \)-calibrated curve \( \gamma^c : (-\infty, 0] \to \tilde{M} \) with \( \gamma^c(0) \in N_{r,0} \) satisfies \( \gamma^c(-m) \in V \), \( \forall m \in \mathbb{N} \).

To prove this, assume by contradiction that there exist a sequence of \( u_{\gamma_{c,1}} \)-calibrated curves \( \gamma_j^c : (-\infty, 0] \to \tilde{M} \) and a sequence of positive integers \( T_j \) such that \( \gamma_j^c(-T_j) \notin V \), \( \gamma_j^c(-m) \in V \), \( m \in \{0, 1, \ldots, T_j - 1\} \) and \( \lim_{j \to \infty} \text{dist}(\gamma^c(0), \pi_q \circ T_{c,1}) = 0 \).

We set \( \eta_j^c(t) := \gamma_j^c(t - T_j) \), then \( \eta_j^c : (-\infty, T_j] \to \tilde{M} \) is still a calibrated curve and

\[
\eta_j^c(0) \notin V, \quad \eta_j^c(m) \in V, \quad m \in \{1, 2, \ldots, T_j\}
\]
and

$$\lim_{j \to \infty} \text{dist}(\eta_j(T), \pi_q \circ \Upsilon_{c,l}) = 0. \quad (6.8)$$

Extracting a subsequence if necessary, we suppose that \((\eta_j(t), \dot{\eta}_j(t))\) converges uniformly to a limit curve \((\eta^-(t), \dot{\eta}^-(t)) : I \to M\) on any compact interval of \(\mathbb{R}\). Here, the interval \(I\) is either \((-\infty, T]\) or \(\mathbb{R}\) (\(T\) is a positive integer). Obviously, \(\eta^-(t)\) is still calibrated by \(u_{c,l}\) and

$$\eta^-(0) \notin V. \quad (6.9)$$

In the case \(I = (-\infty, T]\), one obtains \(\eta^-(T) \in \pi_q \circ \Upsilon_{c,l}\) as a consequence of (6.8), hence \(\{\eta^-(m)\}_{m \in \mathbb{Z} \subseteq I} \subseteq \pi_q \circ \Upsilon_{c,l}\), which contradicts (6.9). In the case \(I = \mathbb{R}\), it follows from (6.7) that the \(\omega\)-limit set of \(\{\eta^-(m)\}_{m \in \mathbb{Z}}\) lies in \(\pi_q \circ \Upsilon_{c,l}\), then \(\{\eta^-(m)\}_{m \in \mathbb{Z}} \subseteq \pi_q \circ \Upsilon_{c,l}\) since the \(\omega\)-limit set and \(\alpha\)-limit set belong to the same Aubry class, which contradicts (6.9).

**Step 3.** By what we have proved above, for each \(u_{c,l}\)-calibrated curve \(\gamma^-\) with \(\gamma^-(0) \in N_{q,l+1}\), \(\{(\gamma^-(m), \dot{\gamma}^-(-m))\}_{m \in \mathbb{Z}^+}\) would always stay in a small neighborhood of the cylinder \(\mathcal{L}_2 \Sigma_{H,\text{int}}(0)\). This also means that the \(\alpha\)-limit set of \(\{\mathcal{L}^{-1}(\gamma^-(m), \dot{\gamma}^-(-m))\}_{m \in \mathbb{Z}^+}\) lies in \(\Upsilon_{c,l}\). By normal hyperbolicity, \(\{\mathcal{L}^{-1}(\gamma^-(m), \dot{\gamma}^-(-m))\}_{m \in \mathbb{Z}^+} \subset W^{u,\text{loc}}_{\Upsilon_{c,l}}\). Thus, for each \(q \in N_{q,l}\), there is a unique \(u_{c,l}\)-calibrated curve \(\gamma^- : (-\infty, 0] \to M\) with \(\gamma^- (0) = q\) since \(W^{u,\text{loc}}_{\Upsilon_{c,l}}\) is a Lipschitz graph over \(N_{q,l} \subset V\). By weak KAM theory, \(u_{c,l}\) is therefore \(C^{1.1}\) in \(N_{q,l}\).

Moreover, proposition 3.2 implies that

$$(q, c + \partial_q u_{c,l}(q, 0)) = \mathcal{L}^{-1}(\gamma^-(0), \dot{\gamma}^-(0)) \in \mathcal{W}^{u,\text{loc}}_{\Upsilon_{c,l}}.$$  

This completes the proof. \(\square\)

In [24], the authors introduced an ‘area’ parameter \(\sigma\) to parameterize an invariant circle lying on the NHIC so that the invariant circle \(\Gamma_\sigma\) is \(\frac{1}{2}\)-Hölder continuous with respect to \(\sigma\), namely

$$\|\Gamma_{\sigma_1} - \Gamma_{\sigma_2}\|_{C^0} \leq C|\sigma_1 - \sigma_2|^\frac{1}{2}.$$  

However, this result can be improved by taking advantage of the tools in weak KAM theory. Roughly speaking, the ‘area’ parameter \(\sigma\) is, to some extent, the cohomology class \(c\) (see lemma 6.5 and theorem 6.6 below). Similar results could be found in [7].

Recall that the invariant circle \(\Upsilon_{c,l}\), where \(c \in S\) and \(l = 1, u\), can be viewed as a Lipschitz graph over \(q_1\). More precisely, by abuse of notation, we continue to write \(\Upsilon_{c,l}\) for this Lipschitz function

$$\Upsilon_{c,l} : \mathbb{T} \to \Sigma_{H,\text{int}}(0) \subset \mathbb{T}^2 \times \mathbb{R}^2$$

$$q_1 \mapsto (q_1, \pi_{q_1} \circ \Upsilon_{c,l}(q_1), \pi_{p_1} \circ \Upsilon_{c,l}(q_1), \pi_{p_2} \circ \Upsilon_{c,l}(q_1))$$

with \(\pi_{q_1} \circ \Upsilon_{c,l}(q_1) = q_1\) and \(l = 1, u\). Then, we have:
Lemma 6.5 ((\frac{1}{2})-Hölder regularity). There exists a positive constant C such that for any \( c, c' \in S \),

(a) \( \max_{q_1} \| \Upsilon_{c,l}(q_1) - \Upsilon_{c',l}(q_1) \| < C \| c - c' \|^{\frac{1}{2}} \),

(b) \( \max_{q_1} \| \Upsilon_{c,u}(q_1) - \Upsilon_{c',u}(q_1) \| < C \| c - c' \|^{\frac{1}{2}} \).

Proof. We only prove item (a) and the other one is similar. Recall that lemma 6.4 tells us that the elementary weak KAM solution \( u_{c,l} \) is \( C^{1,1} \) in \( N_{r_0,l} \). Now, in the 4-dimensional space \( T^*N_{r_0,l} = N_{r_0,l} \times \mathbb{R}^2 \), we define two 1-forms \( \omega_1 = (c_1 + \partial_1 u_{c,l}(q,0)) dq_1 + \partial_2 u_{c,l}(q,0) dq_2 \) and \( \omega_2 = p_1 dq_1 + p_2 dq_2 \). Note that \( \omega_1 \big|_{\Upsilon_{c,l}} = \omega_2 \big|_{\Upsilon_{c,l}} \) as a consequence of proposition 3.2 and lemma 6.4. Then,

\[
\int_{\Upsilon_{c,l}} \omega_2 = \int_{\Upsilon_{c,l}} \omega_1 = \int_{\Upsilon_{c,l}} [c_1 dq_1 + du_{c,l}(q_1,0)] = \int_{\Upsilon_{c,l}} c_1 dq_1 = c_1.
\]

For \( c, c' \in S \), we may assume \( c' > c_1 \). Let \( D \) be the region on the cylinder \( \Sigma_{H,l}(0) \) between \( \Upsilon_{c,l} \) and \( \Upsilon_{c',l} \) (see figure 4). By Stoke’s theorem,

\[
\int_D \sum_{i=1}^2 dp_i \wedge dq_i = \int_{\Upsilon_{c,l}} \omega_2 - \int_{\Upsilon_{c',l}} \omega_2 = c_1 - c'_1. \tag{6.10}
\]

Then (6.2) and (6.10) together imply

\[
|c_1 - c'_1| = \left| \int_D \sum_{i=1}^2 dp_i \wedge dq_i \right| = \left| \int_D \left( 1 + \frac{\partial(p_2, q_2)}{\partial(p_1, q_1)} \right) dp_i \wedge dq_i \right|
\geq \frac{1}{4} \left| \int_D dp_i \wedge dq_i \right| = \frac{1}{4} \left| \int_{\Upsilon_{c,l}} p_1 dq_1 - \int_{\Upsilon_{c',l}} p_1 dq_1 \right|
= \frac{1}{4} \left| \int_T \pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c',l}(q_1) \ dq_1 \right|. \tag{6.11}
\]
As the Lipschitz functions $\pi_{p_1} \circ \Upsilon_{c,l}$, $\pi_{p_1} \circ \Upsilon_{c,l'} : \mathbb{T} \to \mathbb{R}$ satisfy $\pi_{p_1} \circ \Upsilon_{c,l} > \pi_{p_1} \circ \Upsilon_{c,l'}$, we have

$$
\int_{\mathbb{T}} \pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1) \, dq_1 \geq \frac{1}{4C_L} \left( \max_{q_1} |\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)| \right)^2,
$$

(6.12)

where $C_L$ is the Lipschitz bound of the functions $\pi_{p_1} \circ \Upsilon_{c,l}$ and $\pi_{p_1} \circ \Upsilon_{c,l'}$.

Recall that the function $p_2(q_1, p_1)$ is at least $C^1$, then there exists a constant $K > 0$ such that

$$
\|p_p \circ \Upsilon_{c,l}(q_1) - p_p \circ \Upsilon_{c,l'}(q_1)\| \\
= |\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)| + |\pi_{p_2} \circ \Upsilon_{c,l}(q_1) - \pi_{p_2} \circ \Upsilon_{c,l'}(q_1)| \\
= |\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)| \\
+ |p_2(q_1, \pi_{p_1} \circ \Upsilon_{c,l}(q_1)) - p_2(q_1, \pi_{p_1} \circ \Upsilon_{c,l'}(q_1))| \\
\leq (1 + K)|\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)|.
$$

(6.13)

Thus, combining (6.11), (6.12) with (6.13), one obtains

$$
\|c - c'\| \geq |c_1 - c'_1| \geq \frac{1}{16C_L} \left( \max_{q_1} |\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)| \right)^2 \\
\geq \frac{1}{16C_L(1 + K)^2} \left( \max_{q_1} \|p_p \circ \Upsilon_{c,l}(q_1) - p_p \circ \Upsilon_{c,l'}(q_1)\| \right)^2,
$$

which implies

$$
\max_{q_1} \|p_p \circ \Upsilon_{c,l}(q_1) - p_p \circ \Upsilon_{c,l'}(q_1)\| \leq 4\sqrt{C_L}(1 + K)\|c - c'\|^2.
$$

Next, since the function $q_2(q_1, p_1)$ is at least $C^1$, there exists a constant $\tilde{C} > 0$ such that

$$
\max_{q_1} |q_{2}(q_1, \pi_{p_1} \circ \Upsilon_{c,l}(q_1)) - q_{2}(q_1, \pi_{p_1} \circ \Upsilon_{c,l'}(q_1))| \\
= \max_{q_1} |q_{2}(q_1, \pi_{p_1} \circ \Upsilon_{c,l}(q_1)) - q_{2}(q_1, \pi_{p_1} \circ \Upsilon_{c,l'}(q_1))| \\
\leq C\max_{q_1} |\pi_{p_1} \circ \Upsilon_{c,l}(q_1) - \pi_{p_1} \circ \Upsilon_{c,l'}(q_1)| \\
\leq 4\sqrt{C_L}\sqrt{C_L}(1 + K)\|c - c'\|^2.
$$

Consequently, item (a) follows immediately by setting $C = 4\sqrt{C_L}(1 + \tilde{C})(1 + K)$. \hfill \Box

We also mention that, for the Peierls barriers restricted on the NHIC, one can even obtain the H"older continuity with respect to perturbations [18].

The result below is analogous to [25, lemma 6.4] and will be crucial for the proof of genericity.

**Theorem 6.6.** Let $r_0$ be the constant given in lemma 6.4, and we fix two points $z_l \in N_{r_0,1}, z_o \in N_{r_0,2}$. Let $u_{c,l}^z(q,t)$, $u_{c,l}^{\alpha}(q,t)$ be the elementary weak KAM solutions satisfying
\begin{align*}
\uL{\mathcal{W}}{L}(z, 0) = u(L)^{+}(z, 0) \equiv \text{constant, for all } c \in S. \text{ Then there exists } C_{h} > 0 \text{ such that for any } c, c' \in S \text{,}
\end{align*}
\begin{align*}
|u_{c, d}^{L}(q, 0) - u_{c', d}^{L}(q, 0)| & \leq C_{h}(\|c - c'\|^2 + \|c' - c\|), \quad \forall q \in M\setminus N_{r_{0}}, \\
\text{and}
\end{align*}
\begin{align*}
|u_{c, a}^{L}(q, 0) - u_{c', a}^{L}(q, 0)| & \leq C_{h}(\|c - c'\|^2 + \|c' - c\|), \quad \forall q \in M\setminus N_{r_{0}}.
\end{align*}

\textbf{Remark.} By adding suitably constants, we can take \( u_{c, d}^{L}(z, 0) = u_{c, d}^{L}(z_{a}, 0) = 0 \) for all \( c \in S \), since any elementary weak KAM solution plus a constant is still an elementary weak KAM solution.

\textbf{Proof.} We only prove the case for \( u_{c, 1} \) and the others are similar. The normal hyperbolicity guarantees the smooth dependence of the unstable leaf \( W_{q, \text{loc}}^{u} \) with respect to the base point \( q \in \Sigma_{U}(0) \). By lemma 6.5, the local unstable manifold \( W_{q, l}^{-} \) of \( \Upsilon_{c, l} \) is also \( \frac{1}{2} \)–Hölder continuous in \( c \in S \). Then, lemma 6.4 implies that some constant \( C_{1} > 0 \) exists such that
\begin{align*}
\| (c + \partial_{c} u_{c, 1}^{L}(q, 0)) - (c' + \partial_{q} u_{c', 1}^{L}(q, 0)) \| \leq C_{1} \|c - c'\|^\frac{1}{2}, \quad \forall q \in N_{r_{0}}, \forall c, c' \in S.
\end{align*}
Further, using integration we obtain that for all \( c, c' \in S \) and all \( q \in N_{r_{0}} \),
\begin{align*}
\left| (u_{c, 1}^{L}(q, 0) - u_{c, 1}^{L}(z, 0) + \langle c, q - z \rangle) - (u_{c', 1}^{L}(q, 0) - u_{c', 1}^{L}(z, 0) + \langle c', q - z \rangle) \right|
\leq C_{1} \|c - c'\|^\frac{1}{2}.
\end{align*}
Since we have chosen \( u_{c, 1}^{L}(z, 0) \equiv \text{constant, for all } c \in S \), we get that \( \forall c, c' \in S \) and \( \forall q \in N_{r_{0}} \),
\begin{align*}
|u_{c, 1}^{L}(q, 0) - u_{c', 1}^{L}(q, 0)| \leq C_{1} \|c - c'\|^\frac{1}{2} + \|c - c\|. \quad (6.14)
\end{align*}

Next, for each \( z \in M\setminus N_{r_{0}}, \) there exists a backward calibrated curve \( \gamma_{c, l} \) with \( \gamma_{c, l}(0) = z \), which is negatively asymptotic to \( \pi_{q} \circ \Upsilon_{c, l} \). Since the duration of \( \gamma_{c, l} \) staying outside \( N_{r_{0}}, \) is uniformly bounded, denoted by \( T_{l} \in \mathbb{Z}^{+} \), we have \( \gamma_{c, l}(-k) \in N_{r_{0}} \) for every integer \( k > T_{l} \). Then,
\begin{align*}
\uL{\mathcal{W}}{L}(\gamma_{c, l}(0), 0) - u_{c, l}^{L}(\gamma_{c, l}(-T_{l}), -T_{l})
& = \int_{-T_{l}}^{0} L(\gamma_{c, l}(s), \dot{\gamma}_{c, l}(s), s - \langle c, \dot{\gamma}_{c, l}(s) \rangle + \alpha(c) \) ds, \\
\uL{\mathcal{W}}{L}(\gamma_{c, l}(0), 0) - u_{c', l}^{L}(\gamma_{c, l}(-T_{l}), -T_{l})
& \leq \int_{-T_{l}}^{0} L(\gamma_{c, l}(s), \dot{\gamma}_{c, l}(s), s - \langle c', \dot{\gamma}_{c, l}(s) \rangle + \alpha(c') \) ds.
\end{align*}
Subtracting the first formula from the second one, one deduces from inequality (6.14) that

\[
\begin{align*}
    u_{c,\tilde{c}}(z,0) - u_{\tilde{c},c}(z,0) & \leq u_{c,\tilde{c}}(\gamma_{c,\tilde{c}}(-T), -T) - u_{\tilde{c},c}(\gamma_{\tilde{c},c}(-T), -T) \\
    & + \int_{-T}^{0} \langle c - c', \dot{\gamma}_{c,\tilde{c}}(s) \rangle + \alpha(c') - \alpha(c) \, ds \\
    & \leq u_{c,\tilde{c}}(\gamma_{c,\tilde{c}}(-T), 0) - u_{\tilde{c},c}(\gamma_{\tilde{c},c}(-T), 0) + C_2 \| c' - c \| \\
    & \leq C_1 \| c' - c \|^{{\frac{1}{2}}} + \| c' - c \| + C_2 \| c' - c \|. 
\end{align*}
\]

Here, the second inequality follows from the fact that \( \| \dot{\gamma}_{c,\tilde{c}} \| \) is uniformly bounded and Mather’s \( \alpha \)-function is Lipschitz continuous. So we conclude that there exists \( C_0 > 0 \) such that

\[
    u_{c,\tilde{c}}(z,0) - u_{\tilde{c},c}(z,0) \leq C_0(\| c' - c \|^{{\frac{1}{2}}} + \| c' - c \|), \quad \forall z \in \mathcal{M} \setminus \mathcal{N}_{r_0,u}.
\]

In a similar way, we can prove that \( u_{c,\tilde{c}}(z,0) - u_{\tilde{c},c}(z,0) \leq C_0(\| c' - c \|^{{\frac{1}{2}}} + \| c' - c \|) \) for all \( z \in \mathcal{M} \setminus \mathcal{N}_{r_0,u} \), which completes the proof. \( \square \)

6.3. Choice of the Gevrey space

In what follows, we assume \( \alpha > 1 \) and \( M = \mathbb{T}^2 \). As we will see later, our proof of genericity is not always valid for all Gevrey space \( \mathcal{G}^{\alpha,L} (L > 0) \), but only for \( \mathcal{G}^{\alpha,L} \) with \( L \) bounded by a positive constant \( L_0 \). This is caused by Gevrey approximation, and we will explain it and show how to choose \( L_0 \) below.

Let us first look at the unperturbed Lagrangian \( L_0 = l_1(v_1) + l_2(q_2, v_2) \) in (1.3). For each \( c = (c_1, 0), |c_1| \leq R - 1 \), the Aubry set and Mañé set are

\[
    \mathcal{A}_{L_0}(c, M) = \mathcal{N}_{L_0}(c, M) = \{(q_1, 0, Dh_1(c_1), 0, t) \in TM \times \mathbb{T} : q_1 \in \mathbb{T}, t \in \mathbb{T}\}.
\]
which is a $\delta_{L_{0,\epsilon}}$-invariant circle. Next, we work in the covering space $\tilde{M}$ and consider $L_0 : T\tilde{M} \to \mathbb{R}$. Restricted on the time section $\{t = 0\}$, the lift of the Aubry set has two copies:

$$\tilde{A}_{L_{0,\epsilon}}(c, \tilde{M})|_{t=0} = \{(q_1, 0, Dh_1(c_1), 0) \in T\tilde{M} : q_1 \in \mathbb{T}\},$$

$$\tilde{A}_{L_{0,\epsilon}}(c, \tilde{M})|_{t=0} = \{(q_1, 1, Dh_1(c_1), 0) \in T\tilde{M} : q_1 \in \mathbb{T}\},$$

and they lie on the following two invariant cylinders respectively

$$\mathcal{L}^c\Sigma_0(0) = \{(q_1, 0, Dh_1(p_1), 0) \in T\tilde{M} : q_1 \in \mathbb{T}, \|p_1\| \leq R - 1\},$$

$$\mathcal{L}^c\Sigma_0(0) = \{(q_1, 1, Dh_1(p_1), 0) \in T\tilde{M} : q_1 \in \mathbb{T}, \|p_1\| \leq R - 1\}.$$

Notice that $\pi_q \circ \mathcal{L}^c\Sigma_0(0) = \mathbb{T} \times \{0\}$ and $\pi_q \circ \mathcal{L}^c\Sigma_0(0) = \mathbb{T} \times \{1\}$.

Denote by $u_{c,\epsilon,\Omega}^\pm, u_{c,\epsilon,\Omega}^\pm$ the elementary weak KAM solutions of $L_0$ with respect to the cohomology class $c$. Recall that $\kappa < r_0$. For each $x \in U_{\kappa,\lambda}$, there exists a unique $u_{c,\epsilon,\Omega}^\pm$ calibrated curve $\xi_{c,\epsilon}^\pm(t): (-\infty, 0] \to \tilde{M}$ such that $\xi_{c,\epsilon}^\pm(0) = x$, and it is negatively asymptotic to $A_{L_{0,\epsilon}}(c)$. We pick and fix a constant $T_c = T_c(\kappa, L_0) > 0$ small enough, then we obtain a local neighborhood

$$V_{c,\kappa} = \{\xi_{c,\epsilon}^\pm(t), t \in \tilde{M} \times \mathbb{T} : x \in U_{\kappa,\lambda}, \ -T_c \leq t \leq 0\}$$

which is diffeomorphic to $U_{c,\lambda} \times [-T_c, 0]$ (see figure 5), namely there is a diffeomorphism

$$f : U_{c,\lambda} \times [-T_c, 0] \to V_{c,\kappa}$$

such that $f(x, t) = (\xi_{c,\epsilon}^\pm(t), t)$ and $V_{c,\kappa} \cap (N_{\kappa,\lambda,\kappa,\lambda} \times \mathbb{T}) = \emptyset$, this is guaranteed by $T_c \ll 1$. Notice that $V_{c,\kappa}$ would vary in $c$.

Recall that $\tilde{M} = \mathbb{T} \times (0, 2]/\sim$, where the equivalence relation $\sim$ is defined by identifying $0$ with $2$ in the $q_2$-coordinate. In the sequel, we will fix, once and for all, a sufficiently small constant $\delta > 0$, which is smaller than $\kappa/4$. Thanks to theorem 5.3 there exists a Gevrey-$\gamma_0^\lambda$ diffeomorphism

$$\Psi_{c,\kappa} : U_{c,\lambda} \times [-T_c, 0] \to V_{c,\kappa}$$

such that $\|\Psi_{c,\kappa} - f\|_{C^0(U_{c,\lambda} \times [-T_c, 0])} \leq \delta/2$, where $V_{c,\kappa} \subseteq \mathbb{T} \times (0, 1) \times \mathbb{T}$ and $V_{c,\kappa} \cap (N_{\kappa,\lambda,\kappa,\lambda} \times \mathbb{T}) = \emptyset$, $\lambda_\gamma = \lambda_\gamma(\kappa, L_0) \ll 1$. It means that $\Psi_{c,\kappa}(x, \cdot)$ remains $\delta/2$--close to $\xi_{c,\epsilon}^\pm(\cdot)$ in the following sense:

$$\text{dist}(\Psi_{c,\kappa}(x, t), \xi_{c,\epsilon}^\pm(t)) \leq \delta/2, \quad \forall (x, t) \in U_{c,\lambda} \times [-T_c, 0].$$

Recall that the number $\epsilon_1$ given in lemma 6.3 is small enough, then one can find an small interval $I_c = \{(c_1', 0) : c_1' \in (c_1 - \tau, c_1 + \tau)\}$ depending on $\kappa, L_0$, such that if the perturbation term $\|L_{c_1}\|_{c_2} < 2\epsilon_1$, then the Lagrangian $L = L_0 + L_1$ satisfies; for each $c_1' \in I_c, x \in U_{c,\kappa}$,

- the $u_{c_1',\epsilon,\Omega}^\pm$-calibrated curve $\gamma_{x, c_1',\epsilon,\Omega}(t): (-\infty, 0] \to \tilde{M}$ with $\gamma_{x, c_1',\epsilon,\Omega}(0) = x$ is negatively asymptotic to $A_{L_{1,\epsilon}}(c, \tilde{M})$,
- $\gamma_{x, c_1',\epsilon,\Omega}(\cdot)$ is still $\delta$--close to $\Psi_{c,\kappa}(x, \cdot)$ in the sense that

$$\text{dist}(\Psi_{c,\kappa}(x, t), \gamma_{x, c_1',\epsilon,\Omega}(t)) \leq \delta, \quad \forall -T_c \leq t \leq 0.$$

(6.15)
These properties are guaranteed by the upper semi-continuity. By the finite covering theorem, there exist finitely many intervals \( \{I_i\}_{i=0}^n \) such that
\[
\bigcup_{0 \leq i \leq m} I_i \supset [-R + 1, R - 1] \times \{0\},
\]
and the corresponding diffeomorphism \( \Psi_{c,l} : U_{c,l} \times [-T_{c,l}, 0] \to V_{c,l} \) is Gevrey-\((\alpha, \lambda, \lambda')\), and the positive number \( T_{c,l} \ll 1 \). According to theorem 5.2, some constant \( \lambda' < \lambda_c \) exists such that \( \Psi_{c,l}^{-1} \) is Gevrey-\((\alpha, \lambda, \lambda')\) smooth.

In what follows, we set
\[
L_0 := \min\{\chi_i' : i = 0, \ldots, m\},
\]
and the corresponding diffeomorphism \( \Psi_{c,l}^{-1} : V_{c,l} \to U_{c,l} \times [-T_{c,l}, 0] \)
is Gevrey-\((\alpha, L_0)\) smooth, for all \( L \leq L_0 \). We also point out that \( L_0 \) is independent of the perturbation Hamiltonian \( H_1 \), and it depends only on \( H_0, R \) and \( \alpha \) since the choice of \( \kappa \) depends only on \( H_0 \) and \( R \), and \( L_0 \) depends only on \( H_0 \).

Similarly, these procedures can be carried out for the region \( U_{c,u} \), and one can get the corresponding Gevrey diffeomorphism \( \Psi_{c,u} : U_{c,u} \times [-T_{c,u}, 0] \to V_{c,u} \). For simplicity, we will assume the same interval covering \( \bigcup_{i=0}^m I_i \) as (6.16) and \( \Psi_{c,u}^{-1} : V_{c,u} \to U_{c,u} \times [-T_{c,u}, 0] \) is Gevrey-\((\alpha, L)\) smooth for all \( L \leq L_0 \), where each \( \Psi_{c,u}^{-1}(i = 0, \ldots, m) \) possesses the property analogous to (6.15).

### 6.4. Total disconnectedness

Let \( \alpha > 1 \), we will study the topological structure of the set of minimal points for
\[
R_{c,l,u}(x, \tau) = u_{c,l}^u(x, \tau) - u_{c,l}^{\pm}(x, \tau), \quad \text{and} \quad R_{c,u}(x, \tau) = u_{c,u}^u(x, \tau) - u_{c,u}^{\pm}(x, \tau)
\]
defined in (3.5), where \( u_{c,l}^\pm \) (\( i = 1, \ldots, m \)) are the elementary weak KAM solutions. Actually, we will show that the minimal set is totally disconnected for generic Lagrangian systems. Inspired by the technique in [20, section 4.2], we will perturb directly a Lagrangian by small potential functions. Compared with the perturbative techniques used in [24, 25] which perturb the generating functions to create genericity, our technique in the current paper provides more information, we can prove the genericity not only in the usual sense but also in the sense of Mane.

Let \( L = L_0 + I_0 \) be our Lagrangian given in (1.3), where \( \|L_1\|_2 < \varepsilon_1 \). Recall the interval covering \( \bigcup_{0 \leq i \leq m} I_i \) given in section 6.3, one can always suppose that the length of each interval \( I_i \) is less than 1. Then theorem 6.6 implies that for any \( c, c' \in I_i \cap S \) and \( q \in U_\kappa = U_{c,l} \cup U_{c,u} \),
\[
|u_{c,l}^\pm(q, 0) - u_{c,l}^{\pm}(q, 0)| \leq 2C_h\|c' - c\|^2, \quad |u_{c,u}^\pm(q, 0) - u_{c,u}^{\pm}(q, 0)| \leq 2C_h\|c' - c\|^2.
\]
(6.18)

Fixing \( L \in (0, L_0] \) and \( \varepsilon_0 \in (0, \varepsilon_1) \), we consider the following set in \( \mathbb{G}^{n,L}(M \times T) \) with \( M = \mathbb{T}^2 \):
\[
\mathcal{P} := \{ P \in \mathbb{G}^{n,L}(M \times T) : \|P\|_{\alpha,L} < \varepsilon_0, \supp P \cap (\pi N_{\varepsilon_1/2} \times T) = \emptyset \}.
\]
(6.19)
Then it is easily seen that each potential perturbation $P \in \mathcal{V}$ to the Hamiltonian $H$ would not affect the NHIC since $\Sigma_H \subset \Sigma \subset \mathbb{R}^2 \times \mathbb{T}$, see (6.4). We also point out that, by a natural extension, any function in $C^{1, L}(M \times \mathbb{T})$ can be viewed as a function defined on $M \times \mathbb{T}$.

**Theorem 6.7.** Let $\alpha > 1$, $L \leq L_0$. There exists a residual set $\mathcal{W} \subset \mathcal{V}$ such that for each Gevrey potential function $P \in \mathcal{W}$, the Lagrangian $L + P : TM \times \mathbb{T} \to \mathbb{R}$ satisfies: for each $c \in \mathcal{S}$, the sets

$$\arg \min B_{c, \alpha} |U_{c, \alpha} \cup U_{c, a}$$

are both totally disconnected. Here, $\arg \min B_{c, \alpha} |U_{c, \alpha} \cup U_{c, a}$ stands for $\arg \min B_{c, \alpha} \cap (U_{c, \alpha} \cup U_{c, a})$, and $\arg \min B_{c, \alpha} |U_{c, \alpha} \cup U_{c, a}$ stands for $\arg \min B_{c, \alpha} \cap (U_{c, \alpha} \cup U_{c, a})$.

**Proof.** For $c \in \mathcal{S}$, we first study the set $\arg \min B_{c, \alpha}$ restricted on the region $U_{c, \alpha} \subset \tilde{M} \times \{t = 0\}$. Let $r_0 > 0$ be the constant given in lemma 6.4. Recall that we have $r_0 > \kappa$. Then, to prove the total disconnectedness of $\arg \min B_{c, \alpha}$, it is enough to verify that

the set $\arg \min B_{c, \alpha} |N_{r_0} \setminus N_{r_0}$ is totally disconnected, (6.20)

where $N_{r_0} \setminus N_{r_0}$ is the closure of $N_{r_0}$. To explain this, we recall that, according to proposition 3.6, a point $(x, 0) \in \arg \min B_{c, \alpha}$ if and only if there exists a $c$-semi static curve $\gamma_{x, c} : \mathbb{R} \to \tilde{M}$ with $\gamma_{x, c}(0) = x$, such that the orbit $\{\gamma_{x, c}(n) : n \in \mathbb{Z}\}$ is negatively asymptotic to $\pi_\gamma \circ \Upsilon_{c, \alpha}$ and positively asymptotic to $\pi_\gamma \circ \Upsilon_{c, 0}$. Then, by letting $\kappa$ be suitably small if necessary, the orbit $\{\gamma_{x, c}(n) : n \in \mathbb{Z}\}$ has to pass through the region $N_{r_0} \setminus N_{r_0}$ when it approaches $\pi_\gamma \circ \Upsilon_{c, 1}$, as a result of the normal hyperbolicity. Consequently, in what follows, we only need to check (6.20).

We first focus on the subinterval $I_{ \leq c}$ in the interval covering $\bigcup_{0 \leq c \leq c} I_c$. Let us pick a 2-dim disk

$$D = \{(x_1, x_2, t) \in \tilde{M} \times \mathbb{T} : t = 0, |x_1 - x_1| \leq d, |x_2 - x_2| \leq d\} \subset N_{r_0} \setminus N_{r_0}$$

which is centered at the point $(x_1, x_2, t)$, and $d$ is small. We also set

$$D + d_1 = \{(x_1, x_2, t) \in \tilde{M} \times \mathbb{T} : t = 0, |x_1 - x_1| \leq d + d_1, |x_2 - x_2| \leq d + d_1\} \subset N_{r_0} \setminus N_{r_0}$$

with $0 < d_1 \ll 1$ (see figure 5).

Now let $\mu$ be suitably small, for the index $i = 1$ or 2, we consider the following space

$$\mathcal{V}_i := \left\{ \mu \left( \sum_{i=1,2} a_{i, \ell} \cos 2\ell \pi(x_i - x_{i,0}) + b_{i, \ell} \sin 2\ell \pi(x_i - x_{i,0}) \right) : a_{i, \ell}, b_{i, \ell} \in [1, 2] \right\}.$$

(6.21)

Obviously, $\mathcal{V}_1, \mathcal{V}_2 \subset C^\infty(M)$. Next, we will construct perturbations based on potential functions of the form in $\mathcal{V}_i$. To this end, we use the notation given in section 6.3. Fixing a sufficiently large constant $\mathcal{E} \gg L$, by lemma 5.1 one can construct a function $\rho(x, t) = g(x) \chi(t) : \tilde{M} \times \mathbb{T} \to \mathbb{R}$ such that $\chi : \mathbb{T} \to \mathbb{R}$ and $g(x) : \tilde{M} \to \mathbb{R}$ are both non-negative Gevrey-$\alpha$ functions. We choose

$$\chi(t) = \begin{cases} > 0, & t \in (-T_\alpha, 0) \\ = 0, & t \in T \setminus (-T_\alpha, 0) \end{cases}$$

490
where $T_{\rho} \ll 1$ is given in section 6.3, and require that $g|_{D} \equiv 1$ and $\operatorname{supp} g \subset D + d_{1} \subset N_{\rho,1} \setminus N_{\rho,2}$. We set

$$
C := \{ \Psi_{\rho,1}(x, t) \mid (x, t) \in (D + d_{1}) \times [-T_{\rho}, 0] \},
$$

then $C \subset V_{\rho,1} \subseteq T \times (0, 1) \times T$, and therefore $C \cap (N \times 2 \times T) = \emptyset$.

- With each $V \in \mathfrak{U}_{1}$ or $\mathfrak{U}_{2}$, which can also be viewed as a function on $M$, one can define $\tilde{V} \in C^{\infty}(M \times T)$ as follows: on the ‘lower’ domain $T \times [0, 1] \times T \subset M \times T$,

$$
\tilde{V}(z) = \begin{cases} 
(\rho V) \circ \Psi_{\rho,1}^{-1}(z) = \rho(x, t)V(x), & z \in C, \\
0, & z \in (T \times [0, 1] \times T) \setminus C.
\end{cases}
$$

Then we extend symmetrically the function to the ‘upper’ domain $T \times [1, 2] \times T$ such that

$$
\tilde{V}(y, t) = \tilde{V}(y - e_{2}, t)
$$

with $e_{2} = (0, 1)$. The support of $\tilde{V}$ satisfies

$$
\operatorname{supp} \tilde{V} \subset C \cup (C + e_{2}). \tag{6.22}
$$

Since $\mathfrak{E} \gg \mathfrak{L}$, according to the properties (G1), (G3) in section 1, we have

$$
\tilde{V} \in G^{0,1}(\hat{M} \times T). \tag{6.23}
$$

- We also remark that, by the symmetry of $\tilde{V} \in G^{0,1}(\hat{M} \times T)$, $\tilde{V}$ can also be viewed as a function on $M \times T$. By abuse of notation, we continue to write $\tilde{V} \in G^{0,1}(\hat{M} \times T)$. Thus, $\tilde{V} \in \mathfrak{V}$. As a result of the construction above, some constant $C_{1} > 0$ exists such that

$$
\int_{-T_{\rho}}^{0} \tilde{V}(\Psi_{c,1}(x, t)) dt = V(x) \int_{-T_{\rho}}^{0} g(x) \chi(t) dt 
$$

$$
= V(x) \int_{-T_{\rho}}^{0} \chi(t) dt = C_{1} V(x), \quad \text{for } x \in D. \tag{6.24}
$$

Here, we have used the fact $g|_{D} \equiv 1$.

Let $\Pi_{i}$, $i = 1, 2$, be the standard projection to the $i$th coordinate of $\hat{M}$. For the Lagrangian $L : TM \times T \rightarrow \mathbb{R}$, we consider two elementary weak KAM solutions $u_{c,1}^{-}(q, t), u_{c,1}^{+}(q, t)$, and denote by $u_{c,1}^{-}(q, t), u_{c,1}^{+}(q, t)$ the elementary weak KAM solutions of the perturbed Lagrangian $L + \tilde{V}$. Then the following result holds:

**Lemma 6.8.** There exists an open and dense set $\mathcal{U}_{D} \subset \mathfrak{V}$ (see (6.19)) such that for each $\tilde{V} \in \mathcal{U}_{D}$,

$$
\Pi_{i} \left( \arg \min \left( u_{c,1}^{-}(x, 0) - u_{c,1}^{+}(x, 0) \right) |_{D} \right) \subseteq [x_{i,0} - d, x_{i,0} + d],
$$

for all $c \in I_{0} \cap \mathbb{S}$, \tag{6.25}

where $i = 1, 2$. 491
Proof. We start with the perturbation \( \tilde{V} \) of the form (6.23), where \( V \in \mathcal{D}_1 \cup \mathcal{D}_2 \). Note that under such a potential perturbation, the cylinders \( \Sigma_{H,I}(0) \) and \( \Sigma_{H,0}(0) \) remain unchanged, hence the Aubry set \( \mathcal{A}_{l,V}^0(c,M) = \mathcal{A}_l(c,M) \).

Step 1. For \( c \in I_0 \cap \mathbb{S} \), the projected Aubry set \( \mathcal{A}_l(c,M) \subset M \times \mathbb{T} \) has two copies, denoted by \( \mathcal{A}_{l,1}(c,M) \) and \( \mathcal{A}_{l,2}(c,M) \). Each set \( \mathcal{A}_{l,i}(c,M) \), \( i = 1, 2 \) is diffeomorphic to \( \mathbb{T}^2 \) since \( c \in \mathbb{S} \). For each \( x \in D \) and each \( u^+_{c,x,V} \)-calibrated curve \( \gamma^+_x(t) : [0, +\infty) \to M \) with \( \gamma^+_x(0) = x \), the minimizing curve \( (\gamma^+_x(t), t) : \mathbb{R}^+ \to M \times \mathbb{T} \) is positively asymptotic to \( \mathcal{A}_{l,0}(c,M) \). Now, we claim that

\[
\text{supp} \tilde{V} \cap \left( \bigcup_{t > 0} (\gamma^+_x(t), t) \right) = \emptyset,
\]

as long as \( D \) is small enough. In fact, according to (6.22), the support of \( \tilde{V} \) has two copies in the lower and upper region respectively and \( \text{supp} \tilde{V} \subset C \cup (C + \mathbf{e}_2) \). It is clear that the minimizing orbit \( (\gamma^+_x(t), t) \) never intersects itself, then \( C \cap (\bigcup_{t > 0} (\gamma^+_x(t), t)) = \emptyset \) since \( D \) is a small neighborhood of \( x \). Moreover, observe that the sets \( \mathcal{A}_{l,1}(c,M) \) and \( \mathcal{A}_{l,2}(c,M) \), which are diffeomorphic to \( \mathbb{T}^2 \), divide the 3-dimensional configuration space \( M \times \mathbb{T} \) into two connected components, then the minimizing curve \( (\gamma^+_x(t), t) : \mathbb{R}^+ \to M \times \mathbb{T} \) always stays in the lower region, which means \( (C + \mathbf{e}_2) \cap (\bigcup_{t > 0} (\gamma^+_x(t), t)) = \emptyset \). This proves our claim (6.26).

Consequently,

\[
u^+_{c,x,V}(x) = u^+_{c,x}(x), \quad \text{for all } x \in D.
\]

But the function \( u^+_{c,x,V} \) would undergo a small perturbation. Indeed, for \( x \in D \), we can take a \( u^-_{c,x,V} \)-calibrated curve \( \gamma^-_x : (-\infty, 0] \to M \) with \( \gamma^-_x(0) = x \), then for \( m \in \mathbb{Z}^+ \),

\[
u^+_c(\gamma^-_x(0), 0) - u^-_{c,x,V}(\gamma^-_x(-m), -m) \leq \int_{-m}^{0} (L - \eta_c + \tilde{V})(d\gamma^-_x(t), t) + \alpha(c) \, dt.
\]

(6.28)

For another perturbation \( \tilde{V} \), we have

\[
u^+_c(\gamma^-_x(0), 0) - u^-_{c,x,V}(\gamma^-_x(-m), -m) \leq \int_{-m}^{0} (L - \eta_c + \tilde{V})(d\gamma^-_x(t), t) + \alpha(c) \, dt.
\]

(6.29)

By normal hyperbolicity, there exists a uniform upper bound \( T \in \mathbb{Z}^+ \), \( T > T_0 \), such that the orbit \( \{ (\gamma^-_x(t), t) \}_{t \geq T} \) shall retreat into the small neighborhood \( N_{s/2,1} \times \mathbb{T} \). As the supports of \( \tilde{V} \) and \( \tilde{V}' \) have empty intersection with \( N_{s/2,1} \times \mathbb{T} \), we have \( u^-_{c,L,V} = u^-_{c,L,V'} \) on \( N_{s/2,1} \times \mathbb{T} \). Then (6.28) and (6.29) imply that

\[
u^+_{c,L,V}(x, 0) - u^-_{c,L,V}(x, 0) \leq \int_{-T}^{0} (\tilde{V}' - \tilde{V})(\gamma^-_x(t), t) \, dt.
\]

Conversely, we can prove similarly that

\[
u^-_{c,L,V}(x, 0) - u^+_{c,L,V}(x, 0) \geq \int_{-T}^{0} (\tilde{V}' - \tilde{V})(\gamma^-_x(t), t) \, dt.
\]
where $\gamma_{x,-V}^-$ denotes the backward $u_{c,+V}^-$-calibrated curve with $\gamma_{x,-V}^-(0) = x$. Since $x$ lies in the region $D$ where $u_{c,+V}$ is differentiable (see lemma 6.4), one has $\|\gamma_{x,-V}^-(t) - \gamma_{x,-V}^+(t)\| \to 0$ as $\|\tilde{V}' - \tilde{V}\| \to 0$, which is guaranteed by the upper semi-continuity. Therefore, for $c \in I_{\delta} \cap S$,

$$u_{c,+V}(x, 0) - u_{c,+V}(x, 0) = \mathcal{K}_{x,\delta}(\tilde{V}' - \tilde{V})(x) + \mathcal{R}_c(\tilde{V}' - \tilde{V})(x), \quad x \in D,$$

(6.30)

where the operator

$$\mathcal{K}_{x,\delta}(\tilde{V}' - \tilde{V})(x) = \int_{-T}^0 (\tilde{V}' - \tilde{V}) (\gamma_{x,-V}^-(t), t) \, dt,$$

(6.31)

and the remainder

$$\mathcal{R}_c(\tilde{V}' - \tilde{V}) = o(\|V' - V\|_{c_0})$$

since $V, V' \in \mathcal{U}_1 \cup \mathcal{U}_2$ are linear combinations of trigonometric functions.

**Step 2.** Now, we claim that there exists an arbitrarily small perturbation $\tilde{V} \in \mathcal{U}$ of the form (6.23), such that

$$\Pi_1 \left( \arg \min \left( u_{c,+V}(x, 0) - u_{c,+V}(x, 0) \right) \right) \subseteq [x_{1,0} - d, x_{1,0} + d],$$

(6.32)

for all $c \in I_{\delta} \cap S$.

To prove this claim, we construct a grid for the parameters $(a_i, b_i, a_{1,2}, b_{1,2})$ in $\mathcal{U}_1$ by splitting the domain $[1, 2]^4$ equally into a family of 4-dimensional cubes whose side length is $\mu^2$, namely

$$\Delta a_{1,\ell} = \Delta b_{1,\ell} = \mu^2, \quad \ell = 1, 2.$$

There are as many as $[\mu^{-8}]$ cubes.

In the sequel, we use the symbol $\text{Osc}_{x \in D} f$ to denote the oscillation of $f$, which describes the difference between the supremum and infimum of $f$ on $D$. According to (6.15), for each $c \in I_{\delta} \cap S$ and $x \in D$, the backward $c$-semi static curve $\gamma_{x,+V}(t)$ will stay in the $\delta$-neighborhood of the curve $\Psi_{c,\delta}(t)$ for $t \in [-T_c, 0]$ provided that $\mu$ is small enough. Besides, since minimizing orbits have no self intersections, by letting $D$ be suitably small if necessary, the minimizing curve

$$\left( \gamma_{x,+V}(t), t \right): [-T_c, -T_c] \to M \times \mathbb{T}$$

does not intersect the support of $\tilde{V}' - \tilde{V}$. This is guaranteed by using arguments similar to the proof of (6.26). Therefore, (6.24) and the above estimate imply that

$$\mathcal{K}_{x,\delta}(\tilde{V}' - \tilde{V})(x) = \int_{-T_c}^0 (\tilde{V}' - \tilde{V})(\Psi_{c,\delta}(x, t)) \, dt + O(\mu \delta)$$

$$= C_1 (V'(x) - V(x)) + O(\mu \delta).$$

(6.33)

Then some constant $C_2 > 0$ exists such that

$$\text{Osc}_{x \in D}(\mathcal{K}_{x,\delta}(\tilde{V}' - \tilde{V})) > \frac{1}{4} C_1 \text{Osc}_{x \in D}(V - V') > C_2 \mu \Delta$$

(6.34)
where $\Delta = \max \{|a_{1,\ell} - a_{1,\ell}'|, |b_{1,\ell} - b_{1,\ell}'|: \ell = 1, 2\}$. This is guaranteed by (6.33) and the fact that $V$ is a finitely linear combination of $\{\sin 2\pi x_1, \cos 2\pi x_1, \sin 4\pi x_1, \cos 4\pi x_1\}$.

Next, we split the interval $L_{0\ell}$ into $|K_{\ell}|^{-6}$ subintervals, where $K_{\ell} = L_{0\ell} \frac{24C_6}{\mu}$ and $L_{0\ell} = |I_{0\ell}|$. $C_6$ is the constant given in (6.18) and $C_2$ is given in (6.34). We pick up the subinterval that has non-empty intersection with $S$, and then denote all such kinds of subintervals by $\{J_i\}_{i \in \mathbb{Z}}$. Clearly, the cardinality of $\mathbb{Z}$ is less than $|K_{\ell}|^{-6}$.

Let us fix a $c^* \in J_i \cap S$. If for some parameter $(a_{1,\ell}', b_{1,\ell}')$, $\ell = 1, 2$ and its corresponding perturbation $V^* \in \mathcal{Q}_1$, formula (6.32) does not hold, then $\min_{i \in \mathbb{Z}} \left\{ u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right\}$ is identically equal to a constant for all $x \in D$, hence

$$\text{Osc}_{c \in D} \min_{i \in \mathbb{Z}} \left\{ u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right\} = 0. \quad (6.35)$$

Next, for another $V' = \mu \left( \sum_{\ell=1,2} a_{1,\ell}' \cos 2\ell\pi(x_1 - x_{1,0}) + b_{1,\ell}' \sin 2\ell\pi(x_1 - x_{1,0}) \right) \in \mathcal{Q}_1$, and the corresponding perturbation $V''$, it follows from (6.27) and (6.30) that for all $c \in J_i \cap S$ and $x \in D$,

$$\left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right) - \left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right)$$

$$= \left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right) - \left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right)$$

$$+ (\mathcal{H}_{c} V' + \mathcal{R}_{c}) (V'' - V^*) (x). \quad (6.36)$$

As the length of $J_i$ is $\frac{|I_{0\ell}|}{|K_{\ell}|^{-6}}$ and $c, c^* \in J_i \cap S$, formula (6.18) then implies that

$$\left| \left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right) - \left( u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right) \right|$$

$$\leq 4C_6 \|c - c^*\|^{\frac{1}{2}} \leq 4C_6 \left( \frac{L_{0\ell}}{|K_{\ell}|^{-6}} \right)^{\frac{1}{2}} \leq C_2 \mu^3 \left( \frac{C_6}{6} \right). \quad (6.37)$$

Since $\mu \ll 1$, one has $\|V'' - V^*\| \ll 1$ and

$$\|\mathcal{R}_{c^*} (V'' - V^*)\| \leq \frac{1}{6} \|.\mathcal{H}_{c^*} (V'' - V^*)\|. \quad (6.38)$$

Regarding the potential function $V'$, if its parameter $(a_{1,\ell}', b_{1,\ell}')$, $\ell = 1, 2$ satisfies

$$\max \{|a_{1,\ell}' - a_{1,\ell}|, |b_{1,\ell}' - a_{1,\ell}'|: \ell = 1, 2\} = \mu^2, \quad (6.39)$$

then inequalities (6.34)–(6.38) together give rise to

$$\text{Osc}_{c \in D} \min_{i \in \mathbb{Z}} \left\{ u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right\} \geq C_2 \frac{\mu^3}{3} > 0. \quad (6.40)$$

So we can conclude that for each $c \in J_i \cap S$ and $V' \in \mathcal{Q}_1$ satisfying (6.39), we have

$$\text{Osc}_{c \in D} \min_{i \in \mathbb{Z}} \left\{ u_{c_{1,\ell},\ell}^-(x, 0) - u_{c_{1,\ell},\ell}^+(x, 0) \right\} > 0. \quad (6.40)$$

Consequently, for each $J_i$, we only need to cancel out at most $2^4$ cubes from the grid $\{\Delta a_{1,\ell}, \Delta b_{1,\ell}: \ell = 1, 2\}$ so that formula (6.40) holds for all other cubes. Let the index $i$ range
over \( J \), we therefore obtain a set \( P_1 \subseteq \{(a_{1,1}, a_{1,2}, b_{1,1}, b_{1,2}) : a_{1,\ell}, b_{1,\ell} \in [1,2], \ell = 1,2\} \) with Lebesgue measure

\[
\text{meas} P_1 \geq 1 - 2^4(\mu^2)^j |J| \geq 1 - 2^4K_1 \mu^2 > 0,
\]

such that formula (6.40) holds for any \( V' \) with parameter in \( P_1 \) and any \( c \in I_0 \cap \mathbb{S} \). As \( \mu \) is small enough, the claim (6.32) is now evident from what we have proved.

**Step 3.** Actually, the arguments above in step 2 also ensure that formula (6.32) has density in \( \mathbb{P} \). The openness is obvious, so there is an open-dense set \( U_{D,1} \) in \( \mathbb{P} \) such that formula (6.32) holds for each perturbed Lagrangian \( L + \tilde{V} \) with \( \tilde{V} \in U_{D,1} \).

Analogously, we can consider a potential function \( V \in \mathbb{P}_2 \) and its associated perturbation \( \tilde{V} \). By repeating similar arguments as in step 2, we also obtain an open-dense set \( U_{D,2} \subset \mathbb{P} \), such that for each perturbed Lagrangian \( L + \tilde{V} \) where \( \tilde{V} \in U_{D,2} \),

\[
\Pi_2 \left( \arg \min \left( u_{i,k}^{-}(x,0) - u_{i,n}^{+}(x,0) \right) \right) \subseteq [x_{2,0} - d, x_{2,0} + d],
\]

for all \( c \in I_0 \cap \mathbb{S} \). Thus, the proof of lemma 6.8 is now completed by taking a set \( U_D = U_{D,1} \cap U_{D,2} \). \( \square \)

Now we continue to prove theorem 6.7.

- From lemma 6.8 we see that for each small disk \( D \subseteq \mathbb{N}_{e1} \setminus N_{2,1} \), there exists an open-dense set \( U_D \subset \mathbb{P} \) such that formula (6.25) holds for each perturbed Lagrangian \( L + \tilde{V} \) with \( \tilde{V} \in U_D \).

Now, we take a countable topology basis \( \bigcup \{D_j \} \) for \( \mathbb{N}_{e1} \setminus N_{2,1} \) where the diameter of \( D_j \) approaches to 0 as \( j \to \infty \), and therefore obtain an open-dense set \( U_D \) for each \( j \). Clearly, \( U_{D,j} = \bigcap \{D_j \} \) is a residual set in \( \mathbb{P} \), and the set \( \arg \min \left( u_{c,i,j}^{-}(x,0) - u_{c,a,p}^{+}(x,0) \right) \big|_{\mathbb{N}_{e1} \setminus N_{2,1}} \) is totally disconnected for each \( P \in U_{D,j} \) and \( c \in \mathbb{S} \cap I_0 \).

The technique above also works for other subintervals \( I_i, i = 1, \ldots, m \), we can then obtain the corresponding residual sets \( U_{D,i} \), \( i = 1, \ldots, m \). So the intersection \( U_D = \bigcap_{i=0}^{m} U_{D,i} \) is residual, and \( \arg \min \left( u_{c,i,j}^{-}(x,0) - u_{c,a,p}^{+}(x,0) \right) \big|_{\mathbb{N}_{e1} \setminus N_{2,1}} \) is totally disconnected for each \( P \in U_D \) and \( c \in \mathbb{S} \). By what have shown at the beginning of the proof, this is equivalent to saying that \( \arg \min \left( u_{c,i,j}^{-}(x,0) - u_{c,a,p}^{+}(x,0) \right) \big|_{U_{2,1}} \) is totally disconnected for each \( P \in U_D \) and \( c \in \mathbb{S} \).

- Similarly, one can prove that there exists a residual set \( U_a \subset \mathbb{P} \), such that the set

\[
\arg \min \left( u_{e,c,i}^{-}(x,0) - u_{e,c,a,p}^{+}(x,0) \right) \big|_{U_{a,c}}
\]

is totally disconnected for each \( P \in U_a \) and \( c \in \mathbb{S} \).

- Conversely, by applying the technique above to \( u_{c,a,p}^{+}(x,0) - u_{e,c,i}^{-}(x,0) \), we can also obtain two residual sets \( V_l \) and \( V_u \) in \( \mathbb{P} \), such that the set \( \arg \min \left( u_{c,a,p}^{+}(x,0) - u_{e,c,i}^{-}(x,0) \right) \big|_{U_{2,1}} \) is totally disconnected for each \( c \in \mathbb{S} \) and \( P \in V_l \), and the set \( \arg \min \left( u_{c,a,p}^{+}(x,0) - u_{e,c,i}^{-}(x,0) \right) \big|_{U_{a,c}} \) is totally disconnected for each \( c \in \mathbb{S} \) and \( P \in V_u \).

Therefore, the proof of theorem 6.7 is now completed by taking \( V = U_l \cap U_a \cap V_l \cap V_u \). \( \square \)
6.5. Proof of theorems 1.2 and 1.3

Now, we are able to prove our main results. Let \( R > 1, \alpha > 1 \) and \( 0 < L \leq L_0 = L_0(H_0, \alpha, R) \), where the constant \( L_0 \) is given in (6.17) and independent of \( H_1 \).

**Proof.** In our problem \( M = \mathbb{T}^2 \), \( s > 0, \gamma_\ell \in [-R + 1, R - 1] \times \{ 0 \}, \ell \in \{ 1, \ldots , k \} \). Let \( \varepsilon_0 \) be a small positive number satisfying

\[
\varepsilon_0 < \min \left\{ 1, L^0, \frac{L^0}{2l\alpha}, \frac{L^3\alpha}{3l\alpha} \right\} \varepsilon_1,
\]

where \( \varepsilon_1 \) is chosen as in lemma 6.3. Now, \( \| H_1 \|_{o_L} < \varepsilon_0 \) implies \( \| H_1 \|_{c_0} < \varepsilon_1 \), hence the Hamiltonian \( H = H_0 + H_1 \) has a persistent NHIC, and the globally minimal set \( \tilde{G}_1(c) \) lies in this NHIC for each \( c = (c_1, 0) \) with \( |c_1| \leq R - 1 \). Here, \( L = L_0 + L_1 \) is the Lagrangian associated to \( H = H_0 + H_1 \). Also, observe that for the Mañé set \( \tilde{N}_{L_0}(c) \),

\[
\pi_p \circ \mathcal{L}^{-1}(\tilde{N}_{L_0}(c)) = c, \quad \text{for each } c = (c_1, 0), |c_1| \leq R - 1. \quad (6.41)
\]

By letting \( \varepsilon_0 \) be small enough, the perturbation term \( L_1 \) is also sufficiently small, then the upper semi-continuity (see proposition 2.1) implies that for each \( c = (c_1, 0), |c_1| \leq R - 1, \)

\[
d(\pi_p \circ \mathcal{L}^{-1}(\tilde{N}_{L_0}(c)), \pi_p \circ \mathcal{L}^{-1}(\tilde{N}_{L_0}(c))) < s/2, \quad \text{whenever } \| H_1 \|_{o_L} < \varepsilon_0 \quad (6.42)
\]

With this fact, one can find that \( \varepsilon_0 = \varepsilon_0(H_0, \alpha, R, s, L) > 0 \) depends on the Hamiltonian \( H_0 \) and the constants \( \alpha, R, L \) and \( s \).

**Density.** For each Hamiltonian \( H_0 + H_1 \) with \( \| H_1 \|_{o_L} < \varepsilon_0 \), we will prove that there exists an arbitrarily small perturbation \( V \in \mathcal{G}^{o_L}(M \times \mathbb{T}) \) such that \( \| H_1 + V \|_{o_L} < \varepsilon_0 \), and the perturbed Hamiltonian \( H_0 + H_1 + V \) has an orbit \((q(t), p(t))\) and times \( t_1 < \cdots < t_k \) such that the action variables \( p(t) \) pass through the ball \( B_{\epsilon}(\gamma) \) at the time \( t = t_i \). To this end, we will establish a generalized transition chain along which one is able to apply theorem 4.3.

Let \( d \in (0, \varepsilon_0 - \| H_1 \|_{o_L}) \) be arbitrarily small.

• First, by applying the genericity property in corollary 6.2 to the Tonelli Lagrangian \( L_0 + L_1 \), one can always choose a small perturbation

\[
\phi \in \mathcal{G}^{o_L}(M \times \mathbb{T}), \quad \| \phi \|_{o_L} < \frac{d}{2}
\]

such that for each rational homology class \( h = (\ell, 0) \), the perturbed Lagrangian \( L_0 + L_1 + \phi \) has only one minimal measure with the rotation vector \( h \). Next, for the irrational case, it is well known in Aubry–Mather theory that for homology \( h = (h_1, 0) \) with \( h_1 \in \mathbb{R} \setminus \mathbb{Q} \), only one minimal measure with the rotation vector \( h \) exists. So it is easily seen that the Aubry class is unique for each \( c = (c_1, 0) \) with \( |c_1| \leq R - 1 \), as a result of property (2.1).

Then the uniqueness of Aubry class implies

\[
\tilde{N}_{L_0+L_1+\phi}(c) = \tilde{N}_{L_0+L_1+\phi}(c). \quad (6.43)
\]

By the Legendre transformation, the associated Hamiltonian is exactly \( H_0 + H_1 - \phi \) with \( \| H_1 - \phi \|_{o_L} < \varepsilon_0 \), so (6.42) and (6.43) imply that

\[
\pi_p \circ \mathcal{L}^{-1}(\tilde{A}_{L_0}(c)) = c, \quad d(\pi_p \circ \mathcal{L}^{-1}(\tilde{A}_{L_0+L_1+\phi}(c)), \pi_p \circ \mathcal{L}^{-1}(\tilde{A}_{L_0}(c))) < s/2, \quad (6.44)
\]

for each \( c = (c_1, 0), |c_1| \leq R - 1 \).
Recall that the set $\tilde{N}_{L_0 + L_1 + \phi(c)}|_{\tau=0}$ lies in the NHIC, so it is either homologically trivial or not. In the homologically trivial case, it is well known that the $c$-equivalence holds inside $(c_1 - \delta_c, c_1 + \delta_c) \times \{0\}$ for some small $\delta_c > 0$, which satisfies condition (a) in definition 4.2.

In the latter case, $\tilde{N}_{L_0 + L_1 + \phi(c)}|_{\tau=0}$ must be an invariant curve as a result of (6.43). Then we define as (6.5) the set

$$S := \{(c_1, 0): |c_1| \leq R - 1, \tilde{\gamma}_c \text{ is an invariant circle on the NHIC}\}.$$ 

Applying theorem 6.7 to $L_0 + L_1 + \phi$, we can find a small potential perturbation with compact support $P \in G^{\alpha L}(M \times \mathbb{T})$, $\|P\|_{\alpha L} < \frac{d}{2}$, such that the Lagrangian $L_0 + L_1 + \phi + P: TM \times \mathbb{T} \to \mathbb{R}$, defined in the double covering space, satisfies: for all $c \in S$, the sets

$$\text{arg min } B_{c,1}|_{U_{c,1},U_{c,2}}, \quad \text{arg min } B_{c,1}|_{U_{c,1},U_{c,2}}$$

are both totally disconnected. Then propositions 3.5 and 3.6 together yield that for each $c \in S$, there exists a small $\delta_c > 0$ such that the set

$$\tilde{\mathcal{N}}(c, M)|_{t=0}(A(c, M)|_{t=0} + \delta_c)$$

is totally disconnected, which satisfies condition (b) in definition 4.2. Consequently, the corresponding Hamiltonian is exactly $H_0 + H_1 - \phi - P$ where $\|\phi + P\|_{\alpha L} < d$ and $\|H_1 - \phi - P\|_{\alpha L} < \varepsilon_0$.

- Next, we take $V = -\phi - P$.

Then the arguments above imply that there exists a generalized transition chain inside $[-R + 1, R - 1] \times \{0\} \subset H^1(M, \mathbb{R})$ for the Lagrangian $L_0 + L_1 - V$. Thus, we conclude from theorem 4.3 and (6.44) that the perturbed Hamiltonian $H_0 + H_1 + V$ has an orbit ($q(t), p(t)$) whose action variables $p(t)$ pass through the ball $B_{c_1}(y)$ at the time $t = t_1$, where $t_1 < t_2 < \cdots < t_k$. Finally, thanks to $|V|_{\alpha L} < d$ and the arbitrariness of $d$, we complete the proof of density in $\mathcal{W}^{1}_{\alpha L, R}$.

**Openness.** It only remains to verify the openness. Since the time for the aforementioned trajectory ($q(t), p(t)$) passing through the balls $B_{c_1}(y_1), \cdots, B_{c_1}(y_k)$ is finite, the smooth dependence of solutions of ODEs on parameters can guarantee the openness in $\mathcal{W}^{1}_{\alpha L, R}$, Theorem 1.2 is now evident from what we have proved.

Notice that in the proof of density part above, the perturbation we constructed is Gevrey potential function. Combining with the obvious openness property, theorem 1.3 is also true. □

**Acknowledgments**

We sincerely thank the anonymous referees for their insightful comments and valuable suggestions on improving our results. The first author was partially supported by China Postdoctoral Science Foundation (Grant No. 2018M641500). The second author was partially supported by
Appendix A. Normally hyperbolic theory

In this appendix, we review some classical results in the theory of normally hyperbolic manifolds. We only give a less general introduction which is better applied to our problem, and refer the reader to [40, 41, 50, 71] for the proofs and more detailed introductions.

Definition A.1. Let $M$ be a smooth Riemannian manifold and $f : M \to M$ be a $C^r(r \geq 1)$ diffeomorphism. Let $N \subset M$ be a submanifold (probably with boundary) which is invariant under $f$. Then $N$ is called a normally hyperbolic invariant manifold (NHIM) if there is an $f$-invariant tangent bundle splitting such that, for every $x \in N$

$$T_NM = T_xN \oplus E^s_x \oplus E^u_x,$$

and there exist a constant $C > 0$, rates $0 < \lambda < 1 < \mu$ with $\lambda \mu < 1$ such that

$$v \in T_xN \iff \|Df^k(x)v\| \leq C\mu^k\|v\|, \quad k \in \mathbb{Z},$$
$$v \in E_x^s \iff \|Df^k(x)v\| \leq C\lambda^k\|v\|, \quad k \geq 0,$$
$$v \in E_x^u \iff \|Df^k(x)v\| \leq C\lambda^k\|v\|, \quad k \leq 0. \quad (A.1)$$

In what follows, $N$ is assumed to be compact and connected. Let $U$ be a tubular neighborhood of the NHIM $N$. In both [40, 50] the existence of local stable and unstable manifolds in $U$, denoted by $W^s_{N,loc}$ and $W^u_{N,loc}$ respectively, are obtained by using the method of Hadamard’s graph transform. Moreover, the local stable and unstable manifolds can be characterized as follows:

$$W^s_{N,loc} = \{ y \in U \mid \text{dist}(f^k(y), N) \leq \tilde{C}_s(\lambda + \tilde{\varepsilon})^k, \quad \text{for all } k \geq 0\},$$
$$W^u_{N,loc} = \{ y \in U \mid \text{dist}(f^k(y), N) \leq \tilde{C}_u(\lambda + \tilde{\varepsilon})^k, \quad \text{for all } k \leq 0\}, \quad (A.2)$$

where the constant $\tilde{C}_s > 0$, and $\tilde{\varepsilon} > 0$ is a small constant satisfying $\lambda + \tilde{\varepsilon} < 1/\mu$. For each $x \in N$, the corresponding local stable and unstable leaves are defined as follows:

$$W^s_{x,loc} = \{ y \in U \mid \text{dist}(f^k(x), f^k(y)) \leq \tilde{C}_s(\lambda + \tilde{\varepsilon})^k, \quad \text{for all } k \geq 0\},$$
$$W^u_{x,loc} = \{ y \in U \mid \text{dist}(f^k(x), f^k(y)) \leq \tilde{C}_u(\lambda + \tilde{\varepsilon})^k, \quad \text{for all } k \leq 0\}. \quad (A.3)$$

where $\tilde{C}_{s,u} > 0$ is a constant. Then we have the following properties (see [50]):

Theorem A.2. Let $N$ be a NHIM given in (A.1), and we define an integer $l := \max\{ k \in \mathbb{Z} \mid 1 \leq k \leq r \text{ and } k < \frac{\log \lambda}{\log \mu} \}$, then

(a) $N$, $W^s_{N,loc}$ and $W^u_{N,loc}$ are $C^l$ manifolds. For each $x \in N$, the manifolds $W^s_{x,loc}$ and $W^u_{x,loc}$ are $C^l$ and $T_xW^s_{x,loc} = E^s_x$, $T_xW^u_{x,loc} = E^u_x$.

(b) $W^s_{N,loc}$ and $W^u_{N,loc}$ are foliated by the stable and unstable leaves respectively, i.e.

$$W^s_{N,loc} = \bigcup_{x \in N} W^s_{x,loc}, \quad W^u_{N,loc} = \bigcup_{x \in N} W^u_{x,loc}.$$

Moreover, if $x \neq x'$, then $W^s_{x,loc} \cap W^d_{x',loc} = \emptyset$ and $W^u_{x,loc} \cap W^d_{x',loc} = \emptyset.$

498
(c) The unstable foliation \( \{ W_{\mu,\text{loc}}^{u,x} : x \in N \} \) is \( C^l \) in the sense that \( \bigcup_{x \in N} T^k_x W_{\mu,\text{loc}}^{u,x} \) is a continuous bundle for each \( 1 \leq k \leq l \), where \( T^k \) denotes the \( k \)-th order tangent. Analogous result holds for the stable foliation.

**Remark.**

(a) We point out that for the models studied in the current paper, we have smoothness as high as the time-1 map since the dynamics on \( N \) is close to integrable.

(b) We can also define the global stable (unstable) sets \( W_N^{s,u} \) and \( W_x^{s,u} \), just by replacing \( U \) with \( M \) in (A.2) and (A.3). But \( W_N^{s,u}, W_x^{s,u} \) may fail to be embedded manifolds.

The normal hyperbolicity has stability under perturbations. Roughly speaking, the normally hyperbolic invariant manifold persists under small perturbations.

**Theorem A.3 (Persistence of normally hyperbolic invariant manifolds).** Suppose that \( N \subset M \) is a NHIM for the \( C^r (r \geq 1) \) diffeomorphism \( f \) and \( \varepsilon > 0 \) is sufficiently small. Then for any \( C^r \) diffeomorphism \( f_\varepsilon : M \to M \) satisfying \( \| f_\varepsilon - f \|_{C^1} < \varepsilon \), there exists a NHIM \( N_\varepsilon \) that is \( C^l \) diffeomorphic and close to \( N \) where \( l = \max \{ k : 1 \leq k \leq r \text{ and } k < \frac{\log \lambda}{\log \mu} \} \). Moreover, the local stable manifold \( W_{N_\varepsilon}^{s,\text{loc}} \) and local unstable manifold \( W_{N_\varepsilon}^{u,\text{loc}} \) are \( C^l \) close to those of \( N \).

**Appendix B. Variational construction of global connecting orbits**

The goal of this section is to prove theorem 4.3, which could be achieved by modifying the arguments and techniques in [25]. We also refer the reader to [21] or [19] for more details. Throughout this section, we assume \( M = \mathbb{T}^n \). Our diffusing orbits are constructed by shadowing a sequence of local connecting orbits, along each of them the Lagrangian action attains a ‘local minimum’.

**B.1. Local connecting orbits**

Let \( d_t \gamma (t) = (\gamma (t), \dot{\gamma} (t)) \). An orbit \( d_t \gamma (t) : \mathbb{R} \to TM \times \mathbb{T} \) is said to connect one Aubry set \( \tilde{\mathcal{A}} (c) \) to another one \( \tilde{\mathcal{A}} (c') \) if the \( \alpha \)-limit set of the orbit is contained in \( \tilde{\mathcal{A}} (c) \) and the \( \omega \)-limit set is contained in \( \tilde{\mathcal{A}} (c') \). We will introduce two types of local connecting orbits: type-\( c \) and type-\( h \), the former corresponds to Mather’s cohomology equivalence, while the latter corresponds to the variational interpretation of Arnold’s mechanism. Before that, we need some preparations.

**B.1.1. Time-step Lagrangian and upper semi-continuity.** Both types of local connecting orbits depend on the upper semi-continuity of minimal curves of a modified \( C^r \) Lagrangian \( L^* : T\mathbb{T}^n \times \mathbb{R} \to \mathbb{R} \) which is defined as follows: let \( L^+, L^- \) be two time-1 periodic Tonelli Lagrangians,

\[
L^*(\cdot, t) := \begin{cases} 
L^-(\cdot, t), & t \in (-\infty, 0] \\
L^+(\cdot, t), & t \in [1, +\infty),
\end{cases}
\]

and \( L^* \) is superlinear and positive definite in the fibers. Notice that \( L^* \) is not periodic in time \( t \), instead, it is periodic when restricted on either \(( -\infty, 0] \) or \([1, +\infty) \). We call such a modified Lagrangian \( L^* \) a **time-step** Lagrangian.
For a time-step Lagrangian $L^*$, a curve $\gamma : \mathbb{R} \to \mathbb{T}^n$ is called minimal if for any $t < t' \in \mathbb{R}$,
\[
\int_t^{t'} L^*(\gamma(s), \dot{\gamma}(s), s) \, ds = \min_{\substack{\xi(t) = \gamma(t), \dot{\xi}(t) = \dot{\gamma}(t), t \in [t, t'] \subset \mathbb{R} \ni \xi \in C^0([t, t'], \mathbb{T}^n)}} \int_t^{t'} L^*(\xi(s), \dot{\xi}(s), s) \, ds.
\]
So we denote by $\mathcal{G}(L^*)$ the set of all minimal curves and $\tilde{\mathcal{G}}(L^*) = \bigcup_{\gamma \in \mathcal{G}(L^*)} \{\gamma(t), \dot{\gamma}(t), t\}$.

Let $\alpha^\pm$ denote Mather’s minimal average action of $L^\pm$. For $m_0, m_1 \in \mathbb{T}^n$ and $T_0, T_1 \in \mathbb{Z}_+$, we define
\[
h_{T_0, T_1}^L(m_0, m_1) := \inf_{\gamma(-T_0) = m_0, \gamma(T_1) = m_1} \int_{-T_0}^{T_1} L^*(\gamma(t), \dot{\gamma}(t), t) \, dt + T_0\alpha^- + T_1\alpha^+,
\]
and
\[
h_{L}^\infty(m_0, m_1) := \lim_{T_0, T_1 \to +\infty} h_{T_0, T_1}^L(m_0, m_1),
\]
which are bounded. We take any two sequences of positive integers $\{T_0^i\}_{i \in \mathbb{Z}_+}$ and $\{T_1^i\}_{i \in \mathbb{Z}_+}$ with $T_1^i \to +\infty$ ($\ell = 0, 1$) as $i \to +\infty$ and the associated minimal curve $\gamma_i(t) : [-T_0^i, T_1^i] \to \mathbb{T}^n$ connecting $m_0$ to $m_1$ such that
\[
h_{L}^\infty(m_0, m_1) = \lim_{i \to +\infty} h_{T_0^i, T_1^i}^L(m_0, m_1) = \lim_{i \to +\infty} \int_{-T_0^i}^{T_1^i} L^*(\gamma_i(t), \dot{\gamma}_i(t), t) \, dt + T_0^i\alpha^- + T_1^i\alpha^+.
\]

The following lemma shows that any accumulation point $\gamma$ of $\{\gamma_i\}_i$ is a pseudo curve playing an analogous role as a semi-static curve. For the proof, see [24] or [25].

**Lemma B.1.** Let $\gamma : \mathbb{R} \to \mathbb{T}^n$ be an accumulation point of $\{\gamma_i\}_i$, as shown above. Then for any $s \geq 0, t \geq 1$,
\[
\int_{-s}^{t} L^*(\gamma(\tau), \dot{\gamma}(\tau), \tau) \, d\tau + sa^- + t\alpha^+ = \inf_{\substack{\xi \in C^0([-s, t], \mathbb{T}^n) \ni \
\xi(-s) \equiv \gamma(-s), \xi(t) \equiv \gamma(t) \\{\xi \in \mathbb{Z}, \xi - r \in \mathbb{Z} \\forall r \in \mathbb{Z}, \xi \equiv m_0, \xi \equiv m_1 \\forall 1 \leq i \leq t_1 \geq 1}} \int_{-s}^{t} L^*(\xi(\tau), \dot{\xi}(\tau), \tau) \, d\tau + sa^- + t \alpha^+,
\]
where the minimum is taken over all absolutely continuous curves.

This leads us to define the set of pseudo connecting curves
\[
\mathcal{C}(L^*) := \{\gamma \mid \gamma \in \mathcal{G}(L^*) \text{ and } (B.1) \text{ holds}\}
\]
Clearly, for each $\gamma \in \mathcal{C}(L^*)$ the orbit $\{\gamma(t), \dot{\gamma}(t), t\}$ would negatively approach the Aubry set $A(L^-)$ of the Lagrangian $L^-$ and positively approach the Aubry set $\tilde{A}(L^+)$ of $L^+$. This is why we call it a pseudo connecting curve. Define the following sets
\[
\tilde{C}(L^*) := \bigcup_{\gamma \in \mathcal{C}(L^*)} \{\gamma(t), \dot{\gamma}(t), t\}, \quad C(L^*) := \bigcup_{\gamma \in \mathcal{C}(L^*)} \{\gamma(t), t\}.
\]
Notice that if \( L^* \) is time-1 periodic, then \( \tilde{C}(L^*) \) is exactly the Mañé set and \( \hat{C}(L^*) \) is exactly the projected Mañé set. So we can prove the following property:

**Proposition B.2.** The set-valued map \( L^* \mapsto \mathcal{E}(L^*) \) is upper semi-continuous, namely if \( L^*_i \to L^* \) in the \( C^2 \) topology, then we have the set inclusion

\[
\lim \sup_i \mathcal{E}(L^*_i) \subset \mathcal{E}(L^*).
\]

Consequently, the map \( L^* \mapsto \tilde{C}(L^*) \) is also upper semi-continuous.

**Proof.** Let \( L^*_i \to L^* \) in the \( C^2 \) topology. If \( \gamma_i \) converges \( C^0 \)-uniformly to a curve \( \gamma \) on each compact interval of \( \mathbb{R} \) with \( \gamma_i \in \mathcal{E}(L^*_i) \). We claim that \( \gamma \in \mathcal{E}(L^*) \).

Indeed, there exists \( K > 0 \) such that \( \| \gamma_i(t) \| \leq K \) for all \( t \in \mathbb{R} \), so the set \( \{ \gamma_i \} \) is compact in the \( C^1 \) topology. Since each \( \gamma_i \) satisfies the Euler–Lagrange equation of \( L_i \), by using the positive definiteness of \( L^*_i \), one can write the Euler–Lagrange equation in the form of \( \tilde{x} = f_i(x, \dot{x}, t) \) for some \( f_i \), which implies that \( \{ \gamma_i \} \) is compact in the \( C^2 \) topology. By the Arzelà–Ascoli theorem, extracting a subsequence if necessary, we can assume that \( \gamma_i \) converges \( C^1 \)-uniformly to a \( C^1 \) curve \( \eta \) on each compact interval of \( \mathbb{R} \). Obviously, \( \eta = \gamma \).

Next, if \( \gamma \notin \mathcal{E}(L^*) \), there would be some \( s \geq 0, t \geq 1 \), a curve \( \tilde{\gamma} : [-s - n_1, t + n_2] \to M \) and \( \delta > 0 \) such that the action

\[
\int_{-s - n_1}^{t + n_2} L^*(\tilde{\gamma}(\tau), \dot{\tilde{\gamma}}(\tau), \tau) \, d\tau + n_1 \alpha^- + n_2 \alpha^+ < \int_{-s}^{t} L^*(\gamma(\tau), \dot{\gamma}(\tau), \tau) \, d\tau - \delta
\]

where \( s, s + n_1 \geq 0, t, t + n_2 \geq 1 \) and \( \tilde{\gamma}(-s - n_1) = \gamma(-s), \tilde{\gamma}(t + n_2) = \gamma(t) \). Since we have shown that \( \gamma \) is an accumulation point of \( \gamma_i \) in the \( C^1 \) topology, for any small \( \varepsilon > 0 \), there would be a sufficiently large \( i \) such that \( \| \gamma - \gamma_i \|_{C^1(\mathbb{R})} \leq \varepsilon \) and a curve \( \tilde{\gamma}_i : [-s - n_1, t + n_2] \to M \) with \( \tilde{\gamma}_i(-s - n_1) = \gamma(-s), \tilde{\gamma}_i(t + n_2) = \gamma(t) \) such that

\[
\int_{-s - n_1}^{t + n_2} L^*_i(\tilde{\gamma}_i(\tau), \dot{\tilde{\gamma}}_i(\tau), \tau) \, d\tau + n_1 \alpha^- + n_2 \alpha^+ \leq \int_{-s}^{t} L^*_i(\gamma(\tau), \dot{\gamma}(\tau), \tau) \, d\tau - \frac{\delta}{2}
\]

By (B.1), \( \gamma_i \notin \mathcal{E}(L^*_i) \), which is a contradiction. This proves \( \gamma \in \mathcal{E}(L^*) \).

Finally, the upper semi-continuity of \( L^* \mapsto \tilde{C}(L^*) \) is a consequence of what we have shown above. \( \square \)

**B.12. Local connecting orbits of type-c.** In condition of the cohomology equivalence (see definition 4.1), we will show how to construct local connecting orbits based on Mather’s variational mechanism. This idea of construction was first proposed by Mather in [67].

**Theorem B.3.** Let \( L : \mathbb{T}^n \times \mathbb{T} \to \mathbb{R} \) be a Tonelli Lagrangian and \( c, c' \in H^1(\mathbb{T}^n, \mathbb{R}) \) be cohomology equivalent through a path \( \Gamma : [0, 1] \to H^1(\mathbb{T}^n, \mathbb{R}) \). Then there would exist \( \tilde{c} = c_0, c_1, \ldots, c_k = c' \) on the path \( \Gamma \), closed 1-forms \( \eta_i \) and \( \mu_i \) on \( M \) with \( [\eta_i] = c_i, [\mu_i] = c_{i+1} - c_i \) and a smooth function \( \rho_i(t) : [0, 1] \to [0, 1] \) for each \( i = 1, \ldots, k \), such that the time-step Lagrangian

\[
L_{\eta_i, \mu_i} = L - \eta_i - \mu_i, \quad \text{with} \quad \mu_i = \rho_i(t)\mu_i
\]

possesses the following properties:

For each curve \( \gamma \in \mathcal{E}(L_{\eta_i, \mu_i}) \), it determines a trajectory \((d\gamma(t), t)\), connecting \( \tilde{\Lambda}(c_i) \) to \( \tilde{\Lambda}(c_{i+1}) \), of the Euler–Lagrange flow \( \phi^t_L \).
Proof. By definition 4.1, it is obvious that there exist \( c = c_0, c_1, \ldots, c_k = c' \) on the path \( \Gamma \), closed 1-forms \( \eta_i \) and \( \mu_i \) on \( M \) with \( [\eta_i] = c_i, [\mu_i] = c_{i+1} - c_i \in \mathbb{V}_{c_i}^+ \) for each \( i = 1, \ldots, k \).

By the arguments in section 4, there is also a neighborhood \( U_i \) of the projected Mañé set \( \mathcal{N}_0(c_i) \) such that \( \forall c_i = l_{U_i} H_1(U_i, \mathbb{R}) \).

In particular, we can suppose \( \mu_i = 0 \) on \( U_i \). Indeed, as \( [\tilde{\mu}_i] \in \mathbb{V}_{c_i}^+ \), \( \tilde{\mu}_i \) is exact when restricted on \( U_i \) and then there is a smooth function \( f : M \to \mathbb{R} \) satisfying \( d\tilde{f} = \tilde{\mu}_i \) on \( U_i \), hence we can replace \( \tilde{\mu}_i \) by \( \mu_i - df \).

As \( \mathcal{N}_0(c_i) \subset U_i \), there exists \( \delta_i \ll 1 \) such that \( \mathcal{N}_0(c_i) \subset U_i \) for all \( t \in [0, \delta_i] \). Let \( \rho_i : \mathbb{R} \to [0, 1] \) be a smooth function such that \( \rho_i(t) = 0 \) for \( t \in (-\infty, 0) \), \( \rho_i(t) = 1 \) for \( t \in [\delta_i, +\infty) \). We set \( \mu_i = \rho_i(t)\tilde{\mu}_i \) and introduce a time-step Lagrangian

\[
L_{\eta, \mu_i} = L - \eta_i - \mu_i : T\mathbb{T}^n \times \mathbb{R} \to \mathbb{R}.
\]

For each orbit \( \gamma \in \mathcal{C}(L_{\eta, \mu_i}) \), by the upper semi-continuity in proposition B.2,

\[
\gamma(t) \in U_i, \quad \forall t \in [0, \delta_i]
\]

holds provided that \( |\tilde{\mu}_i| \) is small enough.

Clearly, \( (\gamma(t), \dot{\gamma}(t)) \) solves the Euler–Lagrange equation of \( L_{\eta, \mu_i} \). To verify it solves the Euler–Lagrange equation of \( L \), we see that \( \gamma(t) \in U_i \) and \( L_{\eta, \mu_i} = L - \eta_i \) on \( U_i \), where \( \eta_i \) is a closed 1-form, so \( \gamma(t) \) solves the Euler–Lagrange equation of \( L \) for \( t \in [0, \delta_i] \). On the other hand, for \( t \in (-\infty, \delta_i] \) we have \( L_{\eta, \mu_i} = L - \eta_i \), then \( \gamma(t) \) is a \( \gamma \)-semi static curve of \( L \) on the interval \( (-\infty, \delta_i] \). Similarly, \( \gamma(t) \) is a \( \gamma \)-semi static curve of \( L \) for \( t \in [\delta_i, +\infty) \). Thus, \( (\gamma(t), \dot{\gamma}(t)) : \mathbb{R} \to T\mathbb{T}^n \) solves the Euler–Lagrange equation of \( L \), and by section 2, this orbit connects \( A(c_i) \) to \( A(c_{i+1}) \).

\[ \square \]

B.1.3. Local connecting orbits of type-\( h \). Next, we will discuss the so-called local connecting orbits of type-\( h \), it can be thought of as a variational version of Arnold’s mechanism, the condition of geometric transversality is extended to the total disconnectedness of minimal points of barrier function. It is used to handle the situation where the cohomology equivalence does not always exist. Usually, it is applied to the case where the Aubry set lies in a neighborhood of a lower dimensional torus, in that case, we let \( \hat{\pi} : M \to \mathbb{T}^n \) be a finite covering of \( \mathbb{T}^n \). Denote by \( \hat{N}(c, \hat{M}), \hat{A}(c, \hat{M}) \) the Mañé set and Aubry set with respect to \( \hat{M} \), then \( \hat{A}(c, \hat{M}) \) would have more than one Aubry classes. In fact, for the construction of type-\( h \) local connecting orbits in our proofs of theorems 1.2 and 1.3, it only involves two Aubry classes (see section 6).

Thus, we only need to deal with the situation where the Tonelli Lagrangian \( L : T\mathbb{T}^n \times \mathbb{T} \to \mathbb{R} \) contains more than one Aubry classes. Let \( A(c) \lvert_{t=0} \) denote the time-0 section of the projected Aubry set \( A(c) \), i.e. \( A(c) \cap (\mathbb{T}^n \times \{ t = 0 \}) \), then we obtain the local connecting orbits of type-\( h \) as follows:

**Theorem B.4.** Let the projected Aubry set \( A(c) = A(c_1) \cup \cdots \cup A(c_k) \) consists of \( k(k \geq 2) \) Aubry classes. Let \( U : = \mathbb{T}^n \setminus (A(c) \lvert_{t=0} + \kappa) \) be an open set where \( \kappa > 0 \) is small. If \( U \cap (\mathbb{T} \times \{ t = 0 \}) \) is non-empty and totally disconnected. Then for any \( c' \) sufficiently close to \( c \), there exists an orbit of the Euler–Lagrange flow \( \phi_t \) whose \( \alpha \)-limit set lies in \( \hat{A}(c) \) and \( \omega \)-limit set lies in \( \hat{A}(c') \).

**Proof.** As the number of Aubry class of \( A(c) \) is finite, it is well known that if \( c' \) is sufficiently close to \( c \), the projected Aubry set \( A(c') \) will be contained in a small neighborhood of \( A(c) \), see e.g. [5].

Since each Aubry class is compact and disjoint with each other, we have \( \text{dist}(A(c_i), A(c_{i'})) > 0 \) for any \( i \neq i' \), and there exist open neighborhoods \( N_1, \ldots, N_k \subset M \) such that \( A(c_i) \lvert_{t=0} \subset N_i \) for
Then we set $\mu$ each $1 \leq i \leq k$ and $\text{dist}(N_i, N_{ip}) > 0$ for $i \neq p$. So $A(c')_{|t=0} \subset \bigcup_i N_i$. From proposition 3.5 and definition (3.6) we know $N(c) = \bigcup_f N_{f,c}$, hence there is a pair $(j, f)$ such that $A(c')_{|t=0} \cap N_f \neq \emptyset$ and $U \cap N_{f,c} \neq \emptyset$.

By the total disconnectedness assumption, we can find simply connected open sets $F$ and $O$ such that $F \subset O \subset U$, $\text{dist}(O, \bigcup_{i=1}^N N_i) > 0$ and $\emptyset \neq O \cap (N_{f,c} | t=0) \subset F$. Then some $\delta > 0$ exists such that

$$O \bigcap (N_{f,c} | 0 \leq t \leq \delta) \subset F.$$  \hfill (B.3)

Let $\eta$ and $\bar{\mu}$ be closed 1-forms such that $[\eta] = c, [\bar{\mu}] = c' - c$, and let $\rho : \mathbb{R} \rightarrow [0, 1]$ be a smooth function such that $\rho(t) = 0$ for $t \leq 0$, $\rho = 1$ for $t \geq \delta$. Note that by the simple connectedness of $O$, we are able to choose $\bar{\mu}$ such that $\text{supp} \bar{\mu} \cap \partial O = \emptyset$. Next, we can construct a smooth function $\psi(x, t) = \varepsilon \psi_1(x)\psi_2(t) : M \times \mathbb{T} \rightarrow [-1, 1]$ where $\varepsilon > 0$, such that

$$\psi_1(x) = \begin{cases} 1, & x \in F, \\ < 1, & x \in O \setminus F, \\ 0, & x \in \bigcup_{i \neq j, f} N_i, \\ = 0, & \text{elsewhere}. \end{cases}$$

and

$$\psi_2(t) = \begin{cases} > 0, & t \in (0, \delta), \\ = 0, & t \in (-\infty, 0] \cup [\delta, +\infty). \end{cases}$$

Then we set $\mu = \rho(t)\bar{\mu}$ and introduce a time-step Lagrangian

$$L_{\eta, \mu, \psi} = L - \eta - \mu - \psi : \mathbb{T}^{d} \times \mathbb{R} \rightarrow \mathbb{R}.$$  

Let us first suppose $\mu = 0$. Since $\psi(x, t) = 0$ for $(x, t) \in \mathcal{A}_{c,j} \cup \mathcal{A}_{c,f}$, and $\psi(x, t) < 0$ for $(x, t) \in \bigcup_{i \neq j, f} N_i$ with $t \in (0, \delta)$, the Lagrangian $L_{\eta, \mu, \psi}$ contains only two Aubry classes which are exactly $\mathcal{A}_{c,j}$ and $\mathcal{A}_{c,f}$ provided the number $\varepsilon > 0$ is small enough. The set $\mathcal{C}(L_{\eta, \mu, \psi})$ then satisfies:

(a) $\mathcal{A}_{c,j} \cup \mathcal{A}_{c,f} \subset \mathcal{C}(L_{\eta, \mu, \psi})$.

(b) $\mathcal{C}(L_{\eta, \mu, \psi}) \setminus (\mathcal{A}_{c,j} \cup \mathcal{A}_{c,f})$ is non-empty. For each pseudo connecting curve $\xi \in \mathcal{C}(L_{\eta, \mu, \psi}) \setminus (\mathcal{A}_{c,j} \cup \mathcal{A}_{c,f})$, we have $\xi(t) \in F$ for $0 \leq t \leq \delta$, but its integer translation $K^t \xi(t) := \xi(t - K)$ with $K \in \mathbb{Z} \setminus 0$ does not belong to $\mathcal{C}(L_{\eta, \mu, \psi})$ since $L_{\eta, \mu, \psi}$ is not periodic in $t$.

(c) $\mathcal{C}(L_{\eta, \mu, \psi})$ does not contain any other curves.

These properties follow directly from (B.3) and the fact that $\psi(x, 0)$ attains its maximum if and only if $x \in F$, and the upper semi-continuity of $(\eta, \mu, \psi) \mapsto \mathcal{C}(L_{\eta, \mu, \psi})$.

If $\mu \neq 0$. For $m_0 \in \mathcal{A}_{c,j}|_{t=0} m_1 \in \mathcal{A}_{c,f}|_{t=0}$, let $T_j^k, T_f^k \rightarrow +\infty$ be two sequences of positive integers such that

$$\lim_{k \rightarrow \infty} h_{T_j^k, T_f^k}^k(m_0, m_1) = h_{T_j^k, T_f^k}^k(m_0, m_1).$$

Let $\gamma_2(t) : [-T_j^k, T_f^k] \rightarrow M$ be a minimizer associated with $h_{T_j^k, T_f^k}^k(m_0, m_1)$ and $\gamma$ be any accumulation point of $\{\gamma_k\}_k$, then $\gamma \in \mathcal{C}(L_{\eta, \mu, \psi})$. If $\mu$ and $\varepsilon$ are small enough, we deduce from the
the Euler–Lagrangian equation of \( L_{\eta,\mu,\psi} \), but we still need to verify that it solves the Euler-Lagrangian equation of \( L \). In fact, \( L_{\eta,\mu,\psi} = L - \eta \) for \( t \leq 0 \) and \( L_{\eta,\mu,\psi} = L - \eta + \mu \) for \( t \geq \delta \) where \( \eta, \mu \) are closed 1-forms, so \( \gamma(t) \) solves the Euler–Lagrangian equation of \( L \) for \( t \in (-\infty, 0) \cup [\delta, +\infty) \). This also implies that \( \gamma: (-\infty, 0] \rightarrow \mathbb{T}^n \) is a \( c \)-semi static curve of \( L \) and \( \gamma: [\delta, +\infty) \rightarrow M \) is a \( c' \)-semi static curve of \( L \), then

\[
\gamma(t) \in F, \quad \forall t \in [0, \delta].
\] (B.4)

Obviously, \((\gamma, \dot{\gamma})\) satisfies the Euler–Lagrangian equation of \( L_{\eta,\mu,\psi} \), but we still need to verify that it solves the Euler-Lagrangian equation of \( L \). In fact, \( L_{\eta,\mu,\psi} = L - \eta \) for \( t \leq 0 \) and \( L_{\eta,\mu,\psi} = L - \eta + \mu \) for \( t \geq \delta \) where \( \eta, \mu \) are closed 1-forms, so \( \gamma(t) \) solves the Euler–Lagrangian equation of \( L \) for \( t \in (-\infty, 0) \cup [\delta, +\infty) \). This also implies that \( \gamma: (-\infty, 0] \rightarrow \mathbb{T}^n \) is a \( c \)-semi static curve of \( L \) and \( \gamma: [\delta, +\infty) \rightarrow M \) is a \( c' \)-semi static curve of \( L \), then

\[
\alpha(d\gamma(t), t) \subset \widetilde{A}(c), \quad \omega(d\gamma(t), t) \subset \widetilde{A}(c').
\]

Besides, for \( t \in [0, \delta] \), we deduce from (B.4) that the Euler–Lagrangian equation \((\frac{\partial}{\partial t} - \partial_x)\ln c_{\mu,\psi} = 0 \) is equivalent to \((\frac{\partial}{\partial t} - \partial_x)\ln L = 0 \) along the curve \( \gamma(t) \) within \( 0 \leq t \leq \delta \), which therefore shows that \((\gamma, \dot{\gamma})\) solves the Euler–Lagrangian equation of \( L \) for \( t \in [0, \delta] \). This completes our proof.

From the proof of theorem B.4 we see that the connecting orbit \((\gamma, \dot{\gamma})\) obtained in this theorem is locally minimal in the following sense:

**Local minimum.** There are two open balls \( V^- \), \( V^+ \subset \mathbb{T}^n \) and \( k^-, k^+ \in \mathbb{Z}^+ \) such that \( V^- \subset N_j A_{1} \rangle \alpha(\gamma)|_{t=0} \) and \( V^+ \subset N_j A_{c'} \rangle \alpha(\gamma)|_{t=0} \gamma(-k^-) \in V^- \), \( \gamma(k^+) \in V^+ \) and

\[
\begin{align*}
\lim_{k^-, k^+ \to \infty} \int_{k^+}^{k^-} L_{\eta,\mu,\psi}(\gamma(t), \dot{\gamma}(t), t) dt + k^- \alpha(c) + k^+ \alpha(c')
\end{align*}
\] (B.5)

holds for all \((m_0, m_1) \in \partial(V^- \times V^+) \), \( x^- \in N_j \cap \alpha(\gamma)|_{t=0} \), \( x^+ \in N_j \cap \omega(\gamma)|_{t=0} \), where \( k^- \), \( k^+ \) are the sequences such that \( \gamma(-k^-) \rightarrow x^- \) and \( \gamma(k^+) \rightarrow x^+ \).

The set of curves starting from \( V^- \) and reaching \( V^+ \) within time \( k^- + k^+ \) would make up a neighborhood of the curve \( \gamma \) in the space of curves. If it touches the boundary of this neighborhood, the action of \( L_{\eta,\mu,\psi} \) along a curve \( \xi \) will be larger than the action along \( \gamma \). Besides, the connecting orbit of type-\( c \) also has local minimal property. In this case, the modified Lagrangian has the form \( \bar{L}_{\eta,\mu,\psi} \). The local minimality is crucial in the variational construction of global connecting orbits.

**B.2. Global connecting orbits**

Now, we are ready to prove theorem 4.3 from a variational viewpoint. Intrinsically, we construct a global connecting orbit by shadowing a sequence of local connecting orbits.

**Sketch of the proof of theorem 4.3.** The proof parallels to that of [25] by a small modification. Here we only give a sketch of the basic idea, and the reader can refer to [25, section 5], [19, 21] for more details. For the generalized transition chain \( \Gamma: [0, 1] \rightarrow H^1(\mathbb{T}^n, \mathbb{R}) \) with \( \Gamma(0) = c \) and \( \Gamma(1) = c' \), by definition there exists a sequence \( 0 = s_0 < s_1 < \cdots < s_m = 1 \) such that

\[
s_i \text{ is sufficiently close to } s_{i+1} \text{ for each } 0 \leq i \leq m - 1, \text{ and } A(\Gamma(s_i)) \text{ could be connected to } A(\Gamma(s_{i+1})) \text{ by a local minimal orbit of either type-}c \text{ (as theorem B.3) or type-}h \text{ (as theorem B.4). Then the global connecting orbits are just constructed by shadowing these local ones.}
\]

For simplicity, we set \( c_j = \Gamma(s_i) \).
For each $i \in \{0, 1, \ldots, m - 1\}$, we take $\eta_i, \mu_i, \psi_i$ and $\delta_i > 0$ as that in the proof of theorems B.3 and B.4, where $\psi_i = 0$ in the case of type-c. Then we choose $k_i \in \mathbb{Z}_+$ with $k_0 = 0$ and $k_{i+1} - k_i$ is suitably large for each $i \in \{0, 1, \ldots, m - 1\}$, and introduce a modified Lagrangian

$$L^* := L - \eta_0 - \sum_{i=0}^{m-1} k_i^*(\mu_i + \psi_i).$$

Here, $k_i^*$ denotes a time translation operator such that $k_i^* f(x, t) = f(x, t - k_i)$, and $\psi_i = 0$ in the case of type-c. By this definition, we see that $L^* = L - \eta_0$ for $t \leq k_0 = 0$, $L^* = L - \eta_m$ for $t \geq k_m + \delta_m$, and for each $i \in \{0, 1, \ldots, m-2\}$, $L^* = L - \eta_i - k_i^*(\mu_i + \psi_i)$ on $t \in [k_i, k_i + \delta_i]$ and $L^* = L - \eta_{i+1}$ for $t \in [k_i + \delta_i, k_{i+1}]$.

For integers $T_0, T_m \in \mathbb{Z}^+$ and $x_0, x_m \in \mathbb{T}^n$, we define

$$h_{T_0:T_m}(x_0, x_m) = \inf_{\xi} \int_{-T_0}^{T_m + k_{m-1}} L^*(\xi(s), \dot{\xi}(s), s) \, ds + \sum_{i=1}^{m-1} (k_i - k_{i-1})\alpha(c_i) + T_0\alpha(c_0) + T_m\alpha(c_m),$$

where the infimum is taken over all absolutely continuous curves $\xi$ defined on the interval $[-T_0, T_m + k_{m-1}]$ under some boundary conditions. By carefully setting boundary conditions and using standard arguments in variational methods, one could find that the minimizer $\gamma(t; T_0, T_m, m_0, m_1)$ of the action $h_{T_0:T_m}(x_0, x_m)$ is smooth everywhere, along which the term $k_i^*(\mu_i + \psi_i)$ would not contribute to the Euler–Lagrange equation. Hence the minimizer produces an orbit of the flow $\phi_t^*$, which passes through the $\varepsilon$-neighborhood of $\bar{A}(c_i)$ at some time $t = t_i$. Let $T_0, T_m \to +\infty$, we can then get an accumulation curve $\gamma(t) : \mathbb{R} \to \mathbb{T}^n$ of the sequence $\{\gamma(t; T_0, T_m, m_0, m_1)\}$ such that the $\alpha$-limit set of $\langle d\gamma(t), t \rangle$ lies in $\bar{A}(c)$ and the $\omega$-limit set of $\langle d\gamma(t), t \rangle$ lies in $\bar{A}(c')$. This completes the proof. \( \square \)

ORCID iDs

Qinbo Chen https://orcid.org/0000-0002-2283-4220

References

[1] Arnol’d V I 1964 Instability of dynamical systems with many degrees of freedom Dokl. Akad. Nauk SSSR 156 9–12
[2] Bernard P 2002 Connecting orbits of time dependent Lagrangian systems Ann. Inst. Fourier 52 1533–68
[3] Bernard P 2008 The dynamics of pseudographs in convex Hamiltonian systems J. Am. Math. Soc. 21 615–69
[4] Bernard P 2010 Large normally hyperbolic cylinders in a priori stable Hamiltonian systems Ann. Henri Poincaré 11 929–42
[5] Bernard P 2010 On the Conley decomposition of Mather sets Rev. Matematica Iberoana, 26 115–32
[6] Bernard P and Contreras G 2008 A generic property of families of Lagrangian systems Ann. of Math. 167 1099–108
[7] Bernard P, Kaloshin V and Zhang K 2016 Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders Acta Math. 217 1–79
[8] Berti M, Biasco L and Bolle P 2003 Drift in phase space: a new variational mechanism with optimal diffusion time J. Math. Pures Appl. 82 613–64
[9] Berti M and Bolle P 2002 A functional analysis approach to Arnold diffusion Annales de l’Institut Henri Poincare (C) Non Linear Anal. 19 395–450
[10] Bessi U 1996 An approach to Arnold’s diffusion through the calculus of variations Nonlinear Anal. Theory, Methods Appl. 26 1115–35
[11] Bessi U, Chierchia L and Valdinoci E 2001 Upper bounds on Arnold diffusion times via Mather theory Journal de Mathématiques Pures et Appliquées 80 105–29
[12] Bounemoura A 2011 Effective stability for Gevrey and finitely differentiable prevalent Hamiltonians Commun. Math. Phys. 307 157–83
[13] Bounemoura A 2013 Normal forms, stability and splitting of invariant manifolds: I. Gevrey Hamiltonians Regul. Chaot. Dyn. 18 237–60
[14] Bounemoura A and Féjoz J 2017 KAM, α-Gevrey regularity and α-Bruno–Rüssmann condition (arXiv:1705.06909v2)
[15] Bounemoura A and Marco J-P 2011 Improved exponential stability for near-integrable quasi-convex Hamiltonians Nonlinearity 24 97–112
[16] Bourgain J and Kaloshin V 2005 On diffusion in high-dimensional Hamiltonian systems J. Funct. Anal. 229 1–61
[17] Bolotin S and Treschev D 1999 Unbounded growth of energy in nonautonomous Hamiltonian systems Nonlinearity 12 365–87
[18] Chen Q and Cheng C-Q 2017 Regular dependence of the Peierls barriers on perturbations J. Differ. Equ. 262 4700–23
[19] Cheng C-Q 2012 Arnold diffusion in nearly integrable Hamiltonian systems (arXiv:1207.4016)
[20] Cheng C-Q 2017 Dynamics around the double resonance Camb. J. Math. 5 153–228
[21] Cheng C-Q 2019 The genericity of Arnold diffusion in nearly integrable Hamiltonian systems Asian J. Math. 23 401–38
[22] Cheng C-Q 2017 Uniform hyperbolicity of invariant cylinder J. Differ. Geom. 106 1–43
[23] Cheng C-Q and Xue J 2015 Existence of diffusion orbits in a priori unstable Hamiltonian systems J. Differ. Geom. 82 457–517
[24] Cheng C-Q and Yan J 2009 Arnold diffusion in Hamiltonian systems: a priori unstable case J. Differ. Geom. 82 229–77
[25] Cheng C-Q and Zhou M 2016 Global normally hyperbolic invariant cylinders in Lagrangian systems Math. Res. Lett. 23 685–705
[26] Contreras G, Iturriaga R, Paternain G P and Paternain M 2000 The Palais–Smale condition and Mañé’s critical values Ann. Henri Poincaré 1 655–84
[27] Delshams A, de la Llave R and Seara T M 2000 A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of T² Commun. Math. Phys. 209 353–92
[28] Delshams A, de la Llave R and Seara T M 2006 A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: announcement of results Electron. Res. Announc. Am. Math. Soc. 9 125–34
[29] Delshams A, de la Llave R and Seara T M 2008 Geometric properties of the scattering map of a normally hyperbolic invariant manifold Adv. Math. 217 1096–153
[30] Delshams A, de la Llave R and Seara T M 2016 Instability of high dimensional Hamiltonian systems: multiple resonances do not impede diffusion Adv. Math. 294 689–755
[31] Davletshin M and Treschev D 2018 Arnold diffusion in multidimensional a priori unstable Hamiltonian systems (arXiv:1807.07832)
[32] Fathi A 2008 Weak KAM theorem in Lagrangian dynamics preliminary version number 10 (www.math.u-bordeaux.fr/~pthieull/Recherche/KamFaible/Publications/Fathi2008_01.pdf)
[33] Fayad B and Sauzin D 2018 KAM tori are no more than sticky (arXiv:1812.04163)
[40] Fenichel N 1971 Persistence and smoothness of invariant manifolds for flows Indiana Univ. Math. J. 21 193–226
[41] Fenichel N 1977 Asymptotic stability with rate conditions: II Indiana Univ. Math. J. 26 81–93
[42] Fontich E, Martín P and Martín P 2001 Arnold diffusion in perturbations of analytic integrable Hamiltonian systems Discrete Continuous Dyn. Syst. 7 61–84
[43] Gelfreich V and Turaev D 2008 Unbounded energy growth in Hamiltonian systems with a slowly varying parameter Commun. Math. Phys. 283 769–94
[44] Gelfreich V and Turaev D 2017 Arnold diffusion in a priori chaotic symplectic maps Commun. Math. Phys. 353 507–47
[45] Gevrey M 1918 Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire Ann. Sci. École Norm. Sup. 35 129–90
[46] Gidea M, de la Llave R and Seara T M 2020 A general mechanism of diffusion in Hamiltonian systems: qualitative results Comm. Pure Appl. Math. 73 150–209
[47] Gidea M and Marco J-P 2017 Diffusion along chains of normally hyperbolic cylinders (arXiv:1708.08314)
[48] Gidea M and Robinson C 2007 Shadowing orbits for transition chains of invariant tori alternating with Birkhoff zones of instability Nonlinearity 20 1115–43
[49] Guardia M, Kaloshin V and Zhang J 2016 A second order expansion of the separatrix map for trigonometric perturbations of a priori unstable systems Commun. Math. Phys. 348 321–61
[50] Hirsch M W, Pugh C C and Shub M 1977 Invariant manifolds (Lecture Notes in Mathematics vol 583) (Berlin: Springer)
[51] Komatsu H 1979 The implicit function theorem for ultradifferentiable mappings Proc. Japan Acad. A 55 69–72
[52] Kaloshin V and Levi M 2008 Geometry of Arnold diffusion SIAM Rev. 50 702–20
[53] Kaloshin V and Zhang K 2013 A strong form of Arnold diffusion for two and a half degrees of freedom (arXiv:1212.1150)
[54] Kaloshin V and Zhang K 2014 A strong form of Arnold diffusion for three and a half degrees of freedom (Announcement of result) (www.math.umd.edu/~vkaloshi/papers/announce-three-and-half.pdf)
[55] Kaloshin V and Zhang K 2014 Dynamics of the dominant Hamiltonian, with applications to Arnold diffusion (arXiv:1410.1844)
[56] Kaloshin V and Zhang K 2015 Arnold diffusion for smooth convex systems of two and a half degrees of freedom Nonlinearity 28 2699–720
[57] Lopes Dias J and Gaivão J P 2017 Renormalization of Gevrey vector fields with a Brjuno type arithmetical condition (arXiv:1706.04510)
[58] Lazzarini L, Marco J-P and Sauzin D 2019 Measure and capacity of wandering domains in Gevrey near-integrable exact symplectic systems Memoir. Am. Math. Soc. 257 vi+110
[59] Li X and Cheng C-Q 2010 Connecting orbits of autonomous Lagrangian systems Nonlinearity 23 119–41
[60] Lochak P and Marco J-P 2005 Diffusion times and stability exponents for nearly integrable analytic systems Cent. Eur. J. Math. 3 342–97
[61] Mañé R 1996 Generic properties and problems of minimizing measures of Lagrangian systems Nonlinearity 9 273–310
[62] Marco J-P 2016 Arnold diffusion for cusp-generic nearly integrable convex systems on \( \mathbb{R}^3 \) (arXiv:1602.02403)
[63] Marco J-P 2016 Chains of the compact cylinders for cusp-generic nearly integrable convex systems on \( \mathbb{R}^2 \) (arXiv:1602.02399)
[64] Marco J-P and Sauzin D 2003 Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems Publ. Math. IHÉS 96 199–275
[65] Marco J-P and Sauzin D 2004 Wandering domains and random walks in Gevrey near-integrable systems Ergod. Theory Dyn. Syst. 24 1619–66
[66] Mather J 1991 Action minimizing invariant measures for positive definite Lagrangian systems Math. Z. 207 169–207
[67] Mather J N 1993 Variational construction of connecting orbits Ann. Inst. Fourier 43 1349–86
[68] Mather J Graduate course at Princeton, 95-96, and Lectures at Penn State, Spring 96, Paris, Summer 96, Austin, Fall 96.
[69] Mather J 2003 Arnold diffusion: I. Announcement of results Sovrem. Mat. Fundam. Napravl. 2 116–30
[70] Mather J 2012 Arnold diffusion by variational methods Essays in Mathematics and its Applications (Heidelberg: Springer) pp 271–85
[71] Pesin Y 2004 Lectures on Partial Hyperbolicity and Stable Ergodicity. Zurich Lectures in Advanced Mathematics (Zürich: European Mathematical Society (EMS))
[72] Popov G 2004 KAM theorem for Gevrey Hamiltonians Ergod. Theory Dyn. Syst. 24 1753–86
[73] Treschev D 2002 Multidimensional symplectic separatrix maps J. Nonlinear Sci. 12 27–58
[74] Treschev D 2004 Evolution of slow variables in a priori unstable Hamiltonian systems Nonlinearity 17 1803–41
[75] Treschev D 2012 Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems Nonlinearity 25 2717–57
[76] Wang K and Yan J 2012 A new kind of Lax–Oleinik type operator with parameters for time-periodic positive definite Lagrangian systems Commun. Math. Phys. 309 663–91
[77] Wang L 2015 Destruction of invariant circles for Gevrey area-preserving twist map J. Dyn. Differ. Equ. 27 283–95
[78] Zhang J and Cheng C-Q 2013 Asymptotic trajectories of KAM torus (arXiv:1312.2102)
[79] Zhang K 2011 Speed of Arnold diffusion for analytic Hamiltonian systems Invent. Math. 186 255–90