SCIENCE CHINA
Mathematics

Progress of Projects Supported by NSFC
• ARTICLES •

August 2023 Vol. 66 No. 8: 1649–1712
https://doi.org/10.1007/s11425-022-2118-1

Arnold diffusion for nearly integrable Hamiltonian systems

Chong-Qing Cheng[1,*] & Jinxin Xue[2]

[1]*School of Mathematical Sciences, Nanjing Normal University, Nanjing* 210023*, China;*
[2]*Department of Mathematical Sciences, Tsinghua University, Beijing* 100084*, China*

Email: chengcq@nju.edu.cn, jxue@tsinghua.edu.cn

Received April 2, 2022; accepted February 24, 2023; published online May 11, 2023

Abstract In this paper, we prove that the nearly integrable system of the form

$$H(x,y) = h(y) + \varepsilon P(x,y), \quad x \in \mathbb{T}^n, \quad y \in \mathbb{R}^n, \quad n \geqslant 3$$

admits orbits that pass through any finitely many prescribed small balls on the same energy level $H^{-1}(E)$ provided that $E > \min h$, if h is convex, and εP is typical. This settles the Arnold diffusion conjecture for convex systems in the smooth category. We also prove the counterpart of Arnold diffusion for the Riemannian metric perturbation of the flat torus.

Keywords Arnold diffusion, normal form, Aubry set, normally hyperbolic invariant cylinder, cohomological equivalence, ladder

MSC(2020) 37J25, 37J40, 37J51

Citation: Cheng C-Q, Xue J X. Arnold diffusion for nearly integrable Hamiltonian systems. Sci China Math, 2023, 66: 1649–1712, https://doi.org/10.1007/s11425-022-2118-1

1 Introduction

In this paper, we consider nearly integrable Hamiltonian systems of the form

$$H(x,y) = h(y) + \varepsilon P(x,y), \quad (x,y) \in T^*\mathbb{T}^n, \quad n \geqslant 3, \tag{1.1}$$

where h is strictly convex, i.e., the Hessian matrix $\frac{\partial^2 h}{\partial y^2}$ is positive definite. It is also assumed that both h and P are the C^r-functions with $7 \leqslant r \leqslant \infty$ and $\min h = 0$.

1.1 The problem and the theorems

The problem of studying the (in)stability of the above system H was considered to be the fundamental problem of Hamiltonian dynamics by Poincaré. According to the celebrated Kolmogorov-Arnold-Moser (KAM) theorem, there exists a large measure Cantor set of Lagrangian tori on which the dynamics is conjugate to irrational rotations and the oscillation of the slow variable (or called the action variable) y is at most $O(\sqrt{\varepsilon})$. The KAM theorem also excludes the possibility of large oscillations of y in the case

* Corresponding author

of $n = 2$ since each energy level, which is three-dimensional, is laminated by two-dimensional KAM tori and each orbit either stays on a KAM torus or is confined between two tori. For $n \geqslant 3$, there does not exist topological obstruction for the slow variable y to have the $O(1)$ oscillation. Thus Arnold [3,4] made the following conjecture.

Conjecture (See [3,4]). For any two points y' and y'' on the connected level hypersurface of h in the action space, there exist orbits of (1.1) connecting an arbitrary small neighborhood of the torus $y = y'$ with an arbitrary small neighborhood of the torus $y = y''$, provided that $\varepsilon \neq 0$ is sufficiently small and P is generic.

In this paper, we prove the conjecture for convex Hamiltonians in the smooth category in the sense of cusp-residual genericity. We next give the precise statement of our main theorem.

By adding a constant to H and introducing a translation $y \to y + y_0$, one can assume $\min h(y) = h(0) = 0$. For $E > 0$, let $H^{-1}(E) = \{(x, y) : H(x, y) = E\}$ denote the energy level set, and $B \subset \mathbb{R}^n$ denote a ball in $\mathbb{R}^n$ such that $\bigcup_{E' \leqslant E+1} h^{-1}(E') \subset B$.

Let $r_0 \geqslant 7$ be a positive number and $r_0 \leqslant r \leqslant \infty$. Let $\mathfrak{S}^r \subset C^r(\mathbb{T}^n \times B)$ (resp. $\mathfrak{B}^r \subset C^r(\mathbb{T}^n \times B)$) be the set of functions $f \in C^r(\mathbb{T}^n \times B)$ satisfying $\|F\|_{C^{r_0}} = 1$ (resp. $\|F\|_{C^{r_0}} < 1$). For a perturbation P independent of y that we call the Mañé perturbation, we use the same notation $\mathfrak{S}^r \subset C^r(\mathbb{T}^n)$ (resp. $\mathfrak{S}^r \subset C^r(\mathbb{T}^n)$) for the set of C^r functions on $\mathbb{T}^n$ with C^{r_0}-norm 1 (resp. < 1).

Definition 1.1 (The cusp-residual set). A set $\mathfrak{C}$ is said to be *cusp-residual* in $\mathfrak{B}^r$ if there exists a residual set $\mathfrak{R} \subset \mathfrak{S}^r$ in the C^r topology such that for each $P \in \mathfrak{R}$, there are a number $a_P \in (0, 1]$ that depends on $P \in C^{r_0}$ continuously, and a residual set $R_P \subset (0, a_P)$ such that $\mathfrak{C} = \cup\{\lambda P \mid P \in \mathfrak{R}, \lambda \in R_P\}$.

Let Φ_H^t denote the Hamiltonian flow determined by H. Given an initial value (x, y), $\Phi_H^t(x, y)$ generates an orbit of the Hamiltonian flow $(x(t), y(t))$. An orbit $(x(t), y(t))$ is said to *visit* $B_\varrho(y_0) \subset \mathbb{R}^n$ if there exists a $t \in \mathbb{R}$ such that $y(t) \in B_\varrho(y_0)$, a ball centered at y_0 with radius ϱ. Our main theorem is as follows.

Theorem 1.2. *Given any small $\varrho > 0$, there exist an ε and a cusp-residual set $\mathfrak{C} \subset C^r(\mathbb{T}^n \times B)$ for $n \geqslant 3$ such that for each $P \in \mathfrak{C}$ and given finitely many small balls $B_\varrho(y_i) \subset \mathbb{R}^n$, where $y_i \in h^{-1}(E)$ with $E > \min h$, there exists the Hamiltonian flow Φ_H^t admitting orbits which visit the balls $B_\varrho(y_i)$ in any prescribed order. Moreover, the theorem still holds if we replace the function space $C^r(\mathbb{T}^n \times B)$ by $C^r(\mathbb{T}^n)$.*

In particular, the theorem implies that for any small ϱ, there exist ϱ-dense orbits on the energy level provided that ε is small enough and P is chosen in the cusp-residual set. We also refer readers to the paper[1)] for a survey of our series of works, including in particular an outline of the proof of the main theorem.

Note that in the above theorem, we allow the perturbation to be in $C^r(\mathbb{T}^n)$, i.e., depending only on the angular variables, which is called the Mañé perturbation (see [35]). This enables us to prove the following theorem on Arnold diffusion for toral geodesic flows.

Let $S \in \mathrm{SL}(n, \mathbb{R})$ be a matrix of determinant 1 and $\mathbb{T}_S^n := \mathbb{R}^n/(S\mathbb{Z}^n)$ be a torus determined by S. The flat metric on $\mathbb{R}^n$ naturally induces a flat metric $ds^2 = \sum_{i=1}^n dx_i^2$ on $\mathbb{T}_S^n$, where x_i $(i = 1, 2, \dots, n)$ are coordinates on $\mathbb{R}^n$. We consider the perturbation of the metric ds^2 of the form

$$ds^2_{\varepsilon\Delta} = \sum_{i=1}^n dx_i^2 + \sum_{i,j} \varepsilon d_{ij}(x) dx_i dx_j,$$

where $\Delta(x) = (d_{ij}(x)) \in \mathrm{Sym}^r(\mathbb{T}_S^n)$ and $\mathrm{Sym}^r(\mathbb{T}_S^n)$ is the space of symmetric $n \times n$ matrices whose entries are in $C^r(\mathbb{T}_S^n)$. We define the C^r-norm on $\mathrm{Sym}^r(\mathbb{T}_S^n)$ as $\|\Delta\|_{C^r} = \sum_{i \leqslant j} |d_{ij}|_{C^r}$, where the $|\cdot|_{C^r}$-norm is the usual one. With the C^r-norm, we introduce the sphere $\mathfrak{S}^r$ and the ball $\mathfrak{B}^r$ in $\mathrm{Sym}^r(\mathbb{T}_S^n)$ as before.

Theorem 1.3. *Given any small $\varrho > 0$, there exist an ε and a cusp-residual set $\mathfrak{C} \subset \mathrm{Sym}^r(\mathbb{T}_S^n)$ for $n \geqslant 3$ such that for each $\Delta \in \mathfrak{C}$ and given finitely many small balls of radius ϱ in the unit tangent bundle of $(\mathbb{T}_S^n, ds^2_{\varepsilon\Delta})$, there exists an orbit of the geodesic flow visiting the balls in any prescribed order.*

1) MSRI proceeding of "Hamiltonian systems, from topology to applications through analysis", to appear.

This theorem shows that a typical perturbation of the geodesic flow of the flat torus $\mathbb{T}^n_S$ admits almost dense orbits in the unit tangent bundle.

If we restrict only to conformal perturbations, i.e., the perturbed metrics of the form

$$ds^2_{\varepsilon V} = (1+\varepsilon V)\sum_{i=1}^{n} dx_i^2 \quad \text{on } \mathbb{T}^n_S,$$

where $V \in C^r(\mathbb{T}^n_S)$, we have the following theorem.

Theorem 1.4. *Given any small $\varrho > 0$, there exist an ε and a cusp-residual set $\mathfrak{C} \subset \mathrm{C}^r(\mathbb{T}^n_S)$ for $n \geqslant 3$ such that for each $V \in \mathfrak{C}$ and given finitely many small balls of radius ϱ in the unit tangent bundle of $(\mathbb{T}^n_S, ds^2_{\varepsilon V})$, there exists an orbit of the geodesic flow visiting the balls in any prescribed order.*

1.2 Historical remarks

Arnold first introduced the system

$$H(I,\theta,y,x,t) = \frac{y^2}{2} + \frac{I^2}{2} + (\cos x - 1)(1+\varepsilon(\sin\theta + \sin t)) \tag{1.2}$$

with $(I,\theta;y,x;t) \in T^*\mathbb{T}^1 \times T^*\mathbb{T}^1 \times \mathbb{T}^1$ in [1] and proved that the orbit exists with $I(0) < A$ and $I(T) > B$ for any $A < B$. The idea is that when $\varepsilon = 0$, the subsystem $H_0 = \frac{y^2}{2} + (\cos x - 1)$ admits a hyperbolic fixed point $(y,x) = (0,0)$ with stable and unstable manifolds (W^s and W^u) both the homoclinic orbit on the energy level $H_0 = 0$. When viewed in the phase space, this gives a normally hyperbolic invariant cylinder (NHIC, see Appendix B) $\mathcal{C} = \{(I,\theta,0,0)\} \subset T^*\mathbb{T}^2$ foliated by circles with constant I. Note that the perturbation $\varepsilon(\cos x - 1)(\sin\theta + \sin t)$ is special since it vanishes on the NHIC. When the perturbation is turned on, the NHIC is untouched but the stable and unstable manifolds (W^s and W^u) are made intersect transversely for each I, and so are all nearby I and I'. Then the diffusion orbit can be found by following the stable-unstable paths.

As the perturbation in Arnold's example is special, it was then natural to ask if the similar result holds if we replace the perturbation by a generic one and assume the presence of an NHIC. Such systems are called of *a priori* unstable type in literature. The NHIC $\mathcal{C}$ persists under the perturbation, but the main difficulty lies in the fact that the dynamics on the perturbed NHIC $\mathcal{C}_\varepsilon$ has a Cantor set of invariant curves and nearby curves may have gaps of width $O(\sqrt{\varepsilon})$ by the KAM theory, while the stable and unstable manifolds of each invariant curve can split at most in order ε, and thus Arnold's mechanism itself is not sufficient. Orbits crossing the gaps were known to Birkhoff [11, Chapter VIII] and Mather [38], but to utilize Arnold's mechanism to cross the invariant curves, one has to show that stable and unstable manifolds of all the invariant curves intersect transversely for generic perturbations, which is a highly nontrivial problem since there are uncountably many invariant curves. This genericity problem was solved by Cheng and Yan [17, 18] exploiting the regularity of weak KAM solutions with respect to the cohomology classes. We refer the readers to the ICM talk of Bernard [8] for a survey of difficulties of *a priori* unstable systems and Cheng-Yan's contribution, and [7, 22, 24, 28, 42, 43] for other works on *a priori* unstable systems.

The study of systems of *a priori* unstable type is important for understanding the above conjecture, since such a system can be obtained via a normal form in the part of the phase space where $\boldsymbol{\omega}(y) := \partial_y h(y)$ admits $n-2$ resonances, which gives rise to the result [10].

Definition 1.5 (Resonance). A frequency $\boldsymbol{\omega}(y) = \frac{\partial h}{\partial y} \neq 0$ is said to admit *a resonance relation*, if there exists an integer vector $\boldsymbol{k} \in \mathbb{Z}^n \setminus \{0\}$ such that $\langle \boldsymbol{k}, \boldsymbol{\omega}(y)\rangle = 0$ at the point y. The number of linearly independent resonance relations is called the *multiplicity* of the resonance. A nonzero frequency is called a *complete resonance point* if the multiplicity is $n-1$.

However, these regions of $(n-2)$-resonances are not connected and it is not avoidable to encounter complete resonances. In particular, when $n = 3$, only single and double resonances occur and *a priori* unstable systems can be used to model the single resonance. An energy level $\{h(y) = E\}$ is a 2-sphere, on which a resonance relation $\langle \boldsymbol{k}, \boldsymbol{\omega}(y)\rangle = 0$ determines a curve and for some y there is a second resonance $\boldsymbol{k}'$. Arnold [3] identified that the main difficulty for proving the conjecture is the double resonance, where the problem is reduced to a nonperturbative mechanical system of the form of two degrees of freedom

$$\mathsf{G}(\mathsf{x}, \mathsf{y}) = \frac{1}{2}\langle A\mathsf{y}, \mathsf{y}\rangle + V(\mathsf{x}), \quad (\mathsf{x}, \mathsf{y}) \in T^*\mathbb{T}^2, \tag{1.3}$$

where A is positive definite and $\max V = 0$. We derive this subsystem in Section 6.

Mather [39] announced a version of Theorem 1.2 in the case of $n = 3$ for C^r perturbations with $r = 3, 4, \ldots, \infty, \omega$, and made the following main contributions among others: (1) developing the variational method for twist maps and Tonelli Lagrangian systems, now known as the Aubry-Mather theory and the Mather theory, respectively; (2) formulating a notion of cusp-residual genericity; (3) the partial result on dynamics around the strong double resonance, such as minimal homoclinic orbits for degenerate Jacobi metrics [40], corresponding to the case of the zero energy level of the system (1.3). However, Mather was not able to fulfill his announcement. Theorem 1.2 with $n = 3$ was proved by the first author in a series of works [13–15, 20]. In particular, the author discovered a mechanism for crossing the strong double resonance (see [13] and Figure 1) which also plays an important role in the present paper. There is also another work [32] following the strategy outlined by Mather.

We emphasize the role played by the genericity argument of [17, 18] (reproduced in Appendix E) in our work, which is the only currently known method to establish the genericity in the variational approach. It relies crucially on the Aubry-Mather theory for twist maps in $T^*\mathbb{T}^1$ (essentially Lemma E.8), which is why we have to perform $n - 2$ steps of reductions of orders to search for NHICs homeomorphic to $T^*\mathbb{T}^1$. An improvement of the argument to Mañé perturbations (see [13, Theorem 4.2]) allows us to get Theorems 1.3 and 1.4, as well as the $C^r(\mathbb{T}^n)$ part of Theorem 1.2. It needs to construct bump functions as perturbations, which is the main obstruction to proving Arnold diffusion for analytic perturbations. It also plays an important role in other experts' work such as [10, 32].

1.3 The outline of the proof

In this subsection, we give an outline of the proof.

The starting point of the proof is a normal form package which we expect to have wide applications beyond Arnold diffusion. The package includes a KAM normal form around resonances and a number of symplectic transformations and reductions.

The KAM normal form dictates that close to $y^\star$ where $\boldsymbol{\omega}(y^\star)$ is resonant against integer vectors $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m$, a symplectic transformation can be introduced to transform the Hamiltonian to the one with the same unperturbed part but with the perturbation dominated by $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} P$ where $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m}$ is the L^2-projector to the Fourier modes in $\mathrm{span}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$. In particular, for the single resonance, i.e., $m = 1$, the normal form package gives us a Hamiltonian of *a priori* unstable type and the nondegenerate global maximum of $\Pi_{\boldsymbol{k}_1} P(y, \cdot)$ corresponds to an NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$ when y is varied.

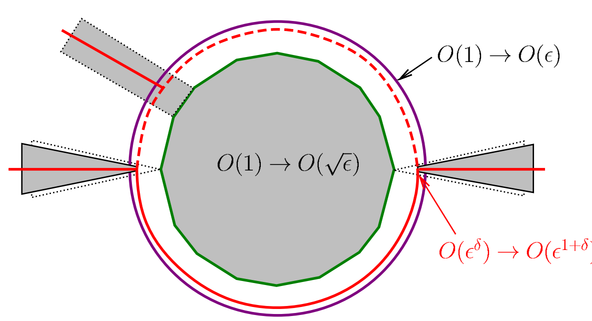

Figure 1 (Color online) The $n = 3$ case. The green curve encloses the flat of the α-function of (1.3), and the red curve is of cohomology equivalence

The $n = 3$ case was outlined above, and we proceed to the case of $n > 3$. We find normally hyperbolic invariant cylinders (NHICs) are $\sqrt{\varepsilon}$-away from the complete resonance (the resonance with multiplicity $n-1$) and study the dynamics within a $\sqrt{\varepsilon}$-neighborhood of the complete resonance.

First, away from the complete resonance, using a scheme of reductions of orders, we find two-dimensional NHICs restricted to which the time-1 map of the system is a twist map and construct diffusion orbits as in *a priori* unstable systems. The idea of the scheme is to consider the frequency path along which there are at least $(n-2)$ linearly independent resonant integer vectors $\boldsymbol{k}', \boldsymbol{k}'', \ldots, \boldsymbol{k}^{(n-2)} \in \mathbb{Z}^n$ forming a hierarchy $|\boldsymbol{k}^{(i)}| \ll |\boldsymbol{k}^{(i+1)}|$, $i = 1, \ldots, n-3$ except that for finitely many points, there are $(n-1)$ linearly independent resonant integer vectors forming such a hierarchy $|\boldsymbol{k}^i| \ll |\boldsymbol{k}^{i+1}|$, $i \neq j$ with two vectors having comparable lengths. We show that for any two balls in the frequency space of a given energy level, there is such a frequency curve with a hierarchy structure shadowing a path with the Diophantine property (see Lemma 13.1). The hierarchy structure allows us to treat $\Pi_{\boldsymbol{k}',\boldsymbol{k}'',\ldots,\boldsymbol{k}^{(i+1)}} P$ as a small perturbation of $\Pi_{\boldsymbol{k}',\boldsymbol{k}'',\ldots,\boldsymbol{k}^{(i)}} P$ so that the NHICs in the latter system persists also in the former. With the persistence and symplecticity of the NHICs (see [23]), we restrict the Hamiltonian to the NHICs to get a system of fewer degrees of freedom. By repeated reductions of orders utilizing the hierarchy structure, we eventually obtain two-dimensional NHICs.

Second, the dynamics near the complete resonance, with the existence of NHICs unknown, is much more delicate. In particular, repeated order reductions are not allowed near the complete resonance due to the lack of regularity of the NHICs after the first step of the order reduction (generically only C^{1+} by Theorem B.2 of the normally hyperbolic invariant manifold (NHIM)). The mechanisms of crossing the double resonance in the $n = 3$ case are not sufficient to cross the complete resonance here. Indeed, when viewed in the space of cohomological classes, the two channels corresponding to two NHICs in the phase space that we would like to find orbits to connect typically have a misalignment in the extra dimensions so that they cannot be connected by the paths constructed in [13] (see Figure 1). To overcome this difficulty, we find a new mechanism to bridge the channels complementary to the paths obtained by the mechanism of [13] (the blue path of Figure 2).

A main disadvantage of our proof is that the speed of the diffusion orbits gets slowed down as the dimension n gets larger, which is unnatural considering statistical physics. This is because the argument relies crucially on the complete understanding of the Aubry-Mather sets in the two-dimensional case. So we propose the following open problem.

Open problem 1. Find an effective proof of the main result, which does not rely on the Aubry-Mather theory for twist maps and gives more abundant and faster diffusion orbits as n gets larger.

In the conclusion part of the main theorems above, we only get ϱ-dense orbits for small enough ε. One may wonder whether we can get dense orbits, which would verify the quasi-ergodic hypothesis. The following problem was asked by Herman [29].

Open problem 2 (See [29])**.** Find a C^∞-Hamiltonian H in any small C^k-neighborhood $k \geqslant 2$ of $H_0(r) = 1/2\|r\|^2$ such that the Hamiltonian flow has a dense orbit on $H^{-1}(1/2)$. One may even ask whether a generic Hamiltonian does so.

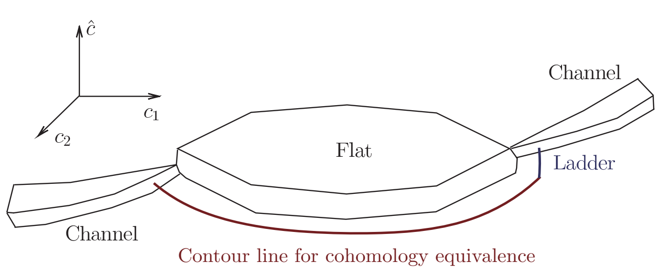

Figure 2 (Color online) The pizza and the ladder climbing in the $n > 3$ case, which project to Figure 1 on the c_1-c_2 plane

The analytic case was announced in [39] but remains open until now. Built on our work on the phase space dynamics, the main difficulty is to extend the genericity argument of [17, 18] to allow analytic perturbations. Furthermore, the convexity is assumed for the purpose of applying the variational method. However, this excludes lots of interesting integrable systems with singular fibers (see [25]). Thus we propose to study the following open problem.

Open problem 3. Study the problem of Arnold diffusion for perturbations of more general integrable systems, for example, Hitchin integrable systems [31] and quantum integrable systems, etc., in the analytic category.

2 The proof of the main theorem

In this section, we prove the main theorem based on some propositions. Please refer to Appendix A for a brief introduction of the Mather theory where more notations are given including the Tonelli Hamiltonian H and Lagrangian L, minimal measures, cohomology class c, rotation vector h, Mather sets $\tilde{\mathcal{M}}(c)$ and $\tilde{\mathcal{M}}_h$, Aubry set $\tilde{\mathcal{A}}(c)$, Mañé set $\tilde{\mathcal{N}}(c)$, α- and β-functions, weak KAM solutions $u_c^{\pm}$ and barrier functions B_c, etc. To proceed, it is important to keep in mind the dictionary in Table 1.

Notation 1. Our convention of using $|\cdot|$ is as follows:

- It is the usual absolute value when applied to real or complex numbers.
- It is the ℓ_1 norm when applied to an integer vector $\boldsymbol{k}\in\mathbb{Z}^n$ which is a *row vector*.
- It is the ℓ_∞ norm when applied to a frequency $\omega\in\mathbb{R}^n$ which is a *column vector*.

We use $\|\cdot\|$ to denote the Euclidean norm.

2.1 The general strategy

Given two points y' and y'' with $h(y')=h(y'')$ and a small number $\varrho>0$, we wish to find an orbit $\{(x(t),y(t)),\ t\in[0,T]\}$ of the system (1.1) for some time T such that $|y(0)-y'|<\varrho$ and $|y(T)-y''|<\varrho$. Let $B_R(0)$ be an open ball of radius R centered at the origin where R is chosen such that $y',y''\in B_R(0)$. Then the map $\boldsymbol{\omega}:=\partial h: B_R(0)\to\boldsymbol{\omega}(B_R(0))$ is a diffeomorphism by the convexity of h. For the integrable system ($\varepsilon=0$ in (1.1)), the vector $\boldsymbol{\omega}(y)$ is the frequency of the rotation $x\mapsto x+\boldsymbol{\omega}(y)t \bmod \mathbb{Z}^n$, $t\in\mathbb{R}$ on the torus $\mathbb{T}^n\times\{y\}$, so we call $\boldsymbol{\omega}(B_R(0))$ the frequency space. Therefore to find the path $\{y(t)\}_{t=0}^T$ connecting neighborhoods of y' and y'' in the perturbed system, we can instead find a path in the frequency space connecting neighborhoods of $\boldsymbol{\omega}(y')$ and $\boldsymbol{\omega}(y'')$.

For the nearly integrable system (1.1), the action variable y is not conserved, nor is $\boldsymbol{\omega}(y)$. Instead, we use the cohomology class $c\in H^1(\mathbb{T}^n,\mathbb{R})$ as a substitute of the action variable y, the homology class (or called the rotation vector of the Mather set) $h\in H_1(\mathbb{T}^n,\mathbb{R})$ as a substitute of the frequency $\boldsymbol{\omega}(y)$ and the α-function as a substitute of the unperturbed Hamiltonian $h(\cdot)$. In fact, when $\varepsilon=0$ and we choose $y=c$, then the unperturbed Hamiltonian $h(y)=\alpha(c)$ and $\boldsymbol{\omega}(y)=\partial h(y)=\partial\alpha(c)$, and the Aubry set $\tilde{\mathcal{A}}(c)$ is exactly the invariant torus $\mathbb{T}^n\times\{y\}$ and it has the rotation vector $h=\partial\alpha(c)=\boldsymbol{\omega}(y)$. When $\varepsilon>0$, though $y(t)$ is no longer conserved, we can still associate an Aubry set $\tilde{\mathcal{A}}(c)$ with each cohomology class c and there are a well-defined α-function $\alpha: H^1(\mathbb{T}^n,\mathbb{R})\to\mathbb{R}$ and its Legendre dual $\beta: H_1(\mathbb{T}^n,\mathbb{R})\to\mathbb{R}$.

Table 1 The dictionary of hyperbolic and variational objects

Hyperbolic objects	Variational objects
The hyperbolic invariant set	The Aubry or Mather set
Stable/unstable manifolds	Graphs of differentials of weak KAMs
Intersections of stable/unstable manifolds	Critical points of the barrier function
Minimal homo- or hetero-clinic orbits	The Mañé set\the Mather set

2.2 The abstract variational framework

In this subsection, we explain our variational mechanisms of constructing orbits connecting two Aubry sets.

Roughly speaking, a *generalized transition chain* is such a path Γ: $[0,1] \to H^1(M,\mathbb{R})$ that for any $s, s' \in [0,1]$ with $|s-s'| \ll 1$, the Aubry sets $\tilde{\mathcal{A}}(\Gamma(s))$ and $\tilde{\mathcal{A}}(\Gamma(s'))$ are connected by an orbit. Let us formulate the definition of the generalized transition chain for the autonomous Tonelli Hamiltonian $H: T^*M \to \mathbb{R}$ where $M = \mathbb{T}^n$ with $n \geqslant 3$ (the nonautonomous version is given in Appendix D). To use the variational method, we perform a Legendre transformation to obtain a Tonelli Lagrangian system $L: TM \to \mathbb{R}$ which defines the Euler-Lagrange flow $\Phi_L^t: TM \to TM$, $t \in \mathbb{R}$. An orbit $(\gamma, \dot\gamma)$ of the Euler-Lagrange flow Φ_L^t is said to *connect two Aubry sets* if the α-limit set of the orbit is contained in one Aubry set and the ω-limit set is contained in the other.

Definition 2.1 (The generalized transition chain: The autonomous case)**.** Two cohomology classes $c, c' \in H^1(M,\mathbb{R})$ are said to be joined by a *generalized transition chain* if there exists a continuous path $\Gamma: [0,1] \to H^1(M,\mathbb{R})$ such that $\Gamma(0) = c$, $\Gamma(1) = c'$ and $\alpha(\Gamma(s)) \equiv E > \min\alpha$, and for each $s \in [0,1]$, at least one of the following cases takes place:

(H1) In some finite covering manifold $\check\pi: \check M \to M$, the Aubry set $\mathcal{A}(\Gamma(s))$ consists of two classes $\mathcal{A}_1(\Gamma(s))$ and $\mathcal{A}_2(\Gamma(s))$. There are two open domains N_1 and N_2 with $\bar N_1 \cap \bar N_2 = \emptyset$, a decomposition $\check M = M_1 \times \mathbb{T}^\ell$, $(n-\ell-1)$-dimensional disks $\{O_m \subset M_1\}$ with $\bar O_m \cap \bar O_{m'} = \emptyset$, an $(n-1)$-dimensional disk D_s and two small numbers $\delta_s, \delta_s' > 0$ such that

(i) the Aubry sets $\mathcal{A}_1(\Gamma(s)) \subset N_1$, $\mathcal{A}_2(\Gamma(s)) \subset N_2$ and $\mathcal{A}(\Gamma(s')) \subset (N_1 \cup N_2)$ for each $|s'-s| < \delta_s$;

(ii) $\check\pi\mathcal{N}(\Gamma(s), \check M)\,|_{D_s} \backslash (\mathcal{A}(\Gamma(s)) + \delta_s')$ is non-empty, of which each connected component is contained in $O_m \times \mathbb{T}^\ell$, where $+\delta_s'$ means a δ_s'-neighborhood;

(iii) $\langle \Gamma(s') - \Gamma(s), g\rangle = 0$ holds for each $g \in H_1(\check M, M_1, \mathbb{R})$.

(H2) For each $s' \in (s-\delta_s, s+\delta_s)$, the cohomology class $\Gamma(s')$ is equivalent to $\Gamma(s)$: some section Σ_s and some small neighborhood U of $\mathcal{N}(\Gamma(s)) \cap \Sigma_s$ exist such that $\langle \Gamma(s') - \Gamma(s), g\rangle = 0$ holds for each $g \in H_1(U, \mathbb{Z})$.

Remark 2.2. (1) Item (H1) with $\ell = 0$ is a variational reformulation of Arnold's mechanism (see [1]) which relies on the existence of NHICs on which Aubry sets lie and (H1)(ii) is the variational characterization of the transverse intersection of the stable and unstable manifolds of the Aubry sets.

(2) Item (H2) is called cohomological equivalence (c-equivalence), which will be used in Section 7.

(3) The case (H1) with $\ell > 0$ is a generalization of Arnold's mechanism by allowing the stable and unstable sets of the Aubry sets to have incomplete intersections in the sense that the stable and unstable sets are allowed to merge in the $\mathbb{T}^\ell$ components and are only required to intersect transversely in the M_1 component.

Once such a generalized transition chain exists, one can construct connecting orbits.

Theorem 2.3 (See [13, 34])**.** *If c is connected to c' by a generalized transition chain Γ as in Definition* 2.1*, then*

(1) *there exists an orbit of the Lagrange flow $d\gamma := (\gamma, \dot\gamma): \mathbb{R} \to TM$ which connects the Aubry set $\tilde{\mathcal{A}}(c)$ to $\tilde{\mathcal{A}}(c')$, namely, the α-limit set $\alpha(d\gamma) \subseteq \tilde{\mathcal{A}}(c)$ and the ω-limit set $\omega(d\gamma) \subseteq \tilde{\mathcal{A}}(c')$;*

(2) *for any $c_1, c_2, \ldots, c_k \in \Gamma$ and arbitrarily small $\delta > 0$, there exist times $t_1 < t_2 < \cdots < t_k$ such that the orbit $d\gamma$ passes through the δ-neighborhood of the Aubry set $\tilde{\mathcal{A}}(c_i)$ at the time $t = t_i$.*

For the sake of completeness, we include a proof in Appendix D.3.

The framework is a bit abstract. We illustrate it using Arnold's Example 1.2. We lift the system to $T^*\mathbb{T} \times T^*2\mathbb{T} \times \mathbb{T}$, and thus the NHIC has two copies $\mathcal{C}_l = \{(I, \theta, 0, 0)\}$ and $\mathcal{C}_r = \{(I, \theta, 0, 2\pi)\}$, where $I \in [A, B]$ and $\theta \in \mathbb{T}^1$. We consider the path $\Gamma: [0,1] \to H^1(\mathbb{T}^2, \mathbb{R})$ of the form $\{(I, 0)\}$ and for each I, the Aubry set $\tilde{\mathcal{A}}(I, 0)$ consists of the two circles $\{(I, \theta, 0, 0)\}$ and $\{(I, \theta, 0, 2\pi)\}$, where $\theta \in \mathbb{T}^1$. The Mañé set satisfies Item (H1)(ii) if the ε-perturbation is turned on by the transverse intersection of the stable and unstable manifolds of the two components of the Aubry set. Thus by Theorem 2.3, we get the diffusion orbit in Arnold's example.

2.3 Existence of the generalized transition chain

We have the following more elaborative statement on the existence of the generalized transition chain.

Theorem 2.4. *Let the Hamiltonian system $H = h + \varepsilon P \in C^r(T^*\mathbb{T}^n, \mathbb{R})$ be as in* (1.1) *restricted to the energy level $E > \min h$. For any $\varrho > 0$ and any M open balls $B_1, \dots, B_M$ of radius ϱ centered on $h^{-1}(E)$, there exist some $\varepsilon_0 > 0$ and an open and dense set $\mathfrak{R} \subset \mathfrak{S}^r$ such that for each $P \in \mathfrak{R}$, there exist an ε_P depending on $P \in C^{r_0}$ continuously and a residual set $R_P \subset (0, \min\{\varepsilon_P, \varepsilon_0\})$ such that for all $\varepsilon \in R_P$, the following hold*:

(1) *There exists a continuous frequency path $\{\omega(t)\} \subset H_1(\mathbb{T}^n, \mathbb{R})$ with $\partial\beta(\omega(t)) \in \alpha^{-1}(E)$, $t \in [1, M]$ satisfying*

(*a*) $\boldsymbol{\omega}^{-1}(\omega(i)) \cap B_i \neq \emptyset$, $i = 1, 2, \dots, M$;

(*b*) *each point $\omega(t)$ is resonant with multiplicity at least $n-2$. There are finitely many marked points on $\omega(t)$ denoted by $\omega_1, \dots, \omega_m$, where m is independent of ε, that are resonant with multiplicity $n-1$.*

(2) *On the energy level E, there are finitely many disjoint C^{r-1} normally hyperbolic invariant cylinders* (*NHICs, see Appendix B*) *homeomorphic to $T^*\mathbb{T} \times \mathbb{T}$.*

(3) *For each ω_i $(i = 1, \dots, m)$, there exists a $\lambda_i > 0$ such that*

(*a*) *the Mather sets of rotation vectors $\omega(t)$ with $|\omega(t) - \omega_i| \geqslant \lambda_i\sqrt{\varepsilon}$ for all $i = 1, 2, \dots, m$, lie in the NHICs*;

(*b*) *any continuous curve lying in the interior of $\{\partial\beta(\omega(t)) \mid |\omega(t) - \omega_i| \geqslant \lambda_i\sqrt{\varepsilon}\} \subset \alpha^{-1}(E)$ is a generalized transition chain satisfying* (H1);

(*c*) *the two neighboring connected components $\{\partial\beta(\omega(t)) \mid |\omega(t) - \omega_i| \geqslant \lambda_i\sqrt{\varepsilon}\} \subset \alpha^{-1}(E)$ near $\partial\beta(\omega_i)$ are joined by a generalized transition chain satisfying* (H2).

Next, we explain how the main Theorem 1.2 follows from this theorem.

Proof of Theorem 1.2. Indeed, given balls $B_1, \dots, B_M$ of radius ϱ centered on $h^{-1}(E)$, we first construct a frequency path $\omega(t)$ $(t \in [0, M])$ as stated. By Items (3)(b) and (3)(c) of the above theorem, there exists a continuous curve of the generalized transition chain visiting small neighborhoods of $\partial\beta(\omega(i)) \subset \alpha^{-1}(E)$, $i = 1, \dots, M$. By Theorem 2.3, we see that once a generalized transition chain is known to exist, an orbit can be constructed by shadowing Aubry sets whose cohomology classes are on the chain. By Item (1)(a) and Subsection 2.1, such an orbit necessarily visits the two balls $B_1, \dots, B_M$ as ordered. This proves Theorem 1.2. □

The remaining part of this paper is devoted to proving Theorem 2.4. In the main body, we complete a proof for $n = 4$ in Section 11, which contains essentially all the main ideas. The proof for general $n \geqslant 4$ is given in Appendix C.6.

3 The KAM normal form

In this section, we derive the KAM normal form (see Proposition 3.3), which roughly says that if the frequency vector $\omega^\star$ admits only resonances $\boldsymbol{k}_1, \dots, \boldsymbol{k}_m$, then only Fourier modes in the span of these integer vectors would dominate. The normal form belongs to our general normal form package, which allows us to reveal the rotator + pendulum structure near a resonance. The package has wide applications beyond the purpose of Arnold diffusion. It includes the following:

(1) *The $\sqrt{\varepsilon}$-blowup* (see Subsection 3.1): this is to restrict to an $O(\sqrt{\varepsilon})$ neighborhood in the action space and blow it up to $O(1)$ size. The outcome is Proposition 3.2 below, which has the structure of kinetic energy + potential energy up to an $O(\sqrt{\varepsilon})$-perturbation.

(2) *The resonant normal form* (see Proposition 3.2): this is to introduce a symplectic transformation to suppress all the nonresonant Fourier modes in the potential energy.

(3) *The linear symplectic transformation and the shear transformation* (see Subsection 5.1): this is to perform two linear symplectic transformations to reduce the normal form into a slow-fast system up to a perturbation, where the slow part is a nonperturbative mechanical system independent of ε (see, e.g., (1.3)) and the fast part is integrable whose dynamics is rotating with speed $O(\varepsilon^{-1/2})$.

(4) *Normal hyperbolicity*: in the slow part obtained in Item (3), a nondegenerate maximum point of the potential gives rise to a hyperbolic fixed point in the slow part which further gives rise to an NHIC in the full system. The NHIC persists under small perturbations by the classical hyperbolic theory (see Appendix B).

We remark that this normal form package is completely elementary, which does not involve the variational theory. We apply the variational theory to construct the diffusion orbit in the system after applying the normal form. It is known in [6] that the variational invariant sets such as the Mather set, the Aubry set and the Mañé set are invariant under symplectic transformations.

3.1 The $\sqrt{\varepsilon}$-blowup

We first introduce the C^r-norm ($r \in \mathbb{N}$) as follows.

Definition 3.1 (The C^r-norm)**.** For a function $f(x,y)$ defined on a domain $\mathcal{D} \times \mathbb{T}^n$, we define the C^r norm as

$$|f|_{C^r} := \sup_{y\in\mathcal{D}} \Bigg(\sum_{\boldsymbol{k}\in\mathbb{Z}^n} \sum_{|\alpha|+|\beta|\leqslant r} \bigg| \frac{\partial^{|\alpha|} f^{\boldsymbol{k}}}{\partial y^\alpha}(y) \bigg| (|k^\beta| + 1) \Bigg),$$

where $f^{\boldsymbol{k}}$ is the $\boldsymbol{k}$-th Fourier coefficient and we use the multi-index notation $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{Z}^n$, $\alpha_i, \beta_i \geqslant 0$ and $i = 1, 2, \ldots, n$.

Fix $y^\star \in h^{-1}(E)$. We introduce the $\sqrt{\varepsilon}$-blowup

$$y - y^\star := \sqrt{\varepsilon} Y, \quad t = \tau/\sqrt{\varepsilon}, \quad H(x,y) = \varepsilon \mathsf{H}(x, Y), \tag{3.1}$$

where Y, τ and H are the blowed-up action variable, time and Hamiltonian, respectively. The $\sqrt{\varepsilon}$-blowup is done in the region $\|y - y^\star\| < \sqrt{\varepsilon}\Lambda$ so that $\|Y\| < \Lambda$.

Notation 2. We use the notation $|\cdot|_r$ to denote the C^r norms with respect to the variables x and Y in the blowup system.

Proposition 3.2. *The Hamiltonian after the $\sqrt{\varepsilon}$-blowup becomes*

$$\mathsf{H}(x,Y) = \frac{h(y^\star)}{\varepsilon} + \frac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \frac{1}{2} \langle \mathsf{A} Y, Y \rangle + \mathsf{V}(x) + \mathsf{P}(x, \sqrt{\varepsilon} Y), \tag{3.2}$$

where

(1) $\frac{h(y^\star)}{\varepsilon} + \frac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \frac{1}{2} \langle AY, Y \rangle$ *is the first three terms of the Taylor expansion of* $h(y)$ *around* $y^\star$;

(2) $\omega^\star = \frac{\partial h}{\partial y}(y^\star)$;

(3) $\mathsf{A} = \frac{\partial^2 h}{\partial y^2}(y^\star)$ *is a positive definite constant matrix*;

(4) $\mathsf{V}(x) = P(x, y^\star)$;

(5) $|\mathsf{P}|_{r-3} \leqslant C_{r,\Lambda}(|P|_{C^r} + |h|_{C^r})\sqrt{\varepsilon}$.

Proof. The first four items are evident. We now focus on Item (5). The term P has a decomposition

$$\mathsf{P} = \mathsf{P}_I + \mathsf{P}_{II},$$

where

$$\begin{aligned} \mathsf{P}_I &= \frac{1}{\varepsilon} \bigg(h(y^\star + \sqrt{\varepsilon} Y) - h(y^\star) - \sqrt{\varepsilon} \langle \omega, Y \rangle - \frac{\varepsilon}{2} \langle AY, Y \rangle \bigg) \\ &= \frac{\sqrt{\varepsilon}}{6} \sum_{1\leqslant i,j,k\leqslant n} Y_i Y_j Y_k \int_0^1 \frac{\partial^3 h}{\partial y_i \partial y_j \partial y_k} (t\sqrt{\varepsilon} Y + y^\star) t^2 dt, \\ \mathsf{P}_{II} &= P(x, y^\star + \sqrt{\varepsilon} Y) - P(x, y^\star) = \sqrt{\varepsilon} \bigg\langle Y, \int_0^1 \frac{\partial P}{\partial y}(x, t\sqrt{\varepsilon} Y + y^\star) dt \bigg\rangle. \end{aligned} \tag{3.3}$$

We have the following estimates:

$$\begin{aligned} &\bigg| \frac{\partial^{|\alpha|+|\beta|} \mathsf{P}_{II}}{\partial x^\alpha \partial Y^\beta} \bigg| \leqslant C_{\beta,\Lambda} |P|_{C^r} \sqrt{\varepsilon}^{|\beta|+1}, \quad 0 \leqslant |\beta| \leqslant r-1, \\ &\bigg| \frac{\partial^\beta \mathsf{P}_I}{\partial Y^\beta} \bigg| \leqslant C_{\beta,\Lambda} |h|_{C^r} \sqrt{\varepsilon}^{|\beta|+1}, \quad 0 \leqslant |\beta| \leqslant r-3. \end{aligned} \tag{3.4}$$

Putting these estimates together we get the estimate in Item (5). □

In the following, we assume that $|P|_{C^{r_0}} \leqslant 1$.

3.2 The KAM normal form

In this subsection, we work out a general normal form.

Notation 3. (1) Given a collection of linearly independent irreducible integer vectors $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m \in \mathbb{Z}^n$, $m < n$ and a function $f \in C^r(\mathbb{T}^n)$, we denote by $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} f$ the function consisting of Fourier modes of f in $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$.

(2) We denote by $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} C^r(\mathbb{T}^n)$ the space of C^r-functions on $\mathbb{T}^n$ consisting of Fourier modes in $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$, and it is similar for $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} C^r(T^*\mathbb{T}^n)$.

(3) We use C^r to refer to either of the function spaces $C^r(T^*\mathbb{T}^n)$ and $C^r(\mathbb{T}^n)$.

Proposition 3.3 (The KAM normal form)**.** *Let $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m$ be m $(< n)$ linearly independent irreducible integer vectors. Given any small δ, there exists an $\varepsilon_0 = \varepsilon_0(\delta, \Lambda)$ such that for all $\varepsilon < \varepsilon_0$, the following holds. Let $\omega^\star = \partial h(y^\star)$ satisfy the following*:

$$|\langle \boldsymbol{k}, \omega^\star \rangle| > \varepsilon^{1/3}, \quad \forall\, \boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}, \quad K = (\delta/3)^{-\frac{1}{2}}. \tag{3.5}$$

Then there exists a symplectic transformation ϕ defined on $B_\Lambda(0) \times \mathbb{T}^n$ satisfying $|\phi - \mathrm{id}|_{r_0-1} = O(\varepsilon^{1/6})$ and sending the Hamiltonian H in the equation (3.2) *to the following form*:

$$\mathsf{H} \circ \phi(x, Y) = \frac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \frac{1}{2} \langle \mathsf{A} Y, Y \rangle + \Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \mathsf{V} + \delta \mathsf{R}(x, Y), \tag{3.6}$$

where

(1) *the remainder $\delta\mathsf{R}(x, Y) = \delta\mathsf{R}_I(x) + \delta\mathsf{R}_{II}(x, Y)$, and $\delta\mathsf{R}_I$ consists of all the Fourier modes of V not in the set $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\} \cup \mathbb{Z}^n_K$*;

(2) *the remainders R_I and R_{II} satisfy $|\mathsf{R}_I|_{r_0-2} \leqslant 1$ and $|\mathsf{R}_{II}|_{r_0-5} \leqslant 1$.*

The point of the normal form is that if the frequency vector $\omega^\star$ admits only resonances $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m$, then Fourier modes $\Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \mathsf{V}$ would dominate.

Proof of Proposition 3.3. We decompose the Hamiltonian (3.2) as follows:

$$\mathsf{H} = \frac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \frac{1}{2} \langle \mathsf{A} Y, Y \rangle + \Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \mathsf{V} + R_{\leqslant}(x) + R_{>}(x) + \mathsf{P}(x, \sqrt{\varepsilon} Y),$$

where

(1) $R_{\leqslant}(x) + R_{>}(x)$ consists of all the Fourier modes of $\mathsf{V}(x)$ in $\mathbb{Z}^n \backslash \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$;

(2) the Fourier modes with

$$\boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$$

are put in $R_{\leqslant}$ and those in $\mathbb{Z}^n \setminus (\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\} \cup \mathbb{Z}^N_K)$ are put in $R_>$.
We have the estimate $|R_>|_{r_0-2} \leqslant \delta$ since we have $K = (\delta/3)^{-1/2}$.

Only one KAM iteration is enough. We use a new Hamiltonian $\sqrt{\varepsilon} F$ whose induced time-1 map $\phi^1_{\sqrt{\varepsilon}F}$ gives rise to a symplectic transformation

$$\begin{aligned}
\mathsf{H} \circ \phi^1_{\sqrt{\varepsilon}F} &= \mathsf{H} + \sqrt{\varepsilon}\{\mathsf{H}, F\} + \frac{\varepsilon}{2} \int_0^1 (1-t)\{\{\mathsf{H}, F\}, F\}(\Phi^t_{\sqrt{\varepsilon}F}) dt \\
&= \frac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \frac{1}{2} \langle \mathsf{A} Y, Y \rangle + \Pi_{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_m} \mathsf{V} + \left\langle \omega^\star, \frac{\partial F}{\partial x} \right\rangle \\
&\quad + R_{\leqslant}(x) + R_{>}(x) + \mathsf{P}(x, \sqrt{\varepsilon} Y) \\
&\quad + \sqrt{\varepsilon} \left\langle \mathsf{A} Y + \frac{\partial \mathsf{P}}{\partial Y}, \frac{\partial F}{\partial x} \right\rangle + \frac{\varepsilon}{2} \int_0^1 (1-t)\{\{\mathsf{H}, F\}, F\}(\Phi^t_{\sqrt{\varepsilon}F}) dt,
\end{aligned}$$

where F satisfies the cohomological equation

$$R_{\leqslant}(x) + \left\langle \omega^{\star}, \frac{\partial F}{\partial x} \right\rangle = 0.$$

Notice that F is a function of x only. Notice also that $|P|_{C^{r_0}} \leqslant 1$ and $\mathsf{V}(x) = P(x, y^{\star})$, so we get $\sqrt{\varepsilon}|F|_{r_0} \leqslant \varepsilon^{1/6}$ by solving the cohomological equation with (3.5).

Let $\delta \mathsf{R}_I = R_{>}$. So we have $|\mathsf{R}_I|_{r_0-2} \leqslant 1$. Let

$$\delta \mathsf{R}_{II} = \mathsf{P}(x, \sqrt{\varepsilon}Y) + \sqrt{\varepsilon}\left\langle \mathsf{A}Y + \frac{\partial \mathsf{P}}{\partial Y}, \frac{\partial F}{\partial x} \right\rangle + \frac{\varepsilon}{2}\int_0^1 (1-t)\{\{\mathsf{H}, F\}, F\}(\phi^t_{\sqrt{\varepsilon}F})dt.$$

We have

(1) $|\mathsf{P}|_{r_0-3} \leqslant C\varepsilon^{1/2}$ by Proposition 3.2;

(2) using the derivative estimates of F and the fact that $\|Y\| \leqslant \Lambda$, we find

$$\left| \sqrt{\varepsilon}\left\langle \mathsf{A}Y + \frac{\partial \mathsf{P}}{\partial Y}, \frac{\partial F}{\partial x} \right\rangle \right|_{r_0-4} = O(\varepsilon^{1/6});$$

(3) since $\{\mathsf{H}, F\} = \{\frac{1}{\sqrt{\varepsilon}}\langle \omega^{\star}, Y\rangle + \frac{1}{2}\langle \mathsf{A}Y, Y\rangle + \mathsf{V}(x) + \mathsf{P}, F\}$, we find

$$\left| \frac{\varepsilon}{2}\int_0^1 (1-t)\{\{\mathsf{H}, F\}, F\}(\phi^t_{\sqrt{\varepsilon}F})dt \right|_{r_0-5} = O(\varepsilon^{1/3}).$$

Therefore, we have $|\delta \mathsf{R}_{II}|_{r_0-5} = O(\varepsilon^{1/6})$ and can make the term $\delta \mathsf{R}_{II}$ less than δ in the C^{r_0-5} norm by decreasing ε. The proof is now completed. □

4 The initial construction of the frequency line

The frequency path $\omega(\cdot)$ in Theorem 2.4 is constructed inductively. In this section, we perform the first step of the construction.

Definition 4.1 (The Diophantine vector)**.** We say that a vector $v \in \mathbb{R}^d$, $d > 1$ is *Diophantine*, if there exist $\alpha, \tau > 0$ such that

$$|\langle v, \boldsymbol{k}\rangle| \geqslant \frac{\alpha}{|\boldsymbol{k}|^{\tau}}, \quad \forall\, \boldsymbol{k} \in \mathbb{Z}^d \setminus \{0\}. \tag{4.1}$$

We define $v \in \mathrm{DC}(d, \alpha, \tau)$.

In this section, we start with a Diophantine vector $\omega^* = (\omega_1^*, \omega_2^*, \ldots, \omega_n^*)$ up to a scalar multiple and find in its ϱ-neighborhood a frequency path ω_- admitting at most two resonances along which we move the first entry arbitrarily. Once we know how to move the first entry arbitrarily, the same strategy can be applied to moving the other entries one by one, and eventually, we can arrive at the ϱ-neighborhood of another Diophantine vector.

Throughout the paper, we use the following notations.

Notation 4 (The *hat* notation and the *tilde* notation)**.** We fix the meaning of the *hat* notation and the *tilde* notation throughout the paper. For a vector $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, we write $v = (\tilde{v}_i, \hat{v}_{n-i})$, where we have $\hat{v}_{n-i} = (v_{i+1}, v_{i+2}, \ldots, v_n) \in \mathbb{R}^{n-i}$ and $\tilde{v}_i = (v_1, \ldots, v_i) \in \mathbb{R}^i$ for $1 < i < n$. We omit the subscript i if $i = 2$.

4.1 The choice of the frequency path and its number-theoretic property

For given $\varrho, \tau > 0$, let $\alpha > 0$ be a small positive number. Consider the frequency segment $\omega_- \in \mathbb{R}^n$ of the form

$$\omega_a = \rho_a\left(a, \frac{P}{Q}\omega_2^*, \frac{p}{q}\omega_2^*, \hat{\omega}^*_{n-3}\right)^t, \quad P, Q, p, q \in \mathbb{Z}, \quad a \in [\omega_1^{*i} - \varrho, \omega_1^{*f} + \varrho], \tag{4.2}$$

$\hat{\omega}^*_{n-3} = (\omega^*_3, \ldots, \omega^*_n) \in \mathrm{DC}(n-3, \alpha, \tau)$ and $\hat{\omega}^*_{n-2} = (\omega^*_2, \hat{\omega}^*_{n-3}) \in \mathrm{DC}(n-2, \alpha, \tau)$. Here, the scalar ρ_a is chosen such that $h(\boldsymbol{\omega}^{-1}(\omega_-)) = E$. We choose $\frac{P}{Q}$ and $\frac{p}{q}$ such that $|\frac{P}{Q} - 1| < \varrho/2$ and $|\frac{p}{q}\omega^*_2 - \omega^*_3| < \varrho/2$ and in addition g.c.d.$(pQ, Pq) = 1$.

The scalar ρ_a does not affect the resonance relations. Since we know that y lies on an energy level E and the energy hypersurface $h^{-1}(E)$ encloses a convex set containing the origin, the equation $h(\boldsymbol{\omega}^{-1}(\omega_-)) = E$, $\boldsymbol{\omega}(y) = \partial h(y)$ determines uniquely ρ_a. For example, when $h(y) = \frac{1}{2}\|y\|^2$, we easily see that

$$\rho_a = \frac{\sqrt{2E}}{\|(a, \frac{P}{Q}\omega^*_2, \frac{p}{q}\omega^*_2, \hat{\omega}^*_{n-3})\|}.$$

Since we assume $\hat{\omega}^*_{n-2} \in \mathrm{DC}(n-2, \alpha, \tau)$, we have at most two resonances as a varies in an interval. We always have the first resonance given by the integer vector $\boldsymbol{k}' = (0, Qp, -qP, \hat{0}_{n-3})$. The g.c.d. of all the components of $\boldsymbol{k}'$ is 1. Then we have

$$\begin{bmatrix} 1 & 0 & 0 & \hat{0}_{n-3} \\ 0 & Qp & -qP & \hat{0}_{n-3} \\ 0 & r & s & \hat{0}_{n-3} \\ \hat{0}_{n-3} & \hat{0}_{n-3} & \hat{0}_{n-3} & \mathrm{id}_{n-3} \end{bmatrix} \begin{bmatrix} a \\ \frac{P}{Q}\omega^*_2 \\ \frac{p}{q}\omega^*_2 \\ \hat{\omega}^*_{n-3} \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ \frac{1}{qQ}\omega^*_2 \\ \hat{\omega}^*_{n-3} \end{bmatrix}, \tag{4.3}$$

where r and s are such that $sQp + rqP = 1$. We denote the $n \times n$ matrix by $M' \in \mathrm{SL}(n, \mathbb{Z})$.

4.2 The double resonance, away from triple or more resonances

In this subsection, we consider the vector (4.2) at the double resonance. We fix some large number K and define $\mathbb{Z}^n_K = \{\boldsymbol{k} \in \mathbb{Z}^n \mid |\boldsymbol{k}| < K\}$. As a varies in an interval, we may encounter *double resonance* points

$$\{\omega_a \mid \langle \boldsymbol{k}, \omega_a \rangle = 0 \ \text{ for some } \boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}'\}\}.$$

There are finitely many such double resonance points, whose number depends only on K.

In this paper, we only consider those resonant integer vectors that are irreducible.

Definition 4.2. An integer vector $\boldsymbol{k} \in \mathbb{Z}^n \setminus \{0\}$ is called *irreducible* if its entries have no common divisor except 1.

The next lemma shows that for fixed K, points along the frequency line ω_- are uniformly bounded away from triple or more resonances.

Lemma 4.3. *Let an irreducible vector $\boldsymbol{k}^o \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}'\}$ be the second resonance of ω_- at some point $a = a^o$, i.e., $\langle \boldsymbol{k}^o, \omega^o \rangle = 0$ for $\omega^o := \omega_{a^o}$. Then there exists a $\mu = \mu(q, Q, K, \tau, \alpha, M')$ such that for all $\boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}^o\}$, we have the estimate*

$$|\langle \boldsymbol{k}, \omega^o \rangle| \geqslant 2nK\mu. \tag{4.4}$$

Proof. We use the linear transformation (4.3) to convert ω_a to the vector

$$\omega'_a = M'\omega_a = \rho_a \left(a, 0, \frac{1}{qQ}\omega^*_2, \hat{\omega}^*_{n-3}\right)^t.$$

Define $\tilde{\boldsymbol{k}}^o = (\tilde{k}^o_1, \tilde{k}^o_2, \ldots, \tilde{k}^o_n) := \boldsymbol{k}^o M'^{-1}$ so that we have

$$0 = \langle \boldsymbol{k}^o, \omega^o \rangle = \langle \boldsymbol{k}^o M'^{-1}, M'\omega^o \rangle =: \langle \tilde{\boldsymbol{k}}^o, \omega'^o \rangle.$$

We have $\tilde{k}^o_1 \neq 0$ since otherwise $\langle \tilde{\boldsymbol{k}}^o, \omega'_a \rangle = 0$ for all a, which is impossible considering that $\hat{\omega}^*_{n-2}$ is Diophantine. We want to bound $|\langle \boldsymbol{k}, \omega^o \rangle|$ from below for all

$$\boldsymbol{k} = (k_1, k_2, \ldots, k_n) \in \mathbb{Z}^n_K \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}^o\}.$$

We set $\tilde{\boldsymbol{k}} = (\tilde{k}_1, \tilde{k}_2, \ldots, \tilde{k}_n) := \boldsymbol{k}M'^{-1}$ to get $\langle \boldsymbol{k}, \omega_a\rangle = \langle \boldsymbol{k}M'^{-1}, M'\omega_a\rangle = \langle \tilde{\boldsymbol{k}}, \omega'_a\rangle$.

We introduce a new vector $\bar{\boldsymbol{k}} = \tilde{\boldsymbol{k}} - \frac{\tilde{k}_1}{\tilde{k}^o_1}\tilde{\boldsymbol{k}}^o = \frac{1}{\tilde{k}^o_1}(\tilde{k}^o_1\tilde{\boldsymbol{k}} - \tilde{k}_1\tilde{\boldsymbol{k}}^o)$. The new vector

$$\tilde{k}^o_1\bar{\boldsymbol{k}} = \tilde{k}^o_1\tilde{\boldsymbol{k}} - \tilde{k}_1\tilde{\boldsymbol{k}}^o := (0, \bar{k}_2, \bar{k}_3, \hat{\bar{\boldsymbol{k}}}_{n-3}) \in \mathbb{Z}^n$$

has zero first entry. We further introduce a new vector $\bar{\bar{\boldsymbol{k}}} = (0, \bar{k}_2, \bar{k}_3, qQ\hat{\bar{\boldsymbol{k}}}_{n-3}) \in \mathbb{Z}^n$. We estimate the norm of $\bar{\bar{\boldsymbol{k}}}$ as

$$|\bar{\bar{\boldsymbol{k}}}| \leqslant qQ|\tilde{k}^o_1\tilde{\boldsymbol{k}} - \tilde{k}_1\tilde{\boldsymbol{k}}^o| \leqslant 2qQ|\tilde{\boldsymbol{k}}^o| \cdot |\tilde{\boldsymbol{k}}| \leqslant 2qQ(\|M'\|_\infty \cdot K)^2.$$

Using the Diophantine conditions and the fact that ω'_a has zero second entry, we have

$$\begin{aligned} |\langle \boldsymbol{k}, \omega^o\rangle| &= |\langle \tilde{\boldsymbol{k}}, \omega'^o\rangle| = |\langle \bar{\boldsymbol{k}}, \omega'^o\rangle| = \left|\frac{1}{\tilde{k}^o_1}\langle(\tilde{k}^o_1\tilde{\boldsymbol{k}} - \tilde{k}_1\tilde{\boldsymbol{k}}^o), \omega'^o\rangle\right| \\ &= \frac{\rho_a}{\tilde{k}^o_1 qQ}|\langle \bar{\bar{\boldsymbol{k}}}, (0, 0, \omega^*_2, \hat{\omega}^*_{n-3})\rangle| \geqslant \frac{\inf_a \rho_a}{\tilde{k}^o_1 qQ}\frac{\alpha}{|\bar{\bar{\boldsymbol{k}}}|^\tau} \\ &\geqslant \frac{\alpha \inf_a \rho_a}{2^\tau (qQ)^{\tau+1}(\|M'\|_\infty K)^{2\tau+1}}. \end{aligned} \tag{4.5}$$

To get the statement, it is enough to define the right-hand side by $2nK\mu$. □

Finally, we have the following fact.

Lemma 4.4. *Let $\boldsymbol{k}^o$ and ω^o be as in Lemma 4.3. Then there exists a matrix $M^o \in \mathrm{SL}(n, \mathbb{Z})$ such that $M'' := M^oM' \in \mathrm{SL}(n, \mathbb{Z})$ has the first row $\boldsymbol{k}^o$ and the second row $\boldsymbol{k}'$.*

Proof. Define $\omega'^o = M'\omega^o$ and $\tilde{\boldsymbol{k}}^o = \boldsymbol{k}^oM'^{-1}$. We have $\langle \boldsymbol{k}^o, \omega^o\rangle = \langle \boldsymbol{k}^oM'^{-1}, \omega'a^o\rangle = 0$. We set the second entry of $\tilde{\boldsymbol{k}}^o$ to be zero and treat it as a vector in $\mathbb{Z}^{n-1}$. We claim that we can find $n-2$ integer vectors in $\mathbb{Z}^{n-1}$ spanning unit volume together with $\tilde{\boldsymbol{k}}^o$. Indeed, suppose, without loss of generality, the first two entries k_1, k_2 of $\tilde{\boldsymbol{k}}^o$ are nonzero and have common divisor 1. This is always possible after permutations of entries. Then using the Euclidean algorithm, we find two numbers s_1 and s_2 such that $k_1s_2 - k_2s_1 = 1$. Extend s_1 and s_2 by adding zeros to a vector in $\mathbb{Z}^{n-1}$ as the second row of the matrix and for the remaining rows, we put 1's on the diagonal and zeros off the diagonal. This gives the desired matrix.

By adding 0 as their second entries, we extend these vectors to be n-dimensional and put these vectors together to get an $n\times n$ matrix M^o whose first row is $\tilde{\boldsymbol{k}}^o := \boldsymbol{k}^o(M')^{-1}$, and the second row is $(0, 1, 0, \ldots, 0)$, and it satisfies the properties stated in the lemma. □

4.3 Resonance submanifolds and their neighborhoods

Next, we find a number μ as the size of the neighborhood of the single resonance manifold to apply the KAM normal forms.

Notation 5. We use the notation $B_\mu(a)$ to denote a ball of radius μ centered at a and the notation $B_\mu(A) := \bigcup_{a\in A} B_\mu(a)$ to denote the μ-neighborhood of a set A.

Notation 6. Denote by ω^o_i $(i = 1, \ldots, m)$ all the double resonances along ω_- for given K above. Without the subscript i, we use ω^o to denote any one of them. It is similar for $\boldsymbol{k}^o$ and $\boldsymbol{k}^o_i$.

Lemma 4.5. *Let ω_-, K, $\boldsymbol{k}'$ and $\boldsymbol{k}^o$ be as above. Let $(\boldsymbol{k}^o)^\perp$ be the $(n-1)$-dimensional space orthogonal to the vector $\boldsymbol{k}^o$. Then there exists a μ such that*

(1) *for all ω in the neighborhood $\mathcal{D}' := B_\mu(\omega_-) \setminus \bigcup_i B_{\varepsilon^{1/3}}(\omega^o_i + (\boldsymbol{k}^o_i)^\perp)$ and for sufficiently small ε, we have*

$$|\langle \boldsymbol{k}, \omega\rangle| > \varepsilon^{1/3}, \quad \forall\, \boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_\mathbb{Z}\{\boldsymbol{k}'\};$$

(2) *for all ω in $\mathcal{D}'' := B_\mu(\omega_-) \cap B_{\varepsilon^{1/3}}(\omega^o + (\boldsymbol{k}^o)^\perp)$ and for all $\boldsymbol{k} \in \mathbb{Z}^n_K \setminus \mathrm{span}_\mathbb{Z}\{\boldsymbol{k}', \boldsymbol{k}^o\}$, we have*

$$|\langle \boldsymbol{k}, \omega\rangle| \geqslant nK\mu. \tag{4.6}$$

Proof. Part 1. We consider two cases depending on whether $\boldsymbol{k}$ in the assumption is one of the double resonant vectors $\boldsymbol{k}^o$ or not.

First we suppose $\boldsymbol{k} = \boldsymbol{k}^o$. Then we get $|\langle \boldsymbol{k}, \omega\rangle| = |\langle \boldsymbol{k}, \omega^o\rangle + \langle \boldsymbol{k}, \omega - \omega^o\rangle| = |\langle \boldsymbol{k}, \omega - \omega^o\rangle|$. By the assumption, the projection of $\omega - \omega^o$ to the vector $\boldsymbol{k}^o$ has the length at least $\varepsilon^{1/3}$. This completes the proof in the case $\boldsymbol{k} = \boldsymbol{k}^o$, since we have $|\boldsymbol{k}| \geqslant 1$.

Next, suppose $\boldsymbol{k} \neq \boldsymbol{k}^o$. Consider the case where the first entry k_1 of $\boldsymbol{k}$ is 0. We have that the vector $\boldsymbol{k}M'^{-1}$ has zero first entry and $M'\omega_a = (a, 0, \frac{1}{qQ}\omega_2^*, \hat{\omega}_{n-3}^*)$ has zero second entry. We have the estimate

$$|\langle \boldsymbol{k}, \omega_a\rangle| = |\langle \boldsymbol{k}M'^{-1}, M'\omega_a\rangle| \geqslant 2nK\mu \tag{4.7}$$

using the Diophantine property of $(\omega_2^*, \hat{\omega}_{n-3}^*)$. We get

$$|\langle \boldsymbol{k}, \omega\rangle| \geqslant |\langle \boldsymbol{k}, \omega_a\rangle| - |\langle \boldsymbol{k}, \omega - \omega_a\rangle| \geqslant 2nK\mu - nK\mu \gg \varepsilon^{1/3}. \tag{4.8}$$

Next, consider the case $k_1 \neq 0$. We change the first entry a of ω_a to $a^o := a - \frac{\langle \boldsymbol{k}, \omega_a\rangle}{k_1}$ to get another frequency vector $\omega^o = \omega_{a^o}$. We have by definition $\langle \boldsymbol{k}, \omega^o\rangle = 0$. This contradicts the assumption that $\boldsymbol{k} \neq \boldsymbol{k}^o$.

Part 2. For given ω as assumed, we have $|\omega - \omega^o| \leqslant \mu$. As $\boldsymbol{k} \in \mathbb{Z}_K^n \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}^o\}$, we have the following estimate:

$$|\langle \boldsymbol{k}, \omega\rangle| = |\langle \boldsymbol{k}, \omega^o\rangle + \langle \boldsymbol{k}, \omega - \omega^o\rangle| \geqslant |\langle \boldsymbol{k}, \omega^o\rangle| - |\langle \boldsymbol{k}, \omega - \omega^o\rangle| \geqslant 2nK\mu - nK\mu \geqslant nK\mu,$$

where in the first inequality, we apply Lemma 4.3 and in the second inequality, we apply the definition of μ. □

5 The reduction of orders for single resonances

In this section, we carry out the last step in the normal form package and obtain the NHIC of codimension 2 in the single resonance regime. Let ω_-, $\boldsymbol{k}'$ and $\boldsymbol{k}^o$ be as in Section 4.

Definition 5.1 (Resonance submanifolds)**.** (1) We define the *single resonance submanifold* associated with the vector $\boldsymbol{k}'$:

$$\Sigma(\boldsymbol{k}') := \{y \in h^{-1}(E) \mid \langle \boldsymbol{k}', \boldsymbol{\omega}(y)\rangle = 0\}.$$

(2) In the single resonance submanifold, we define the *double resonance submanifold* for the resonant vectors $\boldsymbol{k}'$ and $\boldsymbol{k}^o$:

$$\Sigma(\boldsymbol{k}', \boldsymbol{k}^o) := \{y \in h^{-1}(E) \mid \langle \boldsymbol{k}', \boldsymbol{\omega}(y)\rangle = \langle \boldsymbol{k}^o, \boldsymbol{\omega}(y)\rangle = 0\}.$$

Consider the μ-neighborhood $B_\mu(\omega_-)$ of the frequency line ω_-. In the space of action variables, its preimage under the frequency map $\boldsymbol{\omega}$ is $\boldsymbol{\omega}^{-1}(B_\mu(\omega_-))$. We fix a large constant $\Lambda > 0$ and cover the set $\boldsymbol{\omega}^{-1}(B_\mu(\omega_-))$ by balls of radius $\Lambda\sqrt{\varepsilon}$. We choose the covering to be locally finite and the Lebesgue number of the covering to be $0.1\Lambda\sqrt{\varepsilon}$ so that any ball of radius $1/20\Lambda\sqrt{\varepsilon}$ lies entirely in the $\Lambda\sqrt{\varepsilon}$-ball that it intersects.

5.1 The main result

The next result establishes the existence of NHICs.

Proposition 5.2. *There exists an open and dense set $\mathcal{O}_1 = \mathcal{O}_1(\boldsymbol{k}') \subset \Pi_{\boldsymbol{k}'}C^r$ such that for each $P \in C^r$ with $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1$, there exists a $\delta_1 = \delta_1(\Pi_{\boldsymbol{k}'}P)$ such that for all $0 < \delta < \delta_1$,*

(1) *the system* (5.1) (*see Lemma* 5.3 *below*) *based at a point $y^\star \in \Sigma(\boldsymbol{k}')$ and defined on $B_\Lambda(0) \times \mathbb{T}^n$ admits a C^{r-1} NHIC $\mathcal{C}(\boldsymbol{k}')$ homeomorphic to $T^*\mathbb{T}^{n-1}$ with uniform normal hyperbolicity, independent of δ or ε;*

(2) *Mather sets with rotation vectors in $\{\varepsilon^{-1/2}\boldsymbol{\omega}(y^\star + \sqrt{\varepsilon}Y), \|Y\| \leqslant 0.9\Lambda\}$ and perpendicular to $\boldsymbol{k}'$ lie inside $\mathcal{C}(\boldsymbol{k}')$.*

The proposition follows from the normal form for the single resonance and a parametric transversality theorem.

We apply the normal form Proposition 3.3 to the case (1) of Lemma 4.5.

Lemma 5.3. *Let ω_- and μ be as in Lemma 4.5, where $K=(\delta/3)^{-1/2}$ for a small δ. Then there exists an $\varepsilon_1=\varepsilon_1(\delta,\Lambda)$ such that for $\varepsilon<\varepsilon_1$, the following holds. Let $\omega^\star\in\mathcal{D}'$ be as in the case (1) of Lemma 4.5. Then there exists a symplectic transformation ϕ defined on $B_\Lambda(0)\times\mathbb{T}^n$ that is $o_{\varepsilon\to0}(1)$ close to identity in the C^{r_0} norm, such that*

$$\mathsf{H}\circ\phi(x,Y)=\frac{1}{\sqrt{\varepsilon}}\langle\omega^\star,Y\rangle+\frac{1}{2}\langle\mathsf{A}Y,Y\rangle+V(\langle\boldsymbol{k}',x\rangle)+\delta\mathsf{R}(x,Y),\tag{5.1}$$

where

(1) $V(\langle\boldsymbol{k}',x\rangle)=\Pi_{\boldsymbol{k}'}\mathsf{V}$;

(2) $\delta\mathsf{R}(x,Y)=\delta\mathsf{R}_I(x)+\delta\mathsf{R}_{II}(x,Y)$, *where R_I consists of Fourier modes of V not in $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}'\}\cup\mathbb{Z}^n_K$, and we have $|\mathsf{R}_I|_{r_0-2}\leqslant1$ and $|\mathsf{R}_{II}|_{r_0-5}\leqslant1$.*

Using Formula (4.3), we introduce a linear symplectic transformation denoted by $\mathfrak{M}':T^*\mathbb{T}^n\to T^*\mathbb{T}^n$, i.e., $\mathfrak{M}'(x,Y)=(M'x,(M')^{-t}Y)=:(x',Y')$.

In (5.1), we choose $y^\star\in\Sigma(\boldsymbol{k}')$ such that $\omega'^\star=M'\omega^\star$ has zero as the second entry. Applying the symplectic transformation $\mathfrak{M}'$ to the normal form (5.1), we get the following system up to an additive constant:

$$\mathsf{H}'_\delta:=\mathfrak{M}'^{-1*}\mathsf{H}\circ\phi=\frac{1}{\sqrt{\varepsilon}}\langle\omega'^\star,Y'\rangle+\frac{1}{2}\langle AY',Y'\rangle+V(x_2')+\delta R(x',Y'),\tag{5.2}$$

where $A=M'\mathsf{A}M'^t$ and $R(x',Y')=\mathfrak{M}'^{-1*}\mathsf{R}(x,Y)$.

For the purpose of getting the NHIC in Proposition 5.2, we need $V(\cdot)$ to have a nondegenerate global maximum. Since $V(\cdot)=\mathfrak{M}'^{-1*}\Pi_{\boldsymbol{k}'}P(y^\star,\cdot)$, where $y^\star\in\boldsymbol{\omega}^{-1}(\omega_-)$, we need the following result from [20] in order to find NHICs in the system H'_δ.

Proposition 5.4 (See [20, Theorem 3.1]). *Let $F(\cdot,\zeta)\in C^r(\mathbb{T}^1,\mathbb{R})$ with $r\geqslant4$ be Lipschitz in the parameter $\zeta\in[0,1]$. Then there exists an open and dense set $\mathfrak{V}\subset C^r(\mathbb{T}^1,\mathbb{R})$ so that for each $V\in\mathfrak{V}$, it holds simultaneously for all $\zeta\in[0,1]$ that the global maximum of $F(\cdot,\zeta)+V$ is nondegenerate. Moreover, given $V\in\mathfrak{V}$, there are finitely many $\zeta_i\in[0,1]$ such that $F(\cdot,\zeta)+V$ has only one global maximum for $\zeta\neq\zeta_i$ and has two global maxima if $\zeta=\zeta_i$.*

The last result guarantees that $V(\cdot)$ has a nondegenerate global maximum for all $y^\star$ along the path $\boldsymbol{\omega}^{-1}(\omega_-)$. Note that when P is independent of y, the desired property for V follows easily from the Morse lemma.

5.2 The proof of Proposition 5.2

With the above normal form and the parametric transversality result, we are ready to give the proof of Proposition 5.2.

Proof of Proposition 5.2. We first apply Proposition 5.4 to the function $\Pi_{\boldsymbol{k}'}P$ along the segment $y\in\boldsymbol{\omega}^{-1}(\omega_-)$ to get an open and dense set $\mathcal{O}_1$ in $\Pi_{\boldsymbol{k}'}C^r$ such that for each y, the function $\Pi_{\boldsymbol{k}'}P(y,\cdot)\in\mathcal{O}_1$ admits a nondegenerate global maximum up to finitely many bifurcations, where there are two nondegenerate global maxima. Let us now choose a P with $\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_1$ and determine $V(x_2')$ from $\Pi_{\boldsymbol{k}'}P$ by applying the $\sqrt{\varepsilon}$-blowup and Lemma 5.3 so that V has a nondegenerate global maximum.

In (5.2), we neglect the remainder δR to get that the remaining system

$$\mathsf{H}'_0:=\frac{1}{\sqrt{\varepsilon}}\langle\omega'^\star,Y'\rangle+\frac{1}{2}\langle AY',Y'\rangle+V(x_2')$$

admits an NHIC given by

$$\left\{\dot{Y}_2'=\frac{\partial V}{\partial x_2'}=0,\ \dot{x}_2'=\frac{1}{2}\frac{\partial\langle AY',Y'\rangle}{\partial Y_2'}=\sum_{i=1}^n A_{2i}Y_i'=0\right\}.\tag{5.3}$$

The normal hyperbolicity depends only on A and the second-order derivative of V at the global maximum, and hence does not depend on ε or δ. Restricted to the NHIC, we get a system with one fewer degree of freedom due to Theorem B.4.

Let us now make preparations for the application of the theorem of the NHIM. The system H (without the linear transformation) is defined in a Λ-ball in the Y variables since the $\sqrt{\varepsilon}$-blowup is done in a $\Lambda\sqrt{\varepsilon}$ ball. We introduce a C^∞ bump function χ supported in $B_\Lambda(0)$ satisfying $\chi(Y) = 1$ if $\|Y\| < 0.95\Lambda$ and is zero for $\|Y\| > 0.98\Lambda$. To apply the NHIM theorem, we replace the remainder $\delta\mathsf{R}$ in (5.1) by $\chi(Y)(\delta\mathsf{R})$. The modified perturbation vanishes in the region $\{\|Y\| \geqslant 0.98\Lambda\}$ so that the dynamics therein is integrable when restricted to the NHIC which is the unperturbed NHIC. We will show below how to apply the theorem of the NHIM to obtain an NHIC for the modified system. Since the modified system agrees with the original system on $\{\|Y\| \leqslant 0.95\Lambda\}$, the NHIC for the modified system is indeed an NHIC for the original system in the region $\{\|Y\| \leqslant 0.95\Lambda\}$.

We next apply the NHIM Theorems B.2 and B.4. However, there is a subtle point. In the Hamiltonian equations, the vector field in the center manifold is very fast: $\dot{x} = \frac{\omega^\star}{\sqrt{\varepsilon}} + O(1)$. This is a nonstandard setting where the NHIM theorems are applicable. We present the statement and the proof in Appendix B.2. The conclusions of the NHIM theorem still hold since the large term $\frac{\omega^\star}{\sqrt{\varepsilon}}$ is constant, which contributes neither to the variational equation, nor the derivatives of the Hamiltonian flow, and hence has nothing to do with the normal hyperbolicity. The perturbation $\delta\mathsf{R}$ is δ-small in the C^{r_0-5} norm, so its perturbation to the Hamiltonian vector field is δ-small in the C^{r_0-6} norm. By the assumption $r_0 \geqslant 7$ and applying the NHIM theorem (see Theorem B.5), we get an NHIC which is C^{r-1} and is δ-close to the unperturbed one in the C^1-topology as the center Lyapunov exponents are zero.

In this case, we apply Theorem B.4 to restrict the system to the NHIC to get a Hamiltonian system with one fewer degree of freedom. Note that here δ_1 is determined by the normal hyperbolicity which comes from the second-order derivative of V at the global maximum, and hence δ_1 is determined by $\Pi_{\boldsymbol{k}'}P$.

Finally, we study the oscillation of the action variables of orbits in the Mather set. First, we know that for the modified system, all the Mather sets with cohomology classes $\|c\| \leqslant \Lambda$ and rotation vectors perpendicular to $\boldsymbol{k}'$ lie inside the NHIC, since these Mather sets necessarily lie in a small neighborhood of the NHIC if δ is small and a Mather set does not lie on the NHIC, the normal hyperbolicity will push it away from the NHIC violating the invariance of Mather sets. We next show that within the NHIC, the action variables of orbits in the Mather sets have the $O(\sqrt{\delta})$ oscillation. Write the Lagrangian as

$$\mathsf{L}_c(x,\dot{x}) = \frac{1}{2}\langle A^{-1}(\dot{x} - \varepsilon^{-1/2}\omega^\star - c), (\dot{x} - \varepsilon^{-1/2}\omega^\star - c)\rangle - V(x_2) - \delta\chi R - \frac{1}{2}\langle Ac, c\rangle + \alpha(c),$$

where $\mathsf{L}_c(x,\dot{x}) := \mathsf{L}(x,\dot{x}) - \langle c, \dot{x}\rangle + \alpha(c)$ and

$$\mathsf{L}(x,\dot{x}) := \frac{1}{2}\langle A^{-1}(\dot{x} - \varepsilon^{-1/2}\omega^\star), (\dot{x} - \varepsilon^{-1/2}\omega^\star)\rangle - V(x_2) - \delta\chi R.$$

Let μ be a measure in the Mather set of the cohomology class c. Fix a large number C and decompose $\mu = \mu_1 + \mu_2$ such that $\mathrm{supp}\mu_1 \subset \{\|\dot{x} - \varepsilon^{-1/2}\omega^\star - c\| \leqslant C\sqrt{\delta}\} \times \mathbb{T}^n$ and $\mathrm{supp}\mu_2$ lies in the complement. Denote by $\hat{\mu}_i = \frac{1}{m_i}\mu_i$ the normalization of μ_i, where $m_i = \int d\mu_i$, $i = 1, 2$. So we get the action $A_c(\mu) := \int \mathsf{L}_c\, d\mu = m_1 \int \mathsf{L}_c d\hat{\mu}_1 + m_2 \int \mathsf{L}_c d\hat{\mu}_2$. We always have $\int \mathsf{L}_c d\hat{\mu}_1 \geqslant 0$. For the second term, we have

$$\frac{1}{2}\langle A^{-1}(\dot{x} - \varepsilon^{-1/2}\omega^\star - c), (\dot{x} - \varepsilon^{-1/2}\omega^\star - c)\rangle > C^2\|A\|^{-1}\delta/2$$

by the definition of μ_2 and $|V(x_2)|_{\mathrm{supp}\mu_2}| \leqslant \ell\delta^2$ for some constant ℓ, since the Mather set lies on the NHIC and the NHIC undergoes an $O(\delta)$ perturbation from the unperturbed one given by x_2^*, a nondegenerate global maximum of V. We denote by μ_0 the Haar measure supported on the torus $\{\dot{x} = \varepsilon^{-1/2}\omega^\star + c\} \times \{x_2 = x_2^*\}$ and we have $A_c(\mu_0) = \int \mathsf{L}_c\, d\mu_0 = -\frac{1}{2}\langle Ac, c\rangle + \alpha(c) + O(\delta) \geqslant 0$. We also have $\sup|R| \leqslant 1$, so we conclude

$$\int \mathsf{L}_c d\hat{\mu}_2 - A_c(\mu_0) \geqslant \frac{1}{2}C^2\|A\|^{-1}\delta - \ell\delta^2 - \delta.$$

Choosing C large and δ small, we find that $A_c(\mu) \geqslant m_2 A_c(\hat{\mu}_2) > 0$ violating the definition of minimal measure. Part (2) of this proposition is proved since Mather sets intersecting the region $\{\|Y\| \leqslant 0.9\Lambda\}$ have to stay in $\{\|Y\| \leqslant 0.95\Lambda\}$, where the modified system agrees with the original system. □

6 Dynamics around strong double resonances: NHICs

The number of double resonances depends on δ. However, most of the double resonances are weak and can be treated as single resonances. The number of strong double resonances is independent of δ and ε. In this section, we first give a criterion to distinguish weak and strong double resonances, and next show that there exist NHICs close to the strong double resonances.

6.1 Distinguishing weak and strong double resonances

We apply the normal form Proposition 3.3 to the case (2) of Lemma 4.5 to obtain the following lemma.

Lemma 6.1. *Let ω_- and μ be as in Lemma 4.5, where $K = (\delta/3)^{-1/2}$ for a small δ. Then there exists an $\varepsilon_2 = \varepsilon_2(\delta, \Lambda)$ such that for $\varepsilon < \varepsilon_2$, the following holds. Let $\omega^\star \in \mathcal{D}''$ be as in the case (2) of Lemma 4.5. Then there exists a symplectic transformation ϕ defined on $\{|Y| \leqslant \Lambda\} \times \mathbb{T}^n$ that is $o_{\varepsilon\to 0}(1)$ close to identity in the C^{r_0-1} norm such that*

$$\mathsf{H} \circ \phi(x, Y) = \frac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \frac{1}{2}\langle \mathsf{A}Y, Y\rangle + V(\langle \boldsymbol{k}', x\rangle, \langle \boldsymbol{k}^o, x\rangle) + \delta \mathsf{R}(x, Y), \tag{6.1}$$

where

(1) $V(\langle \boldsymbol{k}', x\rangle, \langle \boldsymbol{k}^o, x\rangle) = \Pi_{\boldsymbol{k}',\boldsymbol{k}^o}\mathsf{V}$;

(2) $\delta\mathsf{R}(x, Y) = \delta\mathsf{R}_I(x) + \delta\mathsf{R}_{II}(x, Y)$*, where* R_I *consists of Fourier modes of* V *not in* $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}^o\} \cup \mathbb{Z}^n_K$*, and we have* $|\mathsf{R}_I|_{r_0-2} \leqslant 1$ *and* $|\mathsf{R}_{II}|_{r_0-5} \leqslant 1$.

Consider a double resonance associated with the vector $\boldsymbol{k}^o$. We decompose $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}\mathsf{V}(x)$ in (6.1) in Lemma 6.1 as

$$\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}\mathsf{V}(x) = Z'(\langle \boldsymbol{k}', x\rangle) + Z''(\langle \boldsymbol{k}', x\rangle, \langle \boldsymbol{k}^o, x\rangle), \tag{6.2}$$

where Z' includes all the Fourier harmonics in the $\mathrm{span}\{\boldsymbol{k}'\}$ and Z'' contains the rest.

Notice that Z'' must contain at least one term with $\boldsymbol{k}^o$. Since $\boldsymbol{k}'$ does not depend on δ, we get $|Z''|_{C^{r_0-2}} \leqslant \frac{C}{|\boldsymbol{k}^o|^2}$ for some constant C independent of δ. We first treat $Z'' + \delta\mathsf{R}$ as a perturbation to the truncated Hamiltonian $\frac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \frac{1}{2}\langle \mathsf{A}Y, Y\rangle + Z'(\langle \boldsymbol{k}', x\rangle)$, which has an NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$ following from exactly the same reasoning as in Proposition 5.2. There is a threshold denoted by $\boldsymbol{\delta}$, independent of δ and ε, i.e., the maximal allowable C^1 norm of the perturbation for applying the NHIM theorem (see Appendix B.2) to the NHIC in the truncated Hamiltonian.

Definition 6.2 (Weak and strong double resonances)**.** Suppose that the frequency vector $\omega^\star$ admits two resonances $\boldsymbol{k}'$ and $\boldsymbol{k}^o$. It is called a *weak double resonance*, if we have $\boldsymbol{\delta} > 2\frac{C}{|\boldsymbol{k}^o|^2}$ and the theorem of the NHIM (see Theorem B.2) can be applied to yielding the NHIC as in Proposition 5.2. Otherwise, it is called a *strong double resonance*.

We summarize the above analysis into the following lemma.

Lemma 6.3. *The total number of strong double resonance points is bounded by $(\frac{2C}{\boldsymbol{\delta}})^{n/r'}$, which is independent of ε and δ for given $P \in \mathcal{O}_1$.*

In the following, we treat the weak double resonances in the same way as the single resonances in the last section, and focus on strong double resonances.

6.2 The shear transformation for strong double resonances

In this subsection, we perform the next step of our normal form package, i.e., the linear symplectic transformation and the shear transformation, to write the double resonance normal form into a decoupled form. In particular, we separate a mechanical system of two degrees of freedom.

6.2.1 *The linear symplectic transformation*

In the $\sqrt{\varepsilon}$-blowup and Lemma 6.1, we choose the base point $y^\star \in \Sigma(\boldsymbol{k}')$ so that $\omega^\star = \boldsymbol{\omega}(y^\star) \in \boldsymbol{k}'^\perp$. We introduce the matrix $M'' \in \mathrm{SL}(n,\mathbb{Z})$ in Lemma 4.4 whose first two rows are $\boldsymbol{k}^o$ and $\boldsymbol{k}'$, respectively, and introduce the linear symplectic transformation

$$\mathfrak{M}'' : T^*\mathbb{T}^n \to T^*\mathbb{T}^n, \quad (x,Y) \mapsto (M''x, M''^{-t}Y) =: (x'', Y''). \tag{6.3}$$

We also keep track of the frequency vector $\omega''_a = M''\omega_a = (\nu(a), 0, *, \ldots, *)$ where $\nu(a)$ satisfies $\nu(a^o) = 0$, where a^o is such that $\omega^o = \omega_{a^o}$ is a strong double resonance. We get a Hamiltonian system

$$\mathsf{H}''_\delta := (\mathfrak{M}''^{-1})^* \mathsf{H} \circ \phi = \frac{1}{\sqrt{\varepsilon}}\langle \omega''_a, Y''\rangle + \frac{1}{2}\langle A''Y'', Y''\rangle + V(x''_1, x''_2) + \delta \mathsf{R}'' \tag{6.4}$$

by applying $\mathfrak{M}''$ term by term to (6.1).

6.2.2 *The shear transformation*

In the next lemma, we are going to introduce a linear symplectic transformation called the *shear transformation*, induced by a matrix in $\mathrm{SL}(2n,\mathbb{R})$ but not in $\mathrm{SL}(2n,\mathbb{Z})$ so that it is not a symplectic transformation on $T^*\mathbb{T}^n$. We introduce the following notation.

Notation 7. Given a matrix $S \in \mathrm{SL}(n,\mathbb{R})$, we denote by $\mathbb{T}^n_S$ the torus $\mathbb{R}^n/(S\mathbb{Z}^n)$, where $S\mathbb{Z}^n = \{S\boldsymbol{k} \mid \boldsymbol{k} \in \mathbb{Z}^n\}$.

Lemma 6.4. *There is a linear symplectic transformation from $T^*\mathbb{T}^n$ to $T^*\mathbb{T}^n_S$ defined by*

$$\mathfrak{S}\mathfrak{M}'' : (x,y) \mapsto (SM''x, (SM'')^{-t}y) =: (\mathsf{x},\mathsf{y}) \in T^*\mathbb{T}^n_S,$$

where $S \in \mathrm{SL}(n,\mathbb{R})$ is in (6.11)*, such that the Hamiltonian system $\mathsf{H}\circ\phi$ in Lemma* 6.1 *is reduced to the following Hamiltonian defined on $(SM'')^{-t}B_\Lambda(0) \times \mathbb{T}^n_S \subset T^*\mathbb{T}^n_S$, up to an additive constant*:

$$\mathsf{H}_{S,\delta} := (\mathfrak{S}\mathfrak{M}'')^{-1*}\mathsf{H}\circ\phi = \tilde{\mathsf{G}}(\tilde{\mathsf{x}},\tilde{\mathsf{y}}) + \hat{\mathsf{G}}(\hat{\mathsf{y}}_{n-2}) + \delta R(\mathsf{x},\mathsf{y}), \tag{6.5}$$

where

$$\begin{aligned} \tilde{\mathsf{G}}(\tilde{\mathsf{x}},\tilde{\mathsf{y}}) &= \varepsilon^{-1/2}\omega_{S,1}\mathsf{y}_1 + \frac{1}{2}\langle \tilde{A}\tilde{\mathsf{y}},\tilde{\mathsf{y}}\rangle + V(\tilde{\mathsf{x}}) : T^*\mathbb{T}^2 \to \mathbb{R}, \\ \hat{\mathsf{G}}(\hat{\mathsf{y}}_{n-2}) &= \frac{1}{2}\langle \hat{\mathsf{y}}_{n-2}, B\hat{\mathsf{y}}_{n-2}\rangle + \frac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}_{S,n-2}, \hat{\mathsf{y}}_{n-2}\rangle, \end{aligned} \tag{6.6}$$

where

(1) $\omega_S = SM''\omega^o = (\tilde{\omega}_S, \hat{\omega}_{S,n-2})$ *with* $\omega_{S,2} = 0$ *since* $y^\star \in \Sigma(\boldsymbol{k}')$*, and* $\tilde{\omega}_S = (\omega_{S,1}, \omega_{S,2}) = 0$ *if* $y^\star \in \Sigma(\boldsymbol{k}', \boldsymbol{k}^o)$;

(2) *the two matrices* $\tilde{A}$ *and* $B = (\hat{A} - \breve{A}^t\tilde{A}^{-1}\breve{A})$ *are positive definite, where* $\tilde{A}$, $\breve{A}$ *and* $\hat{A}$ *in* $\mathbb{R}^{2^2}$, $\mathbb{R}^{2\times(n-2)}$ *and* $\mathbb{R}^{(n-2)^2}$*, respectively, form the matrix*

$$A = \begin{pmatrix} \tilde{A} & \breve{A} \\ \breve{A}^t & \hat{A} \end{pmatrix}; \tag{6.7}$$

(3) *the remainder* $R(\mathsf{x},\mathsf{y}) = (\mathfrak{S}\mathfrak{M}'')^{-1*}\mathsf{R}$ *satisfies* $|R|_{r_0-5} < C$*, where the constant C is determined by M'' and hence S is independent of ε or δ.*

Proof. In the proof, for simplicity of notations and without causing confusion, we also remove the $''$ in (6.3). Let

$$\mathsf{G}(Y,x) = \frac{1}{\sqrt{\varepsilon}}\langle \omega''_{a^o}, Y\rangle + \frac{1}{2}\langle AY, Y\rangle + V(x_1,x_2). \tag{6.8}$$

We write the matrix A in the block form of (6.7). We also denote by $\tilde{v} = (v_1, v_2)$ the first two entries of a vector $v \in \mathbb{R}^n$. Next, we have the following formal derivation:

$$\mathsf{G}(Y,x) = \frac{1}{\sqrt{\varepsilon}}\langle \omega, Y\rangle + \frac{1}{2}\langle AY, Y\rangle + V(\tilde{x})$$

$$
\begin{aligned}
&= \frac{1}{2}\langle \tilde{A}\tilde{Y}, \tilde{Y}\rangle + \langle \tilde{Y}, \breve{A}\hat{Y}_{n-2}\rangle + V(\tilde{x}) + \frac{1}{\sqrt{\varepsilon}}\langle \tilde{\omega}, \tilde{Y}\rangle \\
&\quad + \frac{1}{2}\langle \hat{A}\hat{Y}_{n-2}, \hat{Y}_{n-2}\rangle + \frac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}_{n-2}, \hat{Y}_{n-2}\rangle \\
&= \frac{1}{2}\langle \tilde{A}(\tilde{Y} + \tilde{A}^{-1}\breve{A}\hat{Y}_{n-2}), (\tilde{Y} + \tilde{A}^{-1}\breve{A}\hat{Y}_{n-2})\rangle + V(\tilde{x}) + \frac{1}{\sqrt{\varepsilon}}\langle \tilde{\omega}, \tilde{Y}\rangle \\
&\quad - \frac{1}{2}\langle \breve{A}\hat{Y}_{n-2}, \tilde{A}^{-1}\breve{A}\hat{Y}_{n-2}\rangle + \frac{1}{2}\langle \hat{A}\hat{Y}_{n-2}, \hat{Y}_{n-2}\rangle + \frac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}_{n-2}, \hat{Y}_{n-2}\rangle \\
&= \frac{1}{2}\langle \tilde{A}(\tilde{Y} + \tilde{A}^{-1}\breve{A}\hat{Y}_{n-2}), (\tilde{Y} + \tilde{A}^{-1}\breve{A}\hat{Y}_{n-2})\rangle + V(\tilde{x}) + \frac{1}{\sqrt{\varepsilon}}\langle \tilde{\omega}, \tilde{Y}\rangle \\
&\quad + \frac{1}{2}\langle \hat{Y}_{n-2}, (\hat{A} - \breve{A}^t\tilde{A}^{-1}\breve{A})\hat{Y}_{n-2}\rangle + \varepsilon^{-1/2}\langle \hat{\omega}_{n-2}, \hat{Y}_{n-2}\rangle. \qquad (6.9)
\end{aligned}
$$

We perform the following linear shear symplectic transformation denoted by $\mathfrak{S}$:

$$
\begin{bmatrix} \tilde{\mathsf{y}} \\ \hat{\mathsf{y}}_{n-2} \end{bmatrix} = \begin{bmatrix} \mathrm{id}_2 & \tilde{A}^{-1}\breve{A} \\ 0 & \mathrm{id}_{n-2} \end{bmatrix} \begin{bmatrix} \tilde{Y} \\ \hat{Y}_{n-2} \end{bmatrix}, \quad \begin{bmatrix} \tilde{\mathsf{x}} \\ \hat{\mathsf{x}}_{n-2} \end{bmatrix} = \begin{bmatrix} \mathrm{id}_2 & 0 \\ -\breve{A}^t\tilde{A}^{-t} & \mathrm{id}_{n-2} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \hat{x}_{n-2} \end{bmatrix} \qquad (6.10)
$$

so that the blowup system in the new coordinates is written in the form $\mathsf{G} = \tilde{\mathsf{G}} + \hat{\mathsf{G}}$ stated in the lemma. Here, the variables x are local coordinates on $\mathbb{T}^n_S$ and can be viewed as global coordinates on the universal covering space $\mathbb{R}^n$.

We define

$$
S = \begin{bmatrix} \mathrm{id}_2 & 0 \\ -\breve{A}^t\tilde{A}^{-t} & \mathrm{id}_{n-2} \end{bmatrix}, \quad S^{-t} = \begin{bmatrix} \mathrm{id}_2 & \tilde{A}^{-1}\breve{A} \\ 0 & \mathrm{id}_{n-2} \end{bmatrix} \qquad (6.11)
$$

so that the above symplectic transformation $\mathfrak{S}$ is simplified to $\mathsf{x} = Sx$ and $\mathsf{y} = S^{-t}Y$.

Since A is positive definite and the linear symplectic transformation $\mathfrak{S}$ does not change the signature, we see that both $\tilde{A}$ and $B = (\hat{A} - \breve{A}^t\tilde{A}^{-1}\breve{A})$ are positive definite.

Notice that the above matrix S is identity in the $\tilde{x}$ component, and hence the Hamiltonian $\tilde{\mathsf{G}}$ depends on $\tilde{\mathsf{x}}$ $\mathbb{Z}^2$-periodically. So $\tilde{\mathsf{G}}$ is a Hamiltonian defined on $T^*\mathbb{T}^2$. □

Remark 6.5. This lemma implies that configuration space dynamics on $\mathbb{T}^n$ of the system H''_δ ($\delta = 0$) in (6.4) has a skew product structure. The base dynamics is given by the configuration space dynamics on $\mathbb{T}^2$ of $\tilde{\mathsf{G}} : T^*\mathbb{T}^2 \to \mathbb{R}$. Each fiber is a $\mathbb{T}^{n-2}$. The dynamics on each fiber at the point $\tilde{x}$ depends on the base point $\tilde{x}$ by (6.10).

For $\omega^\star \in \mathcal{D}''$, we further distinguish two cases depending on if $\omega^\star$ is in $B_\mu(\omega_-) \cap B_{\Lambda\varepsilon^{1/2}}(\omega^o + (\boldsymbol{k}^o)^\perp)$ or not. If $\omega^\star$ lies in the set, then when choosing the covering defining the $\sqrt{\varepsilon}$-blowup, we require $y^\star$ is such that $\omega^\star = \boldsymbol{\omega}(y^\star)$ is at the strong double resonance. In the following, we will focus mainly on this case. The other case is easy and will be studied in Subsection 6.4.2.

6.3 Hamiltonian systems of two degrees of freedom

Suppose that $\boldsymbol{\omega}(y^\star)$ is a strong double resonance. So in (6.6), the frequency $\tilde{\omega}_S = 0$ and we have obtained a mechanical system

$$
\tilde{\mathsf{G}}(\tilde{\mathsf{x}}, \tilde{\mathsf{y}}) = \frac{1}{2}\langle \tilde{A}\tilde{\mathsf{y}}, \tilde{\mathsf{y}}\rangle + V(\tilde{\mathsf{x}}), \quad (\tilde{\mathsf{x}}, \tilde{\mathsf{y}}) \in T^*\mathbb{T}^2. \qquad (6.12)
$$

We normalize V such that $\max V = 0$ and assume that the maximum is attained at 0 and is nondegenerate. Thus, we see that the point $(\tilde{\mathsf{x}}, \tilde{\mathsf{y}}) = (0, 0)$ is a hyperbolic fixed point of the system.

The system was studied intensively in [13, 14, 20], whose main results will be recalled next. We divide the study of the system into the following three regimes:

(1) low energy regime: energy levels $E \in [0, \Delta)$ for some small Δ;

(2) intermediate energy regime: $E \in [E_-, E_+]$ for any $0 < E_- < E_+ < \infty$;

(3) high energy regime: $E \in [E_*, \infty)$ for some E_* sufficiently large.

Item (1) will be studied in the next subsection. In this subsection, we study Items (2) and (3), and show that there exists an NHIC, bifurcating at most finitely many times and foliated into hyperbolic periodic orbits with the homology class $g = \pm(1, 0)$ in each energy level $[E_-, +\infty)$. Going back to the full system H_δ, we see that such NHICs would give rise to an NHIC of codimension 2 containing Mather sets whose rotation vector is perpendicular to $\boldsymbol{k}'$.

To see why we choose the homology class $g = \pm(1, 0)$, we consider the frequency vector $\omega'' = (\nu(a), 0, *, \ldots, *)$ after the transformation (6.3). The double resonance corresponds to $\nu(a) = 0$ and the crossing double resonance means that we should let $\nu(0)$ go from a positive value to a negative value. Our diffusion orbit, when projected to the subsystem $\tilde{\mathsf{G}}$, will move along the NHIC foliated by periodic orbits with the homology class $g = (1, 0)$, going from high energy to low energy all the way to some small positive energy, and then by the mechanism in the next section, moving to the NHIC foliated by periodic orbits with the homology class $g = -(1, 0)$ and going all the way up to very high energy levels.

6.3.1 *The NHIC in the low energy region*

We cite the following result from [20], which gives the existence of the NHIC for low energy levels for Tonelli Hamiltonian systems of two degrees of freedom.

Theorem 6.6 (See [20, Theorem 2.1]). *Consider a C^r Tonelli Hamiltonian $H : T^*\mathbb{T}^2 \to \mathbb{R}$ normalized such that $\min \alpha_H = 0$ by adding a constant. Given a class $g \in H_1(\mathbb{T}^2, \mathbb{Z})$ and a closed interval $[E_-, E_+] \subset (0, \infty)$, there exists an open and dense set $\tilde{\mathcal{O}}_2(E_-, E_+) \subset C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V \in \tilde{\mathcal{O}}_2(E_-, E_+)$ normalized by adding a constant such that $\min \alpha_H = \min \alpha_{H+V} = 0$, it holds simultaneously for all $E \in [E_-, E_+]$ that the Mather set $\tilde{\mathcal{M}}(E, g)$ on the energy level E having the homology class g for $H + V$ consists of hyperbolic periodic orbits. Moreover, except for finitely many $E_j \in [E_-, E_+]$ where the Mather set consists of two hyperbolic periodic orbits, for all other $E \in [E_-, E_+]$, the Mather set is exactly one hyperbolic periodic orbit.*

We denote by $\nu_\pm(1, 0)$ the rotation vectors of the Mather set on the energy levels $E_\pm$ with the homology class $(1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$. The next lemma shows that each hyperbolic periodic orbit corresponds to a one-dimensional flat in $H^1(\mathbb{T}^2, \mathbb{R})$.

Lemma 6.7. *Let $H(\tilde{\mathsf{x}}, \tilde{\mathsf{y}}) : T^*\mathbb{T}^2 \to \mathbb{R}$ be a Tonelli Hamiltonian and $c^* \in H^1(\mathbb{T}^2, \mathbb{R})$. We assume that the Mather set $\tilde{\mathcal{M}}(c^*)$ is supported on a hyperbolic periodic orbit with the rotation vector νg for $g \in H_1(\mathbb{T}^2, \mathbb{Z})$ and $\nu \neq 0$. Then the set $\partial\beta_H(\nu g)$ is an interval $\{c^* + sc_g \mid s_- \leqslant s \leqslant s_+\} \subset H^1(\mathbb{T}^2, \mathbb{R})$ with $s_- < s_+$, $s_- \leqslant 0 \leqslant s_+$, $c_g \perp g$ and $\|c_g\| = 1$ such that for each $c \in \{c^* + sc_g \mid s_- < s < s+\}$, we have $\tilde{\mathcal{A}}(c) = \tilde{\mathcal{M}}(c^*)$.*

The proof is postponed to the end of this section.

6.3.2 *The high energy regime*

We first show the high energy level case is the same as the system $\tilde{\mathsf{G}}$ in (6.6) with a linear term in y_1. In (6.12), we consider $\tilde{\mathsf{y}}^\star$ and Λ_* such that $\tilde{A}\tilde{\mathsf{y}}^\star = \nu(1, 0)$ for some large ν with $\nu\|\tilde{A}^{-1}(1, 0)\| > \Lambda_*$ and $\frac{1}{2}\|\tilde{A}^{-1}\|^{-1}\Lambda_*^2 + \min V \geqslant E_*$, and introduce $\tilde{\mathsf{y}} - \tilde{\mathsf{y}}^\star = \tilde{\mathsf{Y}}$. In the coordinates $(\tilde{\mathsf{x}}, \tilde{\mathsf{Y}})$, the Hamiltonian becomes

$$\tilde{\mathsf{G}}(\tilde{\mathsf{x}}, \tilde{\mathsf{Y}}) = \frac{1}{2}\langle \tilde{A}\tilde{\mathsf{y}}^\star, \tilde{\mathsf{y}}^\star\rangle + \nu\tilde{\mathsf{Y}}_1 + \frac{1}{2}\langle \tilde{A}\tilde{\mathsf{Y}}, \tilde{\mathsf{Y}}\rangle + V(\tilde{\mathsf{x}}). \tag{6.13}$$

This is of the form of the Hamiltonian $\tilde{\mathsf{G}}$ in (6.6).

We next cite a result from [14] concerning the high energy regime of the system $\tilde{\mathsf{G}}$ in (6.12). We define $[V](\mathsf{x}_2) = \int_{\mathbb{T}^1} V(\mathsf{x}_1, \mathsf{x}_2)\, d\mathsf{x}_1$. Suppose that $[V]$ has a unique nondegenerate global maximum at a point denoted by x_2^*, which is a C^2 open and dense condition.

Theorem 6.8 (See [14, Theorem 3.1 and Proposition 3.1]). *Suppose that the potential V of the system $\tilde{\mathsf{G}}$ in (6.12) satisfies that $[V]$ has a unique nondegenerate global maximum at x_2^*. Then there exists an $E_* > 0$ such that*

(1) *the action minimizing periodic orbits in the homology class* $g = (1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$ *on the energy levels* $\{E > E_*\}$ *form a unique* C^r *NHIC homeomorphic to* $T^*\mathbb{T}$ *with uniform normal hyperbolicity*;

(2) *as* $E \to \infty$, *the periodic orbit* $\{(\tilde{\mathsf{x}}_E(t), \tilde{\mathsf{y}}_E(t))\}$ *on the energy level* E *has the following uniform convergence*: $\mathsf{x}_{2,E}(t) \to \mathsf{x}_2^*$ *and* $\dot{\mathsf{x}}_{2,E}(t) \to 0$.

By the reversibility of the system $\tilde{\mathsf{G}}$, the same conclusion holds for the homology class $g = (-1, 0)$. In fact, the periodic orbit in the Mather set $\tilde{\mathcal{M}}_{\nu(-1,0)}$ is the time reversal of $\tilde{\mathcal{M}}_{\nu(1,0)}$.

Here, we only sketch the proof and the details can be found in [14, Theorem 3.1 and Proposition 3.1].

Sketch of the proof. The main idea of the proof is that on the high energy level, the fast oscillation in the x_1 component (see the equation (6.13), which implies $\dot{\mathsf{x}}_1 = \nu + O(1) \gg 1$) will effectively average out the dependence on x_1 in V, so the Hamiltonian system is effectively $\frac{1}{2}\langle\tilde{\mathsf{y}}, \tilde{A}\tilde{\mathsf{y}}\rangle + [V](\mathsf{x}_2)$ as $E \to \infty$. So we get that the normal hyperbolicity is determined by $\tilde{A}$ and the second-order derivative $[V]''(\mathsf{x}_2^*)$ hence is independent of the energy levels. The genericity assumption on V is to guarantee that $[V](\mathsf{x}_2)$ has a nondegenerate global maximum. □

6.4 NHICs in the double resonance regime

In this subsection, we embed the NHICs of the system $\tilde{\mathsf{G}}$ into the full system to get NHICs of codimension 2.

6.4.1 *The NHIC in the* $\Lambda\sqrt{\varepsilon}$*-neighborhood of the double resonance*

In this subsection, we give the existence of the NHIC in $B_{\Lambda\sqrt{\varepsilon}}(y^\star) \times \mathbb{T}^n$ near the strong double resonance up to a neighborhood of the strong double resonance. This corresponds to Theorem 6.6.

Proposition 6.9. *Let* $y^\star \in \Sigma(\boldsymbol{k}', \boldsymbol{k}^o)$ *be such that* $\omega^\star = \boldsymbol{\omega}(y^\star)$ *is at the strong double resonance with integer vectors* $\boldsymbol{k}'$ *and* $\boldsymbol{k}^o$. *Then for any* $\lambda > 0$, *there is an open and dense set* $\mathcal{O}_2 = \mathcal{O}_2(\boldsymbol{k}', \boldsymbol{k}^o; \lambda, \Lambda) \subset \Pi_{\boldsymbol{k}',\boldsymbol{k}^o} C^r(\mathbb{T}^n)/\mathbb{R}$ *such that for each* P *with* $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P(x, y^\star) \in \mathcal{O}_2$ *normalized such that* $\max \Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P(x, y^\star) = 0$, *there exists a* $\delta_2 = \delta_2(\Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P(x, y^\star), \lambda) > 0$ *such that for all* $0 < \delta < \delta_2$ *and all* $0 < \varepsilon < \varepsilon_2(\delta, \Lambda)$ *as in Lemma* 6.1, *the following hold*:

(1) *The Hamiltonian system* (6.1) *admits an NHIC* $\mathcal{C}(\boldsymbol{k}')$ *homeomorphic to* $T^*\mathbb{T}^{n-1}$, *up to finitely many bifurcations, entering a* λ*-neighborhood of* $\Sigma(\boldsymbol{k}', \boldsymbol{k}^o) \times \mathbb{T}^n$.

(2) *The NHIC has uniform normal hyperbolicity, independent of* δ *or* ε.

(3) *Mather sets lying in* $B_{0.9\Lambda}(0) \times \mathbb{T}^n$ *and with rotation vectors perpendicular to* $\boldsymbol{k}'$ *and of distance* λ*-away from* $-\varepsilon^{-1/2}\omega^\star + (\boldsymbol{k}^o)^\perp$, *are contained in* $\mathcal{C}(\boldsymbol{k}')$.

Proof. In the system $\mathsf{H}_{S,\delta}$ (6.5), we first discard the δ-perturbation and consider the system (6.8) $\mathsf{H}_{S,0} = \tilde{\mathsf{G}}(\hat{\mathsf{x}}, \hat{\mathsf{y}}) + \hat{\mathsf{G}}(\hat{\mathsf{y}}_{n-2}) : T^*\mathbb{T}^n_S \to \mathbb{R}$.

First, the system $\tilde{\mathsf{G}}(\hat{\mathsf{x}}, \hat{\mathsf{y}})$ admits an NHIC by Theorem 6.6 with the homology class $(1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$ for V chosen in an open and dense subset $\mathcal{O}_2(E_-, E_+)$ of $C^r(\mathbb{T}^2)/\mathbb{R}$. Here, we choose $E_- = \alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda(1, 0)))$ and E_+ to be the highest possible energy level for $\|Y\| \leqslant \Lambda$. This gives the open and dense set $\mathcal{O}_2(\boldsymbol{k}^o, \boldsymbol{k}'; \lambda, \Lambda)$, since V is obtained from $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}'} P(y^\star, x)$ after a linear transformation. We next show that the system $\mathsf{H}_{S,0}$ admits an NHIC. Indeed, given a periodic orbit $\tilde{\gamma} = (\tilde{\mathsf{x}}_E(t), \tilde{\mathsf{y}}_E(t))$ of the system $\tilde{\mathsf{G}}$ in the Mather set $\tilde{\mathcal{M}}(E)$, it gives rise to an orbit of the system $\mathsf{H}_{S,0}$:

$$(\tilde{\mathsf{x}}_E(t), \hat{\mathsf{x}}(0) + (\varepsilon^{-1/2}\hat{\omega}^\star + B\hat{\mathsf{y}}(0))t, \tilde{\mathsf{y}}_E(t), \hat{\mathsf{y}}(0)) \subset T^*\mathbb{T}^n_S, \quad t \in \mathbb{R}.$$

Taking union over all the periodic orbits and all the initial conditions $\hat{\mathsf{x}}(0) \in (-\breve{A}^t\tilde{A}^{-t}\tilde{\mathsf{x}} + \mathbb{T}^{n-2})$ and $\|\hat{\mathsf{y}}(0)\| \leqslant \Lambda$, we get an NHIC for the system $\mathsf{H}_{S,0}$ that is homeomorphic to $T^*\mathbb{T}^{n-1}_{\bar{S}}$, where $\bar{S}$ is obtained from S by removing the second row and the second column. Going back to the system (6.1) with $\delta = 0$ by inverting the symplectic transformation $\mathfrak{S}\mathfrak{M}''$, we get an NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$.

Due to the uniform hyperbolicity, when the δ-perturbation in (6.5) is turned on, we get the persistence of the NHIC as we did in the proof of Proposition 5.2. Here, the modification of the $\delta\mathsf{R}$ should be done as follows in addition to that used in the proof of Proposition 5.2 in order to smoothen the Hamiltonian in the region $\{0 \leqslant \tilde{\mathsf{G}}(\hat{\mathsf{x}}, \hat{\mathsf{y}}) < E_-\}$. We introduce a C^∞ monotone cutoff function $\rho : [0, \infty) \to [0, 1]$ satisfying

$\rho(x)=0$ for $x\leqslant 1/3$ and $\rho(x)=1$ if $x>2/3$. We next modify $\delta\mathsf{R}$ to $\rho(\frac{\tilde{\mathsf{G}}(\hat{\mathsf{x}},\hat{\mathsf{y}})}{\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda(1,0)))})\chi(\|Y\|/\Lambda)\delta\mathsf{R}$. Now the proposition follows from the same argument as in Proposition 5.2. □

6.4.2 *The high energy regime*

In this subsection, we give the existence of the NHIC in $B_{\Lambda\sqrt{\varepsilon}}(y^\star)\times\mathbb{T}^n$ for $y^\star\in\Sigma(\boldsymbol{k}')$ and

$$\omega^\star=\boldsymbol{\omega}(y^\star)\in\mathcal{D}''\setminus B_{\Lambda\varepsilon^{1/2}}(\omega^o+(\boldsymbol{k}^o)^\perp) \tag{6.14}$$

that is $\Lambda\varepsilon^{1/2}$-away from but $\varepsilon^{1/3}$-close to the strong double resonance. This corresponds to Theorem 6.8.

Proposition 6.10. *Let $P\in\mathcal{O}_1$ and δ_1 be as in Proposition* 5.2. *Then there exists a Λ_* such that for all $\Lambda>\Lambda_*$ and $y^\star$ being such that $\omega^\star=\boldsymbol{\omega}(y^\star)$ is as in* (6.14)*, all $0<\delta<\delta_1$ and all $0<\varepsilon<\varepsilon_2(\delta,\Lambda)$, the Hamiltonian system* (6.1) *defined in $B_\Lambda(0)\times\mathbb{T}^n$ admits a C^{r-1} NHIC $\mathcal{C}(\boldsymbol{k}')$ homeomorphic to $T^*\mathbb{T}^{n-1}$ with the following properties*:

(1) *The normal hyperbolicity is uniform, independent of Λ,δ or ε.*

(2) *Mather sets lying in $\{\|Y\|\leqslant 0.9\Lambda\}\times\mathbb{T}^n$ with rotation vectors perpendicular to $\boldsymbol{k}'$ lie inside $\mathcal{C}(\boldsymbol{k}')$.*

Proof. By the previous Theorem 6.8, we get the existence of the NHIC in the system $\tilde{\mathsf{G}}$ in (6.6). By the same argument as the proof of Proposition 6.9, we get the existence of the NHIC in the original system (6.1). The assumption in Theorem 6.8 on the nondegeneracy of $[V]$ turns out to be the nondegeneracy of the global maximum of $\Pi_{\boldsymbol{k}'}P(y^\star,x)$ which is guaranteed by $\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_1\subset\Pi_{\boldsymbol{k}'}C^r$. The remaining statements are proved in the same way as in Proposition 5.2. □

6.5 Proofs

In this subsection, we give the proof of Lemma 6.7.

Proof of Lemma 6.7. The proof is a variant of [13, Proposition 2.1]. As the system is autonomous with two degrees of freedom, $\partial\beta_H(\nu g)$ is either an interval or a point since $\partial\beta_H(\nu g)$ lies on an energy level $\alpha^{-1}(E)$, which is a closed curve. In the case of the interval, some $c_g\in H^1(\mathbb{T}^2,\mathbb{R})$ exists such that $\partial\beta_H(\nu g) = \{c^*+sc_g\mid s_-\leqslant s\leqslant s_+\}$. It follows from [36] that for all the classes in the set $\{c^*+sc_g\mid s_-<s<s_+\}$, the Aubry sets $\tilde{\mathcal{A}}(c)$ are the same. Let us show that $s_-<s_+$ and $\tilde{\mathcal{A}}(c)=\tilde{\mathcal{M}}(c^*)$ for $c\in\{c^*+sc_g\mid s_- < s<s_+\}$.

Given any absolutely continuous curve γ, its Lagrange action is defined as follows:

$$A_c(\gamma)=\int(L_H(\dot\gamma,\gamma)-\eta_c+\alpha_H(c))dt,\quad [\eta_c]=c.$$

Denote by γ_0 the hyperbolic periodic orbit. We consider *minimal* homoclinic orbits to γ_0, which is located in the intersection of the stable and unstable manifolds of $(\dot\gamma_0,\gamma_0)$. A homoclinic orbit $(\dot\gamma,\gamma)$ is called *minimal* if the lift of γ, $\check\gamma$: $\mathbb{R}\to\check M$ is semi-static for the class c^*, where $\check M$ is the largest covering space of $\mathbb{T}^2$ so that $\pi_1(\check M)=\pi_1(U)$ holds for each open neighborhood of $\mathcal{M}(c^*)$. Because of the topology of $\mathbb{T}^2$, there are only two types of minimal homoclinic orbits, denoted by $(\dot\gamma^\pm,\gamma^\pm)$. Given a point $x\in\gamma_0$, there are four sequences of time $t^\pm_{i,\pm}$ such that $\gamma^-(t^\pm_{i,-})\to x$ as $t^\pm_{i,-}\to\pm\infty$ and $\gamma^+(t^\pm_{i,+})\to x$ as $t^\pm_{i,+}\to\pm\infty$ and $t^\pm_{i,-}\to\pm\infty$ as $i\to\infty$. We define

$$A_c(\gamma^-,x)=\liminf_{i\to\infty}\int_{t^-_{i,-}}^{t^+_{i,-}}(L_H(\dot\gamma^-,\gamma^-)-\langle c,\dot\gamma^-\rangle+\alpha_H(c))dt,$$
$$A_c(\gamma^+,x)=\liminf_{i\to\infty}\int_{t^-_{i,+}}^{t^+_{i,+}}(L_H(\dot\gamma^+,\gamma^+)-\langle c,\dot\gamma^+\rangle+\alpha_H(c))dt.$$

We obviously have $A_{c^*}(\gamma^\pm,x)\geqslant 0$. Next, we claim that $A_{c^*}(\gamma^+,x)+A_{c^*}(\gamma^-,x)>0$. Otherwise, we would have $A_{c^*}(\gamma^\pm)=0$ for both $\pm$, which implies that $\gamma^\pm\subset\tilde{\mathcal{A}}(c^*)$. However, this violates the graph property of the Aubry set since in the first relative homology group $H_1(\mathbb{T}^2,\gamma_0,\mathbb{Z})$ we have $[\gamma^+]\neq[\gamma^-]$, and when lifted to $\mathbb{R}^2$, the two curves $\gamma^{\pm}$ lying in the same strip bounded by two neighboring lifts of γ_0 and hence the projections of $\gamma^{\pm}$ on $\mathbb{T}^2$ must intersect. The contradiction proves our claim. Let us assume $A_{c^*}(\gamma^+) > 0$ without loss of generality.

Pick Δc small enough and satisfying

$$\langle \Delta c, [\gamma_0] \rangle = 0, \quad \langle \Delta c, [\gamma^+] \rangle > 0 \quad \text{and} \quad A_{c^*}(\gamma^+) - \langle \Delta c, [\gamma^+] \rangle > 0.$$

According to the upper semi-continuity of the Mañé set in the cohomology class, any minimal measure μ_c is supported by a set lying in a small neighborhood of these homoclinic orbits if $c = c^* + \Delta c$ and $|\Delta c|$ is very small. By the assumption $A_{c^*}(\gamma^+) > 0$, it can only happen that μ_c is supported in a neighborhood of $\gamma^- \cup \gamma_0$.

We claim that the minimal measure μ_c for $c = c^* + \Delta c$ is still supported on the periodic orbit γ_0. First, we show that $\rho(\mu_c) \parallel \rho(\mu_{c^*}) \perp \Delta c$. Otherwise, since $\text{supp}(\mu_c)$ lies in the small neighborhood of γ^-, it follows that $-\langle \Delta c, \rho(\mu_c) \rangle > 0$. On the other hand, as the c^*-minimal measure is uniquely supported on the periodic orbits, the β-function is strictly convex at $\rho(\mu_{c^*})$ and hence the α-function is differentiable at c^* and $\rho(\mu_{c^*}) = \nu[\gamma_0]$ holds for a certain number $\nu \neq 0$. Therefore, we have $\alpha_H(c^* + \Delta c) - \alpha_H(c^*) = o(|\Delta c|)$. Consequently, we obtain from the definition that

$$\begin{aligned} A_c(\mu_c) &= \int (L_H - \eta_{c^*}) d\mu_c + \alpha_H(c^* + \Delta c) - \langle \Delta c, \rho(\mu_c) \rangle \\ &= \int (L_H - \eta_{c^*}) d\mu_c + \alpha_H(c^*) - \langle \Delta c, \rho(\mu_c) \rangle + o(|\Delta c|), \end{aligned}$$

from which we have $A_c(\mu_c) > 0$ as $A_{c^*}(\mu_{c^*}) \geqslant 0$, $-\langle \Delta c, \rho(\mu_c) \rangle > 0$ and $o(|\Delta c|)$ is a higher-order term of $|\Delta c|$. The contradiction implies that $\rho(\mu_c) \perp \Delta c$. Next, by the convexity of α, we have

$$\alpha(c) - \alpha(c^*) \geqslant \langle \Delta c, \rho(\mu_{c^*}) \rangle = 0 \quad \text{and} \quad \alpha(c^*) - \alpha(c) \geqslant \langle -\Delta c, \rho(\mu_c) \rangle = 0,$$

so we have $\alpha(c) = \alpha(c^*)$. We get that the interval $c^* + s\Delta c$, $s \in [0, 1]$ lies entirely on the energy level $\alpha(c^*)$, on which the Mather set in the homology class $g \in H_1(\mathbb{T}^2, \mathbb{Z})$ is known to be the unique hyperbolic periodic orbit and hence the rotation vector is constant for c in the interval. Finally, from the proof, we see that the curves $\gamma^{\pm}$ appear in the Aubry set only when the cohomology class lies on the endpoints of the interval. Otherwise, the Aubry set agrees with the Mather set being the periodic orbit. This completes the proof. □

7 Dynamics around strong double resonances: Cohomology equivalence

In the last section, we have written the system at the strong double resonance into a form of a mechanical system of two degrees of freedom coupled with a fast rotating integrable system. In this section, we first recall the main result of [13] on the existence of cohomological equivalence for Hamiltonian systems of two degrees of freedom near the zero energy level. Next, we generalize it to the full system to build a piece of transition chain.

7.1 Cohomological equivalence for the subsystem of two degrees of freedom

Theorem 7.1 (See [13, Proposition 2.1] and Figure 3). *Let $H : T^*\mathbb{T}^n \to \mathbb{R}$ be a Tonelli Hamiltonian. Given a class $c_0 \in H^1(\mathbb{T}^n, \mathbb{R})$, if the minimal measure is supported on a hyperbolic fixed point, then there exists an n-dimensional convex flat $\mathbb{F}_0 \subset H^1(\mathbb{T}^n, \mathbb{R})$ containing c_0 such that this fixed point supports a c-minimal measure for all $c \in \mathbb{F}_0$.*

In the following, we specialize in the case of $n = 2$. The next theorem is one of the main results in [13].

Theorem 7.2 (See [13, Theorem 3.1]). *There is an open and dense set $\tilde{\mathcal{O}}_3 \subset C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V \in \tilde{\mathcal{O}}_3$ normalized by $\max V = 0$, for each $c \in \partial\tilde{\mathbb{F}}_0$, where $\tilde{\mathbb{F}}_0 = \alpha^{-1}(\min \alpha)$ is the flat of the α-function for $\tilde{\mathsf{G}}$, the Mañé set $\mathcal{N}(c)$ does not cover the whole configuration space $\mathbb{T}^2$, i.e., $\mathcal{N}(c) \subsetneq \mathbb{T}^2$.*

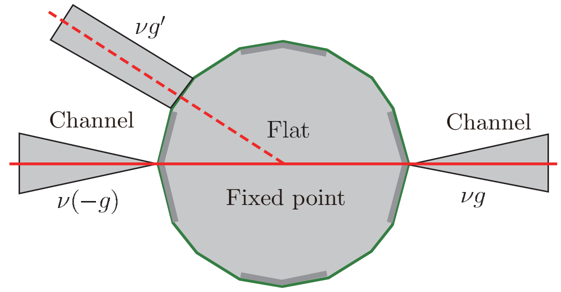

Figure 3 (Color online) Two ways that the flat $\mathbb{F}_0$ connects to the channels

Remark 7.3. [13, Theorem 3.1] gives only a residual set. The openness of the set follows immediately from the upper-semi-continuity of the Mañé set.

This theorem allows us to construct the orbit connecting two Aubry sets $\tilde{\mathcal{A}}(c)$ and $\tilde{\mathcal{A}}(c')$ for any c and c' in $\partial\tilde{\mathbb{F}}_0$ using the mechanism of cohomological equivalence (Item (H2) of Definition 2.1).

Proposition 7.4 (See [13, Theorems 1.1 and 3.2])**.** *Given $V \in \tilde{\mathfrak{O}}_3$ normalized by $\max V = 0$, there exists some positive numbers $\tilde{\Delta}_0 = \tilde{\Delta}_0(V) > 0$ such that for each $E \in (0, \tilde{\Delta}_0)$ and each $c \in \alpha^{-1}(E)$, there exists a circle $\Sigma_c \subset \mathbb{T}^2$ such that all the c-semi-static curves of the system $\tilde{\mathsf{G}}$ pass through that circle transversely and $\mathcal{N}(c) \cap \Sigma_c \subset \bigcup I_{c,i}$, where $I_{c,i} \subset \Sigma_c$ are finitely many disjoint open intervals. Therefore any two cohomology classes c and c' in $\alpha^{-1}(E)$ are c-equivalent.*

7.2 Cohomological equivalence for the full system

We next construct a generalized transition chain using the c-equivalent mechanism in the full system $\mathsf{H}_{S,\delta} : T^*\mathbb{T}^n_S \to \mathbb{R}$ near the strong double resonance. We assume $y^\star \in \Sigma(\boldsymbol{k}', \boldsymbol{k}^o)$. Such a generalized transition chain will give rise to one for the original system H after the linear symplectic transformations. We first study the α-function for $\mathsf{H}_{S,\delta}$. Note that the Mather theory is defined for Tonelli systems on T^*M for a general closed manifold M. Here, we have $H^1(\mathbb{T}^n_S, \mathbb{R})$ isomorphic to $H^1(\mathbb{T}^n, \mathbb{R})$ with the basis vectors transformed by S^{-t} in the same way as the y variables in (6.10).

Lemma 7.5. *The α-functions of $\mathsf{H}_{S,\delta}$ satisfy $\|\alpha_{\mathsf{H}_{S,\delta}} - \alpha_{\mathsf{H}_{S,0}}\|_{C^0} < \delta$, where*

$$\alpha_{\mathsf{H}_{S,0}}(c) = \alpha_{\tilde{\mathsf{G}}}(\tilde{c}) + \frac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}_{S,n-2}, \hat{c}\rangle + \frac{1}{2}\langle B\hat{c}, \hat{c}\rangle, \quad c \in H^1(\mathbb{T}^n_S, \mathbb{R}) = \mathbb{R}^n.$$

Proof. For the δ-estimate of the difference, we denote by L_δ and L_0 the Lagrangians corresponding to $\mathsf{H}_{S,\delta}$ and $\mathsf{H}_{S,0}$, respectively. Then we have $\|\mathsf{L}_\delta - \mathsf{L}_0\|_{C^0} \leqslant \delta$. Given the cohomology class c, we denote by μ_δ and μ_0 the c-minimal measures for L_δ and L_0, respectively. Choose a closed one-form η_c with $[\eta_c] = c$, and then we get

$$-\alpha_{\mathsf{H}_{S,\delta}} = \int (\mathsf{L}_\delta - \eta_c)d\mu_\delta \leqslant \int (\mathsf{L}_\delta - \eta_c)d\mu_0, \quad -\alpha_{\mathsf{H}_{S,0}} = \int (\mathsf{L}_0 - \eta_c)d\mu_0 \leqslant \int (\mathsf{L}_0 - \eta_c)d\mu_\delta.$$

The δ-estimate of the difference is obtained by taking difference.

To determine the form of the α-function for $\mathsf{H}_{S,0}$, let us consider an invariant measure μ in the Mather set with the cohomology class $c = (\tilde{c}, \hat{c})$ of the system $\mathsf{H}_{S,0}$. Denote by $\tilde{\mu}$ the corresponding invariant measure in the cohomology class of $\tilde{c}$ of the subsystem $\tilde{\mathsf{G}}$. By Mather's graph theorem, we know that suppμ is a graph from a subset of $\mathbb{T}^n_S$ to $\mathbb{R}^n$ and supp$\tilde{\mu}$ is a graph from a subset of $\mathbb{T}^2$ to $\mathbb{R}^2$. Next, we know that μ has a skew product structure: for each $\tilde{\mathsf{x}} \in \mathbb{T}^2$, there is a measure $\hat{\mu}_{\tilde{\mathsf{x}}}$ supported on the torus Graph$\tilde{\mu}(\tilde{\mathsf{x}}) \times (-\breve{A}^t\tilde{A}^{-t}\tilde{\mathsf{x}} + \mathbb{T}^{n-2}) \times \{\hat{c}\}$ using the transformation $\mathfrak{S}$ in (6.10) as well as the fact that $\dot{\hat{\mathsf{y}}} = 0$. So the integration with respect to $d\mu$ disintegrates into $d\mu(\mathsf{x}) = d\hat{\mu}_{\tilde{\mathsf{x}}}(\hat{\mathsf{x}})d\tilde{\mu}(\tilde{\mathsf{x}})$. When doing the inner integral with the integrand being the Lagrangian of $\mathsf{H}_{S,0}$, note that the Lagrangian does not depend on $\hat{\mathsf{x}}$, so the integration with respect to $d\tilde{\mu}_{\hat{\mathsf{x}}}(\tilde{\mathsf{x}})$ is effectively the integration with respect to a Haar measure supported on the above torus containing the support of $\hat{\mu}_{\tilde{\mathsf{x}}}$. In particular, in the $\hat{\mathsf{y}}$ component, the measure is Dirac-δ supported on $\{\hat{y} = \hat{c}\}$. This gives the term $\frac{1}{\sqrt{\varepsilon}}\langle\hat{\omega}_{S,n-2}, \hat{c}\rangle + \frac{1}{2}\langle B\hat{c}, \hat{c}\rangle$. Finally, the outer integral with respect to $d\tilde{\mu}$ gives the term $\alpha_{\tilde{\mathsf{G}}}(\tilde{c})$. □

The next lemma shows that in the system $\tilde{\mathsf{G}}$, the NHIC overlaps the region of c-equivalence.

Lemma 7.6. *There exists an open and dense subset $\hat{\mathcal{O}}_3 \subset C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V \in \hat{\mathcal{O}}_3$ with $\max V = 0$, there exists a $\lambda > 0$ such that the following hold for the system $\tilde{\mathsf{G}}$:*

(1) *the system $\tilde{\mathsf{G}}$ admits an NHIC on the energy interval $[\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda, 0)), \infty)$, foliated by hyperbolic periodic orbits in the Mather sets with rotation vectors $\nu(1, 0)$ $(|\nu| > \lambda)$, up to finitely many bifurcations*;

(2) *each curve $\alpha_{\tilde{\mathsf{G}}}^{-1}(E)$ $(E/\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda, 0)) \in [1, 2))$ is a curve of c-equivalence.*

Proof. By Proposition 7.4, there exists an open and dense subset $\tilde{\mathcal{O}}_3$ in $C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V \in \tilde{\mathcal{O}}_3$ with $\max V = 0$, there exists a $\tilde{\Delta}_0(V) > 0$ such that each curve $\alpha_{\tilde{\mathsf{G}}}^{-1}(E)$ $(E \in [0, \tilde{\Delta}_0))$ is a curve of c-equivalence. We introduce a sequence of open sets $\tilde{\mathcal{O}}_{3,\ell}$ $(\ell \in \mathbb{N})$ satisfying $\tilde{\mathcal{O}}_{3,\ell} \subset \tilde{\mathcal{O}}_{3,\ell+1}$ and $\tilde{\mathcal{O}}_3 = \bigcup_{\ell\in\mathbb{N}} \tilde{\mathcal{O}}_{3,\ell}$, where $\tilde{\mathcal{O}}_{3,\ell} := \{V \in \tilde{\mathcal{O}}_3 \mid \tilde{\Delta}_0(V) > 2/\ell\}$. Each set $\tilde{\mathcal{O}}_{3,\ell}$ is open due to the upper-semi-continuity of the Mañé set. Indeed, suppose $V_* \in \tilde{\mathcal{O}}_{3,\ell}$ with $\tilde{\Delta}_0(V_*) > 2/\ell$. So for all c with $\alpha_{\tilde{\mathsf{G}}}(c) < 2/\ell$, the Mañé sets $\mathcal{N}(c)$ are broken in the sense of the conclusion of Proposition 7.4. By the upper-semi-continuity of the Mañé set with respect to the Lagrangian, the same is true for any potential V that is C^r sufficiently close to V_*, so $\tilde{\Delta}_0(V) \geqslant \tilde{\Delta}_0(V_*) > 2/\ell$. This means that there is a C^r-ball, centered at V_* contained in $\tilde{\mathcal{O}}_{3,\ell}$.

Next, we fix large $E_+ = E_*$ (see Theorem 6.8) and choose $E_- = 1/\ell$, and we introduce an open and dense set $\tilde{\mathcal{O}}_{2,\ell} := \tilde{\mathcal{O}}_2(1/\ell, E_+) \subset C^r(\mathbb{T}^2)/\mathbb{R}$ as in Theorem 6.6. Now the intersection $\tilde{\mathcal{O}}_{3,\ell} \cap \tilde{\mathcal{O}}_{2,\ell}$ is open in $C^r(\mathbb{T}^2)/\mathbb{R}$ and the union $\hat{\mathcal{O}}_3 := \bigcup_\ell(\tilde{\mathcal{O}}_{3,\ell} \cap \tilde{\mathcal{O}}_{2,\ell})$ is open and dense in $C^r(\mathbb{T}^2)/\mathbb{R}$. To get the statement, it is enough to set $1/\ell = \alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda, 0))$ if $V \in (\tilde{\mathcal{O}}_{3,\ell} \cap \tilde{\mathcal{O}}_{2,\ell})$. □

Going back to the original system, we have the following proposition.

Proposition 7.7. *Let $y^\star \in \Sigma(\boldsymbol{k}', \boldsymbol{k}^o)$ be such that $\omega^\star = \boldsymbol{\omega}(y^\star)$ is at the strong double resonance with integer vectors $\boldsymbol{k}'$ and $\boldsymbol{k}^o$. Then there exists an open and dense set $\mathcal{O}_3 = \mathcal{O}_3(\boldsymbol{k}', \boldsymbol{k}^o) \subset \Pi_{\boldsymbol{k}',\boldsymbol{k}^o}C^r(\mathbb{T}^n)/\mathbb{R}$ such that for any P with $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y^\star, x) \in \mathcal{O}_3$ normalized by $\max \Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y^\star, x) = 0$, there exist $\lambda = \lambda(\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y^\star, x))$ and $\delta_3 = \delta_3(\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P, \lambda)$ such that for all $0 < \delta < \delta_3$, the following holds. Suppose $c_* = (\tilde{c}_*, \hat{c}_*) \in \mathbb{R}^2 \times \mathbb{R}^{n-2} = H^1(\mathbb{T}^n, \mathbb{R})$ and $\mathsf{c}_* := (\tilde{\mathsf{c}}_*, \hat{\mathsf{c}}_*) = S^{-t}c_*$ satisfy $\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda, 0)) < \alpha_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}_*) < 2\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda, 0))$ and $\|\hat{\mathsf{c}}_*\| \leqslant \Lambda$. Then*

(1) *the path $\Gamma_\delta(\mathsf{c}_*) := \{(\tilde{\mathsf{c}}, \hat{\mathsf{c}}_*) \mid \alpha_{\mathsf{H}_{S,\delta}}(\tilde{\mathsf{c}}, \hat{\mathsf{c}}_*) = \alpha_{\mathsf{H}_{S,\delta}}(\mathsf{c}_*)\}$ is a path of c-equivalence for the system $\mathsf{H}_{S,\delta}$ in* (6.5);

(2) *the path $\Gamma_\delta(\mathsf{c}_*)$ lies in a δ-neighborhood of the curve $\Gamma_0(\mathsf{c}_*) := (\alpha_{\tilde{\mathsf{G}}}^{-1}(\alpha_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}_*)), \hat{\mathsf{c}}_*)$;*

(3) *the path $(SM'')^t\Gamma_\delta(\mathsf{c}_*)$ is a path of c-equivalence for the original system* (3.2).

7.3 Center straightening

Let the Tonelli Hamiltonian $H : T^*\mathbb{T}^2 \to \mathbb{R}$, the homology class $g \in H_1(\mathbb{T}^2, \mathbb{Z})$, the energy interval $[E_-, E_+]$ and the potential $V \in \tilde{\mathcal{O}}_2(E_-, E_+)$ be as in Theorem 6.6. Then we get at most finitely many pieces of NHICs foliated by hyperbolic periodic orbits.

Proposition 7.8. *Let the Tonelli Hamiltonian $H : T^*\mathbb{T}^2 \to \mathbb{R}$, the homology class $g = (1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$, the energy interval $[E_-, E_+]$ and the potential $V \in \tilde{\mathcal{O}}_2(E_-, E_+)$ be as in Theorem* 6.6. *Suppose that on this energy interval, H admits an NHIC N foliated by hyperbolic periodic orbits in the Mather set with rotation vectors νg $(\nu \in [\nu_-, \nu_+] \subset (0, \infty))$. Then the following hold*:

(1) *Restricted on the cylinder N, there exist two numbers $0 < I_- < I_+$ and a symplectic change of variables $\Phi : (I, \varphi) \in [I_-, I_+] \times \mathbb{T} \to (x, y)|_N$ such that the Hamiltonian H can be written as $\Phi^*H = H \circ \Phi = \tilde{h}(I)$, where $\tilde{h}$ is as smooth as H and satisfies*

$$\tilde{h}(I_\pm) = E_\pm, \quad \tilde{h}'(I_\pm) = \nu_\pm \quad \textit{and} \quad \tilde{h}'(I) > 0, \quad \tilde{h}''(I) > 0, \quad \forall\, I \in [I_-, I_+].$$

(2) *There is a neighborhood $\mathcal{U}$ of the c_1 line in $H^1(\mathbb{T}^2, \mathbb{R})$ such that for each $c = (c_1, c_2) \in \mathcal{U}$ with $c_1 \in [I_-, I_+]$, we have $\alpha_H(c) = \tilde{h}(c_1)$.*

(3) *Assume furthermore that H is reversible, i.e., $H(x,y) = H(x,-y)$. Then the Mather set with rotation vectors $-\nu g$ ($\nu \in [\nu_-, \nu_+]$) is the time reversal of that of νg. On the NHIC foliated by Mather sets with rotation vectors $-\nu g$ ($\nu \in [\nu_-, \nu_+]$), the restricted Hamiltonian system $\bar{h} : [-I_+, -I_-] \times \mathbb{T} \to \mathbb{R}$ of one degree of freedom satisfies $\bar{h}(I) = \tilde{h}(-I)$.*

7.4 Proofs

In this subsection, we give the proofs of Propositions 7.7 and 7.8.

Proof of Proposition 7.7. The open and dense set $\mathcal{O}_3$ is obtained by transforming the open and dense set $\hat{\mathcal{O}}_3$ in Lemma 7.6 by the linear transformation M''. Let us now go back to the system $\tilde{\mathsf{G}}$ for which we choose $V \in \hat{\mathcal{O}}_3$ which determines λ.

We define $\tilde{\Gamma}(\tilde{\mathsf{c}}_*) = \alpha_{\tilde{\mathsf{G}}}^{-1}(\alpha_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}_*))$ for given $\tilde{\mathsf{c}}_*$ satisfying $\alpha_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}_*)/\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda,0)) \in (1,2)$. The coordinate change S does not change the $\tilde{\mathsf{x}}$ components, so for each $\hat{\mathsf{c}}$, let $\Gamma_0(\tilde{\mathsf{c}}_*, \hat{\mathsf{c}}) = (\tilde{\Gamma}(\tilde{\mathsf{c}}_*), \hat{\mathsf{c}})$, and the Mañé set $\tilde{\mathcal{N}}_{\mathsf{G}}(\mathsf{c}(t))$ for the system G in (6.8) and the cohomology class $\mathsf{c}(t) \in \Gamma_0(\tilde{\mathsf{c}}_*, \hat{\mathsf{c}})$, when projected to the first $\mathbb{T}^2$ factor, coincides with the Mañé set $\mathcal{N}_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}(t))$ for the system $\tilde{\mathsf{G}}$. So by Proposition 7.4, for each $\mathsf{c}(t) \in \Gamma_0(\tilde{\mathsf{c}}_*, \hat{\mathsf{c}})$, there exist a circle $\Sigma_{\mathsf{c}(t)} \subset \mathbb{T}^2$ and disjoint open intervals $I_{\mathsf{c}(t),i}$ so that all the c-semi-static curves of the system G pass through that circle transversely and

$$\mathcal{N}_{\mathsf{G}}(\mathsf{c}(t)) \cap S(\Sigma_{\mathsf{c}(t)} \times \mathbb{T}^{n-2}) \subset \bigcup S(I_{\mathsf{c}(t),i} \times \mathbb{T}^{n-2}),$$

whose homology is in the set $\{(0,0)\} \times \mathbb{R}^{n-2}$. For each $\hat{\mathsf{c}}$ with $\|\hat{\mathsf{c}}\| < \Lambda$, the curve $(\tilde{\Gamma}(\tilde{\mathsf{c}}_*), \hat{\mathsf{c}})$ is a curve of cohomological equivalence for the system $\mathsf{H}_{S,0}$ since for two points $\mathsf{c}(t)$ and $\mathsf{c}(t')$ on the curve, the difference $\mathsf{c}(t) - \mathsf{c}(t') = (*, \hat{0})$ is perpendicular to the subspace $\{(0,0)\} \times \mathbb{R}^{n-2}$ which contains the homology of $\mathcal{N}_{\mathsf{G}}(\mathsf{c}(t)) \cap S(\Sigma_{\mathsf{c}(t)} \times \mathbb{T}^{n-2})$.

Next, we show that the level set $\{(\tilde{\mathsf{c}}, \hat{\mathsf{c}}_*) \mid \alpha_{\mathsf{H}_{S,\delta}}(\tilde{\mathsf{c}}, \hat{\mathsf{c}}_*) = \alpha_{\mathsf{H}_{S,\delta}}(\mathsf{c}_*)\}$ is $O(\delta)$-close to that of the case $\delta = 0$ which is $\Gamma_0(\tilde{\mathsf{c}}_*, \hat{\mathsf{c}})$. This follows from the following fact about convex functions: given two convex functions α_δ and α_0 with $|\alpha_\delta - \alpha_0|_{C^0} \leqslant \delta$ and $\|D\alpha_0\| \geqslant C > 0$ on the level set $\{\alpha_0(c) = E\}$, then the level sets $\{\alpha_\delta(c) = E\}$ and $\{\alpha_0(c) = E\}$ are $O(\delta)$-close to each other. To prove this fact, it is enough to measure the distance of the intersection points of the two level sets with each radial line. Since the subdifferential $D\alpha$ is bounded away from zero, to maintain constant E, the distance can at most be $O(\delta)$.

By the upper semi-continuity of the Mañé set, since the Hamiltonians and the cohomology paths are $O(\delta)$-close, when we consider the system $\mathsf{H}_{S,\delta}$ with δ small enough, the same conclusion holds. □

Proof of Proposition 7.8. The normal hyperbolicity implies that the symplectic form Ω restricted to the cylinder is still a symplectic form denoted by Ω_N (see [23, Equation (63)]) . Denote by $\mathbb{T} \times [I_-, I_+]$ the standard cylinder where $I_\pm$ are to be determined later, and let Ψ_0: $\mathbb{T} \times [I_-, I_+] \to N$ be a diffeomorphism. Then the pullback $\Psi_0^*\Omega_N$ of Ω_N is a symplectic form on the standard cylinder $\mathbb{T} \times [I_-, I_+]$. As the second de Rham cohomology group of cylinder $\mathbb{T} \times [I_-, I_+]$ is trivial, Moser's argument on the isotopy of symplectic forms shows that a certain diffeomorphism Ψ: $\mathbb{T} \times [I_-, I_+] \to \mathbb{T} \times [I_-, I_+]$ exists such that $\Psi^*\Psi_0^*\Omega_N = dI \wedge d\varphi$. The Hamiltonian H induces a Hamiltonian defined on $\mathbb{T} \times [I_-, I_+]$: $H\Psi_0\Psi(I, \varphi)$.

Restricted to N, the Hamiltonian system has one degree of freedom and hence is integrable. We have a standard method of introducing action-angle coordinates (see [2, Sections 50B and 50C]). Namely, the action variable I is defined as integrating the Poincaré-Cartan one form ydx along the periodic orbits, and an angular variable φ is introduced as symplectic conjugate of I. In the action-angle coordinates, the Hamiltonian depends only on I, so we denote it by $\tilde{h}(I)$. We define $I_\pm$ by $\tilde{h}(I_\pm) = E_\pm$ and $\tilde{h}'(I_\pm) = \nu_\pm$.

It remains to show the twist. We use a result of Carneiro [12] saying that Mather's β-function is differentiable in the radial direction for autonomous systems. Now $\tilde{h}(I)$ is actually Mather's α-function since the Mather set is exactly the periodic orbit γ_ν. The direction of νg is the radial direction as ν varies. The α-function is strictly convex $\frac{d^2\tilde{h}(I)}{dI^2} > 0$, a.e. in order that β is differentiable. It holds that

$$\frac{d\tilde{h}(I)}{dI} = \frac{d\tilde{h}(I_-)}{dI} + \int_{I_-}^{I} \frac{d^2\tilde{h}(t)}{dI^2} dt = \int_{I_-}^{I} \frac{d^2\tilde{h}(t)}{dI^2} dt > 0.$$

Since the symplectic transformation is explicit, we get that $\tilde{h}$ is as smooth as H.

By Lemma 6.7, we get that for each rotation vector $\nu(1,0)$ ($\nu \in [\nu_-, \nu_+]$), its Legendre transformation is a line segment perpendicular to the homology class $(1,0)$. Taking union over all the line segments, we get a two-dimensional strip in $H^1(\mathbb{T}^2, \mathbb{R})$ as the $\mathcal{U}$ in the statement. It remains to locate $\mathcal{U}$. Note that integrating a closed one-form η with the cohomology class c along a loop of the homology class $(1,0)$ will pick out the first entry of c. For the Hamiltonian system of one degree of freedom defined on $T^*\mathbb{T}$, the cohomology class of each periodic orbit γ is given by $\oint_\gamma y\,dx$. In our case, the restricted Hamiltonian system on the NHIC foliated by periodic orbits has one degree of freedom, so we get the cohomology class by integrating the Poincaré-Cartan form $y_1dx_1 + y_2dx_2$ along the periodic orbit. Restricted to the NHIC, the Hamiltonian system is integrable whose α-function is known to be the same as the Hamiltonian.

Finally, to see that the system $\tilde{h}(I)$ is reversible, we notice that the reversibility of the system $H(x,y)$ implies that the Mather sets with rotation vectors νg and $-\nu g$ ($\nu > 0$) are supported on the same periodic orbit with reversed time. Since the Legendre transformation of an even function is also even, we get the Lagrangian $L(\dot{x}, x)$ is even with respect to $\dot{x}$, and hence $p = \frac{\partial L}{\partial \dot{x}}$ gets a negative sign when we reverse the time. The two periodic orbits lie on the same energy level and their corresponding action variables are opposite to each other from the formula $I = \frac{1}{2\pi}\oint_\gamma y\,dx$. The proof is now completed. □

8 Frequency refinement

In this section, we describe the second step of the reduction of orders. We construct NHICs homeomorphic to $T^*\mathbb{T}^{n-2}$ and build generalized transition chains connecting the NHICs crossing the triple resonance. This section gives the major part of the proof of Theorem 2.4 in the case of $n = 4$.

We fix $\boldsymbol{k}'$ and choose a $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1$ and $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P \in \mathcal{O}_3$ for finitely many strong second resonances $\boldsymbol{k}^o$. This determines δ_1, δ_2 and δ_3 by Propositions 5.2, 6.9 and 7.7, respectively. We also fix a δ ($< \min\{\delta_1, \delta_2, \delta_3\}$) so that Propositions 5.2, 6.9 and 7.7 are applicable.

Definition 8.1 (The direct sum decomposition of the function space)**.** Recall Notation 3 for C^r. The space $C^r/(\Pi_{\boldsymbol{k}'}C^r)$ is defined in such a way that each $P \in C^r$ admits a decomposition $P = \Pi_{\boldsymbol{k}'}P + (P - \Pi_{\boldsymbol{k}'}P)$ respecting the L^2 orthogonal decomposition $C^r = \Pi_{\boldsymbol{k}'}C^r \oplus C^r/(\Pi_{\boldsymbol{k}'}C^r)$. The space $C^r/(\Pi_{\boldsymbol{k}'}C^r)$ inherits the norm of C^r. It is similar for $C^r/(\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}C^r)$.

8.1 Frequency refinement

We have been working in a μ-neighborhood of the frequency segment $\omega_a = \rho_a(a, \frac{P}{Q}\omega_2^*, \frac{p}{q}\omega_2^*, \hat{\omega}_{n-3}^*)$, $a \in [\omega_1^{*i} - \varrho, \omega_1^{*f} + \varrho]$. Note that μ is determined by δ through K.

We pick a rational number denoted by $\frac{\bar{p}}{\bar{q}}$ satisfying

$$\left|\frac{\bar{p}}{\bar{q}}\omega_2^* - \omega_4^*\right| < \mu, \quad \text{g.c.d.}(\bar{p}qQ, \bar{q}) = 1, \quad \text{g.c.d.}(\bar{q}p, \bar{p}q) = 1, \tag{8.1}$$

and obtain a new segment of frequency $\bar{\omega}_a := \bar{\rho}_a(a, \frac{P}{Q}\omega_2^*, \frac{p}{q}\omega_2^*, \frac{\bar{p}}{\bar{q}}\omega_2^*, \hat{\omega}_{n-4}^*)$.

Besides $\boldsymbol{k}'$, the frequency $\bar{\omega}_a$ now admits a new resonant integer vector denoted by $\boldsymbol{k}''$ for all a. For μ sufficiently small, the rational number $\bar{p}/\bar{q}$ necessarily has a large denominator bounded from below by $O(\mu^{-1})$. So we get that $|\boldsymbol{k}''|$ is bounded from below by $O(\mu^{-1})$. Thus $|\boldsymbol{k}''| \gg |\boldsymbol{k}'|$ if μ is small enough.

The transformed frequency segment is $M'\bar{\omega}_a = \bar{\rho}_a(a, 0, \frac{\bar{P}}{\bar{Q}}\omega_2^*, \frac{\bar{p}}{\bar{q}}\omega_2^*, \hat{\omega}_{n-4}^*)$, where $\frac{\bar{P}}{\bar{Q}} = \frac{1}{qQ}$ with $\bar{P} = 1$ and $\bar{Q} = qQ$.

For the system restricted to the NHIC in Proposition 5.2, we remove the zero entry in $M'\bar{\omega}$. Now we are in a situation completely parallel to Section 4. Again we encounter the situation of single and double resonances. The new resonant integer vector can be determined from the equation $\boldsymbol{k}''(M')^{-1} = (0, 0, \bar{Q}\bar{p}, -\bar{q}\bar{P}, \hat{0}_{n-4})$, where g.c.d.$(\bar{q}\bar{P}, \bar{p}\bar{Q}) =$ g.c.d.$(\bar{p}qQ, \bar{q}) = 1$.

As we vary a in an interval, a third resonance may appear. We fix $K = (\delta/3)^{-1/2}$ as in Lemma 4.5 by fixing δ. Parallel to Lemma 4.5, we have the following lemma.

Lemma 8.2. *Let ω_-, μ, $\bar{\omega}_-$, K, $\boldsymbol{k}'$ and $\boldsymbol{k}''$ be as above. For any $\bar{K} > \max\{K, |\boldsymbol{k}''|\}$, let $\boldsymbol{k}_i^o$ $(i = 1,\dots,m)$ be the collection of all the irreducible integer vectors in $\mathbb{Z}^n_{\bar{K}} \setminus \mathrm{span}\{\boldsymbol{k}', \boldsymbol{k}''\}$ satisfying $\langle \boldsymbol{k}_i^o, \bar{\omega}^o\rangle = 0$, and $(\boldsymbol{k}_i^o)^\perp$ be the $(n-1)$-dimensional space orthogonal to the vector $\boldsymbol{k}_i^o$, where $a^o \in [\omega_1^{*i} - \varrho, \omega_1^{*f} + \varrho]$, $i = 1,\dots,m$. Then there exists a $\bar{\mu} = \bar{\mu}(\bar{K})$ such that $B_{\bar{\mu}}(\bar{\omega}_-) \subset B_\mu(\omega_-)$ and*

(1) for any small ε and all ω in the neighborhood $\bar{\mathcal{D}}'' = B_{\bar{\mu}}(\bar{\omega}_-) \setminus \bigcup_i B_{\varepsilon^{1/3}}(\bar{\omega}_i^o + (\boldsymbol{k}_i^o)^\perp)$, we have $|\langle \boldsymbol{k}, \omega\rangle| > \varepsilon^{1/3}$, $\forall\, \boldsymbol{k} \in \mathbb{Z}^n_{\bar{K}} \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}''\}$;

(2) for all ω in $\bar{\mathcal{D}}''' := B_{\bar{\mu}}(\bar{\omega}_-) \cap B_{\varepsilon^{1/3}}(\bar{\omega}^o + (\boldsymbol{k}_i^o)^\perp)$, for each i and for all $\boldsymbol{k} \in \mathbb{Z}^n_{\bar{K}} \setminus \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}', \boldsymbol{k}_i^o, \boldsymbol{k}''\}$, we have

$$|\langle \boldsymbol{k}, \omega\rangle| \geqslant n\bar{K}\bar{\mu}. \tag{8.2}$$

Proof. We first perform a reduction and then invoke Lemma 4.5. The transformed frequency segment

$$M'\bar{\omega}_a = \bar{\rho}_a\left(a, 0, \frac{\bar{P}}{\bar{Q}}\omega_2^*, \frac{\bar{p}}{\bar{q}}\omega_2^*, \hat{\omega}^*_{n-4}\right)$$

admits resonant integer vectors $\boldsymbol{k}'(M')^{-1} = (0,1,0,\dots,0)$ and $\boldsymbol{k}''(M')^{-1} = (0,0,\bar{Q}\bar{p}, -\bar{q}\bar{P}, \hat{0}_{n-4})$. If $a = a_i^o$, it also admits the integer vector $\boldsymbol{k}_i^o(M')^{-1}$. Now, remove the zero entry in $M'\bar{\omega}_a$, and then the resulting vector has the form of ω_a but with one fewer dimension. The integer vectors $\pi_{-2}(\boldsymbol{k}''(M')^{-1})$ and $\pi_{-2}(\boldsymbol{k}_i^o(M')^{-1})$ play the role of $\boldsymbol{k}'$ and $\boldsymbol{k}^o$ in Lemma 4.5, respectively, where $\pi_{-2} : \mathbb{R}^n \to \mathbb{R}^{n-1}$ means removing the second entry. Now this lemma follows from Lemma 4.5 up to a linear transformation. □

8.2 The nondegeneracy condition

Similar to Proposition 5.2, we have the following result.

Proposition 8.3. *Given $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1(\boldsymbol{k}') \subset \Pi_{\boldsymbol{k}'}C^r$, we choose δ, μ, $\boldsymbol{k}''$ and $\bar{\omega}_a$ as above. Then there exists an open and dense subset $\mathcal{O}_{1,2} = \mathcal{O}_{1,2}(\boldsymbol{k}', \boldsymbol{k}'')$ in the unit ball of $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r$ such that each $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}P$ with $\Pi_{\boldsymbol{k}'}\Pi_{\boldsymbol{k}',\boldsymbol{k}''}P = \Pi_{\boldsymbol{k}'}P$ and $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}P - \Pi_{\boldsymbol{k}'}P \in \mathcal{O}_{1,2}$ has a unique nondegenerate global maximum along the segment $y \in \boldsymbol{\omega}^{-1}(\bar{\omega}_-)$, up to finitely many bifurcations, where there are two nondegenerate global maxima. Moreover, the curves $\{\mathrm{Argmax}\Pi_{\boldsymbol{k}',\boldsymbol{k}''}P(y,\cdot), y \in \boldsymbol{\omega}^{-1}(\bar{\omega}_-)\}$, when projected to the set $\{\langle \boldsymbol{k}', x\rangle, x \in \mathbb{T}^n\} \times \mathbb{R}^n$, are within $O(\mu)$ Hausdorff distance of the curves $\{\mathrm{Argmax}\Pi_{\boldsymbol{k}'}P(y,\cdot), y \in \boldsymbol{\omega}^{-1}(\omega_-)\}$.*

Proof. The statement (without the "Moreover" part) can be obtained directly by applying the main theorem of [19] which is a higher-dimensional generalization of Proposition 5.4. Here, we give an argument using only Proposition 5.4. Since we have $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1$, $\Pi_{\boldsymbol{k}'}P$ has a nondegenerate global maximum up to finitely many bifurcations, where there are two nondegenerate global maxima. Moreover, $\Pi_{\boldsymbol{k}'}P$ determines δ, μ, $\bar{\omega}_a$ and $\boldsymbol{k}''$. We next decompose $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}P(y,x) = \Pi_{\boldsymbol{k}'}P(y,x) + \bar{P}(y, \langle \boldsymbol{k}', x\rangle, \langle \boldsymbol{k}'', x\rangle)$ induced by the decomposition $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r = \Pi_{\boldsymbol{k}'}C^r \oplus \Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r$. So we get $|\bar{P}|_{C^2} \leqslant C\frac{1}{|\boldsymbol{k}''|^2} \leqslant C\mu^2$ since $|\boldsymbol{k}''| \geqslant C\mu^{-1}$. We next make a linear coordinate change in x so that $Z_2(y, \mathsf{x}_1, \mathsf{x}_2) = Z(y, \mathsf{x}_1) + \bar{P}(y, \mathsf{x}_1, \mathsf{x}_2)$, where $\langle \boldsymbol{k}', x\rangle =: \mathsf{x}_1$, $\langle \boldsymbol{k}'', x\rangle =: \mathsf{x}_2$, $Z(y, \mathsf{x}_1) = \Pi_{\boldsymbol{k}'}P(y,x)$ and $Z_2(y, \mathsf{x}_1, \mathsf{x}_2) = \Pi_{\boldsymbol{k}',\boldsymbol{k}''}P(y,x)$.

By the choice of $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1$, for each y, we have that $\max_x Z(y, \mathsf{x}_1)$ is nondegenerate and attained at $\mathsf{x}_1^*(y)$. Then by the implicit function theorem, for small enough μ, the global maximum of Z_2 is attained at a point $(\bar{\mathsf{x}}_1^*, \bar{\mathsf{x}}_2^*)(y)$ with $|\bar{\mathsf{x}}_1^*(y) - \mathsf{x}_1^*(y)| \leqslant C\mu^2$. To see the nondegeneracy of the global maximum for Z_2, we consider that for each y and x_2, the function $Z_2(y, \cdot, \mathsf{x}_2)$ attains the global maximum at a point $\bar{\bar{\mathsf{x}}}_1^*(y, \mathsf{x}_2)$ that is within μ^2-distance from x_1^* by the implicit function theorem. Now the function $Z_2(y, \bar{\bar{\mathsf{x}}}_1^*(y, \mathsf{x}_2), \mathsf{x}_2)$ becomes a function of y and x_2. We then apply Proposition 5.4 to Z_2 to get an open and dense set $\tilde{\mathcal{O}}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ such that Z_2 has a nondegenerate global maximum along $\boldsymbol{\omega}^{-1}(\omega_-)$. The nondegeneracy can be achieved by adding a function $f \in C^r(\mathbb{T})$ of x_2 only. This induces an open and dense set $\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ in the unit ball of $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r$. □

In the proposition, each $\mathcal{O}_{1,2}$ depends on $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1$, so we define $\mathcal{O}_{1,2} = \mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$.

Lemma 8.4. *The union $\bigcup_{\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_1}(\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P) \oplus (C^r/\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r))$ intersects the unit ball of C^r in an open and dense subset of the latter.*

Proof. We first decompose $C^r = \Pi_{\boldsymbol{k}'}C^r \oplus (C^r/\Pi_{\boldsymbol{k}'}C^r)$ for each irreducible $\boldsymbol{k}' \in \mathbb{Z}^n$. Applying the following Kuratowski-Ulam Theorem 8.5, we get that the union $\bigcup_{\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_1}\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ is a set of the second category in C^r. Indeed, we first divide $\mathcal{O}_1 = \mathcal{O}_1(\boldsymbol{k}')$ into the union of the form $\mathcal{O}_1 = \bigcup_{\boldsymbol{k}''}\mathcal{O}_{1,\boldsymbol{k}''}$ such that each $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_{1,\boldsymbol{k}''}$ admits the frequency segment $\bar{\omega}_-$ having a second resonance $\boldsymbol{k}''$ (see Subsection 8.1, and note that $\Pi_{\boldsymbol{k}'}P$ determines δ and hence μ). Each $\mathcal{O}_{1,\boldsymbol{k}''}$ is open (may be empty). We use the notation $\mathfrak{B}_1(E)$ to denote the unit ball of a Banach space E.

Now each $\Pi_{\boldsymbol{k}'}P \in \mathcal{O}_{1,\boldsymbol{k}''}$ determines an open and dense subset $\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ in $\mathfrak{B}_1(\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r)$ by Proposition 8.3. So by the following Kuratowski-Ulam Theorem 8.5, the union $\bigcup_{\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_{1,\boldsymbol{k}''}}\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ is of the second category in the product space $\mathcal{O}_{1,\boldsymbol{k}''}\times\mathfrak{B}_1(\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r)$. Next, since each $\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ is open in $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r$, we get that the union $\bigcup_{\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_{1,\boldsymbol{k}''}}\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)$ is also open in $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r/\Pi_{\boldsymbol{k}'}C^r$. So we get that $\bigcup_{\Pi_{\boldsymbol{k}'}P\in\mathcal{O}_{1,\boldsymbol{k}''}}(\mathcal{O}_{1,2}(\Pi_{\boldsymbol{k}'}P)\times\mathfrak{B}_1(C^r/\Pi_{\boldsymbol{k}',\boldsymbol{k}''}C^r))$ is open and dense in $\mathcal{O}_{1,\boldsymbol{k}''}\times\mathfrak{B}_1(C^r/\Pi_{\boldsymbol{k}'}C^r)$.

Taking union over all the $\boldsymbol{k}''$, we get the statement in the lemma. □

Theorem 8.5 (Kuratowski-Ulam [41, Theorem 15.1])**.** *Let X and Y be two topological spaces, where Y has a countable bases. If $E \subset X\times Y$ is a set of the first category, then $E\cap\{x\}\times Y$ is the first category in Y for all x except a set of the first category.*

8.3 The KAM normal forms

Lemma 8.2 allows us to apply Proposition 3.3 in its two cases to obtain the following normal forms.

Lemma 8.6. *Let $\bar{\delta}$ be a small number satisfying $\bar{\delta} < \min\{3(|\boldsymbol{k}''|)^{-2}, \delta\}$ and $\bar{K} = (\bar{\delta}/3)^{-1/2}$. Then there exists an $\bar{\varepsilon}_1 = \bar{\varepsilon}_1(\bar{\delta}, \Lambda)$ such that for all $\varepsilon < \bar{\varepsilon}_1$, the following holds. Suppose $\omega^\star \in \bar{\mathcal{D}}''$ as in the case (1) of Lemma 8.2. Then there exists a symplectic transformation $\bar{\phi}$ defined on $B_\Lambda(0)\times\mathbb{T}^n$ that is $o_{\varepsilon\to0}(1)$ close to identity in the C^{r_0-1} norm such that*

$$\mathsf{H}\circ\bar{\phi}(x,Y) = \frac{1}{\sqrt{\varepsilon}}\langle\omega^\star, Y\rangle + \frac{1}{2}\langle\mathsf{A}Y, Y\rangle + \Pi_{\boldsymbol{k}',\boldsymbol{k}''}\mathsf{V} + \bar{\delta}\bar{\mathsf{R}}(x,Y), \tag{8.3}$$

where

(1) $\Pi_{\boldsymbol{k}',\boldsymbol{k}''}\mathsf{V} = V(\langle\boldsymbol{k}',x\rangle) + \delta\bar{V}(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}'',x\rangle)$ with A and V the same as those in Lemma 5.3, and $|\bar{V}(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}'',x\rangle)|_{r_0-2} \leqslant 1$;

(2) $\bar{\mathsf{R}}(x,Y) = \bar{\mathsf{R}}_I(x) + \bar{\mathsf{R}}_{II}(x,Y)$, where $\bar{\mathsf{R}}_I$ consists of Fourier modes of V not in the set $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}',\boldsymbol{k}''\}\cup\mathbb{Z}^n_{\bar{K}}$, and we have $|\bar{\mathsf{R}}_I|_{r_0-2}\leqslant 1$ and $|\bar{\mathsf{R}}_{II}|_{r_0-5}\leqslant 1$.

Lemma 8.7. *Let $\bar{\delta}$ and $\bar{K}$ be as in the previous lemma. Then there exists an $\bar{\varepsilon}_2 = \bar{\varepsilon}_2(\bar{\delta},\Lambda)$ such that for all $\varepsilon < \bar{\varepsilon}_2$, the following holds. Suppose $\omega^\star \in \bar{\mathcal{D}}'''$ as in the case (2) of Lemma 8.2. Then there exists a symplectic transformation $\bar{\phi}$ defined on $B_\Lambda(0)\times\mathbb{T}^n$ that is $o_{\varepsilon\to0}(1)$ close to identity in the C^{r_0-1} norm such that*

$$\mathsf{H}\circ\bar{\phi}(x,Y) = \frac{1}{\sqrt{\varepsilon}}\langle\omega^\star, Y\rangle + \frac{1}{2}\langle\mathsf{A}Y, Y\rangle + \Pi_{\boldsymbol{k}',\boldsymbol{k}'',\boldsymbol{k}^o}\mathsf{V} + \bar{\delta}\bar{\mathsf{R}}(x,Y), \tag{8.4}$$

where

(1) (a) if $|\boldsymbol{k}^o| < K$, we have $\Pi_{\boldsymbol{k}',\boldsymbol{k}'',\boldsymbol{k}^o}\mathsf{V} = V(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}^o,x\rangle) + \delta\bar{V}(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}'',x\rangle,\langle\boldsymbol{k}^o,x\rangle)$ with A and V the same as those in Lemma 6.1 or

(b) if $|\boldsymbol{k}^o| \geqslant K$, we have $\Pi_{\boldsymbol{k}',\boldsymbol{k}'',\boldsymbol{k}^o}\mathsf{V} = V(\langle\boldsymbol{k}',x\rangle) + \delta\bar{V}(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}'',x\rangle,\langle\boldsymbol{k}^o,x\rangle)$ with A and V the same as those in Lemma 5.3;

in both cases, we have $|\bar{V}(\langle\boldsymbol{k}',x\rangle,\langle\boldsymbol{k}'',x\rangle,\langle\boldsymbol{k}^o,x\rangle)|_{r_0-2}\leqslant 1$;

(2) $\bar{\mathsf{R}}(x,Y) = \bar{\mathsf{R}}_I(x) + \bar{\mathsf{R}}_{II}(x,Y)$, where $\bar{\mathsf{R}}_I$ consists of Fourier modes of V not in the set $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''\}\cup\mathbb{Z}^n_{\bar{K}}$, and we have $|\bar{\mathsf{R}}_I|_{r_0-2}\leqslant 1$ and $|\bar{\mathsf{R}}_{II}|_{r_0-5}\leqslant 1$.

Now there are several subcases. We assume that $\langle\omega^\star,\boldsymbol{k}'\rangle = \langle\omega^\star,\boldsymbol{k}''\rangle = 0$.

(1) $\omega^\star$ is as in Lemma 8.6. The same argument as Proposition 5.2 gives that there is a C^{r-1} NHIC homeomorphic to $T^*\mathbb{T}^{n-2}$ if $\bar{\delta}$ is sufficiently small. The normal hyperbolicity is independent of ε or $\bar{\delta}$, but may depend on δ. This NHIC is a subset of the NHIC in Proposition 5.2.

(2) $\omega^\star$ is as in Item (1)(b) of Lemma 8.7. This case occurs when $|\boldsymbol{k}^o|\geqslant K$. We first apply Proposition 5.2 to reduce the Hamiltonian system to a system defined on $T^*\mathbb{T}^{n-1}$. The restricted system to the NHIC

would depend on x through $\langle \boldsymbol{k}^o, x\rangle$ and $\langle \boldsymbol{k}'', x\rangle$ up to a $\bar{\delta}$ perturbation. That means that the restricted system is at the double resonance. If the double resonance is weak, then it is treated as a single resonance given by $\boldsymbol{k}''$. Otherwise, we apply Proposition 6.9 to find an NHIC homeomorphic to $T^*\mathbb{T}^{n-2}$ and a Proposition 7.7 to find a generalized transition chain connecting two neighboring NHICs.

(3) $\omega^\star$ is as in Item (1)(a) of Lemma 8.7. This case occurs when $|\boldsymbol{k}^o| < K$, i.e., the vectors $\boldsymbol{k}'$ and $\boldsymbol{k}^o$ give rise to a strong double resonance for the first step of the reduction of orders. We call this case a strong triple resonance.

In the following, without loss of generality, we focus on the third case to explain how to introduce the extra resonance $\boldsymbol{k}''$. The other two cases can be reduced to Propositions 5.2, 6.9 and 7.7.

9 Dynamics around the triple resonance: NHICs

In this section, we perform the reduction of orders around the triple resonance. We find NHICs getting close to the triple resonance.

Definition 9.1 (The triple resonance submanifold)**.** Given three irreducible integer vectors $\boldsymbol{k}^o$, $\boldsymbol{k}'$ and $\boldsymbol{k}''$, we define the triple resonance submanifold as

$$\Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'') := \{y \mid \langle \boldsymbol{k}', \boldsymbol{\omega}(y)\rangle = \langle \boldsymbol{k}'', \boldsymbol{\omega}(y)\rangle = \langle \boldsymbol{k}^o, \boldsymbol{\omega}(y)\rangle = 0\}.$$

We assume $\omega^\star \in \Sigma(\boldsymbol{k}', \boldsymbol{k}'')$ and is within $\varepsilon^{1/3}$ distance of $\Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')$. Again there are two subcases depending on whether $\omega^\star$ is within $\Lambda\varepsilon^{1/2}$ distance of $\Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')$ or not. The case of $\mathrm{dist}(\omega^\star, \Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')) > \Lambda\varepsilon^{1/2}$ can be treated in the same way as Theorem 6.8 and Proposition 6.10, which is essentially reduced to the case of Lemma 8.6, so we skip this case and focus on the case of $\mathrm{dist}(\omega^\star, \Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')) \leqslant \Lambda\varepsilon^{1/2}$. Without loss of generality, we assume $y^\star \in \Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')$ so that $\omega^\star = \boldsymbol{\omega}(y^\star)$ is perpendicular to $\boldsymbol{k}'$, $\boldsymbol{k}^o$ and $\boldsymbol{k}''$.

9.1 The shear transformation

Similar to Lemma 4.4, there exists a matrix $M''' \in \mathrm{SL}(n, \mathbb{Z})$ whose first three rows are $\boldsymbol{k}^o$, $\boldsymbol{k}'$ and $\boldsymbol{k}''$, respectively. The matrix M''' induces a symplectic transformation

$$\mathfrak{M}''' : T^*\mathbb{T}^n \to T^*\mathbb{T}^n, \quad (x, Y) \mapsto (M'''x, M'''^{-t}Y), \quad A = M'''\mathsf{A}M'''^t.$$

We define $\boldsymbol{\omega} = M'''\omega^\star$, which has 0 as the first three entries since $y^\star \in \Sigma(\boldsymbol{k}^o, \boldsymbol{k}', \boldsymbol{k}'')$. By the symplectic transformation $\mathfrak{M}'''$, one obtains the Hamiltonian

$$H := \mathfrak{M}'''^{-1*}(\mathsf{H}\bar{\phi}) = \frac{1}{2}\langle AY, Y\rangle + V(x_1, x_2) + \delta\bar{V}(x_1, x_2, x_3) + \frac{1}{\sqrt{\varepsilon}}\langle \hat{\boldsymbol{\omega}}_{n-3}, \hat{Y}_{n-3}\rangle + \bar{\delta}\bar{R}(x, Y), \tag{9.1}$$

where $\bar{R} = \mathfrak{M}'''^{-1*}\bar{\mathsf{R}}$. The matrix M''' depends on δ through $\boldsymbol{k}''$ but is independent of $\bar{\delta}$.

We next introduce the shear transformation as we did in Lemma 6.4 to block diagonalize A. Let $A, S''' \in \mathrm{SL}(n, \mathbb{R})$ be defined as follows:

$$A = \begin{bmatrix} \tilde{A}_3 & \breve{A}_3 \\ \breve{A}_3^t & \hat{A}_3 \end{bmatrix}, \quad S''' = \begin{bmatrix} \mathrm{id}_3 & 0 \\ -\breve{A}_3^t\tilde{A}_3^{-t} & \mathrm{id}_{n-3} \end{bmatrix}, \tag{9.2}$$

where $\tilde{A}_3$, $\breve{A}_3$ and $\hat{A}_3$ are 3×3, $3\times(n-3)$ and $(n-3)\times(n-3)$, respectively. With the shear matrix, we introduce a symplectic transformation

$$\mathfrak{S}''' : T^*\mathbb{T}^n \to T^*\mathbb{T}^n_{S'''}, \quad (x, Y) \mapsto (S'''x, S'''^{-t}Y) =: (\mathsf{x}, \mathsf{y}), \quad \omega_{S'''} := S'''\boldsymbol{\omega},$$

which transforms the Hamiltonian into the following form defined on $T^*\mathbb{T}^n_{S'''}$:

$$\mathsf{H}_{S'''} := (\mathfrak{S}'''\mathfrak{M}''')^{-1*}(\mathsf{H}\circ\phi) = \left[\frac{1}{2}\langle \tilde{A}_3\tilde{\mathsf{y}}_3, \tilde{\mathsf{y}}_3\rangle + V(\tilde{\mathsf{x}}) + \delta\bar{V}(\tilde{\mathsf{x}}_3)\right]$$

$$+\frac{1}{\sqrt{\varepsilon}}\langle\hat{\omega}_{S''',n-3},\hat{\mathsf{y}}_{n-3}\rangle+\frac{1}{2}\langle B_3\hat{\mathsf{y}}_{n-3},\hat{\mathsf{y}}_{n-3}\rangle+\mathfrak{S}'''^{-1*}(\bar{\delta}\bar{R}(\mathsf{x},\mathsf{y})),\tag{9.3}$$

where we define $B_3=(\hat{A}_3-\breve{A}_3^t\tilde{A}_3^{-1}\breve{A}_3)$, $\tilde{\mathsf{x}}_3=(\mathsf{x}_1,\mathsf{x}_2,\mathsf{x}_3)$, $\tilde{\mathsf{y}}_3=(\mathsf{y}_1,\mathsf{y}_2,\mathsf{y}_3)$ and $\tilde{\mathsf{x}}=(\mathsf{x}_1,\mathsf{x}_2)$. The norms of the matrices B_3 and S''' depend on δ but not on $\bar{\delta}$.

9.2 The existence of NHICs

To understand the full system $\mathsf{H}_{S'''}$, we first need to understand its bracketed subsystem in (9.3). The next lemma shows the existence of the NHIC of dimension 2 in the subsystem.

Lemma 9.2. *For any $\lambda>0$, there exists an open and dense subset $\tilde{\mathcal{O}}_2\subset C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V\in\tilde{\mathcal{O}}_2$ normalized by $\max V=0$, there exist a $\tilde{\delta}_2=\tilde{\delta}_2(V)$ and an open and dense subset $\tilde{\mathcal{O}}_{2,*}$ in the $\tilde{\delta}_2$-ball of $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$ such that for each $\delta\bar{V}\in\tilde{\mathcal{O}}_{2,*}$,*

(1) *the subsystem*

$$\mathsf{G}_{3,\delta}:=\frac{1}{2}\langle\tilde{A}_3\tilde{\mathsf{y}}_3,\tilde{\mathsf{y}}_3\rangle+V(\tilde{\mathsf{x}})+\delta\bar{V}(\tilde{\mathsf{x}}_3),\quad T^*\mathbb{T}^3\to\mathbb{R}\tag{9.4}$$

admits a C^{r-1} NHIC homeomorphic to $T^*\mathbb{T}$, up to finitely many bifurcations;

(2) *the NHIC is foliated by hyperbolic periodic orbits as Mather sets with rotation vectors $\nu(1,0,0)$ and $|\nu|\geqslant\lambda$;*

(3) *the absolute values of the normal Lyapunov exponents are bounded away from zero by $C\sqrt{\delta}$ for some constant $C>0$.*

The next proposition establishes the existence of NHICs in the full system.

Proposition 9.3. (1) *Given irreducible $\boldsymbol{k}',\boldsymbol{k}^o\in\mathbb{Z}^n_K$, let $y'^\star\in\Sigma(\boldsymbol{k}^o,\boldsymbol{k}')$.*

(2) *Let λ, $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y'^\star,x)\in\mathcal{O}_2$ and δ be as in the assumption of Proposition* 6.9.

(3) *Let $\boldsymbol{k}''$ be the third resonance given in Subsection* 8.1 *and consider $y''^\star\in\Sigma(\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}'')$ that is μ-close to $y'^\star$.*

Then there exists an open and dense set $\mathcal{O}_{2,*}=\mathcal{O}_{2,*}(\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y'^\star,\cdot),\boldsymbol{k}'')$ in the unit ball of $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}C^r(\mathbb{T}^n)/\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}C^r(\mathbb{T}^n)$ such that for each $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}P(y''^\star,\cdot)$ with

$$\Pi_{\boldsymbol{k}^o,\boldsymbol{k}'}\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}P(y''^\star,\cdot)=\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y''^\star,\cdot),\quad\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}P(y''^\star,\cdot)-\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y''^\star,\cdot)\in\mathcal{O}_{2,*},$$

there exists a $\bar{\delta}_1=\bar{\delta}_1(\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}P(y''^\star,\cdot),\lambda,\delta)>0$ such that for all $0<\bar{\delta}\leqslant\bar{\delta}_1$ and all $0<\varepsilon<\bar{\varepsilon}_2$,

(1) *the Hamiltonian system* (8.4) *admits a C^{r-1} NHIC $\mathcal{C}(\boldsymbol{k}',\boldsymbol{k}'')$ homeomorphic to $T^*\mathbb{T}^{n-2}$ up to finitely many bifurcations*; *the normal hyperbolicity is independent of ε or $\bar{\delta}$, but may depend on δ;*

(2) *Mather sets in the region $B_\Lambda(0)\times\mathbb{T}^n$ with rotation vectors orthogonal to both $\boldsymbol{k}'$ and $\boldsymbol{k}''$ and of distance λ-away from $\varepsilon^{-1/2}\boldsymbol{\omega}(y''^\star)+(\boldsymbol{k}^o)^\perp$ lie inside $\mathcal{C}(\boldsymbol{k}',\boldsymbol{k}'')$.*

Proof. The proof is similar to those of Propositions 5.2 and 6.9. After the linear transformation induced by $S'''M'''$, the problem of finding the NHIC is reduced to Lemma 9.2. The NHIC persists if the $\bar{\delta}$ perturbation is sufficiently small. Here, we only explain two points. First, here we choose $\mathcal{O}_{2,*}$ to be in the unit ball of $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}C^r(\mathbb{T}^n)/\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}C^r(\mathbb{T}^n)$ rather than in a $\tilde{\delta}_2$-ball as in Lemma 9.2. The reason is that $\tilde{\delta}_2$ is determined by the persistence of the NHIC in the subsystem $\tilde{\mathsf{G}}$ of $\mathsf{G}_{3,\delta}$. The theorem of the NHIC requires only C^1 smallness of the perturbation to the Hamiltonian flow and we have that every function in the unit ball of $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''}C^r(\mathbb{T}^n)/\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}C^r(\mathbb{T}^n)$ has the C^{r-2} norm less than $|\boldsymbol{k}''|^{-2}<\delta_2$ in Proposition 6.9.

Next, we explain the difference of $y'^\star$ and $y''^\star$. For each $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}P(y'^\star,x)\in\mathcal{O}_2$, there exists an NHIC $\mathcal{C}(\boldsymbol{k}')$ that is λ-away from the double resonance by Proposition 6.9. If we perform the $\sqrt{\varepsilon}$-blowup based at the point $y''^\star$ that is μ-close to $y'^\star$ the resulting $\tilde{\mathsf{G}}$'s differ by $O(\mu)$ in the C^2 topology. Since the normal hyperbolicity of the NHIC $\mathcal{C}(\boldsymbol{k}')$ is independent of δ and $\mu=o(\delta)$, we see that for small enough δ, Proposition 6.9 that is stated for any $y^\star=y'^\star\in\Sigma(\boldsymbol{k}',\boldsymbol{k}^o)$ remains to hold for another $y''^\star\in\Sigma(\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}'')$ that is μ-close to $y'^\star$. $\square$

9.3 The proof of Lemma 9.2

The remaining part of this subsection is devoted to the proof of Lemma 9.2.

Proof of Lemma 9.2. Applying Theorem 6.6, we get an open and dense subset $\tilde{\mathcal{O}}_2 \subset C^r(\mathbb{T}^2)/\mathbb{R}$ such that for each $V \in \tilde{\mathcal{O}}_2$, the system $\tilde{\mathsf{G}} : T^*\mathbb{T}^2 \to \mathbb{R}$ admits an NHIC foliated by periodic orbits with rotation vectors $\nu(1,0)$ ($|\nu| > \lambda$). Let us now fix such a $V \in \tilde{\mathcal{O}}_2$.

We next block diagonalize the quadratic form $\langle \tilde{A}_3\tilde{\mathsf{y}}_3, \tilde{\mathsf{y}}_3\rangle$ by introducing one more shear transformation

$$S_3 = \begin{bmatrix} \mathrm{id}_2 & 0 \\ -\mathsf{a}_3\tilde{A}^{-1} & 1 \end{bmatrix} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ s_1 & s_2 & 1 \end{bmatrix},$$

where $\mathsf{a}_3 = (a_{31}, a_{32}) \in \mathbb{R}^2$ is the vector formed by the entries of A on the third row to the left of the diagonal. We can verify that

$$S_3\tilde{A}_3S_3^t = \begin{bmatrix} \tilde{A} & 0 \\ 0 & b_3 \end{bmatrix}, \tag{9.5}$$

where $b_3 = a_{33} - \mathsf{a}_3\tilde{A}^{-1}\mathsf{a}_3^t$. This linear transformation induces a linear transformation $\mathfrak{S}_3 : (\tilde{\mathsf{x}}_3, \tilde{\mathsf{y}}_3) \mapsto (S_3\tilde{\mathsf{x}}_3, S_3^{-t}\tilde{\mathsf{y}}_3) =: (\tilde{x}_3, \tilde{y}_3)$ and transforms the Hamiltonian $\mathsf{G}_{3,\delta}$ to the following system $\mathfrak{S}_3^*\mathsf{G}_{3,\delta} := \mathcal{G}_{3,\delta}$ of the form:

$$\mathcal{G}_{3,\delta} = \frac{1}{2}\langle \tilde{A}\tilde{y}, \tilde{y}\rangle + V(\tilde{x}) + \frac{b_3}{2}y_3^2 + \delta\bar{V}(S_3^{-1}\tilde{x}_3), \quad T^*\mathbb{T}^3_{S_3} \to \mathbb{R}. \tag{9.6}$$

In the above system $\mathcal{G}_{3,\delta}$, we apply Theorem 6.6 with the homology class $g = (1,0)$ and find the NHIC in the subsystem $\tilde{\mathsf{G}} := \frac{1}{2}\langle \tilde{A}\tilde{y}, \tilde{y}\rangle + V(\tilde{x})$. Restricted to the NHIC, the subsystem $\tilde{\mathsf{G}}$ is reduced to a system of one degree of freedom denoted by $\tilde{h}(I)$ in action-angle coordinates (see Proposition 7.8). We restrict to the region $|h'(I)| > \lambda$. In the case of $\delta = 0$, restricted to the NHIC, the system $\mathcal{G}_{3,0}$ becomes $\overline{\mathcal{G}_{3,0}} := \tilde{h}(I) + \frac{b_3}{2}y_3^2$ defined on $T^*\mathbb{T}^2_{\bar{S}}$, where

$$\bar{S} = \begin{bmatrix} 1 & 0 \\ s_1 & 1 \end{bmatrix} \in \mathrm{SL}(2,\mathbb{R}) \quad \text{and} \quad \mathbb{T}^2_{\bar{S}} = \mathbb{T}^3_{S_3}/\mathbb{T}^1.$$

When the δ-perturbation in $\mathcal{G}_{3,\delta}$ is turned on, we apply the theorem of the NHIM to get that $\mathcal{G}_{3,\delta}$ admits an NHIC homeomorphic to $T^*\mathbb{T}^2_{\bar{S}}$ for sufficiently small δ and for any $\lambda > 0$, the bound $\tilde{\delta}_2$ is determined in the same way as the proof of Proposition 6.9(2). The restriction of $\mathcal{G}_{3,\delta}$ to the NHIC has the form

$$\bar{\mathcal{G}}_{3,\delta} := \tilde{h}(I) + \frac{b_3}{2}y_3^2 + \delta\bar{Z}(I,\varphi,x_3,y_3), \quad T^*\mathbb{T}^2_{\bar{S}} \to \mathbb{R}, \tag{9.7}$$

where we have $\bar{Z} = \bar{V}(\tilde{x}(I,\varphi), x_3 - (s_1,s_2)\cdot\tilde{x}(I,\varphi)) + O(\delta)$. Indeed, the leading term in $\bar{Z}$ is obtained by evaluating $\bar{V}(S_3^{-1}\tilde{x}_3)$ restricted to the unperturbed NHIC with $\tilde{x} = \tilde{x}(I,\varphi)$. The $O(\delta)$ error is created by the deformation of the NHIC under the perturbation.

Finally, going back to the original system $\mathsf{G}_{3,\delta}$, we obtain an expression for the restricted system to the NHIC which is homeomorphic to $T^*\mathbb{T}^2$. We introduce the following undo-shear transformation:

$$\bar{\mathfrak{S}} : (\varphi, x_3; I, y_3) \mapsto (\bar{S}(\varphi, x_3); \bar{S}^{-t}(I, y_3)) = (\varphi, s_1\varphi + x_3; I - s_1y_3, y_3) =: (\varphi, \mathsf{x}_3; J, \mathsf{y}_3), \tag{9.8}$$

under which, we get the restriction of $\mathsf{G}_{3,\delta}$ to the NHIC

$$\bar{\mathsf{G}}_{3,\delta} := \left[\tilde{h}(J + s_1\mathsf{y}_3) + \frac{1}{2}b_3\mathsf{y}_3^2 + \delta U(J, \mathsf{y}_3, \varphi, \mathsf{x}_3)\right] : T^*\mathbb{T}^2 \to \mathbb{R}, \tag{9.9}$$

where $U(J, \mathsf{y}_3, \varphi, \mathsf{x}_3) = \bar{V}(\tilde{\mathsf{x}}(I,\varphi), \mathsf{x}_3 + (s_1,s_2)\cdot\tilde{\mathsf{x}}(I,\varphi) - s_1\varphi) + O(\delta)$ with $I = J + s_1\mathsf{y}_3$. Moreover, the $O(\delta)$ part depends on the angular variables x_3 and φ in the same way as the leading term. To see that x_3 is defined on $\mathbb{T}^1$, we lift a periodic orbit $\tilde{\mathsf{x}}$ with the homology class $g = (1,0)$ to the universal cover. As

$\varphi \mapsto \varphi + 1$ we get $\tilde{\mathsf{x}} \mapsto \tilde{\mathsf{x}} + (1, 0)$ and after the shear and undo-shear transformations $(s_1, s_2) \cdot \tilde{\mathsf{x}}(I, \varphi) - s_1\varphi \mapsto (s_1, s_2) \cdot \tilde{\mathsf{x}}(I, \varphi) - s_1\varphi$.

We apply the procedure of the order reduction to the system $\bar{\mathsf{G}}_{3,\delta} : T^*\mathbb{T}^2 \to \mathbb{R}$. Namely, we want to apply Theorem 6.6 with the homology class $g = (1, 0)$ to get an NHIC and restrict the system to the NHIC to get a system of one degree of freedom. It is known that all its Mather sets with rotation vectors $\nu(1, 0)$ ($|\nu| > \lambda$) are supported on periodic orbits due to the two-dimensionality. Going back to the system $\mathsf{G}_{3,\delta}$ of three degrees of freedom, we obtain that all its Mather sets with rotation vectors $\nu(1, 0, 0)$ ($|\nu| > \lambda$) are supported on periodic orbits. It remains to show the nondegeneracy and hyperbolicity of the periodic orbits if $\delta\bar{V}$ is chosen in an open and dense subset $\tilde{\mathcal{O}}_{2,*}$ of the $\tilde{\delta}_2$-ball of the quotient $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$. The proof is essentially the same as the proof of [20, Theorem 6.6], but there is a subtle point: here we are only allowed to perturb the potential of the system $\mathsf{G}_{3,\delta}$ of three degrees of freedom but cannot perturb $\bar{\mathsf{G}}_{3,\delta}$ of two degrees of freedom directly.

We next show how to adapt the proof of [20] to our setting. Let us briefly recall the perturbation argument of [20]. For a Tonelli Lagrangian system $L(x, \dot{x}) : T\mathbb{T}^2 \to \mathbb{R}$,

(1) we first pick a section $\{x_1 = 0\}$ and reduce it to a nonautonomous system defined on $T\mathbb{T} \times \mathbb{T} \to \mathbb{R}$, and then introduce the action functional $F(E, x_2) : I \times \mathbb{T} \to \mathbb{R}$ where E is the energy level and I is the energy interval, by evaluating the action along the orbit on the energy level E starting and ending at the same point $x_2 \in \mathbb{T}$;

(2) we then choose the perturbation of the form $A_\ell \cos \ell x_2 + B_\ell \sin \ell x_2$, $A_\ell, B_\ell \in [\epsilon, 2\epsilon]$, $\ell = 1, 2$; by the construction in [20, Section 3], such a perturbation to the Lagrangian becomes a perturbation of the same form to the action functional F;

(3) an open and dense subset of the perturbation can make the global minimum of F nondegenerate uniformly for $E \in I$ (see Proposition 5.4 here and [20, Theorem 3.1]);

(4) nondegenerate periodic orbits are hyperbolic.

Now we show that the above argument applies to the subsystem $\bar{\mathsf{G}}_{3,\delta}$ of two degrees of freedom by perturbing the system $\mathsf{G}_{3,\delta}$ of three degrees of freedom. In place of the above step (1), we pick the section $\{\varphi = 0\}$ in the subsystem $\bar{\mathsf{G}}_{3,\delta}$. Next, consider a perturbation to the system $\mathsf{G}_{3,\delta}$ depending only on x_3 of the form $A_\ell \cos \ell\mathsf{x}_3 + B_\ell \sin \ell\mathsf{x}_3$, $A_\ell, B_\ell \in [\epsilon\delta, 2\epsilon\delta]$, $\ell = 1, 2$ as the above Item (2). Restricted to the section $\{\varphi = 0\}$ in the subsystem $\bar{\mathsf{G}}_{3,\delta}$, we get a perturbation of the same form up to a horizontal translation by a constant (see the expression of U above). Then Items (3) and (4) go through without any change. Since the system $\bar{\mathsf{G}}_{3,\delta}$ is already restricted to an NHIC, its hyperbolic periodic orbit is also hyperbolic in the system $\mathsf{G}_{3,\delta}$. □

Lemma 9.4. *We have the following estimates for the constants* b_3 *and* s_1 *appearing in Equation* (9.9):

$$b_3 = \mathrm{const}_{b_3}|\boldsymbol{k}''|^2, \quad s_1 = \mathrm{const}_{s_1}|\boldsymbol{k}''|,$$

where the constants are independent of δ, $\mathrm{const}_{b_3} > 0$ *and* $\mathrm{const}_{s_1} \in \mathbb{R}$.

Proof. Recall the definition of b_3 (see (9.5)) $b_3 = a_{33} - \mathsf{a}_3\tilde{A}^{-1}\mathsf{a}_3^t$ and s is the first entry of $\mathsf{a}_3\tilde{A}^{-1}$. The (i, j)-th entry of $A = M'''\mathsf{A}M'''^t$ is $m_i\mathsf{A}m_j^t$, where m_i and m_j are the i-th and j-th rows of M''', respectively. Since the first three rows of M''' are $\boldsymbol{k}^o$, $\boldsymbol{k}'$ and $\boldsymbol{k}''$, respectively, we get

$$b_3 = \boldsymbol{k}''\mathsf{A}(\boldsymbol{k}'')^t - (\boldsymbol{k}''\mathsf{A}\boldsymbol{K}^t)(\boldsymbol{K}\mathsf{A}\boldsymbol{K}^t)^{-1}(\boldsymbol{K}\mathsf{A}(\boldsymbol{k}'')^t),$$

and s is the first entry of $\boldsymbol{k}''\mathsf{A}\boldsymbol{K}^t(\boldsymbol{K}\mathsf{A}\boldsymbol{K}^t)^{-1}$, where we denote by $\boldsymbol{K}$ the matrix of $2 \times n$ whose two rows are $\boldsymbol{k}^o$ and $\boldsymbol{k}'$, respectively. Now s_1 is estimated easily as $\mathrm{const}|\boldsymbol{k}''|$ since $\mathsf{A}\boldsymbol{K}^t(\boldsymbol{K}\mathsf{A}\boldsymbol{K}^t)^{-1}$ does not depend on δ.

We focus on b_3 in the following. Since A is positive definite, we decompose $\mathsf{A} = CC^t$ for some $C \in \mathrm{GL}(n, \mathbb{R})$ and let $\boldsymbol{k}''C =: \mathsf{k}$ and $\boldsymbol{K}C =: \mathsf{K}$. This gives us $b_3 = \mathsf{k}(\mathsf{k}^t - \mathsf{K}^t(\mathsf{K}\mathsf{K}^t)^{-1}\mathsf{K}\mathsf{k}^t)$. Now, we recall the Gauss least-squares method. Though the equation $\mathsf{K}^t\mathsf{x} =: \mathsf{k}^t$, in general is not solvable for $\mathsf{x} \in \mathbb{R}^2$, we can seek for a least-squares solution given by $\mathsf{x}_{ls} = (\mathsf{K}\mathsf{K}^t)^{-1}\mathsf{K}\mathsf{k}^t$, which has a geometric interpretation as follows. The vector $\mathsf{K}\mathsf{x}_{ls} = \mathsf{K}(\mathsf{K}\mathsf{K}^t)^{-1}\mathsf{K}\mathsf{k}^t$ is the projection of k to the linear space spanned by the column vectors of K^t. Hence, $(\mathsf{k}^t - \mathsf{K}^t(\mathsf{K}\mathsf{K}^t)^{-1}\mathsf{K}\mathsf{k}^t)$ is the projection of k to the orthogonal complement of the linear space spanned by the column vectors of K^t. We see from the construction of the vectors $\boldsymbol{k}^o$, $\boldsymbol{k}'$ and $\boldsymbol{k}''$ that $\boldsymbol{k}''$ forms a nonzero angle with the plane $\mathrm{span}\{\boldsymbol{k}', \boldsymbol{k}^o\}$ independent of δ, since as $\mu \to 0$ one has

$$\bar{\boldsymbol{k}}' = (0, \bar{Q}\bar{p}, -\bar{q}\bar{P}, \hat{0}_{n-4}) \Bigg\| \left(0, 1, -\frac{\bar{q}\bar{P}}{\bar{Q}\bar{p}}, \hat{0}_{n-3}\right) \to \left(0, 1, -\frac{\omega_2^*\bar{P}}{\omega_4^*\bar{Q}}, \hat{0}_{n-3}\right),$$

which is obtained from $\boldsymbol{k}''$ by removing the second entry of $\boldsymbol{k}''M'$, and the matrices $\boldsymbol{K}$, M' and A do not depend on δ. This linear independence relation is preserved by the linear transformation C. Hence, we get $b_3 = c|\boldsymbol{k}''|^2$ for some constant $c > 0$ independent of δ. □

10 Dynamics around the triple resonance: The ladder

In this section, we show how to find the orbit going around the triple resonance. For this purpose, we first describe the special structure of the α-function, whose flat looks like a pizza. We next show that after applying the c-equivalence path of Theorem 7.2, there is a further misalignment, which is the new difficulty caused by the high dimensionality. We next show how to overcome this difficulty by introducing a ladder, and thus prove the existence of the diffusion orbit moving along ω_- for the $n = 4$ case.

10.1 The description of the α-function

Applying Theorem 7.1 to the system $\mathsf{G}_{3,\delta}$, we get a three-dimensional flat on the energy level $\min \alpha_{\mathsf{G}_{3,\delta}}$. Next, applying Lemma 6.7 twice (since we have applied Theorem 6.6 twice in the proof of Lemma 9.2) we see that the NHICs in Lemma 9.2 correspond to two channels

$$\mathbb{C}_\pm = \{\partial\beta_{\mathsf{G}_{3,\delta}}(\nu(1,0,0)) \mid \pm\nu > \lambda\} \subset H^1(\mathbb{T}^3, \mathbb{R}).$$

For each $c \in \mathbb{C}_\pm$, the corresponding Mather set $\tilde{\mathcal{M}}(c)$ lies in the NHIC with $\pm\nu > \lambda > 0$. Lemma 6.7 implies that the Mather set $\tilde{\mathcal{M}}(c)$ remains the same for c in a two-dimensional rectangle. Taking union over all the energy levels, we see that each $\mathbb{C}_\pm$ is a three-dimensional rectangular prism. Moreover, the channels $\mathbb{C}_+$ and $\mathbb{C}_-$ are centrally symmetric to each other since $\mathsf{G}_{3,\delta}$ is reversible.

In the following, since the rationality and irrationality of the rotation vectors do not play a role, for simplicity of notations, we work with the system $\mathcal{G}_{3,\delta} := \mathfrak{S}_3^*\mathsf{G}_{3,\delta} : T^*\mathbb{T}^3_{S_3} \to \mathbb{R}$ (see (9.6)), which is related to the system $\mathsf{G}_{3,\delta} : T^*\mathbb{T}^3 \to \mathbb{R}$ (see (9.4)) by the symplectic transformation induced by S_3. Similarly, we work with $\bar{\mathcal{G}}_{3,\delta} : T^*\mathbb{T}^2_{\bar{S}} \to \mathbb{R}$ (see (9.7)) instead of the system $\bar{\mathsf{G}}_{3,\delta} : T^*\mathbb{T}^2 \to \mathbb{R}$ (see (9.9)) for the system restricted to the NHICs. We first have the following description of the α-functions.

Lemma 10.1. (1) *We have the estimate for the α-function of $\mathcal{G}_{3,\delta}$: $\|\alpha_{\mathcal{G}_{3,\delta}} - \alpha_{\mathcal{G}_{3,0}}\|_{C^0} \leqslant \delta$ with $\alpha_{\mathcal{G}_{3,0}}(c) = \alpha_{\tilde{\mathsf{G}}}(\tilde{c}) + \frac{b_3}{2}c_3^2$.*

(2) *For the α-functions of the Hamiltonian $\bar{\mathcal{G}}_{3,\delta}$ restricted to the NHIC, we have the estimate $\|\alpha_{\bar{\mathcal{G}}_{3,\delta}} - \alpha_{\bar{\mathcal{G}}_{3,0}}\|_{C^0} \leqslant \delta$ with $\alpha_{\bar{\mathcal{G}}_{3,0}}(c_1, c_3) = \tilde{h}(c_1) + \frac{b_3}{2}c_3^2$.*

The proof of this lemma is the same as that of Lemma 7.5, so we skip it.

Proposition 10.2. *Under the assumption of Lemma 9.2, the flat $\mathbb{F}_0 = \{c \mid \alpha_{\mathcal{G}_{3,\delta}}(c) = \min \alpha_{\mathcal{G}_{3,\delta}}\}$ is a three-dimensional convex set lying in a $O(\sqrt{\delta/b_3})$-neighborhood of the disk $\tilde{\mathbb{F}}_0 \times \{c_3 = \hat{0}\}$, where $\tilde{\mathbb{F}}_0 = \mathrm{Argmin}\alpha_{\tilde{\mathsf{G}}}$.*

Proof. The fact that the flat is three-dimensional is given by Theorem 7.1. Since we have $|\mathsf{G}_{3,\delta} - \mathsf{G}_{3,0}|_{C^0} < \delta$, we obtain $|\alpha_{\mathsf{G}_{3,\delta}}(c) - \alpha_{\mathsf{G}_{3,0}}(c)| \leqslant \delta$, $\forall\, c \in H^1(\mathbb{T}^3, \mathbb{R})$ (see Lemma 7.5). After the same linear transformation S_3^t, this gives

$$|\alpha_{\mathcal{G}_{3,\delta}} - \alpha_{\mathcal{G}_{3,0}}|_{C^0} \leqslant \delta. \tag{10.1}$$

Since we have $\alpha_{\mathcal{G}_{3,0}}(c) = \alpha_{\tilde{\mathsf{G}}}(\tilde{c}) + \frac{b_3}{2}c_3^2$, we get $\alpha_{\mathcal{G}_{3,0}}(c) > 2\delta$, if $|c_3| > 2\sqrt{\delta/b_3}$ and $\tilde{c} \in \tilde{\mathbb{F}}_0$. As $\alpha_{\tilde{\mathsf{G}}}$ is non-negative, it follows from (10.1) that $\alpha_{\mathcal{G}_{3,\delta}}(c) > \delta$. Also due to (10.1), we have $\min \alpha_{\mathcal{G}_{3,0}} \leqslant \delta$. Therefore, $\alpha_{\mathcal{G}_{3,\delta}}(c) > \min \alpha_{\mathcal{G}_{3,\delta}}$, if $|c_3| > 2\sqrt{\delta/b_3}$. This completes the proof for the $O(\sqrt{\delta/b_3})$ estimate. □

Therefore, the flat looks like a pizza, horizontal in the direction of $\tilde{c}$ with small thickness of order $O(\sqrt{\delta/b_3})$ (see Figure 2).

10.2 Construction of the ladder

The generalized transition chain built by the application of the c-equivalence mechanism (see Proposition 7.4) does not connect the channels $\mathbb{C}_\pm$ mainly due to the misalignment in the c_3 component. In this subsection, we show how the misalignment appears and how to overcome it to build a transition chain connecting $\mathbb{C}_\pm$, which is called a ladder (see Figure 2).

The next result gives the existence of the generalized transition chain in the subsystem $\mathsf{G}_{3,\delta}$.

Proposition 10.3. *Let $V \in \hat{\mathcal{O}}_3 \cap \tilde{\mathcal{O}}_2 \subset C^r(\mathbb{T}^2)/\mathbb{R}$, normalized by $\max V = 0$ (see Lemma 7.6 for $\hat{\mathcal{O}}_3$ and Lemma 9.2 for $\tilde{\mathcal{O}}_2$), and λ be as in Lemma 7.6. Then there exist $\tilde{\delta}_3$ ($\leqslant \tilde{\delta}_2$) and an open and dense subset $\hat{\mathcal{O}}_{3,*}$ ($\subset \tilde{\mathcal{O}}_{2,*}$) in the $\tilde{\delta}_3$-ball of $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$ such that for each $\delta\bar{V} \in \hat{\mathcal{O}}_{3,*}$, the following holds. For any point $\mathsf{c}^* = (\tilde{\mathsf{c}}^*, \mathsf{c}_3^*) \in \mathbb{C}_+$ satisfying $\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda(1,0))) < \alpha_{\tilde{\mathsf{G}}}(\tilde{\mathsf{c}}^* + (s_1, s_2)\mathsf{c}_3^*) < 2\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda(1,0)))$ on the energy level of $\alpha_{\mathsf{G}_{3,\delta}}(\mathsf{c}^*)$, there exists a generalized transition chain connecting $\mathsf{c}^* \in \mathbb{C}_+$ to $-\mathsf{c}^* \in \mathbb{C}_-$.*

Similar to Proposition 7.7, we have the following result extending the generalized transition chain of the system $\mathsf{G}_{3,\delta}$ to the full system.

Proposition 10.4. *For each $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P \in \mathcal{O}_3 \subset \Pi_{\boldsymbol{k}^o,\boldsymbol{k}'} C^r(\mathbb{T}^n)/\mathbb{R}$, as in Proposition 7.7, let δ_3 be as in Proposition 7.7. For any $\delta < \delta_3$ and $|\boldsymbol{k}''| > K = (\delta/3)^{-1/2}$, there exists an open and dense subset $\mathcal{O}_{3,*}$ in the unit ball of $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''} C^r(\mathbb{T}^n)/\Pi_{\boldsymbol{k}',\boldsymbol{k}^o} C^r(\mathbb{T}^n)$ such that for each $\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''} P(y^\star, \cdot)$ satisfying*

$$\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}(\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''} P) = \Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P \quad \text{and} \quad \Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''} P - \Pi_{\boldsymbol{k}',\boldsymbol{k}^o} P \in \mathcal{O}_{3,*},$$

there exists a $\bar{\delta}_2 = \bar{\delta}_2(\Pi_{\boldsymbol{k}^o,\boldsymbol{k}',\boldsymbol{k}''} P) > 0$ such that for all $0 < \bar{\delta} \leqslant \bar{\delta}_2$ and each $\hat{c}^* \in \mathbb{R}^{n-3}$ satisfying $\|\hat{c}^*\| < \Lambda$, there exists a generalized transition chain of the Hamiltonian system (8.4) connecting the two channels corresponding to the NHICs $\mathcal{C}(\boldsymbol{k}', \boldsymbol{k}'')$.

We next prove Proposition 10.3, which is reduced to the following two lemmas.

Lemma 10.5. *For each $V \in \hat{\mathcal{O}}_3 \subset C^r(\mathbb{T}^2)/\mathbb{R}$ normalized by $\max V = 0$ as in Lemma 7.6, let λ be as in Lemma 7.6. Then there exists a $\tilde{\delta}_3$ ($< \tilde{\delta}_2$ in Lemma 9.2) such that for any $\delta\bar{V}$ in the $\tilde{\delta}_3$-ball of $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$ and any c^* as in Proposition 10.3, there is a generalized transition chain of the system $\mathsf{G}_{3,\delta}$ connecting the point $\mathsf{c}^* = S_3^t c^* = S_3^t(\tilde{c}^*, c_3^*)$ to a point $S_3^t(\tilde{c}^\sharp, c_3^*)$, where $\tilde{c}^\sharp$ is δ-close to $-\tilde{c}^*$ and satisfies $\alpha_{\mathcal{G}_{3,\delta}}(\tilde{c}^\sharp, c_3^*) = \alpha_{\mathcal{G}_{3,\delta}}(c^*) = \alpha_{\mathsf{G}_{3,\delta}}(\mathsf{c}^*)$.*

Proof. Given c^*, we fix $c_3 = c_3^*$ and define the path $\Gamma_\delta(c_3^*) = \{\tilde{c} \in \mathbb{R}^2 \mid \alpha_{\mathcal{G}_{3,\delta}}(\tilde{c}, c_3^*) = \alpha_{\mathcal{G}_{3,\delta}}(c^*)\}$. In the case of $\delta = 0$, this path lies on the constant energy level of $\alpha_{\tilde{\mathsf{G}}}$ and in the small positive δ case, the path undergoes a δ-perturbation as proved in Proposition 7.7.

On $\Gamma_\delta(c_3^*)$, we find a point that is closest to $(-\tilde{c}^*, c_3^*)$ and denote it by $(\tilde{c}^\sharp, c_3^*)$ where $|\tilde{c}^\sharp + \tilde{c}^*| \leqslant C\delta$.

The fact that the path $\Gamma_\delta(c_3^*)$ is a generalized transition chain follows from Proposition 7.4 (see also Proposition 7.7) and the upper-semi-continuity of the Mañé set. □

Lemma 10.6. *Let $\lambda, \tilde{\mathcal{O}}_2 \subset C^r(\mathbb{T}^2)/\mathbb{R}$, $\tilde{\delta}_2$ be as in Lemma 9.2, $\hat{\mathcal{O}}_3$ be as in Lemma 7.6, and $V \in \tilde{\mathcal{O}}_2 \cap \hat{\mathcal{O}}_3$. Then there exists an open and dense subset $\hat{\mathcal{O}}_{3,*}$ in the $\tilde{\delta}_3$-ball of $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$ such that for each $\delta\bar{V} \in \hat{\mathcal{O}}_{3,*}$, let c^* and $\tilde{c}^\sharp$ be as in Lemma 10.5, and there is a generalized transition chain of the system of $\mathsf{G}_{3,\delta}$ lying on the energy level $\alpha_{\mathsf{G}_{3,\delta}}(\mathsf{c}^*)$ connecting $S_3^t(\tilde{c}^\sharp, c_3^*)$ to $-\mathsf{c}^* \in \mathbb{C}_-$.*

Proof. We fix the energy level $E = \alpha_{\mathsf{G}_{3,\delta}}(\mathsf{c}^*)$ in the system $\mathsf{G}_{3,\delta}$, on which there exists an NHIM restricted to which the Hamiltonian system is $\bar{\mathsf{G}}_{3,\delta}$. We define (see Figure 4)

$$\bar{\mathbb{L}}_\delta := \{(c_1, c_3) \mid \alpha_{\bar{\mathcal{G}}_{3,\delta}}(c_1, c_3) = E,\ |\tilde{h}'(c_1)| > \lambda\}.$$

The variable c_2 does not appear due to Proposition 7.8(2).

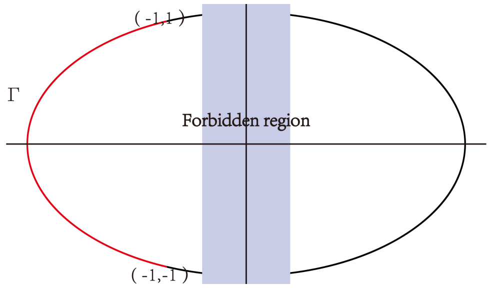

Figure 4 (Color online) The ladder for $\alpha_{\bar{\mathcal{G}}_{3,\delta}}$

To see the path $\bar{\mathbb{L}}_\delta$ clearly, we introduce the following coordinate change $\mathcal{R}:(c_1,c_3)\mapsto(c_1,\frac{1}{s_1}c_3)$. In the new coordinates, the α-function for the restricted system has the form

$$\mathcal{R}^*\alpha_{\bar{\mathcal{G}}_{3,0}}(c_1,c_3):=\tilde{h}(c_1)+\frac{b_3}{2s_1^2}c_3^2.$$

From Lemma 9.4, we see that b_3/s_1^2 is of order one as $\delta\to 0$. The frequency $\nu(1,0)$ for the system $\bar{\mathsf{G}}_{3,0}$ is transformed to $\nu(1,1)$ under the linear transformation $\bar{S}$ followed by $\mathcal{R}$. Its Legendre transformation of $\nu(1,1)$ is solved from the equation

$$\left(\tilde{h}'(c_1),\frac{b_3}{s_1^2}c_3\right)=\nu(1,1)$$

for $\mathcal{R}^*\alpha_{\bar{\mathcal{G}}_{3,0}}$. For the α-function $\mathcal{R}^*\alpha_{\bar{\mathcal{G}}_{3,\delta}}$, we get that the projection of $\mathbb{C}_\pm$ to the (c_1,c_3) plane is δ-close to the set $\{(\tilde{h}'(c_1),\frac{b_3}{s_1^2}c_3)=\nu(1,1)\}$.

The path $\bar{\mathbb{L}}_\delta$ is δ-close to $\bar{\mathbb{L}}_0$ which is the path on the level set $\alpha_{\bar{\mathcal{G}}_{3,0}}(c_1,c_3)=E$ connecting the point $(-c_1^*,c_3^*)$ to $(-c_1^*,-c_3^*)$ symmetric around the c_1-axis. If we choose $c_3=c_3^*$, we get a unique point on $\bar{\mathbb{L}}_\delta$ near $-c_1^*$. In the full system $\mathsf{G}_{3,\delta}$, adding back the c_2 variable by Proposition 7.8(2), we obtain a two-dimensional channel $\mathbb{L}_\delta$ in $\alpha^{-1}_{\mathsf{G}_{3,\delta}}(E)$. By the definition of the point $\tilde{c}^\sharp$ in the proof of the previous lemma, we get the point $(\tilde{c}^\sharp,c_3^*)\in\Gamma_\delta(c_3^*)\cap\mathbb{L}_\delta$. This shows that $\mathbb{L}_\delta\cap\Gamma_\delta(c_3^*)\neq\emptyset$ and $\mathbb{L}_\delta\cap\mathbb{C}_-\neq\emptyset$.

We claim that *for $\delta\bar{V}$ chosen in an open and dense subset $\hat{\mathcal{O}}_{3,*}$ of the $\tilde{\delta}_3$-ball of $C^r(\mathbb{T}^3)/C^r(\mathbb{T}^2)$, any continuous curve in the interior of $\mathbb{L}_\delta$ is a generalized transition chain.*

We introduce a subset $\Delta\subset\mathbb{L}_\delta$ in the following way: $\sigma\in\Delta$ if and only if the weak KAM u_σ of the restricted system $\bar{\mathsf{G}}_{3,\delta}$ on the NHIM is C^1 (must also be $C^{1,1}$ due to [5]), i.e., the Mañé set is an invariant 2-torus. For $\sigma\notin\Delta$, a certain section Σ_σ of 2-torus exists such that $\mathcal{N}_\sigma\cap\Sigma_\sigma$ is shrinkable so that Definition 2.1(H2) can be verified. To prove that $\mathbb{L}_\delta$ is a generalized transition chain, it remains to prove the following in order to verify Definition 2.1(H1):

For $\delta\bar{V}$ in $\tilde{\mathcal{O}}_{2,\star}$ and for all $\sigma\in\Delta$, each connected component of $\mathrm{Argmin}\{B_\sigma,\Sigma_{0,\sigma}\backslash\bigcup_m N_m\}$ *is contained in a certain disk $O_m\subset\Sigma_{0,\sigma}$,*

where B_σ is the barrier function of the system $\mathsf{G}_{3,\delta}$, $\Sigma_{0,\sigma}$ is a 2-dimensional section of $\mathbb{T}^3$ which is transversal to σ-semi-static curves, $\bigcup_m N_m$ denotes a neighborhood of the Aubry set in the finite covering space, and $\mathrm{Argmin}\{B_\sigma,\Sigma_0\backslash\bigcup_m N_m\}$ denotes the set of minimal points of B_σ which fall into the set $\Sigma_{0,\sigma}\backslash\bigcup_m N_m$.

This is given by Theorem E.2 in Appendix E (the autonomous case and the Mañé perturbation case). The argument is similar to the case of *a priori* unstable systems (see [17,18]). □

In the next remark, we explain our mechanism of ladder climbing.

Remark 10.7 (The mechanism for the ladder climbing). Here, we have employed a variant of Arnold's example (1.2). Consider an autonomous Hamiltonian system of three degrees of freedom of the form

$$H=\frac{y_1^2}{2}+\frac{y_2^2}{2}+\frac{y_3^2}{2}+(\cos x_3-1)(1+\varepsilon(\cos x_1+\sin x_2)).$$

In this system, there exists a diffusion orbit for each $E > 0$ such that (y_1, y_2) stays close to the circle $\{y_1^2 + y_2^2 = 2E\}$ and $\arctan\frac{y_1}{y_2}$ achieves any value in $[0, 2\pi)$, which can be proved by computing the Melnikov integral as in the Arnold's example. In our case, the system $\mathsf{G}_{3,\delta}$ plays the role of H here and the system $\bar{\mathsf{G}}_{3,\delta}$ plays the role of $y_1^2 + y_2^2$.

11 The proof of Theorem 2.4 for $n = 4$

We are now ready to give the proof of Theorem 2.4 in the case of $n = 4$ along one frequency segment, i.e., $\omega(t)$ in Theorem 2.4 lies within $B_\mu(\omega_-)$. To prove Theorem 2.4 in full generality, we need the following:

(1) transition from one frequency segment to the next;

(2) the $n > 4$ case.

(1) will be addressed in Section 13 and (2) will be addressed in Appendix C. Both are algorithmic but not dynamical, based on the proof in this section.

The general case is very similar and we postpone the induction argument to Appendix C, where the proof of Theorem 2.4 for general $n > 3$ is completed in Appendix C.6.

Proof of Theorem 2.4 *for* $n = 4$ *along one frequency segment.* Let us first review the results that we have obtained up to now. We start from the frequency line segment ω_- of the form (4.2), along which we treat two different regimes separately (the way of distinguishing the two cases was given in Subsection 6.1):

(a) single or weak double resonances;

(b) strong double resonances.

In the case (a), we apply Proposition 5.2 to yield an NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$ restricted to which the Hamiltonian system has one fewer degree of freedom. In the case (b), we obtain the same result in Propositions 6.9 and 6.10 up to a $\lambda\sqrt{\varepsilon}$-neighborhood of the strong double resonance.

We next modify slightly the frequency line segment to $\bar{\omega}_-$ in (8.1), along which we perform the second step of the order reduction. For the reduced system after the previous step, we again distinguish the above two cases. For the case (a), similar to Proposition 5.2, we obtain the NHIC homeomorphic to $T^*\mathbb{T}^{n-2}$ and for the case (b), which is now the triple resonance case, we obtain the same result in Proposition 9.3. This completes the proof of the parts (1), (2) and (3)(a) of Theorem 2.4 for $n = 4$.

We next prove the statements (3)(b) and (3)(c) on the generalized transition chain. The generalized transition chain of type Definition 2.1(H2) connecting nearby NHICs crossing the triple resonance is proved in Proposition 10.4, and hence (3)(c) is proved. Item (3)(b) can be reduced to the *a priori* unstable systems studied in [17, 18] due to the presence of NHICs. Indeed, we perform the standard energetic reduction procedure to reduce the autonomous system of n degrees of freedom to a nonautonomous one with $n - 1/2$ degrees of freedom as follows. We consider the Hamiltonian (9.1) in the region that is the $\Lambda\sqrt{\varepsilon}$-near triple resonance noting that the first three entries of the frequency vector $\boldsymbol{\omega}$ vanish due to the triple resonance. We fix the energy level E and solve the equation $H(x, Y) = E$ for $\mathcal{Y} := \boldsymbol{\omega}_4 Y_4/\sqrt{\varepsilon}$ (let $\boldsymbol{\omega}_4 = \hat{\boldsymbol{\omega}}_{n-3}$) and treat $\mathcal{Y}$ as the new Hamiltonian and $x_4\sqrt{\varepsilon}/\boldsymbol{\omega}_4 =: \tau$ as the new time to yield a system of the following form:

$$\mathcal{Y} = \frac{1}{2}\langle A_3\tilde{Y}_3, \tilde{Y}_3\rangle + V(\tilde{x}_2) + \delta\bar{V}(\tilde{x}_3) + \bar{\delta}\hat{R}\bigg(\tilde{x}_3, \tilde{Y}_3, \frac{\boldsymbol{\omega}_4\tau}{\sqrt{\varepsilon}}\bigg).$$

Outside the $\Lambda\sqrt{\varepsilon}$-neighborhood of the triple resonance, we obtain

$$\mathcal{Y} = \frac{\boldsymbol{\omega}_1 Y_1}{\sqrt{\varepsilon}} + \frac{1}{2}\langle A_3\tilde{Y}_3, \tilde{Y}_3\rangle + V(\tilde{x}_2) + \delta\bar{V}(\tilde{x}_3) + \bar{\delta}\hat{R}\bigg(\tilde{x}_3, \tilde{Y}_3, \frac{\boldsymbol{\omega}_4\tau}{\sqrt{\varepsilon}}\bigg).$$

The system admits NHICs restricted to which the time-1 map of the system is a twist map (this follows from Lemma 9.2 for the former and in addition Proposition 6.8 for the latter). Thus, they can be considered as *a priori* unstable systems studied in [17, 18]. The generalized transition chain satisfying Definition 2.1(H1) can be constructed for generic perturbations P (see Theorem E.2 in the appendix, and note that Theorem E.2 is applied to the original system (1.1) without applying the normal form). This gives Theorem 2.4(3)(b) for $n = 4$. We have considered only one frequency segment. The switch from one frequency segment to the next is the same as crossing a complete resonance and hence is the same as (3)(c) (see Appendix C.5).

Finally, we consider the genericity condition. We need a genericity assumption in the proof which can be classified into two classes: (HT) hyperbolicity type and (TT) transversality type. The conditions (HT) are used to guarantee the existence of NHICs and (TT) are used to guarantee that the transverse intersection of stable and unstable manifolds of Aubry sets (or destruction of Mañé sets), i.e., Definition 2.1(H1)(ii) as well as (H2). In Definition 1.1 of the cusp-residual set, (HT) is responsible for the "cusp" part, i.e., existence of $\mathfrak{R} \subset \mathfrak{S}^r$ and a_P for each $P \in \mathfrak{R}$, and (TT) is responsible for the residual part (see Theorem E.2). More explicitly, we have

(HT) construction of the NHICs (see Propositions 5.2, 6.9 and 6.10), whose genericity comes from Proposition 5.4 and Theorem 6.6;

(TT) (a) the c-equivalence path around the strong double resonances (see Proposition 7.7), whose genericity comes from Proposition 7.4;

(b) verification of Definition 2.1(H1)(ii), which is needed in the ladder construction of Proposition 10.4, as well as in the proof of Theorem 2.4(3)(b); the genericity is given by Theorem E.2.

With these, we verify the cusp-residual condition (see Definition 1.1).

First, we have finitely many open and dense conditions from the propositions in the above (HT). Denote by $\mathcal{O} \subset \mathfrak{B}^r$ the open and dense set obtained by taking the intersection of the finitely many open and dense sets. This P determines ε_P such that for $\varepsilon < \varepsilon_P$, (HT) holds. The fact that ε_P is continuous in P with the C^{r_0} norm follows from that a C^{r_0} small perturbation of P gives rise to a C^{r_0-5} small error term $\delta\mathsf{R}$ in the normal form, which does not destroy the NHICs.

Second, for this purpose of Item (TT), we fix an $\varepsilon < \varepsilon_P$ and apply Theorem E.2, which gives us an $\varepsilon' = \varepsilon'(\varepsilon P)$ and a residual set $\mathfrak{R}_{\varepsilon'}(\varepsilon P) \subset \varepsilon' \mathfrak{B}^r$ such that Definition 2.1(H1)(ii) is satisfied for $\varepsilon P + \varepsilon' P'$ for any $\varepsilon' P' \in \mathfrak{R}_{\varepsilon'}(\varepsilon P)$. Finally, apply Kuratowski-Ulam Theorem 8.5 to the set $\bigcup_{P\in\mathcal{O}} \bigcup_{\varepsilon<\varepsilon_P} \mathfrak{R}_{\varepsilon'}(\varepsilon P)$, which gives the stated form of the cusp-residual set in Definition 1.1. This completes the proof. □

12 The Riemannian metric perturbation of the flat torus

In this section, we show how to reduce Theorems 1.3 and 1.4 to Theorems 2.4 and 1.2.

For convenience, we apply the linear map $S : \mathbb{T}^N \to \mathbb{T}^N_S$ to converting the torus $\mathbb{T}^N_S$ to the standard torus $\mathbb{T}^N$, which pulls back the metric $ds^2 = \sum dx_i^2$ on $\mathbb{T}^N_S$ to $ds^2 = \sum c_{ij} dx_i dx_j$ on $\mathbb{T}^N$, where the matrix $C = (c_{ij}) = S^{\mathrm{T}} S$.

Proof of Theorem 1.4. This follows from Maupertuis's principle (see [2, Subsection 45.D]), which identifies the Hamiltonian flow of a mechanical Hamiltonian system $H(x, y) = \langle C^{-1} y, y\rangle + V(x)$ restricted to energy level 1 with the geodesic flow of the Riemannian metric $(1 - V(x)) \sum_{ij} c_{ij} dx_i dx_j$. Thus, Theorem 1.4 follows immediately from Theorem 1.2. □

We next work on the proof of Theorem 1.3.

Proof of Theorem 1.3. We consider the Riemannian metric on $\mathbb{T}^N$ of the form

$$ds_\varepsilon^2 = \sum_{i,j} (c_{ij} + \varepsilon d_{ij}(x)) dx_i dx_j.$$

We treat ds_ε^2 as twice of the Lagrangian and perform a Legendre transformation to get the Hamiltonian of the form

$$H_\varepsilon(x, y) = \frac{1}{2} \sum_{i,j} (a_{ij} + \varepsilon b_{ij}(x)) y_i y_j = \frac{1}{2}\langle Ay, y\rangle + \frac{\varepsilon}{2}\langle B(x)y, y\rangle =: h(y) + \varepsilon P(x, y), \tag{12.1}$$

where we define $A = (a_{ij}) = C^{-1}$ and $B(x) = (b_{ij}(x))$, and the matrix $A + \varepsilon B(x)$ is the inverse of $C + \varepsilon\Delta(x)$. Thus solving the equation

$$(A + \varepsilon B)(C + \varepsilon\Delta(x)) = \mathrm{Id},$$

we get $B(x) = -C^{-1}\Delta(x)C^{-1} + O(\varepsilon)$.

Thus H_ε has the form of the nearly integrable system with $h(y) = \frac{1}{2}\langle Ay, y\rangle$ and $P(x, y) = \frac{1}{2}\langle B(x)y, y\rangle$. The frequency map $\boldsymbol{\omega}(y) = \partial_y h(y) = Ay$.

We next pick $y^\star$ and perform a $\sqrt{\varepsilon}$-blowup in $B_{\Lambda\sqrt{\varepsilon}}(y^\star)$ for some Λ large, as we did in Subsection 3.1. Corresponding to (3.2),

$$\mathsf{H}(x, Y) = \varepsilon^{-1/2}\langle\omega^\star, Y\rangle + \frac{1}{2}\langle AY, Y\rangle + \mathsf{V}(x) + \mathsf{P}(x, \sqrt{\varepsilon}Y), \tag{12.2}$$

where $\omega^\star = Ay^\star$, $\mathsf{V}(x) = \frac{1}{2}\langle B(x)y^\star, y^\star\rangle$ and $\mathsf{P}(x, \sqrt{\varepsilon}Y) = \sqrt{\varepsilon}\langle B(x)y^\star, Y\rangle + \frac{1}{2}\varepsilon\langle B(x)Y, Y\rangle$, and the system is defined on $B_1(0) \times \mathbb{T}^N$.

The system (12.2) is formally the same as (3.2), and thus we can repeat the proof of Theorem 2.4 formally.

We next consider the genericity conditions. As we have analyzed in the last section, there are two types of such conditions: (HT) used to guarantee the existence of NHICs and (TT) used to destroy the Mañé sets (guarantee Definition 2.1(H1)). One way to verify the cusp-residual condition in the Riemannian metric perturbation setting is to adapt Proposition 5.4 and Theorem E.2 to the present setting, which can be done by modifying the proofs of Proposition 5.4 and [13, Theorem 4.2]. However, there is a simpler argument as follows.

We first note that both conditions are open conditions. Indeed, we first consider (TT), which is used to guarantee Definition 2.1(H1). Note that Definition 2.1(H1) is an open condition, since we cover Mañé sets by finitely many open balls. Once it is satisfied, by the upper-semi-continuity of Mañé sets, it remains to hold for C^r any small perturbation of $\varepsilon\Delta$. Thus, it is enough to get the denseness part, i.e., there is an arbitrarily small $\varepsilon'\Delta' \in \mathrm{Sym}^r(\mathbb{T}^n)$ such that $\varepsilon\Delta + \varepsilon'\Delta'$ satisfies Definition 2.1(H1). This follows from Theorem E.2 with the Mañé perturbation, since a Mañé perturbation $\varepsilon' V \in C^r(\mathbb{T}^n)$ to the Hamiltonian H_ε can be turned into a perturbation to the Riemannian metric by Maupertuis's principle. Next, (HT) is also open by the persistence of NHICs, and thus we first add an arbitrarily small Mañé perturbation to H_ε to guarantee the existence of NHICs as in Theorem 2.4, and then apply the same Maupertuis trick. This completes the proof. □

13 Construction of the global frequency path

In the previous sections, we have shown how to find a frequency path along which only the first entry can be arbitrarily moved along an orbit. In this section, we show how to repeat the strategy to move the second, third, ..., entries. In the frequency space, the picture is similar to moving along edges of a cube to connect different vertices.

13.1 The choice of the Diophantine path

To find the frequency path $\{\omega(t)\}$ as described above, we first find a guiding frequency path with certain Diophantine properties, which will be shadowed by $\{\omega(t)\}$.

Lemma 13.1. *Given any $\varrho > 0$, $\tau > n$ and any finitely many frequency vectors $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_M \in \partial h(h^{-1}(E))$, $E > \min h$ and $M > 1$, there exist a constant $\alpha > 0$ and vectors*

$$\boldsymbol{\omega}_i^* = (\omega_{i,1}^*, \ldots, \omega_{i,n}^*) \in \partial h(h^{-1}(E))$$

satisfying $|\boldsymbol{\omega}_i - \boldsymbol{\omega}_i^*| < \varrho$, $i = 1, \ldots, M$ and

$$\boldsymbol{\omega}_{i,[j]}^* := (\omega_{i+1,1}^*, \ldots, \omega_{i+1,j}^*, \omega_{i,j+1}^*, \ldots, \omega_{i,n}^*) \in \mathrm{DC}(n, \alpha, \tau)$$

for all $i = 1, \ldots, M-1$ and $j = 0, 1, 2, \ldots, n$.

The proof of this lemma is postponed to the end of this section. From the Diophantine vectors $\boldsymbol{\omega}^*_{i,[j]}$, we construct $n(M-1)$ frequency segments

$$\Omega_{i,[j]}(t)=\rho_{i,[j]}(t)(\omega^*_{i+1,1},\ldots,\omega^*_{i+1,j-1},t,\omega^*_{i,j+1},\ldots,\omega^*_{i,n}),\quad t\in[\omega^*_{i,j},\omega^*_{i+1,j}]\cup[\omega^*_{i+1,j},\omega^*_{i,j}],$$

$j=1,\ldots,n$, $i=1,\ldots,M-1$, where the scalar multiple $\rho_{i,[j]}(t)$ is determined by requiring that the segment $\Omega_{i,[j]}$ lie on $\partial h(h^{-1}(E))$. By the construction, the end point of $\Omega_{i,[j]}$ agrees with the starting point of $\Omega_{i,[j+1]}$ (for $j<n$) and the end point of $\Omega_{i,[n]}$ agrees with the starting point of $\Omega_{i+1,[1]}$ for all $i=1,\ldots,M-1$. So the segments are concatenated into a connected curve in $\partial h(h^{-1}(E))$ connecting $\boldsymbol{\omega}^*_1$ to $\boldsymbol{\omega}^*_M$ and passing by the points $\boldsymbol{\omega}^*_i,\ i=1,\ldots,M$.

13.2 Construction of the frequency segments

In the previous sections, we have been considering $\Omega_{1,[1]}$ and found in its ϱ-neighborhood a frequency vector of the form

$$\omega^{(2)}(a)=\rho^{(2)}_a\left(a,\frac{p_2}{q_2}\omega^{*i}_2,\frac{p_3}{q_3}\omega^{*i}_2,\frac{p_4}{q_4}\omega^{*i}_2,\omega^{*i}_5,\ldots,\omega^{*i}_n\right),\quad \omega^{*i}_k=\omega^{*i}_{1,k}$$

that admits two resonant integer vectors $\boldsymbol{k}'$ and $\boldsymbol{k}''$ for all a, and found NHICs $\mathcal{C}(\boldsymbol{k}',\boldsymbol{k}'')$ homeomorphic to $T^*\mathbb{T}^{n-2}$ away from triple resonances (see Proposition 9.3).

We next formulate the induction rule to perform the procedure inductively. Suppose that we have done ℓ steps with $\ell=1,\ldots,n-3$ and found in the ϱ-neighborhood of $\Omega_{1,[1]}$ a frequency vector of the form

$$\omega^{(\ell)}_a=\left(a,\frac{p_2}{q_2}\omega^*_2,\frac{p_3}{q_3}\omega^*_2,\ldots,\frac{p_{\ell+1}}{q_{\ell+1}}\omega^*_2,\frac{p_{\ell+2}}{q_{\ell+2}}\omega^*_2,\omega^*_{\ell+3},\ldots,\omega^*_n\right)$$

with resonant integer vectors $\boldsymbol{k}',\boldsymbol{k}'',\ldots,\boldsymbol{k}^{(\ell)}$ whose span is denoted by $\boldsymbol{K}^{(\ell)}$ and the Hamiltonian system admits NHICs $\mathcal{C}(\boldsymbol{K}^{(\ell)})$ homeomorphic to $T^*\mathbb{T}^{n-\ell}$ up to a $\lambda\sqrt{\varepsilon}$-neighborhood of strong $(\ell+1)$ resonances. We define the induction rule.

Definition 13.2 (The induction rule). Suppose that $\Pi_{\boldsymbol{K}^{(\ell)}}P(\cdot,y)$ has a nondegenerate global maximum for all $y\in\omega^{(\ell)}_-$. We modify $\omega^*_{\ell+3}$ into $\frac{p_{\ell+3}}{q_{\ell+3}}\omega^*_2$ such that $|q_{\ell+3}|$ is so large and the new resonant integer vector $\boldsymbol{k}^{(\ell+1)}$ has a so large norm that the term $\Pi_{\boldsymbol{K}^{(\ell+1)}}P(\cdot,y)-\Pi_{\boldsymbol{K}^{(\ell)}}P(\cdot,y)$ with the C^{r_0-2} norm less than $C/|\boldsymbol{k}^{(\ell+1)}|^2$ does not spoil the NHICs $\mathcal{C}(\boldsymbol{K}^{(\ell)})$.

Proposition 13.3. *For generic $P\subset\mathfrak{B}^r$, there exists an $\varepsilon_P>0$ such that for all $0<\varepsilon<\varepsilon_P$, the Hamiltonian system* (1.1) *admits a frequency path $\boldsymbol{\omega}^\sharp_{1,[1]}$ lying in a ϱ-neighborhood of $\Omega_{1,[1]}$ such that the induction rule is satisfied for all $n-2$ steps.*

The proposition can be proved by applying the refinement procedure in Section 8 repeatedly. More details will be provided in Appendix C.

Thus we get a frequency segment of the form

$$\boldsymbol{\omega}^\sharp_{1,[1]}(a)=\omega^{(n-2)}(a)=\rho^{(n-2)}_a\omega^{*i}_2\left(\frac{a}{\omega^{*i}_2},\frac{p_2}{q_2},\frac{p_3}{q_3},\ldots,\frac{p_n}{q_n}\right),$$

along which we have $n-2$ resonant integer vectors $\boldsymbol{k}',\boldsymbol{k}'',\ldots,\boldsymbol{k}^{(n-2)}$ for all a and $n-1$ for some a. We get NHICs $\mathcal{C}(\boldsymbol{k}',\boldsymbol{k}'',\ldots,\boldsymbol{k}^{(n-2)})$ homeomorphic to $T^*\mathbb{T}^2$ $\lambda\sqrt{\varepsilon}$-away from complete resonances.

13.3 The resonant frequency path shadowing all $\Omega_{i,[j]}$

We next find the resonant frequency path $\omega(t)$ in Theorem 2.4 shadowing all the segments $\bigcup_{i,j}\Omega_{i,[j]}$.

Applying the last proposition, we have constructed a frequency path $\boldsymbol{\omega}^\sharp_{1,[1]}$ shadowing $\Omega_{1,[1]}$ satisfying the induction rules. We next do the same thing to shadow $\Omega_{2,[1]}$. Note that $\Omega_{1,[1]}$ has the same entries as $\Omega_{2,[1]}$ except the first two, so should $\boldsymbol{\omega}^\sharp_{1,[1]}$ and $\boldsymbol{\omega}^\sharp_{2,[1]}$ do. Thus, we cannot do the induction procedures of $\Omega_{1,[1]}$ and $\Omega_{2,[1]}$ independently. Let us introduce the following definition.

Definition 13.4 (The connection rule)**.** The frequency paths $\boldsymbol{\omega}^\sharp_{i,[j]}$ and $\boldsymbol{\omega}^\sharp_{i+1,[j]}$ obtained from Proposition 13.3 shadowing $\Omega_{i,[j]}$ and $\Omega_{i+1,[j]}$, respectively, are said to satisfy the connection rule if their all except the i-th and $(i+1)$-th entries are identical, and it is similar for $(i,[j])$ and $(i+1,[j])$ replaced by $(n,[j])$ and $(1,[j+1])$, respectively.

Proposition 13.5. *For generic $P \in \mathfrak{B}^r$, there exists an $\varepsilon_P > 0$ such that for all $0 < \varepsilon < \varepsilon_P$, the Hamiltonian system* (1.1) *admits a continuous frequency path lying in a ϱ-neighborhood of all the segments $\bigcup_{i,j} \Omega_{i,[j]}$ and satisfying the induction rule and the connection rule for all i and j.*

Proof. Let $\boldsymbol{\omega}^*_i$ $(i=1,\ldots,M)$ be the frequencies and $\Omega_{i,[j]}$ $(j=1,\ldots,n$ and $i=1,\ldots,M-1)$ be the frequency segments defined in Subsection 13.1. In order to apply the last proposition, we always perform a permutation such that the varying entry becomes the first one. Up to permutations of entries and a scalar multiple, we can list the frequency segments as follows:

$$\begin{array}{lcccccccc}
\Omega_{1,[1]}(a) = & a & \omega^*_{1,2} & \omega^*_{1,3} & \cdots & \omega^*_{1,n} & & & \\
\Omega_{1,[2]}(a) = & & a & \omega^*_{1,3} & \cdots & \omega^*_{1,n} & \omega^*_{2,1} & & \\
\vdots & & & & \vdots & \vdots & \vdots & & \\
\Omega_{1,[n]}(a) = & & & & & a & \omega^*_{2,1} & \omega^*_{2,2} \ \cdots & \omega^*_{2,n-1} \\
\Omega_{2,[1]}(a) = & & & & & & a & \omega^*_{2,2} \ \cdots & \omega^*_{2,n-1} \ \ \omega^*_{2,n} \\
\vdots & & & & & & \vdots & \vdots &
\end{array} \tag{13.1}$$

The rules are as follows:

(1) In $\Omega_{i,[j]}(a)$, the j-th entry is $a \in [\omega^*_{i,[j]} - \varrho, \omega^*_{i+1,[j]} + \varrho]$. The entries with subscripts less than j coincide with those of $\boldsymbol{\omega}^*_{i+1}$ and the entries with subscripts greater than j coincide with those of $\boldsymbol{\omega}^*_i$.

(2) We permute entries of $\Omega_{i,[j]}$ in such a way that a is the leading entry and the entries with subscripts less than j are placed after its last entry.

(3) The vectors $\Omega_{i,[j]}$ $(i=1,\ldots,M-1$ and $j=1,\ldots,n)$ are arrayed in a parallelogram such that $\Omega_{i,[j]}$ is placed on the $((i-1)n+j)$-th row with the leading entry a placed at the $((i-1)n+j)$-th column.

We inductively refine the frequency segment such that after $Mn-3$ steps, all the above $\omega^*_{i,j}$ become a rational multiple of $\omega^*_{1,2}$. Denoting the resulting vector by $\boldsymbol{\omega}^\sharp_{i,[j]}(a)$, we have $|\boldsymbol{\omega}^\sharp_{i,[j]}(a) - \Omega_{i,[j]}(a)| < \varrho$. It holds that

$$\begin{array}{lcccccccc}
\boldsymbol{\omega}^\sharp_{1,[1]}(a) = & \frac{a}{\omega^*_{1,2}} & \frac{p_2}{q_2} & \frac{p_3}{q_3} & \cdots & \frac{p_n}{q_n} & & & \\
\boldsymbol{\omega}^\sharp_{1,[2]}(a) = & & \frac{a}{\omega^*_{1,2}} & \frac{p_3}{q_3} & \cdots & \frac{p_n}{q_n} & \frac{p_{n+1}}{q_{n+1}} & & \\
\vdots & & & & \vdots & \vdots & \vdots & & \\
\boldsymbol{\omega}^\sharp_{1,[n]}(a) = & & & & & \frac{a}{\omega^*_{1,2}} & \frac{p_{n+1}}{q_{n+1}} & \frac{p_{n+2}}{q_{n+2}} \ \cdots & \frac{p_{2n-1}}{q_{2n-1}} \\
\boldsymbol{\omega}^\sharp_{2,[1]}(a) = & & & & & & \frac{a}{\omega^*_{1,2}} & \frac{p_{n+2}}{q_{n+2}} \ \cdots & \frac{p_{2n-1}}{q_{2n-1}} \ \ \frac{p_{2n}}{q_{2n}} \\
\vdots & & & & & & \vdots & \vdots &
\end{array}$$

The frequency refinements are done inductively as follows. We introduce the superscript (ℓ) with $1 \leqslant \ell \leqslant Mn-3$ counting the step of refinements. During the ℓ-th step of the order reduction, we modify the Diophantine number in the $(\ell+2)$-th column into $\frac{p_{\ell+1}}{q_{\ell+1}}\omega^*_{1,2}$, where the number $\frac{p_{\ell+1}}{q_{\ell+1}}$ is to be determined.

Notation 8. (1) For each $1 \leqslant \ell \leqslant Mn-3$, there is an index set $\mathcal{I}(\ell)$ such that for each $(i,[j]) \in \mathcal{I}(\ell)$, $\Omega_{i,[j]}$ intersects the $(\ell+2)$-th column of the table (13.1) not at the a entry. If $(i,[j]) \in \mathcal{I}(\ell)$, we denote the frequency vector by $\omega^{(\ell)}_{i,[j]}$.

(2) If the frequency vector with the subscript $(i,[j])$ lies entirely to the left of the $(\ell+2)$-th column without intersecting it at the step (ℓ), this vector has completed its refinement and has all the entries being rational multiples of $\omega^*_{1,2}$ except the leading a, so we denote it by $\boldsymbol{\omega}^\sharp_{i,[j]}$.

(3) The frequency vector $\Omega_{i,[j]}$ lying to the right of the $(\ell+2)$-th column without intersecting it will maintain its notation $\Omega_{i,[j]}$.

For example, in the case of $\ell \leqslant n-2$, we have that $\omega^{(\ell)}_{1,[1]}(a), \omega^{(\ell)}_{1,[2]}(a), \ldots, \omega^{(\ell)}_{1,[\ell]}(a)$ are the following, respectively:

$$\begin{array}{ccccccccccc} a & \frac{p_{1,2}}{q_{1,2}}\omega^*_{1,2} & \frac{p_{1,3}}{q_{1,3}}\omega^*_{1,2} & \cdots & \frac{p_{1,\ell+1}}{q_{1,\ell+1}}\omega^*_{1,2} & \frac{p_{1,\ell+2}}{q_{1,\ell+2}}\omega^*_{1,2} & \omega^*_{1,\ell+3} & \cdots & \omega^*_{1,n} & & \\ & a & \frac{p_{1,3}}{q_{1,3}}\omega^*_{1,2} & \cdots & \frac{p_{1,\ell+1}}{q_{1,\ell+1}}\omega^*_{1,2} & \frac{p_{1,\ell+2}}{q_{1,\ell+2}}\omega^*_{1,2} & \omega^*_{1,\ell+3} & \cdots & \omega^*_{1,n} & \omega^*_{2,1} & \\ & & \vdots & & & \vdots & & \vdots & & & \\ & & & & a & \frac{p_{1,\ell+2}}{q_{1,\ell+2}}\omega^*_{1,2} & \omega^*_{1,\ell+3} & \cdots & \omega^*_{1,n} & \omega^*_{2,1} & \cdots\ \omega^*_{2,\ell-3}. \end{array}$$

The proposition is then proved by repeating this procedure. $\square$

Notation 9. To simplify the notations and for clarity, instead of using the double subscripts $(i,[j]) \in \mathcal{I}(\ell)$, we introduce a single subscript $\kappa = 0, 1, \ldots, \sharp\mathcal{I}(\ell)-1$ ($\kappa \leqslant n-2$) counting the number of independent irreducible resonant integer vectors for each $\omega^{(\ell)}_\kappa(a)$ for all a.

13.4 The proof of Lemma 13.1

In this subsection, we give the proof of the number-theoretic Lemma 13.1.

Proof of Lemma 13.1. Fix $\varrho > 0$ and $\tau > n$. We prove the lemma by induction from $j+1$ to j. First, for $j = n$, it is easy to find two Diophantine numbers ω^i_n and ω^f_n. Suppose that we already have

$$(\omega^{*i}_{j+1}, \ldots, \omega^{*i}_n) \in \mathrm{DC}(n-j, \alpha, \tau).$$

We claim that *given ω^i_j and ω^f_j, there are numbers ω^{*i}_j and ω^{*f}_j satisfying*

$$|\omega^{*i}_j - \omega^i_j| < \varrho, \quad |\omega^{*f}_j - \omega^f_j| < \varrho \quad \text{and} \quad (\omega^{*i,f}_j, \omega^{*i}_{j+1}, \ldots, \omega^{*i}_n) \in \mathrm{DC}(n-j+1, \alpha, \tau)$$

for sufficiently small $\alpha > 0$.

Indeed, by the assumption, we already have

$$|\langle \hat{\omega}^{*i}_{n-j}, \hat{\boldsymbol{k}}_{n-j}\rangle| \geqslant \frac{\alpha}{|\hat{\boldsymbol{k}}_{n-j}|^\tau}, \quad \forall\, \hat{\boldsymbol{k}}_{n-j} \in \mathbb{Z}^{n-j} \setminus \{0\}.$$

We want to show that all those $\omega_j \in \mathbb{R}$ which satisfy the condition

$$|\langle (\omega_j, \hat{\omega}^{*i}_{n-j}), \hat{\boldsymbol{k}}_{n-j+1}\rangle| \geqslant \frac{\alpha}{|\hat{\boldsymbol{k}}_{n-j+1}|^\tau}, \quad \forall\, \hat{\boldsymbol{k}}_{n-j+1} \in \mathbb{Z}^{n-j+1} \setminus \{0\} \tag{13.2}$$

form a ϱ-dense set provided that α is small enough. Given $\hat{\boldsymbol{k}}_{n-j}$, we consider all k_j and $\omega^\dagger_j$ satisfying

$$k_j \omega^\dagger_j + \langle \hat{\omega}^{*i}_{n-j}, \hat{\boldsymbol{k}}_{n-j}\rangle = 0.$$

Formula (13.2) is satisfied automatically for $\hat{\boldsymbol{k}}_{n-j+1} = (k_j, \hat{\boldsymbol{k}}_{n-j})$ when $k_j = 0$, so we assume $k_j \neq 0$. In order to guarantee (13.2), we need to remove an interval of measure $\frac{2\alpha}{k_j(|\hat{\boldsymbol{k}}_{n-j}|+|k_j|)^\tau}$ centered at $\omega^\dagger_j$ so that (13.2) is satisfied for all ω_j in the complement for this k_j. The total measure of these intervals when k_j ranges over $\mathbb{Z} \setminus \{0\}$ is

$$\sum_{k_j} \frac{2\alpha}{|k_j|(|\hat{\boldsymbol{k}}_{n-j}| + |k_j|)^\tau} \leqslant 2\int_1^\infty \frac{2\alpha}{x(|\hat{\boldsymbol{k}}_{n-j}| + x)^\tau}\, dx.$$

Next, the total measure of these intervals when $\hat{\boldsymbol{k}}_{n-j}$ ranges over $\mathbb{Z}^{n-j} \setminus \{0\}$ is

$$\begin{aligned} \sum_{\hat{\boldsymbol{k}}_{n-j}} \sum_{k_j} \frac{4\alpha}{|k_j|(|\hat{\boldsymbol{k}}_{n-j}| + |k_j|)^\tau} &\leqslant \sum_{\hat{\boldsymbol{k}}_{n-j}} \int_1^\infty \frac{4\alpha}{x(|\hat{\boldsymbol{k}}_{n-j}| + x)^\tau} dx \\ &\leqslant \int_{\mathbb{S}^{n-j-1}} \int_1^\infty \int_1^\infty \frac{4\alpha}{x(r+x)^\tau} dx r^{n-j-1} dr d\mathbb{S}^{n-j-1} \\ &\overset{y=x/r}{=\!=\!=} 4\alpha C \int_1^\infty r^{n-j-\tau-1} \int_{1/r}^\infty \frac{1}{y(1+y)^\tau} dy dr, \end{aligned}$$

where the constant $C = \frac{2\pi^{(n-j-1)/2}}{\Gamma((n-j-1)/2)}$ is the area of the sphere $\mathbb{S}^{n-j-1}$. The inner integral converges for large y and has the asymptote $\log r$ for r large and y close to $1/r$. Hence the iterated integral can be estimated as

$$\int_1^\infty r^{n-j-\tau-1}\int_{1/r}^\infty \frac{1}{y(1+y)^\tau}dydr \leqslant 2\int_1^\infty r^{n-j-\tau-1}(\log r+\text{const})dr,$$

where the right-hand side is convergent since $\tau > n$. The assertion above is proven if $\alpha > 0$ is chosen small enough. □

Acknowledgements The first author was supported by National Natural Science Foundation of China (Grant Nos. 11790272 and 11631006). The second author was supported by National Natural Science Foundation of China (Grant Nos. 11790273 and 12271285) and the Xiaomi endowed professorship of Tsinghua University.

References

1 Arnold V I. Instability of dynamical systems with several degrees of freedom. Soviet Math Dokl, 1964, 5: 581–585
2 Arnold V I. Mathematical Methods of Classical Mechanics. New York: Springer-Verlag, 1989
3 Arnold V I. Mathematical problems in classical physics. In: Trends and Perspectives in Applied Mathematics. Applied Mathematical Sciences, vol. 100. New York: Springer, 1994, 1–20
4 Arnold V I. The stability problem and ergodic properties for classical dynamical systems. In: Vladimir I. Arnold—Collected Works, Volume II. Berlin: Springer-Verlag, 2014, 107–113
5 Bernard P. Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds. Ann Sci Éc Norm Supér (4), 2007, 40: 445–452
6 Bernard P. Symplectic aspects of Mather theory. Duke Math J, 2007, 136: 401–420
7 Bernard P. The dynamics of pseudographs in convex Hamiltonian systems. J Amer Math Soc, 2008, 21: 615–669
8 Bernard P. Arnold's diffusion: From the *a priori* unstable to the *a priori* stable case. In: Proceedings of the International Congress of Mathematicians, vol. 3. New Delhi: Hindustan Book Agency, 2010, 1680–1700
9 Bernard P, Contreras G. A generic property of families of Lagrangian systems. Ann of Math (2), 2008, 167: 1099–1108
10 Bernard P, Kaloshin V, Zhang K. Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders. Acta Math, 2016, 217: 1–79
11 Birkhoff G D. Dynamical Systems. Providence: Amer Math Soc, 1960
12 Carneiro M J D. On minimizing measures of the action of autonomous Lagrangians. Nonlinearity, 1995, 8: 1077–1085
13 Cheng C-Q. Dynamics around the double resonance. Camb J Math, 2017, 5: 153–228
14 Cheng C-Q. Uniform hyperbolicity of invariant cylinder. J Differential Geom, 2017, 106: 1–43
15 Cheng C-Q. The genericity of Arnold diffusion in nearly integrable Hamiltonian systems. Asian J Math, 2019, 23: 401–438
16 Cheng C-Q, Xue J X. Order property and modulus of continuity of weak KAM solutions. Calc Var Partial Differential Equations, 2018, 57: 65
17 Cheng C-Q, Yan J. Existence of diffusion orbits in *a priori* unstable Hamiltonian systems. J Differential Geom, 2004, 67: 457–517
18 Cheng C-Q, Yan J. Arnold diffusion in Hamiltonian systems: *A priori* unstable case. J Differential Geom, 2009, 82: 229–277
19 Cheng C-Q, Zhou M. Non-degeneracy of extremal points in multi-dimensional space. Sci China Math, 2015, 58: 2255–2260
20 Cheng C-Q, Zhou M. Hyperbolicity of minimal periodic orbits. Math Res Lett, 2016, 23: 685–705
21 Delshams A, de la Llave R, Seara T M. A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of $\mathbb{T}^2$. Comm Math Phys, 2000, 209: 353–392
22 Delshams A, de la Llave R, Seara T M. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: Heuristics and rigorous verification on a model. Mem Amer Math Soc, 2006, 179: 1–141
23 Delshams A, de la Llave R, Seara T M. Geometric properties of the scattering map of a normally hyperbolic invariant manifold. Adv Math, 2008, 217: 1096–1153
24 Delshams A, de la Llave R, Seara T M. Instability of high dimensional Hamiltonian systems: Multiple resonances do not impede diffusion. Adv Math, 2016, 294: 689–755
25 Eliasson L H. Normal forms for Hamiltonian systems with Poisson commuting integrals—elliptic case. Comment Math Helv, 1990, 65: 4–35
26 Fathi A. Weak KAM theorem in Lagrangian dynamics, preliminary version number 10. Https://www.math.u-bordeaux.fr/~pthieull/Recherche/KamFaible/Publications/Fathi2008_01.pdf
27 Fenichel N. Persistence and smoothness of invariant manifolds for flows. Indiana Univ Math J, 1972, 21: 193–226

28 Gidea M, de la Llave R, Seara T M. A general mechanism of diffusion in Hamiltonian systems: Qualitative results. Comm Pure Appl Math, 2020, 73: 150–209
29 Herman M. Some open problems in dynamical systems. In: Proceedings of the International Congress of Mathematicians, vol. 2. Bielefeld: Doc Math, 1998, 797–808
30 Hirsch M W, Pugh C C, Shub M. Invariant Manifolds. Lecture Notes in Mathematics, vol. 583. Cham: Springer-Verlag, 1977
31 Hitchin N. Stable bundles and integrable systems. Duke Math J, 1987, 54: 91–114
32 Kaloshin V, Zhang K. A Strong Form of Arnold Diffusion for Two and A Half Degrees of Freedom. Princeton: Princeton University Press, 2021
33 Li X. On c-equivalence. Sci China Ser A, 2009, 52: 2389–2396
34 Li X, Cheng C-Q. Connecting orbits of autonomous Lagrangian systems. Nonlinearity, 2009, 23: 119–141
35 Mañé R. Generic properties and problems of minimizing measures of Lagrangian systems. Nonlinearity, 1996, 9: 273–310
36 Massart D. On Aubry sets and Mather's action functional. Israel J Math, 2003, 134: 157–171
37 Mather J N. Action minimizing invariant measures for positive definite Lagrangian systems. Math Z, 1991, 207: 169–207
38 Mather J N. Variational construction of connecting orbits. Ann Inst Fourier (Grenoble), 1993, 43: 1349–1386
39 Mather J N. Arnold diffusion. I: Announcement of results. J Math Sci (NY), 2004, 124: 5275–5289
40 Mather J N. Shortest curves associated to a degenerate Jacobi metric on $\mathbb{T}^2$. In: Progress in Variational Methods. Nankai Series in Pure, Appled Mathematics and Theoretical Physics, 7. Hackensack: World Scientific, 2011, 126–168
41 Oxtoby J C. Measure and Category: A Survey of the Analogies between Topological and Measure Spaces. Graduate Texts in Mathematics, vol. 2. New York-Heidelberg-Berlin: Springer-Verlag, 2013
42 Treschev D. Evolution of slow variables in *a priori* unstable Hamiltonian systems. Nonlinearity, 2004, 17: 1803–1841
43 Treschev D. Arnold diffusion far from strong resonances in multidimensional *a priori* unstable Hamiltonian systems. Nonlinearity, 2012, 25: 2717–2757

Appendix A A brief introduction to the variational theory

In this appendix, we give a brief introduction to the Mather theory and the weak KAM theory.

Appendix A.1 Minimizing measures and α- and β-functions

The theory is established for the Tonelli Lagrangian (see [38]).

Definition A.1 (Tonelli Lagrangian)**.** Let M be a closed manifold. A C^2-function L: $TM\times\mathbb{T}\to\mathbb{R}$ is called *Tonelli Lagrangian* if it satisfies the following conditions:

(1) *Positive definiteness.* For each $(x,t)\in M\times\mathbb{T}$, the Lagrangian function is strictly convex in velocity: the Hessian $\partial^2_{\dot x\dot x}L$ is positive definite.

(2) *Superlinear growth.* We assume that L has fiber-wise superlinear growth: for each $(x,t)\in M\times\mathbb{T}$, we have $L/\|\dot x\|\to\infty$ as $\|\dot x\|\to\infty$.

(3) *Completeness.* All the solutions of the Lagrangian equations are well defined for the whole $t\in\mathbb{R}$.

For autonomous systems, the completeness is automatically satisfied, as each orbit entirely stays in a certain compact energy level set.

Given a closed 1-form $\langle\eta_c(x),dx\rangle$ with the first cohomology class $[\langle\eta_c(x),dx\rangle]=c$, we introduce a Lagrange multiplier $\eta_c=\langle\eta_c(x),\dot x\rangle$. Without danger of confusion, we also call it a closed 1-form.

For each C^1 curve γ: $\mathbb{R}\to M$ with period k, there is a unique probability measure μ_γ on $TM\times\mathbb{T}$ so that the following holds:

$$\int_{TM\times\mathbb{T}} f d\mu_\gamma=\frac{1}{k}\int_0^k f(d\gamma(s),s)ds$$

for each $f\in C^0(TM\times\mathbb{T},\mathbb{R})$, where we use the notation $d\gamma=(\gamma,\dot\gamma)$. Let

$$\mathfrak{H}^*=\{\mu_\gamma\mid\gamma\in C^1(\mathbb{R},M)\text{ is periodic with the period }k\in\mathbb{N}\}.$$

The set $\mathfrak{H}$ of *holonomic probability measures* is the closure of $\mathfrak{H}^*$ in the vector space of continuous linear functionals. One sees that $\mathfrak{H}$ is convex.

For each $\nu \in \mathfrak{H}$, the action $A_c(\nu)$ is defined as $A_c(\nu) = \int (L - \eta_c) d\nu$. It is proved in [35, 37] that for each cohomology class c, there exists at least one invariant probability measure μ_c minimizing the action over $\mathfrak{H}$:

$$A_c(\mu_c) = \inf_{\nu \in \mathfrak{H}} \int (L - \eta_c) d\nu,$$

which is called a *c-minimal measure*.

Definition A.2 (The Mather set). (1) Let $\mathfrak{H}_c \subset \mathfrak{H}$ be the set of c-minimal measures. The *Mather set* $\tilde{\mathcal{M}}(c)$ is defined as $\tilde{\mathcal{M}}(c) = \bigcup_{\mu_c \in \mathfrak{H}_c} \mathrm{supp}\mu_c$.

(2) The *α-function* is defined as $\alpha(c) = -A_c(\mu_c) : H^1(M, \mathbb{R}) \to \mathbb{R}$. It is convex and finite everywhere with superlinear growth.

(3) Its Legendre transformation $\beta : H_1(M, \mathbb{R}) \to \mathbb{R}$ is called a *β-function*

$$\beta(\omega) = \max_c (\langle \omega, c \rangle - \alpha(c)).$$

It is also convex and finite everywhere with superlinear growth (see [37]).

Note that $\int \lambda d\mu_\gamma = 0$ holds for each exact 1-form λ and each $\mu_\gamma \in \mathfrak{H}^*$.

Definition A.3 (The rotation vector). For each measure $\mu \in \mathfrak{H}$, one can define its *rotation vector* $\omega(\mu) \in H_1(M, \mathbb{R})$ such that $\langle [\lambda], \omega(\mu) \rangle = \int \lambda d\mu$ holds for every closed 1-form λ on M.

We have the relation $c \in \partial\beta(\rho) \Leftrightarrow \alpha(c) + \beta(\rho) = \langle c, \rho \rangle$.

Appendix A.2 (Semi)-static curves, the Aubry set and the Mañé set

The concept of semi-static curves is introduced by Mather [38] and Mañé [35].

Definition A.4 (The c-semi-static curve). A curve γ: $\mathbb{R} \to M$ is called *c-semi-static* if

(1) in the time-1-periodic case, we have

$$[A_c(\gamma)\,|_{[t,t']}] = F_c((\gamma(t), t), (\gamma(t'), t')),$$

where

$$\begin{aligned}
[A_c(\gamma)\,|_{[t,t']}] &= \int_t^{t'} (L(d\gamma(t), t) - \eta_c(d\gamma(t)))dt + \alpha(c)(t' - t), \\
F_c((x, t), (x', t')) &= \inf_{\substack{\tau = t \bmod 1 \\ \tau' = t' \bmod 1}} h_c((x, \tau), (x', \tau')), \\
h_c((x, \tau), (x', \tau')) &= \inf_{\substack{\xi \in C^1 \\ \xi(\tau) = x \\ \xi(\tau') = x'}} [A_c(\xi)\,|_{[\tau,\tau']}];
\end{aligned}$$

(2) in the autonomous case, the semi-static curve is defined as

$$[A_c(\gamma)\,|_{(t,t')}] = F_c(\gamma(t), \gamma(t')), \quad \text{where } F_c(x, x') = \inf_{\tau > 0} h_c((x, 0), (x', \tau)).$$

Definition A.5 (The c-static curve). A semi-static curve $\gamma \in C^1(\mathbb{R}, M)$ is called *c-static* if, in addition, the relation $[A_c(\gamma)\,|_{(t,t')}] = -F_c((\gamma(t'), \tau'), (\gamma(t), \tau))$ holds in the time-1-periodic case and $[A_c(\gamma)\,|_{(t,t')}] = -F_c(\gamma(t'), \gamma(t))$ holds in the autonomous case.

Definition A.6 (The Mañé set and the Aubry set). We call the *Mañé set* $\tilde{\mathcal{N}}(c)$ the union of c-semi-static orbits

$$\tilde{\mathcal{N}}(c) = \bigcup \{d\gamma : \gamma \text{ is } c\text{-semi-static}\}$$

and call the *Aubry set* $\tilde{\mathcal{A}}(c)$ the union of c-static orbits

$$\tilde{\mathcal{A}}(c) = \bigcup \{d\gamma : \gamma \text{ is } c\text{-static}\}.$$

Notation 10. We use $\mathcal{M}(c)$, $\mathcal{A}(c)$ and $\mathcal{N}(c)$ to denote the standard projections of $\tilde{\mathcal{M}}(c)$, $\tilde{\mathcal{A}}(c)$ and $\tilde{\mathcal{N}}(c)$ from $TM \times \mathbb{T}$ to $M \times \mathbb{T}$, respectively.

They satisfy the inclusion relations $\tilde{\mathcal{M}}(c) \subseteq \tilde{\mathcal{A}}(c) \subseteq \tilde{\mathcal{N}}(c)$. It is showed in [37, 38] that the inverse of the projection is Lipschitz when it is restricted to $\mathcal{A}(c)$ as well as to $\mathcal{M}(c)$. By adding the subscript s to $\mathcal{N}$, i.e., $\mathcal{N}_s$, we define its time-s-section. This principle also applies to $\tilde{\mathcal{N}}(c)$, $\tilde{\mathcal{A}}(c)$, $\tilde{\mathcal{M}}(c)$, $\mathcal{A}(c)$ and $\mathcal{M}(c)$ to denote their time-s-sections, respectively. For autonomous systems, these sets are defined without the time component.

On the time-1-section of the Aubry set, a *pseudo-metric* d_c was introduced by Mather [38], whose definition relies on the quantity h_c^∞. Define

$$h_c^\infty((x,s),(x',s')) = \liminf_{\substack{s=t \bmod 1 \\ t'=s' \bmod 1 \\ t'-t\to\infty}} h_c((x,t),(x',t')).$$

For the autonomous system,

$$h_c^\infty(x,x') = \liminf_{\tau\to\infty} h_c((x,0),(x',\tau)).$$

The pseudo-metric d_c on the Aubry set is defined as

$$d_c((x,t),(x',t')) = h_c^\infty((x,t),(x',t')) + h_c^\infty((x',t'),(x,t)).$$

With the pseudo-metric d_c, one defines equivalence classes in an Aubry set. The equivalence $(x,t) \sim (x',t')$ implies $d_c((x,t),(x',t')) = 0$, with which one can define the *quotient Aubry set* $\mathcal{A}(c)/\sim$. Its element is called the *Aubry class*, denoted by $\mathcal{A}_i(c)$ or $\mathcal{A}_{c,i}$, whose lift to $TM \times \mathbb{T}$ is denoted by $\tilde{\mathcal{A}}_i(c)$. Thus,

$$\mathcal{A}(c) = \bigcup_{i\in\Lambda} \mathcal{A}_i(c), \quad \tilde{\mathcal{A}}(c) = \bigcup_{i\in\Lambda} \tilde{\mathcal{A}}_i(c).$$

Although Mather constructed an example with a quotient Aubry set homeomorphic to an interval, it is generic that each c-minimal measure contains not more than $n+1$ ergodic components if the system has n degrees of freedom (see [9]). In this case, each Aubry set contains at most $n+1$ classes. It is also known that the Mather set, the Aubry set and the Mañé set are invariant under symplectic transformations, which allows us to use the normal form (see [6]).

Appendix A.3 Elementary weak KAM theory

The concept of c-semi-static curves can be extended to the curves only defined on $\mathbb{R}^\pm$, which are called *forward* (*backward*) *c-semi-static curves*, denoted by $\gamma_c^\pm$, respectively. A curve γ_c^- (resp. γ_c^+) produces a backward (resp. forward) semi-static orbit $(\gamma_c^-, \dot{\gamma}_c^-)$ (resp. $(\gamma_c^+, \dot{\gamma}_c^+)$).

Proposition A.7. *If the Lagrangian L is of Tonelli type, for each point $(x,\tau) \in M \times \mathbb{T}$, there is at least one $\gamma_c^\pm(t,x,\tau)$ which is a forward* (*backward*) *semi-static curve.*

Since both the ω-limit set of $d\gamma_c^+$ and the α-limit set of $d\gamma_c^-$ are in the Aubry set, one defines

$$W_c^\pm = \bigcup_{(x,\tau)\in M\times\mathbb{T}} \left\{x, \tau, \frac{d\gamma_c^\pm(\tau,x,\tau)}{dt}\right\},$$

and calls W_c^+ the stable set, and W_c^- the unstable set of the c-minimal measures, respectively. If $\dot{\gamma}^-(\tau,x,\tau) = \dot{\gamma}^+(\tau,x,\tau)$ holds for some $(x,\tau) \in M \times \mathbb{T}$, passing through the point $(x,\tau,\dot{\gamma}_c^-(\tau,x,\tau))$, the orbit is either in the Aubry set or homoclinic to this Aubry set.

If the Aubry set consists of one class, the stable as well as the unstable set has its own generating function $u_c^\pm$ such that $W_c^\pm = \mathrm{Graph}(du_c^\pm)$ holds almost everywhere [26]. These functions are *weak KAM solutions*. We use $u_c^\pm$ to denote the weak KAM solution for the Lagrangian $L - \eta_c$, where η_c is a closed form with $[\eta_c] = c$. These functions are Lipschitz, and thus differentiable almost everywhere. At each differentiable point (x,τ), $(x,\tau,\partial_x u^-(x,\tau))$ uniquely determines the backward c-semi-static curve γ_x^-:

$(-\infty,\tau]\to M$ such that $\gamma_x^-(\tau)=x$ and $\dot\gamma_x^-(\tau)=\partial_y H(x,\tau,\partial_x u^-(x,\tau))$. Similarly, $(x,\tau,\partial_x u^+(x,\tau))$ uniquely determines the forward c-semi-static curve γ_x^-: $[\tau,\infty)\to M$ such that $\gamma_x^+(\tau)=x$ and $\dot\gamma_x^+(\tau)=\partial_y H(x,\tau,\partial_x u^+(x,\tau))$.

If two or more Aubry classes exist, there are infinitely many weak KAM solutions, among which we are interested in the so-called *elementary weak KAM solution*, obtained from the function h_c^∞. Indeed, treated as the function of (x,t), the function $h_c^\infty((x,t),(x',t'))$ is a weak KAM solution that determines orbits approaching the Aubry set as the time approaches infinity, and treated as the function of (x',t'), the function $h_c^\infty((x,t),(x',t'))$ is a weak KAM solution that determines orbits approaching the Aubry set as the time approaches minus infinity. Letting (x,t) range over an Aubry class, denoted by $\mathcal{A}_{c,i}$, one has a decomposition

$$h_c^\infty((x,t),(x',t'))=u_{c,i}^-(x',t')-u_{c,i}^+(x,t),\quad \forall\,(x',t')\in\mathbb{T}^n\times\mathbb{T},$$

where $u_{c,i}^+$ is a constant, and $u_{c,i}^-$ is called the elementary weak KAM solution with respect to $\mathcal{A}_{c,i}$. Similarly, letting (x',t') range over an Aubry class, one obtains an elementary weak KAM solution $u_{c,i}^-$. Again, for the autonomous system, one skips the time component.

Appendix B The theory of normally hyperbolic invariant manifolds

In this appendix, we introduce the theory of the normally hyperbolic invariant manifold (NHIM).

Appendix B.1 Normally hyperbolic invariant manifolds for Hamiltonian systems

Definition B.1. Let $f: M\to M$ be a C^r-diffeomorphism on a smooth manifold M with $r>1$. Let $N\subset M$ be a submanifold invariant under f, i.e., $f(N)=N$. We say that N is a *normally hyperbolic invariant manifold* (*NHIM*) if there exist a constant $C>0$, rates $0<\lambda<\mu^{-1}<1$ and a splitting $T_xM=E_x^s\oplus E_x^u\oplus T_xN$ for every $x\in N$ in such a way that

$$\begin{aligned} v\in E_x^s &\Leftrightarrow |Df^k(x)v|\leqslant C\lambda^k|v|,\quad k\geqslant 0,\\ v\in E_x^u &\Leftrightarrow |Df^k(x)v|\leqslant C\lambda^{|k|}|v|,\quad k\leqslant 0,\\ v\in T_xN &\Leftrightarrow |Df^k(x)v|\leqslant C\mu^k|v|,\quad k\in\mathbb{Z}. \end{aligned}$$

Notation 11. In the paper, we use the phrase "with uniform normal hyperbolicity independent of ε", which means that neither the normal Lyapunov exponents nor the splitting angle between E^s and E^u depends on ε.

Theorem B.2 (See [21, Theorem A.14]). *Let $N_X\subset M$—not necessarily compact—be normally hyperbolic invariant for the map f_X generated by the vector field X, which is uniformly C^r in a neighborhood U of N_X such that* $\mathrm{dist}(M\setminus U,N_X)>0$. *Let f_Y be the C^r-map generated by another vector field Y which is sufficiently close to X in the C^1-topology. Then we can find a manifold N_Y which is normally hyperbolic for Y and close to N_X in the $C^{\min\{r,|\frac{\ln\lambda}{\ln\mu}|-\varepsilon\}}$ topology for any small ε. The Lyapunov exponents for N_Y are arbitrarily close to those of N_X if Y is sufficiently close to X in the C^1 topology. The manifold N_Y is the only C^1 manifold close to N_X in the C^0 topology, and invariant under the flow of Y.*

We give a proof of the result in Appendix B.2 in a special setting adapted to the need of the paper.

Remark B.3. The NHICs constructed in the paper are not invariant under the original Hamiltonian flow, but invariant under a modified Hamiltonian by a cutoff function χ (see the proof of Proposition 5.2). These NHICs vary as χ does but on Aubry sets lying entirely in the region where $\chi=1$, and they all coincide. The role played by the NHICs is auxiliary. They are used to localize Aubry sets and run the Cheng-Yan genericity argument. Once the generalized transition chains are known to exist, they are no longer needed. Thus the NHICs obtained by choosing one cutoff function χ are sufficient for our purpose.

When the normally hyperbolic flow is Hamiltonian, we have the following theorem saying that the restriction of the Hamiltonian system to the central manifold is also Hamiltonian with fewer degrees of freedom.

Theorem B.4 (See [23, Theorems 23 and 26])**.** *Suppose that M is endowed with a (an exact) symplectic form ω. Let $f_\varepsilon : M \to M$ be a C^r family of Hamiltomorphisms, $r \geqslant 2$ preserving ω. Assume that $N \subset M$ is an NHIM for f_0 with rates λ and μ.*

(1) *Then for sufficiently small ε, there exist C^ℓ-families of diffeomorphisms k_ε and r_ε with $\ell \leqslant \min\{r, |\frac{\ln\lambda}{\ln\mu}|\}$ satisfying $f_\varepsilon \circ k_\varepsilon = k_\varepsilon \circ r_\varepsilon$, where k_ε is the map such that $k_\varepsilon(N) = N_\varepsilon$ and $r_\varepsilon : N \to N$ is the restricted map on N.*

(2) *We denote by $\mathcal{R}_\varepsilon$ the generating vector field corresponding to r_ε defined by $\frac{d}{d\varepsilon} r_\varepsilon = \mathcal{R}_\varepsilon \circ r_\varepsilon$. Then we have*

- *$k_\varepsilon^* \omega = \omega_N$ is a (an exact) symplectic form on N; it is independent of ε;*
- *the vector field $\mathcal{R}_\varepsilon$ is the (exactly) Hamiltonian vector field with respect to ω_N; moreover, its (global) Hamiltonian is $R_\varepsilon = F_\varepsilon \circ k_\varepsilon$, where F_ε is the Hamiltonian for f_ε.*

Appendix B.2 A special normally hyperbolic invariant manifold theorem

In this paper, we need a special version of the theorem of the NHIM. Here, we present its detailed proof using the graph transformation method. The statement given below is adapted to the setting needed in the paper and we do not pursue generality.

Theorem B.5. *Let $N = (\mathbb{R}^m/S\mathbb{Z}^m) \times \mathbb{R}^{m'}$, $S \in \mathrm{SL}(m, \mathbb{R})$ be a submanifold of a (non compact) manifold $M^{m+m'+k}$. Given $\Lambda > 0$, we define $N_\Lambda = (\mathbb{R}^m/S\mathbb{Z}^m) \times B_\Lambda(0)$ and choose a small neighborhood $U(\subset M)$ of N_Λ. Let $V = (\dot q, \dot p, \dot n)$ be a C^2 vector field compactly supported in U satisfying the following properties:*

(1)

$$\begin{cases} \dot q = \varepsilon^{-1/2}\omega^\star + a(p), \\ \dot p = 0, \end{cases}$$

where $(q, p) \in N_\Lambda$, $\omega^\star \in \mathbb{R}^m$ is constant, $a(p) \in C^2$ and $\varepsilon > 0$.

(2) *Restricted on the normal bundle $\bigcup_{x\in N_\Lambda} E_x^s \oplus E_x^u$, we have $\dot n = An$, where $A \in \mathbb{R}^{k\times k}$ is a constant matrix all of whose eigenvalues lie off the imaginary axis.*

Then there exists a γ_0 such that any vector field $V_{\gamma,\varepsilon}$ compactly supported in U and satisfying that $\|V_{\gamma,\varepsilon} - V\|_{C^1} \leqslant \gamma_0$, admits an NHIC that is a graph over N_Λ.

Remark B.6. We see in the following proof that ε does not play any role, since the large term $\varepsilon^{-1/2}\omega^\star$ is constant and does not appear in the derivative of the time-1 map of the flow; on the other hand, only derivative information matters in the proof (see (B.2)). The vector field $V_{\gamma,\varepsilon}$ is also allowed to depend on t periodically, and even with fast oscillations for example it depends on $t/\varepsilon^\alpha \in \mathbb{T}$, for any $\alpha > 0$, in which case the $\|\cdot\|_{C^1}$ norm does not include the derivative with respect to t.

Proof of Theorem B.5. We follow the proof of Fenichel [27] (see [30] for another approach). In the proof, for clarity of the ideas, we consider first the contracting case, namely, $E^u = 0$ in the splitting of T_xM (see Definition B.1), i.e., all the eigenvalues of A have negative real parts.

We denote by f (resp. f_γ) the time-1 map generated by the vector field V (resp. $V_{\gamma,\varepsilon}$). We now introduce coordinates. We cover a neighborhood U_d ($d > 0$) of the center manifold N_Λ by balls of the form $B_{2d}(p_i)$ with $p_i \in N_\Lambda$ using any preferred Riemannian metric. In each of the ball $B_{2d}(p_i)$, we choose local coordinates given by $\exp_{p_i} : T_{p_i}N_\Lambda \oplus E^s_{p_i} \to B_{2d}(p_i)$ with

$$\exp_{p_i}(x, 0) \in B_{2d}(p_i) \cap N_\Lambda \quad \text{and} \quad \exp_{p_i}(0, 0) = p_i. \tag{B.1}$$

In coordinates, the map f^n can be written as

$$F_{j,i} := \exp_{p_j}^{-1} \circ f^n \circ \exp_{p_i} : T_{p_i}N \oplus E^s_{p_i} \to T_{p_j}N \oplus E^s_{p_j},$$

if $p = f^{-n}(p')$ for $p \in B_{2d}(p_i)$ and $p' \in B_{2d}(p_j)$, where the number of iterations n will be determined later. We suppress the subscripts i and j for simplicity and define $F(x,y) = (X(x,y), Y(x,y))$, where $Y(x,0) = 0$. We set

$$dF = \begin{pmatrix} \partial_x X & \partial_y X \\ \partial_x Y & \partial_y Y \end{pmatrix} =: \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

We have by definition that

$$\begin{aligned} C(x,0) &= \partial_x Y(x,0) = 0, \\ D(x,0) &= \partial_y Y(x,0) = dF\,|_{E^s}, \\ A(x,0) &= \partial_x X(x,0) = dF\,|_{E^c}. \end{aligned}$$

Now the normal hyperbolicity assumption implies the following important bounds:

$$\|D\|_{C^0}\|A^{-1}\|_{C^0} < 1/2, \quad \|D\|_{C^0} < 1/2, \quad \|C\|_{C^0} < \eta \ll 1 \tag{B.2}$$

by choosing n large and the neighborhood U_d small enough. The derivative dF is obtained by integrating the variational equation derived from the ODE of $V_{\gamma,\varepsilon}$. Note that the term $\varepsilon^{-1/2}\omega^\star$ does not appear in the variational equation since $\omega^\star$ is a constant. Moreover, if $V_{\gamma,\varepsilon}$ depends on t explicitly, the variational equation does not involve derivatives with respect to t. Since the map f_γ is γ-close to f in the C^1 norm, we define F_γ from f_γ^n in the same way as F from f^n. For fixed n and small enough γ, the above bounds (B.2) also hold for F_γ in the domain U_d. In the following, we suppress the subscript γ and work exclusively with F_γ instead of F.

Define first the set $\mathcal{S}$ of Lipschitz sections $S : T_{p_i}N_\Lambda \to T_{p_i}N_\Lambda \oplus E^s_{p_i}$. Next, we define

$$\mathcal{S}_\delta := \{S \in \mathcal{S} \mid \mathrm{Lip}(S) \leqslant \delta\}.$$

The graph transformation is defined to be

$$G : \mathcal{S}_\delta \to \mathcal{S}, \quad (G(S))(X(x, S(x))) = Y(x, S(x)). \tag{B.3}$$

Lemma B.7. *For sufficiently small η and δ, the image of the graph transformation G lies in $\mathcal{S}_\delta$, i.e., $G : \mathcal{S}_\delta \to \mathcal{S}_\delta$.*

Proof. Suppose that $\xi = X(x, S(x))$ and $\xi' = X(x', S(x'))$ are sufficiently close. The injectivity of $X(\cdot, S(\cdot))$ will be shown below. Then we have

$$\begin{aligned} \|(G(S))(\xi) - (G(S))(\xi')\| &= \|Y(x, S(x)) - Y(x', S(x'))\| \\ &\leqslant \|C\|_{C^0}\|x - x'\| + \delta\|D\|_{C^0}\|x - x'\|. \end{aligned} \tag{B.4}$$

Next, we bound $\|x - x'\|$ using $\|\xi - \xi'\|$. It holds that

$$\begin{aligned} \|\xi - \xi'\| &= \|X(x, S(x)) - X(x', S(x'))\| \\ &\geqslant \|X(x, S(x)) - X(x', S(x))\| - \|X(x', S(x)) - X(x', S(x'))\| \\ &\geqslant \|A^{-1}\|^{-1}_{C^0}\|x - x'\| - \|B\|_{C^0}\|S(x) - S(x')\| \\ &\geqslant (\|A^{-1}\|^{-1}_{C^0} - \delta\|B\|_{C^0})\|x - x'\|. \end{aligned}$$

Let $c = \frac{\|C\|_{C^0} + \delta\|D\|_{C^0}}{\|A^{-1}\|^{-1}_{C^0} - \delta\|B\|_{C^0}}$. Combining (B.4), we get

$$\|(G(S))(\xi) - (G(S))(\xi')\| \leqslant c\|\xi - \xi'\|.$$

We can make $\|C\|_{C^0}$ as small as we wish by choosing η small, and hence for small δ, the leading term in c is given by $\delta\|D\|_{C^0}\|A^{-1}\|_{C^0} \leqslant \delta/2$. □

Lemma B.8. *The graph transformation $G : \mathcal{S}_\delta \to \mathcal{S}_\delta$ is a contraction in the C^0 norm, i.e., $\|G(S) - G(S')\|_{C^0} \leqslant \lambda\|S - S'\|_{C^0}$ for some $0 < \lambda < 1$.*

Proof. For $S, S' \in \mathcal{S}_\delta$, choosing x and x' with $\xi = X(x, S(x)) = X(x', S'(x'))$, we get

$$\begin{aligned}\|(G(S))(\xi) - (G(S'))(\xi)\| &= \|Y(x, S(x)) - Y(x', S'(x'))\| \\ &\leqslant \|C\|_{C^0}\|x - x'\| + \|D\|_{C^0}(\|S(x) - S'(x)\| + \|S'(x) - S'(x')\|) \\ &\leqslant (\|C\|_{C^0} + \delta\|D\|_{C^0})\|x - x'\| + \|D\|_{C^0}\|S - S'\|_{C^0}.\end{aligned}$$

Since $\|C\|_{C^0} < \eta$ can be as small as we wish, $\|D\|_{C^0} < 1/2$ due to the contraction. The proof will be completed if we can show $\|x - x'\| \leqslant c\|S - S'\|_{C^0}$ for some constant c. We have

$$\|X(x, S(x)) - X(x', S(x))\| \geqslant \|A\|_{C^0}\|x - x'\|$$

and

$$\|X(x', S'(x')) - X(x', S(x))\| \leqslant \|B\|_{C^0}(\delta\|x - x'\| + \|S - S'\|_{C_0}).$$

Since we have $\xi = X(x, S(x)) = X(x', S'(x'))$, combining the two estimates, we get $\|x-x'\| \leqslant c\|S-S'\|_{C^0}$ for some constant c. This completes the proof. □

By the contracting mapping theorem, there exists a unique δ-Lipschitz solution S to the graph transformation, $S = G(S)$. By the uniqueness of the fixed point of G, we get that $\exp S$ is invariant under f_γ.

For the hyperbolic splitting, i.e., the matrix A has both positive and negative eigenvalues, we introduce coordinates respecting the splitting and write the map f_γ^n in coordinates as before, i.e.,

$$F(x, y, z) = (X(x, y, z), Y(x, y, z), Z(x, y, z)) \in E^c_{p'} \oplus E^u_{p'} \oplus E^s_{p'}, \tag{B.5}$$

where $(x, y, z) \in E^c_p \oplus E^u_p \oplus E^s_p$ and $f_\gamma^{-n}(p') = p$ with the derivative control for sufficiently small η, and in a sufficiently small neighborhood U_d,

$$\begin{aligned}&\|\partial_x X^{-1}\|^k_{C^0}\|\partial_z Z\|_{C^0} < 1/2, \quad \|\partial_x X\|^k_{C^0}\|(\partial_z Y)^{-1}\|_{C^0} < 1/2, \\ &\|\partial_x Z\|_{C^0},\ \|\partial_x Y\|_{C^0},\ \|\partial_y Z\|_{C^0},\ \|\partial_z Y\|_{C^0} < \eta.\end{aligned}$$

The graph transformation is defined as follows: for $S(x) = (S^u(x), S^s(x))$, a section in $E^c_{p'} \to E^c_{p'} \oplus E^u_{p'} \oplus E^s_{p'}$, we assign $S' = (S'^u(x), S'^s(x)) = G(S)$, where $S'^s(X(x, S^u(x), S^s(x))) = Z(x, S^u(x), S^s(x))$ and $S'^u(x)$ is solved implicitly from

$$S^u(X(x, S'^u(x), S^s(x))) = Y(x, S'^u(x), S^s(x)).$$

The solution exists since $Y(x, 0, 0) = 0$ and $\partial_y Y \neq 0$. One can verify that the graph transformation G is a contraction from $\mathcal{S}_\delta \to \mathcal{S}_\delta$, and hence there is a unique solution (S^u, S^v) satisfying

$$\begin{aligned}S^s(X(x, S^u(x), S^s(x))) &= Z(x, S^u(x), S^s(x)), \\ S^u(X(x, S^u(x), S^s(x))) &= Y(x, S^u(x), S^s(x)).\end{aligned}$$

Here, we only show how to prove the existence of the NHIC. We see from the above proof that the $\varepsilon^{-1/2}\omega^\star$ term does not play a role since it disappears in the derivative of the map. It turns out that the conclusion of the standard normally hyperbolic invariant manifold theorem holds in our setting. For more information such as the regularity of the center manifold and the existence and regularity of stable and unstable manifolds, we refer the readers to [27]. □

Appendix C Induction and dynamics around complete resonances

The main construction in this paper was done in the previous section for the $n = 4$ case. In this appendix, we perform induction for the general $n > 4$ case. Theorem 2.4 will be proved in this appendix in full generality.

Appendix C.1 Two types of resonances and normal forms

In this appendix, we provide more details of the proof of Proposition 13.3. Suppose that we have completed step ℓ and are about to work on the $(\ell+1)$-th step of the induction. At step ℓ, we are handed with the following data (we use the same notations as in Section 13):

(1) for each $\kappa=0,\dots,\sharp\mathcal{I}(\ell)-1$, we have a frequency segment $\omega_\kappa^{(\ell)}(a)$;

(2) a number $\mu^{(\ell)}$: the size of the neighborhood of $\omega_\kappa^{(\ell)}(a)$ for all κ;

(3) associated with each $\omega_\kappa^{(\ell)}(a)$ for all a, a collection of irreducible resonant integer vectors $\boldsymbol{K}_\kappa^{(\ell)}=\{\boldsymbol{k}_\kappa^{\prime(\ell)},\dots,\boldsymbol{k}_\kappa^{(\kappa),(\ell)}\}$. For some a, there is one more denoted by $\boldsymbol{k}_\kappa^{o,(\ell)}$. We define $\boldsymbol{K}_\kappa^{o,(\ell)}=\boldsymbol{K}_\kappa^{(\ell)}\cup\{\boldsymbol{k}_\kappa^{o,(\ell)}\}$. By definition, we have $\sharp\boldsymbol{K}_\kappa^{(\ell)}=\kappa$ and $\sharp\boldsymbol{K}_\kappa^{o,(\ell)}=\kappa+1$.

We next pick a rational number $\frac{p_{\ell+1}}{q_{\ell+1}}$ such that $\frac{p_{\ell+1}}{q_{\ell+1}}\omega_{1,2}^*$ is within $\mu^{(\ell)}$-distance of the irrational number on the $(\ell+3)$-th column of the table (13.1).

When we update ℓ to $(\ell+1)$, the subscript $(i,[j])$ remains unchanged, but the subscript κ associated with each $(i,[j])$ will also be updated to $\kappa+1$. The $\kappa=0$ case was handled in Sections 3, 5 and 6, and the $\kappa=1$ case was done in Section 9. The $\kappa=n-2$ case means that the frequency segment $\omega_{i,[j]}^{(\ell)}(a)$ has completed the reduction of orders so it becomes $\boldsymbol{\omega}_{i,[j]}^\sharp$ and it will be treated in Appendix C.4. So in the remaining part of this appendix till Appendix C.4, we will consider the range $\ell=0,\dots,Mn-3$ and $\kappa=0,1,\dots,\min\{\sharp\mathcal{I}(\ell)-1,n-3\}$.

Appendix C.1.1 *Two types of resonances*

The following lemma is an analogue of Lemmas 4.5 and 8.2.

Lemma C.1. *Let $\omega_\kappa^{(\ell)}(a)$, $\mu^{(\ell)}$ and $\omega_{\kappa+1}^{(\ell+1)}(a)$, $\boldsymbol{K}_{\kappa+1}^{(\ell+1)}$ $(\kappa=0,1,\dots,\min\{\sharp\mathcal{I}(\ell)-1,n-3\})$ be as above. For any $K^{(\ell+1)}>\max_\kappa|\boldsymbol{K}_\kappa^{(\ell)}|$, let $\boldsymbol{k}_{\kappa+1,i}^o$ $(i=1,\dots,m_\kappa)$ be the collection of all the integer vectors in $\mathbb{Z}^n_{K^{(\ell+1)}}\setminus\mathrm{span}\boldsymbol{K}_{\kappa+1}^{(\ell+1)}$ satisfying $\langle\boldsymbol{k}_{\kappa+1,i}^o,\omega_{\kappa+1}^{(\ell+1)}(a_i^o)\rangle=0$ for some a_i^o, and $(\boldsymbol{k}_{\kappa+1,i}^o)^\perp$ be the $(n-1)$-dimensional space orthogonal to the vector $\boldsymbol{k}_{\kappa+1,i}^o$. Then there exists a $\mu^{(\ell+1)}$ with $B_{\mu^{(\ell+1)}}(\omega_{\kappa+1}^{(\ell+1)}(a))\subset B_{\mu^{(\ell)}}(\omega_\kappa^{(\ell)}(a))$ and*

(1) for all $\omega\in B_{\mu^{(\ell+1)}}(\omega_{\kappa+1}^{(\ell+1)}(a))\setminus\bigcup_i B_{\varepsilon^{1/3}}(\omega_{\kappa+1}^{(\ell+1)}(a_i^o)+(\boldsymbol{k}_{\kappa+1,i}^o)^\perp)$ and for sufficiently small ε, we have

$$|\langle\boldsymbol{k},\omega\rangle|>\varepsilon^{1/3},\quad\forall\,\boldsymbol{k}\in\mathbb{Z}^n_{K^{(\ell+1)}}\setminus\mathrm{span}_{\mathbb{Z}}\boldsymbol{K}_{\kappa+1}^{(\ell+1)};$$

(2) for all $\omega\in B_{\mu^{(\ell+1)}}(\omega_{\kappa+1}^{(\ell+1)}(a))\cap B_{\varepsilon^{1/3}}(\omega_{\kappa+1}^{(\ell+1)}(a_i^o)+(\boldsymbol{k}_{\kappa+1,i}^o)^\perp)$, for each i and for all $\boldsymbol{k}\in\mathbb{Z}^n_{K^{(\ell+1)}}\setminus\mathrm{span}_{\mathbb{Z}}\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}$, we have

$$|\langle\boldsymbol{k},\omega\rangle|\geqslant nK^{(\ell+1)}\mu^{(\ell+1)}.\tag{C.1}$$

Note that in the lemma, our choices of $\mu^{(\ell+1)}$ and $K^{(\ell+1)}$ are independent of the subscript κ. We will next introduce a small parameter $\delta^{(\ell+1)}$, independent of κ, to determine $K^{(\ell+1)}$ and hence $\mu^{(\ell+1)}$.

Appendix C.1.2 *The KAM normal forms*

Now we determine the resonance submanifold as

$$\Sigma(\boldsymbol{K}_\kappa^{o,(\ell)}):=\{y\mid\langle\boldsymbol{k}_\kappa^{o,(\ell)},\boldsymbol{\omega}(y)\rangle=\langle\boldsymbol{k}_\kappa^{\prime(\ell)},\boldsymbol{\omega}(y)\rangle=\cdots=\langle\boldsymbol{k}_\kappa^{(\kappa),(\ell)},\boldsymbol{\omega}(y)\rangle=0\}.$$

Lemma C.1 allows us to apply Proposition 3.3 in the two cases in Lemma C.1 to obtain the following normal forms.

Lemma C.2. *Let $\delta^{(\ell+1)}$ be a small number satisfying $\delta^{(\ell+1)}<\min_\kappa\{3(|\boldsymbol{K}_\kappa^{(\ell)}|)^{-2}\}$ and define $K^{(\ell+1)}=(\delta^{(\ell+1)}/3)^{-1/2}$. Then there exists an $\varepsilon_1^{(\ell+1)}=\varepsilon_1^{(\ell+1)}(\delta^{(\ell+1)},\Lambda)$ such that for all $\varepsilon<\varepsilon_1^{(\ell+1)}$, the following holds. Suppose that $\omega^\star$ is as in the case (1) in Lemma C.1. Then there exists a symplectic transformation $\phi_{\kappa+1}^{(\ell+1)}$ defined on $B_\Lambda(0)\times\mathbb{T}^n$ that is $o_{\varepsilon\to0}(1)$ close to identity in the C^{r_0-1} norm such that*

$$\mathsf{H}^{(\ell+1)}_{\kappa+1,\delta^{(\ell+1)}}:=\mathsf{H}\circ\phi_{\kappa+1}^{(\ell+1)}(x,Y)$$

$$= \frac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \frac{1}{2}\langle \mathsf{A}Y, Y\rangle + \Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}\mathsf{V} + \delta^{(\ell+1)}\mathsf{R}_{\kappa+1}^{(\ell+1)}(x, Y), \tag{C.2}$$

where

(1) A *and* V *are the same as those in Lemma* 5.3;

(2) $\mathsf{R}_{\kappa+1}^{(\ell+1)}(x, Y) = \mathsf{R}_{\kappa+1,I}^{(\ell+1)}(x) + \mathsf{R}_{\kappa+1,II}^{(\ell+1)}(x, Y)$, *where* $\mathsf{R}_{\kappa+1,I}^{(\ell+1)}$ *consists of Fourier modes of* V *not in* $\mathbb{Z}^n_{K^{(\ell+1)}} \cup \mathrm{span}_{\mathbb{Z}}\boldsymbol{K}_{\kappa+1}^{(\ell+1)}$, *and we have* $|\mathsf{R}_{\kappa+1,I}^{(\ell+1)}|_{r_0-2} \leqslant 1$ *and* $|\mathsf{R}_{\kappa+1,II}^{(\ell+1)}|_{r_0-5} \leqslant 1$.

Lemma C.3. *Let* $\delta^{(\ell+1)}$ *and* $K^{(\ell+1)}$ *be as in the previous lemma. Then there exists an* $\varepsilon_2^{(\ell+1)} = \varepsilon_2^{(\ell+1)}(\delta^{(\ell+1)}, \Lambda)$ *such that for all* $\varepsilon < \varepsilon_2^{(\ell+1)}$ *and any* $y^\star$ *such that* $\omega^\star = \boldsymbol{\omega}(y^\star)$ *is as in the case* (2) *in Lemma* C.1, *there exists a symplectic transformation* $\phi_{\kappa+1}^{(\ell+1)}$ *defined on* $B_\Lambda(0) \times \mathbb{T}^n$ *that is* $o_{\varepsilon\to 0}(1)$ *close to identity in the* C^{r_0-1} *norm such that*

$$\begin{aligned}\mathsf{H}_{\kappa+1,\delta^{(\ell+1)}}^{(\ell+1)} &:= \mathsf{H} \circ \phi_{\kappa+1}^{(\ell+1)}(x, Y)\\ &= \frac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \frac{1}{2}\langle \mathsf{A}Y, Y\rangle + \Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}\mathsf{V} + \delta^{(\ell+1)}\mathsf{R}_{\kappa+1}^{(\ell+1)}(x, Y),\end{aligned} \tag{C.3}$$

where

(1) A *and* V *are the same as those in Lemma* 5.3;

(2) $\mathsf{R}_{\kappa+1}^{(\ell+1)}(x, Y) = \mathsf{R}_{\kappa+1,I}^{(\ell+1)}(x) + \mathsf{R}_{\kappa+1,II}^{(\ell+1)}(x, Y)$, *where* $\mathsf{R}_{\kappa+1,I}^{(\ell+1)}$ *consists of Fourier modes of* V *not in* $\mathbb{Z}^n_{K^{(\ell+1)}} \cup \mathrm{span}_{\mathbb{Z}}\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}$, *and we have* $|\mathsf{R}_{\kappa+1,I}^{(\ell+1)}|_{r_0-2} \leqslant 1$ *and* $|\mathsf{R}_{\kappa+1,II}^{(\ell+1)}|_{r_0-5} \leqslant 1$.

Appendix C.2 NHICs away from strong resonances

The following result is an analogue of Proposition 8.3, which will be used to establish the existence of the NHIC.

Proposition C.4. *Suppose that there exists an open and dense set* $\mathcal{O}_\kappa^{(\ell)} \subset \Pi_{\boldsymbol{K}_\kappa^{(\ell)}}C^r$ *such that each* $\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P(y, \cdot) \in \mathcal{O}_\kappa^{(\ell)}$ *has a unique nondegenerate global maximum along the segment* $y \in \boldsymbol{\omega}^{-1}(\omega_\kappa^{(\ell)}(a))$, *up to finitely many bifurcations, where there are two nondegenerate global maxima.*

Then for each $\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P \in \mathcal{O}_\kappa^{(\ell)}$ *there exists a* $\delta_{0,\kappa}^{(\ell)} = \delta_{0,\kappa}^{(\ell)}(\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P)$ *such that defining* $\delta_0^{(\ell)} = \min_\kappa \delta_{0,\kappa}^{(\ell)}$ *for any* $\delta^{(\ell)} < \delta_0^{(\ell)}$, $K^{(\ell)} = (\delta^{(\ell)}/3)^{1/2}$ *and* $\mu^{(\ell)} = \mu^{(\ell)}(K^{(\ell)})$ *as in Lemma* C.1, *and choosing* $\omega_{\kappa+1}^{(\ell+1)}(a) \subset B_{\mu^{(\ell)}}(\omega_\kappa^{(\ell)}(a))$ *associated with irreducible* $\boldsymbol{K}_{\kappa+1}^{(\ell+1)}$, *we have an open and dense set* $\mathcal{O}_{\kappa+1}^{(\ell+1)} = \mathcal{O}_{\kappa+1}^{(\ell+1)}(\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P)$ *in the unit ball of* $\Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}C^r/\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}C^r$ *such that for each* $\Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}P$ *with*

$$\Pi_{\boldsymbol{K}_\kappa^{(\ell)}}(\Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}P) = \Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P \quad \textit{and} \quad \Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}P - \Pi_{\boldsymbol{K}_\kappa^{(\ell)}}P \in \mathcal{O}_{\kappa+1}^{(\ell+1)},$$

we have that $\Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}P(y, \cdot)$ *has a unique nondegenerate global maximum along the segment* $y \in \boldsymbol{\omega}^{-1}(\omega_{\kappa+1}^{(\ell+1)}(a))$, *up to finitely many bifurcations, where there are two nondegenerate global maxima.*

In the cases of Lemmas C.1(1) and C.2, we can repeat the argument of Proposition 5.2 to find NHICs $\mathcal{C}(\boldsymbol{K}_{\kappa+1}^{(\ell+1)})$ homeomorphic to $T^*\mathbb{T}^{n-\kappa-1}$ along the frequency $\omega_{\kappa+1}^{(\ell+1)}(a)$ with the help of Proposition C.4. In the cases of Lemmas C.1(2) and C.3 in the presence of an extra resonance $\boldsymbol{k}_{\kappa+1}^{o,(\ell+1)}$, the NHICs may or may not exist. When the NHICs do not exist, we denote the corresponding resonance submanifolds by $\Sigma(\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)})$. We have the following result.

Proposition C.5. *Let* $\Pi_{\boldsymbol{K}_{\kappa+1}^{(\ell+1)}}P$ *be as in the conclusion of Proposition* C.4. *Then for any* $\lambda > 0$, *there exists a* $\delta_1^{(\ell+1)}$ *such that in the system* $\mathsf{H}_{\kappa+1,\delta^{(\ell+1)}}^{(\ell+1)}$, *for all* $0 < \delta^{(\ell+1)} < \delta_1^{(\ell+1)}$, *we have the following*:

(1) *there exists a* C^{r-1} *NHIC* $\mathcal{C}(\boldsymbol{K}_{\kappa+1}^{(\ell+1)})$ *homeomorphic to* $T^*\mathbb{T}^{n-\kappa-1}$ *up to finitely many bifurcations*;

(2) *the Mather set lying in* $B_\Lambda(0) \times \mathbb{T}^n$ *and with the rotation vector orthogonal to* $\boldsymbol{K}_{\kappa+1}^{(\ell+1)}$ *lies inside* $\mathcal{C}(\boldsymbol{K}_{\kappa+1}^{(\ell+1)})$, *provided that the rotation vector does not intersect the* $\lambda\sqrt{\varepsilon}$-*neighborhood of* $\partial h(\Sigma(\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}))$;

(3) *the normal hyperbolicity is independent of* ε *or* $\delta^{(\ell+1)}$.

Appendix C.3 NHICs around strong resonances

We next focus on the case (2) of Lemma C.1, i.e., a generalization of that in Sections 9 and 10. In this appendix, we perform the reduction of orders in the presence of an extra resonance $\boldsymbol{k}_{\kappa+1}^{o,(\ell+1)}$. For given $(i,[j])$, the extra resonance may appear during the κ-th step of the reduction of orders. Without loss of generality, we assume that we encounter the extra resonance point during the $\kappa=0$ step of the reduction of orders. In this case, $\boldsymbol{k}_{\kappa+1}^{\prime(\ell+1)}$ and $\boldsymbol{k}_{\kappa+1}^{o,(\ell+1)}$ have comparable lengths and are much shorter than other vectors in $\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}$.

Appendix C.3.1 *The linear symplectic transformation and the Hamiltonian normal form*

We construct a matrix $M_{\kappa+1}^{(\ell+1)}\in\mathrm{SL}(n,\mathbb{Z})$, $\kappa=0,\ldots,\min\{\sharp\mathcal{I}(\ell),n-3\}$, whose first $\kappa+2$ rows are exactly the vectors in $\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}$ ordered as $\boldsymbol{k}_{\kappa+1}^{o,(\ell+1)},\boldsymbol{k}_{\kappa+1}^{\prime(\ell+1)},\ldots,\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}$. This is always possible by applying Lemma 4.4 repeatedly. The matrix $M_{\kappa+1}^{(\ell+1)}$ induces a symplectic transformation

$$\mathfrak{M}_{\kappa+1}^{(\ell+1)}:(x,Y)\mapsto(M_{\kappa+1}^{(\ell+1)}x,(M_{\kappa+1}^{(\ell+1)})^{-t}Y).$$

We define $\mathsf{A}_{\kappa+1}^{(\ell+1)}=M_{\kappa+1}^{(\ell+1)}\mathsf{A}(M_{\kappa+1}^{(\ell+1)})^{t}$. Then the (i,j)-th entry of $\mathsf{A}_{\kappa+1}^{(\ell+1)}$ is given by $\boldsymbol{k}_{\kappa+1}^{(i-1),(\ell+1)}\mathsf{A}(\boldsymbol{k}_{\kappa+1}^{(j-1),(\ell+1)})^{t}$, $i,j=1,\ldots,\kappa+2$, and we count o as 0.

We choose the base point $y^\star$ such that the frequency vector $\omega^\star=\boldsymbol{\omega}(y^\star)\in\Sigma(\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)})$, and then we get the transformed frequency vector $M_{\kappa+1}^{(\ell+1)}\omega^\star$ has zero as the first $\kappa+2$ entries. We define $M_{\kappa+1}^{(\ell+1)}\omega^\star=(0,\hat\omega_{\kappa+1}^{(\ell+1)})\in\mathbb{R}^n$ for some vector $\hat\omega_{\kappa+1}^{(\ell+1)}\in\mathbb{R}^{n-\kappa-2}$.

The Hamiltonian (C.3) under the transformation becomes

$$\begin{aligned}&(\mathfrak{M}_{\kappa+1}^{(\ell+1)})^{-1*}\mathsf{H}_{\kappa+1,\delta^{(\ell+1)}}^{(\ell+1)}\\&\quad=\frac{1}{\sqrt{\varepsilon}}\langle\hat\omega_{\kappa+1}^{(\ell+1)},\hat Y\rangle+\frac12\langle\mathsf{A}_{\kappa+1}^{(\ell+1)}Y,Y\rangle+V_{\kappa+1}^{(\ell+1)}(x_1,\ldots,x_{\kappa+2})+\delta^{(\ell+1)}R_{\kappa+1}^{(\ell+1)}(x,Y),\end{aligned}\tag{C.4}$$

where $V_{\kappa+1}^{(\ell+1)}=(\mathfrak{M}_{\kappa+1}^{(\ell+1)})^{-1}\Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}\mathsf{V}$ and $R_{\kappa+1}^{(\ell+1)}=(\mathfrak{M}_{\kappa+1}^{(\ell+1)})^{-1*}\mathsf{R}_{\kappa+1}^{(\ell+1)}$.

We denote by $A_{\kappa+1}^{(\ell+1)}$ the first $(\kappa+2)\times(\kappa+2)$ block of $\mathsf{A}_{\kappa+1}^{(\ell+1)}$ and by $A_\kappa^{(\ell)}$ the first $(\kappa+1)\times(\kappa+1)$ block of $\mathsf{A}_{\kappa+1}^{(\ell+1)}$. Note that $A_\kappa^{(\ell)}$ depends only on A and $\boldsymbol{K}_\kappa^{o,(\ell)}$ but does not depend on $\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}$.

Next, we introduce two subsystems

$$\begin{aligned}&\mathsf{G}_\kappa^{(\ell)}=\frac12\langle\mathsf{A}_\kappa^{(\ell)}Y_\kappa^{(\ell)},Y_\kappa^{(\ell)}\rangle+V_\kappa^{(\ell)}(x_\kappa^{(\ell)}),\quad T^*\mathbb{T}^{\kappa+1}\to\mathbb{R},\\&\mathsf{G}_{\kappa+1}^{(\ell+1)}=\frac12\langle\mathsf{A}_{\kappa+1}^{(\ell+1)}Y_{\kappa+1}^{(\ell+1)},Y_{\kappa+1}^{(\ell+1)}\rangle+V_{\kappa+1}^{(\ell+1)}(x_{\kappa+1}^{(\ell+1)}),\quad T^*\mathbb{T}^{\kappa+2}\to\mathbb{R}.\end{aligned}\tag{C.5}$$

Defining $\delta^{(\ell)}\bar V_{\kappa+1}^{(\ell+1)}(x_{\kappa+1}^{(\ell+1)}):=V_{\kappa+1}^{(\ell+1)}-V_\kappa^{(\ell)}$, we have

$$\|\delta^{(\ell)}\bar V_{\kappa+1}^{(\ell+1)}\|_{C^{r_0-2}}\leqslant\frac{1}{|\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}|^2}\leqslant\delta^{(\ell)}/3,$$

since the difference comes from the Fourier modes in V containing $\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}$ whose length is greater than $K^{(\ell)}=(\delta^{(\ell)}/3)^{1/2}$.

Appendix C.3.2 *NHICs around strong double resonances*

In the following, without loss of generality, *we fix λ such that $\alpha_{\tilde{\mathsf{G}}}(\partial\beta_{\tilde{\mathsf{G}}}(\lambda(1,0)))<\tilde\Delta_0$, where $\tilde\Delta_0$ and $\tilde{\mathsf{G}}$* (*see* (6.12)) *depend only on $\Pi_{\boldsymbol{k}',\boldsymbol{k}^o}\mathsf{V}$ but not on other resonant integer vectors* (*see Proposition* 7.4). *This is assumed in Propositions* 7.7 *and* 10.3.

Proposition C.6. *There exists an open and dense set $\mathcal{O}_\kappa^{(\ell)}$ in the unit ball of $\Pi_{\boldsymbol{K}_\kappa^{o,(\ell)}}C^r(\mathbb{T}^n)$ such that for each V with $\Pi_{\boldsymbol{K}_\kappa^{o,(\ell)}}\mathsf{V}\in\mathcal{O}_\kappa^{(\ell)}$, there exists a $\delta_2^{(\ell)}=\delta_2^{(\ell)}(\Pi_{\boldsymbol{K}_\kappa^{o,(\ell)}}\mathsf{V})$ such that for all $0<\delta^{(\ell)}\leqslant\delta_2^{(\ell)}$,*

$K^{(\ell)}=(\delta^{(\ell)}/3)^{1/2}$ and any $\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}$ with $|\boldsymbol{k}_{\kappa+1}^{(\kappa+1),(\ell+1)}|>K^{(\ell)}$, there exists an open and dense set $\mathcal{O}_{\kappa+1}^{(\ell+1)}=\mathcal{O}_{\kappa+1}^{(\ell+1)}(\Pi_{\boldsymbol{K}_{\kappa}^{o,(\ell)}}\mathsf{V})$ in the unit ball of $\Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}C^r(\mathbb{T}^n)/\Pi_{\boldsymbol{K}_{\kappa}^{o,(\ell)}}C^r(\mathbb{T}^n)$ such that for each $\Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}\mathsf{V}\in\mathcal{O}_{\kappa+1}^{(\ell+1)}$ with $\Pi_{\boldsymbol{K}_{\kappa}^{o,(\ell)}}(\Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}\mathsf{V})=\Pi_{\boldsymbol{K}_{\kappa}^{o,(\ell)}}\mathsf{V}$ and $\Pi_{\boldsymbol{K}_{\kappa+1}^{o,(\ell+1)}}\mathsf{V}-\Pi_{\boldsymbol{K}_{\kappa}^{o,(\ell)}}\mathsf{V}\in\mathcal{O}_{\kappa+1}^{(\ell+1)}$, the system $\mathsf{G}_{\kappa+1}^{(\ell+1)}$, $\kappa=0,1,\ldots,\min\{\sharp\mathcal{I}(\ell)-1,n-3\}$ satisfies the following:

(1) *up to finitely many bifurcations, there exists an NHIC homeomorphic to $T^*\mathbb{T}^1$ foliated by Mather sets of the rotation vector $\nu(1,0,\ldots,0)\in H_1(\mathbb{T}^{\kappa+2},\mathbb{R})$, $|\nu|>\lambda$; each Mather set is a periodic orbit, and at each bifurcation point, the Mather set consists of two periodic orbits*;

(2) *the normal hyperbolicity is independent of $\delta^{(\ell+1)}$*;

(3) *there is a generalized transition chain connecting the channels*

$$\mathbb{C}_{\kappa+1,\pm}^{(\ell+1)}:=\{\partial\beta_{\mathsf{G}_{\kappa+1}^{(\ell+1)}}(\nu(1,0,\ldots,0))\mid \pm\nu>\lambda\}\subset H^1(\mathbb{T}^{\kappa+2},\mathbb{R}).$$

Appendix C.4 Dynamics around complete resonances

Suppose for the frequency segment with the subscript $(i,[j])$ that we have completed all the reductions of orders, and hence it becomes the frequency $\boldsymbol{\omega}_{i,[j]}^{\sharp}(a)$, for which there are $(n-2)$ resonant integer vectors $\boldsymbol{k}'_{i,[j]},\ldots,\boldsymbol{k}_{i,[j]}^{(n-2)}$ for all a, and for finitely many a's, there is one more resonant integer vector $\boldsymbol{k}_{i,[j]}^{o}$. We assume that each vector is irreducible. In the above Proposition C.6, we take $\kappa+1=n-2$.

The complete resonance on the energy level $E>\min h$

$$\Sigma(\boldsymbol{K}_{i,[j]}^{o})=\{y\in h^{-1}(E)\mid\langle\boldsymbol{k}_{i,[j]}^{o},\boldsymbol{\omega}(y)\rangle=\langle\boldsymbol{k}_{i,[j]}^{1},\boldsymbol{\omega}(y)\rangle=\cdots=\langle\boldsymbol{k}_{i,[j]}^{(n-2)},\boldsymbol{\omega}(y)\rangle=0\}$$

is a point. We choose $y^\star\in\Sigma(\boldsymbol{K}_{i,[j]}^{o})$ so that $\omega^\star=\boldsymbol{\omega}(y^\star)$ is such a complete resonance point. In the remaining part of this appendix, we omit the subscript $(i,[j])$ for simplicity.

We introduce a matrix $M^\sharp\in\mathrm{SL}(n,\mathbb{Z})$ whose first $n-1$ rows are $\boldsymbol{k}^o,\boldsymbol{k},\ldots,\boldsymbol{k}^{(n-2)}$. We first apply Lemma C.3 to get a Hamiltonian normal form. We next introduce a linear symplectic transformation

$$\mathfrak{M}^\sharp:(x,Y)\mapsto(M^\sharp x,(M^\sharp)^{-t}Y).$$

We set $\mathsf{A}^\sharp=M^\sharp\mathsf{A}(M^\sharp)^t$. The transformed frequency has the form $M^\sharp\omega^\star=(0,\ldots,0,\omega_n)$, $\omega_n\neq0$. Applying the symplectic transformation $\mathfrak{M}^\sharp$ to the normal form, one obtains a Hamiltonian of the following form:

$$H(x,Y)=\frac{1}{\sqrt{\varepsilon}}\omega_nY_n+\frac{1}{2}\langle\mathsf{A}^\sharp Y,Y\rangle+V(x_1,x_2,\ldots,x_{n-1})+\delta^\sharp R(x,Y)\tag{C.6}$$

defined on $T^*\mathbb{T}^n$, where the remainder $R(x,Y)$ is bounded in C^2 and $V=(\mathfrak{M}^\sharp)^{-1*}\Pi_{\boldsymbol{K}_{i,[j]}^{o}}\mathsf{V}$, $\mathsf{V}(\cdot)=P(y^\star,\cdot)$.

Next, we perform a standard energetic reduction to reduce it to a system of $n-1/2$ degrees of freedom. We update the notations $x=(x_1,\ldots,x_{n-1})$ and $y=(Y_1,\ldots,Y_{n-1})$. Removing the last row and the column of $\mathsf{A}^\sharp$, we get a matrix $A^\sharp\in\mathrm{GL}(n-1,\mathbb{R})$. As $\omega_n\neq0$ and $\varepsilon>0$ is very small, one has the function $Y_n(x,x_n,y)$ as the solution of the equation

$$H(x,x_n,y,Y_n(x,x_n,y))=E>\min\alpha_H,$$

which takes the form $Y_n=-Y_{\delta^\sharp}\frac{\sqrt{\varepsilon}}{\omega_n}$, where

$$Y_{\delta^\sharp}=\frac{1}{2}\langle A^\sharp y,y\rangle+V(x_1,\ldots,x_{n-1})+\delta^\sharp\hat{R}\bigg(x,-\frac{x_n\omega_n}{\sqrt{\varepsilon}},y\bigg)\tag{C.7}$$

is defined on $T^*\mathbb{T}^{n-1}\times\mathbb{T}$ and the remainder $\hat{R}(x,\tau,y)$ is bounded in C^2.

Applying Propositions C.5 and C.6 inductively, we get the following result.

Proposition C.7. *There exists an open and dense set $\mathcal{O}$ in the unit ball of $C^r(\mathbb{T}^n)$ such that for each $\mathsf{V}(\cdot) = P(y^\star, \cdot) \in \mathcal{O}$, there exists a $\delta_0^\sharp$ such that for all $0 < \delta^\sharp < \delta_0^\sharp$, there exists an $\varepsilon_0^\sharp = \varepsilon_0^\sharp(\delta^\sharp)$ such that for all $0 < \varepsilon < \varepsilon_0^\sharp$, we have the following for the Hamiltonian system $Y_{\delta^\sharp}$:*

(1) *There exists a collection of NHICs homeomorphic to $T^*\mathbb{T}^1$, restricted to which the time-1 map of the system $Y_{\delta^\sharp}$ is a twist map. Any Mather set with rotation vectors $\omega^\sharp$ lies on the NHICs, if the rotation vector does not lie in the $\lambda\sqrt{\varepsilon}$-neighborhood of $\Sigma(\boldsymbol{K}^o_{i,[j]})$.*

(2) *The normal hyperbolicity is independent of ε or $\delta^\sharp$.*

(3) *There exists a generalized transition chain connecting the two channels*

$$\mathbb{C}^\sharp_\pm := \{\partial\beta_{Y_{\delta^\sharp}}(\nu(1,0,\ldots,0)) \mid \pm\nu > \lambda\} \subset H^1(\mathbb{T}^{n-1},\mathbb{R}),$$

corresponding to two neighboring NHICs.

Appendix C.5 Switching from one frequency line to another

In this appendix, we explain how to move from one frequency segment to the next.

We set up the problem as follows. From the construction, our frequency segments have a hierarchy structure. We consider the switch from $\boldsymbol{\omega}^\sharp_{1,[1]}$ to $\boldsymbol{\omega}^\sharp_{1,[2]}$. For simplicity, we use the subscript $[i]$ instead of $(1,[i])$ for $i = 1, 2$. We need to switch from

$$\boldsymbol{\omega}^\sharp_{[1]}(a) = \rho^\sharp_{[1],a}\omega^{*i}_{1,2}\left(\frac{a}{\omega^*_{1,2}}, \frac{p_2}{q_2}, \ldots, \frac{p_n}{q_n}\right) \quad \text{to} \quad \boldsymbol{\omega}^\sharp_{[2]}(b) = \rho^\sharp_{[2],b}\omega^*_{1,2}\left(\frac{p_{n+1}}{q_{n+1}}, \frac{b}{\omega^*_{1,2}}, \frac{p_3}{q_3}, \ldots, \frac{p_n}{q_n}\right).$$

The switch occurs near the complete resonances $\boldsymbol{\omega}^\sharp_{[1]} \cap \boldsymbol{\omega}^\sharp_{[2]} = (\frac{p_{n+1}}{q_{n+1}}, \frac{p_2}{q_2}, \ldots, \frac{p_n}{q_n})$ up to a positive multiple. When a is moved to $\frac{p_{n+1}}{q_{n+1}}\omega^*_{1,2}$, since $\frac{p_{n+1}}{q_{n+1}}$ is much closer to a Diophantine number than other rational numbers, the new resonance introduced by $\frac{p_{n+1}}{q_{n+1}}$ is a weak resonance and the NHIC $\mathcal{C}(\boldsymbol{k}'_{[1]}, \ldots, \boldsymbol{k}^{(n-2)}_{[1]})$ (homeomorphic to $T^*\mathbb{T}^2$) exists. So moving a to $\frac{p_{n+1}}{q_{n+1}}\omega^*_{1,2}$ along $\boldsymbol{\omega}^\sharp_{[1]}$ is standard as in *a priori* unstable systems. However, it is not clear if it is possible to move b to $\frac{p_2}{q_2}\omega^*_{1,2}$ along $\boldsymbol{\omega}^\sharp_{[2]}$, since $\frac{p_2}{q_2}$ introduces a new strong resonance $\boldsymbol{k}^o_{[2]}$, so the NHIC $\mathcal{C}(\boldsymbol{k}'_{[2]}, \ldots, \boldsymbol{k}^{(n-2)}_{[2]})$ does not exist near $\boldsymbol{\omega}^\sharp_{[1]} \cap \boldsymbol{\omega}^\sharp_{[2]}$.

In the next proposition, we solve the problem by combining and applying repeatedly the c-equivalence mechanism (see Proposition 7.4) and the ladder mechanism (see Lemma 10.6).

Proposition C.8. *Under the assumption of Proposition* C.7*, there exists a generalized transition chain connecting the two channels $\mathbb{C}^\sharp_{i,[j]}(a) := \partial\beta_H(\boldsymbol{\omega}^\sharp_{i,[j]}(a))$ and $\mathbb{C}^\sharp_{i',[j']}(b) := \partial\beta_H(\boldsymbol{\omega}^\sharp_{i',[j']}(b))$ near the complete resonance $\boldsymbol{\omega}^\sharp_{i,[j]} \cap \boldsymbol{\omega}^\sharp_{i',[j']}$, $(i',[j']) = (i,[j+1])$ for $j = 1,\ldots,n-1$, or $(i',[j']) = (i+1,[0])$, $j = n$ and $i = 1,\ldots,M-1$.*

Proof. Without loss of generality, we study only the case of switching from $\boldsymbol{\omega}^\sharp_{[1]}(a)$ to $\boldsymbol{\omega}^\sharp_{[2]}(b)$ as above. All other cases are similar. By the construction in the previous subsection, there exists an NHIC $\mathcal{C}(\boldsymbol{K}_{[1]})$ with $\boldsymbol{K}_{[1]} = \{\boldsymbol{k}'_{[1]}, \ldots, \boldsymbol{k}^{(n-2)}_{[1]}\}$ along the frequency segment $\boldsymbol{\omega}^\sharp_{[1]}(a)$, since by the choice of p_{n+1}/q_{n+1}, the point $\boldsymbol{\omega}^\sharp_{[1]}(a)$ with $a = p_{n+1}/q_{n+1}\omega^*_{1,2}$ is always a point of weak resonance during each reduction of orders along the segment $\boldsymbol{\omega}^\sharp_{[1]}(a)$.

When viewed along the frequency segment $\boldsymbol{\omega}^\sharp_{[2]}(b)$, the complete resonance point $\boldsymbol{\omega}^\dagger := \boldsymbol{\omega}^\sharp_{[1]} \cap \boldsymbol{\omega}^\sharp_{[2]}$ admits an extra resonance $\boldsymbol{k}^o_{[2]}$ which is shorter than any of $\boldsymbol{k}^{(i)}_{[2]}$. So the NHIC $\mathcal{C}(\boldsymbol{K}_{[2]})$ with $\boldsymbol{K}_{[2]} = \{\boldsymbol{k}'_{[2]}, \ldots, \boldsymbol{k}^{(n-2)}_{[2]}\}$ may not exist near the complete resonance $\Sigma(\boldsymbol{k}^o_{[2]}, \boldsymbol{K}_{[2]})$, and the Mather set with the rotation vector $\boldsymbol{\omega}^\dagger$ does not lie on $\mathcal{C}(\boldsymbol{K}_{[2]})$.

We want to move a point on $\boldsymbol{\omega}^\sharp_{[2]}$ to $\boldsymbol{\omega}^\sharp_{[1]}$. The argument goes as follows. We first move along $\mathcal{C}(\boldsymbol{K}_{[2]})$ to arrive at a point $\omega^i \in \boldsymbol{\omega}^\sharp_{[2]}$ with $\mathrm{dist}(\omega^i, \boldsymbol{\omega}(\Sigma(\boldsymbol{k}^o_{[2]}, \boldsymbol{k}'_{[2]}))) < \lambda$. By Proposition 7.7, we get a convex loop $\boldsymbol{\omega}^\dagger + \ell(\boldsymbol{k}^o_{[2]}, \boldsymbol{k}'_{[2]})$ enclosing 0 on the plane $\boldsymbol{\omega}^\dagger + (SM''_{[2]})^{-1}\mathrm{span}\{e_1, e_2\}$ whose Legendre transform is a generalized transition chain of Proposition 7.7 (essentially due to Proposition 7.4). We first find a point ω' on $\boldsymbol{\omega}^\dagger + \ell(\boldsymbol{k}^o_{[2]}, \boldsymbol{k}'_{[2]}) \in (\boldsymbol{k}'_{[1]})^\perp \cap (\partial\alpha(\alpha^{-1}(E)))$.

Complementing $\boldsymbol{K}_{[1]} = \{\boldsymbol{k}'_{[1]}, \ldots, \boldsymbol{k}^{(n-2)}_{[1]}\}$, the rational number p_{n+1}/q_{n+1} introduces one more resonant integer vector denoted by $\boldsymbol{k}^o_{[1]}$ who is much longer than any one in $\boldsymbol{K}_{[1]}$. We introduce a normal form (C.6) at this complete resonance $\boldsymbol{\omega}^\dagger$ as in Appendix C.4. Here, the $n-1$ rows of the matrix $M^\sharp \in \mathrm{SL}(n,\mathbb{Z})$ are ordered as $\boldsymbol{k}'_{[1]}, \ldots, \boldsymbol{k}^{(n-2)}_{[1]}, \boldsymbol{k}^o_{[1]}$. We permute the variables to $x = (x_2, x_3, \ldots, x_n, x_1)$ and $y = (y_2, y_3, \ldots, y_n, y_1)$. In this new coordinate system, the frequency ω' has the form $(0, O(\lambda), \ldots, O(\lambda), O(\varepsilon^{-1/2}))$ since $\omega' \in (\boldsymbol{k}'_{[1]})^\perp \cap (\partial\alpha(\alpha^{-1}(E)))$. As in Subsection 9.1 after a shear transformation S''' in (9.2), we separate a subsystem $\mathsf{G}_{3,\delta}$ (see (9.4)) of three degrees of freedom (corresponding to the first three coordinates) from the full system. We want to kill the second entry $O(\lambda)$. Note that the system $\mathsf{G}_{3,\delta}$ admits an NHIC which is due to the NHIC $\mathcal{C}(\boldsymbol{k}'_{[1]})$ in the original system. Restricted to the NHIC, we get a system $\bar{\mathsf{G}}_{3,\delta}$ (see (9.9)) of two degrees of freedom. We remark that the NHIC here is not near the strong double resonance. By Lemma 10.6 and Remark 10.7, under the generic perturbation, all the cohomology classes on a level set of $\alpha_{\bar{\mathsf{G}}_{3,\delta}}$ lie in a generalized transition chain, along which the frequency vector moves on a convex curve enclosing 0 on the plane $\mathrm{span}\{e_2, e_3\}$. In this way, we kill the second entry $O(\lambda)$ of ω'. Define the resulting frequency ω''. Now ω'' lies on $(\boldsymbol{k}'_{[1]})^\perp \cap (\boldsymbol{k}''_{[1]})^\perp \cap (\partial\alpha(\alpha^{-1}(E)))$. We next perform a shear transformation to separate a subsystem of four degrees of freedom and restricted to its NHIC $\mathcal{C}(\boldsymbol{k}'_{[1]}, \boldsymbol{k}''_{[1]})$, we again get a subsystem of two degrees of freedom of the form $\bar{\mathsf{G}}$ above. We then kill the next $O(\lambda)$ entry using again Lemma 10.6 and Remark 10.7. This procedure can be done repeatedly to obtain a resulting frequency vector having the first $n-2$ entries vanished. In the original coordinates, this means that the frequency is orthogonal to $\boldsymbol{K}_{[1]}$ so it lies on $\boldsymbol{\omega}^\sharp_{[1]}$. The proof is now completed. □

Appendix C.6 Completing the proof of Theorem 2.4

In this appendix, we complete the proof of Theorem 2.4 for general $N \geqslant 3$. The $N = 4$ case was already given in Section 11.

Proof of Theorem 2.4. When the induction in Appendix C is completed, we obtain a collection of frequency segments $\boldsymbol{\omega}^\sharp_{i,[j]}$ ($i = 1, \ldots, M-1$ and $j = 1, \ldots, n$), which is concatenated into a connected curve $\omega(t) : [0, M] \to (\partial\alpha_H)(\alpha_H^{-1}(E))$ lying in the ϱ-neighborhood of the union of $\Omega_{i,[j]}$. Next, the existence of the NHIC (see the parts (2) and (3)(a) of Theorem 2.4) is given by Proposition C.7(1). Neighboring NHICs near a complete resonance are connected by a generalized transition chain by Proposition C.7(2), which proves the part (3)(c) of Theorem 2.4. Next, the existence of the transition chain switching from one frequency line segment to the next is done by Proposition C.8. Propositions C.5–C.7 give the existence of NHICs along which the generalized transition chain can be constructed by a similar proof as in the $n = 4$ case (see Section 11 and Appendix E). This gives the part (3)(b). The cusp-residual genericity follows from the same argument as the $n = 4$ case (see Section 11). This completes the proof. □

Appendix D Variational construction of global connecting orbits

Global connecting orbits are constructed by shadowing a sequence of local connecting orbits. There are two types of local connecting orbits, one is called *type-h* as it looks like a “heteroclinic” orbit, and another one is called *type-c* as it is constructed by using “cohomology equivalence”, corresponding to the assumptions (H1) and (H2) in Definition 2.1, respectively.

Appendix D.1 Local connecting orbits of type-*h* with incomplete intersections

For an Aubry set, if its stable set “intersects” its unstable set transversely, this Aubry set is connected to any other Aubry set nearby by local minimal orbits. It can be thought of as a variational version of Arnold’s mechanism, and the condition of geometric transversality is replaced by the total disconnectedness of minimal points of the barrier function. However, in our case it may happen that the stable set coincides with the unstable set on a subset with nontrivial first homology, which we call an *incomplete intersection*. In this appendix, we design a new method to handle this problem. Let us first formulate a version for the time-periodic dependent Lagrangian.

Recall the definition of the function h_c^∞ introduced in [38]:

$$h_c^\infty(x,x') = \liminf_{k\to\infty} \inf_{\substack{\gamma(-k)=x\\ \gamma(k)=x'}} \int_{-k}^{k} (L(\gamma(t),\dot\gamma(t),t) - \langle c,\dot\gamma\rangle + \alpha(c))dt.$$

This function is closely related to the weak KAM. Indeed, for $x \in \mathcal{A}_{c,i}|_{t=0}$ (the time-1-section of the Aubry class $\mathcal{A}_{c,i} \subset \mathcal{A}(c)$), we have

$$h_c^\infty(x,x') = u_{c,i}^-(x') - u_{c,i}^+(x),$$

where both $u_{c,i}^-$ and $u_{c,i}^+$ are the time-1-sections of the backward and forward elementary weak KAM, respectively (see Appendix A.3 for details). It inspires us to introduce a barrier function for two Aubry classes $\mathcal{A}_{c,i}$ and $\mathcal{A}_{c,j}$:

$$B_{c,i,j}(x) = u_{c,j}^-(x) - u_{c,i}^+(x).$$

Passing through its minimal point, we see that there is a semi-static curve connecting these two classes, provided that this point does not lie in the Aubry set.

If the Aubry set contains only one class, we work in a certain finite covering space so that there are two classes. For example, if the configuration space is $\mathbb{T}^{j+k+\ell}$ and the time-1-section of the Aubry set stays in a neighborhood of a certain lower-dimensional torus, $\mathcal{A}_0(c) \subset \mathbb{T}^{j+\ell} + \delta$, we introduce a covering space $\mathbb{T}^{j+\ell} \times \mathbb{T}^{k-1} \times 2\mathbb{T}$. With respect to this covering space the Aubry set contains two classes.

We introduce some notations and conventions.

Notation 12. (1) For the product space $\mathbb{T}^{j+k+\ell}$, we use $\mathbb{T}^{j+\ell} = \{x \in \mathbb{T}^{j+k+\ell} : x_i = 0, \forall\, i = j+1,\ldots,j+k\}$.

(2) Given a set S, a point x and a number δ, $S+x$ denotes the translation of S by x, i.e., $S+x = \{x'+x : x' \in S\}$ and $S+\delta$ denotes the δ-neighborhood of S, i.e., $S+\delta = \{x : d(x,S) \leqslant \delta\}$.

(3) A set N is called a *neighborhood of* (j,ℓ)*-torus* if it is homeomorphic to an open neighborhood of a $(j+\ell)$-dimensional torus whose first homology group is generated by $\{e_i : i = 1,\ldots,j, j+k+1, \ldots, j+k+\ell\}$.

(4) Given a function B, we use $\mathrm{Argmin}\{B,S\} = \{x \in S : B(x) = \min B\}$ to denote the set of those minimal points of B which are contained in the set S.

Theorem D.1. *For a time-periodic C^2-Lagrangian $L : T\mathbb{T}^{j+k+\ell} \times \mathbb{T} \to \mathbb{R}$ and a first cohomology class $c \in H^1(\mathbb{T}^{j+k+\ell},\mathbb{R})$, we assume the conditions as follows*:

(1) *the Aubry set $\mathcal{A}(c)$ contains two classes $\{\mathcal{A}_{c,i},\mathcal{A}_{c,i'}\}$, which lie in a neighborhood of (j,ℓ)-torus, i.e., $\mathcal{A}_{c,i}|_{t=0} \subset N_i$ and $\mathcal{A}_{c,i'}|_{t=0} \subset N_{i'}$; these neighborhoods are separated, i.e., $\bar N_i \cap \bar N_{i'} = \emptyset$;*

(2) *there exist topological balls $\{O_m \subset \mathbb{T}^{j+k}\}$ with $\bar O_m \cap \bar O_{m'} = \emptyset$ for $m \neq m'$, and each connected component of* $\mathrm{Argmin}\{B_{c,i,i'}, \mathbb{T}^{j+k+\ell}\backslash N_i \cup N_{i'}\}$ *is contained in certain $O_m \times \mathbb{T}^\ell$.*

Then for $c' \in H^1(\mathbb{T}^{j+k+\ell},\mathbb{R})$ satisfying the following conditions:

(1) $\langle c'-c, g\rangle = 0$ *holds,* $\forall\, g \in H_1(\mathbb{T}^{j+k+\ell},\mathbb{T}^{j+k},\mathbb{Z})$ *and* $|c'-c| \ll 1$;

(2) *the Aubry set $\mathcal{A}(c') \subset N_i \cup N_{i'}$,*

there exists an orbit $(\gamma,\dot\gamma)$ connecting $\tilde{\mathcal{A}}(c)$ to $\tilde{\mathcal{A}}(c')$.

Remark D.2. The assumption of this theorem is the nonautonomous version in Definition 2.1(H1). If $\ell = 0$, the set $\mathrm{Argmin}\{B_{c,i,i'}, \mathbb{T}^{j+k+\ell}\backslash N_i \cup N_{i'}\}$ is topologically trivial, and it implies that the stable set intersects the unstable set topologically transversely. Therefore, it turns out to be a variational version of Arnold's mechanism. The case of $\ell > 0$ is a generalization of Arnold's mechanism allowing the separatrix to remain non-splitting on the $\mathbb{T}^\ell$ component.

Proof of Theorem D.1. It is proved by exploiting the upper semi-continuity of the Mañé set with respect to the perturbation on the Lagrangian. As $\mathcal{A}(c') \subset N_i \cup N_{i'}$, without loss of generality we assume $\mathcal{A}(c') \cap N_{i'} \neq \emptyset$.

Given a ball O_m, there exists a small ϵ such that $O_m+\epsilon$ does not touch other balls. Let τ_1: $\mathbb{R}\to[0,\varepsilon]$ be a smooth function such that $\tau_1(t)=0$ for $t\in(-\infty,0]\cup[1,\infty)$, $\tau_1(t)\geqslant 0$ for $t\in[0,1]$ and $\max\tau_1=1$. Let $\tau_2:\mathbb{R}\to[0,1]$ be a smooth function such that $\tau_2(t)=0$ for $t\leqslant 0$ and $\tau_2(t)=1$ for $t\geqslant 1$. Let $v:\mathbb{T}^{j+k+\ell}\to[0,\varepsilon]$ so that $v(x)=0$ if $x\notin(O_m+\epsilon)\times\mathbb{T}^\ell$ and $v(x)=\varepsilon$ if $x\in O_m\times\mathbb{T}^\ell$. As $\langle c'-c,g\rangle=0$ for each $g\in H_1(\mathbb{T}^{j+k+\ell},\mathbb{T}^{j+k},\mathbb{Z})$, there exists a smooth function $u\in\mathbb{T}^{j+k+\ell}\to\mathbb{R}$ so that $\partial u=c'-c$, when it is restricted in $(O_m+\epsilon)\times\mathbb{T}^\ell$ and $\partial u=0$ if $x\notin(O_m+2\epsilon)\times\mathbb{T}^\ell$.

We introduce a modified Lagrangian

$$L_{c,v,u}(\dot x,x,t)=L(\dot x,x,t)-\langle c,\dot x\rangle-\tau_1(t)v(x)-\tau_2(t)\langle c'-c-\partial u,\dot x\rangle$$

and consider the minimizer γ_{k^-,k^+}: $[-k^-,k^+]\to\bar M$ of the action

$$h^{k^-,k^+}_{c,v,u}(x^-,x^+)=\inf_{\substack{\gamma(-k^-)=x^-\\ \gamma(k^+)=x^+}}\int_{-k^-}^{k^+}L_{c,v,u}(\gamma(t),\dot\gamma(t),t)dt+k^-\alpha(c)+k^+\alpha(c'),$$

where $x\in\mathcal{A}_{c,i}\,|_{t=0}$ and $x'\in\mathcal{A}_{c',i'}\,|_{t=0}$. As the Lagrangian is Tonelli, for any large T, the set of the curves $\{\gamma_k\,|_{[-T,T]}:k^-,k^+\geqslant T\}$ is C^2-bounded; therefore it is C^1-compact. Letting $T\to\infty$, by the diagonal extraction argument, we can find a subsequence of γ_{k_i}, which converges C^1-uniformly on each compact interval to a C^1-curve γ: $\mathbb{R}\to\bar M$ which is a minimizer of $L_{c,v,u}$ on any compact interval of $\mathbb{R}$.

Let $\mathscr{C}(L_{c,v,u})$ denote the set of minimal curves of $L_{c,v,u}$. It follows from the above argument that the set $\mathscr{C}(L_{c,v,u})$ is non-empty. Restricted on $(-\infty,0]$ as well as on $[1,\infty)$, each curve in $\mathscr{C}(L_{c,v,u})$ satisfies the Euler-Lagrange equation for L since $\tau_1=0$ and $\langle c'-c-\partial u,\dot x\rangle$ is closed. We are going to show that it also satisfies the equation for $t\in[0,1]$.

If both τ_1 and τ_2 vanish, each curve in the set $\mathscr{C}(L_{c,v,u})$ is nothing else but a c-semi-static curve of L. These curves produce orbits which connect $\mathcal{A}_{c,i}$ to $\mathcal{A}_{c,i'}$. Consider all the semi-static curves which intersect $O_m\times\mathbb{T}^\ell$ at $t=0$. As $O_m\times\mathbb{T}^\ell$ is open, the set of semi-static curves is closed, and there exists a small $t_\delta>0$ such that these curves intersect $O_m\times\mathbb{T}^\ell$ also for $t\in[0,t_\delta]$. If we set $\tau_1=0$ for $t\in(-\infty,0]\cup[t_\delta,\infty)$ and set $\tau_2\equiv 0$, these semi-static curves satisfy the Euler-Lagrange equation produced by $L_{c,v,u}$. As a matter of fact, along these curves the function v keeps constant when $\tau_1\neq 0$, and the term $\tau_1 v$ does not contribute to the equation. Clearly, the action of $L_{c,v,u}$ along these curves is smaller than that along those semi-static curves which do not pass through $O_m\times\mathbb{T}^\ell$ around $t=0$. Since $L_{c,v,u}$ is no longer time-periodic, a time-1-translation of its minimal curve is not necessarily minimal, i.e., $\gamma\in\mathscr{C}(L_{c,v,u})$ does not guarantee $k^*\gamma\in\mathscr{C}(L_{c,v,u})$ for $k\in\mathbb{Z}$, where k^* denotes a translation operator such that $k^*\tau(t)=\tau(t+k)$.

Next, let us recover the term τ_2. Because of upper semi-continuity, the minimal curve of $L_{c,v,u}$ must pass through $O_m\times\mathbb{T}^\ell$ if c' is sufficiently close to c. Again, along these curves, the term $\tau_2\partial u$ does not contribute to the Euler-Lagrange equation, and along these curves $\partial u=c'-c$ when $\tau_2\in(0,1)$.

Obviously, the orbit produced by each curve in the set $\mathscr{C}(L_{c,v,u})$ takes $\tilde{\mathcal{A}}(c)$ as its α-limit set and takes $\tilde{\mathcal{A}}(c')$ as its ω-limit set. □

The orbit $(\gamma,\dot\gamma)$ obtained in this theorem is locally minimal in the following sense.

Local minimum. *There are open balls V_i^- and $V_{i'}^+$ and positive integers t^- and t^+ such that $\bar V_i^-\subset N_i\backslash\mathcal{A}_0(c)$, $\bar V_{i'}^+\subset N_{i'}\backslash\mathcal{A}_0(c')$, $\gamma(-k^-)\in V_i^-$, $\gamma(k^+)\in V_{i'}^+$ and*

$$\begin{aligned}&h_c^\infty(x^-,m_0)+h^{k^-,k^+}_{c,v,u}(m_0,m_1)+h^\infty_{c'}(m_1,x^+)\\&\quad-\liminf_{\substack{k_i^-\to\infty\\ k_i^+\to\infty}}\int_{-k_i^-}^{k_i^+}L_{c,v,u}(d\gamma(t),t)dt-k_i^-\alpha(c)-k_i^+\alpha(c')>0\end{aligned}\tag{D.1}$$

holds, $\forall$ $(m_0,m_1)\in\partial(V_i^-\times V_{i'}^+)$, $x^-\in N_i\cap\pi_x(\alpha(d\gamma))_{t=0}$ and $x^+\in N_{i'}\cap\pi_x(\omega(d\gamma))\,|_{t=0}$, where $k_i^-,k_i^+\in\mathbb{Z}^+$ are the sequences such that $\gamma(-k_i^-)\to x^-$ and $\gamma(k_i^+)\to x^+$.

The set of curves starting from V_i^- and reaching $V_{i'}^+$ with time k^-+k^+ makes up a neighborhood of the curve γ in the space of curves. If it touches the boundary of this neighborhood, the action of $L_{c,v,u}$

along a curve ξ will be larger than the action along γ. The local minimality is crucial in the variational construction of global connecting orbits.

Next, we formulate the theorem for the autonomous Lagrangian. The idea is to treat one of the angular variables as the time and repeat the proof in the time dependent case. Given a first cohomology class, some coordinate system $G_c^{-1}x$ exists such that $\omega_1(\mu) > 0$ for each ergodic c-minimal measure μ if $\alpha(c) > \min\alpha$, where we use $\omega(\mu) = (\omega_1(\mu), \ldots, \omega_n(\mu))$ to denote the rotation vector of the invariant measure. For this purpose, we work in a covering space $\bar{\pi} : \bar{M} = \mathbb{R} \times \pi_{-1}\check{M}$, where π_{-1} denotes the operation to eliminate the first entry, $\pi_{-1}(x_1, x_2, \ldots, x_m) = (x_2, \ldots, x_m)$, the dimension $\mathbb{R}$ is for the coordinate x_1, $\check{M} = \mathbb{T}^{j+\ell} \times \mathbb{T}^{k-1} \times 2\mathbb{T}$ if the Aubry set consists of only one class which stays in a neighborhood of (j,ℓ)-torus and $\check{M} = \mathbb{T}^{j+k+\ell}$ if the Aubry set contains two classes.

Theorem D.3. *For the autonomous C^2-Lagrangian $L : T\mathbb{T}^{j+k+\ell} \to \mathbb{R}$ and the first cohomology class $c \in H^1(\mathbb{T}^{j+k+\ell}, \mathbb{R})$, we assume the conditions as follows*:

(1) $\omega_1(\mu) > 0$ *holds for each ergodic c-minimal measure*;

(2) *the Aubry set $\mathcal{A}(c, \check{M})$ contains two classes $\{\mathcal{A}_{c,i}, \mathcal{A}_{c,i'}\}$, and both stay in a neighborhood of (j,ℓ)-torus, i.e., $\mathcal{A}_{c,i} \subset N_i$ and $\mathcal{A}_{c,i'} \subset N_{i'}$; these neighborhoods are separated, i.e., $\bar{N}_i \cap \bar{N}_{i'} = \emptyset$; the lifts of both N_i and $N_{i'}$ to $\bar{M}$ are still connected and extend to $x_1 = \pm\infty$*;

(3) *there exist topological disks $\{O_m \subset \pi_{-1}(\mathbb{T}^j \times \mathbb{T}^k)\}$ with $\bar{O}_m \cap \bar{O}_{m'} = \emptyset$ for $m \neq m'$ such that each connected component of* Argmin$\{B_{c,i,i'}, \Sigma_0 \backslash N_i \cup N_{i'}\}$ *is contained in certain $\{x_1 = 0\} \times O_m \times \mathbb{T}^\ell$, where $\Sigma_0 = \{x_1 = 0\} \times \pi_{-1}\check{M}$ is a section of $\bar{M}$.*

Then for $c' \in H^1(\mathbb{T}^{j+k+\ell}, \mathbb{R})$ satisfying the following conditions:

(1) $\alpha(c') = \alpha(c)$;

(2) $\langle c' - c, g\rangle = 0$ *holds*, $\forall\, g \in H_1(\mathbb{T}^{j+k+\ell}, \mathbb{T}^{j+k}, \mathbb{Z})$ *and* $|c' - c| \ll 1$;

(3) *the Aubry set $\mathcal{A}(c') \subset N_i \cup N_{i'}$,*

there exists an orbit $(\gamma, \dot{\gamma})$ connecting $\tilde{\mathcal{A}}(c)$ to $\tilde{\mathcal{A}}(c')$.

The proof is similar to the autonomous case and we skip it. Details can be found in [33]. There is an analogous local minimality statement to (D.1) that we also skip.

Remark D.4. For the autonomous system, the barrier function keeps constant along the minimal curve. The intersection of minimal curves of the autonomous system with a codimension-1 section is an analogy of $\mathcal{A}_0(c)$ and $\mathcal{N}_0(c)$ for the time-periodic system.

Appendix D.2 Local connecting orbits of type-*c*

Theorem D.5 (Connecting orbits of type-c)**.** *Assume that the cohomology class c^* is c-equivalent to the class c' through the path Γ: $[0,1] \to H^1(\mathbb{T}^n, \mathbb{R})$. For each $s \in [0,1]$, the following are assumed*:

(1) *there exists a coordinate system $G_s^{-1}x$, where the first component of the rotation vector is positive, and $\omega_1(\mu_{\Gamma(s)}) > 0$ for each ergodic $\Gamma(s)$-minimal measure $\mu_{\Gamma(s)}$*;

(2) *for the covering space $\bar{M}_s = \mathbb{R} \times \mathbb{T}^{n-1}$ in the coordinate system, the lift of the nondegenerately embedded codimension-one torus $\Sigma_{\Gamma(s)}$ has infinitely many connected and compact components, each of which is also a codimension-one torus.*

Then there exist some classes $c^* = c_0, c_1, \ldots, c_k = c'$ on this path such that there exists an orbit $(\gamma, \dot{\gamma})$ connecting $\tilde{\mathcal{A}}(c_i)$ to $\tilde{\mathcal{A}}(c_{i+1})$.

Appendix D.3 Global connecting orbits

In this appendix, we explain how to construct globally connecting orbits from local ones, i.e., Theorem 2.3.

Sketch of the proof of Theorem 2.3. The proof of this theorem is the same as that in [34]. We only sketch the idea of the proof here, and the readers can refer to [34], [17, Section 5] and [18, Section 5] for the details. Because of the condition of the generalized transition chain, there is a sequence $0 = s_0 < s_1 < \cdots < s_k = 1$ such that for each $0 \leqslant j < k$, $\tilde{\mathcal{A}}(\Gamma(s_j))$ is connected to $\tilde{\mathcal{A}}(\Gamma(s_{j+1}))$ by the local minimal orbit either of type-h with incomplete intersections or of type-c. The global connecting orbits are constructed by shadowing such a sequence of orbits.

Recall the construction of the local connecting orbit as above. For each $i \in \{0, 1, \dots, k\}$, let $\eta_i(x, \dot{x}) = \langle c_i, \dot{x} \rangle$ and

$$\mu_i(x, \dot{x}) = w_i \langle \partial(\tau_i \circ s_i'), \dot{x} \rangle, \quad \psi_i(x, \dot{x}) = \chi_i \langle c_{i+1} - c_i - \partial u_i, \dot{x} \rangle$$

in a certain coordinate system $G_i^{-1}x$ such that the first component of the rotation vector of $\tilde{M}(\Gamma(s_i))$ is positive. If it is for type-c, we set $\mu_i = 0$. For each integer k, we introduce a translation operator on functions $k^* f(x_1, x_2, \dots, x_n) = f(x_1 - k, x_2, \dots, x_n)$.

Let $\tilde{\pi}$: $\mathbb{R}^n \to M$ be the universal covering space. For a curve $\tilde{\gamma} : [-K, K'] \to \mathbb{R}^n$, let $\gamma = \tilde{\pi}\tilde{\gamma}$: $[-K, K'] \to M$. Let $\vec{t} = (t_0^-, t_1^\pm, \dots, t_{k-1}^\pm, t_k^+)$ and $\vec{x} = (\tilde{x}_0^-, \tilde{x}_1^\pm, \dots, \tilde{x}_{k-1}^\pm, \tilde{x}_k^+)$ with $t_i^+ < t_i^- < t_{i+1}^+$, $t_0^- = -K$ and $t_k^+ = K'$. We consider the minimal action

$$\begin{aligned} h_L^{K,K'}(m, m', \vec{x}, \vec{t}) = \inf \sum_{i=0}^{k} \int_{t_i^+}^{t_i^-} (L - \eta_i)(d\tilde{\gamma}_i^-(t))dt \\ + \sum_{i=0}^{k-1} \int_{t_i^-}^{t_{i+1}^+} (L - \eta_i - (k_i G_i)^*(\mu_i + \psi_i))(d\tilde{\gamma}_i^+(t))dt, \end{aligned}$$

where the infimum is taken over all the absolutely continuous curves $\tilde{\gamma} : [-K, K'] \to \mathbb{R}^n$ satisfying the boundary conditions $\tilde{\gamma}^-(t_i^-) = \tilde{\gamma}_i^+(t_i^-) = \tilde{x}_i^-$ and $\tilde{\gamma}_i^+(t_{i+1}^+) = \tilde{x}_{i+1}^+$ for $i = 0, 1, \dots, i_{k-1}$, $\gamma(-K) = m$ and $\gamma(K') = m'$. By the carefully set boundary condition, we find that the minimizer is smooth everywhere, along which the term $(k_i G_i)^*(\mu_i + \psi_i)$ does not contribute to the Euler-Lagrange equation. It is guaranteed by the local minimality of (D.1) and setting the translation $k_{i+1} - k_i$ sufficiently large. Therefore, the minimizer produces an orbit $(\tilde{\gamma}, \dot{\tilde{\gamma}})$ which has the properties stated in the theorem. □

Appendix E The proof of genericity

In this appendix, we present a proof of the genericity property of (H1) type generalized transition chains needed in the proofs of (3)(b) and (3)(c) of Theorem 2.4. The theory was well established in [17,18]. We reproduce it here for readers' convenience. Moreover, we need an autonomous version of the argument for the proof of Lemma 10.6 and also a version for the Mañé perturbation, so we include these variants in this appendix.

Appendix E.1 The settings

We consider two settings, the nonautonomous (A) and the autonomous (B) cases:

(A) Given a Tonelli Hamiltonian $H(p, q, t) : T^*\mathbb{T}^n \times \mathbb{T} \to \mathbb{R}$,

(i) there exists an NHIC $\tilde{\Pi}$, which is a deformation of a standard cylinder $\{(p, q, t) \in T^*\mathbb{T}^n \times \mathbb{T} : (\hat{p}_{n-1}, \hat{q}_{n-1}) = 0\}$;

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

(B) Given a Tonelli Hamiltonian $H(p, q) : T^*\mathbb{T}^n \to \mathbb{R}$ and an energy level $E > \min \alpha_H$,

(i) there is a subsystem $G : N \to \mathbb{R}$, where $N \subset T^*\mathbb{T}^n$ is an NHIM of the Hamiltonian flow of H; coordinates can be given such that G is a Tonelli system defined on $T^*\mathbb{T}^2$;

(ii) there exists a continuous path $\Gamma_c : [0, 1] \to H^1(\mathbb{T}^n, \mathbb{R})$ such that for any $c \in \Gamma_c$, the Aubry set entirely lies in the level set $\tilde{\Pi} := G^{-1}(E)$.

Notation 13. (1) Let $\check{\pi} : \check{M} \to \mathbb{T}^n$ be a double covering space of $\mathbb{T}^n$ such that the lift of $\tilde{\Pi}$ to $T^*\check{M} \times \mathbb{T}$ consists of two copies, denoted by $\tilde{\Pi}_\ell$ and $\tilde{\Pi}_r$. For $c \in \Gamma_c$, if the Aubry set $\tilde{\mathcal{A}}(c)$ is an invariant torus $\tilde{\Upsilon}_c \subset \tilde{\Pi}$, its lift also consists of two components, $\tilde{\Upsilon}_{c,\ell} \subset \tilde{\Pi}_\ell$ and $\tilde{\Upsilon}_{c,r} \subset \tilde{\Pi}_r$.

(2) Denote by π the projection such that $\pi(p, q, t) = (q, t)$, and let $\Upsilon = \pi\tilde{\Upsilon}$.

(3) Let $\Gamma_c^* \subset \Gamma_c$ such that

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

We allow the following two types of perturbations:

(a) perturbations depending on all the variables, and

(b) the Mañé perturbation: perturbations depending only on the angular variables.

Definition E.1 (The c-minimal curve and the c-minimal orbit)**.** Given the cohomology class $c \in H^1(M,\mathbb{R})$ where M is a closed manifold, a curve $\gamma:\mathbb{R}\to M$ is called *c-minimal* if for any curve $\xi:\mathbb{R}\to M$ and for any $t_0,t_1,t_1'\in\mathbb{R}$ with $t_1'=t_1 \mod 1$, $\gamma(t_0)=\xi(t_0)$ and $\gamma(t_1)=\xi(t_1')$, one has

$$\int_{t_0}^{t_1}(L(\gamma(t),\dot\gamma(t),t)-\langle c,\dot\gamma(t)\rangle+\alpha(c))dt \leqslant \int_{t_0}^{t_1'}(L(\xi(t),\dot\xi(t),t)-\langle c,\dot\xi(t)\rangle+\alpha(c))dt,$$

where the Tonelli Lagrangian L is assumed time-1-periodic: $L(\cdot,t)=L(\cdot,t+1)$. If a curve γ is c-minimal, then $d\gamma:=(\gamma,\dot\gamma)$ is called a *c-minimal orbit*.

Appendix E.2 The main theorem

Let $B_D\in\mathbb{R}^n$ denote a ball about the origin of radius D. We assume that $D>0$ is suitably large such that for all $c\in\Gamma_c$, the c-minimal orbits of H entirely stay in $B_D\times\mathbb{T}^{n+1}$. Let $\mathfrak{B}_\epsilon\subset C^r(B_D\times\mathbb{T}^{n+1},\mathbb{R})$ (or $\mathfrak{B}_\epsilon\subset C^r(\mathbb{T}^{n+1})$ in the Mañé perturbation case) denote a ball about the origin of radius $\epsilon>0$. In the autonomous case, we define $\mathfrak{B}_\epsilon$ similarly as subsets in $C^r(B_D\times\mathbb{T}^n)$ or $C^r(\mathbb{T}^n)$.

Theorem E.2. *Let H be a C^r Tonelli Hamiltonian as in the above case* (A) *or* (B)*. Then there exists an $\epsilon_0=\epsilon_0(H)$ such that for all $\epsilon<\epsilon_0$ and any small $d_1>0$, there exists a set $\mathfrak{O}$ open and dense in $\mathfrak{B}_\epsilon$ such that for each $H_\delta\in\mathfrak{O}$, it holds for $H+H_\delta$ and simultaneously for all $c\in\Gamma_c^*$ that the diameter of each connected component of the set $\mathcal{N}(c,\check M)\,|_{t=0}\backslash(\mathcal{A}(c,\check M)+\delta)\,|_{t=0}\neq\emptyset$ is not larger than d_1.*

Proof of Theorem E.2 *in the case* (A) *and* (a)*.*

Step 1. *Relating the Mañé set to the minimum of the barrier functions.*

Given an Aubry class for $c\in\Gamma_c$ we can define its elementary weak KAM solution. In the covering space $\check M$, there are two Aubry classes for $c\in\Gamma_c$, $\tilde\Upsilon_{c,\ell}$ and $\tilde\Upsilon_{c,r}$. We introduce the elementary weak KAM solution $u^\pm_{c,\ell}$ for $\tilde\Upsilon_{c,\ell}$ as in Appendix A.3.

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

Definition E.3 (The barrier function)**.** The barrier functions for $c\in\Gamma_c$ are defined as follows:

$$B_c^\ell(q,t)=u^-_{c,\ell}(q,t)-u^+_{c,r}(q,t),\quad B_c^r(q,t)=u^-_{c,r}(q,t)-u^+_{c,\ell}(q,t).$$

In the following, we only study B_c^ℓ. The arguments for B_c^r are the same. The following two lemmas are standard.

Since the backward weak KAM is semi-concave and the forward weak KAM is semi-convex, the barrier function is semi-concave. Therefore, we have the following lemma.

Lemma E.4. *At each minimal point of B_c^ℓ, both $u^-_{c,r}$ and $u^+_{c,\ell}$ are differentiable.*

At the global minimum of B_c^ℓ, we have $\partial u^-_{c,\ell}=\partial u^+_{c,r}$, and hence the backward minimal curve $\gamma^-_{c,x}$ is joined smoothly to the forward minimal curve $\gamma^+_{c,x}$. So we have the following lemma.

Lemma E.5. *If $(q,t)\in\check M\times\mathbb{T}\backslash((\Upsilon_{c,\ell}\cup\Upsilon_{c,r})+\delta)$ is a global minimal point of B_c^ℓ, then $(q,t)\subset\mathcal{N}(c,\check M)$, i.e., passing through the point (q,t) there is a c-semi-static curve in the covering space $\check M\times\mathbb{T}$.*

Step 2. *Localization.*

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

Lemma E.6. *There is a finite partition of Γ_c : $\Gamma_c = \cup I_k$, and each I_k is a segment of Γ_c. For each I_k, there are an annulus $N_k \subset \mathbb{A}_{c,r}|_{t=0}$, and two numbers $\delta > 0$ and $d > 0$ such that for each $c \in I_k \cap \Gamma_c^*$,*

(1) $\mathrm{dist}(N_k, \Upsilon_{c,\ell} \cup \Upsilon_{c,r}) \geqslant \delta$;

(2) *each curve $(\gamma(t), t)$ lying in $(\mathcal{N}(c, \check{M}) \backslash \mathcal{A}(c, \check{M})) \cap \mathbb{A}_{c,r}$ passes through N_k;*

(3) *for each backward (forward) c-minimal curve γ, let $\{q_i = \gamma(2i\pi) \in N_k\}$, then $|q_i - q_j| \geqslant d$ if $i \neq j$.*

Step 3. *The main perturbation lemma.*

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

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

Lemma E.7 (The main perturbation lemma). *For any small $\epsilon > 0$, there is a set $\mathfrak{O}$ open and dense in $\mathfrak{B}_\epsilon$ such that for each $H_\delta \in \mathfrak{O}$, letting $B^\ell_{c,\delta}$ be the barrier function for the Hamiltonian $H + H_\delta$ and the class c, we have that simultaneously for all $c \in I_k \cap \Gamma_c^*$ the set $\mathrm{Argmin}(\mathbb{S}_{d_1}(q^*), B^\ell_{c,\delta})$ is trivial for $\mathbb{S}_{d_1}(q^*)$ provided that $\mathbb{S}_{d_1}(q^*) \subset N_k$ and $d_1 < d/3$ is suitably small.*

Step 4. *Completing the proof.*

Let π_i be the projection so that $\pi_i(q_1, q_2) = q_i$ $(i = 1, 2)$. A connected set V is said to be non-trivial for $\mathbb{S}_{d_1}(q^*)$ if $\pi_i V \cap \mathbb{S}_{d_1}(q^*) = \pi_i \mathbb{S}_{d_1}(q^*)$ holds for $i = 1$ or $i = 2$. Otherwise, it is said to be trivial for $\mathbb{S}_{d_1}(q^*)$. To finish the proof of Theorem E.2, we split the annulus N_k equally into squares $\{\mathbb{S}_j = |q - q_j| \leqslant \frac{d_1}{5}\}$. By Lemma E.7, for each $\mathbb{S}_j$, there exists an open and dense set $\mathfrak{O}_{k,j} \subset \mathfrak{B}_\epsilon$, and for each $H_\delta \in \mathfrak{O}_{k,j}$, it holds simultaneously for all $c \in I_k \cap \Gamma_c^*$ that the set $\mathrm{Argmin}(\mathbb{S}_j, B^\ell_{c,\epsilon})$ is trivial for $\mathbb{S}_j$. The intersection $\cap \mathfrak{O}_{k,j}$ is still open and dense in $\mathfrak{B}_\epsilon$. For each $H_\delta \in \cap_{k,j} \mathfrak{O}_{k,j}$, it holds simultaneously for all $c \in \Gamma_c^*$ that the diameter of each connected component of the Mañé set is not larger than $\frac{4}{5} d_1$ if it keeps away from the Aubry set. □

We prove Lemma E.6.

Proof of Lemma E.6. Because Γ_c is compact, the speed of each c-minimal orbit is uniformly upper bounded for all $c \in \Gamma_c^*$. Given an integer $m > 0$, there will be a small $\delta_c > 0$ such that the period for each c-minimal curve to cross the annulus $N_c = \mathbb{A}_{c,r} \backslash ((\Upsilon_{c,\ell} \cup \Upsilon_{c,r}) + \delta_c)$ is not shorter than $4m\pi$. Because of the upper semi-continuity of the Mañé set in c, there exists some $\delta'_c > 0$ such that $\Upsilon_{c',\ell} \cup \Upsilon_{c',r}$ does not touch N_c and the period for each c'-minimal curve to cross the annulus N_c is not shorter than $2m\pi$ provided that $|c - c'| \leqslant \delta'_c$ and $c' \in \Gamma_c^*$. The first two items are then proved if we notice Γ_c^* is compact.

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

Appendix E.3 The proof of the main perturbation lemma in the case (A) and (a)

In this appendix, we prove Lemma E.7. The proof is based on the following three lemmas whose proofs can be found in [17, Section 6].

The next lemma on the regular dependence on a certain parameter of the invariant circles of the twist map is the key observation to establish the genericity.

Lemma E.8. *There exist a constant C_L and a parametrization $\sigma \mapsto c(\sigma) \in I_k \cap \Gamma_c^*$ such that the invariant curve $\tilde{\Upsilon}_{c(\sigma),0}(q)$ on the NHIC forms a 1/2-Hölder family in the C^0 norm with respect to the parameter σ:*

$$\max_q |\tilde{\Upsilon}_{c(\sigma),0}(q) - \tilde{\Upsilon}_{c(\sigma'),0}(q)| \leqslant \sqrt{2C_L|\sigma - \sigma'|}. \tag{E.1}$$

Each invariant circle corresponds to a unique $c \in \Gamma_c$ such that the Aubry set is the circle. The parameter σ is usually defined on a Cantor set, denoted by Σ. We next use the normal hyperbolicity of the cylinder to extend the Hölder estimate to barrier functions defined on $\mathbb{T}^n$.

Lemma E.9. *For $\sigma, \sigma' \in \Sigma$, let $c = c(\sigma)$ and $c' = c(\sigma')$. If $c, c' \in I_k$ and $q \in N_k$, then*

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

Recall the quantities defined in Lemma E.6 such as the annulus N_k and the number $d > 0$.

Lemma E.10. *For any $\epsilon > 0$ small enough, there exists a δ such that if $S_\delta(q)$ is a C^r-function such that* $\max\{|q - q'| : q, q' \in \text{supp}S_\delta\} \leqslant d$, $\text{supp}S_\delta \subset N_k$ *and* $\|S_\delta\|_{C^r} \leqslant \delta$, *then restricted on I_k, there exists a perturbation $H \to H' = H + H_\delta$ with $\|H_\delta\|_{C^r} < \epsilon$ and the barrier function is subject to a translation*

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

Let us now give the proof of Lemma E.7.

Proof of Lemma E.7. The openness is obvious. To show the denseness, by Lemma E.10, we construct the perturbations $H_\delta \in \mathfrak{B}_\epsilon$ such that the barrier function is under a translation $B_c(q, 0) \mapsto B_c(q, 0) + S_\delta(q)$ for all $c \in I_k \cap \Gamma_c^*$ and $q \in \text{supp}S_\delta$.

Recall the number $d > 0$ defined in Lemma E.6. Given a square $\mathbb{S}_{d_1}(q^*) \subset N_k$ with $3d_1 < d$, we consider the space of C^r-functions $\mathfrak{S}_1$. A function $S \in \mathfrak{S}_1$ if it satisfies the conditions that $\text{supp}S \subset B_{d/2}(q^*)$ and S is constant in q_2 when it is restricted in $\mathbb{S}_{d_1}(q^*)$. Similarly, we can define $\mathfrak{S}_2$ such that $S \in \mathfrak{S}_2$ implies that $\text{supp}S \subset B_{d/2}(q^*)$ and it is constant in q_1 when it is restricted in $\mathbb{S}_{d_1}(q^*)$.

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

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

$$\pi_1 \text{Argmin}(\mathbb{S}_{d_1}(q^*), B_c^\ell + S_\delta) \subsetneqq [q_1^* - d_1, q_1^* + d_1]. \tag{E.2}$$

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

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

where $q^* = (q_1^*, q_2^*)$.

Should the denseness do not hold, there would be a small $\epsilon > 0$, and for each $S_\delta \in \mathfrak{B}_{1,\epsilon}$, some $c \in \Gamma_c^*$ exists such that $B_c^\ell + S_\delta \in \mathfrak{Z}_1$. Let $\mathfrak{B}_{1,\epsilon}^k$ be the intersection of $\mathfrak{B}_{1,\epsilon}$ with a k-dimensional subspace. The box-dimension of $\mathfrak{B}_{1,\epsilon}^k$ in C^0-topology will not be smaller than k.

For any $B_c^\ell \in \mathfrak{F}_c$, there is only one $S_\delta \in \mathfrak{B}_{1,\epsilon}$ such that $B_c^\ell + S_\delta \in \mathfrak{Z}_1$. Otherwise, there would be $S'_\delta \neq S_\delta$ such that $B_c^\ell + S'_\delta \in \mathfrak{Z}_1$ too. We have $B_c^\ell + S'_\delta = B_c^\ell + S_\delta + S'_\delta - S_\delta$ where $B_c^\ell + S_\delta \in \mathfrak{Z}_1$ and $S'_\delta \sim S_\delta$, which contradicts the definition of $\mathfrak{S}_1$. For $S_\delta \in \mathfrak{B}_{1,\epsilon}$, let $\mathfrak{S}_{S_\delta} = \{B_c^\ell \in \mathfrak{F}_c : B_c^\ell + S_\delta \in \mathfrak{Z}_1\}$. Should the denseness do not hold, $\mathfrak{S}_{S_\delta}$ would be non-empty. For any $S_\delta, S'_\delta \in \mathfrak{B}_{1,\epsilon}^k$, each $B_c^\ell \in \mathfrak{S}_{S_\delta}$ and each $B_{c'}^\ell \in \mathfrak{S}_{S'_\delta}$, one has

$$\begin{aligned} d(B_c^\ell, B_{c'}^\ell) &= \max_{q \in \mathbb{S}_{d_1}(q^*)} |B_c^\ell(q, 0) - B_{c'}^\ell(q, 0)| \\ &\geqslant \max_{|q_1 - q_1^*| \leqslant d_1} \Big| \min_{|q_2 - q_2^*| \leqslant d_1} B_c^\ell(q, 0) - \min_{|q_2 - q_2^*| \leqslant d_1} B_{c'}^\ell(q, 0) \Big| \\ &= \max_{|q_1 - q_1^*| \leqslant d_1} |S_\delta(q) - S'_\delta(q)| = d(S_\delta, S'_\delta), \end{aligned} \tag{E.3}$$

where $q = (q_1, q_2)$ and $d(\cdot, \cdot)$ denotes the C^0-metric. It implies that the box-dimension of the set $\mathfrak{F}_c$ is not smaller than the box-dimension of $\mathfrak{B}_{1,\epsilon}^k$ in C^0-topology. Guaranteed by the modulus continuity of Lemma E.9, the box-dimension of the set $\mathfrak{F}_c$ is not larger than 3. Therefore, we obtain an absurdity if we choose $k \geqslant 4$.

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

$$\pi_2 \text{Argmin}(\mathbb{S}_{d_1}(q^*), B_c^\ell + S_\delta) \subsetneq [q_2^* - d_1, q_2^* + d_1]. \tag{E.4}$$

Therefore, there exists an arbitrarily small $S_{i,\delta} \in \mathfrak{B}_{i,\epsilon}$ such that $\pi_i \text{Argmin}(\mathbb{S}_{d_1}(q^*), B_c^\ell + S_{1,\delta} + S_{2,\delta})$ is trivial for $\mathbb{S}_{d_1}(q^*)$ and for all $c \in I_k \cap \Gamma_c^*$. Due to Lemma E.10, we obtain the density. □

Appendix E.4 The autonomous case

To prove Theorem E.2 for autonomous systems, i.e., the combination (B) and (a), we replace Lemmas E.8 and E.9 by the following two theorems respectively.

Theorem E.11 (See [16, Theorem 1.1]). *Let $G : T^*\mathbb{T}^2 \to \mathbb{R}$ be a Tonelli Hamiltonian, the set $\mathcal{E}_E$ be the set of extremal points of the convex set $\bigcup_{E' \leqslant E}\{\alpha_G^{-1}(E')\}$, $E > \min \alpha_G$, and $u_c^\pm : \mathbb{R}^2 \to \mathbb{R}$, $c \in \mathcal{E}_E$ be lifted elementary weak KAM solutions to $\mathbb{R}^2$ normalized by $u_c^\pm(0) = 0$. For a given bounded domain $\Omega \subset \mathbb{R}^2$, there exist a constant $C(\Omega, G)$ depending only on Ω and G, and a one-to-one parametrization of the elementary weak KAM solutions of cohomology classes in $\mathcal{E}_E$ by a number $\sigma \in \Sigma \subset [0, 1]$ such that we have the following Hölder regularity*:

$$\|u^\pm_{c(\sigma)} - u^\pm_{c(\sigma')}\|_{C^0(\Omega)} \leqslant C(\Omega, G)(\|c(\sigma) - c(\sigma')\| + |\sigma - \sigma'|^{\frac{1}{3}}), \quad \forall\, \sigma \in \Sigma \quad \text{and} \quad c \in c(\Sigma) = \mathcal{E}_E.$$

Theorem E.12 (See [16, Theorem 6.1]). *Let $\mathbb{T}^k \times \mathbb{R}^k (\subset \mathbb{T}^n \times \mathbb{R}^n)$, $k < n$ be a normally hyperbolic invariant manifold for the Hamiltonian flow with $k \geqslant 2$ and $u^\pm_{c(\sigma)}$ be elementary weak KAM solutions defined on $\mathbb{T}^n$ for $c(\cdot) : \Sigma \to H^1(\mathbb{T}^k, \mathbb{R})$ continuous and one-to-one, where Σ is a compact subset of $\mathbb{R}^k$. If $\bar{u}^\pm_{c(\sigma)} := \bar{u}^\pm_{c(\sigma)}|_{\mathbb{T}^k}$ is ν-Hölder continuous in σ, then the weak KAM solutions $u^\pm_{c(\sigma)}$ satisfy the following estimate*:

$$\|u^\pm_{c(\sigma)} - u^\pm_{c(\sigma')}\|_{C^0(\mathbb{T}^n)} \leqslant C(\|\sigma - \sigma'\|^\nu + \|c(\sigma) - c(\sigma')\|)$$

for some constant C.

Appendix E.5 Mañé perturbations (b)

To prove Theorem E.2 for Mañé perturbations (b), we replace Lemma E.7 by [13, Theorem 4.2]. Thus, we complete the proofs of all the cases of Theorem E.2.