Hamiltonian Systems
MSRI Publications
Volume **72**, 2024

Some recent developments in Arnold diffusion

CHONG-QING CHENG AND JINXIN XUE

This survey paper grows out of the lecture notes of the nine-hour lectures that the authors delivered in the special semester on Hamiltonian dynamics at MSRI in the fall of 2018. It can be considered as an introduction to our work on Arnold diffusion.

1. Introduction

In the Hamiltonian formalism of the classical mechanics, a smooth Hamiltonian function H on a symplectic manifold (M, ω) is given, and defines a vector field X through $\omega(\cdot, X) = dH$. The main problem is to study the dynamics, which is the long time behavior, of the solution of the differential equation $x' = X(x)$, $x \in M$, determined by the vector field X. The dynamics of a Hamiltonian system in general can be very complicated and deny analytical approaches. From dynamical perspectives, the most well-understood class of Hamiltonian systems is integrable systems. The classical Liouville–Arnold theorem states as follows.

Theorem 1.1 (Liouville–Arnold). *Let $H_1 = H : M^{2n} \to \mathbb{R}$ be a Hamiltonian and suppose there are $H_2, \dots, H_n : M \to \mathbb{R}$ satisfying*:

(a) $\{H_i, H_j\} \equiv 0$, *for all* $i, j = 1, \dots, n$.

(b) *The level set* $M_{\mathbf{a}} := \{(q, p) \in M \mid H_i(q, p) = a_i, i = 1, \dots, n\}$ *is compact.*

(c) *At each point of* $M_{\mathbf{a}}$, *the* n *vectors* DH_i, $i = 1, \dots, n$ *are linearly independent.*

Then:

(1) $M_{\mathbf{a}}$ *is diffeomorphic to* $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ *and is invariant under the Hamiltonian flow of each* H_i.

(2) $M_{\mathbf{a}}$ *is a Lagrangian submanifold, i.e., for any* $u, v \in T_x M_{\mathbf{a}}$, $\forall x \in M_{\mathbf{a}}$, *we have* $\omega(u, v) = 0$.

MSC2020: 00A05.
Keywords: Arnold diffusion.

(3) *In a neighborhood U of $M_{\mathbf{a}}$, there is a symplectic transform $\Phi(q, p) = (\theta, I)$ such that*

$$\Phi(U) = \mathbb{T}^n \times (-\delta, \delta)^n$$

for some $\delta > 0$.

(4) *In the new coordinates, each $K_i := H_i \circ \Phi^{-1}$ is a function of I only so the Hamiltonian equation is*

$$\dot{\theta} = \omega_i(I) := \frac{\partial K_i}{\partial I}, \quad \dot{I} = 0.$$

Integrable systems are also important in algebraic geometry, representation theory, etc; see Hitchin [27]. For the purpose of studying dynamics of Hamiltonian systems, the Liouville–Arnold theorem gives a good description of the dynamics of integrable systems. Each regular level set $M_{\mathbf{a}}$ is an invariant torus under the Hamiltonian flow and the dynamics on it is linear flow. However, integrable systems are very rare. In nature, a system always undergoes some internal or external perturbations. Therefore the next interesting and natural class of Hamiltonian systems is nearly integrable systems which are small perturbations of integrable systems. This class of systems models many interesting natural phenomena including in particular the Newtonian N-body problem. It turns out that this class of systems has rich dynamics and also approachable to a large extent by analytic tools. From the Liouville–Arnold theorem, we see that the natural phase space for studying nearly integrable systems is the symplectic manifold $T^*\mathbb{T}^n$ or its subsets endowed with the standard symplectic structure. We will call a system of the following form a nearly integrable system

$$H(x, y) = h(y) + \varepsilon P(x, y), \quad (x, y) \in \mathbb{T}^n \times \mathbb{R}^n = T^*\mathbb{T}^n \tag{1-1}$$

which gives rise to the Hamiltonian equation

$$\begin{cases} \dot{x} = \partial_y h(y) + \varepsilon \partial_y P(x, y), \\ \dot{y} = -\varepsilon \partial_x P(x, y). \end{cases} \tag{1-2}$$

The natural regularity assumption on H is C^r, $r \geq 2$ including ∞ and ω (meaning analytic).

The celebrated Kolmogorov–Arnold–Moser theorem says that under certain isoenergetic nondegeneracy condition, when ε is sufficiently small, most volume of the phase space is occupied by invariant Lagrangian tori, each of which carries irrational flow with Diophantine frequency. Systems with $n = 1$ are integrable, so are well understood by Liouville–Arnold theorem. For systems with $n = 2$, the KAM theorem gives lots of disjoint two-dimensional tori separating the three dimensional level set, so each orbit is either on a invariant torus or trapped between two nearby tori. We may use the oscillation of the action variable y

along an orbit as a measurement of the instability of the system. In both the cases $n = 1$ and $n = 2$, we see that the oscillation of y along each orbit is small ($o(1)$ as $\varepsilon \to 0$) so we think the systems as being stable. However, the cases of $n \geq 3$ is drastically different. The KAM tori are n-dimensional invariant sets of codimension $n - 1 > 1$ on each energy level set, so the complement of the union of the KAM tori is connected (there may be other Lagrangian tori which are not given by the KAM theorem). This leaves room for the possibility of having orbits wandering in the complements of these tori. Moreover, Arnold constructed an example in which one action variable can indeed oscillate as far as possible. Thus Arnold proposed the following conjecture.

Conjecture [2; 4]. *For any two points y' and y'' on the connected level hypersurface of h in the action space there exist orbits of (1-2) 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.*

The statement can be found in [2; 4], as well as in the book [5] in Problems 1963-1, 1966-3, 1994-33 etc. We make the following remarks concerning the statement of the conjecture.

Remark 1.2. (1) In some circumstances the statement of Arnold talks about only "general" or "typical" systems without specifying the regularity of the Hamiltonian; see [2]. In [4], Arnold considered generic analytic Hamiltonians. The KAM theory applies to both analytic Hamiltonians and smooth (C^r, r large or ∞) Hamiltonians. However, when talking about genericity in differential topology and Riemannian geometry, the C^∞ category is the one used since it allows to construct bump functions and partition of unity. So for this conjecture, the analytic category and the smooth category are essentially different.

(2) The genericity as in this conjecture is not the usual Baire second category, since the smallness parameter ε may depend on the perturbation P. We will introduce a cusp-residual genericity similar to Mather [35].

(3) In [4] Arnold also talked about generic unperturbed part h. In particular, he mentioned Lorentzian type mechanical systems as the first step to study the conjecture for nonconvex h. Our variational method applies only to the convex case i.e., D^2h is positive definite, which includes already lots of physical models, since mechanical systems (kinetic energy plus potential energy) have positive definite kinetic energy part.

The conjecture is in essence asking for an understanding of the global dynamics in the complement of the KAM tori, where the dynamics is expected to be very chaotic but is very resistant to analysis. A related problem called standard map

conjecture states that the system

$$(x, y) \mapsto (x + y + k \sin x, y + k \sin x)$$

has positive Lyapunov exponent for a positive Lebesgue measure set for all or some parameter $k \in \mathbb{R}$. The conjecture expect that the complement of the KAM curves are nonuniformly hyperbolic in the sense of Pesin theory. Having an orbit with wandering action variable is a way to measure the instability. It was foreseen by Arnold [4] that the major difficulty is caused by the double resonance, where the system can be reduced to a nonperturbative mechanical system.

We next state our main result as follows, which is an answer to the above conjecture in the smooth category for convex systems and in the sense of cusp-residual genericity. In the main body of the survey, we will sketch the main ingredients of our proof in a series of papers [16; 9; 8; 10; 11]. Readers are also referred to [14; 15; 6; 18; 19; 37; 7; 35; 28; 24; 32; 37] for other relevant work. Denote by $\mathfrak{S}_1$ the unit sphere in $C^r(T^*\mathbb{T}^n)$ (or $C^r(\mathbb{T}^n)$) with $r \in [7, \infty]$, we have:

Theorem 1.3. *Given any small $\delta > 0$, and finitely many small balls $B_\delta(y_i) \subset \mathbb{R}^n$, where $y_i \in h^{-1}(E)$ with $E > \min h(y)$, there exists a residual set $\mathfrak{C} \subset \mathfrak{S}_1$ such that the following holds for the system (1-1). For each $P \in \mathfrak{C}$ there exists an $\varepsilon_P > 0$, such that there is a residual set of ε in $(0, \varepsilon_P)$, the Hamiltonian flow admits orbits visiting the balls $B_\delta(y_i)$ in any prescribed order.*

The paper is organized as follows. In the main body of the paper, we explain Arnold diffusion in *a priori* unstable systems and the proof of Theorem 1.3 in the case of $n = 3$. We postpone the general $n > 3$ case to Appendix C due to its technicality. Though a bit technical, Appendix C may still serve as a road-map of our paper [11] for readers who want to understand the detailed proof. In Section 2, we explain Arnold's example. In Section 3, we explain the variational method using the pendulum and Arnold's example. We also provide our mechanism of changing cohomology classes. In Section 4, we explain the main difficulties in the proof of Arnold diffusion for *a priori* unstable systems and how to overcome them. In Section 5, we derive the resonant normal form. Section 6 is the important section, which is about mechanical systems with two degrees of freedom. In Section 7, we explain how to construct diffusing orbit in systems with three degrees of freedom, in particular how to overcome the main difficulty of the strong double resonance. In Section 8, we briefly discuss the issue of genericity. Finally, we have three appendices. In Appendix A, we provide the basic concepts in Mather theory and in Appendix B, we provide the theorem of normally hyperbolic invariant manifolds. In Appendix C, we explain how to construct diffusing orbit in systems of arbitrary degrees of freedom.

2. Arnold's example

2A. ***The pendulum.*** The mathematical pendulum is prominent in the study of Arnold diffusion. The Hamiltonian is

$$H_0(x, y) = \tfrac{1}{2}y^2 + (\cos 2\pi x - 1), \quad (x, y) \in T^*\mathbb{T}.$$

First, as a system of one degree of freedom, the Liouville–Arnold theorem can be applied to the regular values of H. Thus the phase space dynamics is further determined by the critical values of H.

Near the fixed points $O = (0, 0)$, the Hamiltonian can be linearized as $y^2/2 - (2\pi x)^2/2$. The linearized Hamiltonian equation is

$$\begin{cases} \dot{x} = y, \\ \dot{y} = 4\pi^2 x, \end{cases}$$

so the fixed point is hyperbolic. Let O be the hyperbolic fixed point and ϕ^t, $t \in \mathbb{R}$, be the flow generated by the Hamiltonian vector field, we define the stable (W^s) and unstable (W^u) manifolds of the fixed point O as

$$W^s(O) = \{z \in T^*\mathbb{T} \mid \phi^t(z) \to O, \text{ as } t \to +\infty\},$$
$$W^u(O) = \{z \in T^*\mathbb{T} \mid \phi^t(z) \to O, \text{ as } t \to -\infty\}.$$

For the pendulum, we see that $W^s(O)$ coincides with $W^u(O)$ consisting of two entire homoclinic orbits denoted by $\{(x_0(t), \pm y_0(t)), t \in \mathbb{R}\}$ with $(x_0(t), \pm y_0(t)) \to (0, 0)$ as $t \to \pm\infty$.

It was discovered by Poincaré that the stable and unstable manifold will split (i.e., will not coincide) if a generic time-periodic perturbation is added. Let us consider the perturbed Hamiltonian

$$H_\varepsilon(x, y, t) = \tfrac{1}{2}y^2 + (\cos 2\pi x - 1) + \varepsilon H_1(x, y, t), \quad (x, y) \in T^*\mathbb{T},$$

where $H_1(x, y, t) = H_1(x, y, t + 1)$ and $\partial_x H_1(0, 0, t) = \partial_y H_1(0, 0, t) = 0$ for all $(x, y, t) \in T^*\mathbb{T} \times \mathbb{T}$. The latter assumption on H_1 implies that O remains a fixed point for the perturbed system H_ε. In this case, the Hamiltonian equation is time-dependent, so its solution is not an $\mathbb{R}$-action on $T^*\mathbb{T}$. Instead, we consider the time-1 map denoted by ϕ_ε^1, whose iterations give rise to a $\mathbb{Z}$-action on $T^*\mathbb{T}$, due to the 1-periodic dependence on t of H_1. We redefine the stable and unstable manifolds as

$$W_\varepsilon^s(O) = \{z \in T^*\mathbb{T} \mid \phi_\varepsilon^n(z) \to O, \text{ as } n \to +\infty\},$$
$$W_\varepsilon^u(O) = \{z \in T^*\mathbb{T} \mid \phi_\varepsilon^n(z) \to O, \text{ as } n \to -\infty\}.$$

The splitting of $W_\varepsilon^s(O)$ and $W_\varepsilon^u(O)$ is one of the main mechanisms responsible for the nonintegrability of the perturbed system. The general method of measuring

the separatrix splitting to the first order is the Melnikov function

$$\begin{aligned} \mathcal{M}(\alpha) &= \int_{\mathbb{R}} \{H_0, H_1\}(x_0(t), y_0(t), t+\alpha)\, dt, \\ \{H_0, H_1\} &= \partial_x H_0 \cdot \partial_y H_1 - \partial_y H_0 \cdot \partial_x H_1. \end{aligned} \tag{2-1}$$

2B. ***Arnold's example.*** Arnold in [1] constructed the following example for which he first discovered the phenomena called now Arnold diffusion

$$\begin{aligned} &H(\theta, I, x, y, t) \\ &\quad = \frac{I^2}{2} + \frac{y^2}{2} + \varepsilon(\cos(2\pi x) - 1)(1 + \mu(\cos(2\pi\theta) + \sin(2\pi t))), \end{aligned} \tag{2-2}$$

where $(\theta, I; x, y; t) \in T^*\mathbb{T}^1 \times T^*\mathbb{T}^1 \times \mathbb{T}$. It is proved in [1] that

Theorem 2.1 (Arnold). *In the system (2-2), for any given $A < B$, $\varepsilon > 0$, there is an orbit $\{(\theta(t), I(t), x(t), y(t))\}$ of the system and time $t_1 < t_2$ with $I(t_1) \le A$ and $I(t_2) \ge B$, provided $\mu > 0$ is small enough.*

We first consider the case of $\mu = 0$. The resulting system has two degrees of freedom. Away from the set $\{y^2/2 + \varepsilon(\cos(2\pi x) - 1) = 0\}$, the system is integrable in the Liouville–Arnold sense.

The product of the hyperbolic fixed point $O = (0, 0)$ of the system H_0 and the phase space of the subsystem $\tilde{H} = I^2/2$ gives rise to the following cylinder in the product space $T^*\mathbb{T}^2$

$$\mathcal{C} = \{(\theta, I, x, y) = (\theta, I, 0, 0), I \in \mathbb{R}, \theta \in \mathbb{T}^1\}.$$

Each circle $\mathcal{C}(I) := \{I = \text{const}, \theta \in \mathbb{T}, x = 0, y = 0\}$ in the cylinder is invariant under the Hamiltonian flow of H_0. When restricted to $\mathcal{C}$, the resulting Hamiltonian system is given by the integrable Hamiltonian $\tilde{H} = I^2/2$. The frequency ω along $\mathcal{C}$ has the form $(I, 0)$ (item (4) of Theorem 1.1), so the cylinder on which the Liouville–Arnold theorem does not apply has resonant frequency, i.e., for all integer vector $k \in \mathbb{Z}^2$ of the form $(0, *)$, we have $\omega \cdot k = 0$. Each circle $\mathcal{C}(I)$ also has stable and unstable manifolds denoted by $W_I^{u,s}$.

When the time-dependent perturbation is turned on, using the Melnikov function (2-1) in the previous subsection, it can be verified that the stable W_I^s and unstable manifolds W_I^u of $\mathcal{C}(I)$ intersect transversely for all I. Therefore the transversality implies that W_I^u intersects $W_{I'}^s$ transversely if I and I' is sufficiently close. Then orbits can be found to shadow a sequence of $W^{u/s}$ chain to have large oscillation of I. We refer readers to [25] for a shadowing lemma developed recently. It is important to point out that the particularly chosen perturbation in Arnold's example gives a vanishing perturbation to the Hamiltonian vector field on the cylinder $\mathcal{C}$, so that the dynamics on $\mathcal{C}$ remains unperturbed. It is not the

case for a generic perturbation, which constitutes the main difficulty for *a priori* unstable systems.

3. The variational method

In this section, we briefly introduce the variational method after Mather [33; 34] and Mañé [31]. Formal definitions of the objects in this theory are summarized in the appendix. Here we only illustrate some of the key points using mainly the pendulum.

3A. ***Variational methods in terms of rotator and pendulum.*** Let $L : T\mathbb{T}^n \to \mathbb{R}$ be a Tonelli Lagrangian system. Let η be a closed 1-form with cohomology class $[\eta] = c \in H^1(\mathbb{T}^n, \mathbb{R})$. We take infimum among all the invariant probability measures μ supported on $T\mathbb{T}^n$

$$-\alpha(c) := \inf_{\mu} \int L(x, \dot{x}) - \eta \, d\mu.$$

We define Mather set as $\widetilde{\mathcal{M}}(c) := \cup \operatorname{supp} \mu$ where the union is taken over all the measures attaining the above infimum.

Let us first give an illuminating example. In the pendulum the hyperbolic fixed point $O = (0, 0)$ the Hamiltonian is linearized as $H_0 = \frac{1}{2}y^2 - (2\pi)^2 x^2$, so after Legendre transform, the corresponding Lagrangian is $L_0 = \frac{1}{2}\dot{x}^2 + (1/(2\pi)^2)x^2$. The probability measure minimizing the action $\inf_\mu \int L_0 \, d\mu$ is easily seen to be the Dirac-δ supported at O. So we see the link

Minimal measure (Dirac-δ)
$\leftrightarrow$ *Nondegenerate global maximum of the Hamiltonian*
$\rightarrow$ *Hyperbolic fixed point.*

This is a guiding principle for us to locate the Mather set with hyperbolicity for nearly integrable systems.

In the example of the mathematical pendulum, the Mather set $\tilde{\mathcal{M}}(c)$ is supported on the hyperbolic fixed point when $c = 0$ and there are orbits homoclinic to the hyperbolic fixed point. In variational methods, we introduce the Aubry set and Mañé set to capture the homoclinic orbits and heteroclinic orbits. The Aubry set $\tilde{\mathcal{A}}(c)$ is the lift to $T^*\mathbb{T}^n$ of the following projected Aubry set

$$\mathcal{A}(c) = \{x \in \mathbb{T}^n \mid h_c(x, x) = 0\},$$

where

$$h_c(x, y) := \limsup_{t\to\infty} \inf_{\gamma} \int_0^t L(\gamma(s), \dot{\gamma}(s)) - \eta + \alpha(c) \, ds$$

and $\gamma : [0, t] \to \mathbb{T}^n$ is a C^1 curve with $\gamma(0) = x$ and $\gamma(t) = y$, $x, y \in \mathbb{T}^n$. On $\mathcal{A}(c)$, we can introduce an equivalence relation: $x \sim y$ if $h_c(x, y) = 0$. Then we get the quotient $\bar{\mathcal{A}}(c)$ called the Aubry class.

Denoting by $\phi^t : T\mathbb{T}^n \to T\mathbb{T}^n$, $t \in \mathbb{R}$, the Lagrangian flow, we can again define the stable and unstable sets analogous to that in hyperbolic dynamics so we can introduce

$$W_c^u = \{z \in T\mathbb{T}^n \mid \phi^t z \to \tilde{\mathcal{A}}(c), t \to -\infty\},$$
$$W_c^s = \{z \in T\mathbb{T}^n \mid \phi^t z \to \tilde{\mathcal{A}}(c), \ t \to +\infty\}.$$

Here we use the notations $W_c^{u,s}$, though the Aubry set $\tilde{\mathcal{A}}(c)$ may not be hyperbolic.

In case when the Aubry class consists of a single point, these sets are defined as graphs of the gradients of the backward/forward weak KAM solutions u_c^- and $u_c^+ : \mathbb{T}^n \to \mathbb{R}$ respectively, which are known to be unique up to an additive constant. We call the difference $B_c(x) = u_c^-(x) - u_c^+(x)$ the barrier function, whose critical points correspond to the intersection of W_c^u and W_c^s. If $z = (x, y) \in W_c^u \cap W_c^s$ and x is a global minimal point of B_c, then by definition $\phi^t(z)$ approaches $\widetilde{\mathcal{M}}(c)$ in both the future and the past, such an orbit is a prototypical orbit in the Mañé set $\tilde{\mathcal{N}}(c)$. The barrier function is in general only known to be Lipschitz, however, it has the remarkable property of being differentiable at its global minimal points. We refer readers to the appendix for formal definitions of the Mather set $\widetilde{\mathcal{M}}(c)$, Aubry set $\tilde{\mathcal{A}}(c)$, Mañé set $\tilde{\mathcal{N}}(c)$ and weak KAM solutions $u_c^{\pm}$ and their basic properties. We also refer readers to [13] for how to realize these objects in the pendulum.

We next explain the effect of changing cohomology class. In Liouville–Arnold theorem, the action variable I is constructed by integrating the Liouville 1-form pdq along a basis of the first homology group $H_1(M_{\mathbf{a}}, \mathbb{Z})$. In variational methods, the changing of the cohomological class has the effect of selecting the corresponding action variable for integrable systems hence selecting the Lagrangian torus. Let $H(I) : T^*\mathbb{T}^n \to \mathbb{R}$ be a convex integrable Hamiltonian independent of the angular variable $\theta \in \mathbb{T}^n$. The corresponding Lagrangian is denoted by $L(\dot{\theta})$. We next show how to find the minimizer of the variational problem $\inf_\mu \int L(\dot{\theta}) - \eta \, d\mu$ with $[\eta] = c \in H^1(\mathbb{T}^n, \mathbb{Z}) \in \mathbb{R}^n$. For simplicity, we take $\eta = cd\theta$, so the minimization problem is solved by Legendre transform as

$$-\alpha(c) := \inf_{\dot{\theta} \in \mathbb{R}^n} L(\dot{\theta}) - c \cdot \dot{\theta} = -H(c).$$

The infimum is attained as the point $\dot{\theta} = \partial_c H(c)$ and $c = \partial_{\dot{\theta}} L(\dot{\theta}) = I$. So we see that for integrable systems the cohomology class c agrees with the action variable I, the Mather set is the corresponding invariant torus $\{\dot{\theta} = \partial_c H(c)\} \times \mathbb{T}^n \subset T\mathbb{T}^n$.

3B. ***Arnold's example in the variational language.*** Recall that in Arnold's example, orbits are found to shadow different circles $\mathcal{C}(I)$s. Each circle can be realized as the Aubry set with cohomology class $c = (I, 0) \in H^1(\mathbb{T}^2, \mathbb{R})$. Thus in variational terms, Arnold diffusing orbit corresponds to orbits shadowing different Aubry sets $\tilde{\mathcal{A}}(c)$s. To find orbit shadowing Aubry sets with different cohomology classes, we have the following variational version of Arnold's mechanism of constructing diffusing orbit using the intersection of the stable and unstable manifolds.

Theorem 3.1 (Type-h orbit). *Let $\Gamma : [0, 1] \to H^1(M, \mathbb{R})$ be a continuous curve. Suppose there exist a certain finite covering $\breve{\pi} : \breve{M} \to M$, two open domains $N_1, N_2 \subset \breve{M}$ with $d(N_1, N_2) > 0$, and for each $s \in [0, 1]$, there exist a codimenion one disk D_s and small numbers $\delta_s, \delta'_s > 0$ such that*:

(1) *The projected Aubry sets satisfy $\mathcal{A}(\Gamma(s)) \cap N_1 \neq \varnothing$, $\mathcal{A}(\Gamma(s)) \cap N_2 \neq \varnothing$ and $\mathcal{A}(\Gamma(s')) \cap (N_1 \cup N_2) \neq \varnothing$ for each $|s' - s| < \delta_s$.*

(2) *$\breve{\pi}\mathcal{N}(\Gamma(s), \breve{M})|_{D_s} \setminus (\mathcal{A}(\Gamma(s)) + \delta'_s)$ is nonempty and totally disconnected, where the $+\delta'_s$ notation means a δ'_s neighborhood.*

Then there exists an orbit $d\gamma \subset T^*M$ such that $\boldsymbol{\alpha}(\gamma) \subset \tilde{\mathcal{A}}(\Gamma(0))$ and $\boldsymbol{\omega}(\gamma) \subset \tilde{\mathcal{A}}(\Gamma(1))$.

We call orbits in the theorem type-h, standing for *heteroclinic*. Bernard [6] introduced a similar variational mechanism called forcing relations. The way the theorem applies to Arnold's example is as follows. We treat the x-variable in (2-2) as being defined on $\mathbb{R}/(2(2\pi\mathbb{Z}))$, in other words, we lift the pendulum component to the double covering space of $T^*\mathbb{T}$. Thus, the Aubry set for each cohomology class $c = (I, 0)$ has two copies and the second assumption is satisfied due to the transversal intersection of the stable manifold of one component of the Aubry set and the unstable manifold of the other component, for each cohomology class $(I, 0)$, hence the theorem applies to Arnold's example. This advantage of the lifting procedure here is that it produces orbits in the Mañé set but not in the Aubry set. The last point is subtle, and we refer readers to [13] for the description of Mañé set and Aubry set in the pendulum example.

Summarizing the above, we have the following dictionary:

hyperbolic objects	variational objects
hyperbolic set	Aubry set
stable/unstable manifold	graph of differential of weak KAM
intersection of stable/unstable manifolds	critical point of the barrier function
homo- or hetero-clinic orbits	Mañé set\Mather set
Arnold's orbit shadowing different $\mathcal{C}(I)$s	type-h orbit

The main advantage of the variational methods is that the existence of the variational objects such as Mather sets, Aubry sets and Mañé sets are always assured. Without requiring good regularities of the variational objects, diffusing orbits can still be constructed.

This dictionary goes far beyond Arnold's example and *a priori* unstable systems. In the next subsection, we will introduce a new way of changing cohomology classes, which will be used crucially in the proof of Arnold diffusion in *a priori* stable systems, in addition to Arnold's mechanism.

3C. ***Type-c orbit.*** In this section, we introduce the second way of finding orbits shadowing Aubry sets with different cohomology classes, which we call type-c, standing for *cohomology equivalence*. The basic idea is that the cohomology class can be changed in the orthogonal complement of the homology of a section of the Mañé set. This mechanism first appeared in [34] proved in [14] for nonautonomous systems. The version for autonomous systems that we are going to give here was first established in [29; 30]; see Section 3.1 of [8].

We suppose that there exists Σ_c nondegenerately embedded $(n-1)$-dimensional torus on $\mathbb{T}^n$ given by an embedding $\varphi\colon \mathbb{T}^{n-1} \to \mathbb{T}^n$ with $\Sigma_c = \varphi(\mathbb{T}^{n-1})$ the image of φ, and the induced map $\varphi_*\colon H_1(\mathbb{T}^{n-1}, \mathbb{Z}) \hookrightarrow H_1(\mathbb{T}^n, \mathbb{Z})$ is an injection. We can simply choose Σ_c in the nonautonomous setting to be the configuration space with $\{t = 0\}$.

Let $\mathfrak{C} \subset H^1(\mathbb{T}^n, \mathbb{R})$ be a connected set. For each class $c \in \mathfrak{C}$, we assume that there exists a nondegenerate embedded $(n-1)$-dimensional torus $\Sigma_c \subset \mathbb{T}^n$ such that each c-semistatic curve γ transversally intersects Σ_c. Let

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

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

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

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

Definition 3.2 (c-equivalence). We say that $c, c' \in H^1(M, \mathbb{R})$ are *c-equivalent* if there exists a continuous curve $\Gamma\colon [0, 1] \to \mathfrak{C}$ such that $\Gamma(0) = c$, $\Gamma(1) = c'$, $\alpha(\Gamma(s))$ keeps constant for all $s \in [0, 1]$, and for each $s_0 \in [0, 1]$ there exists $\epsilon > 0$ such that $\Gamma(s) - \Gamma(s_0) \in \mathbb{V}_{\Gamma(s_0)}^\perp$ whenever $s \in [0, 1]$ and $|s - s_0| < \epsilon$.

Theorem 3.3 (Type-c orbit). *Suppose that two cohomology classes c and c' are c-equivalent, then there exists an orbit whose $\boldsymbol{\alpha}$-limit set is in $\tilde{\mathcal{A}}(c)$ and $\boldsymbol{\omega}$-limit set is $\tilde{\mathcal{A}}(c')$, and vice versa.*

Proof. We only prove the case when $|c - c'|$ is sufficiently small. The details of changing c in the large scale (global connecting orbits) can be found in Section 5 of [14]. We first denote by U^c a small neighborhood of $\mathcal{N}(c) \cap \Sigma_c$. We next modify the Lagrangian $L_c := L(\gamma, \dot{\gamma}) - \langle c, \dot{\gamma} \rangle$ to $L_{c+\eta\rho} := L(\gamma, \dot{\gamma}) - \langle c + \eta\rho(t), \dot{\gamma} \rangle$ where η is a de Rham closed one-form with cohomology class $[\eta] = c' - c$ whose support lies in U^c and $\rho(t) \in C^\infty$ satisfies $\rho(t) = 0$ for $t \leq 0$ and $\rho(t) = 1$ for $t \geq \varepsilon$ for ε small. Such a closed one-form exists following from the definition of the c-equivalence, $c - c' \perp H_1(\mathcal{N}(c)|_\Sigma, \mathbb{R})$. The free time global minimizer (defined as the semistatic curves, see appendix) of the action $\int L_{c+\eta\rho}\, dt$ is taken over all the curves with endpoints in $\widetilde{\mathcal{M}}(c)$ and $\widetilde{\mathcal{M}}(c')$. First it is known that the minimizer stays close to the Mañé set $\tilde{\mathcal{N}}(c)$ if $|c - c'|$ is small enough so it also passes through U. The proof of this fact is the same as that of the upper-semicontinuity of the Mañé set; see Section 2 of [14]. We claims that the minimizer satisfies the Euler–Lagrange equation. As we know, adding a closed 1-form to the Lagrangian does not change the E-L equation. If we keep track of the orbit, before entering U, the Lagrangian is L_c. In U, since $\mathrm{supp}\eta \cap U = \varnothing$, the Lagrangian is still L_c. When the orbit gets outside of U, for ε small enough, the Lagrangian is now actually $L_{c'}$. In all the cases, the E-L equation is the same as that of L so we have constructed an orbit $d\gamma$. Since orbits in $\tilde{\mathcal{N}}(c) \setminus \tilde{\mathcal{M}}(c)$ does not recur, as $t \to \infty$, the orbit $d\gamma$ stays in a region where the cohomology class is $\eta(c')$ hence the $\boldsymbol{\omega}$-limit set is $\tilde{\mathcal{A}}(c')$ and similarly, the $\boldsymbol{\alpha}$-limit set is $\tilde{\mathcal{A}}(c)$. □

4. a priori unstable systems

We now explain the main difficulty of constructing diffusing orbit in the so-called *a priori* unstable system, which are generalizations of Arnold's example but maintaining the structure of normally hyperbolic invariant cylinder (NHIC). We refer readers to the appendix for the definition of normally hyperbolic invariant manifold (NHIM) and a theorem on its persistence under perturbations. A prototypical form of the *a priori* unstable system is

$$H = \frac{I^2}{2} + \frac{y^2}{2} + (\cos(2\pi x) - 1) + \varepsilon P(\theta, I, x, y, t), \tag{4-1}$$

where $(\theta, I, x, y, t) \in T^*\mathbb{T}^2 \times \mathbb{T}^1$. This kind of system appears as the single resonance normal form (see Section 5 below), thus the following problem is the first step towards the conjecture of Arnold diffusion.

For C^r-generic P with $2 \le r \le \infty$ and for any $A < B$, the system (4-1) admits an orbit $\{(\theta(t), I(t), x(t), y(t)), t \in \mathbb{R}\}$ and t_1, t_2 such that $I(t_1) \le A$ and $I(t_2) \ge B$ provided ε is sufficiently small.

The problem is solved by different authors using different methods. In this section, we briefly explain the main difficulties and the solution of Cheng and Yan [14; 15].

First, note that when $\varepsilon = 0$, the system admits a NHIC given by

$$\mathcal{C} = \{x = y = 0, (\theta, I) \in \mathbb{T} \times \mathbb{R}\}.$$

In order to apply the theorem of NHIM, we first replace the perturbation εP by $\varepsilon \chi P$ where $\chi : T^*\mathbb{T}^2 \to \mathbb{R}$ such that $\chi = 1$ for $|I| < R$ and $|y| < 10$, and $\chi = 0$ for $|I| > R + 1$ and $|y| > 11$ for some large R with $R > \max\{|A|, |B|\}$. Our orbit will stay within the region where $\chi = 1$ so it is also an orbit of the original system. For the perturbed system, we shall consider the time-1 map denoted by $\phi_\varepsilon^1 : T^*\mathbb{T}^2 \to T^*\mathbb{T}^2$. By the theorem of NHIM so we get a NHIC $\mathcal{C}_\varepsilon$ close to $\mathcal{C}$ and is invariant under ϕ_ε^1. Restricted to $\mathcal{C}_\varepsilon$, the map ϕ_ε^1 is a twist map, so we can then apply KAM theorem to get that for ε small enough, there are uncountably many invariant circles on $\mathcal{C}_\varepsilon$ that are homologically nontrivial. In general there are also other homologically nontrivial invariant circles that are not given by the KAM theorem.

Here comes the first main difficulty. The distances between two neighboring circles may be of order $\sqrt{\varepsilon}$. However, the size of separatrix splitting is only of order ε. This means that Arnold's mechanism of utilizing the intersection of stable and unstable manifolds fails to find orbit crossing the $\sqrt{\varepsilon}$-gaps. This is called the "big gap" problem.

The way to overcome this problem is to invoke the cohomology equivalence mechanism in Section 3C. The reason is that for each $c \in H^1(\mathbb{T}^2, \mathbb{R})$ such that $\tilde{\mathcal{N}}(c)$ lies in the gap, the Mañé set is contractible so Definition 3.2 is verified. Now the general strategy is to apply the c-equivalence mechanism whenever there is a big gap and to apply Arnold's mechanism (variationally type-h orbit) whenever nearby invariant circles are close enough to have transversely intersecting stable and unstable manifolds.

Here comes the second main difficulty. There are uncountably many invariant circles for which we want their stable and unstable manifolds to intersect transversely in order to implement Arnold's mechanism. It is easy to add a perturbation to create the intersection for one such circle. However, it is not allowed to add uncountably many perturbations for the consideration of genericity.

The key to this problem is the following regularity result, which holds for nearly integrable twist maps on $T^*\mathbb{T}^1$ or equivalently nearly integrable Hamiltonian systems of one and a half degrees of freedom.

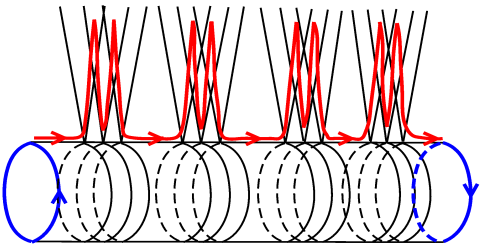

Figure 1. Phase space dynamics of the mechanism in [14; 15].

Theorem 4.1. *Let $H : T^*\mathbb{T} \times \mathbb{T} \to \mathbb{R}$ be a time-dependent nearly integrable Tonelli Hamiltonian system and $u_c^{\pm} : \mathbb{T} \to \mathbb{R}$, $c \in H^1(\mathbb{T}, \mathbb{R})$ be its weak KAM solutions. Then for all c, c' in a bounded subset of $H^1(\mathbb{T}, \mathbb{R})$ and such that $\widetilde{\mathcal{M}}(c)$ and $\widetilde{\mathcal{M}}(c')$ are invariant circles, there is a uniform constant C such that*

$$\|u_c^{\pm}(\cdot) - u_{c'}^{\pm}(\cdot)\|_{C^0} \le C\|c - c'\|^{1/2}.$$

A similar regularity result holds for the barrier function of the full system. We thus see that the set $\mathcal{S}$ of all weak KAM solutions corresponding to invariant circles on $\mathcal{C}_\varepsilon$ is a set of finite box dimension in $C^0(\mathbb{T}^2, \mathbb{R})$. We have seen in Section 3A that the intersection of the stable and unstable manifolds can be interpreted as the minimal point of the barrier function. Thus we need the barrier function to be nonconstant outside $\mathcal{C}_\varepsilon$. Note that a C^0 function $f : \mathbb{T} \to \mathbb{R}$ being constant on an interval $J \subset \mathbb{T}$ is of infinite codimension. Since $\mathcal{S}$ has finite box dimension, it is easy to find an arbitrarily small $\tilde{u}$ such that the entire set $\tilde{u} + \mathcal{S}$ avoids the infinite codimensional space of functions that are constants over some sets of the form $\mathbb{T} \times J$, where $J \subset \mathbb{T}$ is an interval. In this way, we thus have verified assumption (2) of Theorem 3.1 and can facilitate Arnold's mechanism. We refer readers to [14; 15] for details of this argument. We finally emphasize that the regularity of the weak KAM solutions of the form 4.1 is the essential ingredient in the proof of the genericity.

We finally remark on the literature. The use of box dimension to the genericity problem goes back to Moeckel [36] in which the author studied the iteration of a pair of twist maps where the regularity problem is straightforward. The regularity result Theorem 4.1, its generalization to the full system and the genericity argument for *a priori* unstable systems first appeared in [14; 15]. Bernard [6] gave a different mechanism for constructing diffusing orbit using only Arnold's mechanism designed for the variational objects, without a genericity argument. The regularity result adapted to the mechanism of [6] was given in [38].

5. The normal form

When applied to nearly integrable systems, the variational method is greatly enhanced by the normal form theory. The basic objects such as Mather sets, Aubry sets and Mañé sets are invariant under symplectomorphisms. In the normal form theory, we will apply a symplectic transformation to reduce the Hamiltonian to a normal form to reveal the rotator-pendulum structure. We have seen from Arnold's example that the rotator-pendulum structure is intimately related to the appearance of resonances. The normal form theory reveals this link in this section.

5A. ***Homogenization.*** For nearly integrable systems of the form (1-1), the natural scale to work with is $\sqrt{\varepsilon}$ in the space of action variables. In this section, we introduce a procedure called homogenization used to blow up a $O(\sqrt{\varepsilon})$ ball in the space of action variables to the unit size. The main outcome of the homogenization procedure is a mechanical system with a fast drift and a small perturbation.

Consider an autonomous Hamiltonian H defined on $T^*\mathbb{T}^n$. Picking a point $y^\star \in \mathbb{R}^n$, we introduce the homogenization operator

$$\mathfrak{H}: \quad y - y^\star := \sqrt{\varepsilon} Y, \quad H(x, y) = \varepsilon \mathsf{H}(x, Y), \tag{5-1}$$

where Y, τ, H are the homogenized action variable and Hamiltonian respectively. We will simultaneously rescale the time t to the new time $\tau = t\sqrt{\varepsilon}$, The Hamiltonian (1-1) becomes

$$\mathsf{H}(x, Y) = \tfrac{h(y^\star)}{\varepsilon} + \tfrac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \tfrac{1}{2}\langle \mathsf{A}Y, Y\rangle + \mathsf{V}(x) + \mathsf{P}(x, \sqrt{\varepsilon} Y), \tag{5-2}$$

where:

(1) $\tfrac{h(y^\star)}{\varepsilon} + \tfrac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle + \tfrac{1}{2}\langle \mathsf{A}Y, Y\rangle$ is the first three terms of the Taylor expansion of $h(y)$ around $y^\star$.

(2) $\omega^\star = \tfrac{\partial h}{\partial y}(y^\star)$.

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

(4) $\mathsf{V}(x) = P(x, y^\star)$ is the constant term in the Taylor expansion of $P(x, y)$ with respect to the variable y.

(5) $\mathsf{P}(x, \sqrt{\varepsilon} Y)$ consists of all the remaining terms and we have the estimate $\|\mathsf{P}\|_{C^r} = O(\sqrt{\varepsilon})$ if $\|Y\| < C$.

5B. ***Normal form.*** We next state a normal form proposition for the homogenized system. Simply put, the normal form deduces the rotator+pendulum structure from a resonance.

Definition 5.1. A frequency vector $\omega \in \mathbb{R}^n \setminus \{0\}$ is said to be *resonant* if we have $\langle \omega, \mathbf{k}_i \rangle = 0$ for some linearly independent $\mathbf{k}_1, \dots, \mathbf{k}_m \in \mathbb{Z}^n \setminus \{0\}$, $1 \le m \le n-1$. The number m is called the *multiplicity of the resonance*. We call ω a *complete resonance* if $m = n-1$, in which case ω is a nonzero multiple of a rational vector.

Proposition 5.2 [11, Proposition 3.10]. *For any $\delta > 0$, there exists ε_0 such that for all $\varepsilon < \varepsilon_0$ the following holds. Suppose $\omega^\star \in \mathbb{R}^n$ admits m independent resonance relations $\langle \mathbf{k}_i, \omega^\star \rangle = 0$ with $|\mathbf{k}_i| \le \delta^{-1/2}$, $i = 1, \dots, m$, and $|\langle \mathbf{k}, \omega^\star \rangle| > \varepsilon^{1/3}$ for any $|\mathbf{k}| \le \delta^{-1/2}$ and $\mathbf{k} \notin \operatorname{span}_{\mathbb{Z}}\{\mathbf{k}_1, \dots, \mathbf{k}_m\}$. Then there exists a symplectic transformation ϕ, which is $o_\varepsilon(1)$ close to identity in the C^{r-2} norm in the domain $\{|Y| \le 1\}$, such that the Hamiltonian system (5-2) is transformed to the following*

$$\begin{aligned} &\mathsf{H} \circ \phi(x, Y) \\ &\quad = \tfrac{1}{\varepsilon} h(y^\star) + \tfrac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \tfrac{1}{2} \langle \mathsf{A} Y, Y \rangle + V(\langle \mathbf{k}_1, x \rangle, \dots, \langle \mathbf{k}_m, x \rangle) + \delta \mathsf{R}(x, Y), \end{aligned} \tag{5-3}$$

where:

(1) *V consists of all the Fourier modes of V in $\operatorname{span}_{\mathbb{Z}}\{\mathbf{k}_1, \dots, \mathbf{k}_m\}$.*

(2) *The remainder $\delta \mathsf{R}(x, Y) = \mathsf{R}_I(x) + \sqrt{\varepsilon} \mathsf{R}_{II}(x, Y)$, where $\delta \mathsf{R}_I$ consists of all the Fourier modes in V with $|\mathbf{k}| > \delta^{-1/2}$.*

(3) *If the perturbation P in (1-1) satisfies $\|P(x, y)\|_{C^r} \le 1$, then the norms of V, R_I, R_{II} satisfy $\|V\|_{C^r}$, $\|\mathsf{R}_I\|_{C^{r-2}}$, $\|\mathsf{R}_{II}(x, Y)\|_{C^{r-2}} \le 1$.*

Sketch of proof. We sketch an argument to give the main idea of the proof and refer readers to [11]. We consider the pullback of H by the time-1 map $\phi^1_{\sqrt{\varepsilon} F}$ of another Hamiltonian $\sqrt{\varepsilon} F$ to be determined. Then we get by the definition of the Poisson bracket $\left(\frac{d}{dt}|_{t=0} \mathsf{H}(\phi^t_{\sqrt{\varepsilon} F})\right) := \{\mathsf{H}, \sqrt{\varepsilon} F\} = \frac{\partial \mathsf{H}}{\partial x} \frac{\partial \sqrt{\varepsilon} F}{\partial y} - \frac{\partial \mathsf{H}}{\partial y} \frac{\partial \sqrt{\varepsilon} F}{\partial x}$ and Taylor expansion that

$$\begin{aligned} \mathsf{H} \circ \phi^1_{\sqrt{\varepsilon} F} &= \mathsf{H} + \{\mathsf{H}, \sqrt{\varepsilon} F\} + \varepsilon \int_0^1 (1-t) \{\{\mathsf{H}, F\}, F\} \circ \phi^t_F \, dt \\ &= \tfrac{1}{\varepsilon} h(y^\star) + \tfrac{1}{\sqrt{\varepsilon}} \langle \omega^\star, Y \rangle + \tfrac{1}{2} \langle \mathsf{A} Y, Y \rangle + \mathsf{V}(x) + \langle \omega^\star, \partial_x F \rangle + O(\varepsilon^{1/3}). \end{aligned} \tag{5-4}$$

We next decompose $\mathsf{V}(x) = V + \tilde{V}$ where V consists of all the Fourier modes of V in $\operatorname{span}_{\mathbb{Z}}\{\mathbf{k}_1, \dots, \mathbf{k}_m\}$, and $\tilde{V}$ consists of the rest. We further decompose $\tilde{V} = \tilde{V}_1 + \tilde{V}_2$ where $\tilde{V}_1$ consists of those Fourier modes with $|\mathbf{k}| \le \delta^{-1/2}$ and $\tilde{V}_2$ consists of the rest. Note that $\tilde{V}_2$ has C^{r-2} norm less than δ by the decay of Fourier coefficients. Then we can solve the equation $\tilde{V}_1 + \langle \omega^\star, \partial_x F \rangle = 0$ by taking Fourier expansion and using the assumption $|\langle \mathbf{k}, \omega^\star \rangle| > \varepsilon^{1/3}$ for any $\mathbf{k} \in \mathbb{Z}^n$ with $|\mathbf{k}| \le \delta^{-1/2}$ and $\mathbf{k} \notin \operatorname{span}_{\mathbb{Z}}\{\mathbf{k}_1, \dots, \mathbf{k}_m\}$. We thus obtain the normal form. □

Once we have the normal form, we can then find a matrix $M \in \mathrm{SL}(n, \mathbb{Z})$ whose first m rows are $\mathbf{k}_1, \dots, \mathbf{k}_m$ such that $M\omega^\star$ has 0 first m entries. Thus, after a linear symplectic transformations $(x, Y) \mapsto (M^{-1}x, M^t Y)$, the potential V is a function of $x_1, \dots, x_m$. This allows us to reduce the normal form to a "pendulum+rotator" structure as in Arnold's example.

5C. ***The pendulum+rotator structure near resonance.*** For example, we consider the following Hamiltonian

$$H(x, y) = \tfrac{1}{2}\|y\|^2 + \varepsilon P(x), \quad (x, y) \in T^*\mathbb{T}^n.$$

We remark that we have chosen the kinetic energy part of the form $\frac{1}{2}\|y\|^2$ to simplify the discussion. A general kinetic energy of the form $\frac{1}{2}\langle \mathsf{A}Y, Y\rangle$ will create some new difficulty in separating the rotator and the pendulum. We have developed systematic tools (shear transformation and undo-shear etc) in [11] to deal with this issue. We avoid this complication by restricting ourselves to the simple example and refer interested readers to [11] for more details in the general case.

Suppose $y^\star = (0, \hat{\omega}^\star)$ where $\hat{\omega}^\star \in \mathbb{R}^{n-m}$ is Diophantine. Then the Hamiltonian has the following normal form up to an additive constant

$$\mathsf{H}(x, Y) = \tfrac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}^\star, \hat{Y}\rangle + \tfrac{1}{2}\|\hat{Y}\|^2 + \tfrac{1}{2}\|\tilde{Y}\|^2 + V(\tilde{x}) + \delta \mathsf{R}(x, Y), \tag{5-5}$$

where we use notation $x = (\tilde{x}, \hat{x})$ and $Y = (\tilde{Y}, \hat{Y})$, where $\tilde{}$ means the first m variables and $\hat{}$ means the last $n - m$ variables. This Hamiltonian system is split naturally into a product system if we discard the $\delta\mathsf{R}$ term. The subsystem $\hat{\mathsf{H}}(\hat{x}, \hat{Y}) = \frac{1}{\sqrt{\varepsilon}}\langle \hat{\omega}^\star, \hat{Y}\rangle + \frac{1}{2}\|\hat{Y}\|^2$ is integrable and can be considered as a rotator. Suppose V has a nondegenerate global maximum at 0 so the subsystem

$$\tilde{\mathsf{H}}(\tilde{x}, \tilde{Y}) = \tfrac{1}{2}\|\tilde{Y}\|^2 + V(\tilde{x}) \tag{5-6}$$

has a hyperbolic fixed point $(0, 0)$. The Hamiltonian H now has the form of "pendulum+rotator" structure as in Arnold's example. In particular, single resonance normal form gives rise to an *a priori* unstable system of the form (4-1) (the pendulum subsystem has one degree of freedom).

We warn the readers that the normal form becomes singular as $\varepsilon \to 0$, which is reflected in the term $\frac{1}{\sqrt{\varepsilon}}\langle \omega^\star, Y\rangle$ in (5-3) implying that the dynamics on the NHIC is fast rotating ($\dot{\hat{x}} = O(\varepsilon^{-1/2})$ in example (5-5)). This presents a technical difficulty in the proof. The way we solve the problem is to notice that its contribution to the variational equation disappears since $\omega^\star$ is a constant, hence it has no contribution to the differential of the time-1 map of the Hamiltonian system! This fact enables us to perform the graph transform as in [21; 26] to obtain a version of the theorem of NHIM in this setting, with which we turn on

the coupling $\delta \mathsf{R}$ in (5-5) (recall that in Theorem B.2, the assumptions are made on the differential of the map $f : M \to M$ but not on the vector field generating the flow). Details of the statement and the proof can be found in Appendix E of [11].

6. Mechanical systems with two degrees of freedom

In this section, we study the dynamics of a mechanical system of two degrees of freedom of the following form where A is positive definite

$$\mathsf{G}(x, Y) = \tfrac{1}{2}\langle \mathsf{A}Y, Y\rangle + V(x), \quad (x, Y) \in T^*\mathbb{T}^2. \tag{6-1}$$

This system appears naturally in the double resonance normal form. We will see in the next section that such a system is inevitable in constructing diffusing orbits for system (1-1) with $n = 3$.

This system is hard to analyze in general due to its nonperturbative nature. However, the two-dimensionality enables us to obtain enough information on the structure of Mather sets and Mañé sets so that diffusing orbit can be constructed passing through the double resonance.

6A. ***Two degrees of freedom: positive energy levels.*** We have the following theorem describing Mather sets of rational rotation vectors.

Theorem 6.1 [16, Theorem 2.1]. *Given a Tonelli Hamiltonian $H : T^*\mathbb{T}^2 \to \mathbb{R}$ and a class $g \in H_1(\mathbb{T}^2, \mathbb{Z})$ and a closed interval $[\nu^-, \nu^+]$ with $\nu^+ > \nu^- > 0$, there exists an open-dense set $\mathfrak{V} \subset C^r(\mathbb{T}^2, \mathbb{R})/\mathbb{R}$ with $r \geq 5$ such that for each $V \in \mathfrak{V}$ normalized by* $\max V = 0$, *it holds simultaneously for all $\nu \in [\nu^-, \nu^+]$ that the Mather set $\mathcal{M}_{\nu g}$ for $H + V$ consists of hyperbolic periodic orbits. Indeed, except for finitely many ν_j, the Mather set consists of two hyperbolic periodic orbits, for all other $\nu \in [\nu^-, \nu^+]$ it consists of exactly one hyperbolic periodic orbit.*

For each fixed positive energy level, the existence of periodic orbits as the Mather set and its generic uniqueness were known in the Aubry–Mather theory for twist maps. However, it is highly nontrivial to show that these periodic orbits form smooth families and the finiteness of the bifurcations when varying energy levels. This theorem completely describes the structure of the Mather set with rotation vectors along a rational ray in the frequency space. These Mather sets constitute a NHIC. At the bifurcation values ν_j, the two components of the Mather sets are connected by heteroclinic orbits in the Mañé set. When the system is perturbed by a time-periodic perturbation, again we have a system of *a priori* unstable type.

6B. ***Two degrees of freedom: the zeroth energy level.*** We next study the zero frequency case. This was done in [8]. Instead of studying the dynamics in the

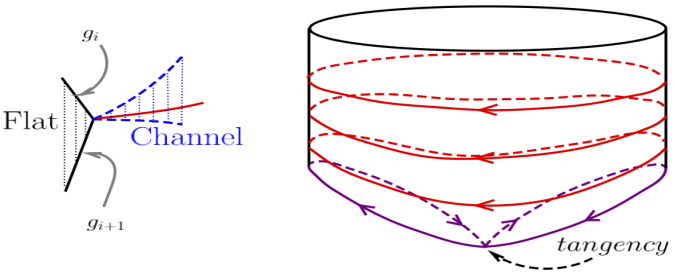

Figure 2. Channel connected to a vertex of $\partial\mathbb{F}_0$.

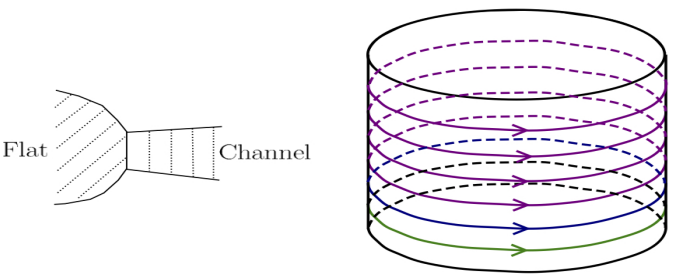

Figure 3. Channel connected to an edge of $\partial\mathbb{F}_0$.

frequency space, we switch to the space of cohomology classes. This passage has the effect of blowing up singularities.

We denote $\mathbb{F}_0 = \alpha_{\mathsf{G}}^{-1}(0)$ and call it the flat, which is a convex set and by $\partial\mathbb{F}_0$ the boundary of $\mathbb{F}_0$. A simple example is the product of two identical pendulums whose $\mathbb{F}_0$ is a square. We next introduce

$$\partial^*\mathbb{F}_0 = \{c \in \partial\mathbb{F}_0 : \mathcal{M}(c)\backslash\{x = 0\} \neq \varnothing\}.$$

The set $\partial^*\mathbb{F}_0$ can be nonempty. An example of a system with $\partial^*\mathbb{F}_0 \neq \varnothing$ was given in Section 2 of [9]. When $\partial^*\mathbb{F}_0 \neq \varnothing$ happens, then $\partial\mathbb{F}_0$ has infinitely many edges; see [39].

We next introduce a subset $G_{m,c} \subset H_1(\mathbb{T}^2, \mathbb{Z})$ be a subset that $g \in G_{m,c}$ if there is a minimal homoclinic orbit $(\gamma, \dot\gamma)$ in $\tilde{\mathcal{A}}(c)$ with $[\gamma] = g$. Given an edge $\mathbb{E}_i$, we define $G_{m,\mathbb{E}_i} = G_{m,c}$ for each $c \in \operatorname{int}\mathbb{E}_i$ since all classes in $\operatorname{int}\mathbb{E}_i$ share the same Aubry set.

The following theorem was proved in [8].

Theorem 6.2. *Let $\mathbb{F}_0 = \alpha_{\mathsf{G}}^{-1}(\min\alpha_{\mathsf{G}})$ be a two dimensional flat, and $\mathcal{M}(c_0)$ be a singleton for $c_0 \in \operatorname{int}\mathbb{F}_0$. Let $\mathbb{E}_i$ denote an edge of $\mathbb{F}_0$ (not a point), then*:

(1) *Either $\mathbb{E}_i \cap \partial^*\mathbb{F}_0 = \varnothing$ or $\mathbb{E}_i \subset \partial^*\mathbb{F}_0$.*

(2) *If $\mathbb{E}_i \cap \partial^*\mathbb{F}_0 = \varnothing$, then $G_{m,\mathbb{E}_i}$ contains exactly one element, if $\mathbb{E}_i \subset \partial^*\mathbb{F}_0$, all curves in $\mathcal{M}(\mathbb{E}_i)\backslash\{0\}$ have the same rotation vector.*

(3) *If $c \in \partial \mathbb{E}_i$ and $c \notin \partial^* \mathbb{F}_0$ then $G_{m,c}$ contains exactly two elements.*

(4) *If $\mathbb{E}_i, \mathbb{E}_j \subset \partial^* \mathbb{F}_0$, then either $\mathbb{E}_i = \mathbb{E}_j$, or $\mathbb{E}_i$ and $\mathbb{E}_j$ are disjoint.*

(5) *If $\mathbb{E}_i \subset \partial^* \mathbb{F}_0$, $\mathcal{M}(c) = \mathcal{M}(c')$ holds for $c \in \partial \mathbb{E}_i$ and $c' \in int \mathbb{E}_i$.*

(6) *If μ_c is an ergodic c-minimal measure for $c \in \partial^* \mathbb{F}_0$ and $\omega(\mu_c)$ is irrational, then the class c is an extremal point of $\mathbb{F}_0$.*

(7) *If $c \in \partial \mathbb{F}_0 \backslash \partial^* \mathbb{F}_0$ and $\tilde{\mathcal{A}}(c)$ consists of the fixed point and one homoclinic orbit $(\gamma, \dot{\gamma})$ only, then c is located in the interior of certain edge $\mathbb{E}_i$.*

(8) *Each edge in $\partial \mathbb{F}_0 \backslash \partial^* \mathbb{F}_0$ is joined by two edges in $\partial \mathbb{F}_0 \backslash \partial^* \mathbb{F}_0$.*

The result is summarized in the following dictionary. For each cohomology class c in the right column, the corresponding Aubry set $\tilde{\mathcal{A}}(c)$ is in the left column:

phase space	$H^1(\mathbb{T}^2, \mathbb{R})$
hyperbolic fixed point	convex disk int $\mathbb{F}_0$
homoclinic or periodic orbit	edge of $\mathbb{F}_0$
two homoclinic orbits	vertex of $\mathbb{F}_0$
homology class of homoclinic orbits	$\perp$ edge
NHIC foliated by periodic orbits	channel connected to $\mathbb{F}_0$

To relate this dictionary to Theorem 6.1, we see that the NHIC foliated by periodic orbits given in Theorem 6.1 gives the channel in the last row of the dictionary. We may let the energy level to approach zero. The fact that there is no infinite bifurcation in this limiting procedure is proved in the following Theorem 6.4. In the limit, the Mather set may remain a periodic orbit or degenerate to a homoclinic orbit as in the second line of the dictionary (see Figure 2) or degenerate to two homoclinic orbits as in the third line of the dictionary (see Figure 3). The purpose of a careful study of the structure of $\partial \mathbb{F}_0$ and its dynamical correspondence is to understand the dynamics on small positive energy levels, as we shall talk about in the next subsection.

6C. ***Dynamics around the strong double resonance.*** It is shown in [8] that for all $c \in \partial \mathbb{F}_0$, the projected Mañé set $\mathcal{N}(c)$ (the projection of $\tilde{\mathcal{N}}(c)$ from $T^*\mathbb{T}^2$ to $\mathbb{T}^2$) does not cover the two torus.

Theorem 6.3 [8, Theorem 3.1]. *Consider the Hamiltonian* G *of the type (6-1). There exists a residual set in $C^r(\mathbb{T}^2)/\mathbb{R}$, $r \geq 2$ such that for each V in the set normalized by* $\max V = 0$, *and for each $c \in \partial \mathbb{F}_0$, the Mañé set $\mathcal{N}(c)$ does not cover the whole configuration space, i.e., $\mathcal{N}(c) \subsetneq \mathbb{T}^2$ for all $c \in \partial \mathbb{F}_0$. Moreover, the upper-semicontinuity of $\mathcal{N}(c)$ with respect to c implies that there is a $\Delta = \Delta(V) > 0$ such that the same conclusion holds for all $c \in \alpha_{\mathsf{G}}^{-1}([0, \Delta])$.*

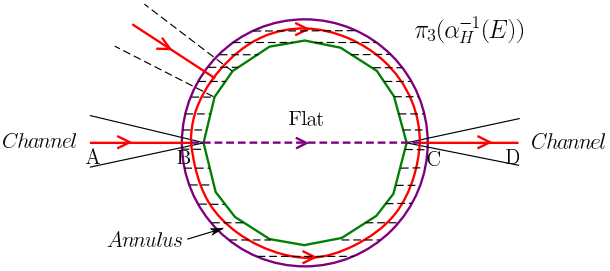

Figure 4. Turning around strong double resonance.

This theorem can be understood as that along the circle $\partial\mathbb{F}_0$, the dynamics is similar to that of the Birkhoff instability region of the twist map where there is no invariant circles. However, the facts that the dynamics on the zero energy level of G cannot be written as a twist map, the destruction of $\mathcal{N}(c)$ for all $c \in \partial\mathbb{F}_0$ and the nonperturbative nature of the system G make the result highly nontrivial. We refer readers to [8] for details.

The theorem implies that any two cohomology classes $c, c' \in \partial(\alpha_0^{-1}(E))$, $E \in [0, \Delta]$, are equivalent which gives rise to an orbit shadowing Mather sets $\widetilde{\mathcal{M}}(c)$ and $\widetilde{\mathcal{M}}(c')$; see Theorem 3.3. When viewed in the frequency space, this implies in particular that for any two rational rays starting from 0, there is an orbit shadowing Mather sets with rotation vectors lying on the two rays.

In general, it seems not easy to see the dynamical picture of orbits constructed here in the phase space. On very small energy levels when two channels are close, it seems natural that orbit shadows heteroclinics between periodic orbits on two channels, but the dynamics seems much richer when $\partial^*\mathbb{F}_0 \neq \varnothing$. We also remark that the number Δ here is obtained by the upper-semicontinuity of the Mañé set, hence does not admit an estimate, but it is certainly independent of ε. The numerical experiment of [23] seems to indicate that for dynamics around double resonance should mostly follow the mechanism here.

6D. ***Cylinders with a hole.*** The phase space picture of the dynamics near double resonance was studied by many authors. The idea is that the NHICs in Theorem 6.1 can reach the zero energy level and even extend slightly to the negative energy levels. On the zero energy level, the Poincaré return map takes infinitely long time to return. This makes it hard to verify the smoothness of the cylinder near the zero energy level. The classical theorem of normally hyperbolic invariant manifold does not apply since the cylinder here is constructed from periodic orbits but not perturbed from a known cylinder, while the theorem of normally hyperbolic invariant manifold is a theorem about the persistence of center manifolds under perturbations. The problem of the C^1 regularity of the cylinders was addressed in [17]. We state the main result of [17] in the

general setting.

$$H(x, y) = \tfrac{1}{2}\langle Ay, y\rangle + V(x), \quad z = (x, y) \in \mathbb{T}^n \times \mathbb{R}^n, \tag{6-2}$$

where the matrix A is positive definite and the smooth potential V attains its maximum at a unique point $x_0 \in \mathbb{T}^n$. In this case, $z_0 = (x_0, 0)$ is a fixed point of the Hamiltonian flow Φ_H^t and there exist some orbits homoclinic to the fixed point known after the work of Bolotin. Be aware that the system admits a symmetry $\mathbf{s} : (x, y) \to (x, -y)$, we see that if $z^+(t) = (x^+(t), y^+(t))$ is an orbit, $z^-(t) = \mathbf{s}z^+(t) = (x^+(-t), -y^+(-t))$ is also an orbit. Hence, nonshrinkable homoclinic orbits emerge paired.

To formulate our result, by a translation of variables $x \to x - x_0$ and $V \to V - V(x_0)$ we assume $x_0 = 0$, $V(0) = 0$. We study k pairs of homoclinic orbits $\{z_1^\pm(t), \dots, z_k^\pm(t)\}$ and denote by $\Gamma_i^\pm$ the closure of $\{z_i^\pm(t) \mid t \in \mathbb{R}\}$. A periodic orbit $z^+(t)$ is said to *shadow the orbits* $\{z_1^+(t), \dots, z_k^+(t)\}$ if the period admits a partition $[0, T] = [0, t_1] \cup [t_1, t_2] \cup \cdots \cup [t_{k-1}, T]$ such that $z^+(t)|_{[t_{i-1}, t_i]}$ falls into a small neighborhood of $z_i^+(t)$. In this case, its $\mathbf{s}$-symmetric counterpart $z^-(t) = \mathbf{s}z^+(t)$ shadows the orbits $\{z_k^-(t), \dots, z_1^-(t)\}$.

The case of $k = 1$ will be studied in the original phase space $\mathbb{T}^n \times \mathbb{R}^n$. To study the case $k \geq 2$, we work in the covering spaces $\bar{\pi}_h$: $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{T}_h^n \times \mathbb{R}^n$ and $\pi_h : \mathbb{T}_h^n \times \mathbb{R}^n \to \mathbb{T}^n \times \mathbb{R}^n$, where $\mathbb{T}_h^n = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_i \bmod h_i \in \mathbb{N}\backslash 0\}$. To decide the class $h = (h_1, \dots, h_n)$, we let $\bar{z}_1(t)$ be the lift of $z_1^+(t)$ to $\mathbb{R}^{2n}$ such that $\lim_{t\to-\infty} \bar{z}_1(t) = 0$, then choose a lift $\bar{z}_2(t)$ of $z_2^+(t)$ with $\lim_{t\to-\infty} \bar{z}_2(t) = \lim_{t\to\infty} \bar{z}_1(t)$. In the way, we get successively a lift $\bar{z}_i(t)$ of $z_i^+(t)$ for each i and define h to be the integer vector $\lim_{t\to\infty} \bar{z}_k(t)$. Let $\overline{\Gamma}$ be the closure of $\{\cup_{i\leq k} \bar{z}_i(t)) \mid t \in \mathbb{R}\}$, then we construct a shift $\sigma\overline{\Gamma}$ as follows. We define $\bar{z}_1'(t) \subset \sigma\overline{\Gamma}$ to be the lift of $z_1^+(t)$ such that $\lim_{t\to-\infty} \bar{z}_1'(t) = \lim_{t\to\infty} \bar{z}_k(t)$. Other $\bar{z}_i'(t)$, $i = 2, \dots, k$, is successively constructed. Let $\sigma\overline{\Gamma}$ be the closure of $\{\cup_i \bar{z}_i'(t)) \mid t \in \mathbb{R}\}$. We make the following assumption:

> *For k pairs of homoclinic orbits $\{z_1^\pm(t), \dots, z_k^\pm(t)\}$, there exists a nonnegative integer ℓ and a covering space $\bar{\pi}_h$: $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{T}_h^n \times \mathbb{R}^n$ such that $\bar{\pi}_h(\overline{\Gamma} \cup \sigma\overline{\Gamma} \cup \cdots \cup \sigma^\ell\overline{\Gamma})$ is a closed curve without self-intersection.* **(H)**

Theorem 6.4. *Under certain genericity assumptions including* **(H)** *(see* [17, Theorem 1.1]*) there exists a continuation of periodic orbits from the homoclinic orbits $\{z_1^\pm(t), \dots, z_k^\pm(t)\}$. More precisely, some $E_0 > 0$ exists such that*:

(1) *For any $E \in (0, E_0]$, on the energy level E there exist unique periodic orbit $z_E^+(t)$ and its $\mathbf{s}$-symmetric orbit $z_E^-(t) = \mathbf{s}z_E^+(t)$ shadowing the orbits $\{z_1^+(t), \dots, z_k^+(t)\}$ and $\{z_k^-(t), \dots, z_1^-(t)\}$ respectively. The set $\{z_E^\pm(t) \mid t \in \mathbb{R}\}$ depending on E approaches $\cup_i \Gamma_i^\pm$ in Hausdorff metric as $E \downarrow 0$.*

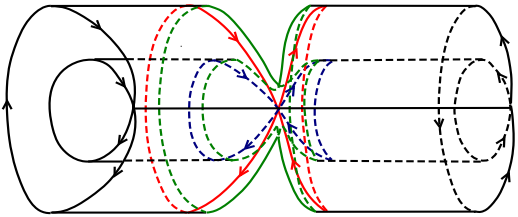

Figure 5. Cylinder with a hole.

(2) *For any* $E \in [-E_0, 0)$ *there exists a unique periodic orbit* $z_{E,i}$ *shadowing the orbits* $\{z_i^+(t), z_i^-(t)\}$ *for* $i = 1, \dots, k$. *As a set depending on* E, $\{z_{E,i}(t) \mid t \in \mathbb{R}\}$ *approaches* $\Gamma_i^+ \cup \Gamma_i^-$ *in Hausdorff metric as* $E \uparrow 0$.

Let $\Pi = \Pi^+ \cup_{1 \le i \le k} (\Pi_i^- \cup \Gamma_i^+ \cup \Gamma_i^-)$ *where* $\Pi^+ = \cup_{E>0}(\cup_t z_E^+(t) \cup z_E^-(t))$ *and* $\Pi_i^- = \cup_{E<0} \cup_t z_{E,i}(t)$. *For* $k = 1$, *it makes up a* C^1*-NHIC with one hole. For* $k \ge 2$, *each connected component in the pull-back* $\pi_h^{-1}\Pi$ *of* Π *to* $\mathbb{T}_h^n \times \mathbb{R}^n$ *is a* C^1*-NHIC with* $(\ell + 1)k$ *holes. The homoclinic orbits are contained inside of the manifold.*

In [17], to which readers are referred to, the authors give two more mechanisms of crossing the double resonance utilizing the geometric structure of cylinders with a hole. Compared to the first mechanism of turning around double resonance using c-equivalence, in this mechanism using cylinders with holes, orbits has to cross zero energy level hence we expect that the orbit should be much slower than that in the first mechanism hence is less likely. This is an interesting subject for future study.

7. Systems with three degrees of freedom

In this section, we give an overview of the proof of Theorem 1.3 in the case of $n = 3$.

7A. ***Design resonance paths and separate single and double resonances.*** We first show how to apply the homogenization and normal form to design resonance paths along which we construct diffusing orbit. We consider the case of three degrees of freedom for simplicity. Let $\varepsilon = 0$ in (1-1), now the frequency vector $\omega(y) := Dh(y) : \mathbb{R}^3 \to \mathbb{R}^3$ has range defined on a sphere when h is restricted to an energy level $E > \min h$. For any integer vector $\mathbf{k} \in \mathbb{Z}^3 \setminus \{0\}$, the resonance condition $\langle \mathbf{k}, \omega(y) \rangle = 0$ defines a circle on the sphere. Given two balls on the sphere, one can connect them by some of the resonant circles (in general at least 2). Along each resonance circle $S_{\mathbf{k}} := \{\langle \mathbf{k}, \omega(y) \rangle = 0, h(y) = E\}$, we show that the perturbation $P(\cdot, y) : \mathbb{T}^3 \to \mathbb{R}$, $y \in S_{\mathbf{k}}$ generically has a unique nondegenerate global max, up to finitely many bifurcation points where there

are two nondegenerate global max (see Proposition 8.1 below for a version of this type of parametric transversality result). As we have explained in Section 5, we will cover such resonant circles by balls of radius $\Lambda\sqrt{\varepsilon}$ centered at $y^\star \in S_{\mathbf{k}}$, and perform the homogenization in each ball. The global max of $P(\cdot, y^\star)$ gives rise to a hyperbolic fixed point in a pendulum-like subsystem as we have seen in Section 5C. In particular, the normal hyperbolicity is uniform for all the homogenized systems around $S_{\mathbf{k}}$ due to the uniform nondegeneracy of the global max of $P(\cdot, y)$, $y \in S_{\mathbf{k}}$. The uniform normal hyperbolicity gives a bound d_0 of the maximal allowable C^1-norm of the perturbation so that the theorem of NHIM is valid. Note that this d_0 is independent of ε but depends only on P.

Let δ be a small number but independent of ε and apply the normal form Proposition 5.2 and we consider only finitely many integer vectors of lengths less than $\delta^{-1/2}$. Along $S_{\mathbf{k}}$, there might be a second resonance, i.e., there is $\mathbf{k}'$ with $|\mathbf{k}'| < \delta^{-1/2}$ linearly independent of $\mathbf{k}$ such that $\langle \omega(y_*), \mathbf{k}'\rangle = \langle \omega(y_*), \mathbf{k}\rangle = 0$, $y_* \in S_{\mathbf{k}}$. For each such point y_*, outside of its $O(\varepsilon^{1/3})$-neighborhood, we can apply Proposition 5.2 with single resonance ($m = 1$), and within such an $O(\varepsilon^{1/3})$-neighborhood, we apply Proposition 5.2 with double resonance ($m = 2$). In the former case, the problem is essentially reduced to the *a priori* unstable case after some highly nontrivial work (recall example (5-5) with $m = 1$). In the latter case, the potential $V(\langle \mathbf{k}, x\rangle, \langle \mathbf{k}', x\rangle)$ in (5-3) can be decomposed into $\overline{V}(\langle \mathbf{k}, x\rangle) + \tilde{V}(\langle \mathbf{k}, x\rangle, \langle \mathbf{k}', x\rangle)$, where $\tilde{V}$ depends on $\langle \mathbf{k}', x\rangle$ nontrivially and its C^2 norm is estimated as $C|\mathbf{k}'|^{-(r-2)}$ for fixed $\mathbf{k}$ and $P \in C^r$. When $\|\tilde{V}\|_{C^2} < d_0$, we can still apply the theorem of NHIM by treating $\tilde{V}$ as a perturbation so we call this case *weak double resonance* and treat it in a similar manner as a single resonance. Then the remaining case $\|\tilde{V}\|_{C^2} \geq d_0$ is called *strong double resonance*. Note that there are only finitely many of them, whose number is independent of ε, δ. The $O(\varepsilon^{1/3})$-neighborhood of a strong double resonance can be further divided into the $O(\varepsilon^{1/2})$-neighborhood and the region outside the $O(\varepsilon^{1/2})$-neighborhood. The former case is reduced to the setting of Section 6 as we will see in the next subsection. The latter case is the regime of transiting from single to double resonance regimes. It can be treated as the high energy level sets in the system (6-1) and in the normal form Proposition 5.2, the frequency $\omega^\star/\sqrt{\varepsilon}$ goes from $O(\varepsilon^{-1/6})$ to $O(1)$. It is shown in [8] that the cylinder in Theorem 6.1 in the high energy level regime consists of a single piece, without bifurcation and the normal hyperbolicity is uniform, so this transiting regime can also be treated as a system of *a priori* unstable type. In the following, we focus on the strong double resonances.

7B. ***The double resonance.*** In a $\Lambda\sqrt{\varepsilon}$ ball centered at a double resonance, we apply the homogenization and the normal form followed by a linear symplectic

transform to get the following Hamiltonian at double resonance

$$\mathsf{H}(x, Y) = \tfrac{\omega_3}{\sqrt{\varepsilon}} Y_3 + \tfrac{1}{2}\langle \mathsf{A}Y, Y\rangle + V(x_1, x_2) + \delta \mathsf{R}(x, Y),$$
$$(x; Y) = (x_1, x_2, x_3; Y_1, Y_2, Y_3).$$

We perform a standard energetic reduction (fixing an energy level and solve for $\mathsf{Y}_{\text{double}} := (\omega_3/\sqrt{\varepsilon})Y_3$ as the new Hamiltonian and its conjugate $\tau = (\sqrt{\varepsilon}/\omega_3)x_3$ as the new time) to get

$$\mathsf{Y}_{\text{double}} = \tfrac{1}{2}\langle \tilde{\mathsf{A}}\tilde{Y}, \tilde{Y}\rangle + V(\tilde{x}) + \delta\tilde{\mathsf{R}}\left(\tilde{x}, \frac{\omega_3\tau}{\sqrt{\varepsilon}}, \tilde{Y}\right), \quad (\tilde{x}; \tilde{Y}) = (x_1, x_2; Y_1, Y_2) \quad (7\text{-}1)$$

where $\tilde{\mathsf{A}}$ is obtained from A by removing the third row and column, which are absorbed in $\delta\tilde{\mathsf{R}}$ during the reduction. We thus arrive at a system that is a small time-dependent perturbation of the nonperturbative mechanical system G of two degrees of freedom. Note that the τ-dependence in $\delta\tilde{\mathsf{R}}$ is fast oscillating as $\varepsilon \to 0$. This singular behavior does not invalidate the theorem of NHIM since it does not enter the estimate of the differential of the time-1 map for the similar reason to the discussion near the end of Section 5. We again refer readers to Appendix E of [11] for this point. We remark that Arnold [4] already identified this as the main difficulty for Arnold diffusion.

Let us now see how the action variable changes if a diffusion orbit is to be constructed. Suppose we want to move y along the resonant circles determined by

$$S_1 := \{\langle \omega(y), \mathbf{k}_1\rangle = 0\} \quad \text{and} \quad S_2 := \{\langle \omega(y), \mathbf{k}_2\rangle = 0\}.$$

For simplicity we assume $\mathbf{k}_1 = (0, 1, 0)$ and $\mathbf{k}_2 = (1, 0, 0)$ hence along S_1 the frequency has the form $\omega(y) = (\omega_1(y), 0, \omega_3(y))$ and along S_2 we have $\omega(y) = (0, \omega_2(y), \omega_3(y))$. When two resonances occur simultaneously we have $\omega(y) = (0, 0, \omega_3(y))$. Along the resonant circle S_1, we apply the normal form Proposition 5.2 with $m = 1$, then by the argument following Proposition 5.2, we reduce the problem to an *a priori* unstable system in Section 4 so that we can move y freely on the resonant circle S_1 provided there is no second resonance. When the second resonance appears, in a neighborhood of $S_1 \cap S_2$ we apply Proposition 5.2 with $m = 2$ to yield the normal form (7-1) after the energetic reduction. The energetic reduction treats the third angular variable x_3 as the new time, hence for the system (7-1), the frequency vector is obtained by removing the third entry from $\omega(y)$. So along S_1 the reduced frequency vector has the form $\omega_a := (a, 0)$, $a \in \mathbb{R}$ and along S_2 it has the form $\omega_b := (0, b)$, $b \in \mathbb{R}$, and the double resonance corresponds to the frequency vector $(0, 0)$. To cross the strong double resonance $S_1 \cap S_2$, we have to construct orbit moving along ω_a close to $(0, 0)$ then along ω_b.

Equipped with the knowledge in Section 6 on the mechanical system G of two degrees of freedom, we are ready to construct diffusing orbits moving around the double resonance. We interpret the frequency ω_a or ω_b as the rotation vector of a Mather set, which is the velocity averaged against the minimizing measure. Theorem 6.1 enables us to move the frequencies along ω_a and ω_b using the mechanism of *a priori* unstable systems up to a small neighborhood of 0. Next, we apply Theorem 6.3 and the c-equivalence mechanism to find orbit shadowing Mather sets with rotation vectors on the frequency segments ω_a and ω_b. Therefore we overcome the difficulty of strong double resonance and global diffusing orbits are constructed in the case of $n = 3$.

8. The genericty

The genericity of the perturbations is a central issue and is closely related to the dynamics. For *a priori* unstable systems, we have outlined the genericity argument in Section 4, which is also applicable to *a priori* stable systems in the regime of single resonances and transition from single to double resonances, where the problem is essentially reduced to *a priori* unstable systems after some work, though highly nontrivial. Near double resonance, the genericity of the perturbations are given in Theorem 6.1 and 6.3. In Theorem 6.1, the genericity originates from the following parametric transversality result; see [16].

Proposition 8.1. *Let $F_s : \mathbb{T}^1 \to \mathbb{R}$, $s \in [0, 1]$ be a family of C^r, $r > 4$ functions that is Lipschitz in s. Then there is an open and dense subset $\mathcal{R}$ of $C^r(\mathbb{T})$ such that for each $V \in \mathcal{R}$, the function $F_s + V$ admits a unique nondegenerate global minimum for all but finitely many parameters $s_1, \ldots, s_n$ for which $F_{s_i} + V$, $i = 1, \ldots, n$ admits two nondegenerate global minimums.*

The genericity in Theorem 6.3 is achieved by only finitely many perturbations in the proof. Since the proof is a bit involved, we refer readers to [9] for details.

Note that in the main terms of the normal form (5-3) as well as the system (6-1), the system is of the form of mechanical systems (kinetic energy+potential energy). In particular, the potential part depends only on the angular variables. That is why we consider only Mañé perturbations (perturbations depending only on angular variables) in the statements of Theorem 6.1 and 6.3, which are the only allowed perturbations.

In the statement of Theorem 1.3, the perturbation P can either depend on all variables or simply Mañé. To achieve the genericity of Mañé perturbation, one of the main difficulties is to do it for *a priori* unstable systems, considering that Theorem 6.1 and 6.3 are stated for Mañé perturbations. Indeed, this is exactly the content of Section 4.2 of [9], where we refer interested readers.

Appendix A: Preliminary Hamiltonian dynamics

In this section, we give some preliminaries on Hamiltonian dynamics.

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

(1) Positive definiteness: for each $(x, t) \in M \times \mathbb{T}$, the Lagrangian function is C^2 strictly convex in velocity, i.e., the Hessian $\partial_{\dot{x}\dot{x}} L$ is positive definite.

(2) Super-linear 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 solutions of the Euler–Lagrange equation are well defined for the whole $t \in \mathbb{R}$.

We have the following remarks:

- (Euler–Lagrange equation) Given a Lagrangian L, its Lagrangian flow is solved from the *Euler–Lagrange equation* $\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$; see [3, Chapter 3].
- (autonomous, nonautonomous, twist maps) We say that L is *autonomous* if it does not depend on t, otherwise it is called *nonautonomous.* A nonautonomous system $L : T\mathbb{T}^n \times \mathbb{T} \to \mathbb{R}$ will be said to have $n + \frac{1}{2}$ degrees of freedom. When $n = 1$, a Tonelli Lagrangian of 1.5 degrees of freedom has time-1 map defined on $T\mathbb{T}$.
- (Tonelli Hamiltonian) A Hamiltonian $H : T^*M \times \mathbb{T} \to \mathbb{R}$ is called Tonelli, if it is the Legendre transform of a Tonelli Lagrangian, i.e.,

$$H(x, y, t) = \max_{\dot{x}} \langle y, \dot{x} \rangle - L(x, \dot{x}, t).$$

  For instance, any *mechanical Hamiltonian* of the form $H(x, y) = \frac{1}{2}\|y\|^2 + V(x)$, $(x, y) \in T^*\mathbb{T}^n$ is *Tonelli*, since it is the Legendre transform of $L(x, \dot{x}) = \frac{1}{2}\|\dot{x}\|^2 - V(x)$, $(x, \dot{x}) \in T\mathbb{T}^n$.
- (energetic reduction) Given an autonomous Hamiltonian $H : T^*\mathbb{T}^n \to \mathbb{R}$, denoting $x = (\hat{x}, x_n) \in \mathbb{T}^{n-1} \times \mathbb{T}$, $y = (\hat{y}, y_n) \in U (\subset \mathbb{R}^{n-1} \times \mathbb{R})$ where U is a bounded domain, if we know $\frac{\partial H}{\partial y_n} \neq 0$, then we can apply the implicit function theorem to the Hamiltonian $H(\hat{x}, x_n, \hat{y}, y_n) = E$ restricted to the constant energy level E, to solve for $y_n = y_n(\hat{x}, x_n, \hat{y})$. Now y_n can be considered as a nonautonomous Hamiltonian of $n - \frac{1}{2}$ degrees of freedom with angular variables $\hat{x}$, action variables $\hat{y}$ and x_n as the time variable; see [3, Section 45, Chapter 9].

Appendix B: The theorem of normally hyperbolic invariant manifolds

In this section we give the version of normally hyperbolic invariant manifold theorem that we used in the proof of our main theorem. The standard references are [26; 22]. Readers are also referred to [20]. There are some subtleties of applying the theorem in the proof concerning the fast oscillatory nature in the nonresonant degrees of freedom. Readers can find more details in the appendix of [11].

Definition B.1 (NHIM). Let $N \subset M$ be a submanifold (maybe noncompact) invariant under f, $f(N) = N$. We say that N is a normally hyperbolic invariant manifold if there exist a constant $C > 0$, rates $0 < \lambda < \mu^{-1} < 1$ and an invariant (under Df) splitting for every $x \in N$

$$T_x M = E^s(x) \oplus E^u(x) \oplus T_x N$$

in such a way that

$$\begin{aligned} v \in E^s(x) &\Leftrightarrow |Df^n(x)v| \le C\lambda^n |v|, \quad n \ge 0, \\ v \in E^u(x) &\Leftrightarrow |Df^n(x)v| \le C\lambda^{|n|} |v|, \quad n \le 0, \\ v \in T_x N &\Leftrightarrow |Df^n(x)v| \le C\mu^n |v|, \quad n \in \mathbb{Z}. \end{aligned}$$

Here the Riemannian metric $|\cdot|$ can be any prescribed one, which may change the constant C but not λ, μ.

Theorem B.2. *Suppose N is a NHIM under the C^r, $r > 1$, diffeomorphism $f : M \to M$. Denote $\ell = \min\{r, |\ln\lambda|/|\ln\mu|\}$. Then for any C^r f_ϵ that is sufficiently close to f in the C^1 norm*:

(1) *There exists a NHIM N_ϵ that is a C^ℓ graph over N.*

(2) (*Invariant splitting*) *There exists a splitting for $x \in N_\epsilon$*

$$T_x M = E^u_\epsilon(x) \oplus E^s_\epsilon(x) \oplus T_x N_\epsilon \tag{B-1}$$

invariant under the map f_ϵ. The bundle $E^{u,s}_\epsilon(x)$ is $C^{\ell-1}$ in x.

(3) *There exist C^ℓ stable and unstable manifolds $W^s(N_\epsilon)$ and $W^u(N_\epsilon)$ that are invariant under f and are tangent to $E^s_\epsilon \oplus TN_\epsilon$ and $E^u_\epsilon \oplus TN_\epsilon$ respectively.*

(4) *The stable and unstable manifolds $W^{u,s}(N_\epsilon)$ are fibered by the corresponding stable and unstable leaves $W^{u,s}_{x,\epsilon}$:*

$$W^u(N_\epsilon) = \cup_{x\in N_\epsilon} W^u_{x,\epsilon}, \quad W^s(N_\epsilon) = \cup_{x\in N_\epsilon} W^s_{x,\epsilon}.$$

(5) *The maps $x \mapsto W^{u,s}_{x,\epsilon}$ are $C^{\ell-j}$ when $W^{u,s}_{x,\epsilon}$ is given C^j topology.*

(6) *If f and f_ϵ are Hamiltonian and* $\dim E^s = \dim E^u$, *then N_ϵ is symplectic and the map f_ϵ restricted to N_ϵ is also Hamiltonian* [20].

Appendix C: Systems of arbitrary degrees of freedom

In this appendix, we illustrate how to construct diffusing orbit in the general $n > 3$ case. The main difficulty is that it is not avoidable to study the dynamics around the complete resonance where the system is reduced to a mechanical system of $(n-1)$ degrees of freedom, which is in general nonperturbative. The high dimensional and nonperturbative nature of the problem creates serious difficulties in general. For example, for a nonperturbative mechanical system of $(n-1)$ degrees of freedom, Mather sets with rational rotation vectors may not be periodic orbits and when they are periodic they are not necessarily hyperbolic. So the theory in Section 6 cannot be recovered in this case. However, it turns out that close to codimension 1 and 2 KAM tori, we can find a connected set where perturbative techniques can be applied to reduce the problem to a multiscale system such that in each scale we have only single or double resonances. In this way, the methods in the previous sections can be applied to construct diffusing orbits. The general strategy is as follows:

(1) Try to find NHIC homeomorphic to $T^*\mathbb{T}\times\mathbb{T}$ to apply the method of *a priori* unstable system (4-1).

(2) Apply the mechanism of c-equivalence when there is a strong double resonance.

(3) Introduce new ideas to cross resonances of higher multiplicity.

In case (1), we require the NHIC to be homeomorphic to $T^*\mathbb{T}$ instead of $T^*\mathbb{T}^k$, $k > 1$, mainly because the regularity Theorem 4.1 is only established in the case of $T^*\mathbb{T}$.

C1. ***Choosing the frequency path.*** We describe an algorithm to choose the frequency lines along which the diffusion orbits are constructed.

The diffusing orbit will be constructed along some resonant path in order to utilize the resonant normal form. We design a procedure to construct a frequency path with special hierarchy structure. In the first step we start with a frequency segment of the form

$$\omega_a = \rho_a\left(a, \frac{p_2}{q_2}\omega_2^*, \frac{p_3}{q_3}\omega_2^*, \hat{\omega}_{n-3}^*\right) \in \mathbb{R}^n, \tag{C-1}$$

where $(\omega_2^*, \hat{\omega}_{n-3}^*) = (\omega_2^*, \omega_4^*, \omega_5^*, \ldots, \omega_n^*) \in \mathbb{R}^{n-2}$ is a Diophantine vector in $\mathbb{R}^{n-2}$, and a lies in an interval, p_2/q_2, $p_3/q_3 \in \mathbb{Q}$ irreducible. For all a, the vector ω_a admits a resonant integer vector $\mathbf{k}_1 = (0, q_2p_3, -q_3p_2, 0, \ldots, 0)$. After a linear transform by a matrix in $\mathrm{SL}(n, \mathbb{Z})$, we get $\check{\omega}_a = \rho_a(a, 0, \frac{p}{q}\omega_2^*, \hat{\omega}_{n-3}^*) \in \mathbb{R}^n$. We want to show that a can be moved arbitrarily. More precisely, for any a', $a'' \in \mathbb{R}$ and δ sufficiently small, there is an orbit $(x(t), y(t))$, $t \in [t', t'']$, such

that $\omega(y(t))|_{[t',t'']}$ lies in a δ neighborhood of ω_a, $a \in [a', a'']$. The frequency $\check{\omega}_a$ has at most two resonance relations, one of which is always $(0, 1, 0, \dots, 0)$, so the normal form Proposition 5.2 with either $m = 1$ or $m = 2$ applies.

We first consider region where Proposition 5.2 with $m = 1$ applies. In this case the function V in (5-3) is defined on $\mathbb{T}^1$. Similar to Section 3A and the example after the statement of Proposition 5.2, we see that the system admits a NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$ corresponding to the global max of V. In the case of $m = 2$, the function V is defined on $\mathbb{T}^2$. We can then separate a subsystem of the form (6-1), similar to the example after Proposition 5.2, and apply Theorem 6.1 to it to find a NHIC, which gives rise to a NHIC homeomorphic to $T^*\mathbb{T}^{n-1}$ for the full system slightly away from the strong double resonance.

To proceed, we need the following observations:

(1) The Hamiltonian system restricted to the NHIC is still Hamiltonian with one less degree of freedom.

(2) The hyperbolicity of the NHIC is determined by the nondegeneracy of the global max of the potential V.

(3) The remainder $\delta \mathsf{R}$ can be made as small as we wish.

(4) The normal form Proposition 5.2 with $m = 1$ or 2 holds in a neighborhood U of the frequency segment ω_a. The size of the neighborhood depends on δ.

Using (2) and (3), we choose δ so small that the perturbation $\delta \mathsf{R}$ does not destroy the NHIC constructed above. We then fix δ to proceed to the next step. Using (1), we obtain a Hamiltonian system restricted to the NHIC which is still nearly integrable has $n - 1$ degrees of freedom. Item (4) implies that the restricted Hamiltonian has frequencies (or rotation vectors of Mather sets, more precisely) in a neighborhood of $\rho_a\big(a, \frac{p}{q}\omega_2^*, \hat{\omega}_{n-3}^*\big) \in \mathbb{R}^{n-1}$ which is obtained from $\check{\omega}_a$ by removing the zero entry corresponding to the normal to the NHIC. So we can modify in the neighborhood U the first component ω_4^* of the vector $\hat{\omega}_{n-3}^*$ to a rational multiple of ω_2^*, so that the new frequency segment denoted by $\bar{\omega}_a = \big(a, \frac{p}{q}\omega_2^*, \frac{p_4}{q_4}\omega_2^*, \hat{\omega}_{n-4}^*\big)$ has a similar structure as ω_a so we can repeat the above procedure. Note that the rational $\frac{p_4}{q_4}$ necessarily has large denominator depending on δ. In the original system the means that we modify the frequency segment ω_a to $\omega_a' = \rho_a\big(a, \frac{p_2}{q_2}\omega_2^*, \frac{p_3}{q_3}\omega_2^*, \frac{p_4}{q_4}\omega_2^*, \hat{\omega}_{n-4}^*\big)$ hence introduces a second resonant integer vector $\mathbf{k}_2$ such that $\langle \mathbf{k}_2, \omega_a' \rangle = 0$ for all a. We have $|\mathbf{k}_2| \gg |\mathbf{k}_1|$ (more precisely, as $\delta \to 0$, we have $|\mathbf{k}_1|$ fixed but $|\mathbf{k}_2| \to \infty$) and moreover, the two vectors are not determined at once, instead, after $\mathbf{k}_1$ is determined and δ is fixed, we can then determine $\mathbf{k}_2$ by choosing p_4/q_4.

After $n-3$ steps, the above algorithm gives a final frequency segment of the form

$$\boldsymbol{\omega}_a^\sharp = \rho_a^\sharp\left(a, \frac{p_2}{q_2}, \frac{p_3}{q_3}, \ldots, \frac{p_n}{q_n}\right) \in \mathbb{R}^n$$

with a special *hierarchy structure*: for fixed p_i/q_i, we then choose p_{i+1}/q_{i+1} sufficiently close to a Diophantine number $\omega_{i+1}^*/\omega_2^*$. We choose p_{i+1}/q_{i+1} so close to $\omega_{i+1}^*/\omega_2^*$ that the resonance integer vector $\mathbf{k}_i$ has a large norm and the Fourier modes $\Pi_{\mathbf{k}_i} P$ are so small that it does not destroy the NHIC obtained in the previous step applying the theorem of NHIM. Now the frequency vector $\boldsymbol{\omega}_a^\sharp$ has at least $n-2$ resonant integer vectors $\mathbf{k}_i$, $i = 1, 2, \ldots, n-2$ with $|\mathbf{k}_i| \ll |\mathbf{k}_{i+1}|$ for all a in an interval. For some a, there might be another resonant vector $\mathbf{k}''$ whose length is comparable to one of $\mathbf{k}_i$. We note that the vectors $\mathbf{k}_i$, $i = 1, \ldots, n-2$ are not determined at once, instead, we determine $\mathbf{k}_{i+1}$ after $\mathbf{k}_i$ is fixed.

Suppose we finish moving a and want to move the second component of the frequency vector. The idea is to send a close to a Diophantine number that is much closer than $|p_n/q_n - \omega_n^*/\omega_2^*|$ and start moving p_2/q_2 applying the above algorithm.

Carrying out the above procedure, we get the existence of NHICs outside a small neighborhood of the complete resonance. With the presence of the NHICs, we can consider the system as an *a priori* unstable system and construct diffusing orbit. We thus have the following statement (except part (3)(c) to be explained in the next subsection).

Theorem C.1 [11, Theorem 2.9]. *Let the Hamiltonian system $H = h + \varepsilon P \in C^r(T^*\mathbb{T}^n, \mathbb{R})$, $7 \le r \le \infty$, be as in (1-1) restricted to the energy level $E > \min h$. For any $\varrho > 0$, and any M open balls $B_1, \ldots, B_M$ of radius ϱ centered on $h^{-1}(E)$, there exist some $\varepsilon_0 > 0$ and an open-dense set $\mathfrak{R} \subset \mathfrak{S}_1$, such that for each $P \in \mathfrak{R}$ there exist ε_P 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 $\omega(t)$ with $\partial\beta(\omega(t)) \in \alpha^{-1}(E)$, $t \in [1, M]$ satisfying*:
   (a) $(\partial h)^{-1}(\omega(i)) \cap B_i \ne \varnothing$, $i = 1, 2, \ldots, M$.
   (b) *Each point $\omega(t)$ is resonant with multiplicity at least $n-2$. There are finitely many marked points on $\omega(t)$ denoted by $\omega_1, \ldots, \omega_m$, 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 normally hyperbolic invariant cylinders homeomorphic to $T^*\mathbb{T} \times \mathbb{T}$.*

(3) *For each ω_i, $i = 1, \ldots, m$, there exists $\lambda_i > 0$ such that*:
   (a) *The Mather sets of rotation vectors $\omega(t)$ with $|\omega(t) - \omega_i| \ge \lambda_i \sqrt{\varepsilon}$ for all $i = 1, 2, \ldots, m$, lie in the NHICs.*

(b) *Any continuous curve lying in the interior of* $\{\partial\beta(\omega(t)) \mid |\omega(t)-\omega_i| \geq \lambda_i\sqrt{\varepsilon}\} \subset \alpha^{-1}(E)$ *satisfies Theorem 3.1.*

(c) *The two neighboring connected components* $\{\partial\beta(\omega(t)) \mid |\omega(t)-\omega_i| \geq \lambda_i\sqrt{\varepsilon}\} \subset \alpha^{-1}(E)$ *near* $\partial\beta(\omega_i)$ *are c-equivalent.*

Remark C.2. Each marked point corresponds to a strong double resonance point appearing at some step of the reduction of order where there is a resonant integer vector $\mathbf{k}''$ whose length is comparable to some $\mathbf{k}_i$. We avoid getting too close to the double resonance. The reason is that the NHIC, if it exists, has only $C^{1+\alpha}$ smoothness where $\alpha > 0$ can be close to 0 since $|\ln\lambda/\ln\mu|$ can be close to one in Theorem B.2 near the strong double resonant point. The regularity is too low to perform further reduction of order.

C2. ***Crossing the complete resonance.*** In the previous subsection, we have shown how to construct NHICs away from complete resonances. In this section, we show how to cross the complete resonance hence prove part (3)(c) in Theorem C.1. Similar to the case of $n = 3$, the complete resonance causes essential difficulty to construct diffusing orbit in the higher dimensional case.

The normal form near the complete resonance. Applying Proposition 5.2 repeatedly, we derive the following Hamiltonian normal form at the complete resonant frequency $\boldsymbol{\omega}_a^\sharp$; see Section 7.5 of [11]. After a linear transform in $\mathrm{SL}(n,\mathbb{Z})$, we transform $\boldsymbol{\omega}_a^\sharp$ to $(0,\ldots,0,\omega_n)$

$$H_{n-1} = \tfrac{1}{\sqrt{\varepsilon}}\omega_n Y_n + \tfrac{1}{2}\langle A_{n-1}Y, Y\rangle + \sum_{i=2}^{n-1}\delta_i V_i(x_1,\ldots,x_i) + \delta_n R(x,Y),$$

where $(x, Y) \in T^*\mathbb{T}^n$, $V_i \in C^r$, and $R \in C^{r-2}$. The Hamiltonian has the following properties which originate from the hierarchy structure in the choice of the frequency line in the previous section:

(1) $\delta_{i+1} \ll \delta_i$, $\delta_2 = 1$, *and we have the freedom to choose* δ_{i+1} *as small as we wish once* $\delta_i V_i$ *is fixed, and* V_{i+1} *depends on* δ_{i+1} *but* $\|V_{i+1}\|_{C^r}$ *is uniformly bounded as* $\delta_{i+1} \to 0$. *The number* δ_{i+1} *is chosen so that the* δ_{i+1}*-perturbation does not destroy the NHIC constructed in the previous step whose normal hyperbolicity depends on* δ_i.

(2) *The positive definite matrix* A_{n-1} *depends on* δ_i *in the following way: the first* $i \times i$ *block depends only on* $\delta_2,\ldots,\delta_i$ *but does not depend on* $\delta_{i+1},\ldots,\delta_{n-1}$ *for* $i = 2,\ldots,n-1$. *Such dependence on* δ_i *appears due to our introduction of the linear symplectic map after applying the normal form.*

We next perform a standard energetic reduction to solve for $Y_n(x, x_n, y)$ as the solution of the equation

$$H_{n-1}(x, x_n, y, Y_n(x, x_n, y)) = E^* > \min \alpha_{H_{n-1}}$$

to arrive at the normal form which is a nonautonomous system with $n + 1/2$ degrees of freedom

$$Y_\delta := -Y_n \tfrac{\omega_n}{\sqrt{\varepsilon}} = \tfrac{1}{2}\langle Ay, y\rangle + \sum_{j=2}^{n-1} \delta_i V_i(x_1, \dots, x_i) + \delta_n \tilde{R}\big(x, \tfrac{\tau}{\sqrt{\varepsilon}}, y\big), \qquad \text{(C-2)}$$

where we update the notation $x = (x_1, \dots, x_{n-1})$, $y = (Y_1, \dots, Y_{n-1})$, and A denotes an $(n-1)\times(n-1)$ matrix obtained by removing the last row and column in A_{n-1}.

In these coordinates, one case of crossing the complete resonance is to move the frequency $a(1, 0, \dots, 0) \in \mathbb{R}^{n-1}$ from some positive a to some negative a along an orbit with the obstruction being the zero frequency.

The algorithm of constructing diffusing orbit crossing the complete resonance. For simplicity, we consider the case $n = 4$ and assume $A = \mathrm{Id}_3$. The general case is more complicated and we refer readers to Section 6 of [11] for details. We also discard the term $\delta_4 \tilde{R}$ in Y_δ since it is useless in our argument of passing complete resonance.

Step 1 *The cohomology space picture.* We get the Hamiltonian

$$Y_\delta = \tfrac{1}{2}\|y\|^2 + V(x_1, x_2) + \delta V_3(x_1, x_2, x_3), \quad (x, y) \in T^*\mathbb{T}^3. \qquad \text{(C-3)}$$

We first study the picture of $\mathbb{F}_0 = \alpha_{Y_\delta}^{-1}(0)$ in $H^1(\mathbb{T}^3, \mathbb{R})$. This has the shape of a big pizza (see Figure 6): $O(1)$ in the c_1, c_2 direction and very tiny $O(\sqrt{\delta})$ in the c_3 direction where $c = (c_1, c_2, c_3) \in H^1(\mathbb{T}^3, \mathbb{R})$, since the hyperbolicity of the hyperbolic fixed point is weak in the x_3, y_3 component. Each NHIC (homeomorphic to $T^*\mathbb{T}^1$) provided by Theorem C.1 corresponds in $H^1(\mathbb{T}^3, \mathbb{R})$ to an open set that we call a channel connected to $\mathbb{F}_0$. The correspondence is in the following sense. Each NHIC consists of hyperbolic periodic orbits in the Mather sets with rotation vectors lying in the frequency line $(a, 0, 0)$, $a \in \mathbb{R} \setminus \{0\}$ and the channels are the images of the frequency line under the map $\partial\beta : H_1(\mathbb{T}^3, \mathbb{R}) \to H^1(\mathbb{T}^3, \mathbb{R})$. One case of crossing the complete resonance is to find an orbit shadowing Mather sets with rotation vectors $(a, 0, 0)$ and $(-a, 0, 0)$, $a \neq 0$. Note that the picture of the pizza and channels is centrally symmetric since the system Y_δ is reversible (invariant under the change $y \to -y$). Our goal is to move the cohomology class c from one channel to another, hence by symmetry $c \to -c$. We have the following algorithm.

Step 2 ***a priori*** *unstable dynamics and center-straightening.* The NHIC provided by Theorem C.1 is obtained in the following way. First, since the subsystem $\mathsf{G} = \frac{1}{2}(y_1^2 + y_2^2) + V(x_1, x_2)$ has two degrees of freedom, we apply Theorem 6.1 to get a NHIC foliated by action minimizing periodic orbits in the homology class $g = (1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$. Moreover, action-angle coordinates (θ, I) on the cylinder can be introduced to reduce the subsystem to a system $\tilde{h}(I)$ of one degree of freedom restricted to the NHIC. This reduces the Hamiltonian Y_δ to the form

$$\overline{Y}_\delta = \tilde{h}(I) + \tfrac{1}{2}y_3^2 + \delta Z(\theta, I, x_3), \quad (\theta, I, x_3, y_3) \in T^*\mathbb{T}^1 \times \mathbb{R}^2,$$

to which we can apply Theorem 6.1 again to get a NHIC foliated by action minimizing periodic orbits in the homology class $g = (1, 0) \in H_1(\mathbb{T}^2, \mathbb{Z})$. This gives the NHIC in Theorem C.1. Recovering the $\delta_4 \tilde{R}$ perturbation, diffusing orbits can be constructed moving along the NHIC (channel in $H^1(\mathbb{T}^3, \mathbb{R})$) up to a $o_{\delta_4 \to 0}(1)$-neighborhood of the pizza using the method of *a priori* unstable systems.

Step 3 *The cohomology equivalence.* As a result of the previous step, we have arrived at a neighborhood of the pizza where the cohomology class $c = (c_1, c_2, c_3)$ satisfies $\alpha_{\mathsf{G}}(c_1, c_2) \in (0, \Delta)$ and c_3 close to zero; see Theorem 6.3 for the definition of Δ. We now view the system Y_δ as a small perturbation of the subsystem G. By Theorem 6.3 and the upper-semicontinuity of the Mañé set, for small enough δ, the Mañé set $\tilde{\mathcal{N}}(c)$ when projected to $\mathbb{T}^2(\ni (x_1, x_2))$ does not cover $\mathbb{T}^2$. We apply the c-equivalence mechanism (Theorem 3.3) to get that the cohomology class (c_1, c_2, c_3) is c-equivalent to $(-c_1, -c_2, c_3)$.

Step 4 *The ladder climbing.* Here comes an intrinsic problem due to the high dimensionality. The two channels are *centrally symmetric* due to the reversibility of the mechanical system. Namely, the projection of the two channels to the c_3 coordinate axis, may not overlap. So for $c = (c_1, c_2, c_3)$ in one channel, the point $(-c_1, -c_2, c_3)$ does not lie in the opposite channel. We have to find a way to change c_3 to $-c_3$. The idea is to notice that restricting the system Y_δ to the NHIC (homeomorphic to $T^*\mathbb{T}^2$) obtained by applying Theorem 6.1 to G, we get $\overline{Y}_\delta$. The center manifold which is the phase space of $\overline{Y}_\delta$, has stable and unstable manifolds hence we are in a situation similar to Arnold's example. Restricted to an energy level of the system Y_δ, the energy of the subsystem $\overline{Y}_\delta$ is also fixed, so we get a curve $\alpha_{\overline{Y}_\delta}(c) =$const in $H^1(\mathbb{T}^2, \mathbb{R})$. Along this curve, we can move c_3 significantly by Arnold's mechanism, so we send

$$(-c_1, -c_2, c_3) \to (-c_1, -c_2, -c_3).$$

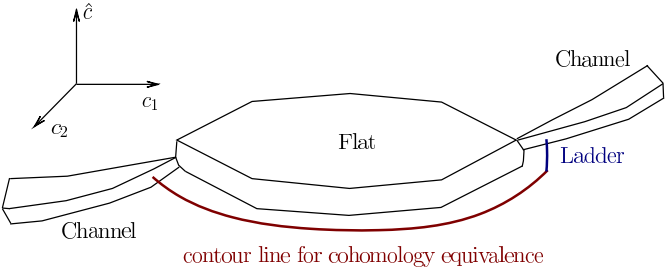

Figure 6. Turning around complete resonance: channel (NHIC), c-equivalence, and ladder.

To see the last mechanism clearly, we modify Arnold's Hamiltonian slightly to yield

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

In this system, for each $E > 0$ there exists diffusing orbit along which (y_1, y_2) moves arbitrarily on the circle $\{y_1^2 + y_2^2 = 2E\}$. In our case, the system Y_δ plays the role of H here and the subsystem $\overline{Y}_\delta$ plays the role of $\frac{1}{2}(y_1^2 + y_2^2)$ which lies on the NHIM $\{x_3 = y_3 = 0\}$.

As in the case of *a priori* unstable systems, we need a regularity result similar to Theorem 4.1 to show that the barrier functions $\overline{B}_c(x)$ of the system $\overline{Y}_\delta$ for $\alpha_{\overline{Y}_\delta}(c)$ =constant can be parametrized into a Hölder family. This is proved in [12].

We complete the sketch of the proof here and refer interested readers to [11] for more details.

Acknowledgments

We greatly appreciate the organizing committee, in particular Albert Fathi, for his efforts to getting people together. We would like to thank the participants of the program for attending our lectures and many valuable questions and comments. C.-Q. Cheng is supported by NNSF of China (No.11790272 and No.11631006). J. Xue is supported by NNSF of China (No.11790273) and Beijing Natural Science Foundation (Z180003).

References

[1] V. I. Arnold, "Instability of dynamical systems with many degrees of freedom", *Dokl. Akad. Nauk SSSR* **156** (1964), 9–12. In Russian; translated in *Sov. Math. Dokl.* **5**(1964), 581–585.

[2] V. I. Arnold, "The stability problem and ergodic properties for classical dynamical systems", pp. 387–392 in *Proceedings of International Congress of Mathematicians* (Moscow, 1966),

vol. 1, edited by I. G. Petrovsky, MIR, Moscow, 1968. Reprinted as pp. 107-113 in his *Collected Works*, *Volume II*, Springer, 2014.

[3] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Graduate Texts in Mathematics **60**, Springer, 1989.

[4] V. I. Arnold, "Mathematical problems in classical physics", pp. 1–20 in *Trends and perspectives in applied mathematics*, edited by L. Sirovich, Appl. Math. Sci. **100**, Springer, 1994.

[5] V. I. Arnold, *Arnold's problems*, Springer, 2004.

[6] P. Bernard, "The dynamics of pseudographs in convex Hamiltonian systems", *J. Amer. Math. Soc.* **21**:3 (2008), 615–669.

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

[8] C.-Q. Cheng, "Dynamics around the double resonance", *Camb. J. Math.* **5**:2 (2017), 153–228.

[9] C.-Q. Cheng, "Uniform hyperbolicity of invariant cylinder", *J. Differential Geom.* **106**:1 (2017), 1–43.

[10] C.-Q. Cheng, "The genericity of Arnold diffusion in nearly integrable Hamiltonian systems", *Asian J. Math.* **23**:3 (2019), 401–438.

[11] C.-Q. Cheng and J. Xue, "Arnold diffusion in nearly integrable Hamiltonian systems of arbitrary degrees of freedom", preprint, 2015. arXiv 1503.04153v2

[12] C.-Q. Cheng and J. Xue, "Order property and modulus of continuity of weak KAM solutions", *Calc. Var. Partial Differential Equations* **57**:2 (2018), Paper No. 65, 27.

[13] C.-Q. Cheng and J. Xue, "Variational approach to Arnold diffusion", *Sci. China Math.* **62**:11 (2019), 2103–2130.

[14] C.-Q. Cheng and J. Yan, "Existence of diffusion orbits in a priori unstable Hamiltonian systems", *J. Differential Geom.* **67**:3 (2004), 457–517.

[15] C.-Q. Cheng and J. Yan, "Arnold diffusion in Hamiltonian systems: a priori unstable case", *J. Differential Geom.* **82**:2 (2009), 229–277.

[16] C.-Q. Cheng and M. Zhou, "Global normally hyperbolic invariant cylinders in Lagrangian systems", *Math. Res. Lett.* **23**:3 (2016), 685–705.

[17] C.-Q. Cheng and M. Zhou, "C^1-continuation of periodic orbits from homoclinics", *Pacific J. Math.* **317**:1 (2022), 67–118.

[18] A. Delshams and G. Huguet, "Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems", *Nonlinearity* **22**:8 (2009), 1997–2077.

[19] A. Delshams, R. de la Llave, and T. M. Seara, "A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model", **179**:844 (2006), viii+141 in.

[20] A. Delshams, R. de la Llave, and T. M. Seara, "Geometric properties of the scattering map of a normally hyperbolic invariant manifold", *Adv. Math.* **217**:3 (2008), 1096–1153.

[21] A. Fathi, "Weak KAM theorem in Lagrangian dynamics", 2008, tinyurl.com/FathiKAM. 10th preliminary version.

[22] N. Fenichel, "Persistence and smoothness of invariant manifolds for flows", *Indiana Univ. Math. J.* **21** (1971/72), 193–226.

[23] V. Gelfreich, C. Simó, and A. Vieiro, "Dynamics of $4D$ symplectic maps near a double resonance", *Phys. D* **243** (2013), 92–110.

[24] M. Gidea and J.-P. Marco, "Diffusion along chains of normally hyperbolic cylinders", preprint, 2017. arXiv 1708.08314

[25] M. Gidea, R. de la Llave, and T. M-Seara, "A general mechanism of diffusion in Hamiltonian systems: qualitative results", *Comm. Pure Appl. Math.* **73**:1 (2020), 150–209.

[26] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant manifolds*, Lecture Notes in Mathematics **583**, Springer, 1977.

[27] N. Hitchin, "Stable bundles and integrable systems", *Duke Math. J.* **54**:1 (1987), 91–114.

[28] V. Kaloshin and K. Zhang, *Arnold diffusion for smooth systems of two and a half degrees of freedom*, Annals of Mathematics Studies **208**, Princeton University Press, 2020.

[29] X. Li, "On c-equivalence", *Sci. China Ser. A* **52**:11 (2009), 2389–2396.

[30] X. Li and C.-Q. Cheng, "Connecting orbits of autonomous Lagrangian systems", *Nonlinearity* **23**:1 (2010), 119–141.

[31] R. Mañé, "Generic properties and problems of minimizing measures of Lagrangian systems", *Nonlinearity* **9**:2 (1996), 273–310.

[32] J.-P. Marco, "Chains of compact cylinders for cusp-generic nearly integrable convex systems on $\mathbb{A}^3$", preprint, 2016. arXiv 1602.02399

[33] J. N. Mather, "Action minimizing invariant measures for positive definite Lagrangian systems", *Math. Z.* **207**:2 (1991), 169–207.

[34] J. N. Mather, "Variational construction of connecting orbits", *Ann. Inst. Fourier* (*Grenoble*) **43**:5 (1993), 1349–1386.

[35] D. N. Mèzer, "Arnold diffusion, I: Announcement of results", *Sovrem. Mat. Fundam. Napravl.* **2** (2003), 116–130.

[36] R. Moeckel, "Generic drift on Cantor sets of annuli", pp. 163–171 in *Celestial mechanics* (Evanston, IL, 1999), edited by A. Chenciner et al., Contemp. Math. **292**, Amer. Math. Soc., Providence, RI, 2002.

[37] D. Treschev, "Evolution of slow variables in a priori unstable Hamiltonian systems", *Nonlinearity* **17**:5 (2004), 1803–1841.

[38] M. Zhou, "Hölder regularity of weak KAM solutions in a priori unstable systems", *Math. Res. Lett.* **18**:1 (2011), 75–92.

[39] M. Zhou, "Infinity of minimal homoclinic orbits", *Nonlinearity* **24**:3 (2011), 931–939.

chengcq@nju.edu.cn *School of Mathematical Sciences, Nanjing Normal University, Nanjing, China*

jxue@tsinghua.edu.cn *Department of Mathematical Sciences, Tsinghua University, Beijing, China*