

Partial averaging and dynamics of the dominant Hamiltonian, with applications to Arnold diffusion

V. Kaloshin*, K. Zhang†

October 8, 2014

Abstract

We discuss partial averaging for fiber convex nearly integrable systems and the properties of its scaling limit. We define a subclass of these limits, which we call dominant, and analyze them from the point of view of the perturbation theory and weak KAM theory. One of the main motivations is the study of Arnold diffusion. In the appendix, using the dominant Hamiltonian systems, we propose a scheme of proving Arnold diffusion in arbitrary degrees of freedom.

Contents

1	Introduction	2
2	Formulation of results	6
2.1	Partial averaging and the slow Hamiltonians	6
2.2	Dominant Hamiltonians	9
2.3	The rescaling limit	10
2.4	The variational aspect of dominant Hamiltonians	12
3	The choice of basis and averaging	14
3.1	The choice of the basis	14
3.2	Estimating the weak potential	17
4	Strong and slow systems of dominant Hamiltonians	18
4.1	Vector fields of dominant Hamiltonians	18
4.2	The slow Lagrangian	22

*University of Maryland at College Park (vadim.kaloshin@gmail.com)

†University of Toronto (kzhang@math.utoronto.edu)

5	Weak KAM solutions of dominant Hamiltonians and convergence	24
5.1	Weak KAM solutions of Tonelli Lagrangian	24
5.2	Minimizers of strong and slow Lagrangians, their a priori compactness	27
5.3	Approximate Lipschitz property of weak KAM solutions	30
5.4	The alpha function and rotation number estimate	30
5.5	Convergence of weak KAM solutions	32
6	The Mañe and the Aubry sets and the barrier function	34
6.1	The Mañe set and barrier function	35
6.2	Semi-continuity of the Aubry set	36
7	Technical estimates on weak KAM solutions	37
7.1	Approximate Lipschitz property in the strong component	38
7.2	Finer decomposition of the slow Lagrangian	41
7.3	Approximate Lipschitz property in the weak component	44
A	Diffusion path with dominant structure	49
A.1	Diffusion path for Arnold diffusion	49
A.2	Nondegeneracy conditions for Arnold diffusion	52
A.3	Properties of the nondegeneracy condition	54
A.4	Construction of a diffusion path and surgery of resonant manifolds	56
A.4.1	An initial step of the induction	57
A.4.2	Step 2 of the induction	58
A.4.3	Step $s + 1$ of the induction	62
A.4.4	Concluding the induction	66
B	Useful facts from linear algebra	67

1 Introduction

Consider a nearly integrable system with $n\frac{1}{2}$ degrees of freedom

$$H_\varepsilon(\theta, p, t) = H_0(p) + \varepsilon H_1(\theta, p, t), \quad \theta \in \mathbb{T}^n, p \in \mathbb{R}^n, t \in \mathbb{T}. \quad (1.1)$$

We will restrict to the case where the integrable part H_0 is strictly convex, more precisely, we assume that there is $D > 1$ such that

$$D^{-1}\text{Id} \leq \partial_{pp}^2 H_0(p) \leq D\text{Id}$$

as quadratic forms, where Id denotes the identity matrix.

The main motivation behind this work is the question of Arnold diffusion, that is, topological instability for the system H_ε . Arnold provided the first example in [Arn64], and asks ([Arn63; Arn68; Arn94]) whether topological instability is “typical”

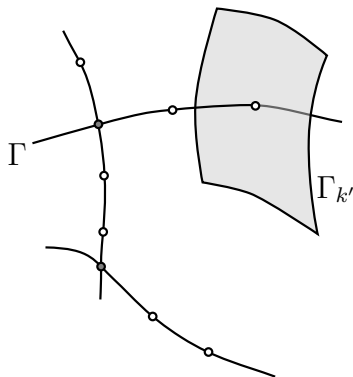


Figure 1: Diffusion path and essential resonances in $n = 3$. The hollow dots requires crossing, while the grey dots requires switching

in nearly integrable systems with $n \geq 2$ (the system is stable when $n = 1$, due to low dimensionality).

It is well known that the instabilities of nearly integrable systems occurs along resonances. Given an integer vector $k = (\bar{k}, k^0) \in \mathbb{Z}^n \times \mathbb{Z}$ with $\bar{k} \neq 0$, we define the resonant submanifold to be $\Gamma_k = \{p \in \mathbb{R}^n : k \cdot (\omega(p), 1) = 0\}$, where $\omega(p) = \partial_p H_0(p)$. More generally, we consider a subgroup Λ of \mathbb{Z}^{n+1} which does not contain vectors of the type $(0, \dots, 0, k^0)$, called a *resonance lattice*. The *rank* of Λ is the dimension of the real subspace containing it. Then for a rank d resonance lattice Λ , we define

$$\Gamma_\Lambda = \bigcap \{\Gamma_k : k \in \Lambda\} = \bigcap_{i=1}^d \Gamma_{k_i},$$

where $\{k_1, \dots, k_d\}$ is any linear independent set in Λ . We call such Γ_Λ a *d-resonance submanifold* (*d-resonance* for short), which is a co-dimension d submanifold of \mathbb{R}^n , and in particular, an n -resonant submanifold is a single point. We say that Λ is *irreducible* if it is not contained in any lattices of the same rank, or equivalently, $\text{span}_{\mathbb{R}} \Lambda \cap \mathbb{Z}^{n+1} = \Lambda$.

We now focus on the diffusion that occurs along a connected net of $(n-1)$ -resonances, with each $(n-1)$ -resonance being a curve in \mathbb{R}^n . Let us first consider diffusion along a single $(n-1)$ -resonance Γ . It is shown in [BKZ11] that generically, diffusion indeed occur along Γ , except for a finite subset of n -resonances (called *the strong resonances*) which divides Γ into disconnected components. A strong resonance can be viewed as the intersection of Γ with a transversal 1-resonance manifold $\Gamma_{k'}$ (see Figure 1).

The main obstacle to proving diffusion along Γ reduces to whether the diffusion can “cross” the strong resonances. In a more general diffusion path that contains two intersecting $(n-1)$ -resonances Γ_1 and Γ_2 , the intersection is an n -resonance which by definition is *essential*. The question is then whether one can travel along Γ_1 and then “switch” to Γ_2 at the intersection. Solution to either problem requires an understanding of the system near an n -resonance.

For an n -resonance $\{p_0\} = \Gamma_\Lambda$, we assume that Λ is irreducible, and $\mathcal{B} = \{k_1, \dots, k_n\}$ is its basis over \mathbb{Z} . The study of diffusion near p_0 reduces to the study of a particular *slow*

system defined on $\mathbb{T}^n \times \mathbb{R}^n$, denoted $H_{p_0, \mathcal{B}}^s$. More precisely, in an $O(\sqrt{\varepsilon})$ -neighborhood of p_0 , the system H_ε admits the normal form

$$H_{p_0, \mathcal{B}}^s(\varphi, I) + \sqrt{\varepsilon}P(\varphi, I, \tau), \quad \varphi \in \mathbb{T}^n, I \in \mathbb{T}^n, \tau \in \sqrt{\varepsilon}\mathbb{T},$$

and therefore can be seen as a fast periodic perturbation to $H_{p_0, \mathcal{B}}^s$. We stress that the slow system depends on the choice of the basis \mathcal{B} . Such averaged systems were studied in [Mat08].

When $n = 2$, the slow system is a 2 degrees of freedom mechanical system, the structure of its (minimal) orbits is well understood. This fact underlies the results on Arnold diffusion in two and half degrees of freedom (see [Mat03], [Mat08], [Mat11], [Che13], [KZ13], [GK14b], [KMV04], [Mar12a], [Mar12b]). This is no longer the case when $n > 2$, which is a serious obstacle to proving Arnold diffusion in higher degrees of freedom. In [KZ14] it is proposed that we can sidestep this difficulty by using *dimension reduction*: using existence of normally hyperbolic cylinders to restrict the system to a lower dimensional manifold. This approach only works when the slow system has a particular *dominant structure*, which is the topic of this paper.

It is more convenient to define the slow system for any p_0 and any d -resonance $d \leq n$. For $p_0 \in \mathbb{R}^n$, an irreducible rank d resonance lattice Λ , and its basis $\mathcal{B} = [k_1, \dots, k_d]$, the slow system is

$$H_{p_0, \mathcal{B}}^s(\varphi, I) = K_{p_0, \mathcal{B}}(I) - U_{p_0, \mathcal{B}}(\varphi), \quad \varphi \in \mathbb{T}^d, I \in \mathbb{T}^d. \quad (1.2)$$

Suppose the fourier expansion of H_1 is $\sum_{k \in \mathbb{Z}^{n+1}} h_k(p) e^{2\pi i k \cdot (\theta, t)}$, then

$$K_{p_0, \mathcal{B}}(I) = \frac{1}{2} \partial_{pp}^2 H_0(p_0) (I_1 \bar{k}_1 + \dots + I_d \bar{k}_d) \cdot (I_1 \bar{k}_1 + \dots + I_d \bar{k}_d), \quad (1.3)$$

$$U_{p_0, \mathcal{B}}(\varphi_1, \dots, \varphi_d) = - \sum_{l \in \mathbb{Z}^d} h_{l_1 k_1 + \dots + l_d k_d}(p) e^{2\pi i (l_1 \varphi_1 + \dots + l_d \varphi_d)}. \quad (1.4)$$

The system $H_{p_0, \mathcal{B}}^s$ is only dynamically meaningful when $p_0 \in \Gamma_\Lambda$. However, the more general set up allows us to embed the meaningful slow systems into a nicer space.

We say that the resonance lattice Λ admits a dominant structure if it contains an irreducible lattice Λ^{st} of rank $m < d$, such that

$$M(\Lambda | \Lambda^{\text{st}}) := \min_{k \in \Lambda \setminus \Lambda^{\text{st}}} |k| \gg \max_{k \in \mathcal{B}^{\text{st}}} |k|, \quad (1.5)$$

where $|k| = \sup_i |k_i|$ is the sup-norm. Given the relation $\Lambda^{\text{st}} \subset \Lambda$, one can choose an *adapted basis* $\mathcal{B} = [k_1, \dots, k_d]$ of Λ , meaning that $\mathcal{B}^{\text{st}} = [k_1, \dots, k_m]$ is a properly ordered basis of Λ^{st} . In this case we have two slow systems $H_{p_0, \mathcal{B}^{\text{st}}}^s$ and $H_{p_0, \mathcal{B}}^s$.

Our main results can be formulated as follows:

Main Result. *For a fixed Λ^{st} of rank m , and each rank d , $m \leq d \leq n$ irreducible lattice $\Lambda \supset \Lambda^{\text{st}}$, there exists an adapted basis \mathcal{B} such that:*

1. (Geometrical) As $M(\Lambda|\Lambda^{\text{st}}) \rightarrow \infty$, the vector field of $H_{p_0, \mathcal{B}}^s$ converges to a trivial lift of the vector field of $H_{p_0, \mathcal{B}^{\text{st}}}^s$. In particular, if $H_{p_0, \mathcal{B}^{\text{st}}}^s$ admits a normally hyperbolic invariant cylinder, so does $H_{p_0, \mathcal{B}}^s$.
2. (Variational) As $M(\Lambda|\Lambda^{\text{st}}) \rightarrow \infty$, the weak KAM solution to $H_{p_0, \mathcal{B}}^s$ converges uniformly to a trivial lift of a weak KAM solution of $H_{p_0, \mathcal{B}^{\text{st}}}^s$. We also obtain corollaries concerning the limits of Mañe, Aubry sets, rotation number of minimal measure, and Peierl's barrier function (whose precise definitions are given later¹ for more details).

The statement that $H_{p_0, \mathcal{B}^{\text{st}}}^s$ approximates $H_{p_0, \mathcal{B}}^s$ is related to the classic result of partial averaging (see for example [AKN06]). The statement $\min_{k \in \Lambda \setminus \Lambda^{\text{st}}} |k| \gg \max_{k \in \Lambda^{\text{st}}} |k|$ says that the resonances in Λ^{st} is much *stronger* than the rest of the resonances in Λ . Partial averaging says that the weaker resonances contributes to smaller terms in a normal form.

However, our treatment of the partial averaging theory is quite different from the classical theory. By looking at the rescaling limit, we study the property of the averaging independent of the small parameter ε . Many subtleties arise, including the sensitive dependence of the slow system on the basis, and special meaning on the type of limit taken.

In [Mat08], John Mather developed a theory of (partial) averaging for a nearly integrable Lagrangian system

$$L_\varepsilon(\theta, v, t) = L_0(v) + \varepsilon L_1(\theta, v, t), \quad \text{where } (\theta, v, t) \in \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T}.$$

In particular, it is shown that the slow system relative to a resonant lattice can be defined on the tangent bundle of a sub-torus $\mathbb{T}^d \times \mathbb{R}^d$. Quantitative estimates on the action of minimizing orbits of the original system versus the slow system are obtained. Our variational result is related to [Mat08], but different in many ways. We work with the scaling limit system, and the small parameter ε does not show up in our analysis. We also avoid quantitative estimates (in the statement of the theorem) and obtain a limit theorem in weak KAM solutions. This allows us to take consecutive limits, which is very useful for our construction of diffusion path in higher dimensions (see Appendix A).

The formulation of the limit theorem in weak KAM solution requires special care. Usually, one uses the Tonelli convergence (convergence of Lagrangian within the Tonelli family, see [Ber10]). In our setup, $H_{p_0, \mathcal{B}^{\text{st}}}^s$ and $H_{p_0, \mathcal{B}}^s$ are defined on different spaces, and $H_{p_0, \mathcal{B}}^s$ converges to a trivial lift of $H_{p_0, \mathcal{B}^{\text{st}}}^s$. Moreover, the corresponding Lagrangians *do not* converge. We nevertheless obtain the convergence of weak KAM solutions.

While this paper is mainly motivated by Arnold diffusion, we hope our treatment of partial averaging is of independent interest.

The plan of the paper is as follows. The rigorous formulation of the results will be presented in section 2. The choice of the basis is handled in section 3, and the estimates

¹ see Theorem 2.3 and Proposition 6.1

of the vector fields, including the geometrical result is in section 4. The variational aspect is more involved, and occupies sections 5 and 6, with some technical estimates deferred to section 7.

As a demonstration of our theory, in Appendix A we show that one can construct a connected net of $(n - 1)$ -resonances, such that all strong resonances have the dominant structure. We state corollaries of having the dominant structure, which include existence of 3-dimensional normally hyperbolic invariant cylinders carrying families of properly chosen Aubry sets. In [KZ13] and [KZ14] using these structures we prove existence of Arnold diffusion. This net of diffusion paths also can be chosen to be γ -dense for any pre-determined $\gamma > 0$. We show that for a “typical” H_ε (for any $n > 2$), such a net exists and expect to prove Arnold diffusion along this net in the future publication.

2 Formulation of results

2.1 Partial averaging and the slow Hamiltonians

As before we write $H_1(\theta, p, t) = \sum_{k \in \mathbb{Z}^{d+1}} h_k(p) e^{2\pi i k \cdot (\theta, t)}$. For an irreducible resonance lattice Λ , we define the resonant component of H_1 relative to Λ by

$$[H_1]_\Lambda(\theta, p, t) = \sum_{k \in \Lambda} h_k(p) e^{2\pi i k \cdot (\theta, t)}. \quad (2.1)$$

If p_0 is maximally resonant and $\Lambda = \Lambda_{p_0}$, then in a $O(\varepsilon^{\frac{1}{2}})$ neighborhood of p_0 there exists a symplectic coordinate change Φ , such that (see for example [Loc92])

$$N_\varepsilon = H_\varepsilon \circ \Phi = H_0(p) + \varepsilon [H_1]_{\Lambda_{p_0}} + O(\varepsilon^{\frac{3}{2}}). \quad (2.2)$$

Consider an irreducible resonant lattice $\Lambda^{\text{st}} \subsetneq \Lambda$. Then

$$[H_1]_\Lambda = [H_1]_{\Lambda^{\text{st}}} - ([H_1]_\Lambda - [H_1]_{\Lambda^{\text{st}}}). \quad (2.3)$$

The function $[H_1]_\Lambda - [H_1]_{\Lambda^{\text{st}}}$ contains only fourier modes $k \in \Lambda \setminus \Lambda^{\text{st}}$, called the *additional resonances*. For $M(\Lambda | \Lambda^{\text{st}})$ defined in (1.5) and any $\delta > 0$, if $M(\Lambda | \Lambda^{\text{st}})$ is large enough, $[H_1]_\Lambda - [H_1]_{\Lambda^{\text{st}}} = O(\delta)$. In particular, if $\Lambda = \Lambda_p$, (2.2) becomes

$$N_\varepsilon = H_0 + \varepsilon [H_1]_{\Lambda^{\text{st}}} + O(\varepsilon \delta).$$

In other words, if the additional resonance relations have large norm, the perturbation consists of the Λ^{st} -resonant component and a small remainder. This is a version of the classical partial averaging, see for example [AKN06].

We introduce the slow system which is related to a scaling limit of (2.2). Let Λ be a rank d resonance lattice and $\mathcal{B} = [k_1, \dots, k_d]$ be a basis over \mathbb{Z} . Denote

$$\varphi_1 = k_1 \cdot (\theta, t), \dots, \varphi_d = k_d \cdot (\theta, t),$$

and let

$$Z_{\mathcal{B}}(\varphi_1, \dots, \varphi_d, p) = \sum_{l \in \mathbb{Z}^d} h_{l_1 k_1 + \dots + l_d k_d}(p) e^{2\pi i(l_1 \varphi_1 + \dots + l_d \varphi_d)}, \quad (2.4)$$

then $[H_1]_{\Lambda}(\theta, p, t) = Z_{\mathcal{B}}(k_1 \cdot (\theta, t), \dots, k_d \cdot (\theta, t), p)$.

We denote $Q_0(p) = \partial_{pp}^2 H_0(p) \in \text{Sym}(n)$, where $\text{Sym}(n)$ denote the space of $n \times n$ symmetric matrices. Define

$$Q(p) = \begin{bmatrix} Q_0(p) & 0 \\ 0 & 0 \end{bmatrix} \in \text{Sym}(n+1). \quad (2.5)$$

To motivate this definition, note that the non-autonomous Hamiltonian $H_{\varepsilon}(\theta, p, t)$ can be written in autonomous form $G(\theta, p, t, E) = H_{\varepsilon}(\theta, p, t) + E$, and $Q(p)$ is the Hessian of $H_0(p) + E$.

For $p_0 \in \Gamma_{\Lambda}$, we define the slow Hamiltonian $H_{p_0, \mathcal{B}}^s(\varphi, I)$, as in (1.2). Note that for a resonance lattice Λ , the vectors $\bar{k}_1, \dots, \bar{k}_d$ are linearly independent, which implies $K_{p_0, \mathcal{B}}$ is *strictly* positive definite. The slow system is a classical mechanical system. We stress that the slow system depends not only on Λ , but also the choice of the basis \mathcal{B} .

When p_0 is maximally resonant and $\Lambda = \Lambda_{p_0}$ has rank n , the system $H_{p_0, \mathcal{B}}^s$ is a scaling limit of (2.2). This no longer holds when $d < n$, nevertheless, the slow system is still related to the averaging property of H_{ε} near p_0 .

In this paper, we study the partial averaging theory of the slow system $H_{p_0, \mathcal{B}}^s$. Fix an irreducible resonance lattice Λ^{st} of rank m , called *the strong lattice*, and an ordered basis $\mathcal{B}^{\text{st}} = [k_1, \dots, k_m]$. Let $\Lambda \supset \Lambda^{\text{st}}$ be an irreducible resonance lattice of rank d , where $m < d \leq n$, and $p_0 \in \Gamma_{\Lambda} \subset \Gamma_{\Lambda^{\text{st}}}$. A basis \mathcal{B} of Λ is called an *adapted basis* if its first m components equals \mathcal{B}^{st} , i.e. $\mathcal{B} = [k_1, \dots, k_m, \dots, k_d]$. When \mathcal{B} is an adapted basis, the corresponding slow system is denoted $H_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^s$, where

$$\begin{aligned} \mathcal{B}^{\text{st}} &= [k_1, \dots, k_m] = [k_1^{\text{st}}, \dots, k_m^{\text{st}}], \\ \mathcal{B}^{\text{wk}} &= [k_{m+1}, \dots, k_d] = [k_1^{\text{wk}}, \dots, k_{d-m}^{\text{wk}}]. \end{aligned}$$

Denote

$$\begin{aligned} \varphi^{\text{st}} &= (\varphi_1, \dots, \varphi_m), & \varphi^{\text{wk}} &= (\varphi_{m+1}, \dots, \varphi_d), \\ I^{\text{st}} &= (I_1, \dots, I_m), & I^{\text{wk}} &= (I_{m+1}, \dots, I_d), \end{aligned}$$

similar to (2.3), we have

$$\begin{aligned} H_{p_0, \mathcal{B}}^s(\varphi, I) &= H_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, I^{\text{st}}, I^{\text{wk}}) \\ &= K_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}(I^{\text{st}}, I^{\text{wk}}) - U_{p_0, \mathcal{B}^{\text{st}}}(\varphi^{\text{st}}) - U_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \end{aligned}$$

where

$$U_{p_0, \mathcal{B}^{\text{st}}}(\varphi^{\text{st}}) = -Z_{\mathcal{B}^{\text{st}}}(\varphi^{\text{st}}, p_0), \quad U_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}) = -(Z_{\mathcal{B}} - Z_{\mathcal{B}^{\text{st}}})(\varphi^{\text{st}}, \varphi^{\text{wk}}, p_0).$$

The term $U_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^{\text{wk}}$ can be further decomposed to separate the angles φ_j^{wk} in a diagonal fashion. To do this we define, for each $1 \leq i \leq d$, $\mathcal{B}_i = [k_1, \dots, k_i]$ is an ordered basis and Λ_i is the corresponding rank i lattice. Then $\mathcal{B}_m = \mathcal{B}^{\text{st}} \subset \dots \subset \mathcal{B}_d = \mathcal{B}$. For $m < i \leq d$, we write

$$\begin{aligned} U_{p_0, \mathcal{B}_{i-1}, \mathcal{B}_i}^{\text{wk}} &= -(Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}})(\varphi_1, \dots, \varphi_i, p_0) \\ &= -(Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}})(\varphi^{\text{st}}, \varphi_1^{\text{wk}}, \dots, \varphi_{i-m}^{\text{wk}}, p_0). \end{aligned}$$

Then the slow system takes the form

$$H_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^s = K_{p_0, \mathcal{B}} - U_{p_0, \mathcal{B}^{\text{st}}} - U_{p_0, \mathcal{B}_m, \mathcal{B}_{m+1}}^{\text{wk}} - \dots - U_{\mathcal{B}_{d-1}, \mathcal{B}_d}^{\text{wk}}. \quad (2.6)$$

We also consider the slow system to Λ^{st}

$$\begin{aligned} H_{p_0, \mathcal{B}^{\text{st}}}^s(\varphi^{\text{st}}, I^{\text{st}}) &= K_{p_0, \mathcal{B}^{\text{st}}}(I^{\text{st}}) - U_{p_0, \mathcal{B}^{\text{st}}}(\varphi^{\text{st}}) \\ &= K_{p_0, \mathcal{B}}(I^{\text{st}}, 0) - U_{p_0, \mathcal{B}^{\text{st}}}(\varphi^{\text{st}}). \end{aligned}$$

The justification of the Main Theorem takes several steps.

- Theorem 2.1 implies that the slow system $H_{p_0, \mathcal{B}}$ belong to a specific class of systems which we call the *dominant Hamiltonians*. This is introduced in section 2.2.
- In section 2.3 and Theorem 2.2 we justify the Main Theorem in the sense of rescaling limit of the vector field.
- In section 2.4 and Theorem 2.3 we provide an variational version of the Main Theorem, and prove semi-continuity of the weak KAM solutions.

Theorem 2.1. *Let $\Lambda^{\text{st}} \subset \Lambda$ be irreducible resonance lattices of rank m and d resp., where $m < d$. Fix an ordered basis $\mathcal{B}^{\text{st}} = [k_1^{\text{st}}, \dots, k_m^{\text{st}}]$ of Λ^{st} . Suppose H_1 is C^r with $r > 3n - 2m + 6$, and $\|H_1\|_{C^r} = 1$. Then there exists a constant $\kappa = \kappa(H_0, \Lambda^{\text{st}}, n) > 1$, integer vectors $\mathcal{B}^{\text{wk}} = [k_1^{\text{wk}}, \dots, k_{d-m}^{\text{wk}}]$ with $[\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}]$ forming an adapted basis, such that the following hold.*

1. For any $1 \leq i < j \leq d - m$, $|k_i^{\text{wk}}| < \kappa(1 + |k_j^{\text{wk}}|)$.
2. For $1 \leq j \leq d - m$, we have

$$\|U_{p_0, \mathcal{B}_{j+m-1}, \mathcal{B}_{j+m}}^{\text{wk}}\|_{C^2} \leq \|Z_{\mathcal{B}_{j+m}} - Z_{\mathcal{B}_{j+m-1}}\|_{C^2} \leq \kappa |k_j^{\text{wk}}|^{-r+n+2(d-m)+4}.$$

Remark. *Item 2 implies that as $M(\Lambda|\Lambda^{\text{st}}) \rightarrow \infty$, for the specifically chosen basis, we have $\|U_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}^{\text{wk}}\|_{C^2} \rightarrow 0$.*

Item 1 says that vectors in \mathcal{B}^{wk} are approximately in an increasing order. This, when combined with item 2, implies that the norm of the weak potentials $U_{p_0, \mathcal{B}_{j+m-1}, \mathcal{B}_{j+m}}$ are approximately in an decreasing order.

This theorem is proven in section 3.

2.2 Dominant Hamiltonians

We start with the following data:

1. A C^2 function $Q_0 : \mathbb{R}^n \rightarrow \text{Sym}(n)$, where $\text{Sym}(n)$ denote the space of $n \times n$ symmetric matrices. Define as before $Q(p) = \begin{bmatrix} Q_0(p) & 0 \\ 0 & 0 \end{bmatrix}$, and assume $D^{-1}\text{Id} \leq Q_0 \leq D\text{Id}$.
2. An irreducible resonance lattice $\Lambda^{\text{st}} \subset \mathbb{Z}^{n+1}$ of rank $1 \leq m < n$, and a basis $\mathcal{B}^{\text{st}} = [k_1^{\text{st}}, \dots, k_m^{\text{st}}]$.
3. Constant $\kappa > 1$ and $q > 1$.

We continue to use the notation $(k_1, \dots, k_m) = (k_1^{\text{st}}, \dots, k_m^{\text{st}})$ and $(k_{m+1}, \dots, k_d) = (k_1^{\text{wk}}, \dots, k_{d-m}^{\text{wk}})$, and apply the same convention to the variables φ and I . Define

$$\Omega^{m,d} := (\mathbb{Z}^{n+1})^d \times \mathbb{R}^n \times C^2(\mathbb{T}^m) \times C^2(\mathbb{T}^{m+1}) \times \dots \times C^2(\mathbb{T}^d), \quad \mathcal{H}^s : \Omega^{m,d} \rightarrow C^2(\mathbb{T}^d \times \mathbb{R}^d),$$

with

$$\begin{aligned} & (\mathcal{B}^{\text{st}} = [k_1^{\text{st}}, \dots, k_m^{\text{st}}], \mathcal{B}^{\text{wk}} = [k_1^{\text{wk}}, \dots, k_{d-m}^{\text{wk}}], p_0, U^{\text{st}}, \mathcal{U}^{\text{wk}} = \{U_1^{\text{wk}}, \dots, U_{d-m}^{\text{wk}}\}) \mapsto \\ \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p_0, U^{\text{st}}, \mathcal{U}^{\text{wk}}) &= K_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}(I) - U^{\text{st}}(\varphi_1, \dots, \varphi_m) - \sum_{j=1}^{d-m} U_j^{\text{wk}}(\varphi_1, \dots, \varphi_{j+m}), \end{aligned}$$

where

$$K_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}(I) = \frac{1}{2} Q(p_0) (k_1 I_1 + \dots + k_d I_d) \cdot (k_1 I_1 + \dots + k_d I_d).$$

We equip $\Omega^{m,d}$ with the product topology, with discrete topology on k_j^{wk} and the standard norms on other components. \mathcal{H}^s is smooth in $p_0, U^{\text{st}}, U_1^{\text{wk}}, \dots, U_{d-m}^{\text{wk}}$. Let $\Omega^{m,d}(\mathcal{B}^{\text{st}})$ be the subset of $\Omega^{m,d}$ with fixed \mathcal{B}^{st} .

We define $\Omega_{\kappa,q}^{m,d}(\mathcal{B}^{\text{st}}) \subset \Omega^{m,d}(\mathcal{B}^{\text{st}})$ to be the tuple $(\mathcal{B}^{\text{wk}}, p_0, U^{\text{st}}, \mathcal{U}^{\text{wk}})$ satisfying the following conditions:

1. For any $1 \leq i < j \leq d - m$, $|k_i^{\text{wk}}| \leq \kappa (1 + |k_j^{\text{wk}}|)$.
2. For each $1 \leq j \leq d - m$, $\|U_j^{\text{wk}}\|_{C^2} \leq \kappa |k_j^{\text{wk}}|^{-q}$.

Each element in $\mathcal{H}^s(\Omega_{\kappa,q}^{m,d})$ is called an (m, d) -dominant Hamiltonian with constants (κ, q) . Define

$$\mu(\mathcal{B}^{\text{wk}}) = \min_{1 \leq j \leq d-m} |k_j^{\text{wk}}|,$$

then in $\Omega_{\kappa,q}^{m,d}(\mathcal{B}^{\text{st}})$, we have $\|U_j^{\text{wk}}\| \leq \kappa \mu(\mathcal{B}^{\text{wk}})^{-q}$, i.e. the weak potential $U^{\text{wk}} = \sum_{j=1}^{d-m} U_j^{\text{wk}} \rightarrow 0$ as $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$.

We now restate Theorem 2.1 using the formal definition.

Theorem (Theorem 2.1 restated). *Under the assumptions of Theorem 2.1, there exists a constant $\kappa = \kappa(H_0, \mathcal{B}^{\text{st}}, n) > 1$, integer vectors $\mathcal{B}^{\text{wk}} = [k_1^{\text{wk}}, \dots, k_{d-m}^{\text{wk}}]$ with $\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}$ forming an adapted basis, such that*

$$(\mathcal{B}^{\text{wk}}, p, U_{p_0, \mathcal{B}^{\text{st}}}, (U_{p_0, \mathcal{B}_m, \mathcal{B}_{m+1}}, \dots, U_{p_0, \mathcal{B}_{d-1}, \mathcal{B}_d})) \in \Omega_{\kappa, r-n-2(d-m)-4}^{m,d}(\mathcal{B}^{\text{st}}).$$

The strong Hamiltonian is defined by the mapping

$$\mathcal{H}^{\text{st}} : \mathbb{R}^n \times C^2(\mathbb{T}^m) \rightarrow C^2(\mathbb{T}^m \times \mathbb{R}^m), \quad \mathcal{H}^{\text{st}}(p_0, U^{\text{st}}) = K_{p_0, \mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}}(I^{\text{st}}, 0) - U^{\text{st}}(\varphi^{\text{st}}).$$

We extend the definition to $\Omega^{m,d}$ by writing $\mathcal{H}^{\text{st}}(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p_0, U^{\text{st}}, \mathcal{U}^{\text{wk}}) = \mathcal{H}^{\text{st}}(p_0, U^{\text{st}})$. In the next section, we formalize the Heuristic Theorem in the sense of rescaling limit of vector fields.

2.3 The rescaling limit

We fix $\mathcal{B}^{\text{st}}, \kappa > 1$ and $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m,d}(\mathcal{B}^{\text{st}})$. Denote

$$H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}), \quad H^{\text{st}} = \mathcal{H}^{\text{st}}(p, U^{\text{st}}).$$

Then

$$H^s(\varphi, I) = K(I) - U^{\text{st}}(\varphi^{\text{st}}) - U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \quad H^{\text{st}}(\varphi^{\text{st}}, I^{\text{st}}) = K(I^{\text{st}}, 0) - U^{\text{st}}(\varphi^{\text{st}}),$$

where $U^{\text{wk}} = \sum_{j=1}^{d-m} U_j^{\text{wk}}$. As $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$, we have $\|U^{\text{wk}}\|_{C^2} \rightarrow 0$. However, $K(I^{\text{st}}, I^{\text{wk}})$ is not a small perturbation of $K(I^{\text{st}}, 0)$, in fact, $K(I^{\text{st}}, I^{\text{wk}})$ becomes unbounded as $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$. To justify the Heuristic Theorem, we perform a coordinate change (2.10), and a rescaling.

We write

$$\partial_{II}^2 K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \quad A = \partial_{I^{\text{st}} I^{\text{st}}}^2 K, B = \partial_{I^{\text{st}} I^{\text{wk}}}^2 K, C = \partial_{I^{\text{wk}} I^{\text{wk}}}^2 K, \quad (2.7)$$

then

$$(A)_{ij} = (k_i^{\text{st}})^T Q k_j^{\text{st}}, \quad (B)_{ij} = (k_i^{\text{st}})^T Q k_j^{\text{wk}}, \quad (C)_{ij} = (k_i^{\text{wk}})^T Q k_j^{\text{wk}}. \quad (2.8)$$

Note in particular that $A = \partial_{I^{\text{st}} I^{\text{st}}}^2 H^{\text{st}}$. The Hamiltonian equation for H^s reads

$$\begin{cases} \dot{\varphi}^{\text{st}} = AI^{\text{st}} + BI^{\text{wk}}, & \dot{\varphi}^{\text{wk}} = B^T I^{\text{st}} + CI^{\text{wk}}, \\ \dot{I}^{\text{st}} = \partial_{\varphi^{\text{st}}} U, & \dot{I}^{\text{wk}} = \partial_{\varphi^{\text{wk}}} U, \end{cases} \quad (2.9)$$

where $U = U^{\text{st}} + U^{\text{wk}}$. We make the coordinate change

$$(\varphi^{\text{st}}, I^{\text{st}}, \varphi^{\text{wk}}, I^{\text{wk}}) \mapsto (\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}}), \quad v^{\text{st}} = AI^{\text{st}} + BI^{\text{wk}}. \quad (2.10)$$

This is a ‘‘half Lagrangian’’ setting in the sense that $(\varphi^{\text{st}}, I^{\text{st}})$ is converted to Lagrangian setup, while $(\varphi^{\text{wk}}, I^{\text{wk}})$ keep the Hamiltonian format. Using $I^{\text{st}} = A^{-1}v^{\text{st}} - A^{-1}BI^{\text{wk}}$, we get

$$\begin{cases} \dot{\varphi}^{\text{st}} = v^{\text{st}}, & \dot{\varphi}^{\text{wk}} = B^T A^{-1}v^{\text{st}} - \tilde{C}I^{\text{wk}}, \\ \dot{v}^{\text{st}} = A\partial_{\varphi^{\text{st}}}U + B\partial_{\varphi^{\text{wk}}}U, & \dot{I}^{\text{wk}} = \partial_{\varphi^{\text{wk}}}U, \end{cases} \quad (2.11)$$

where $\tilde{C} = C - B^T A^{-1}B$. We denote by $X^s(\varphi^{\text{st}}, v^{\text{st}}, \varphi^{\text{wk}}, I^{\text{wk}})$ the vector field of the above equation, defined on the universal cover $\mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^m \times \mathbb{R}^{d-m}$. In this form the strong components decouple from the weak ones as $\partial_{\varphi^{\text{wk}}}U \rightarrow 0$.

The Euler-Lagrange equation for H^{st} is

$$\dot{\varphi}^{\text{st}} = v^{\text{st}}, \quad \dot{v}^{\text{st}} = A\partial_{\varphi^{\text{st}}}U^{\text{st}},$$

whose vector field we denote by $X^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}})$. We also consider its lift to the universal cover $\mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^m \times \mathbb{R}^{d-m}$

$$\begin{cases} \dot{\varphi}^{\text{st}} = v^{\text{st}}, & \dot{\varphi}^{\text{wk}} = 0, \\ \dot{v}^{\text{st}} = 0, & \dot{I}^{\text{wk}} = \partial_{\varphi^{\text{wk}}}U, \end{cases} \quad (2.12)$$

whose vector field we denote by $X_L^{\text{st}}(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}})$. We show that X_L^{st} is a rescaling limit of X^s .

Given $1 \geq \sigma_1 \geq \dots \geq \sigma_{d-m} > 0$, let $\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{d-m}\}$. We define a rescaling coordinate change $\Phi_\Sigma : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$ by

$$\Phi_\Sigma : (\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}}) \mapsto (\varphi^{\text{st}}, \tilde{\varphi}^{\text{wk}}, v^{\text{st}}, \tilde{I}^{\text{wk}}) := (\varphi^{\text{st}}, \Sigma^{-1}\varphi^{\text{wk}}, v^{\text{st}}, \Sigma I^{\text{wk}}). \quad (2.13)$$

The rescaled vector field for X^s is

$$\tilde{X}^s := (\Phi_\Sigma)^{-1}X^s \circ \Phi_\Sigma^{-1}, \quad \tilde{X}^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}}) = (\Phi_\Sigma)^{-1}X^s(\varphi^{\text{st}}, \Sigma^{-1}\varphi^{\text{wk}}, v^{\text{st}}, \Sigma I^{\text{wk}}),$$

while X_L^{st} is unchanged under the rescaling.

Theorem 2.2. *Fix \mathcal{B}^{st} and $\kappa > 1$. Assume that $q > 2$. Then there exists a constant $M = M(\mathcal{B}^{\text{st}}, Q, \kappa, q, d - m) > 1$, such that for $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}})$ and $H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})$, $H^{\text{st}} = \mathcal{H}^{\text{st}}(p, U^{\text{st}})$, such that the following hold.*

For the rescaling parameter $\sigma_j = |k_j^{\text{wk}}|^{-\frac{q+1}{3}}$, uniformly on $\mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^m \times \mathbb{R}^{d-m}$ we have

$$\begin{aligned} \|\Pi_{(\varphi^{\text{st}}, v^{\text{st}})}(\tilde{X}^s - X_L^{\text{st}})\|_{C^0} &\leq M\mu(\mathcal{B}^{\text{wk}})^{-(q-1)}, \\ \|D\tilde{X}^s - DX_L^{\text{st}}\|_{C^0} &\leq M\mu(\mathcal{B}^{\text{wk}})^{-\frac{q-2}{3}}. \end{aligned}$$

In particular, as $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$, the vector field \tilde{X}^s converges to X_L^{st} in C^1 over compact sets. Theorem 2.2 is proven in section 4.1.

2.4 The variational aspect of dominant Hamiltonians

We will develop a similar perturbation theory for the weak KAM solutions of the dominant Hamiltonian. The weak KAM solution is closely related to some important invariant sets of the Hamiltonian system, known as the Mather, Aubry and Mañe sets.

- **Preliminaries in weak KAM solutions** In this section we give only enough concepts to formulate our theorem. A more detailed exposition will be given in Section 5.1. Let

$$H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$$

be a C^3 Hamiltonian satisfying the condition $D^{-1}\text{Id} \leq \partial_{II}^2 H(\varphi, I) \leq D\text{Id}$. The associated Lagrangian $L : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is given by

$$L_H(\varphi, v) = \sup_{I \in \mathbb{R}^n} \{I \cdot v - H(\varphi, I)\}.$$

Let $c \in \mathbb{R}^d \simeq H^1(\mathbb{T}^d, \mathbb{R})$, we define Mather's alpha function to be

$$\alpha_H(c) = - \inf_{\nu} \left\{ \int (L_H - c \cdot v) d\nu \right\},$$

where the infimum is taken over all Borel probability measures on $\mathbb{T}^d \times \mathbb{R}^d$ that is invariant under the Euler-Lagrange flow of L_H .

A continuous function $u : \mathbb{T}^d \rightarrow \mathbb{R}$ is called a (negative) weak KAM solution to $L_H - c \cdot v$ if for any $t > 0$, we have

$$u(x) = \inf_{y \in \mathbb{T}^d, \gamma(0)=y, \gamma(t)=x} \left(u(y) + \int_0^t (L_H(\gamma(t), \dot{\gamma}(t)) - c \cdot \dot{\gamma}(t) + \alpha_H(c)) dt \right),$$

where $\gamma : [0, t] \rightarrow \mathbb{T}^d$ is absolutely continuous. Weak KAM solutions exist and are Lipschitz (see [Fat08], [Ber10]).

- **The relation between Lagrangians** We now turn to the weak KAM solutions of dominant Hamiltonians. Fix \mathcal{B}^{st} and consider

$$(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega^{m,d}(\mathcal{B}^{\text{st}})$$

and write $H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})$, $H^{\text{st}} = \mathcal{H}^{\text{st}}(p, \mathcal{B}^{\text{st}}, U^{\text{st}})$. Note

$$H^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, I^{\text{st}}, I^{\text{wk}}) = K(I^{\text{st}}, I^{\text{wk}}) - U^{\text{st}}(\varphi^{\text{st}}) - U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}),$$

where $U^{\text{wk}} = \sum_{j=1}^{d-m} U_j^{\text{wk}}$, and

$$H^{\text{st}}(\varphi^{\text{st}}, I^{\text{st}}) = K(I^{\text{st}}, 0) - U^{\text{st}}(\varphi^{\text{st}}).$$

Denote $L^s = L_{H^s}$ and $L^{\text{st}} = L_{H^{\text{st}}}$, we have

$$\begin{aligned} L^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, v^{\text{wk}}) &= L_0^s(v^{\text{st}}, v^{\text{wk}}) + U^{\text{st}}(\varphi^{\text{st}}) + U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \\ L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) &= L_0^{\text{st}}(v^{\text{st}}) + U^{\text{st}}(\varphi^{\text{st}}), \end{aligned}$$

where L^s, L^{st} are quadratic functions with $(\partial_{vv}^2 L^s) = (\partial_{II}^2 K)^{-1}$ and $(\partial_{v^{\text{st}}v^{\text{st}}}^2 L^{\text{st}}) = (\partial_{I^{\text{st}}I^{\text{st}}}^2 K)^{-1}$ for $v = \partial_I K$ and $v^{\text{st}} = \partial_{I^{\text{st}}} K$.

Given $c = (c^{\text{st}}, c^{\text{wk}}) \in \mathbb{R}^m \times \mathbb{R}^{d-m} = \mathbb{R}^d$, we show that the weak KAM solution of $L^s - c \cdot v$ is related to the weak KAM solution of $L^{\text{st}} - \bar{c} \cdot v^{\text{st}}$, where \bar{c} is computed using an explicit formula. More precisely, we define

$$\bar{c} = c^{\text{st}} + A^{-1} B c^{\text{wk}},$$

where $\partial_{II}^2 K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ as in (2.8). Then (we refer to section 4.2 for details)

$$\begin{aligned} L^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, v^{\text{wk}}) - (c^{\text{st}}, c^{\text{wk}}) \cdot (v^{\text{st}}, v^{\text{wk}}) &= L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) - \bar{c} \cdot v^{\text{st}} \\ &+ \frac{1}{2}(v^{\text{wk}} - B^T A^{-1} v^{\text{st}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1}(v^{\text{wk}} - B^T A^{-1} v^{\text{st}} - \tilde{C} c^{\text{wk}}) \\ &+ \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}} + U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \end{aligned}$$

where $\tilde{C} = C - B^T A^{-1} B$. The above computation suggests a connection between the Lagrangian $L^s - c \cdot v$ and $L^{\text{st}} - \bar{c} \cdot v^{\text{st}}$. Indeed, in Proposition 5.5 we show

$$\alpha_{H^s}(c) - \|U^{\text{wk}}\|_{C^0} \leq \alpha_{H^{\text{st}}}(\bar{c}) + \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}} \leq \alpha_{H^s}(c) + \|U^{\text{wk}}\|_{C^0}.$$

- **Semi-continuity of weak KAM solutions**

We now state our main variational results. We consider a sequence of dominant Hamiltonians with $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$, and their weak KAM solutions. Then there is always a converging subsequence, and the limit point is the weak KAM solution of the strong Hamiltonian. This is sometimes referred to as upper semi-continuity.

Theorem 2.3. *Fix \mathcal{B}^{st} and $\kappa > 1$. Assume that $q > 2(d - m)$.*

For $p_0 \in \mathbb{R}^n$, $U_0^{\text{st}} \in C^2(\mathbb{T}^m)$ and $\bar{c} \in \mathbb{R}^m$, we consider a sequence

$$(\mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}}), \quad c_i = (c_i^{\text{st}}, c_i^{\text{wk}}) \in \mathbb{R}^m \times \mathbb{R}^{d-m},$$

and let u_i be a weak KAM solution of

$$L_{\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}})} - c_i \cdot v.$$

Denote $K_i = K_{p_i, \mathcal{B}^{\text{st}}, \mathcal{B}_i^{\text{wk}}}$, and

$$A_i = \partial_{I^{\text{st}}I^{\text{st}}}^2 K_i, \quad B_i = \partial_{I^{\text{st}}I^{\text{wk}}}^2 K_i, \quad C_i = \partial_{I^{\text{wk}}I^{\text{wk}}}^2 K_i.$$

Assume:

- $\mu(B_i^{\text{wk}}) \rightarrow \infty$, $p_i \rightarrow p_0$, $U_i^{\text{st}} \rightarrow U_0^{\text{st}}$.
- $c_i^{\text{st}} + A_i^{-1} B_i c_i^{\text{wk}} \rightarrow \bar{c}$.

Then:

1. The sequence $\{u_i\}$ is equi-continuous. In particular, the sequence $\{u_i(\cdot) - u_i(0)\}$ is pre-compact in the C^0 topology.
2. Let u be any accumulation point of the sequence $u_i(\cdot) - u_i(0)$. Then there exists $u^{\text{st}} : \mathbb{T}^m \rightarrow \mathbb{R}$ such that $u(\varphi^{\text{st}}, \varphi^{\text{wk}}) = u^{\text{st}}(\varphi^{\text{st}})$, i.e, u is independent of φ^{wk} .
3. u^{st} is a weak KAM solution of

$$L_{\mathcal{H}^{\text{st}}(p_0, U_0^{\text{st}})} - \bar{c} \cdot v^{\text{st}}.$$

The proof of Theorem 2.3 occupies sections 4 and 5, with some technical statements deferred to section 7.

Using the point of view in [Ber10], the semi-continuity of the weak KAM solution is closely related to the semi-continuity of the Aubry and Mañe sets. These properties have important applications to Arnold diffusion. In section 6 we develop an analog of these results for the dominant Hamiltonians.

3 The choice of basis and averaging

In this section we prove Theorem 2.1. The proof consists of two parts: the choice of the basis and estimates on the norms.

3.1 The choice of the basis

Recall that we have a fixed irreducible lattice $\Lambda^{\text{st}} \subset \mathbb{Z}^{n+1}$ of rank $m < n$, and a fixed basis $\mathcal{B}^{\text{st}} = \{k_1, \dots, k_m\}$ for Λ^{st} . The following proposition describes the choice of the adapted basis.

Proposition 3.1. *Let $\Lambda^{\text{st}} \subset \mathbb{Z}^{n+1}$ be an irreducible lattice of rank $m < n$, and fix a basis k_1, \dots, k_m . Let $\Lambda \supset \Lambda^{\text{st}}$ be an irreducible lattice of rank d , then there exists $k_{m+1}, \dots, k_d \in \mathbb{Z}^{n+1}$ such that k_1, \dots, k_d form a basis of Λ , and the following hold.*

1. For each $m < j \leq d$,

$$|k_j| \leq \bar{M} + (d - m)M_j,$$

where

$$\bar{M} = |k_1| + \dots + |k_m|, \quad \Lambda_j = \text{span}_{\mathbb{Z}}\{k_1, \dots, k_j\}, \quad M_j = M(\Lambda_j | \Lambda_{j-1}).$$

2. For each $m < i < j \leq d$,

$$|k_i| \leq \bar{M} + (d - m)|k_j|.$$

We now describe the choice of the vectors k_{m+1}, \dots, k_d . We define $k'_i = k_i$ for $1 \leq i \leq m$, and define k'_i with $i > m$ inductively using the following procedure. Suppose k'_1, \dots, k'_i are defined, let

$$\Lambda_i = \text{span}_{\mathbb{R}}\{k'_1, \dots, k'_i\} \cap \Lambda, \quad M_{i+1} = \min\{|k| : k \in \Lambda \setminus \Lambda_i\}.$$

We define k'_{i+1} to be a vector reaching the minimum in the definition of M_{i+1} , i.e. $|k'_{i+1}| = M_{i+1}$. We have

$$|k'_i| = M_i, \quad m < i \leq d, \quad |k'_j| \leq |k'_i|, \quad m < j < i \leq d,$$

but k'_1, \dots, k'_d may not form a basis. We turn them into a basis using the following procedure (see [Sie89]).

For each $j = 1, \dots, m$, define

$$c_j = \min\{s_j : s_{j,1}k'_1 + \dots + s_{j,j-1}k'_{j-1} + s_j k'_j \in \Lambda, s_j \in \mathbb{R}^+, s_{j,i} \in \mathbb{R}^+ \cup \{0\}\}. \quad (3.1)$$

We define $c_{j,j-1}$ using a similar minimization given the value c_j :

$$c_{j,j-1} = \min\{s_{j,j-1} : s_{j,1}k'_1 + \dots + s_{j,j-1}k'_{j-1} + c_j k'_j, s_{j,i} \in \mathbb{R}^+ \cup \{0\}\}.$$

We now define $c_{j,i}$ for $1 \leq i \leq j - 2$ inductively as follows. Assume that $c_{j,i}, \dots, c_{j,j-1}$ are all defined, then

$$c_{j,i-1} = \min\{s_{j,i-1} : s_{j,1}k'_1 + \dots + s_{j,i-1}k'_{i-1} + c_{j,i}k'_i + \dots + c_{j,j-1}k'_{j-1} + c_j k'_j \in \Lambda, s_{j,1}, \dots, s_{j,i-1} \in \mathbb{R}^+ \cup \{0\}\}.$$

Finally,

$$k_j = c_{j,1}k'_1 + \dots + c_{j,j-1}k'_{j-1} + c_j k'_j.$$

We have the following lemma from the geometry of numbers.

Lemma 3.2 (see [Sie89]). *Let $\Lambda \subset \mathbb{Z}^{n+1}$ be a lattice of rank $d \leq n$ and let k'_1, \dots, k'_d be any linearly independent set in Λ . Let*

$$k_j = c_{j,1}k'_1 + \dots + c_{j,j-1}k'_{j-1} + c_j k'_j.$$

be defined using the procedure above. Then

1. For each $1 \leq j \leq d$, k_1, \dots, k_j form a basis of $\text{span}_{\mathbb{R}}\{k'_1, \dots, k'_j\} \cap \Lambda$ over \mathbb{Z} . In particular, k_1, \dots, k_d form a basis of Λ .

2. For $1 \leq j < d$ and $1 \leq i \leq j - 1$, we have

$$0 \leq c_{j,i} < 1, \quad 0 < c_j \leq 1.$$

3. If for some $1 \leq m \leq d$, k'_1, \dots, k'_m already form a basis of $\text{span}_{\mathbb{R}}\{k'_1, \dots, k'_m\} \cap \Lambda$ over \mathbb{Z} , then $k_1 = k'_1, \dots, k_m = k'_m$.

Proof. To prove item 1, we proceed by induction. When $i = 1$, we have $k_1 = c_1 k'_1$, where $c_i = \min\{t > 0, tk'_1 \in \Lambda\}$. The choice of k_1 implies any $k \in k'_1 \mathbb{R} \cap \Lambda$ is an integer multiple of k_1 .

Suppose the statement holds for the index j , we prove it for the index $j + 1$. For any $k \in \text{span}_{\mathbb{R}}\{k'_1, \dots, k'_{j+1}\} \cap \Lambda$ with

$$k = a_1 k'_1 + \dots + a_{j+1} k'_{j+1}, \quad a_1, \dots, a_{j+1} \in \mathbb{R},$$

we claim that a_{j+1} is an integer multiple of c_{j+1} , where c_{j+1} reaches the minimum in (3.1) for the index $j + 1$. Assume otherwise, there exist an integer s_{j+1} such that $0 < s_{j+1} + a_{j+1} < c_{j+1}$. Choose $s_1, \dots, s_j \in \mathbb{Z}$, such that $s_i + a_i \geq 0$ for $1 \leq i \leq j$, we have

$$k + \sum_{i=1}^{j+1} s_i k'_i = \sum_{i=1}^{j+1} (s_i + a_i) k'_i \in \Lambda,$$

and $s_i + a_i \geq 0$ for $1 \leq i \leq j$, $0 < s_{j+1} + a_{j+1} < c_{j+1}$, contradicting the definition of c_{j+1} .

Using the claim, there exists $l \in \mathbb{Z}$ such that $a_{j+1} = lc_{j+1}$, so

$$k - lk'_{j+1} \in \text{span}_{\mathbb{R}}\{k'_1, \dots, k'_j\} \cap \Lambda$$

By inductive hypothesis, $k - lk'_{j+1} \in \text{span}_{\mathbb{Z}}\{k_1, \dots, k_j\}$, implying $k \in \text{span}_{\mathbb{Z}}\{k'_1, \dots, k'_{j+1}\}$.

For item 2, note that for any $k_j = c_{j,1} k'_1 + \dots + c_{j,j-1} k'_{j-1} + c_j k'_j \in \Lambda$, we can always subtract an integer from any $c_{j,i}$ or c_j and remain in Λ . If the estimates do not hold, we can get a contradiction by reducing $c_{j,i}$ or c_j .

For item 3, if k'_1, \dots, k'_m is a basis (over \mathbb{Z}) of $\text{span}_{\mathbb{R}}\{k'_1, \dots, k'_m\} \cap \Lambda$, then all coefficients of $k_j = c_{j,1} k'_1 + \dots + c_{j,j-1} k'_{j-1} + c_j k'_j \in \Lambda$ for $j \leq m$ must be integers. Then the constraints of item 2 implies $c_{j,i} = 0$ and $c_j = 1$, namely $k_j = k'_j$. \square

Proof of Proposition 3.1. We choose the basis k_1, \dots, k_d as described. Lemma 3.2 implies $k_j = k'_j$ for $1 \leq j \leq m$. Using

$$0 < c_{j+1} \leq 1, \quad 0 \leq c_{j+1,i} < 1,$$

we get

$$|k_j| \leq |k'_1| + \dots + |k'_j| = |k_1| + \dots + |k_m| + M_{m+1} + \dots + M_j.$$

Since $M_{m+1} \leq \dots \leq M_d$, and $\bar{M} = |k_1| + \dots + |k_m|$, we get

$$|k_j| \leq \bar{M} + (j - m)M_j \leq \bar{M} + (d - m)M_j.$$

Moreover, for $i < j$, we have

$$|k_i| \leq \bar{M} + (d - m)M_i < \bar{M} + (d - m)M_j \leq \bar{M} + (d - m)|k_j|.$$

□

We note that the basis, as chosen in Proposition 3.1, satisfies item 1 of Theorem 2.1 for $\kappa \geq \max\{\bar{M}, d - m\}$.

3.2 Estimating the weak potential

In this section we prove the second item in Theorem 2.1 and conclude its proof. Assume that $H_1 \in C^r(\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T})$ with $r > n + 2d - 2m + 4$. Let the basis k_1, \dots, k_n be chosen as in Proposition 3.1. We show that there exists $\kappa = \kappa(\mathcal{B}^{\text{st}}, Q, n) > 1$ such that for $m < i \leq d$,

$$\|U_{p_0, \mathcal{B}_{i-1}, \mathcal{B}_i}^{\text{wk}}\|_{C^2} \leq \|Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}}\|_{C^2} \leq \kappa |k_i|^{-r+3n-2m+6}.$$

By (2.1) and (2.4),

$$(Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}})(k_1 \cdot (\theta, t), \dots, k_i \cdot (\theta, t), p) = ([H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}})(\theta, p, t),$$

and the norm of $[H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}}$ can be estimated using a standard estimates of the Fourier series.

Lemma 3.3 (c.f. [BKZ11], Lemma 2.1, item 3). *Let $H_1(\theta, p, t) = \sum_{k \in \mathbb{Z}^{n+1}} h_k(p) e^{2\pi i k \cdot (\theta, t)}$ satisfy $\|H_1\|_{C^r} = 1$, with $r \geq n + 3$. There exists a constant C_n depending only on n , such that for any subset $\tilde{\Lambda} \subset \mathbb{Z}^{n+1}$ with $\min_{k \in \tilde{\Lambda}} |k| = M > 0$, we have*

$$\left\| \sum_{k \in \tilde{\Lambda}} h_k(p) e^{2\pi i k \cdot (\theta, t)} \right\|_{C^2} \leq C_n M^{-r+n+4}.$$

Since $\min_{k \in \Lambda_i \setminus \Lambda_{i-1}} |k| = M_i$, we apply Lemma 3.3 to $\Lambda_i \setminus \Lambda_{i-1}$ to get

$$\|[H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}}\|_{C^2} \leq C_n M_i^{-r+n+3}. \quad (3.2)$$

To estimate $Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}}$, we apply a linear coordinate change. Given k_1, \dots, k_i , we choose $\hat{k}_{i+1}, \dots, \hat{k}_{n+1} \in \mathbb{Z}^{n+1}$ to be coordinate vectors (unit integer vectors) such that

$$P_i := \begin{bmatrix} k_1 & \cdots & k_i & \hat{k}_{i+1} & \cdots & \hat{k}_{n+1} \end{bmatrix}$$

is invertible. We extend $(Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}})(\varphi_1, \dots, \varphi_i)$ trivially to a function of $(\varphi_1, \dots, \varphi_{n+1})$, then

$$(Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}}) \left(P_i^T \begin{bmatrix} \theta \\ t \end{bmatrix} \right) = ([H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}})(\theta, t).$$

We get

$$\|Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}}\|_{C^2} \leq (1 + \|P_i^{-1}\|)(1 + \|(P_i^T)^{-1}\|)\|[H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}}\|_{C^2}.$$

Using Lemma B.1, there exists a constant $c_n > 0$ depending only on n such that

$$\|P_i^{-1}\| = \|(P_i^T)^{-1}\| \leq c_n |k_1| \cdots |k_i| |\hat{k}_{i+1}| \cdots |\hat{k}_{n+1}|.$$

We have $|k_1|, \dots, |k_m| \leq \bar{M}$, $|\hat{k}_{i+1}| = \dots = |\hat{k}_{n+1}| = 1$, and from Lemma 3.2, $|k_{m+1}|, \dots, |k_i| \leq \bar{M} + (d - m)M_i$. Hence there exists a constant $c_{n, \bar{M}} > 0$ such that

$$\|P_i^{-1}\| = \|(P_i^T)^{-1}\| \leq c_{n, \bar{M}} M_i^{i-m}.$$

Combine with (3.2), we get for $\kappa = \kappa(n, \bar{M})$,

$$\|Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}}\|_{C^2} \leq \kappa M_i^{-r+n+4+2(i-m)} \leq \kappa M_i^{-r+n+4+2(d-m)} \leq \kappa |k_i|^{-r+n+2d-2m+4}.$$

This implies item 2 of Theorem 2.1. The proof is complete.

4 Strong and slow systems of dominant Hamiltonians

In this section we study the relation between Hamiltonians and the corresponding Lagrangians for dominant systems. We start by comparing the Hamiltonian vector fields and then compare their Lagrangians.

4.1 Vector fields of dominant Hamiltonians

In this section we expand on section 2.3 and prove Theorem 2.2. Fix $\mathcal{B}^{\text{st}}, \kappa > 1$ and let $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}})$, we recall the notations

$$H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}), \quad H^{\text{st}} = \mathcal{H}^{\text{st}}(p, U^{\text{st}}).$$

Then

$$H^s(\varphi, I) = K(I) - U^{\text{st}}(\varphi^{\text{st}}) - U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \quad H^{\text{st}}(\varphi^{\text{st}}, I^{\text{st}}) = K(I^{\text{st}}, 0) - U^{\text{st}}(\varphi^{\text{st}}).$$

Recall from (2.7) that $\partial_{II}^2 K = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, then

$$(A)_{ij} = (k_i^{\text{st}})^T Q k_j^{\text{st}}, \quad (B)_{ij} = (k_i^{\text{st}})^T Q k_j^{\text{wk}}, \quad (C)_{ij} = (k_i^{\text{wk}})^T Q k_j^{\text{wk}}.$$

The vector field $X^s(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}})$ defined on the universal cover $\mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^m \times \mathbb{R}^{d-m}$ is obtained from the Hamiltonian vector field of H^s , with the coordinate change $v^{\text{st}} = AI^{\text{st}} + BI^{\text{wk}}$ (see (2.10)). The vector field $X_L^{\text{st}}(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, I^{\text{wk}})$ is defined

as a trivial extension of the Lagrangian vector field of H^{st} , also defined on the universal cover. More explicitly (see (2.11), (2.12))

$$X^s = \begin{bmatrix} v^{\text{st}} \\ B^T A^{-1} v^{\text{st}} - \tilde{C} I^{\text{wk}} \\ A \partial_{\varphi^{\text{st}}} U + B \partial_{\varphi^{\text{wk}}} U \\ \partial_{\varphi^{\text{wk}}} U \end{bmatrix}, \quad X_L^{\text{st}} = \begin{bmatrix} v^{\text{st}} \\ 0 \\ A \partial_{\varphi^{\text{st}}} U \\ 0 \end{bmatrix}. \quad (4.1)$$

Given $1 \geq \sigma_1 \geq \dots \geq \sigma_{d-m} > 0$, let $\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{d-m}\}$. The rescaling is $\Phi_\Sigma : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$, given by (2.13). We denote by $\tilde{X}^s(\varphi^{\text{st}}, v^{\text{st}}, \tilde{\varphi}^{\text{wk}}, \tilde{I}^{\text{wk}})$ the rescaled X^s . Using (4.1), we have

$$\tilde{X}^s - X_L^{\text{st}} = (\Phi_\Sigma)^{-1} X^s \circ \Phi_\Sigma^{-1} - X_L^{\text{st}} = \begin{bmatrix} 0 \\ \Sigma B^T A^{-1} v^{\text{st}} - \Sigma \tilde{C} \Sigma \tilde{I}^{\text{wk}} \\ A \partial_{\varphi^{\text{st}}} U^{\text{wk}} + B \partial_{\varphi^{\text{wk}}} U^{\text{wk}} \\ \Sigma^{-1} \partial_{\varphi^{\text{wk}}} U^{\text{wk}} \end{bmatrix} \quad (4.2)$$

noting that U^{st} is independent of φ^{wk} , so $\partial_{\varphi^{\text{wk}}} U = \partial_{\varphi^{\text{wk}}} U^{\text{wk}}$. Furthermore

$$D(\tilde{X}^s - X_L^{\text{st}}) = (\Phi_\Sigma)^{-1} D X^s \circ (\Phi_\Sigma)^{-1} - X_L^{\text{st}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \Sigma B^T A^{-1} & \Sigma \tilde{C} \Sigma \\ 0 & A \partial_{\varphi^{\text{st}} \varphi^{\text{wk}}}^2 U^{\text{wk}} \Sigma^{-1} + B \partial_{\varphi^{\text{wk}} \varphi^{\text{wk}}}^2 U^{\text{wk}} \Sigma^{-1} & 0 & 0 \\ \Sigma^{-1} \partial_{\varphi^{\text{wk}} \varphi^{\text{st}}}^2 U^{\text{wk}} & \Sigma^{-1} \partial_{\varphi^{\text{wk}} \varphi^{\text{wk}}}^2 U^{\text{wk}} \Sigma^{-1} & 0 & 0 \end{bmatrix}. \quad (4.3)$$

The quantities in (4.2) and (4.3) are estimated as follows.

Lemma 4.1. *Fix $\mathcal{B}^{\text{st}}, \kappa > 1$. Assume $q > 2$. Then there exists a constant $M_1 = M_1(\mathcal{B}^{\text{st}}, Q, \kappa, q, d - m)$ such that for the parameters $\sigma_j = |k_j^{\text{wk}}|^{-\frac{q+1}{3}}$, uniformly over $\mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^m \times \mathbb{R}^{d-m}$, the following hold.*

1. *For any $1 \leq i \leq m$ and $1 \leq j \leq d - m$, $\|\partial_{\varphi_j^{\text{wk}}} U^{\text{wk}}\|_{C^0}, \|\partial_{\varphi_i^{\text{st}} \varphi_j^{\text{wk}}}^2 U^{\text{wk}}\|_{C^0} \leq M_1 |k_j^{\text{wk}}|^{-q}$;
for any $1 \leq i, j \leq d - m$, $\|\partial_{\varphi_i^{\text{wk}} \varphi_j^{\text{wk}}}^2 U^{\text{wk}}\|_{C^0} \leq M_1 \sup\{|k_i^{\text{wk}}|^{-q}, |k_j^{\text{wk}}|^{-q}\}$.*
2. $\|A \partial_{\varphi^{\text{wk}}} U^{\text{wk}}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-q}\}$.
3. $\|B \partial_{\varphi^{\text{wk}}} U^{\text{wk}}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-(q-1)}\}$.
4. $\|A \partial_{\varphi^{\text{st}} \varphi^{\text{wk}}}^2 U \Sigma^{-1}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-\frac{2q-1}{3}}\}$.
5. $\|B \partial_{\varphi^{\text{st}} \varphi^{\text{wk}}}^2 U\|_{C^0} \leq \|B \partial_{\varphi^{\text{st}} \varphi^{\text{wk}}}^2 U \Sigma^{-1}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-\frac{2q-4}{3}}\}$.
6. $\|\Sigma^{-1} \partial_{\varphi^{\text{wk}} \varphi^{\text{wk}}}^2 U \Sigma^{-1}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-\frac{q-2}{3}}\}$.

$$7. \|\Sigma B^T A^{-1}\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-\frac{q-2}{3}}\}.$$

$$8. \|\Sigma \tilde{C} \Sigma\|_{C^0} \leq M_1 \sup_j \{|k_j^{\text{wk}}|^{-\frac{2q-4}{3}}\}.$$

We first prove Theorem 2.2 using our lemma.

Proof of Theorem 2.2. Noting that $\Pi_{(\varphi^{\text{st}}, v^{\text{st}})}(\tilde{X}^s - X_L^{\text{st}})$ is the first and third line of (4.2), using item 2 and 3 of Lemma 4.1 we get

$$\|\Pi_{(\varphi^{\text{st}}, v^{\text{st}})}(\tilde{X}^s - X_L^{\text{st}})\| \leq M^* \sup_j \{|k_j^{\text{wk}}|^{-(q-1)}\} = M^* \mu(\mathcal{B}^{\text{wk}})^{-(q-1)},$$

for any constant $M^* \geq 2M_1^*$, where M_1^* is from Lemma 4.1.

Since $D(\tilde{X}^s - X_L^{\text{st}})$ is bounded, up to a universal constant, the sum of the norms of all the non-zero blocks in (4.3), using Lemma 4.1 items 4-8, we get

$$\|D\tilde{X}^s - DX_L^{\text{st}}\| \leq M^* \sup_j \{|k_j^{\text{wk}}|^{-\frac{q-2}{3}}\} = M^* \mu(\mathcal{B}^{\text{wk}})^{-\frac{q-2}{3}},$$

where M^* depends only on M_1^* . □

The rest of the section is dedicated to proving Lemma 4.1.

Proof of Lemma 4.1. Denote $\bar{M} = |k_1^{\text{st}}| + \dots + |k_m^{\text{st}}|$, which depends only on \mathcal{B}^{st} .

Item 1. We have

$$\begin{aligned} \|\partial_{\varphi_j^{\text{wk}}} U^{\text{wk}}\|_{C^0} &\leq \sum_{l=1}^{d-m} \|\partial_{\varphi_j^{\text{wk}}} U_l^{\text{wk}}\|_{C^0} \leq \sum_{l \geq j} \|\partial_{\varphi_j^{\text{wk}}} U_l^{\text{wk}}\|_{C^0} \\ &\leq \kappa \sum_{l \geq j} |k_l^{\text{wk}}|^{-q} \leq (d-m) \kappa^{q+1} |k_j^{\text{wk}}|^{-q}, \end{aligned}$$

where the second inequality is due to U_l^{wk} depending only on $(\varphi_1^{\text{wk}}, \dots, \varphi_l^{\text{wk}})$, and the last two inequalities uses the definition of $\Omega_{\kappa, q}^{m, d}$, see section 2.2. By the same reasoning, we have

$$\begin{aligned} \|\partial_{\varphi_i^{\text{st}} \varphi_j^{\text{wk}}}^2 U^{\text{wk}}\| &\leq \sum_{l \geq j} \|U_l^{\text{wk}}\|_{C^2} \leq (d-m) \kappa^{q+1} |k_j^{\text{wk}}|^{-q}, \\ \|\partial_{\varphi_i^{\text{wk}} \varphi_j^{\text{wk}}}^2 U^{\text{wk}}\| &\leq \sum_{l \geq \sup\{i, j\}} \|U_l^{\text{wk}}\|_{C^2} \leq (d-m) \kappa^{q+1} \sup\{|k_i^{\text{wk}}|^{-q}, |k_j^{\text{wk}}|^{-q}\} \end{aligned}$$

the second and third estimate follows.

Item 2. As a matrix,

$$\begin{aligned} |(A \partial_{\varphi^{\text{st}}} U^{\text{wk}})_{ij}| &= |(k_i^{\text{st}})^T Q k_j^{\text{st}} \partial_{\varphi_j^{\text{st}}} U^{\text{wk}}| \leq \bar{M}^2 \|Q\| \|\partial_{\varphi_j^{\text{st}}} U\| \\ &\leq (d-m) \bar{M}^2 \|Q\| \kappa^{q+1} |k_j^{\text{wk}}|^{-q}, \end{aligned}$$

where the last line is due to item 1. Since the matrix norm is bounded by the supremum of all matrix entries, up to a constant depending only on dimension, item 2 follows. In the sequel, we apply the same reasoning and only estimate the supremum of matrix entries.

Item 3. Similar to item 2,

$$\begin{aligned} |(B\partial_{\varphi^{\text{wk}}}U)_{ij}| &= |(k_i^{\text{st}})^T Q k_j^{\text{wk}} \partial_{\varphi_j^{\text{wk}}} U^{\text{wk}}| \leq \bar{M} \|Q\| |k_j^{\text{wk}}| \|\partial_{\varphi_j^{\text{wk}}} U^{\text{wk}}\| \\ &\leq \bar{M} \|Q\| (d-m) \kappa^{q+1} |k_j^{\text{wk}}|^{-(q-1)}, \end{aligned}$$

Item 3 follows.

Item 4. We have

$$\begin{aligned} |(A\partial_{\varphi^{\text{st}}\varphi^{\text{wk}}}^2 U^{\text{wk}} \Sigma^{-1})_{ij}| &= \left| \sum_l (k_i^{\text{st}})^T Q k_l^{\text{st}} \partial_{\varphi_l^{\text{st}}\varphi_j^{\text{wk}}}^2 U \sigma_j^{-1} \right| \leq \bar{M}^2 \|Q\| \sum_l \sigma_j^{-1} |\partial_{\varphi_l^{\text{st}}\varphi_j^{\text{wk}}}^2 U| \\ &\leq (d-m)^2 \bar{M} \|Q\| \kappa^{q+1} \sigma_j^{-1} |k_j^{\text{wk}}|^{-q} \leq (d-m)^2 \bar{M} \|Q\| \kappa^{q+1} |k_j^{\text{wk}}|^{-\frac{2q-1}{3}}, \end{aligned}$$

noting the first inequality of the second line is due to item 1 and the second is due to the choice of σ_j .

Item 5.

$$\begin{aligned} |(B\partial_{\varphi^{\text{wk}}\varphi^{\text{wk}}}^2 U \Sigma^{-1})_{ij}| &= \left| \sum_{l \geq j} (k_i^{\text{st}})^T Q k_l^{\text{wk}} \partial_{\varphi_l^{\text{wk}}\varphi_j^{\text{wk}}}^2 U \sigma_j^{-1} \right| \leq \bar{M} \|Q\| \sum_{l \geq j} |k_l^{\text{wk}}| \sigma_j^{-1} |\partial_{\varphi_l^{\text{wk}}\varphi_j^{\text{wk}}}^2 U| \\ &\leq \bar{M} \|Q\| (d-m)^2 \kappa^{q+2} |k_j^{\text{wk}}| |k_j^{\text{wk}}|^{-q} |k_j^{\text{wk}}|^{\frac{q+1}{3}} = \bar{M} \|Q\| (d-m) \kappa^{q+2} |k_j^{\text{wk}}|^{-\frac{2q-4}{3}}, \end{aligned}$$

where the inequality of the second line uses $|k_l^{\text{wk}}| \leq \kappa |k_j^{\text{wk}}|$, item 1 and the choice of σ_j .

Item 6. Using item 1 and choice of σ_j , we have

$$\begin{aligned} |(\Sigma^{-1} \partial_{\varphi^{\text{wk}}\varphi^{\text{wk}}}^2 U^{\text{wk}} \Sigma^{-1})_{ij}| &= |\sigma_i^{-1} \partial_{\varphi_i^{\text{wk}}\varphi_j^{\text{wk}}}^2 U^{\text{wk}} \sigma_j^{-1}| \\ &\leq (d-m) \kappa^{q+1} \sigma_i^{-1} \sigma_j^{-1} \sup\{|k_i^{\text{wk}}|^{-q}, |k_j^{\text{wk}}|^{-q}\} \\ &\leq (d-m) \kappa^{q+1} \sup\{|k_i^{\text{wk}}|^{-\frac{q-2}{3}}, |k_j^{\text{wk}}|^{-\frac{q-2}{3}}\}. \end{aligned}$$

Item 7. We have

$$|(\Sigma B^T)_{ij}| = |\sigma_i (k_i^{\text{wk}})^T Q k_j^{\text{st}}| \leq \bar{M} \|Q\| \sup_j \{|k_j^{\text{wk}}| \sigma_j\} = \bar{M} \|Q\| \sup_j \{|k_j^{\text{wk}}|^{-\frac{q-2}{3}}\}$$

and uses $\|\Sigma B^T A^{-1}\| \leq \|\Sigma B^T\| \|A^{-1}\|$, noting that $\|A^{-1}\|$ depends only on Q and \mathcal{B}^{st} .

Item 8. Recall $\tilde{C} = C - B^T A^{-1} B$. We have

$$|(\Sigma C \Sigma)_{ij}| = |\sigma_i (k_i^{\text{wk}})^T Q k_j^{\text{wk}} \sigma_j| \leq (\sup_j \sigma_j |k_j^{\text{wk}}|)^2 \|Q\| \leq \|Q\| \sup_j \{|k_j^{\text{wk}}|^{-\frac{2q-4}{3}}\}.$$

Suppose S_1, S_2 are positive definite symmetric matrices with $S_1 \geq S_2$, for any $v \in \mathbb{R}^{d-m}$,

$$v^T S_1 v = v^T (S_1 - S_2 + S_2) v \geq v^T S_2 v,$$

we obtain $\|S_1\| \geq \|S_2\|$. Since $C - B^T A^{-1} B \geq 0$, we have $\Sigma C \Sigma - \Sigma B^T A^{-1} B \Sigma \geq 0$. Apply the observation to the matrices $\Sigma C \Sigma$ and $\Sigma B^T A^{-1} B \Sigma$ we get

$$\|\Sigma \tilde{C} \Sigma\| \leq \|\Sigma C \Sigma\| + \|\Sigma B^T A^{-1} B \Sigma\| \leq 2\|\Sigma C \Sigma\|.$$

Item 8 follows. \square

4.2 The slow Lagrangian

We derive the special form of the slow Lagrangian described in section 2.4. We fix $\mathcal{B}^{\text{st}}, \kappa > 1$ and $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}})$. Denote $H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})$, $H^{\text{st}} = \mathcal{H}^{\text{st}}(p, U^{\text{st}})$ and the associated Lagrangian is denoted L^s and L^{st} .

As before we write

$$H^s(\varphi, I) = K(I) - U^{\text{st}}(\varphi^{\text{st}}) - U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \quad H^{\text{st}}(\varphi^{\text{st}}, I^{\text{st}}) = K(I^{\text{st}}, 0) - U^{\text{st}}(\varphi^{\text{st}}),$$

and

$$L^s(\varphi, v) = L_0^s(v) + U^{\text{st}}(\varphi^{\text{st}}) + U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \quad L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) = L_0^{\text{st}}(v^{\text{st}}) + U^{\text{st}}(\varphi^{\text{st}}),$$

where $\partial_{vv}^2 L_0^s = (\partial_{II}^2 K)^{-1}$, $\partial_{v^{\text{st}} v^{\text{st}}}^2 L_0^{\text{st}} = (\partial_{I^{\text{st}} I^{\text{st}}}^2 K)^{-1}$ for $v = \partial_I K$ and $v^{\text{st}} = \partial_{I^{\text{st}}} K$. Recall the notation

$$\partial_{II}^2 K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \quad A = \partial_{I^{\text{st}} I^{\text{st}}}^2 K, \quad B = \partial_{I^{\text{st}} I^{\text{wk}}}^2 K, \quad C = \partial_{I^{\text{wk}} I^{\text{wk}}}^2 K.$$

Lemma 4.2. *With the above notations we have*

1.

$$L^s(v, \varphi) = L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) + \frac{1}{2}(v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) \cdot \tilde{C}^{-1}(v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) + U^{\text{wk}}(\varphi^{\text{st}}, \varphi^{\text{wk}}), \quad (4.4)$$

where

$$\tilde{C} = C - B^T A^{-1} B.$$

2. Let $c = (c^{\text{st}}, c^{\text{wk}}) \in \mathbb{R}^m \times \mathbb{R}^{d-m}$. We denote ²

$$\bar{c} = c^{\text{st}} + A^{-1} B c^{\text{wk}}, \quad w^{\text{wk}} = v^{\text{wk}} - B^T A^{-1} v^{\text{st}}, \quad (4.5)$$

then

$$L^s(v, \varphi) - c \cdot v = L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) - \bar{c} \cdot v^{\text{st}} + \frac{1}{2}(w^{\text{wk}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1}(w^{\text{wk}} - \tilde{C} c^{\text{wk}}) - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}} + U^{\text{wk}}(\varphi^{\text{wk}}, \varphi^{\text{st}}). \quad (4.6)$$

²We stress here that no coordinate change is performed: w^{wk} is simply an abbreviation for $v^{\text{wk}} - B^T A^{-1} v^{\text{st}}$.

Proof. We have the following identity in block matrix inverse, which can be verified by a direct computation.

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ \text{Id} \end{bmatrix} \tilde{C}^{-1} [-B^T A^{-1} \quad \text{Id}].$$

Then

$$\begin{aligned} L_0^s(v^{\text{st}}, v^{\text{wk}}) &= \frac{1}{2} [(v^{\text{st}})^T (v^{\text{wk}})^T] \left(\begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \text{Id} \\ -A^{-1}B \end{bmatrix} \tilde{C}^{-1} [\text{Id} \quad -B^T A^{-1}] \right) \begin{bmatrix} v^{\text{st}} \\ v^{\text{wk}} \end{bmatrix} \\ &= \frac{1}{2} v^{\text{st}} \cdot A^{-1} v^{\text{st}} + \frac{1}{2} (v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) \cdot \tilde{C}^{-1} (v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) \\ &= L_0^{\text{st}}(v^{\text{st}}) + \frac{1}{2} (v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) \cdot \tilde{C}^{-1} (v^{\text{wk}} - B^T A^{-1} v^{\text{st}}), \end{aligned}$$

and (4.4) follows.

Moreover,

$$\begin{aligned} L_0^s - (c^{\text{st}}, c^{\text{wk}}) \cdot (v^{\text{st}}, v^{\text{wk}}) &= L_0^{\text{st}}(v^{\text{st}}) - (c^{\text{st}} + A^{-1} B c^{\text{wk}}) \cdot v^{\text{st}} + \frac{1}{2} w^{\text{wk}} \cdot \tilde{C}^{-1} w^{\text{wk}} - c^{\text{wk}} \cdot v^{\text{wk}} + A^{-1} B c^{\text{wk}} \cdot v^{\text{st}} \\ &= L_0^{\text{st}}(v^{\text{st}}) - \bar{c} \cdot v^{\text{st}} + \frac{1}{2} w^{\text{wk}} \cdot \tilde{C}^{-1} w^{\text{wk}} - c^{\text{wk}} \cdot (v^{\text{wk}} - B^T A^{-1} v^{\text{st}}) \\ &= L_0^{\text{st}}(v^{\text{st}}) - \bar{c} \cdot v^{\text{st}} + \frac{1}{2} w^{\text{wk}} \cdot \tilde{C}^{-1} w^{\text{wk}} - (\tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1} w^{\text{wk}} \\ &= L_0^{\text{st}}(v^{\text{st}}) - \bar{c} \cdot v^{\text{st}} + \frac{1}{2} (w^{\text{wk}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1} (w^{\text{wk}} - \tilde{C} c^{\text{wk}}) - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}}. \end{aligned}$$

We obtain (4.6). \square

The Euler-Lagrange flow of L^s satisfies the following estimates.

Lemma 4.3. *Fix \mathcal{B}^{st} , $\kappa > 1$. Assume that $q > 1$, $L^s = L_{\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})}$, with $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}})$. Let $\gamma = (\gamma^{\text{st}}, \gamma^{\text{wk}}) : [0, T] \rightarrow \mathbb{T}^d$ satisfy the Euler-Lagrange equation of L^s .*

1. *There exists a constant $M_1 = M_1(\mathcal{B}^{\text{st}}, Q, \kappa, q)$ such that*

$$\|\ddot{\gamma}^{\text{st}} - A \partial_{\varphi^{\text{st}}} U^{\text{st}}(\gamma^{\text{st}})\|_{C^0} \leq M_1 (\mu(\mathcal{B}^{\text{wk}}))^{-(q-1)}.$$

2. *There exists a constant $M_2 = M_2(\mathcal{B}^{\text{st}}, Q, \kappa, q, \|U^{\text{st}}\|)$ such that*

$$\|\ddot{\gamma}^{\text{st}}\|_{C^0} \leq M_2.$$

Proof. Observe that the $(\varphi^{\text{st}}, v^{\text{st}})$ component of the Euler-Lagrange vector field of L^s is precisely the vector field $\Pi_{\varphi^{\text{st}}, v^{\text{st}}} \tilde{X}^s$ in Theorem 2.2. The Euler-Lagrange equation of L^{st} (which is X^{st} in Theorem 2.2) is $\ddot{\varphi}^{\text{st}} = A \partial_{\varphi^{\text{st}}} U^{\text{st}}$. Hence item 1 is a rephrasing of the first conclusion of in Theorem 2.2.

Since $\|A \partial_{\varphi^{\text{st}}} U^{\text{st}}\| \leq \|A\| \|U^{\text{st}}\|$, and $\|A\|$ depends only on \mathcal{B}^{st} and Q , item 2 follows directly from item 1. \square

5 Weak KAM solutions of dominant Hamiltonians and convergence

In this section, we provide some basic information about the weak KAM solution of the dominant system.

In section 5.1, we give an overview on the relevant weak KAM theory. Recall that in section 4.2, we derive the relation between the slow Lagrangian and the strong Lagrangian. In section 5.2, we obtain a compactness result for the strong component of a minimizing curve.

5.1 Weak KAM solutions of Tonelli Lagrangian

For an extensive exposition of the topic, we refer to [Fat08].

Tonelli Lagrangian. The Lagrangian function $L = L(\varphi, v) : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is called Tonelli if it satisfies the following conditions.

1. (smoothness) L is C^r with $r \geq 2$.
2. (fiber convexity) $\partial_{vv}^2 L$ is strictly positive definite.
3. (superlinearity) $\lim_{\|v\| \rightarrow \infty} |L(x, v)|/\|v\| = \infty$.

The Lagrangians considered in this paper are Tonelli.

Minimizers. An absolutely continuous curve $\gamma : [a, b] \rightarrow \mathbb{T}^d$ is called *minimizing* for the Tonelli Lagrangian L if

$$\int_a^b L(\gamma, \dot{\gamma}) dt = \min_{\xi} \int_a^b L(\xi, \dot{\xi}) dt,$$

where the minimization is over all absolutely continuous curves $\xi : [a, b] \rightarrow \mathbb{T}^d$ with $b > a$, such that $\xi(a) = \gamma(a)$, $\xi(b) = \gamma(b)$. The functional

$$\mathbb{A}(\gamma) = \int_a^b L(\gamma, \dot{\gamma}) dt$$

is called *the action functional*. The curve γ is called an *extremal* if it is a critical point of the action functional. A minimizer is extremal, and it satisfies the Euler-Lagrange equation

$$\frac{d}{dt}(\partial_v L(\gamma, \dot{\gamma})) = \partial_\varphi L(\gamma, \dot{\gamma}).$$

Tonelli Theorem and a priori compactness. By the Tonelli Theorem (c.f [Fat08], Corollary 3.3.1), for any $[a, b] \subset \mathbb{R}$ with $b > a$, $\varphi, \psi \in \mathbb{T}^d$, there always exists a C^r

minimizer. Moreover, there exists $D > 0$ depending only on $b - a$ such that $\|\dot{\gamma}\| \leq D$ ([Fat08] Corollary 4.3.2). This property is called the a priori compactness.

The alpha function and minimal measures. A measure μ on $\mathbb{T}^d \times \mathbb{R}^d$ is called a closed measure (see [Sor10], Remark 4.40) if for all $f \in C^1(\mathbb{T}^d)$,

$$\int df(\varphi) \cdot v d\mu(\varphi, v) = 0.$$

This notion is equivalent to the more well known notion of holonomic measure defined by Mañe ([Mañ97]).

For $c \in H^1(\mathbb{T}^d, \mathbb{R}) \simeq \mathbb{R}^d$, the alpha function

$$\alpha_L(c) = - \inf_{\nu} \int (L(\varphi, v) - c \cdot v) d\nu(\varphi, v),$$

where the minimization is over all closed Borel probability measures. When $L = L_H$ we also use the notation $\alpha_H(c)$. A measure μ is called a c -minimizing if it reaches the infimum above. A minimizing measure is invariant under the Euler-Lagrange flow (c.f [Mañ97; Ber08]), and therefore this definition is equivalent to the one given in section 2.4.

Rotation number and the beta function. The rotation number ρ of a closed measure μ is defined by the relation

$$\int (c \cdot v) d\mu(\varphi, v) = c \cdot \rho, \quad \text{for all } c \in H^1(\mathbb{T}^d, \mathbb{R}).$$

For $h \in H_1(\mathbb{T}^d, \mathbb{R}) \simeq \mathbb{R}^d$, the beta function is

$$\beta_L(h) = \inf_{\rho(\nu)=h} \int L(\varphi, v) d\nu(\varphi, v).$$

When $L = L_H$ we use the notation $\beta_H(h)$. The alpha function and beta function are Legendre duals:

$$\beta_L(h) = \sup_{c \in \mathbb{R}^d} \{c \cdot h - \alpha_L(c)\}.$$

The Legendre-Fenichel transform. Define the Legendre-Fenichel transform associated to the beta function

$$\mathcal{LF}_\beta : H_1(\mathbb{T}^d, \mathbb{R}) \rightarrow \tag{5.1}$$

the collection of nonempty, compact convex subsets of $H^1(\mathbb{T}^d, \mathbb{R})$,

defined by

$$\mathcal{LF}_\beta(h) = \{c \in H^1(\mathbb{T}^d, \mathbb{R}) : \beta_L(h) + \alpha_L(c) = c \cdot h\}.$$

Domination and calibration. For $\alpha \in \mathbb{R}$, a function $u : \mathbb{T}^d \rightarrow \mathbb{R}$ is dominated by $L + \alpha$ if for all $[a, b] \subset \mathbb{R}$ and piecewise C^1 curves $\gamma : [0, T] \rightarrow \mathbb{T}^d$, we have

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma, \dot{\gamma}) dt + \alpha(b - a).$$

A piecewise C^1 curve $\gamma : I \rightarrow \mathbb{R}$ defined on an interval $I \subset \mathbb{R}$ is called (u, L, α) -calibrated if for any $[a, b] \subset I$,

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma, \dot{\gamma}) dt + \alpha(b - a).$$

Weak KAM solutions. A function $u : \mathbb{T}^d \rightarrow \mathbb{R}$ is called a weak KAM solution of L if there exists $\alpha \in \mathbb{R}$ such that the following hold.

1. u is dominated by $L + \alpha$.
2. For all $\varphi \in \mathbb{T}^d$, there exists a (u, L, α) -calibrated curve $\gamma : (-\infty, 0] \rightarrow \mathbb{T}^d$ with $\gamma(0) = \varphi$.

This definition of the weak KAM solution is equivalent to the one given in section 2.4 (see [Fat08], Proposition 4.4.8), and the constant $\alpha = \alpha_L(0)$, where α_L is the alpha function.

Peierls' barrier. For $T > 0$, we define the function $h_L^T : \mathbb{T}^d \times \mathbb{T}^d \rightarrow \mathbb{R}$ by

$$h_L^T(\varphi, \psi) = \min_{\gamma(0)=\varphi, \gamma(T)=\psi} \int_0^T (L(\gamma, \dot{\gamma}) + \alpha_L) dt.$$

Peierls' barrier is $h_L(\varphi, \psi) = \lim_{T \rightarrow \infty} h_L^T(\varphi, \psi)$. The limit exists, and the function h_L is Lipschitz in both variables. Denote $h_{L,c} = h_{L-c \cdot v}$.

Mather, Aubry and Mañe sets. These sets are defined by Mather (see [Mat93]). Here we only introduce the projected version. Define the Aubry and the Mañe sets as

$$\mathcal{A}_L(c) = \{x \in \mathbb{T}^d : h_{L,c}(x, x) = 0\},$$

$$\mathcal{N}_L(c) = \left\{ y \in \mathbb{T}^d : \min_{x, z \in \mathcal{A}_L(c)} (h_{L,c}(x, y) + h_{L,c}(y, z) - h_{L,c}(x, z)) = 0 \right\}.$$

The Mather set is $\tilde{\mathcal{M}}_L(c) = \overline{\bigcup_{\mu} \text{supp}(\mu)}$ is the closure of the support of all c -minimal measures. Its projection $\pi \tilde{\mathcal{M}}(c) = \mathcal{M}(c)$ onto \mathbb{T}^d is called the projected Mather set. Then

$$\mathcal{M}_L(c) \subset \mathcal{A}_L(c) \subset \mathcal{N}_L(c).$$

When $L = L_H$ we also use the subscript H to identify these sets.

Static classes. For any $\varphi, \psi \in \mathcal{A}_L(c)$, Mather defined the following equivalence relation:

$$\varphi \sim \psi \text{ if } h_{L,c}(\varphi, \psi) + h_{L,c}(\psi, \varphi) = 0.$$

The equivalence classes defined by this equivalence condition are called the static classes. The static classes are linked to the family of weak KAM solutions, in particular, if there is only one static class, then the weak KAM solution is unique up to a constant.

In this section, we provide a few useful estimates in weak KAM theory, and prove Theorem 2.3. In section 5.2, we prove a projected version of the a priori compactness property. We then introduce an approximate version of Lipschitz property and use it to prove Theorem 2.3.

5.2 Minimizers of strong and slow Lagrangians, their a priori compactness

We prove a version of the a priori compactness theorem for the strong component. Recall that the notations $c = (c^{\text{st}}, c^{\text{wk}})$, $\bar{c} = c^{\text{st}} + A^{-1}Bc^{\text{wk}}$.

Proposition 5.1. *Fix $\mathcal{B}^{\text{st}}, \kappa > 1$. For any $R > 0$, there exists $M = M(\mathcal{B}^{\text{st}}, Q, R, \kappa)$ such that the following hold. For any*

$$(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}}) \cap \{\|U^{\text{st}}\|_{C^2} \leq R\} \text{ and } L^s = L_{\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})},$$

let $T \geq \frac{1}{2}$, $c \in \mathbb{R}^d$ and $\gamma = (\gamma^{\text{st}}, \gamma^{\text{wk}}) : [0, T] \rightarrow \mathbb{T}^d$ be a minimizer of $L^s - c \cdot v$. Then for \bar{c} , given by (4.5), we have

$$\|\dot{\gamma}^{\text{st}} - A\bar{c}\| \leq M.$$

We first state a lemma on the strong component of the action and relate minimizers of the slow system with those of the strong one.

Lemma 5.2. *In the notations of Proposition 5.1 for $T \geq \frac{1}{2}$ and $c \in \mathbb{R}^d$, let $\gamma = (\gamma^{\text{st}}, \gamma^{\text{wk}}) : [0, T] \rightarrow \mathbb{T}^d$ be a minimizer for the lagrangian $L^s - c \cdot v$. Then*

$$\int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}}) dt \leq \min_{\zeta} \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\zeta, \dot{\zeta}) dt + 2T\|U^{\text{wk}}\|_{C^0},$$

where the minimization is over all absolutely continuous $\zeta : [0, T] \rightarrow \mathbb{T}^m$ with $\zeta(0) = \gamma^{\text{st}}(0)$, $\zeta(T) = \gamma^{\text{st}}(T)$.

Proof. Let $\gamma_0^{\text{st}} : [0, T] \rightarrow \mathbb{T}^m$ be such that

$$\int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}}) dt = \min_{\zeta} \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\zeta, \dot{\zeta}) dt$$

with $\zeta(0) = \gamma^{\text{st}}(0)$, $\zeta(T) = \gamma^{\text{st}}(T)$. Define $\gamma_0 = (\gamma_0^{\text{st}}, \gamma_0^{\text{wk}}) : [0, T] \rightarrow \mathbb{T}^d$, by

$$\gamma_0^{\text{wk}}(t) = \gamma^{\text{wk}}(t) - A^{-1}B\gamma^{\text{st}}(t) + A^{-1}B\gamma_0^{\text{st}}(t) = \min_{\zeta \in C^2([0, T] \rightarrow \mathbb{T}^m)} \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\zeta, \dot{\zeta}) dt.$$

Note that

$$\gamma_0^{\text{wk}}(0) = \gamma^{\text{wk}}(0), \quad \gamma_0^{\text{wk}}(T) = \gamma^{\text{wk}}(T), \quad \dot{\gamma}_0^{\text{wk}} - A^{-1}B\dot{\gamma}_0^{\text{st}} = \dot{\gamma}^{\text{wk}} - A^{-1}B\dot{\gamma}^{\text{st}}. \quad (5.2)$$

Using (4.6) and (5.4), we have

$$\begin{aligned} L^s - c \cdot v + \frac{1}{2}c^{\text{wk}} \cdot \tilde{C}^{-1}c^{\text{wk}} &= L^{\text{st}} - \bar{c} \cdot v^{\text{st}} \\ &+ \frac{1}{2}(v^{\text{wk}} - B^T A^{-1}v^{\text{st}} - \tilde{C}^{\text{wk}}) \cdot \tilde{C}(v^{\text{wk}} - B^T A^{-1}v^{\text{st}} - \tilde{C}^{\text{wk}}) + U^{\text{wk}} \end{aligned} \quad (5.3)$$

Since γ is a minimizer for $L^s - c \cdot v$,

$$\int_0^T (L^s - c \cdot v)(\gamma, \dot{\gamma})dt \leq \int_0^T (L^s - c \cdot v)(\gamma_0, \dot{\gamma}_0)dt.$$

By (5.3), we have

$$\begin{aligned} &\int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}})dt + \int_0^T U^{\text{wk}}(\gamma(t))dt \\ &+ \frac{1}{2}(\dot{\gamma}^{\text{wk}} - B^T A^{-1}\dot{\gamma}^{\text{st}} - \tilde{C}^{\text{wk}}) \cdot \tilde{C}(\dot{\gamma}^{\text{wk}} - B^T A^{-1}\dot{\gamma}^{\text{st}} - \tilde{C}^{\text{wk}}) \\ &\leq \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}})dt + \int_0^T U^{\text{wk}}(\gamma_0(t))dt \\ &+ \frac{1}{2}(\dot{\gamma}_0^{\text{wk}} - B^T A^{-1}\dot{\gamma}_0^{\text{st}} - \tilde{C}^{\text{wk}}) \cdot \tilde{C}(\dot{\gamma}_0^{\text{wk}} - B^T A^{-1}\dot{\gamma}_0^{\text{st}} - \tilde{C}^{\text{wk}}). \end{aligned}$$

By (5.2), the second and fourth line of the above inequality cancels, therefore

$$\int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}})dt \leq \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}})dt + 2T\|U^{\text{wk}}\|_{C^0}.$$

□

Proof of Proposition 5.1. First, observe that any segments of a minimizer is still a minimizer. By dividing the interval $[0, T]$ into subintervals, it suffice to prove our proposition for $T \in [\frac{1}{2}, 1)$.

We first produce an upper bound for

$$\min_{\zeta} \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \frac{1}{2}\bar{c} \cdot A\bar{c})(\zeta, \dot{\zeta})dt.$$

By completing the squares as in Lemma 4.2, we have

$$L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \frac{1}{2}\bar{c} \cdot A\bar{c} = \frac{1}{2}(v^{\text{st}} - A\bar{c}) \cdot A^{-1}(v^{\text{st}} - A\bar{c}) + U^{\text{st}}(\varphi^{\text{st}}). \quad (5.4)$$

We then take

$$\zeta_0(t) = \gamma^{\text{st}}(0) + tA\bar{c} + \frac{t}{T}y$$

where $y \in [0, 1]^d$ is such that $\zeta_0(0) + TA\bar{c} + y = \gamma^{\text{st}}(T) \pmod{\mathbb{Z}^m}$. We then have $\dot{\zeta}_0 - A\bar{c} = \frac{1}{T}y$, so

$$\int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \frac{1}{2}\bar{c} \cdot A\bar{c})(\zeta_0, \dot{\zeta}_0) dt \leq \frac{1}{2T} \|A^{-1}\| \|y\|^2 + T \|U^{\text{st}}\|_{C^0} \leq d \|A^{-1}\| + \|U^{\text{st}}\|_{C^0}$$

using $T \in [0, 1)$ and $\|y\|^2 \leq d$.

Using Lemma 5.2, and adding $\frac{1}{2}\bar{c} \cdot A\bar{c}$ to the Lagrangian to both sides, we obtain

$$\begin{aligned} \int_0^T (L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \frac{1}{2}\bar{c} \cdot A\bar{c})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}}) dt &\leq 2T \|U^{\text{wk}}\| + d \|A^{-1}\| + \|U^{\text{st}}\|_{C^0} \\ &\leq d \|A^{-1}\| + \|U^{\text{st}}\|_{C^0} + 2 \|U^{\text{wk}}\|_{C^0} \end{aligned}$$

since $T \in [\frac{1}{2}, 1)$.

We now use the above formula get an L^2 estimate on $(\dot{\gamma}^{\text{st}} - A\bar{c})$ and use the Poincaré estimate to conclude. Using the above formula and (5.4), we have

$$\int_0^T (\dot{\gamma}^{\text{st}} - A\bar{c}) \cdot A^{-1}(\dot{\gamma}^{\text{st}} - A\bar{c}) dt \leq d \|A^{-1}\| + 2 \|U^{\text{st}}\|_{C^0} + 2 \|U^{\text{wk}}\|_{C^0}.$$

Using the fact that A^{-1} is strictly positive definite, we get

$$\|\dot{\gamma}^{\text{st}} - A\bar{c}\|_{L^2} \leq \|A\| (d \|A^{-1}\| + 2 \|U^{\text{st}}\|_{C^0} + 2 \|U^{\text{wk}}\|_{C^0}) =: M_1.$$

Then

$$\left\| \frac{1}{T} \int_0^T (\dot{\gamma}^{\text{st}} - A\bar{c}) dt \right\|^2 \leq \frac{1}{T^2} \int_0^T \|\dot{\gamma}^{\text{st}} - A\bar{c}\|^2 dt \leq 4M_1. \quad (5.5)$$

Moreover, from Lemma 4.3,

$$\|\ddot{\gamma}^{\text{st}}\| \leq M_2(\mathcal{B}^{\text{st}}, Q, \kappa, q, R).$$

The Poincaré estimate gives, for some uniform constant $D > 0$,

$$\left\| (\dot{\gamma}^{\text{st}} - A\bar{c}) - \frac{1}{T} \int_0^T (\dot{\gamma}^{\text{st}} - A\bar{c}) dt \right\|_{L^\infty} \leq \|\ddot{\gamma}^{\text{st}}\|_{L^\infty} \leq DM_2.$$

Combine with (5.5) and we conclude the proof. \square

5.3 Approximate Lipschitz property of weak KAM solutions

The weak KAM solutions of the dominant systems are Lipschitz, however, it is not clear if the Lipschitz constant is bounded as $\mu(\mathcal{B}^{\text{wk}}) \rightarrow \infty$. To get uniform estimates, we consider the following weaker notion.

Definition. For $D, \delta > 0$, a function $u : \mathbb{R}^d \rightarrow \mathbb{R}$ is called (D, δ) approximately Lipschitz if

$$|u(x) - u(y)| \leq D\|x - y\| + \delta, \quad x, y \in \mathbb{R}^d.$$

For $u : \mathbb{T}^d \rightarrow \mathbb{R}$, the approximate Lipschitz property is defined by its lift to \mathbb{R}^d .

In Proposition 5.3 and 5.4 we state the approximate Lipschitz property of a weak KAM solution in weak and strong angles.

Proposition 5.3. Fix \mathcal{B}^{st} , $\kappa > 1$. Assume that $q > 2(d - m)$. For $R > 0$, there exists a constant $M = M(\mathcal{B}^{\text{st}}, Q, \kappa, q, R) > 0$, such that for all

$$(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}}) \cap \{\|U^{\text{st}}\| \leq R\},$$

and

$$\delta(\mathcal{B}^{\text{wk}}) = M\mu(\mathcal{B}^{\text{wk}})^{-(\frac{q}{2}-d+m)},$$

let $u = u(\varphi^{\text{st}}, \varphi^{\text{wk}}) : \mathbb{T}^m \times \mathbb{T}^{d-m} \rightarrow \mathbb{R}$ be a weak KAM solution of

$$L_{\mathcal{H}^s}(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) - c \cdot v.$$

Then for all $\varphi^{\text{st}} \in \mathbb{T}^m$, the function $u(\varphi^{\text{st}}, \cdot)$ is (δ, δ) approximately Lipschitz.

Proposition 5.4. There exists a constant $M' = M'(\mathcal{B}^{\text{st}}, Q, \kappa, q, R) > 0$, let $\delta'(\mathcal{B}^{\text{wk}}) = M'(\mu(\mathcal{B}^{\text{wk}}))^{-(\frac{q}{2}-d+m)}$, and u be the weak KAM solution described in Proposition 5.3. Then for all $\varphi^{\text{wk}} \in \mathbb{T}^{d-m}$, the function $u(\cdot, \varphi^{\text{wk}})$ is (M', δ') approximately Lipschitz.

5.4 The alpha function and rotation number estimate

In this section we provide a few useful estimates in weak KAM theory and prove Theorem 2.3 using Propositions 5.3 and 5.4. Recall that the notations $c = (c^{\text{st}}, c^{\text{wk}})$, $\bar{c} = c^{\text{st}} + A^{-1}Bc^{\text{wk}}$.

Proposition 5.5. We have

$$\left| \alpha_{H^s}(c) - \alpha_{H^{\text{st}}}(\bar{c}) + \frac{1}{2}(\tilde{C}c^{\text{wk}}) \cdot c^{\text{wk}} \right| \leq \|U^{\text{wk}}\|_{C^0},$$

Proof. Let μ be a minimal measure for $L^s - c \cdot v$. Let π denote the natural projection from $(\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, v^{\text{wk}})$ to $(\varphi^{\text{st}}, v^{\text{st}})$. By Lemma 4.2 we have

$$\begin{aligned}
-\alpha_{H^s}(c) &= \int (L^s - c \cdot v) d\mu \\
&= \int (L^{\text{st}} - \bar{c} \cdot v^{\text{st}}) d\mu \circ \pi - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}} \\
&\quad + \int \left(\frac{1}{2} (w^{\text{wk}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1} (w^{\text{wk}} - \tilde{C} c^{\text{wk}}) + U^{\text{wk}} \right) d\mu \\
&\geq -\alpha_{H^{\text{st}}}(\bar{c}) - \|U^{\text{wk}}\|_{C^0} - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C} c^{\text{wk}}.
\end{aligned} \tag{5.6}$$

On the other hand, let μ^{st} be an ergodic minimal measure for $L^{\text{st}} - \bar{c} \cdot v^{\text{st}}$. For an L^{st} -Euler-Lagrange orbit $\varphi^{\text{st}}(t)$ in the support of μ^{st} , and any $\varphi_0^{\text{wk}} \in \mathbb{T}^{d-m}$, define

$$\varphi^{\text{wk}}(t) = \varphi_0^{\text{wk}} + B^T A^{-1} \varphi^{\text{st}}(t) + \tilde{C} c^{\text{wk}} t, \quad t \in \mathbb{R} \tag{5.7}$$

and write $\gamma = (\gamma^{\text{st}}, \gamma^{\text{wk}})$. We take a weak-* limit point μ^s of the probability measures $\frac{1}{T} \gamma|_{[0, T]}$ as $T \rightarrow +\infty$. Then μ^s is a closed measure (see section 5.1).

Since on the support of μ^s , $v^{\text{wk}} - B^T A^{-1} v^{\text{st}} - \tilde{C} c^{\text{wk}} = 0$, we have

$$\begin{aligned}
-\alpha_{H^s}(c) &\leq \int (L^s - c \cdot v) d\mu^s \\
&= \int (L^{\text{st}} - \bar{c} \cdot v^{\text{st}}) d\mu^{\text{st}} + \int U^{\text{wk}} d\mu - \frac{1}{2} \tilde{C} c^{\text{wk}} \cdot c^{\text{wk}} \\
&\leq -\alpha_{H^{\text{st}}}(\bar{c}) + \|U\|_{C^0} - \frac{1}{2} \tilde{C} c^{\text{wk}} \cdot c^{\text{wk}}.
\end{aligned}$$

□

The following proposition establishes relations between rotation numbers of minimal measures of the slow and strong systems.

Proposition 5.6. *Let μ^s be an ergodic minimal measure of $L^s - c \cdot v$, and let $(\rho^{\text{st}}, \rho^{\text{wk}})$ denote its rotation number. Then*

$$0 \leq \frac{1}{2} (\tilde{C}(\rho^{\text{wk}} - B^T A^{-1} \rho^{\text{st}} - \tilde{C} c^{\text{wk}})) \cdot (\rho^{\text{wk}} - B^T A^{-1} \rho^{\text{st}} - \tilde{C} c^{\text{wk}}) \leq \|U^{\text{wk}}\|_{C^0}$$

and

$$0 \leq \alpha_{H^{\text{st}}}(\bar{c}) + \beta_{H^{\text{st}}}(\rho^{\text{st}}) - \bar{c} \cdot \rho^{\text{st}} \leq \|U^{\text{wk}}\|_{C^0}.$$

Proof. Using (5.6) and the conclusion of Proposition 5.5, we have

$$\begin{aligned}
\|U^{\text{wk}}\|_{C^0} &\geq \int (L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c})) d\mu^s \circ \pi \\
&\quad + \int \frac{1}{2} (\tilde{C}^{-1} (w - \tilde{C} c^{\text{wk}})) \cdot (w - \tilde{C} c^{\text{wk}}) d\mu^s \circ \pi.
\end{aligned} \tag{5.8}$$

Note the first of the two integrals is non-negative by definition, we obtain

$$0 \leq \int \frac{1}{2} (w^{\text{wk}} - \tilde{C}c^{\text{wk}}) \cdot \tilde{C}^{-1} (w^{\text{wk}} - \tilde{C}c^{\text{wk}}) d\mu^s \leq \|U^{\text{wk}}\|_{C^0}.$$

Denote $\bar{w}^{\text{wk}} := \int w^{\text{wk}} d\mu^s = \rho^{\text{wk}} - B^T A^{-1} \rho^{\text{st}}$, and rewrite the left hand side of the last formula as

$$\begin{aligned} & \frac{1}{2} (\tilde{C}^{-1}(\bar{w}^{\text{wk}} - \tilde{C}c^{\text{wk}})) \cdot (\bar{w}^{\text{wk}} - \tilde{C}c^{\text{wk}}) + \int \tilde{C}^{-1}(\bar{w}^{\text{wk}} - \tilde{C}c^{\text{wk}}) \cdot (w^{\text{wk}} - \bar{w}) d\mu^s \\ & \quad + \frac{1}{2} \int (\tilde{C}^{-1}(w^{\text{wk}} - \bar{w}^{\text{wk}})) \cdot (w^{\text{wk}} - \bar{w}^{\text{wk}}) d\mu^s. \end{aligned}$$

Note that the second term vanishes and the third term is non-negative. Therefore

$$\frac{1}{2} (\tilde{C}^{-1}(\bar{w}^{\text{wk}} - \tilde{C}c^{\text{wk}})) \cdot (\bar{w}^{\text{wk}} - \tilde{C}c^{\text{wk}}) \leq \|U^{\text{wk}}\|_{C^0}$$

which is the first conclusion.

For the second conclusion, using (5.8), we get

$$\|U^{\text{wk}}\|_{C^0} \geq \int (L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c})) d\mu \circ \pi = \int L^{\text{st}} d\mu \circ \pi - \bar{c} \cdot \rho^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c}).$$

Using $\int L^{\text{st}} d\mu \circ \pi \geq \beta_{H^{\text{st}}}(\rho^{\text{st}})$ we get the upper bound of the second conclusion. The lower bound holds by definition. \square

5.5 Convergence of weak KAM solutions

We now prove Theorem 2.3. Fix \mathcal{B}^{st} and $\kappa > 1$.

Let $(\mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}})$ and $c_i = (c_i^{\text{st}}, c_i^{\text{wk}})$ be a sequence satisfying the assumption of the theorem, namely $\mu(\mathcal{B}_i^{\text{wk}}) \rightarrow \infty$, $p_i \rightarrow p_0$, $U_i^{\text{st}} \rightarrow U_0^{\text{st}}$ in C^2 , and $c^{\text{st}} + A_i^{-1} B_i c_i^{\text{wk}} \rightarrow \bar{c}$.

Item 1. Let u_i be the weak KAM solution to $L_i^s = L_{\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}})} - c_i \cdot v$. We first show the sequence $\{u_i\}$ is equi-continuous.

Let M^* be a constant larger than the constants in both Proposition 5.3 and 5.4. Using both propositions, for any $\varphi = (\varphi^{\text{st}}, \varphi^{\text{wk}})$, $\psi = (\psi^{\text{st}}, \psi^{\text{wk}})$,

$$|u_i(\varphi^{\text{st}}, \varphi^{\text{wk}}) - u_i(\psi^{\text{st}}, \psi^{\text{wk}})| \leq M^* \|\varphi^{\text{st}} - \psi^{\text{st}}\| + \delta_i \|\varphi^{\text{wk}} - \psi^{\text{wk}}\| + 2\delta_i,$$

where $\delta_i = M^*(\mu(\mathcal{B}_i^{\text{wk}}))^{-\frac{q}{2}-d+m}$.

Since $\delta_i \rightarrow 0$ as $i \rightarrow \infty$, for any $0 < \varepsilon < 1$ there exists $M > 0$ such that for all $i > M$, $3\delta_i < \frac{\varepsilon}{2}$. It follows that if $\|\varphi - \psi\| < \frac{\varepsilon}{2D} < 1$, then

$$|u_i(\varphi^{\text{st}}, \varphi^{\text{wk}}) - u_i(\psi^{\text{st}}, \psi^{\text{wk}})| < \varepsilon.$$

Since $\{u_i\}_{i \leq M}$ is a finite family, it is equi-continuous. In particular, there exist $\sigma > 0$ such that

$$|u_i(\varphi) - u(\psi)| < \varepsilon, \quad \text{if } 1 \leq i \leq M, \|\varphi - \psi\| < \sigma.$$

This proves equi-continuity. Moreover, since u_i are all periodic, $u_i - u_i(0)$ are equi-bounded, therefore Ascoli's theorem applies and the sequence is pre-compact in uniform norm.

Item 2. Let u be any accumulation point of $u_i - u_i(0)$, without loss of generality, we assume $u_i - u_i(0)$ converges to u uniformly. Proposition 5.3 implies that

$$\limsup_{i \rightarrow \infty} \sup_{\varphi^{\text{st}}} (\max u_i(\varphi^{\text{st}}, \cdot) - \min_i u_i(\varphi^{\text{st}}, \cdot)) \leq 2 \lim_{i \rightarrow \infty} \delta_i = 0,$$

therefore u is independent of φ^{wk} .

Item 3. From item 2, there exists $u^{\text{st}}(\varphi^{\text{st}}) = \lim_{i \rightarrow \infty} u_i(\varphi^{\text{st}}, \varphi^{\text{wk}})$. We show u^{st} is a weak KAM solution of $L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}} = L_{\mathcal{H}^{\text{st}}(p_0, U_0^{\text{st}})} - \bar{c} \cdot v^{\text{st}}$. Denote $L_i^{\text{st}} = L_{\mathcal{H}^{\text{st}}(p_i, U_i^{\text{st}})}$, we have $L_i^{\text{st}} \rightarrow L_0^{\text{st}}$ in C^2 .

We first show that u^{st} is dominated by $L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}$. Let $\xi^{\text{st}} : [0, T] \rightarrow \mathbb{T}^m$ be an extremal curve of L_0^{st} . In the same way as (5.7) in the proof of Proposition 5.5, we define $\xi_i = (\xi_i^{\text{st}}, \xi_i^{\text{wk}}) : [0, T] \rightarrow \mathbb{T}^m$ such that $\xi_i^{\text{st}}(0) = \xi^{\text{st}}(0)$, $\xi_i^{\text{st}}(T) = \xi^{\text{st}}(T)$ and $\dot{\xi}_i^{\text{st}} - B_i^T A_i^{-1} \dot{\xi} - \tilde{C}_i c_i^{\text{wk}} = 0$. Since for $\bar{c}_i = c_i^{\text{st}} + A_i^{-1} B_i c_i^{\text{wk}}$, u_i are dominated by $L_i^s - c_i \cdot v + \alpha_{H_i^s}(c_i)$, we have

$$\begin{aligned} u_i(\xi_i(T)) - u_i(\xi_i(0)) &\leq \int_0^T (L_i^s - c_i \cdot v^s + \alpha_{H_i^s}(c_i))(\xi_i, \dot{\xi}_i) dt \\ &= \int_0^T (L_i^{\text{st}} - \bar{c}_i \cdot v^{\text{st}})(\xi_i^{\text{st}}, \dot{\xi}_i^{\text{st}}) dt + \int_0^T (U_i^{\text{wk}}(\xi_i) + \alpha_{H_i^s}(c_i) - \frac{1}{2} \tilde{C}_i c_i^{\text{wk}} \cdot c_i^{\text{wk}}) dt, \end{aligned}$$

where the equality is due to $\dot{\gamma}_i^{\text{st}} - B_i^T A_i^{-1} \dot{\gamma} - \tilde{C}_i c_i^{\text{wk}} = 0$. Using the fact that $\|U_i^{\text{wk}}\|_{C^0} \rightarrow 0$, $L_i^{\text{st}} \rightarrow L_0^{\text{st}}$, and from Proposition 5.5, $\alpha_{H_i^s} - \frac{1}{2} \tilde{C}_i c_i^{\text{wk}} \cdot c_i^{\text{wk}} \rightarrow \alpha_{H^{\text{st}}}(\bar{c})$ as $i \rightarrow \infty$, we get

$$u^{\text{st}}(\xi^{\text{st}}(b)) - u^{\text{st}}(\xi^{\text{st}}(a)) \leq \int_a^b (L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c})) dt. \quad (5.9)$$

Therefore u^{st} is dominated by $L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}$.

Secondly, we show that for any $\varphi^{\text{st}} \in \mathbb{T}^m$, there exists a $(u^{\text{st}}, L_0^{\text{st}}, \bar{c})$ -calibrated curve $\gamma^{\text{st}} : (-\infty, 0] \rightarrow \mathbb{T}^m$ with $\gamma^{\text{st}}(0) = \varphi^{\text{st}}$. Because u_i are weak KAM solutions, for each i there exists a (u_i, L_i^s, c_i) -calibrated curve $\gamma_i = (\gamma_i^{\text{st}}, \gamma_i^{\text{wk}}) : (-\infty, 0] \rightarrow \mathbb{T}^d$. By Proposition 5.1, all γ_i^{st} are uniformly Lipschitz, so there exists a subsequence that converges in $C_{loc}^1((-\infty, 0], \mathbb{T}^d)$. Assume without loss of generality that $\gamma_i^{\text{st}} \rightarrow \gamma^{\text{st}}$, since $\gamma_i = (\gamma_i^{\text{st}}, \gamma_i^{\text{wk}})$ is extremal for L_i^s , we have

$$\ddot{\gamma}_i^{\text{st}} = \frac{d}{dt} (A_i I^{\text{st}} + B_i I^{\text{wk}}) = A_i \partial_{\varphi^{\text{st}}} U_i^{\text{st}} + B_i \partial_{\varphi^{\text{st}}} U_i^{\text{wk}}.$$

By our assumption, as $i \rightarrow \infty$, $A_i \rightarrow A := \partial_{v^{\text{st}}}^2 L_0^{\text{st}}$, and by Lemma 4.3 $\|B_i\| \|U_i^{\text{wk}}\|_{C^2} \rightarrow 0$.

$$\ddot{\gamma}^{\text{st}} = A \partial_{\varphi^{\text{st}}} U^{\text{st}}(\gamma^{\text{st}}), \quad (5.10)$$

which is the Euler-Lagrange equation for L_0^{st} .

Then for any $[a, b] \subset (-\infty, 0]$,

$$\begin{aligned} u_i(\gamma_i(b)) - u_i(\gamma_i(a)) &= \int_a^b (L_i^s - c_i \cdot v^s + \alpha_{H_i^s}(c_i))(\gamma_i, \dot{\gamma}_i) dt \\ &\geq \int_a^b (L^{\text{st}} - \bar{c}_i \cdot v^{\text{st}})(\gamma_i^{\text{st}}, \dot{\gamma}_i^{\text{st}}) dt + \int_a^b (U_i^{\text{wk}} + \alpha_{H_i^s}(c_i) - \frac{1}{2} \tilde{C}_i c_i^{\text{wk}} \cdot c_i^{\text{wk}})(\gamma_i, \dot{\gamma}_i) dt \end{aligned}$$

Take limit again to get

$$u^{\text{st}}(\gamma^{\text{st}}(b)) - u^{\text{st}}(\gamma^{\text{st}}(a)) \geq \int_a^b (L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c}))(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}}) dt.$$

Because γ^{st} is an L^{st} extremal curve (see (5.10)), we have (5.9) hold for γ^{st} . Combining with last displayed formula, we get (5.9) becomes an equality. Then γ^{st} is a calibrated curve for $L^{\text{st}} - \bar{c} \cdot v^{\text{st}} + \alpha_{H^{\text{st}}}(\bar{c})$, and u^{st} is a weak KAM solution.

6 The Mañe and the Aubry sets and the barrier function

We prove the following result.

Proposition 6.1. *Fix \mathcal{B}^{st} and $\kappa > 1$. Assume that $(\mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}})$ satisfies the assumptions of Theorem 2.3. Denote $H_i^s = \mathcal{H}_i^s(\mathcal{B}^{\text{st}}, \mathcal{B}_i^{\text{wk}}, p_i, U_i^{\text{st}}, \mathcal{U}_i^{\text{wk}})$, $H_0^{\text{st}} = \mathcal{H}^{\text{st}}(p_0, U_0^{\text{st}})$, $L_i^s = L_{H_i^s}$ and $L_0^{\text{st}} = L_{H_0^{\text{st}}}$.*

1. *Any limit point of $\varphi_i \in \mathcal{N}_{H_i^s}(c_i)$ is contained in $\mathcal{N}_{H_0^{\text{st}}}(\bar{c}) \times \mathbb{T}^{d-m}$.*
2. *If $\mathcal{A}_{H_0^{\text{st}}}(\bar{c})$ contains only finitely many static classes, then any limit point of $\varphi_i \in \mathcal{A}_{H_i^s}(c_i)$ is contained in $\mathcal{A}_{H_0^{\text{st}}}(\bar{c}) \times \mathbb{T}^{d-m}$.*
3. *Assume that $\mathcal{A}_{H^{\text{st}}}(\bar{c})$ contains only one static class. Let $\varphi_i = (\varphi_i^{\text{st}}, \varphi_i^{\text{wk}}) \in \mathcal{A}_{H_i^s}(c_i)$ be such that $\varphi_i^{\text{st}} \rightarrow \varphi^{\text{st}} \in \mathcal{A}_{H_0^{\text{st}}}(\bar{c})$. Then for any $\psi = (\psi^{\text{st}}, \psi^{\text{wk}}) \in \mathbb{T}^d$,*

$$\lim_{i \rightarrow \infty} h_{L_i^s, c_i}(\varphi, \psi) = h_{L_0^{\text{st}}, \bar{c}}(\varphi^{\text{st}}, \psi^{\text{st}}).$$

4. *Let $(\rho_i^{\text{st}}, \rho_i^{\text{wk}})$ be the rotation number of any c_i -minimal measure of L_i^s . Then we have*

$$\lim_{i \rightarrow \infty} (\rho_i^{\text{wk}} - B_i^T A_i^{-1} \rho_i^{\text{st}} - \tilde{C}_i c_i^{\text{wk}}) = 0,$$

and any accumulation point ρ of ρ_i^{st} is contained in the set $\partial \alpha_{H^{\text{st}}}(\bar{c})$.

The proof of item 2 requires additional discussion and is presented in Section 6.2. In Section 6.1 we prove item 1, 3 and 4.

6.1 The Mañe set and barrier function

We first state an alternate definition of the Aubry and Mañe sets due to Fathi (see also [Ber08]). Let u be a weak KAM solution for the Lagrangian L . We define $\bar{\mathcal{G}}(L, u)$ to be the set of points $(\varphi, v) \in \mathbb{T}^d \times \mathbb{R}^d$ such that there exists a (u, L, α_L) -calibrated curve $\gamma : (-\infty, 0] \rightarrow \mathbb{T}^d$, such that $(\varphi, v) = (\gamma(0), \dot{\gamma}(0))$. Let ϕ_t denote the Euler-Lagrange flow of L , then

$$\tilde{\mathcal{I}}(L, u) = \bigcap_{t \leq 0} \phi_t(\bar{\mathcal{G}}(L, u)), \quad \tilde{\mathcal{A}}_L = \bigcap_u \tilde{\mathcal{I}}(L, u), \quad \tilde{\mathcal{N}}_L = \bigcup_u \tilde{\mathcal{I}}(L, u), \quad (6.1)$$

where the union and intersection are over all weak KAM solutions of L . The Aubry set and Mañe set of $c \in H^1(\mathbb{T}^d, \mathbb{R})$ is defined as

$$\tilde{\mathcal{A}}_L(c) = \tilde{\mathcal{A}}_{L-c \cdot v}, \quad \tilde{\mathcal{N}}_L(c) = \tilde{\mathcal{N}}_{L-c \cdot v}.$$

The projected Aubry and Mañe sets are the projection of these sets to \mathbb{T}^d .

We now turn to the setting of Proposition 6.1. Let $L_i^s, L_0^{\text{st}}, c_i, \bar{c}$ be as in the assumption. The strategy of the proof is similar to the one in [Ber10].

Lemma 6.2. *Let u_i be a weak KAM solution of $L_i^s - c_i \cdot v$. Assume that $\tilde{\varphi}_i = (\varphi_i, v_i) \in \tilde{\mathcal{I}}(L_i^s - c_i \cdot v, u_i)$ satisfies $\tilde{\varphi}_i \rightarrow \tilde{\varphi} = (\varphi, v) = (\varphi^{\text{st}}, \varphi^{\text{wk}}, v^{\text{st}}, v^{\text{wk}})$, and $u_i(\varphi^{\text{st}}, \varphi^{\text{wk}}) \rightarrow u^{\text{st}}(\varphi^{\text{st}})$. Then*

$$(\varphi^{\text{st}}, v^{\text{st}}) \in \tilde{\mathcal{I}}(L^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}}).$$

Proof. We first show that $(\varphi_i, v_i) \in \bar{\mathcal{G}}(L_i^s - c_i \cdot v, u_i)$ implies $(\varphi^{\text{st}}, v^{\text{st}}) \in \bar{\mathcal{G}}(L^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}})$. Indeed, there exists $\gamma_i : (-\infty, 0] \rightarrow \mathbb{T}^d$, each $(u_i, L_i^s - c_i \cdot v, \alpha_{L_i^s}(c_i))$ -calibrated, with $(\gamma_i, \dot{\gamma}_i)(0) = (\varphi, v)$. We follow the same line as proof of item 3 in Theorem 2.3 (section 5), then by restricting to a subsequence, γ_i^{st} converges in $C_{loc}^1((-\infty, 0], \mathbb{T}^d)$ to a $(u^{\text{st}}, L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}, \alpha_{H_0^{\text{st}}}(\bar{c}))$ -calibrated curve γ^{st} . In particular $(\gamma_i^{\text{st}}, \dot{\gamma}_i^{\text{st}}) \rightarrow (\gamma^{\text{st}}, \dot{\gamma}^{\text{st}})$, which implies $(\varphi^{\text{st}}, v^{\text{st}}) \in \bar{\mathcal{G}}(L^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}})$.

Let ϕ_t^i denote the Euler-Lagrange flow of L_i^s , and ϕ_t^{st} the flow for L^{st} . Let Π^{st} denote the projection to the strong components $(\varphi^{\text{st}}, v^{\text{st}})$, then from Lemma 4.3 $\Pi^{\text{st}} \phi_t^i \rightarrow \phi_t^{\text{st}}$ uniformly. As a result for a fixed $T > 0$ and $(\varphi_i, v_i) \in \tilde{\mathcal{I}}(L_i^{\text{st}} - c_i \cdot v, u_i)$, we have

$$(\varphi_i, v_i) = (\varphi_i^{\text{st}}, \varphi_i^{\text{wk}}, v_i^{\text{st}}, v_i^{\text{wk}}) \in \phi_{-T}^i(\bar{\mathcal{G}}(L_i^s - c_i \cdot v, u_i)),$$

hence $(\varphi_i^{\text{st}}, v_i^{\text{st}}) \rightarrow (\varphi^{\text{st}}, v^{\text{st}}) \in \varphi_{-T}^{\text{st}}(\bar{\mathcal{G}}(L^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}}))$. Since $T > 0$ is arbitrary, we obtain $(\varphi^{\text{st}}, v^{\text{st}}) \in \tilde{\mathcal{I}}(L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}})$. \square

Proof of Proposition 6.1, part I. We first prove item 1. Suppose $\tilde{\varphi}_i \in \tilde{\mathcal{N}}_{H_i^s}(c_i)$, then there exists weak KAM solutions u_i of $L_i^s - c_i \cdot v$, such that $(\varphi_i, v_i) \in \tilde{\mathcal{I}}(L_i^s - c_i \cdot v, u_i)$. By Theorem 2.3, after restricting to a subsequence, we have $u_i(\varphi^{\text{st}}, \varphi^{\text{wk}}) \rightarrow u^{\text{st}}(\varphi^{\text{st}})$. By Lemma 6.2, $(\varphi_i^{\text{st}}, v_i^{\text{st}}) \rightarrow (\varphi^{\text{st}}, v^{\text{st}})$ implies $(\varphi^{\text{st}}, v^{\text{st}}) \in \tilde{\mathcal{I}}(L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}}) \subset \tilde{\mathcal{N}}_{H_0^{\text{st}}}(\bar{c})$.

For item 3, suppose $\varphi_i = (\varphi_i^{\text{st}}, \varphi_i^{\text{wk}}) \in \mathcal{A}_{H_i^s}(c_i)$ satisfies $\varphi_i^{\text{st}} \rightarrow \varphi^{\text{st}} \in \mathcal{A}_{H_0^{\text{st}}}(\bar{c})$. Then $h_{L_i^s, c_i}(\varphi_i, \cdot)$ is a weak KAM solution of $L_i^s - c \cdot v$ (see [Fat08], Theorem 5.3.6). By

Theorem 2.3, by restricting to a subsequence, there exists a weak KAM solution u^{st} of $L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}$ such that

$$\lim_{i \rightarrow \infty} h_{L_i^s, c_i}(\varphi_i, \psi^{\text{st}}, \psi^{\text{wk}}) - h_{L_i^s, c_i}(\varphi_i, 0, 0) = u^{\text{st}}(\psi^{\text{st}}).$$

We may further assume that $h_{L_i^s, c_i}(\varphi_i, 0, 0) \rightarrow C \in \mathbb{R}$. Since $\mathcal{A}_{H_0^{\text{st}}}(\bar{c})$ has only one static class, there exists a constant $C_1 > 0$ such that

$$u^{\text{st}}(\psi^{\text{st}}) + C_1 = h_{L^{\text{st}}, \bar{c}}(\varphi^{\text{st}}, \psi^{\text{st}}).$$

Using the fact that $\varphi_i \in \mathcal{A}_{H_i^s}(c_i)$, we get $h_{L_i^s, c_i}(\varphi_i, \varphi_i) = 0$. Taking the limit,

$$u^{\text{st}}(\varphi^{\text{st}}) = -C_1 = h_{L^{\text{st}}, \bar{c}}(\varphi^{\text{st}}, \varphi^{\text{st}}) - C = -C.$$

Therefore

$$\lim_{i \rightarrow \infty} h_{L_i^s, c_i}(\varphi_i^{\text{st}}, \varphi_i^{\text{wk}}, \psi^{\text{st}}, \psi^{\text{wk}}) = h_{L^{\text{st}}, \bar{c}}(\varphi^{\text{st}}, \psi^{\text{st}}).$$

Item 4: Let $\rho_i = (\rho_i^{\text{st}}, \rho_i^{\text{wk}})$ be the rotation number of minimal measures of $L_i^s - c_i \cdot v$, then from Proposition 5.6,

$$\lim_{i \rightarrow \infty} \rho_i^{\text{wk}} - B_i^T A_i^{-1} \rho_i^{\text{st}} - \tilde{C}_i c_i^{\text{wk}} = 0.$$

Moreover, assume that $\rho_i^{\text{st}} \rightarrow \rho^{\text{st}} \in \mathbb{R}^m$, then by taking limit in the second conclusion of Proposition 5.6, we get

$$\alpha_{H_0^{\text{st}}}(\bar{c}) + \beta_{H_0^{\text{st}}}(\rho^{\text{st}}) - \bar{c} \cdot \rho^{\text{st}} = 0,$$

using Fenchel duality, ρ^{st} is a subdifferential of the convex function $\alpha_{H_0^{\text{st}}}$ at \bar{c} . \square

6.2 Semi-continuity of the Aubry set

Our strategy of the proof mostly follow [Ber10].

Given a compact metric space \mathcal{X} , a semi-flow ϕ_t on \mathcal{X} , and $\varepsilon, T > 0$, an (ε, T) -chain consists of $x_0, \dots, x_N \in \mathcal{X}$ and $T_0, \dots, T_{N-1} \geq T$, such that $d(\phi_{T_i} x_i, x_{i+1}) < \varepsilon$. We say that $x \mathfrak{C}_{\mathcal{X}} y$ if for any $\varepsilon, T > 0$, there exists an (ε, T) -chain with $x_0 = x$ and $x_N = y$. The relation $\mathfrak{C}_{\mathcal{X}}$ is called the chain transitive relation (see [Con88]).

The family of maps $\bar{\phi}_t = \phi_t$ defines a semi-flow on the set $\overline{\mathcal{G}(L - c \cdot v, u)}$, and therefore defines a chain transitive relation. Given $\varphi, \psi \in \mathbb{T}^d$ and a weak KAM solution u of $L - c \cdot v$, we say that $\varphi \mathfrak{C}_u \psi$ if there exists $\tilde{\varphi} = (\varphi, v), \tilde{\psi} = (\psi, w) \in \mathbb{T}^d \times \mathbb{R}^d$ such that

$$\tilde{\varphi} \mathfrak{C}_{\mathcal{X}} \tilde{\psi}, \text{ where } \mathcal{X} = \overline{\mathcal{G}(L - c \cdot v, u)}.$$

Item 1 in the following Proposition is due to Mañe, and item 2 is due to Mather. The version presented here is contained in [Ber10].

Proposition 6.3. *Let L be a Tonelli Lagrangian, then:*

1. Let $\varphi \in \mathcal{A}_L(c)$ and u be a weak KAM solution of $L - c \cdot v$, we have $\varphi \mathfrak{C}_u \varphi$.
2. Suppose $\mathcal{A}_L(c)$ has only finitely many static classes, and there exists a weak KAM solution u such that $\varphi \mathfrak{C}_u \varphi$. Then $\varphi \in \mathcal{A}_L(c)$.

Proposition 6.3 implies that, when $\mathcal{A}_L(c)$ has finitely many static classes, the Aubry set coincides with the set $\{\varphi : \varphi \mathfrak{C}_u \varphi\}$. We will prove semi-continuity for this set.

Definition. Let \mathcal{X} be a compact metric space with a semi-flow ϕ_t . A family of piecewise continuous curves $x_i : [0, T_i] \rightarrow \mathcal{X}$ is said to accumulate locally uniformly to (\mathcal{X}, ϕ_t) if for any sequence $S_i \in [0, T_i]$, the curves $x_i(t + S_i)$ has a subsequence which converges uniformly on compact sets to a trajectory of ϕ_t .

Lemma 6.4. [Ber10] Suppose $x_i : [0, T_i] \rightarrow \mathcal{X}$ accumulates locally uniformly to (X, ϕ_t) , $x_i(0) \rightarrow x$ and $x_i(T_i) \rightarrow y$, then $x \mathfrak{C}_X y$.

Proof of Proposition 6.1, part II. We prove item 2. Let $\varphi_i = (\varphi_i^{\text{st}}, \varphi_i^{\text{wk}}) \in \mathcal{A}_{H_i^s}(c_i)$ and $\varphi_i^{\text{st}} \rightarrow \varphi^{\text{st}}$, we show that $\varphi^{\text{st}} \in \mathcal{A}_{H_0^{\text{st}}}(\bar{c})$. According to Proposition 6.3, $\varphi_i \mathfrak{C}_u \varphi_i$. Let $\tilde{\varphi}_i$ be the unique point in $\tilde{\mathcal{A}}_{H_i^s}(c_i)$ projecting to φ_i , then there exists weak KAM solutions u_i of $L_i^s - c_i \cdot v$, such that $\tilde{\varphi}_i \mathfrak{C} \tilde{\varphi}_i$ in $\bar{\mathcal{G}}(L_i^s - c_i \cdot v, u_i)$. Fix $\varepsilon_i \rightarrow 0$ and $M_i \rightarrow \infty$, then for each i , there exists

$$T_{i,1} < \cdots < T_{i,N_i}, \quad T_{i,j+1} - T_{i,j} > M_i,$$

and a piecewise C^1 curve $\gamma_i = (\gamma_i^{\text{st}}, \gamma_i^{\text{wk}}) : [0, T_i] \rightarrow \mathbb{T}^d$, satisfying

1. $\gamma_i|_{(T_{i,j}, T_{i,j+1})}$ satisfies the Euler-Lagrange equation of L_i^s ;
- 2.

$$d(((\gamma_i(T_{i,j}-), \dot{\gamma}_i(T_{i,j}-)), ((\gamma_i(T_{i,j}+), \dot{\gamma}_i(T_{i,j}+)))) < \varepsilon_i.$$

Using Lemma 4.3, the projection of the Euler-Lagrange flow of L_i^s to $(\varphi^{\text{st}}, v^{\text{st}})$ converges uniformly over compact interval to the Euler-Lagrange flow of L_0^{st} . This, combined with item 2 and Lemma 6.2, implies that $(\gamma_i^{\text{st}}, \dot{\gamma}_i^{\text{st}})$ accumulates locally uniformly to

$$(\bar{\mathcal{G}}(L_0^{\text{st}} - \bar{c} \cdot v^{\text{st}}, u^{\text{st}}), \phi_{-t}^{\text{st}})$$

where ϕ_t^{st} is the Euler-Lagrange flow of L_0^{st} . Therefore $\varphi_i^{\text{st}} \rightarrow \varphi^{\text{st}}$ implies $\varphi^{\text{st}} \mathfrak{C}_u \varphi^{\text{st}}$. Using Proposition 6.3 again, we get $\varphi^{\text{st}} \in \mathcal{A}_{H_0^{\text{st}}}(\bar{c})$. \square

7 Technical estimates on weak KAM solutions

In this section we prove Proposition 5.3 and 5.4. For $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}) \in \Omega_{\kappa, q}^{m, d}(\mathcal{B}^{\text{st}}) \cap \{\|U^{\text{st}}\| \leq R\}$, recall the notations $H^s = \mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})$, $H^{\text{st}} = \mathcal{H}^{\text{st}}(p, U^{\text{st}})$, $L^s = L_{H^s}$, $L^{\text{st}} = L_{H^{\text{st}}}$.

7.1 Approximate Lipschitz property in the strong component

In this section we show that Proposition 5.3 implies Proposition 5.4. Proposition 5.3 is proven in the next two sections.

We first state a lemma of action comparison between an extremal curve and its “linear drift”.

Lemma 7.1. *Let $L : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a Tonelli Hamiltonian, $T \geq 1$, and $\gamma : [0, T] \rightarrow \mathbb{T}^d$ be an extremal curve. Then for any $1 \leq i \leq d$, $h > 0$, and a unit vector $f \in \mathbb{R}^d$,*

$$\begin{aligned} & \int_0^T L\left(\gamma + \frac{th}{T}f, \dot{\gamma} + \frac{h}{T}f\right)dt - \int_0^T L(\gamma, \dot{\gamma})dt \\ & \leq (\partial_v L(\gamma(T), \dot{\gamma}(T)) \cdot f)h + \left(\|f \cdot (\partial_{vv}^2 L)f\| \frac{1}{T} + \|f \cdot (\partial_{\varphi v}^2 L)f\| + T\|f \cdot (\partial_{\varphi\varphi}^2 L)f\| \right) h^2. \end{aligned}$$

Proof. We compute

$$\begin{aligned} L\left(\gamma + \frac{th}{T}f, \dot{\gamma}\right) - L\left(\gamma, \dot{\gamma} + \frac{h}{T}f\right) & \leq \partial_\varphi L(\gamma, \dot{\gamma}) \cdot \frac{th}{T}f + \partial_v L(\gamma, \dot{\gamma}) \frac{h}{T}f \\ & \quad \|f \cdot (\partial_{vv}^2 L)f\| \frac{h^2}{T^2} + \|f \cdot (\partial_{\varphi v}^2 L)f\| \frac{th^2}{T^2} + \|f \cdot (\partial_{\varphi\varphi}^2 L)f\| \frac{t^2 h^2}{T^2}. \end{aligned}$$

It follows from the Euler-Lagrange equation that

$$\partial_\varphi L(\gamma, \dot{\gamma}) \cdot \frac{th}{T} + \partial_v L(\gamma, \dot{\gamma}) \frac{h}{T} = \frac{d}{dt} \left(\partial_v L \frac{th}{T} \right),$$

and our estimate follows from direct integration. \square

The following lemma establishes a relation between “approximate semi concavity” with approximate Lipschitz property.

Lemma 7.2. *For $D, \delta > 0$, assume that $u : \mathbb{T}^d \rightarrow \mathbb{R}$ satisfies that for all $\varphi \in \mathbb{T}^d$, there exists $l \in \mathbb{R}^d$ such that*

$$u(\varphi + y) - u(\varphi) \leq l \cdot y + D\|y\|^2 + \delta, \quad y \in \mathbb{R}^d,$$

Then $\|l\| \leq \sqrt{d}(D + \delta)$, and u is $(2\sqrt{d}(D + \delta), \delta)$ approximately Lipschitz.

Proof. Assume that $l = (l_1, \dots, l_d)$. For each $1 \leq i \leq d$, we pick $y = -e_i \frac{l_i}{|l_i|}$, where e_i is the coordinate vector in φ_i . Then

$$0 = u(\varphi + e_i) - u(\varphi) \leq -|l_i| + D + \delta,$$

so $|l_i| \leq D + \delta$. As a result $\|l\| \leq \sqrt{d}(D + \delta)$. For any $y \in [0, 1]^d$, we have $\|y\| \leq \sqrt{d}$ and

$$u(\varphi + y) - u(\varphi) \leq (\sqrt{d}(D + \delta) + D\|y\|)\|y\| + \delta < 2\sqrt{d}(D + \delta)\|y\| + \delta.$$

\square

Proof of Proposition 5.4. Since u is a weak KAM solution, for any $\varphi \in \mathbb{T}^d$, let $\gamma = (\gamma^{\text{st}}, \gamma^{\text{wk}}) : (-\infty, 0] \rightarrow \mathbb{T}^d$ be a $(u, L^s - c \cdot v, \alpha_{H^s}(c))$ -calibrated curve with $\gamma(0) = \varphi = (\varphi^{\text{st}}, \varphi^{\text{wk}})$. Then for any $T > 0$

$$u(\varphi) = u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\gamma, \dot{\gamma}) dt.$$

Using (4.6), we get

$$\begin{aligned} u(\varphi) &= u(\gamma(-T)) + \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}}) dt + (\alpha_{H^s}(c) - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C}^{-1} c^{\text{wk}}) T \\ &\quad + \int_{-T}^0 \frac{1}{2} (\dot{\gamma}^{\text{wk}} - B^T A^{-1} \dot{\gamma}^{\text{st}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1} (\dot{\gamma}^{\text{wk}} - B^T A^{-1} \dot{\gamma}^{\text{st}} - \tilde{C} c^{\text{wk}}) + U^{\text{wk}}(\gamma(t)) dt. \end{aligned} \quad (7.1)$$

We now produce an upper bound using a special test curve. Let $\gamma_0^{\text{st}} : [-T, 0] \rightarrow \mathbb{T}^m$ be such that

$$\int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}}) dt = \min_{\zeta} \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\zeta, \dot{\zeta}) dt \quad (7.2)$$

where the minimum is over all $\zeta(-T) = \gamma^{\text{st}}(-T)$ and $\zeta(0) = \gamma^{\text{st}}(0)$.

We define $\xi = (\xi^{\text{st}}, \xi^{\text{wk}}) : [-T, 0] \rightarrow \mathbb{T}^d$ as follows.

1. For $y \in \mathbb{R}^d$,

$$\xi^{\text{st}}(t) = \gamma_0^{\text{st}}(t) + \frac{T+t}{T} y.$$

The curve ξ^{st} is a linear drift over γ_0^{st} with $h = \|y\|$ and $f = \frac{y}{\|y\|}$ (see Lemma 7.1).

2. Define

$$\xi^{\text{wk}}(t) = \gamma^{\text{wk}}(-T) + B^T A^{-1} (\xi^{\text{st}}(t) - \gamma_0^{\text{st}}(-T)) + \tilde{C} c^{\text{wk}}(T+t).$$

We note that $\xi^{\text{wk}}(-T) = \gamma^{\text{wk}}(-T)$ and

$$\dot{\xi}_0^{\text{wk}} - B^T A^{-1} \dot{\xi}^{\text{st}} - \tilde{C} c^{\text{wk}} = 0.$$

Using the fact that u is dominated by $L^s - c \cdot v + \alpha_{H^s}(c)$, we have

$$\begin{aligned} u(\varphi^{\text{st}} + y, \xi^{\text{wk}}(0)) &\leq u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\xi, \dot{\xi}) dt \\ &= u(\gamma(-T)) + \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\xi^{\text{st}}, \dot{\xi}^{\text{st}}) dt \\ &\quad + \int_{-T}^0 \frac{1}{2} (\dot{\xi}^{\text{wk}} - B^T A^{-1} \dot{\xi}^{\text{st}} - \tilde{C} c^{\text{wk}}) \cdot \tilde{C}^{-1} (\dot{\xi}^{\text{wk}} - B^T A^{-1} \dot{\xi}^{\text{st}} - \tilde{C} c^{\text{wk}}) dt \\ &\quad + (\alpha_{H^s}(c) - \frac{1}{2} c^{\text{wk}} \cdot \tilde{C}^{-1} c^{\text{wk}}) T + \int_{-T}^0 U^{\text{wk}}(\xi) dt \end{aligned}$$

and note that the third line in the above formula vanishes, using the definition of ξ^{wk} . Combine with (7.1), we get

$$\begin{aligned} & u(\varphi^{\text{st}} + y, \xi^{\text{wk}}(0)) - u(\varphi^{\text{st}}, \varphi^{\text{wk}}) \\ & \leq \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\xi^{\text{st}}, \dot{\xi}^{\text{st}}) dt - \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma^{\text{st}}, \dot{\gamma}^{\text{st}}) dt + 2\|U^{\text{wk}}\|_{C^0}. \end{aligned}$$

From (7.2) we get

$$\begin{aligned} & u(\varphi^{\text{st}} + y, \xi^{\text{wk}}(0)) - u(\varphi^{\text{st}}, \varphi^{\text{wk}}) \\ & \leq \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\xi^{\text{st}}, \dot{\xi}^{\text{st}}) dt - \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}}) dt + 2\|U^{\text{wk}}\|_{C^0}. \end{aligned}$$

Since γ_0^{st} is an extremal of $L^{\text{st}} - \bar{c} \cdot v^{\text{st}}$, the linear drift lemma (Lemma 7.1) applies. Noting that $\|\partial_{v^{\text{st}}}^2 L^{\text{st}}\| \leq \|A^{-1}\|$, $\|\partial_{\varphi^{\text{st}}}^2 L\| \leq \|U^{\text{st}}\|_{C^2} \leq R$, and $\partial_{\varphi^{\text{st}}}^2 L = 0$. We obtain from Lemma 7.1 that

$$\begin{aligned} & \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\xi^{\text{st}}, \dot{\xi}^{\text{st}}) dt - \int_{-T}^0 (L^{\text{st}} - \bar{c} \cdot v^{\text{st}})(\gamma_0^{\text{st}}, \dot{\gamma}_0^{\text{st}}) dt \\ & \leq l \cdot y + (\|A^{-1}\| + \|U^{\text{st}}\|_{C^2})\|y\|^2, \end{aligned}$$

where $l = \partial_v L^{\text{st}}(\gamma_0^{\text{st}}(0), \dot{\gamma}_0^{\text{st}}(0))$. Note that $\|A^{-1}\| + \|U^{\text{st}}\|_{C^2}$ is a constant depending only on $\mathcal{B}^{\text{st}}, Q, R$.

We now invoke Proposition 5.3 to get

$$|u(\varphi^{\text{st}} + y, \xi^{\text{wk}}(0)) - u(\varphi^{\text{st}} + y, \varphi^{\text{wk}})| \leq \delta |\xi^{\text{wk}}(0) - \varphi^{\text{wk}}| + \delta \leq 2\delta,$$

where $\delta = M_1^* \mu(\mathcal{B}^{\text{wk}})^{-\left(\frac{q}{2} - d + m\right)}$ for some $M_1^* = M_1^*(\mathcal{B}^{\text{st}}, Q, \kappa, q, R)$. Combine all the estimates, we get

$$u(\varphi^{\text{st}} + y, \varphi^{\text{wk}}) - u(\varphi^{\text{st}}, \varphi^{\text{wk}}) \leq l \cdot y + (\|A^{-1}\| + \|U^{\text{st}}\|_{C^2})\|y\|^2 + 2\delta + 2\|U^{\text{wk}}\|.$$

We note that in $\Omega_{\kappa, q}^{m, d}$ we have $\|U^{\text{wk}}\|_{C^2} \leq \sum_{i=1}^{d-m} \|U_i^{\text{wk}}\|_{C^2} \leq (d-m)\kappa(\mu(\mathcal{B}^{\text{wk}}))^{-q}$. We may choose $M_2^* = M_2^*(\mathcal{B}^{\text{st}}, Q, \kappa, q, R)$, such that

$$2\delta + 2\|U^{\text{wk}}\| \leq M_2^* \mu(\mathcal{B}^{\text{wk}})^{-\left(\frac{q}{2} - d + m\right)} =: \delta'.$$

We now apply Lemma 7.2 to get $u(\cdot, \varphi^{\text{wk}})$ is

$$(2\sqrt{d}(\|A^{-1}\| + \|U^{\text{st}}\|_{C^2} + \delta'), \delta')$$

approximately Lipschitz. Define $M' = 2\sqrt{d}(\|A^{-1}\| + \|U^{\text{st}}\|_{C^2} + M_2^*)$, and the Proposition follows. \square

7.2 Finer decomposition of the slow Lagrangian

We need a finer decomposition of the Lagrangian L^s which treat all φ_i^{wk} , $1 \leq i \leq d-m$ separately. First, we have the following linear algebra identity.

Lemma 7.3. *Let $S = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ be a nonsingular symmetric matrix in block form.*

Then

$$\begin{bmatrix} \text{Id} & 0 \\ -B^T A^{-1} & \text{Id} \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} \text{Id} & -A^{-1}B \\ 0 & \text{Id} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & \tilde{C} \end{bmatrix},$$

where $\tilde{C} = C - B^T A^{-1} B$. In particular, \tilde{C} is positive definite if S is.

We write $H^s(\varphi, I) = K(I) - U(\varphi) = K(I) - U^{\text{st}}(\varphi^{\text{st}}) - U^{\text{wk}}(\varphi)$ and $S = \partial_{II}^2 K$. We describe a coordinate change block diagonalizing $\partial_{II}^2 K$. Denote

$$S = \begin{bmatrix} X_{d-m} & y_{d-m} \\ y_{d-m}^T & z_{d-m} \end{bmatrix}, \quad X_{d-m} \in M_{(d-1) \times (d-1)}, \quad y_{d-m} \in \mathbb{R}^{d-1}, \quad z_{d-m} \in \mathbb{R},$$

and for each $1 \leq i \leq d-m-1$,

$$X_{i+1} = \begin{bmatrix} X_i & y_i \\ y_i^T & z_i \end{bmatrix}, \quad X_i \in M_{(m+i-1) \times (m+i-1)}, \quad y_i \in \mathbb{R}^{m+i-1}, \quad z_i \in \mathbb{R}.$$

Note that in this notation, $X_1 = \partial_{I^{\text{st}} I^{\text{st}}}^2 K = A$ (see (2.7)).

Define, for $1 \leq i \leq d-m$,

$$E_i = \begin{bmatrix} \text{Id}_{m+i-1} & -X_i^{-1} y_i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \text{Id}_{d-m-i} \end{bmatrix},$$

where Id_i denote the $i \times i$ identity matrix. Then by Lemma 7.3

$$E_{d-m}^T S E_{d-m} = \begin{bmatrix} \text{Id}_{d-1} & 0 \\ -y_{d-m}^T X_{d-m}^{-1} & 1 \end{bmatrix} \begin{bmatrix} X_{d-m} & y_{d-m} \\ y_{d-m}^T & z_{d-m} \end{bmatrix} \begin{bmatrix} \text{Id}_{d-1} & -X_{d-m}^{-1} y_{d-m} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} X_{d-m} & 0 \\ 0 & \tilde{z}_{d-m} \end{bmatrix},$$

where $\tilde{z}_{d-m} = z_{d-m} - y_{d-m}^T X_{d-m}^{-1} y_{d-m}$. Moreover, for each $1 \leq i \leq d-m-1$,

$$\begin{aligned} & \begin{bmatrix} \text{Id}_{m+i-1} & 0 \\ -y_i^T X_i^{-1} & 1 \end{bmatrix} X_{i+1} \begin{bmatrix} \text{Id}_{m+i-1} & -X_i^{-1} y_i \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \text{Id}_{m+i-1} & 0 \\ -y_i^T X_i^{-1} & 1 \end{bmatrix} \begin{bmatrix} X_i & y_i \\ y_i^T & z_i \end{bmatrix} \begin{bmatrix} \text{Id}_{m+i-1} & -X_i^{-1} y_i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} X_i & 0 \\ 0 & \tilde{z}_i \end{bmatrix}. \end{aligned} \quad (7.3)$$

Let

$$E = E_{d-m} \cdots E_1 = \begin{bmatrix} \text{Id}_m & -X_1^{-1}y_1 & -X_2^{-1}y_2 & \cdots & -X_{d-m}^{-1}y_{d-m} \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}, \quad (7.4)$$

then recursive computation yields

$$E^T S E = E_1^T \cdots E_{d-m}^T S E_{d-m} \cdots E_1 = \begin{bmatrix} X_1 & & & \\ & \tilde{z}_1 & & \\ & & \ddots & \\ & & & \tilde{z}_{d-m} \end{bmatrix} =: \tilde{S}. \quad (7.5)$$

We summarize the characterization of the Lagrangian in the following lemma. For $v = (v^{\text{st}}, v_1^{\text{wk}}, \dots, v_{d-m}^{\text{wk}}) \in \mathbb{R}^m \times \mathbb{R}^d$, we define

$$[v]_0 = v^{\text{st}}, \quad [v]_i = (v^{\text{st}}, v_1^{\text{wk}}, \dots, v_i^{\text{wk}}), \quad 1 \leq i \leq d-m. \quad (7.6)$$

Lemma 7.4. *For $v, c \in \mathbb{R}^d$ we denote $w = E^T v$ and $\eta = E^{-1}c$, where E is defined in (7.4). Explicitly, we have*

$$w = \begin{bmatrix} w^{\text{st}} \\ w_1^{\text{wk}} \\ \vdots \\ w_{d-m}^{\text{wk}} \end{bmatrix} = \begin{bmatrix} v^{\text{st}} \\ v_1^{\text{wk}} - y_1^T X_1^{-1} [v]_0 \\ \vdots \\ v_{d-m}^{\text{wk}} - y_{d-m}^T X_{d-m}^{-1} [v]_{d-m-1} \end{bmatrix}, \quad (7.7)$$

and

$$\eta^{\text{st}} = c^{\text{st}} + A^{-1} B c^{\text{wk}}, \quad \eta = (\eta^{\text{st}}, \eta^{\text{wk}}), \quad c = (c^{\text{st}}, c^{\text{wk}}),$$

where A, B are defined in (2.7). Then we have

$$\begin{aligned} & L^s(\varphi, v) - c \cdot v \\ &= L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) - \eta^{\text{st}} \cdot v^{\text{st}} + \sum_{i=1}^{d-m} \left(\frac{1}{2} \tilde{z}_i^{-1} (w_i^{\text{wk}} - \tilde{z}_i \eta_i^{\text{wk}})^2 - \frac{1}{2} z_i (\eta_i^{\text{wk}})^2 + U_i^{\text{wk}}(\varphi) \right). \end{aligned} \quad (7.8)$$

Remark. *This is a finer version of Lemma 4.2. In particular, the strong component $L^s - \eta^{\text{st}} \cdot v^{\text{st}}$ is identical to the $L^s - \bar{c} \cdot v^{\text{st}}$ defined in Lemma 4.2.*

Proof. Formula (7.7) can be read directly from the definition (7.4) and $w = E^T v$. To show $\eta^{\text{st}} = c^{\text{st}} + A^{-1} B c^{\text{wk}}$, we compute

$$\begin{bmatrix} A & 0 \\ 0 & \tilde{C} \end{bmatrix} \begin{bmatrix} \eta^{\text{st}} \\ \eta^{\text{wk}} \end{bmatrix} = \tilde{S} \eta = \tilde{S} E^{-1} c = E^T S c = \begin{bmatrix} \text{Id}_m & 0 \\ * & * \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} c^{\text{st}} \\ c^{\text{wk}} \end{bmatrix}.$$

The first block of the above equation yields $A\eta^{\text{st}} = Ac^{\text{st}} + Bc^{\text{wk}}$, hence $\eta^{\text{st}} = c^{\text{st}} + A^{-1}Bc^{\text{wk}}$.

We now prove (7.8). We have

$$\begin{aligned} L^s(\varphi, v) - c \cdot v &= \frac{1}{2}v^T S^{-1}v - c^T v + U^{\text{st}} + U^{\text{wk}} \\ &= \frac{1}{2}(E^T v)\tilde{S}^{-1}(E^T v) - (E^{-1}c)^T(E^T v) + U^{\text{st}} + U^{\text{wk}} \\ &= \left(\frac{1}{2}w^{\text{st}} \cdot A^{-1}w^{\text{st}} - \eta^{\text{st}} \cdot w^{\text{st}} + U^{\text{st}}\right) + \sum_{i=1}^{d-m} \left(\frac{1}{2}\tilde{z}_i^{-1}(w_i^{\text{wk}})^2 - \eta_i^{\text{wk}}w_i^{\text{wk}} + U_i^{\text{wk}}\right). \end{aligned}$$

In the above formula, the first group is equal to $L^{\text{st}} - \eta^{\text{st}} \cdot v^{\text{st}}$, noting $w^{\text{st}} = v^{\text{st}}$. Moreover

$$\frac{1}{2}\tilde{z}_i^{-1}(w_i^{\text{wk}})^2 - \eta_i^{\text{wk}}w_i^{\text{wk}} = \frac{1}{2}\tilde{z}_i^{-1}(w_i^{\text{wk}} - \tilde{z}_i\eta_i^{\text{wk}})^2 - \frac{1}{2}\tilde{z}_i(\eta_i^{\text{wk}})^2, \quad 1 \leq i \leq d-m,$$

and (7.8) follows. \square

We derive some useful estimates.

Lemma 7.5. *There exists $M^* = M^*(\mathcal{B}^{\text{st}}, Q, \kappa, q) > 1$ such that, for*

$$L^s = L_{\mathcal{H}^s}(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{st}}), \quad (\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{st}}) \in \Omega_{\kappa, q}^{m, d},$$

the following hold.

1. For each $1 \leq i \leq d-m$, we have $\sum_{j=i}^{d-m} \|U_j^{\text{wk}}\|_{C^2} \leq M^* |k_i^{\text{wk}}|^{-q}$.
2. For each $1 \leq i \leq d-m$, $\tilde{z}_i^{-1} \leq M^* |k_i^{\text{wk}}|^{2i}$.

Proof. For item 1, note that for each $j \geq i$, $|k_i^{\text{wk}}| \leq \kappa |k_j^{\text{wk}}|$, hence

$$\|U_j^{\text{wk}}\|_{C^2} \leq \kappa |k_j^{\text{wk}}|^{-q} \leq \kappa^{1+q} |k_i^{\text{wk}}|^{-q}.$$

Item 1 holds for any $M^* \geq (d-m)\kappa^{1+q}$.

For item 2, inverting (7.3) we get

$$X_{i+1}^{-1} = \begin{bmatrix} \text{Id}_{m+i-1} & -X_i^{-1}y_i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_i^{-1} & 0 \\ 0 & \tilde{z}_i^{-1} \end{bmatrix} \begin{bmatrix} \text{Id}_{m+i} & 0 \\ -y_{i+1}^T X_{i+1}^{-1} & 1 \end{bmatrix}.$$

Denote $f = (0, \dots, 0, 1) \in \mathbb{T}^{m+i}$, then

$$f^T X_{i+1} f = f^T \begin{bmatrix} \text{Id}_{m+i-1} & -X_i^{-1}y_i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_i^{-1} & 0 \\ 0 & \tilde{z}_i^{-1} \end{bmatrix} \begin{bmatrix} \text{Id}_{m+i} & 0 \\ -y_{i+1}^T X_{i+1}^{-1} & 1 \end{bmatrix} f = \tilde{z}_i^{-1}.$$

Moreover, using the definition (see (1.3))

$$S = \partial_{II}^2 K = \begin{bmatrix} k_1^{\text{st}} & \dots & k_m^{\text{st}} & k_1^{\text{wk}} & \dots & k_{d-m}^{\text{wk}} \end{bmatrix}^T Q \begin{bmatrix} k_1^{\text{st}} & \dots & k_m^{\text{st}} & k_1^{\text{wk}} & \dots & k_{d-m}^{\text{wk}} \end{bmatrix},$$

we have

$$\begin{aligned} X_{i+1} &= \begin{bmatrix} k_1^{\text{st}} & \cdots & k_m^{\text{st}} & k_1^{\text{wk}} & \cdots & k_i^{\text{wk}} \end{bmatrix}^T Q \begin{bmatrix} k_1^{\text{st}} & \cdots & k_m^{\text{st}} & k_1^{\text{wk}} & \cdots & k_i^{\text{wk}} \end{bmatrix} \\ &= \begin{bmatrix} \bar{k}_1^{\text{st}} & \cdots & \bar{k}_m^{\text{st}} & \bar{k}_1^{\text{wk}} & \cdots & \bar{k}_i^{\text{wk}} \end{bmatrix}^T Q_0 \begin{bmatrix} \bar{k}_1^{\text{st}} & \cdots & \bar{k}_m^{\text{st}} & \bar{k}_1^{\text{wk}} & \cdots & \bar{k}_i^{\text{wk}} \end{bmatrix}, \end{aligned}$$

where \bar{k} is the first n components of k . We have assumed $Q_0 \geq D^{-1}\text{Id}$ for $D > 1$. By Lemma B.2, there exists a constant $c_n > 1$ depending only on n such that

$$\|X_{i+1}^{-1}\| = \left(\min_{\|v\|=1} v^T X_{i+1} v \right)^{-1} \leq D c_n |k_1^{\text{st}}|^2 \cdots |k_m^{\text{st}}|^2 |k_1^{\text{wk}}|^2 \cdots |k_i^{\text{wk}}|^2 \leq D c_n \bar{M}^m \kappa^{i-1} |k_i^{\text{wk}}|^{2i},$$

where $\bar{M} = |k_1^{\text{st}}| + \cdots + |k_m^{\text{st}}|$ depend only on \mathcal{B}^{st} . \square

7.3 Approximate Lipschitz property in the weak component

In this section we prove Proposition 5.3. We fix $(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{st}}) \in \Omega_{\kappa, q}^{m, d} \cap \{\|U^{\text{st}}\|_{C^2} \leq R\}$, and write $L^s = L_{\mathcal{H}^s}(\mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{st}})$.

For $c \in \mathbb{R}^d$, we define

$$\begin{aligned} L_{c, i}^s(\varphi^{\text{st}}, \varphi_1^{\text{wk}}, \dots, \varphi_i^{\text{wk}}, v^{\text{st}}, v_1^{\text{wk}}, \dots, v_i^{\text{wk}}) &= L_{c, i}^s([\varphi]_i, [v]_i) \\ &= L^{\text{st}}(\varphi^{\text{st}}, v^{\text{st}}) - \eta^{\text{st}} \cdot v^{\text{st}} + \sum_{j=1}^i \left(\frac{1}{2} \tilde{z}_j^{-1} (w_j^{\text{wk}} - \tilde{z}_j \eta_j^{\text{wk}})^2 - \frac{1}{2} z_j (\eta_j^{\text{wk}})^2 + U_j^{\text{wk}}(\varphi) \right), \end{aligned} \quad (7.9)$$

then

$$\begin{aligned} L^s(\varphi, v) - c \cdot v &= L_{c, i}^s([\varphi]_i, [v]_i) + \sum_{j=i+1}^{d-m} \left(\frac{1}{2} \tilde{z}_j^{-1} (w_j^{\text{wk}} - \tilde{z}_j \eta_j^{\text{wk}})^2 - \frac{1}{2} z_j (\eta_j^{\text{wk}})^2 + U_j^{\text{wk}}(\varphi) \right). \end{aligned} \quad (7.10)$$

Our proof of Proposition 5.3 follows an inductive scheme. Following our notational convention, denote $e_i^{\text{wk}} = e_{i+m}$, which is the coordinate vector of φ_i^{wk} .

Lemma 7.6. *Let $u : \mathbb{T}^d \rightarrow \mathbb{R}$ be a weak KAM solution of $L^s - c \cdot v$. Then for*

$$\delta_{d-m} := 2(\tilde{z}_{d-m}^{-1} \|U_{d-m}^{\text{wk}}\|_{C^2})^{\frac{1}{2}},$$

we have u is δ_{d-m} -semi-concave and δ_{d-m} -Lipschitz in $\varphi_{d-m}^{\text{wk}}$.

Proof. First we have

$$\partial_{\varphi_{d-m}^{\text{wk}} \varphi_{d-m}^{\text{wk}}}^2 L^s = \partial_{\varphi_{d-m}^{\text{wk}} \varphi_{d-m}^{\text{wk}}}^2 U_{d-m}^{\text{wk}}, \quad \partial_{\varphi_{d-m}^{\text{wk}} v_{d-m}^{\text{wk}}}^2 L^s = 0, \quad \partial_{v_{d-m}^{\text{wk}} v_{d-m}^{\text{wk}}}^2 L^s = \tilde{z}_{d-m}^{-1}.$$

The first two equality follows directly from the definition, while the last one uses (7.7) and (7.8).

For any $\varphi \in \mathbb{T}^d$, let $\gamma : (-\infty, 0] \rightarrow \mathbb{T}^d$ be a (u, L^s, c) -calibrated curve with $\gamma(0) = \varphi$. Then for any $T > 0$

$$u(\varphi) = u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\gamma, \dot{\gamma}) dt.$$

Using the definition of the weak KAM solution,

$$u(\varphi + h e_i^{\text{wk}}) \leq u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\gamma + \frac{th}{T} e_{d-m}^{\text{wk}}, \dot{\gamma} + \frac{h}{T} e_{d-m}^{\text{wk}}) dt.$$

Subtract the two estimates, and apply Lemma 7.1 to $L^s - c \cdot v + \alpha_{H^s}(c)$ and γ , we get

$$\begin{aligned} u(\varphi + h e_i^{\text{wk}}) - u(\varphi) &\leq (\partial_{v_{d-m}^{\text{wk}}} L^s(\gamma(0), \dot{\gamma}(0)) - c_{d-m}) h \\ &\quad + \left(\|\partial_{v_{d-m} v_{d-m}}^2 L^s\| \frac{1}{T} + \|\partial_{\varphi_{d-m}^{\text{wk}} v_{d-m}^{\text{wk}}}^2 L^s\| + T \|\partial_{\varphi_{d-m}^{\text{wk}} \varphi_{d-m}^{\text{wk}}}^2 L^s\| \right) h^2 \\ &\leq (\partial_{v_{d-m}^{\text{wk}}} L^s(\gamma(0), \dot{\gamma}(0)) - c_{d-m}^{\text{wk}}) h + (\tilde{z}_{d-m}^{-1}/T + \|U_{d-m}^{\text{wk}}\|_{C^2} T) h^2, \end{aligned}$$

Take $T = (\tilde{z}_{d-m} \|U_{d-m}^{\text{wk}}\|_{C^2})^{-\frac{1}{2}}$, and write $l = \partial_{v_{d-m}^{\text{wk}}} L^s(\gamma(0), \dot{\gamma}(0)) - c_{d-m}^{\text{wk}}$, we get

$$u(\varphi + h e_i^{\text{wk}}) - u(\varphi) \leq l h + \frac{1}{2} \delta_{d-m} h^2.$$

The semi-concavity estimate follows. Using the fact that u is \mathbb{Z}^d periodic, we take $h = l/|l|$ to get $|l| \leq \frac{1}{2} \delta_{d-m}$. Therefore for $|h| \leq 1$,

$$|u(\varphi + h e_i^{\text{wk}}) - u(\varphi)| \leq \left(\frac{1}{2} \delta_{d-m} + \frac{1}{2} \delta_{d-m} h \right) h \leq \delta_{d-m} h.$$

This is the Lipschitz estimate. □

We now state the inductive step.

Proposition 7.7. *Let $u : \mathbb{T}^d \rightarrow \mathbb{R}$ be a weak KAM solution of $L^s - c \cdot v$. Assume that for a given $1 \leq i \leq d - m - 1$, u is (δ_j, δ_j) approximately Lipschitz in φ_j^{wk} for all $i + 1 \leq j \leq d - m$. Then for*

$$\sigma_i = \left(\tilde{z}_i^{-1} \sum_{j=i}^{d-m} \|U_j^{\text{wk}}\|_{C^2} \right)^{\frac{1}{2}}, \quad \delta_i = \sqrt{d} (6\sigma_i + 4 \sum_{j=i+1}^{d-m} \delta_j),$$

we have u is (δ_i, δ_i) approximately Lipschitz in φ_i^{wk} .

Proof. The proof is very similar to the proof of Proposition 5.4, but uses the finer decomposition in this section.

Since u is a weak KAM solution, then given any $\varphi \in \mathbb{T}^d$, there exists a calibrated curve $\gamma : (-\infty, 0] \rightarrow \mathbb{T}^d$ with $\gamma(0) = \varphi$. Then for any $T > 0$

$$u(\varphi) = u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\gamma, \dot{\gamma}) dt.$$

Let $h \in \mathbb{R}$, $\chi \in \mathbb{R}^d$, and a C^1 curve $\xi : [-T, 0] \rightarrow \mathbb{T}^d$ satisfies

$$\xi(-T) = \gamma(-T), \quad \xi(0) = \varphi + h e_i^{\text{wk}} + \chi,$$

then

$$\begin{aligned} u(\varphi + h e_i^{\text{wk}} + \chi) &\leq u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\xi, \dot{\xi}) dt \\ &\leq u(\gamma(-T)) + \int_{-T}^0 (L^s - c \cdot v + \alpha_{H^s}(c))(\gamma, \dot{\gamma}) dt \\ &\quad + \int_{-T}^0 (L^s - c \cdot v)(\xi, \dot{\xi}) - \int_{-T}^0 (L^s - c \cdot v)(\gamma, \dot{\gamma}) dt \\ &= u(\varphi) + \int_{-T}^0 (L^s - c \cdot v)(\xi, \dot{\xi}) - \int_{-T}^0 (L^s - c \cdot v)(\gamma, \dot{\gamma}) dt. \end{aligned} \tag{7.11}$$

We will first give the precise definition of ξ , then estimate (7.11), before finally obtain the desired estimate.

Definition of ξ . Recall the Lagrangian $L_{c,i}^s : \mathbb{T}^{m+i} \times \mathbb{R}^{m+i} \rightarrow \mathbb{R}$ defined in (7.9). Let $\xi : [-T, 0] \rightarrow \mathbb{T}^{m+i}$ be an $L_{c,i}^s$ minimizing curve satisfying the constraint

$$\zeta(-T) = [\gamma]_i(-T), \quad \zeta(0) = [\gamma]_i(0),$$

where $[\cdot]_i$ is defined in (7.6). For $h \in \mathbb{R}$, we define ξ in the following way.

1. The first $m + i$ components of ξ is ζ with an added linear drift in e_i^{wk} , more precisely,

$$[\xi]_i(t) = \zeta(t) + \frac{th}{T} e_i^{\text{wk}}. \tag{7.12}$$

2. We define the other components inductively. For $i < j \leq d - m$, suppose $[\xi]_{j-1}(t) = (\xi^{\text{st}}, \xi_1^{\text{wk}}, \dots, \xi_{j-1}^{\text{wk}})(t)$ has been defined. We define

$$\xi_j^{\text{wk}}(t) = \gamma_j^{\text{wk}}(t) + y_j^T X_j^{-1} [\xi]_{j-1}(t) - y_j^T X_j^{-1} [\gamma]_{j-1}(t).$$

For each $i < j \leq d - m$, we have

$$\begin{cases} \xi_j^{\text{wk}}(-T) = \gamma_j^{\text{wk}}(-T), \\ \dot{\xi}_j^{\text{wk}} - y_j^T X_j^{-1} [\dot{\xi}]_{j-1} = \dot{\gamma}_j^{\text{wk}} - y_j^T X_j^{-1} [\dot{\gamma}]_{j-1}. \end{cases} \quad (7.13)$$

We define $\chi = \xi(0) - \varphi - h e_i^{\text{wk}}$, and note that from (7.12),

$$[\chi]_i = [\xi]_i(0) - [\gamma]_i(0) - h e_i^{\text{wk}} = 0.$$

Action comparison. We now compute

$$\begin{aligned} & \int_{-T}^0 (L^s - c \cdot v)(\xi, \dot{\xi}) dt - \int_{-T}^0 (L^s - c \cdot v)(\gamma, \dot{\gamma}) dt \\ &= \int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s([\gamma]_i, [\dot{\gamma}]_i) dt + \sum_{j=i+1}^{d-m} \int_{-T}^0 (U_j^{\text{wk}}(\xi(t)) - U_j^{\text{wk}}(\gamma(t))) dt \\ &+ \frac{1}{2} \sum_{j=i+1}^{d-m} \tilde{z}_j^{-1} \int_{-T}^0 \left((\xi_j^{\text{wk}} - y_j^T X_j^{-1} [\dot{\xi}]_{j-1} - \tilde{z}_j \eta_j^{\text{wk}})^2 - (\gamma_j^{\text{wk}} - y_j^T X_j^{-1} [\dot{\gamma}]_{j-1} - \tilde{z}_j \eta_j^{\text{wk}})^2 \right) \\ &\leq \int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s([\gamma]_i, [\dot{\gamma}]_i) dt + 2T \sum_{j=i+1}^{d-m} \|U_j^{\text{wk}}\|_{C^0}. \end{aligned} \quad (7.14)$$

In the above formula, the equality is due to (7.10). Moreover, observe that from (7.13), the third line of the above formula vanishes. The inequality follows by replacing U_j^{wk} with its upper bound $\|U_j^{\text{wk}}\|_{C^0}$.

We now have

$$\begin{aligned} & \int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s([\gamma]_i, [\dot{\gamma}]_i) dt \\ &= \int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s(\zeta, \dot{\zeta}) dt + \int_{-T}^0 L_{c,i}^s(\zeta, \dot{\zeta}) dt - \int_{-T}^0 L_{c,i}^s([\gamma]_i, [\dot{\gamma}]_i) dt \\ &\leq \int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s(\zeta, \dot{\zeta}) dt, \end{aligned}$$

noting that ζ is minimizing for $L_{c,i}^s$.

Since ζ is minimizing and hence extremal for $L_{c,i}^s$, from the definition of ξ in (7.12), Lemma 7.1 applies. Hence

$$\int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s(\zeta, \dot{\zeta}) dt \leq l \cdot h + \left(\frac{1}{T} \tilde{z}_i^{-1} + T \left\| \sum_{j=i}^{d-m} U_j^{\text{wk}} \right\|_{C^2} \right) h^2,$$

where $l = \partial_{v_i}(L_{c,i}^s)(\zeta(0), \dot{\zeta}(0))$. As in the proof of Lemma 7.6, we choose $T = \left(\tilde{z}_i \sum_{j=i}^{d-m} \|U_j^{\text{wk}}\|_{C^2}\right)^{-\frac{1}{2}}$, we get

$$\int_{-T}^0 L_{c,i}^s([\xi]_i, [\dot{\xi}]_i) dt - \int_{-T}^0 L_{c,i}^s(\zeta, \dot{\zeta}) dt \leq l \cdot h + \sigma_i h^2, \quad \sigma_i = \left(\tilde{z}_i^{-1} \sum_{j=i}^{d-m} \|U_j^{\text{wk}}\|_{C^2}\right)^{\frac{1}{2}}.$$

Combine with (7.14), and use the upper bound $\sum_{j=i+1}^{d-m} \|U_j^{\text{wk}}\|_{C^0} \leq \sum_{j=i}^{d-m} \|U_j^{\text{wk}}\|_{C^2}$, we get

$$\int_{-T}^0 (L^s - c \cdot v)(\xi, \dot{\xi}) dt - \int_{-T}^0 (L^s - c \cdot v)(\gamma, \dot{\gamma}) dt \leq l \cdot h + \sigma_i h^2 + 2\sigma_i.$$

Estimating the weak KAM solution. Combine the last formula with (7.11), we get

$$u(\varphi + h e_i^{\text{wk}} + \chi) - u(\varphi) \leq l \cdot h + \sigma_i h^2 + \sigma_i.$$

Since $[\chi]_i = 0$, using the inductive assumption,

$$|u(\varphi + h e_i^{\text{wk}} + \chi) - u(\varphi + h e_i^{\text{wk}})| \leq 2 \sum_{j=i+1}^{d-m} \delta_j.$$

Therefore

$$u(\varphi + h e_i^{\text{wk}}) - u(\varphi) \leq l \cdot h + \sigma_i h^2 + 2\sigma_i + 2 \sum_{j=i+1}^{d-m} \delta_j.$$

We now use Lemma 7.2 to get for

$$\delta_i = 2\sqrt{d}(3\sigma_i + 2 \sum_{j=i+1}^{d-m} \delta_j),$$

u is (δ_i, δ_i) approximately Lipschitz in φ_i^{wk} . \square

Proof of Proposition 5.3. We have shown by induction that for all $1 \leq i \leq d-m$, u is (δ_i, δ_i) approximately Lipschitz in φ_i^{wk} , where δ_i are defined inductively in Lemma 7.6 and Proposition 7.7.

By Lemma 7.5, for each $1 \leq i \leq d-m$

$$\sigma_i = (\tilde{z}_i^{-1} \|U_i^{\text{wk}}\|_{C^2})^{\frac{1}{2}} \leq M^* |k_i^{\text{wk}}|^{-\frac{q}{2}+i-m}.$$

Then $\delta_{d-m} = 2\sigma_{d-m} \leq M^* |k_{d-m}^{\text{wk}}|^{-\frac{q}{2}+d-m}$. For each $1 \leq i \leq d-m$, we have

$$\delta_i = \sqrt{d}(6\sigma_i + 4 \sum_{j=i+1}^{d-m} \delta_j) \leq (6\sqrt{d})^{i-m} \sum_{j=i}^{d-m} \sigma_j \leq M^* (6\sqrt{d})^{i-m} |k_i^{\text{wk}}|^{-\frac{q}{2}+d-m}.$$

For any $\varphi^{\text{wk}}, \psi^{\text{wk}} \in \mathbb{T}^{d-m}$ and $\varphi^{\text{st}} \in \mathbb{T}^m$,

$$|u(\varphi^{\text{st}}, \varphi^{\text{wk}}) - u(\varphi^{\text{st}}, \psi^{\text{wk}})| \leq \sum_{i=1}^{d-m} \delta_i |\varphi_i^{\text{wk}} - \psi_i^{\text{wk}}| + \sum_{i=1}^{d-m} \delta_i$$

Since $\sum_{i=1}^{d-m} \delta_i \leq (d-m)M^*(6\sqrt{d})^{i-m}(\mu(\mathcal{B}^{\text{wk}}))^{-\frac{q}{2}+d-m}$, the proposition follows by replacing M^* by $(d-m)M^*(6\sqrt{d})^{i-m}$. \square

A Diffusion path with dominant structure

A.1 Diffusion path for Arnold diffusion

Our main motivation is to prove Arnold diffusion for a “typical” nearly integrable system of the form (1.1). The word “typical” here means the cusp residual condition introduced by Mather ([Mat03]).

Definition. For $r \geq 3$, we say that a property \mathcal{G} holds for a cusp residual set of C^r nearly integrable systems $H_\varepsilon = H_0 + \varepsilon H_1$, if:

- \mathcal{G} is an open property in C^r topology;
- There exists an open and dense set $\mathcal{V} \subset \{\|H_1\|_{C^r} = 1\}$, and a positive function $\varepsilon_0 : \mathcal{V} \rightarrow \mathbb{R}^+$, such that \mathcal{G} is C^r -dense on $\mathcal{U} = \{H_0 + \varepsilon H_1 : H_1 \in \mathcal{V}, 0 < \varepsilon < \varepsilon_0(H_1)\}$.

We would like to show that the property of topological instability is cusp residual. Instabilities for multidimensional Hamiltonian systems ($n \geq 3$) are studied in [Moe96; GK14a; CY09; BKZ11; KZ14; DLS13; Tre04; Tre12; Zhe10; Mar12a; Mar12b; KLS14].

We formulate our main conjecture.

Conjecture. There exists $r_0 > 0$ such that for each $n \geq 2$, $\gamma > 0$, $r_0 \leq r < \infty$, for a cusp residual set of C^r nearly integrable system, the system admits an orbit $(\theta_\varepsilon, p_\varepsilon)(t)$ such that $\{p_\varepsilon(t)\}_{t \in \mathbb{R}}$ is γ -dense on the unit ball $B^n := \{\|p\| \leq 1\}$.

The conjecture is a theorem for $n = 2$, we refer the reader to [Che13; KZ13] and reference therein. The proof in $n = 2$ follows two steps:

Step 1, define the set \mathcal{V} , which contains the set of “nondegenerate” H_1 . For $H_1 \in \mathcal{V}$, $H_0 + \varepsilon H_1$ possesses certain open structure of instability, such as normally hyperbolic invariant cylinders and the AM property mentioned below.

Step 2, show that for any $H_0 + \varepsilon H_1$ with $H_1 \in \mathcal{V}$ and ε sufficiently small, one can make an arbitrarily small perturbation to $H_0 + \varepsilon H_1$ such that there exists diffusion orbits.

In Theorem A.1, we prove a weaker version of *Step 1*. The heart of the argument is the construction of a diffusion path, on which all the essential resonances has a dominant structure. We expect the same diffusion path can be used to prove the full

conjecture. To avoid excessive length, we will give an outline of the proof with precise statements, but will not give full details.

A diffusion path \mathcal{P} is a subset in \mathbb{R}^n that the diffusion orbit $p_\varepsilon(t)$ roughly shadows. We pick a diffusion path that travels along a collection of $(n - 1)$ -resonances or, equivalently, along a collection of connected 1-dimensional resonant curves.

Definition. A diffusion path \mathcal{P} is a compact connected subset of

$$\bigcup \{ \Gamma_{\Lambda^{(n-1)}} : \Lambda^{(n-1)} \in \mathcal{L}^{(n-1)} \}$$

where $\mathcal{L}^{(n-1)} := \{ \Lambda_i^{(n-1)} \}_{i=1}^N$ is a collection of rank $n - 1$ irreducible resonant lattices (and each $\Gamma_{\Lambda^{(n-1)}}$ is a 1-dimensional resonant curve).³

We define the AM property of a mechanical system relative to an integer homology class.

- Let $H = K - U$ be a mechanical system on $\mathbb{T}^n \times \mathbb{R}^n$,
- h be an integer homology class,
- $S_E = \{H = E\}$ be energy surface.
- $\min U = 0$, and the minimum is unique.

Denote by $\pi : \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{T} \rightarrow \mathbb{R}^n$ the natural projection onto the action component.

Recall that homology and cohomology are related by Legendre-Fenichel transform $\mathcal{LF}_\beta(h) \subset H^1(\mathbb{T}^n, \mathbb{R})$ (see (5.1)). By a result of Diaz Carneiro [Car95] for each cohomology $c \in H^1(\mathbb{T}^n, \mathbb{R})$ the Aubry set $\mathcal{A}(c) \subset S_{\alpha(c)}$.

Definition. Let $\rho > 0$. We say that (H, h, ρ) has the AM property if for any λ such that $c \in \mathcal{LF}_\beta(\lambda h)$ and $\alpha_H(c) \geq \rho$ the Aubry set $\mathcal{A}(c)$ is a finite union of hyperbolic periodic orbits such that each of these periodic orbits as a closed curve has homology h .

Remark. Note that in the definition we do not consider the energy $0 \leq E < \rho$.

Let $\Lambda^{(n-1)} \subset \Sigma^{(n)}$ be two irreducible lattices of rank $n - 1$ and n respectively. Let $\mathcal{B}_e^{(n)} = [l_1, \dots, l_n]$ be an ordered basis of $\Sigma^{(n)}$ and $\mathcal{B}^{(n-1)} = [k_1, \dots, k_{n-1}]$ is be an ordered basis of $\Lambda^{(n-1)}$. Then $\Lambda^{(n-1)}$ and $\Sigma^{(n)}$ induce a (unique up to a sign) irreducible integer homology class denoted $h(\mathcal{B}^{(n-1)}, \mathcal{B}_e^{(n)}) \in H_1(\mathbb{T}^n, \mathbb{Z}) \simeq \mathbb{Z}^n$ (see (A.1) for details).

We now state the main theorem of this section.

Theorem A.1. *There exists $r_0 > 0, C > 0$ such that for each $n \geq 2, \rho > 0, r_0 \leq r < \infty$ for a cusp residual set of C^r nearly integrable systems $H_0 + \varepsilon H_1$, there exists a diffusion path $\mathcal{P} = \mathcal{P}(H_1, \rho)$ with a finite set \mathcal{E} called the punctures or strong resonances, with the following properties.*

³A remark on notation: the superscript $^{(n-1)}$ is not used as an index, but rather an indication for the rank of the lattice.

1. \mathcal{P} is ρ -dense in B^n , i.e. ρ -neighborhood of \mathcal{P} contains B^n .
2. For each 1-dimensional resonant curve $\Gamma_i \subset \Gamma_{\Lambda^{(n-1)}} \cap \mathcal{P}$ there is a 3-dimensional normally hyperbolic invariant cylinder (NHIC) $\tilde{\mathcal{C}}_i^3$ whose projection onto the action component $\text{dist}(\pi\tilde{\mathcal{C}}_i^3, \Gamma_i) \leq C\sqrt{\epsilon}$.⁴
3. (Away from strong resonances) For each $c \in \Gamma_i$ with $\text{dist}(c, \Sigma_n) \geq C\sqrt{\epsilon}$, we have $\mathcal{A}(c)$ belongs to $\tilde{\mathcal{C}}_i^3$.
4. (At strong resonance) Each puncture $p_0 \in \mathcal{E}$ is given by a rank n irreducible lattice $\Sigma^{(n)}$, i.e. $\{p_0\} = \Gamma_{\Sigma^{(n)}}$.
5. Let $p_0 \in \mathcal{P} \cap \Gamma_{\Sigma^{(n)}}$ be a puncture. Then $p_0 \subsetneq \Gamma_{\Lambda^{(n-1)}} \cap \mathcal{P}$ for some rank $n-1$ irreducible lattice $\Lambda^{(n-1)}$, with bases $\mathcal{B}^{(n-1)}$ and $\mathcal{B}_\epsilon^{(n)}$. For the induced homology $h = h(\mathcal{B}^{(n-1)} | \mathcal{B}_\epsilon^{(n)})$ and the slow mechanical system $H = H_{p_0, \mathcal{B}_\epsilon^{(n)}}$, defined in (1.2), we have that (H, h, ρ) have AM property.

We have the following remarks.

- If $n = 2$, a stronger version of Theorem A.1 hold. Namely, one can prove that for an *fixed* diffusion path, there exists a cusp residue set of systems $H_0 + \epsilon H_1$ for which the theorem hold. Whether this statement generalizes to higher degrees of freedom is an open question.

In our formulation, it is essential that the choice of diffusion path \mathcal{P} *does depend* on the perturbation ϵH_1 .

- Item 3 says that 3-dimensional cylinders $\tilde{\mathcal{C}}_i^3$ are minimal in the sense that they contain the Aubry sets with frequency vector from Γ_i away from maximal essential resonances.
- It turn out that away from strong resonances for each $h \in \Gamma_i$ with $\text{dist}(h, \Sigma_n) \geq C\sqrt{\epsilon}$ and $c \in \mathcal{LF}_\beta(h)$ we have not only that $\mathcal{A}(c)$ belongs to $\tilde{\mathcal{C}}_i^3$, but also it is a Lipschitz graph over a certain 2-torus \mathbb{T}_c^2 , i.e. for some submersion $\pi_c : \mathbb{T}^n \times B^n \times \mathbb{T} \rightarrow \mathbb{T}^2$ we have that $\pi_{\mathcal{A}(c)} : \mathcal{A}(c) \rightarrow \mathbb{T}_c^2$ is one-to-one and the inverse is Lipschitz. This is similar but more involved than what is presented in [BKZ11]. See discussion of $n = 3$ in [KZ14].
- 3-dimensional cylinders $\tilde{\mathcal{C}}_i^3$ for H_ϵ correspond to 2-dimensional cylinders \mathcal{C}^2 for averaged Hamiltonians.
- The cylinder $\tilde{\mathcal{C}}_i^3$ might consists of several connected components. At each maximal essential resonance $\Gamma_{\Lambda^{(n)}}$ this cylinder can have two connected components: one on each local component of $\Gamma_{\Lambda^{(n-1)}} \setminus \Gamma_{\Lambda^{(n)}}$.

⁴the cylinder is only weakly invariant in the sense that the Hamiltonian flow of H_ϵ is tangent to it

- The union of hyperbolic periodic orbits gives rise to a NHIC.
- Notice that at each strong resonance, due to our definition of AM property, we do not discuss the case low energy $0 \leq E < \rho$. This is why Theorem A.1 does not complete Step 1. For $n = 2$ a full description can be done, see [KZ13; Che13] and references therein. For $n = 3$, construction of NHIC for away from critical energy in general and normally hyperbolic invariant manifolds (NHIM) for critical energy for simple homologies is discussed in [KZ14], sect. 6.3. We expect these methods extend to arbitrary $n \geq 3$ (see also [Tur14]).
- We point out that presence of NHIC and NHIM is still not sufficient for diffusion as we need to construct the jump from one homology to another (see sect. 12 [KZ13]). In the case $n = 3$ it requires a lot more work (see sect. 8 [KZ14]). We expect to generalize this construction of the jump from [KZ14] to any $n \geq 3$.

A.2 Nondegeneracy conditions for Arnold diffusion

We now describe the set \mathcal{V} in Theorem A.1, using the conditions [H1] and [H2] to be defined later. Let $\rho > 0$ and $r_0 \leq r < \infty$. We say that $H_1 \in \mathcal{V}$ if $\|H_1\|_{C^r} = 1$, and there *exists* a diffusion path \mathcal{P} that is ρ -dense in B^n , with the following properties.

- For each $\Lambda^{(n-1)} \in \mathcal{L}^{(n-1)}$, and each connected component Γ of $\mathcal{P} \cap \Gamma_{\Lambda^{(n-1)}}$, there exists $\lambda > 0$ such that function H_1 satisfies condition [H1 λ] on Γ .
- For each $\lambda > 0$ and $\Lambda^{(n-1)} \in \mathcal{L}^{n-1}$, there exists a finite set of rank n resonant lattices $\mathcal{ES}(\Lambda^{(n-1)}, \lambda)$, with the property $\Sigma^{(n)} \supset \Lambda^{(n-1)}$ for each $\Sigma^{(n)} \in \mathcal{ES}(\Lambda^{(n-1)}, \lambda)$. Then $\Gamma_{\Sigma^{(n)}}$ is a single point contained in $\Gamma_{\Lambda^{(n-1)}}$. The collection $\mathcal{E} = \{\Gamma_{\Sigma^{(n)}} : \Sigma^{(n)} \in \mathcal{ES}(\Lambda^{(n-1)}, \lambda)\}$ is the set of punctures in Theorem A.1.
- Let $\lambda > 0$ be such that [H1 λ] is satisfied for H_1 . For each $\Sigma^{(n)} \in \mathcal{ES}(\Lambda^{(n-1)}, \lambda)$ and $\Lambda^{(n-1)} \in \mathcal{L}^{(n-1)}$ such that $\Gamma_{\Sigma^{(n)}} \in \mathcal{P}$, we choose basis $\mathcal{B}_e^{(n)}$ and $\mathcal{B}^{(n-1)}$. We say that H_1 satisfies condition [H2] at $\Gamma_{\Sigma^{(n)}}$ if for all such $\Lambda^{(n-1)} \subset \Sigma^{(n)}$,

$$(H_{p_0, \mathcal{B}^{(n)}}^s, h(\mathcal{B}^{(n-1)} | \mathcal{B}^{(n)}), \rho)$$

satisfies the AM property.

The condition [H1 λ], and the definition of $\mathcal{ES}(\Lambda^{(n-1)}, \lambda)$ and $h(\mathcal{B}^{(n-1)} | \mathcal{B}^{(n)})$ will be explained below. For the moment we only remark that for a fixed \mathcal{P} , the condition that [H1 λ] holds for some $\lambda > 0$ is open and dense; for each $\Gamma_{\Sigma^{(n)}}$, the AM property is open but not always dense. However, it is a dense condition if the lattice $\Sigma^{(n)}$ satisfies a domination property. The main idea is then, to pick a particular H_1 -dependent diffusion path \mathcal{P} , such that all the essential resonances on this path has this domination property.

The condition [H1]

We now describe our first set of non-degeneracy condition. For $\Lambda^{(n-1)} \in \mathcal{L}^{(n-1)}$, let us fix a basis \mathcal{B} . For $\lambda > 0$ and a connected compact subset $\Gamma^{(n-1)} \subset \Gamma_{\Lambda^{(n-1)}}$, we say that H_1 satisfies condition [H1 λ] on $\Gamma^{(n-1)}$ if

- For all $p \in \Gamma^{(n-1)}$, the function $Z_{\mathcal{B}}(\cdot, p)$ has at most two global maxima.
- At each global maxima φ^* of $Z_{\mathcal{B}}(\cdot, p)$, the Hessian $\partial_{\varphi\varphi}^2 Z_{\mathcal{B}}(\varphi^*, p) \leq -\lambda \text{Id}$ as quadratic forms.
- Suppose p_0 is such that there are two global maxima $\varphi_1^*(p_0)$ and $\varphi_2^*(p_0)$. Then they extend to local maxima for nearby $p \in \Gamma^{(n-1)}$. We assume that the functions $Z_{\mathcal{B}}(\varphi_1^*(p), p)$ and $Z_{\mathcal{B}}(\varphi_2^*(p), p)$ have different derivatives along $\Gamma^{(n-1)}$, with the difference at least λ .

We say that H_1 satisfies [H1] on $\Gamma^{(n-1)}$ if it satisfies [H1 λ] for some $\lambda > 0$. These conditions are introduced by Mather ([Mat03]) for $n = 2$ and assumed in [BKZ11]. We note that the quantitative version [H1 λ] of the condition depends on the choice of basis, while the qualitative version [H1] does not.

For H_1 satisfying [H1], there exists a finite set of rank n lattices containing $\Lambda^{(n-1)}$, which we will call $\mathcal{ES}(\Lambda^{(n-1)}, \lambda)$. More precisely, assume that the basis for $\Lambda^{(n-1)}$ is $\{k_1, \dots, k_{n-1}\}$ and there exists $M = M(\Lambda^{(n-1)}, \lambda) > 0$ such that

$$\mathcal{ES}(\Lambda^{(n-1)}, \lambda) = \{\Lambda_{k_1, \dots, k_{n-1}, k'} : |k'| \leq M\}.$$

For each $\Lambda^{(n)} \in \mathcal{ES}(\Lambda^{(n-1)}, \lambda)$, $\Gamma_{\Lambda^{(n)}}$ is a point contained in 1-dimensional curve $\Gamma_{\Lambda^{(n-1)}}$. The condition [H1] implies the existence of NHIC away from punctures, see [BKZ11]. It is not hard to see that item 1-5 of Theorem A.1 are direct consequences of our non-degeneracy conditions.

Induce homology and non-degeneracy

Fix $\Lambda^{(n)} \in \mathcal{ES}(\Lambda^{(n-1)}, \lambda)$, and let $\mathcal{B}^{(n)}$ be an ordered bases of $\Lambda^{(n)}$, $\mathcal{B}^{(n-1)}$ is an ordered basis of $\Lambda^{(n-1)}$ and $\{p_0\} = \Gamma_{\Lambda^{(n)}}$. Our second set of non-degeneracy condition concerns the slow system $H_{p_0, \mathcal{B}^{(n)}}^s : \mathbb{T}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, for a particular integer homology class $h(\mathcal{B}^{(n-1)} | \mathcal{B}^{(n)}) \in H^1(\mathbb{T}^n, \mathbb{Z})$, uniquely defined modulo the sign. We give a more general definition here.

Definition. For $2 \leq s \leq n$, irreducible lattices $\Lambda^{(s-1)} \subset \Lambda^{(s)}$, with corresponding basis $\mathcal{B}^{(s-1)} = [k_1, \dots, k_{s-1}]$ and $\mathcal{B}^{(s)} = [l_1, \dots, l_s]$. Since $k_i \in \Lambda^{(s)}$, there exists a unique collection $a_i \in \mathbb{Z}^s \setminus 0$, $i = 1, \dots, s-1$, such that

$$k_i = [l_1 \ \cdots \ l_s] a_i.$$

Then $h(\mathcal{B}^{(s-1)} | \mathcal{B}^{(s)}) \in \mathbb{Z}^s$ is defined by the relations

$$a_i \cdot h(\mathcal{B}^{(s-1)} | \mathcal{B}^{(s)}) = 0, \quad 1 \leq i \leq s-1. \quad (\text{A.1})$$

This definition is determined by the resonance relation

$$k_i \cdot (\dot{\theta}, 1) \approx k_i \cdot (\omega(p), 1) = 0, \quad 1 \leq i \leq n-1,$$

after converting to the variables $\varphi_i = l_i \cdot (\theta, t)$, $1 \leq i \leq n$.

We require the triplet

$$(H_{p_0, \mathcal{B}^{(n)}}^s, h(\mathcal{B}^{(n-1)} | \mathcal{B}^{(n)}), \rho)$$

satisfies the AM property. We have the following consequences of the AM property:

- (*Robustness*) The non-degeneracy condition is open.
- (*Minimality*) The condition guarantees, among other things, existence of an ordered collection of minimal *2-dimensional* normally hyperbolic invariant cylinders with heteroclinic connections of neighbors. Each cylinder is minimal in the sense that it is foliated by periodic orbits minimizing action of a certain variational problem.
- (*Hyperbolicity*) Each cylinder is hyperbolic in the sense that it consists of hyperbolic periodic orbits.

A.3 Properties of the nondegeneracy condition

Suppose $H^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}})$ is a dominant system. Then the AM property extends nicely from the strong system to the slow system. More precisely, the following properties hold.

Property A0. For $H = K - U$, the AM property for (H, h, ρ) is an open condition in both K and U .

Property A1. For a fixed quadratic form K and $h \in \mathbb{Z}^2$, there exists an open and dense set of $U \in C^2(\mathbb{T}^2)$ on which $(H = K - U, h, \rho)$ have AM property.

Property A2. Consider the data $(\mathcal{B}^{\text{st}}, Q_0, \kappa, q, \rho)$ and the space of corresponding dominant system $\Omega_{\kappa, q}^{m, m+2}(\mathcal{B}^{\text{st}})$. Assume that $U_0^{\text{st}} \in \mathbb{T}^m$ admits at most two non-degenerate minima.

Then there exists $M > 0$ and $\delta > 0$ depending only on $\mathcal{B}^{\text{st}}, Q_0, \kappa, q, p_0, U_0^{\text{st}}, \rho$ such that the following hold. For each \mathcal{B}^{wk} with

$$\mu(\mathcal{B}^{\text{wk}}) > M, \quad \|p - p_0\| < \delta, \quad \|U^{\text{st}} - U_0^{\text{st}}\|_{C^2} < \delta, \quad \text{and } h \in \mathbb{Z}^2,$$

for an *open and dense set* of \mathcal{U}^{wk} (in the space $\Omega_{\kappa, q}^{m, m+2}(\mathcal{B}^{\text{st}})$ restricted to fixed $\mathcal{B}^{\text{wk}}, p, U^{\text{st}}$), the triple

$$(\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}), g, \rho), \quad \text{with } g = (0, \dots, 0, h),$$

have AM property.

Property A3. Consider the data $(\mathcal{B}^{\text{st}}, Q_0, \kappa, q, \rho)$ and the space of corresponding dominant system $\Omega_{\kappa, q}^{m, m+1}(\mathcal{B}^{\text{st}})$. Assume that $p_0 \in \mathbb{R}^n$, $U_0^{\text{st}} \in C^r(\mathbb{T}^m)$, $h \in \mathbb{Z}^m$ satisfies

$$(\mathcal{H}^{\text{st}}(p_0, U_0^{\text{st}}), h, \rho) \text{ have AM property.}$$

Then there exists $M > \sup_{k \in \mathcal{B}^{\text{st}}} |k|$, $\delta > 0$ depending only on $\mathcal{B}^{\text{st}}, Q_0, \kappa, q, p_0, U_0^{\text{st}}, h, \rho$ such that the following hold. For each \mathcal{B}^{wk} with

$$\mu(\mathcal{B}^{\text{wk}}) > M, \quad \|p - p_0\| < \delta, \quad \|U^{\text{st}} - U_0^{\text{st}}\|_{C^2} < \delta,$$

and any nonzero pair of integers z, w , the following hold.

For an *open and dense set* of \mathcal{U}^{wk} (in the space $\Omega_{\kappa, q}^{m, m+1}(\mathcal{B}^{\text{st}})$ restricted to a fixed set of $\mathcal{B}^{\text{st}}, p, U^{\text{st}}$), the triple

$$(\mathcal{H}^s(\mathcal{B}^{\text{st}}, \mathcal{B}^{\text{wk}}, p, U^{\text{st}}, \mathcal{U}^{\text{wk}}), g_0, \rho), \text{ where } g_0 = (zh, w), \text{ has AM property.}$$

Remarks:

1. The list of properties A0 - A3 provides a setup for proving non-degeneracy using induction over degrees of freedom. Assume that U^{st} admits a non-degenerate minimum, then property A2 allows to extend this system by two more degrees of freedom, provided the homology g is only nontrivial in the weak variables. If H^{st} is nondegenerate in a nontrivial homology h , property A3 allows to extend by one degree of freedom, provided the new homology g is trivial in the weak variable.
2. Let us explain the proof briefly. Property A1 is a known result. This property is used in Arnold diffusion in $2\frac{1}{2}$ degrees of freedom, and we refer to [Mat10], [Mat11], [KZ13], [Che13] for more details.
3. Property A2 uses the first type of dimension reduction. The assumption ensures that H^{st} admits at most two minimal hyperbolic saddles. An arbitrarily small perturbation ensures that only one of them is minimal. Using Theorem 2.2, one obtain that H^s admits a minimal *four-dimensional* normally hyperbolic invariant cylinder (NHIC) \mathcal{C}^4 . Furthermore, Theorem 2.3 and Proposition 6.1 provide variational characterization for the cylinder. Then the restricted system to \mathcal{C}^4 behaves like a system with *two degrees of freedom*, and an analog of property A1 can be proven. In particular, there will be an ordered collection of minimal two-dimensional NHIC's contained in \mathcal{C}^4 .
4. For property A3, when the triple (H^{st}, h, ρ) satisfies the AM property, the strong system admits a family of *two dimensional* NHICs. Because there is only one weak component, Theorem 2.2 implies that H^s admits a minimal *four-dimensional* NHIC. Similar to the previous case, the idea from property A1 can be applied to prove nondegeneracy.

5. One can say that in the case A2 or A3, the slow system H^s is “dominated” by the strong system H^{st} .
6. While we state the properties A0-A4 for the AM property only, we expect the properties ensuring the full diffusion result also satisfies the same properties. As a result we expect the construction described in the next section to work for the full diffusion theorem as well.

A.4 Construction of a diffusion path and surgery of resonant manifolds

To prove Theorem A.1, it remains to construct a diffusion path with our non-degeneracy conditions.

Proposition A.2. *For each $\gamma > 0$, there exists an open and dense set $\mathcal{V} \subset \{\|H_1\|_{C^r} = 1\}$, such that for any $H_1 \in \mathcal{V}$, there exists a γ -dense diffusion path \mathcal{P} , such that the non-degeneracy conditions [H1] and [H2] are satisfied along \mathcal{P} .*

The proof of Proposition A.2 occupies the rest of this section. Since our nondegeneracy conditions are assumed to be open, it suffices to prove density. We fix an arbitrary relative open set $\mathcal{U}_0 \subset \{\|H_1\|_{C^r} = 1\}$, we will show there exists $H_1 \in \mathcal{U}_0$ such that the conclusions hold. The proof follows an inductive scheme. The strategy is as follows:

1. At step s we have a finite collection of integer irreducible lattices $\mathcal{L}^{(s)} = \{\Lambda_i^{(s)}\}$, i.e. each $\Lambda_i^{(s)} := \text{span}_{\mathbb{Z}}\{k_1^i, \dots, k_s^i\}$ has rank s and is spanned by integer vectors k_1^i, \dots, k_s^i . The union of corresponding codimension s resonant manifolds is called $\mathcal{P}^{(s)}$. We choose $\mathcal{L}^{(s)}$ such that $\mathcal{P}^{(s)}$ is $(1 - 2 \cdot 4^{-s})\gamma$ -dense in B^n , and $\mathcal{P}^{(s)} \subset \mathcal{P}^{(s-1)}$.
2. A set of essential resonances $\mathcal{E}^{(s+1)}$ is a collection of irreducible lattices of rank $s + 1$. This is a finite set given by the union

$$\mathcal{E}^{(s+1)} = \cup \mathcal{E}_j^{(s+1)} \quad \text{for } j = 1, \dots, s.$$

Each lattice $\Sigma^{(s+1)} \subset \mathcal{E}_j^{(s+1)}$ contains $\Lambda^{(j)}$ such that $\Lambda^{(j)} \in \mathcal{L}^{(j)}$ for some $j \leq s$. See diagram 5.

Essential resonances $\mathcal{E}^{(s+1)}$ correspond to codimension $s + 1$ resonant manifolds in B^n and have codimension one inside each resonant manifold from $\mathcal{P}^{(s)}$.

3. A nondegenerate set $\mathcal{N}_s \subset \mathcal{P}^{(s)}$, which is open, connected and $(1 - 2 \cdot 4^{-s})\gamma$ -dense in B^n , such that all the strong systems are nondegenerate, setting up for the next step of the induction.
4. The induction finishes at step $n - 1$, when we obtain an open, connected and γ -dense set $\mathcal{N}_{n-1} \subset \mathcal{P}^{(n-1)}$ in B^n which consists of 1-dimensional resonant manifolds and will be our diffusion path.

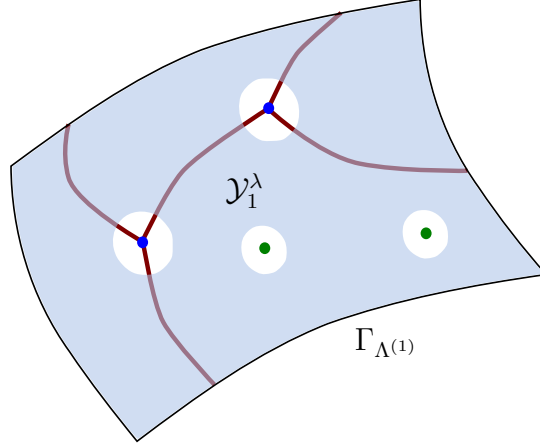


Figure 2: The first nondegeneracy set for $n = 3$: the shaded part is the λ -nondegenerate set. On the bold lines there are two minima for $U_{p, \mathcal{B}^{(1)}}$, at the blue dots there are three minima. At the green dots the minimum is degenerate.

A.4.1 An initial step of the induction

Since the union of all 1-resonant manifolds are dense and locally connected, for each $\gamma > 0$ we can pick $\mathcal{L}^{(1)} = \{\Lambda_i^{(1)}\}$ such that the set

$$\mathcal{P}^{(1)} = \bigcup_{\Lambda^{(1)} \in \mathcal{L}^{(1)}} \Gamma_{\Lambda^{(1)}} \cap B^n$$

is connected and $\gamma/2$ -dense in B^n . For each lattice $\mathcal{L}^{(1)}$ denote its basis by $\mathcal{B}^{(1)} = [k_1]$, i.e. $\mathcal{L}^{(1)} = \text{span}_{\mathbb{R}}(k_1) \cap \mathbb{Z}^{n+1}$. Denote by $\mathcal{K}^{(1)} = \{\mathcal{B}_i^{(1)}\}$ the union of basis vectors.

We define a first non-degeneracy set $\mathcal{Y}_1^\lambda(H_1, \mathcal{P}^{(1)}) \subset \mathcal{P}^{(1)}$ by the following condition: For any $\Lambda^{(1)} \in \mathcal{L}^{(1)}$ with basis $\mathcal{B}^{(1)}$ and $p \in \Gamma_{\Lambda^{(1)}} \cap \mathcal{P}^{(1)}$, the averaged potential $U_{p, \mathcal{B}^{(1)}}$ has at most two λ -nondegenerate minima.

Lemma A.3. *There exists a relative open set $\mathcal{U}_1 \subset \mathcal{U}_0$ and $\lambda_1 > 0$, such that for $H_1 \in \mathcal{U}_1$, the nondegeneracy set $\mathcal{Y}_1^{\lambda_1}(H_1, \mathcal{P}^{(1)})$ is open, connected, $4^{-2\gamma}$ -dense in $\mathcal{P}^{(1)}$, and $(1 - 2 \cdot 4^{-2\gamma})$ -dense in B^n .*

The set $\mathcal{Y}_1^{\lambda_1}(H_1, \mathcal{P}^{(1)})$ is the shaded set on Figure 2. For brevity in what follows we often omit dependence of $\mathcal{Y}_1^{\lambda_1}$ on H_1 and $\mathcal{P}^{(1)}$.

For each $p_0 \in \Gamma_{\Lambda^{(1)}} \cap \mathcal{Y}_1^{\lambda_1}$, where $\Lambda^{(1)}$'s basis is $\mathcal{B}^{(1)}$, the assumption of Property A2 is satisfied for $\mathcal{B}^{\text{st}} = \mathcal{B}^{(1)}$, p_0 and $U^{\text{st}} = U_{p_0, \mathcal{B}^{(1)}}$. Moreover, using compactness, for all

$$m = 1, \quad \Lambda^{(1)} \in \mathcal{L}^{(1)}, \quad p_0 \in \Gamma_{\Lambda^{(1)}} \cap \mathcal{Y}_1^{\lambda_1}, \quad U^{\text{st}} = U_{p_0, \mathcal{B}^{(1)}},$$

there exists a *uniform* $M_1 = M_1(\mathcal{P}^{(1)}, \mathcal{Y}_1^{\lambda_1})$, such that for all $\mathcal{B}^{\text{wk}} = [k_1^{\text{wk}}, k_2^{\text{wk}}]$ with $\mu(\mathcal{B}^{\text{wk}}) > M_1$ the conclusion of Property A2 is satisfied.

We define the first generation of *essential lattices* $\mathcal{E}^{(2)} = \mathcal{E}^{(1)}(\mathcal{P}^{(1)}, \mathcal{Y}_1^{\lambda_1})$. As before $\mathcal{B}^{(1)}$ denotes basis of $\Lambda^{(1)}$. We now have

$$\mathcal{E}^{(2)} := \{\Lambda^{(1)} \vee \{k'\} : \Lambda^{(1)} \in \mathcal{L}^{(1)}, |k'| \leq M_1(\mathcal{P}^{(1)}, \mathcal{Y}_1^{\lambda_1}) + \max_{k_1 \in \mathcal{B}^{(1)}} |k_1|\},$$

where $\Lambda^{(1)} \vee \{k'\} = \text{span}_{\mathbb{R}}(\Lambda^{(1)} \cup \{k'\}) \cap \mathbb{Z}^{n+1}$ is the smallest irreducible lattice containing both sets. By adding $\max_{k_1 \in \mathcal{B}^{(1)}} |k_1|$ we ensure that lattices generated by k_1, \bar{k}_1 are automatically essential. This corresponds to intersection Γ_{k_1} and $\Gamma_{\bar{k}_1}$.

The essential lattice set contain all lattices that is not “dominated” by the lattices in $\mathcal{L}^{(1)}$. Let us also denote

$$\Gamma_{\mathcal{E}^{(2)}} = \bigcup_{\Sigma^{(2)} \in \mathcal{E}^{(2)}} \Gamma_{\Sigma^{(2)}}$$

the union of all resonance manifolds corresponding to the essential lattices.

For each essential lattice $\Sigma^{(2)}$ we fix an ordered basis $\mathcal{B}_e^{(2)}$ (the actual choice is irrelevant). We define a second nondegeneracy set $\mathcal{Z}_1(H_1, \Sigma^{(2)}, \mathcal{Y}_1^{\lambda_1})$ to be the set of $p \in \Gamma_{\Sigma^{(2)}} \cap \mathcal{Y}_1^{\lambda_1}$ such that for each $k_1 \in \Sigma^{(2)}$ with $\mathcal{B}^{(1)} = [k_1] \in \mathcal{K}^{(1)}$, the pair

$$\left(H_{p, \mathcal{B}_e^{(2)}}^s, h(\mathcal{B}^{(1)} | \mathcal{B}_e^{(2)}) \right)$$

is nondegenerate. We then define $\mathcal{Z}_1(H_1, \mathcal{E}^{(2)}, \mathcal{Y}_1^{\lambda_1})$ to be the union of all $\mathcal{Z}_1(H_1, \Sigma^{(2)}, \mathcal{Y}_1^{\lambda_1})$ over essential resonances $\Sigma^{(2)} \in \mathcal{E}^{(2)}$.

Lemma A.4. *There exists a relative open set $\mathcal{U}'_1 \subset \mathcal{U}_1$ and a relative open $\tilde{\mathcal{Z}}_1 \subset \Gamma_{\mathcal{E}^{(2)}} \cap \mathcal{Y}_1^{\lambda_1}$ such that the following hold.*

1. For all $H_1 \in \mathcal{U}'_1$, $\tilde{\mathcal{Z}}_1$ is compactly contained in $\mathcal{Z}_1(H_1, \mathcal{E}^{(2)}, \mathcal{Y}_1^{\lambda_1})$.
2. The set

$$\mathcal{N}_1 := \mathcal{Y}_1^{\lambda_1} \cap \tilde{\mathcal{Z}}_1$$

is open, connected, and $4^{-2\gamma}$ -dense in $\mathcal{P}^{(1)}$.

We choose $\tilde{\mathcal{Z}}_1$ compactly contained in \mathcal{Z}_1 so that the nondegeneracy on $\tilde{\mathcal{Z}}_1$ is *uniform* due to compactness. The idea behind the definition of \mathcal{N}_1 is the following: On the set of essential resonances $\Gamma_{\mathcal{E}^{(2)}}$, domination does not apply, so we should remove it from the nondegeneracy set $\mathcal{Y}_1^{\lambda_1}$. However, in this case the remaining set becomes *disconnected* because the essential resonances divide the space (see Figure 3, left). Instead we only remove only p 's with the nearly degenerate essential resonances, i.e. $\mathcal{Y}_1^{\lambda_1} \cap \Gamma_{\mathcal{E}^{(2)}} \setminus \tilde{\mathcal{Z}}_1$ (see Figure 3, right dashed line).

A.4.2 Step 2 of the induction

We completed step 1 with

- the collection of rank one lattices $\mathcal{L}^{(1)}$, with associated bases $\mathcal{B}^{(1)}$,

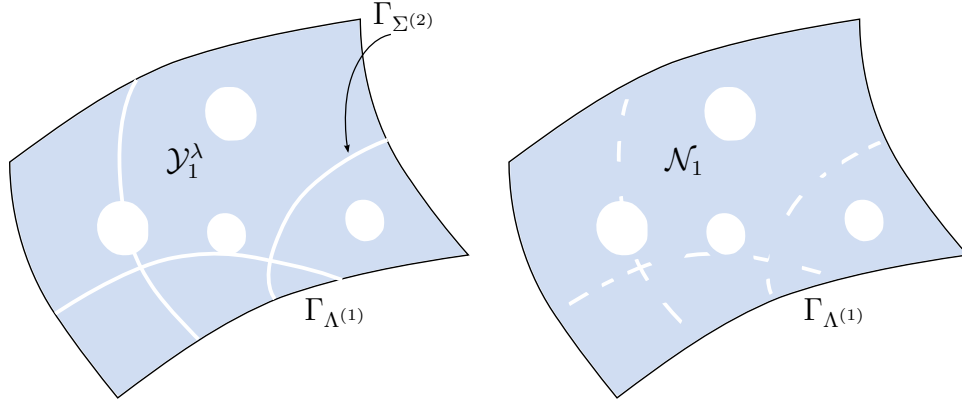


Figure 3: The final nondegeneracy set of step 1: removing all essential resonances results in a disconnected set, but removing only the degenerate part does not destroy connectivity.

- a collection of essential rank two lattices $\mathcal{E}^{(2)}$,
- a dual collection of codimension one resonant manifolds $\mathcal{P}^{(1)}$,
- the nondegenerate set $\mathcal{N}_1 \subset \mathcal{P}^{(1)}$ and is $(1 - 4^{-1}\gamma)$ -dense in B^n .

By step 1, for each essential resonance $\Sigma^{(2)} \supset \mathcal{B}_e^{(2)}$ and $p \in \overline{\Gamma_{\Sigma^{(2)}}} \cap \mathcal{N}_1$ the pair $(H_{p, \mathcal{B}_e^{(2)}}^s, h(\mathcal{B}^{(1)}|\mathcal{B}_e^{(2)}))$ is nondegenerate. Therefore, Property A3 applies with

$$m = 2, \mathcal{B}^{\text{st}} = \mathcal{B}_e^{(2)}, U_0^{\text{st}} = U_{p_0, \mathcal{B}_e^{(2)}}, h = h(\mathcal{B}^{(1)}|\mathcal{B}_e^{(2)}).$$

Moreover, we can choose a *uniform* constant $N_2 = N_2(\mathcal{E}^{(2)}, \mathcal{N}_1)$ over all $\Sigma^{(2)} \in \mathcal{E}^{(2)}$, $\mathcal{B}^{(1)} \subset \Sigma^{(2)}$ and $p \in \overline{\Gamma_{\Sigma^{(2)}}} \cap \mathcal{N}_1$ such that the conclusion of Property A3 hold.

We are now ready to define the set $\mathcal{L}^{(2)}$. We say the pair (k_1, k_2) is *admissible* and $\Lambda^{(2)} = \text{span}_{\mathbb{R}}(k_1, k_2) \cap \mathbb{Z}^{n+1} \in \mathcal{L}^{(2)}$ if the following hold.

1. $\text{span}_{\mathbb{R}}(k_1) \cap \mathbb{Z}^{n+1} \in \mathcal{L}^{(1)}$, and $\Lambda^{(2)}$ is an irreducible lattice of rank 2 over \mathbb{Z} .
2. k_2 cannot be generated by any of the previous generation essential resonances.

$$k_2 \notin \bigvee_{\Sigma^{(2)} \in \mathcal{E}^{(2)}} \Sigma^{(2)},$$

where $\bigvee_{\Sigma^{(2)} \in \mathcal{E}^{(2)}} \Sigma^{(2)}$ is the minimal lattice containing all lattices $\Sigma^{(2)} \in \mathcal{E}^{(2)}$.

3. (ghost property) For each $\Sigma^{(2)} \supset \{k_1\}$, we have

$$M(\Sigma^{(2)} \vee k_2 | \Sigma^{(2)}) > N_2(\mathcal{E}^{(2)}, \mathcal{N}_1).$$

In particular, for \mathcal{B}^{st} being a basis of $\Sigma^{(2)}$, $\mathcal{B}^{\text{wk}} = [k_2]$, and $p \in \overline{\Gamma_{\Sigma^{(2)}}} \cap \mathcal{N}_1$ the conclusion of Property A3 hold⁵.

⁵The name ‘‘ghost’’ comes from the fact that we test k_2 against all possible essential lattices $\Sigma^{(2)}$

The lattice $\Lambda^{(2)} = \text{span}_{\mathbb{R}}(k_1, k_2) \cap \mathbb{Z}^{n+1}$ associated to an admissible pair is also called *admissible*. An ordered basis $\mathcal{B}^{(2)} = [k_1, k_2]$ associated to an admissible $\Lambda^{(2)}$ is called *admissible* too. Denote by $\mathcal{K}^{(2)}$ the union of these ordered bases.

Lemma A.5. *There exists an collection of rank two admissible lattices $\mathcal{L}^{(2)}$ such that*

$$\mathcal{P}^{(2)} = \bigcup_{\Lambda^{(2)} \in \mathcal{L}^{(2)}} \Gamma_{\Lambda^{(2)}} \cap \mathcal{N}_1$$

is connected, $4^{-2}\gamma$ -dense in \mathcal{N}_1 and $(1 - 2 \cdot 4^{-2})\gamma$ dense in B^n .

Similar to step 1, we define the non-degeneracy set $\mathcal{Y}_2^\lambda(H_1, \mathcal{P}^{(2)}) \subset \mathcal{P}^{(2)}$ by the following condition: For any lattice $\Lambda^{(2)} \in \mathcal{L}^{(2)}$, with basis $\mathcal{B}^{(2)}$, and $p \in \Gamma_{\Lambda^{(2)}} \cap \mathcal{P}^{(2)}$, the averaged potential $U_{p, \mathcal{B}^{(2)}}$ has at most two λ -nondegenerate minima.

Lemma A.6. *There exists an open set $\mathcal{U}_2 \subset \mathcal{U}_1$ and $\lambda_2 > 0$, such that for $H_1 \in \mathcal{U}_2$, the nondegeneracy set $\mathcal{Y}_2^{\lambda_2}(H_1, \mathcal{P}^{(1)})$ is open, connected, $4^{-2}\gamma$ -dense in $\mathcal{P}^{(2)}$, and $(1 - 2 \cdot 4^{-2})\gamma$ -dense in B^n .*

Using compactness, we obtain that for

$$m = 2, \quad \mathcal{B}^{\text{st}} = \mathcal{B}^{(2)} \in \mathcal{K}^{(2)}, \quad p_0 \in \Gamma_{\mathcal{B}^{(2)}} \cap \mathcal{Y}_2^{\lambda_2}, \quad U^{\text{st}} = U_{p_0, \mathcal{B}^{(2)}},$$

there exists $M_2(\mathcal{P}^{(2)}, \mathcal{Y}_2^{\lambda_2}) > 0$, such that the conclusion of Property A3 applies for all $\mu(\mathcal{B}^{\text{wk}}) > M_2$.

We now define essential lattices. It suffices to define bases of these lattices. Let

$$\mathcal{E}^{(3)} := \{\mathcal{B}^{(2)} \vee \{k'\} : \mathcal{B}^{(2)} \in \mathcal{K}^{(2)}, |k'| \leq M_2(\mathcal{P}^{(1)}, \mathcal{Y}_1^{\lambda_1}) + \max_{k_1, k_2 \in \mathcal{B}^{(1)}} (|k_1| + |k_2|)\}.$$

At this state, essential resonances comes with a hierarchical structure (see Figure 5).

- *Type (3, 1):* Let $\Sigma^{(3)} \in \mathcal{E}_1^{(3)}$ be an essential resonance that contains $\mathcal{B}^{(2)} = [k_1, k_2]$. We say $\Sigma^{(3)}$ is of type (3, 1) if there exists $\Sigma^{(2)} \in \mathcal{E}^{(2)}$ such that $\Sigma^{(2)} \subset \Sigma^{(3)}$ and $\mathcal{B}^{(2)}$ is basis of $\Sigma^{(2)}$.

This element $\Sigma^{(2)}$ is unique and must contain k_1 , otherwise $\Sigma^{(3)}$ must contain three independent vectors from $\mathcal{E}^{(2)}$, leading to a contradiction with item 2 in the definition of $\mathcal{L}^{(2)}$. Then by item 3 in the definition of $\mathcal{L}^{(2)}$, we have

$$M(\Sigma^{(2)} \vee k_2 | \Sigma^{(2)}) = M(\Sigma^{(3)} | \Sigma^{(2)}) > N_2(\mathcal{E}^{(2)}, \mathcal{N}_1),$$

where $\Sigma^{(2)} \vee k_2$ is the minimal lattice containing both $\Sigma^{(2)}$ and k_2 . Recall that $\Sigma^{(2)}$ comes with fixed basis $\mathcal{B}_e^{(2)}$. We use Proposition 3 to extend this basis to an adapted basis $\mathcal{B}_e^{(3)}$ of $\Sigma^{(3)}$. We take this basis as the fixed basis of $\Sigma^{(3)}$. For each $p \in \Gamma_{\Sigma^{(3)}}$, Property A3 applies. We say that

$$H_{p, \mathcal{B}_e^{(2)}} \text{ dominates } H_{p, \mathcal{B}_e^{(3)}}.$$

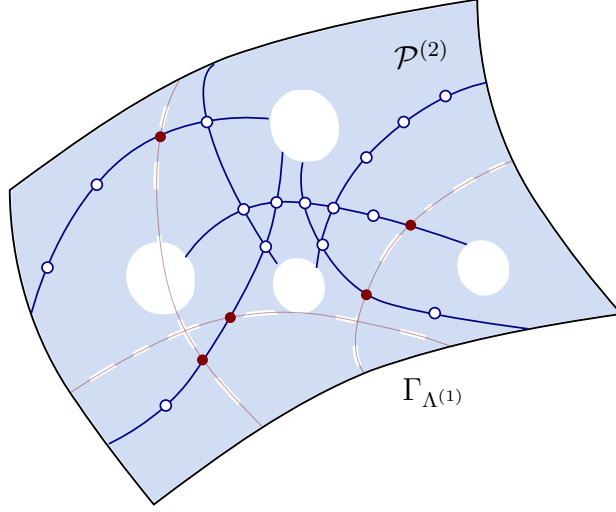


Figure 4: Hierarchy of essential resonances: faint red curves are the previous generation essential resonances, blues line are current generation diffusion path, solid red dots are of type (3, 1), hollow blue dots are of type (3, 2).

- *Type (3, 2)*: Suppose $\Sigma^{(3)} \supset \{k_1, k_2\}$, but the lattice they generate does not contain any element in $\mathcal{E}^{(2)}$. In this case, by definition, $\Sigma^{(3)} \in \mathcal{E}_2^{(3)}$ and we have

$$M(\Sigma^{(3)} | \Lambda_{k_1}) > M_1(\mathcal{P}^{(1)}, \mathcal{Y}_1^{\lambda_1}).$$

We use Proposition 3 to extend k_1 to an adapted basis $\mathcal{B}_e^{(3)}$ of $\Sigma^{(3)}$, taken as the fixed basis for $\Sigma^{(3)}$. Property A2 applies, and we say that

$$H_{p, \mathcal{B}^{(1)}} \text{ dominates } H_{p, \mathcal{B}_e^{(3)}}.$$

We define now the set of essential lattices as the union of type (3, 1) and (3, 2) resonances:

$$\mathcal{E}^{(3)} := \mathcal{E}_1^{(3)} \cup \mathcal{E}_2^{(3)}.$$

For each $\Sigma^{(3)} \in \mathcal{E}^{(3)}$, we define the nondegeneracy set $\mathcal{Z}_2(H_1, \Sigma^{(3)}, \mathcal{Y}_2^{\lambda_2}) \subset \Gamma_{\Sigma^{(3)}} \cap \mathcal{Y}_2^{\lambda_2}$ to be the subset such that for each $\mathcal{B}^{(2)} = [k_1, k_2] \subset \Sigma^{(3)}$, the pair

$$\left(H_{p, \mathcal{B}_e^{(3)}}^s, h(\mathcal{B}^{(2)} | \mathcal{B}_e^{(3)}) \right)$$

is nondegenerate. We then define $\mathcal{Z}_2(H_1, \mathcal{E}^{(3)}, \mathcal{Y}_2^{\lambda_2})$ to be the union of all $\mathcal{Z}_2(H_1, \Sigma^{(3)}, \mathcal{Y}_2^{\lambda_2})$ over essential resonances $\Sigma^{(3)} \in \mathcal{E}^{(3)}$.

Lemma A.7. *There exists a relative open set $\mathcal{U}'_2 \subset \mathcal{U}_2$ and a relative open set $\tilde{\mathcal{Z}}_2$ such that the following hold.*

1. *For each $H_1 \in \mathcal{U}'_2$, $\tilde{\mathcal{Z}}_2$ is compactly contained in $\mathcal{Z}_2(H_1, \mathcal{E}^{(3)}, \mathcal{Y}_2^{\lambda_2})$.*

2. The subset

$$\mathcal{N}_2 := \mathcal{Y}_2^{\lambda_2} \cap \tilde{\mathcal{Z}}_2$$

is open, connected, $4^{-2}\gamma$ -dense in $\mathcal{P}^{(2)}$ and $(1 - 2 \cdot 4^{-2})\gamma$ -dense in B^n .

A.4.3 Step $s + 1$ of the induction

We completed step s with

- the collection of rank s lattices $\mathcal{L}^{(s)}$, with associated bases $\mathcal{B}^{(s)}$,
- a collection of essential rank $s+1$ lattices $\mathcal{E}^{(s+1)}$,
- a dual collection of codimension one resonant manifolds $\mathcal{P}^{(s)}$,
- the nondegenerate set $\mathcal{N}_s \subset \mathcal{P}^{(s)}$, which is $(1 - 4^{-s}\gamma)$ -dense in B^n .

Nonessential resonances.

- We have the collection of lattices $\mathcal{L}^{(1)}, \dots, \mathcal{L}^{(s)}$, with $\mathcal{L}^{(j)} = \{\Lambda_i^{(j)}\}$ are irreducible rank j lattices. For each $\Lambda^{(j)} \in \mathcal{L}^{(j)}$, there exists a *unique* $\Lambda^{(j-1)} \subset \Lambda^{(j)}$ and such that $\Lambda^{(j-1)} \in \mathcal{L}^{(j-1)}$.
- Each $\Lambda^{(s)} \in \mathcal{L}^{(s)}$ has a *ordered* basis defined in the following way. For each $\Lambda^{(1)}$ we fix a basis $\mathcal{B}^{(1)} = \{k_1\}$ which is unique up to a sign. From the previous property, $\Lambda^{(s)}$ comes with the chain of inclusion

$$\Lambda^{(1)} \subset \dots \subset \Lambda^{(s)}, \quad \Lambda^{(j)} \in \mathcal{L}^{(j)}, \quad 1 \leq j \leq s,$$

and we extend the basis $\mathcal{B}^{(1)}$ of $\Lambda^{(1)}$ an increasing set of bases $\mathcal{B}^{(1)} \subset \dots \subset \mathcal{B}^{(s)}$ by consecutive application of Proposition 3.1.

- We use $\mathcal{K}^{(s)}$ to denote the collection of standard bases. For each $\mathcal{B}^{(s)} = [k_1, \dots, k_s]$, we denote $|\mathcal{B}^{(s)}| = \sup_i |k_i|$.
- The diffusion path at step s is

$$\mathcal{P}^{(s)} = \bigcup_{\Lambda^{(s)} \in \mathcal{L}^{(s)}} \Gamma_{\Lambda^{(s)}}.$$

The set $\mathcal{P}^{(s)} \cap B^n$ is connected and $(1 - 2 \cdot 4^{-s})\gamma$ -dense in B^n .

Essential resonances.

- We have the essential lattices $\mathcal{E}^{(2)}, \dots, \mathcal{E}^{(s+1)}$, where for each $1 \leq j \leq s$, $\Sigma^{(j+1)} \in \mathcal{E}^{(j+1)}$ is a rank $j + 1$ irreducible lattice. For each $\Sigma^{(j+1)}$, there exists at least one, and at most two element $\Lambda^{(j)} \in \mathcal{L}^{(j)}$, such that $\Lambda^{(j)} \subset \Sigma^{(j+1)}$.

- If essential lattice $\Sigma^{(s+1)} \in \mathcal{E}^{(s+1)}$ contains only one element $\Sigma^{(s)} \in \mathcal{E}^{(s)}$, then there exists $1 \leq j \leq s$, such that

$$\Sigma^{(j+1)} \subset \dots \subset \Sigma^{(s+1)}, \quad \Sigma^{(t)} \in \mathcal{E}^{(t)}, j+1 \leq t \leq s+1$$

is the longest chain of essential lattices, meaning $\Sigma^{(j+1)}$ does not contain any element of $\mathcal{E}^{(j)}$. We then have the following inclusion

$$\Lambda^{(1)} \subset \dots \subset \Lambda^{(j-1)} \subset \Sigma^{(j+1)} \subset \dots \subset \Sigma^{(s+1)}.$$

We use Proposition 3.1 to obtain the chain of adapted bases (called *ordered basis*):

$$\mathcal{B}^{(1)} \subset \dots \subset \mathcal{B}^{(j-1)} \subset \mathcal{B}_e^{(j+1)} \subset \dots \subset \mathcal{B}_e^{(s+1)},$$

where each $\mathcal{B}^{(t)}$ is a basis of $\Lambda^{(t)} \in \mathcal{L}^{(t)}$ and each $\mathcal{B}_e^{(t)}$ is a basis of $\Sigma^{(t)} \in \mathcal{E}^{(t)}$. Recording the increment of rank in the chain, the essential resonance $\Sigma^{(s+1)}$ is called of type $(s+1, j)$. Denote by $\mathcal{E}_j^{(s+1)}$ the set of essential resonances with this property.

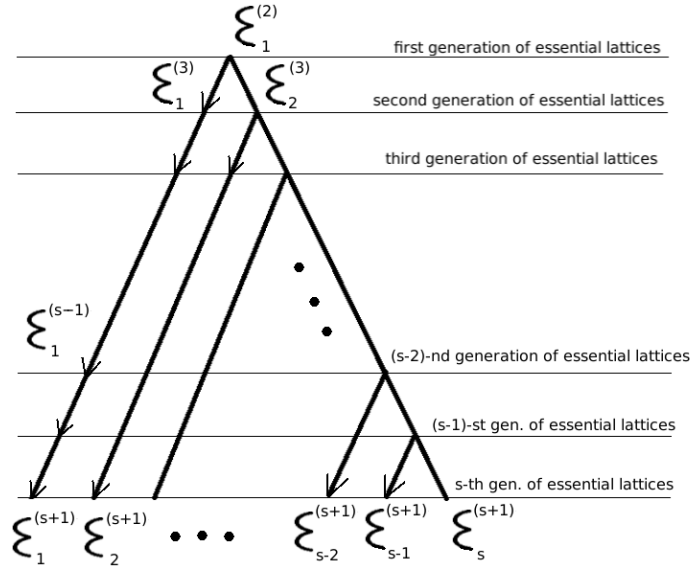


Figure 5: Essential lattices

Strong system and nondegeneracy.

- For each $1 \leq j \leq s$, there exists the nondegeneracy set $\mathcal{N}_j \subset \mathcal{P}^{(j)}$, with the property that each \mathcal{N}_j is relative open, connected and $(1 - 4^{-j}\gamma)$ -dense in $\mathcal{P}^{(j)} \cap B^n$ and $(1 - 2 \cdot 4^{-j})\gamma$ -dense in B^n . The following inclusion hold

$$\mathcal{P}^{(1)} \supset \mathcal{N}_1 \supset \dots \supset \mathcal{P}^{(s)} \supset \mathcal{N}_s.$$

- There exists a sequence of (nonempty) relative open sets

$$\{\|H_1\|_{C^r} = 1\} \supset \mathcal{U}_0 \supset \mathcal{U}_1 \supset \mathcal{U}'_1 \supset \cdots \supset \mathcal{U}_s \supset \mathcal{U}'_s$$

- There exists $\lambda_s > 0$ such that for each $H_1 \in \mathcal{U}_s$, $\Lambda^{(s)} \in \mathcal{L}^{(s)}$ with basis $\mathcal{B}^{(s)}$, and $p \in \overline{\mathcal{N}_s \cap \Gamma_{\Lambda^{(s)}}}$, the strong system $H_{p, \mathcal{B}^{(s)}}$ is nondegenerate in the sense of Property A3 and the averaged potential $U_{p, \mathcal{B}^{(s)}}$ has at most two λ -nondegenerate minima. Using compactness, let

$$M_s(\mathcal{P}^{(s)}, \mathcal{N}_s) > \sup_{\mathcal{B}^{(s)} \in \mathcal{K}^{(s)}} |\mathcal{B}^{(s)}|$$

be a uniform constant such that Property A3 applies.

- For each $H_1 \in \mathcal{U}'_s$, $\Sigma^{(s+1)} \in \mathcal{E}^{(s+1)}$ with basis $\mathcal{B}_e^{(s+1)}$, each $\Lambda^{(s)} \subset \Sigma^{(s+1)}$ with basis $\mathcal{B}^{(s)}$, and $p \in \overline{\mathcal{N}_s \cap \Gamma_{\Sigma^{(s+1)}}}$, the pair

$$\left(H_{p, \mathcal{B}_e^{(s+1)}}, h(\mathcal{B}^{(s)}, \mathcal{B}_e^{(s+1)}) \right)$$

is nondegenerate in the sense of Property A3. Using compactness, let

$$N_s(\mathcal{E}^{(s+1)}, \mathcal{N}_s) > 0$$

be a uniform constant such that Property A3 applies.

Domination Properties

- Let $\Sigma^{(s+1)} \in \mathcal{E}_j^{(s+1)}$ be an essential resonance of type $(s+1, j)$, then we have the chain

$$\Lambda^{(1)} \subset \cdots \subset \Lambda^{(j-1)} \subset \Sigma^{(j+1)} \subset \cdots \subset \Sigma^{(s+1)}.$$

The following domination property holds:

$$M(\Sigma^{(j+1)} | \Lambda^{(j-1)}) > M_{j-1}(\mathcal{P}^{(j-1)}, \mathcal{N}_{j-1}),$$

$$M(\Sigma^{(t+1)} | \Sigma^{(t)}) > N_{t-1}(\mathcal{E}^{(t)}, \mathcal{N}_{t-1}), \quad j+1 \leq t \leq s.$$

- As a corollary of the domination properties, for $\Sigma^{(s+1)}$ with the type $(s+1, j)$, let

$$\mathcal{B}^{(1)} \subset \cdots \subset \mathcal{B}^{(j-1)} \subset \mathcal{B}_e^{(j+1)} \subset \cdots \subset \mathcal{B}_e^{(s+1)}$$

be the chain of basis. Then

- For each $p \in \overline{\Gamma_{\Lambda^{(j-1)}} \cap \mathcal{N}_{j-1}}$, the system $H_{p, \mathcal{B}^{(j-1)}}$ dominates $H_{p, \mathcal{B}_e^{(j+1)}}$ in the sense of Property A2.
- For each $j+1 \leq t \leq s$, $p \in \overline{\Gamma_{\Sigma^{(t)}} \cap \mathcal{N}_{t-1}}$, the system $H_{p, \mathcal{B}_e^{(t)}}$ dominates $H_{p, \mathcal{B}_e^{(t+1)}}$ in the sense of Property A3.

We now define the set $\mathcal{L}^{(s+1)}$. This is essentially an elaboration of step 2. We say the rank $s + 1$ lattice $\Lambda^{(s+1)}$ is *admissible* if the following hold.

1. There exists $\Lambda^{(s)} \in \mathcal{L}^{(s)}$ such that $\Lambda^{(s)} \subset \Lambda^{(s+1)}$.
2. $\Lambda^{(s+1)}$ cannot be generated by any previous generation essential resonances, namely

$$\Lambda^{(s+1)} \not\subset \bigvee \{\Sigma^{(s+1)} \in \mathcal{E}^{(s+1)}\}, \quad (\text{A.2})$$

where \bigvee is the smallest irreducible lattice that contains all lattices $\Sigma^{(s+1)} \in \mathcal{E}^{(s+1)}$.

3. Item 2 ensures that $\Lambda^{(s)} \subset \Lambda^{(s+1)}$ is unique. Otherwise, suppose we have $\Lambda_1^{(s)}, \Lambda_2^{(s)} \subset \Lambda^{(s+1)}$ with bases $\mathcal{B}_1^{(s)}, \mathcal{B}_2^{(s)}$, then $\Lambda^{(s+1)} = \Lambda_1^{(s)} \vee \Lambda_2^{(s)}$, and

$$M(\Lambda^{(s+1)} | \Lambda_1^{(s)}) \leq \max\{|\mathcal{B}_1^{(s)}|, |\mathcal{B}_2^{(s)}|\} \leq M_s(\mathcal{P}^{(s)}, \mathcal{N}_s),$$

hence $\Lambda^{(s+1)} \in \mathcal{E}^{(s+1)}$, which is a violation of item 2.

4. (ghost property) For each $\Lambda^{(s)} \subset \Lambda^{(s+1)}$ and $\Lambda^{(s)} \subset \Sigma^{(s+1)}$, we have

$$M(\Sigma^{(s+1)} \vee \Lambda^{(s+1)} | \Sigma^{(s+1)}) > N_s(\mathcal{E}^{(s+1)}, \mathcal{N}_s). \quad (\text{A.3})$$

Lemma A.8. *There exists an collection $\mathcal{L}^{(s+1)} = \{\Lambda^{(s+1)}\}$ of admissible pairs such that*

$$\mathcal{P}^{(s+1)} = \bigcup_{\Lambda^{(s+1)} \in \mathcal{L}^{(s+1)}} \Gamma_{\Lambda^{(s+1)}} \cap \mathcal{N}_s$$

is connected, $4^{-s-1}\gamma$ -dense in \mathcal{N}_s , and $(1 - 2 \cdot 4^{-s-1})\gamma$ dense in B^n .

The set $\mathcal{Y}_{s+1}^\lambda(H_1, \mathcal{P}^{(s+1)}) \subset \mathcal{P}^{(s+1)}$ is defined by the following condition: For any $\Lambda^{(s+1)} \in \mathcal{L}^{(s+1)}$ with basis $\mathcal{B}^{(s+1)}$, and $p \in \Gamma_{\mathcal{B}^{(s+1)}} \cap \mathcal{P}^{(s+1)}$, there exists $0 < \lambda < \lambda'$ such that the averaged potential $U_{p, \mathcal{B}^{(s+1)}}$ has at most two λ' -nondegenerate minima.

Lemma A.9. *There exists an open set $\mathcal{U}_{s+1} \subset \mathcal{U}'_s$ and $\lambda_{s+1} > 0$, such that for $H_1 \in \mathcal{U}_{s+1}$, the nondegeneracy set $\mathcal{Y}_{s+1}^{\lambda_{s+1}}(H_1, \mathcal{P}^{(s+1)})$ is $4^{-s-1}\gamma$ -dense in $\mathcal{P}^{(s+1)}$ and connected.*

Define

$$M_{s+1}(\mathcal{P}^{(s+1)}, \mathcal{Y}_{s+1}^{\lambda_{s+1}}) > \sup_{\mathcal{B}^{(s+1)} \in \mathcal{K}^{(s+1)}} |\mathcal{B}^{(s+1)}|$$

be the uniform constant over all $H_1 \in \mathcal{U}_{s+1}$, $p \in \mathcal{Y}_{s+1}^{\lambda_{s+1}}$, and $\Lambda^{(s+1)} \in \mathcal{L}^{(s+1)}$. The essential lattice set $\mathcal{E}^{(s+2)}$ is defined as the set of all rank $s + 2$ irreducible lattices $\Sigma^{(s+2)}$ satisfying the following conditions: there exists $\Lambda^{(s+1)} \in \mathcal{L}^{(s+1)}$ such that

$$\Sigma^{(s+2)} \supset \Lambda^{(s+1)}, \text{ and } M(\Lambda^{(s+1)} | \Sigma^{(s+2)}) < M_{s+1}(\mathcal{P}^{(s+1)}, \mathcal{Y}_{s+1}^{\lambda_{s+1}}).$$

We have the following remarks:

- Suppose there exists $\Sigma^{(s+2)} \supset \Sigma^{(s+1)}$ with $\Sigma^{(s+1)} \in \mathcal{E}^{(s+1)}$, then $\Sigma^{(s+1)}$ is unique. Otherwise, suppose $\Sigma^{(s+2)}$ contains both $\Sigma_1^{(s+1)}, \Sigma_2^{(s+1)}$, then there exists $\Lambda^{(s+1)} \subset \Sigma^{(s+2)} = \Sigma_1^{(s+1)} \vee \Sigma_2^{(s+1)}$, this is a violation of (A.2).

- In case that $\Sigma^{(s+2)} \supset \Sigma^{(s+1)}$, then for $\Lambda^{(s+1)} \in \mathcal{L}^{(s+1)}$ with $\Sigma^{(s+2)} \supset \mathcal{L}^{(s+1)}$, we get

$$M(\Sigma^{(s+1)} | \Sigma^{(s+2)}) = M(\Sigma^{(s+1)} | \Sigma^{(s+1)} \vee \Lambda^{(s+1)}) > N_s(\mathcal{E}^{(s+1)}, \mathcal{N}_s)$$

by (A.3).

Finally, for each $\Sigma^{(s+2)} \subset \mathcal{E}^{(s+2)}$, we define the nondegeneracy set $\mathcal{Z}_{s+1}(H_1, \Sigma^{(s+2)}, \mathcal{Y}_{s+1}^{\lambda_{s+1}})$ to be the subset that for each $\mathcal{B}^{(s+1)} = (k_1, \dots, k_{s+1}) \subset \Sigma^{(s+2)}$, the pair

$$\left(H_{p, \mathcal{B}_e}^s, h(\mathcal{B}^{(s+1)} | \mathcal{B}_e^{(s+2)}) \right)$$

is nondegenerate. We then define $\tilde{\mathcal{Z}}_{s+1}(H_1, \mathcal{E}^{(s+2)}, \mathcal{Y}_{s+1}^{\lambda_{s+1}})$ to be the union over essential resonances $\Sigma^{(s+2)} \in \mathcal{E}^{(s+2)}$.

Lemma A.10. *Suppose $s + 2 < n$. There exists an open set $\mathcal{U}'_{s+1} \subset \mathcal{U}_{s+1}$ and a relative open set $\tilde{\mathcal{Z}}_{s+1}$ such that the following hold.*

1. For each $H_1 \in \mathcal{U}'_{s+1}$, $\tilde{\mathcal{Z}}_{s+1}$ is compactly contained in $\mathcal{Z}_{s+1}(H_1, \mathcal{E}^{(s+2)}, \mathcal{Y}_{s+1}^{\lambda_{s+1}})$.

2. The subset

$$\mathcal{N}_{s+1} := \mathcal{Y}_{s+1}^{\lambda_{s+1}} \cap \tilde{\mathcal{Z}}_{s+1}$$

is open, connected, $4^{-s-1}\gamma$ -dense in $\mathcal{P}^{(2)}$ and $(1 - 4^{-s-1})\gamma$ -dense in B^n .

Moreover, if $s + 2 = n$,

$$\bigcup_{\Sigma^{(s+2)} \in \mathcal{E}^{(s+2)}} \Gamma_{\Sigma^{(s+2)}} \cap \mathcal{Y}_{s+1}^{\lambda_{s+1}}$$

is a collection of isolated points. Then the same two points hold with

$$\tilde{\mathcal{Z}}_{s+1} = \bigcup_{\Sigma^{(s+2)} \in \mathcal{E}^{(s+2)}} \Gamma_{\Sigma^{(s+2)}} \cap \mathcal{Y}_{s+1}^{\lambda_{s+1}}.$$

This finishes the construction of the lattices and verification of properties for step $s + 1$.

A.4.4 Concluding the induction

The induction ends when \mathcal{U}'_{n-1} , $\mathcal{L}^{(n-1)}$, $\mathcal{P}^{(n-1)}$, \mathcal{N}_{n-1} and $\mathcal{E}^{(n)}$ are defined. Then $\mathcal{P}^{(n-1)} \cap \mathcal{N}_{n-1}$ is γ -dense diffusion path in \mathcal{B}^n , and for each $p \in \Gamma_{\Lambda^{(n-1)}} \cap \mathcal{N}_{n-1}$, $H_1 \in \mathcal{U}'_{n-1}$, the potential $U_{p, \mathcal{B}^{(n-1)}}$ has at most two λ_{n-1} -nondegenerate minima.

We then have

Lemma A.11. *There exists an open and dense set of $\mathcal{U}''_{n-1} \subset \mathcal{U}'_{n-1}$ such that $[H1\lambda_{n-1}]$ holds for all $H_1 \in \mathcal{U}''_{n-1}$ on $\mathcal{P}^{(n-1)} \cap \mathcal{N}_{n-1}$.*

Moreover, from Lemma A.10 we know that condition [H2] holds on all essential resonances. Therefore, the diffusion path $\mathcal{P}^{(n-1)} \cap \mathcal{N}_{n-1}$ satisfies all the conditions required.

B Useful facts from linear algebra

Lemma B.1. *Given $1 \leq s \leq n + 1$, let $P = [k_1 \ \cdots \ k_s]$ be an integer matrix with linearly independent columns. Then there exists $c_n > 1$ depending only on n such that*

$$\min_{\|v\|=1} \|Pv\| = \min_{\|v\|=1} (v^T P^T P v)^{\frac{1}{2}} = \|(P^T P)^{-1}\|^{-\frac{1}{2}} \geq c_n^{-1} |k_1|^{-1} \cdots |k_m|^{-1}.$$

In particular, if $s = n + 1$, then $\|P^{-1}\| = \|(P^T)^{-1}\| \leq c_n |k_1| \cdots |k_{n+1}|$.

Proof. We only estimate $\|(P^T P)^{-1}\|$. Let $a_{ij} = (P^T P)_{ij}$ and $B_{ij} = (P^T P)^{-1}_{ij}$, then using Cramer's rule and the definition of the cofactor, we have

$$|B_{ij}| \leq \frac{1}{\det(P^T P)} \sum_{\sigma} \prod_{s \neq i} a_{s\sigma(s)},$$

where σ ranges over all one-to-one mappings from $\{1, \dots, m\} \setminus \{i\}$ to $\{1, \dots, m\} \setminus \{j\}$. Since P is a nonsingular integer matrix, we have $\det(P^T P) \geq 1$. Moreover, $a_{ij} = k_i^T k_j \leq n |k_i| |k_j|$. Therefore

$$|B_{ij}| \leq \sum_{\sigma} \prod_{s \neq i} |k_s| |k_{\sigma(s)}| \leq c_n \left(\prod_{s \neq i} |k_s| \right) \left(\prod_{s \neq j} |k_s| \right),$$

where c_n is a constant depending only on n . Using the fact that the norm of a matrix is bounded by its largest entry, up to a factor depending only on dimension, by changing to a different c_n , we have

$$\|(P^T P)^{-1}\| \leq c_n \sup_{i,j} |B_{ij}| \leq c_n \sup_{i,j} \left(\prod_{s \neq i} |k_s| \right) \left(\prod_{s \neq j} |k_s| \right) \leq c_n \left(\prod_{s=1}^m |k_s| \right)^2.$$

The lemma follows from the estimates obtained, by possibly changing c_n again.

If $s = n + 1$, then $\|P^{-1}\| = \|(P^T P)^{-1}\|^{\frac{1}{2}} = \|(P P^T)^{-1}\|^{\frac{1}{2}} = \|(P^T)^{-1}\|$. \square

Lemma B.2. *Given $1 \leq s \leq n$, $D > 1$, let $\bar{k}_1, \dots, \bar{k}_s \in \mathbb{Z}^n$ be linearly independent vectors, and $Q_0 \in \text{Sym}(n)$ satisfying $Q_0 \geq D^{-1} \text{Id}$. There exists $c_n > 1$ depending only on n such that for $\bar{P} = [\bar{k}_1 \ \cdots \ \bar{k}_s]$, we have*

$$\min_{\|v\|=1} v^T \bar{P}^T Q_0 \bar{P} v \geq c_n^{-1} D^{-1} |\bar{k}_1|^{-2} \cdots |\bar{k}_s|^{-2}.$$

Proof. Since

$$v^T \bar{P}^T Q_0 \bar{P} v \geq D^{-1} \|\bar{P} v\|^2 \geq D^{-1} \left(\min_{\|w\|=1} \|\bar{P} w\| \right)^2 \|v\|,$$

the statement follows directly from Lemma B.1. \square

Acknowledgement The first author acknowledges NSF for partial support grant DMS-5237860. The authors would like to thank John Mather, Marcel Guardia, and Abed Bounemoura for useful conversations.

References

- [AKN06] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical aspects of classical and celestial mechanics*. Third. Vol. 3. Encyclopaedia of Mathematical Sciences. [Dynamical systems. III], Translated from the Russian original by E. Khukhro. Springer-Verlag, Berlin, 2006, pp. xiv+518.
- [Arn63] V. I. Arnold. “Small denominators and problems of stability of motion in classical and celestial mechanics”. In: *Uspehi Mat. Nauk* 18.6 (114) (1963), pp. 91–192.
- [Arn64] V. I. Arnold. “Instability of dynamical systems with many degrees of freedom”. In: *Dokl. Akad. Nauk SSSR* 156 (1964), pp. 9–12.
- [Arn68] V. I. Arnold. “A stability problem and ergodic properties of classical dynamical systems”. In: *Proc. Internat. Congr. Math. (Moscow, 1966)*. Izdat. “Mir”, Moscow, 1968, pp. 387–392.
- [Arn94] V. I. Arnold. “Mathematical problems in classical physics”. In: *Trends and perspectives in applied mathematics*. Vol. 100. Appl. Math. Sci. Springer, New York, 1994, pp. 1–20.
- [Ber08] P. Bernard. “Young measures, superposition and transport”. In: *Indiana Univ. Math. J.* 57.1 (2008), pp. 247–275.
- [Ber10] P. Bernard. “On the Conley decomposition of Mather sets”. In: *Rev. Mat. Iberoam.* 26.1 (2010), pp. 115–132.
- [BKZ11] P. Bernard, V. Kaloshin, and K. Zhang. “Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders”. In: *preprint arXiv:1112.2773v2* (2011).
- [Car95] M. J. D. Carneiro. “On minimizing measures of the action of autonomous Lagrangians”. In: *Nonlinearity* 8.6 (1995), pp. 1077–1085.
- [Che13] C.-Q. Cheng. “Arnold diffusion in nearly integrable systems”. In: *preprint* (2013).
- [Con88] C. Conley. “The gradient structure of a flow. I”. In: *Ergodic Theory Dynam. Systems* 8*. Charles Conley Memorial Issue (1988). With a comment by R. Moeckel, pp. 11–26, 9.
- [CY09] C.-Q. Cheng and J. Yan. “Arnold diffusion in Hamiltonian systems: a priori unstable case”. In: *Journal of Differential Geometry* 82, no. 2 (2009), pp. 229–277.
- [DLS13] A. Delshams, R. de la Llave, and T. Seara. “Instability of high dimensional Hamiltonian Systems: Multiple resonances do not impede diffusion”. In: *preprint arXiv:1306.4614 [math.DS]* (2013).
- [Fat08] A. Fathi. *Weak KAM theorem in Lagrangian dynamics, 10th preliminary version*. book preprint, 2008.

- [GK14a] M. Guardia and V. Kaloshin. “Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation”. In: *Journal of the European Mathematical Society* 16.12 (2014).
- [GK14b] M. Guardia and V. Kaloshin. “Orbits of nearly integrable systems accumulating to KAM tori”. In: *preprint* (2014), pp. 1–112.
- [KLS14] V. Kaloshin, M. Levi, and M. Saprykina. “Arnol’d diffusion in a pendulum lattice”. In: *Comm. Pure Appl. Math.* 67.5 (2014), pp. 748–775.
- [KMV04] V. Kaloshin, J. Mather, and E. Valdinoci. “Instability of totally elliptic points of symplectic maps in dimension 4”. In: *Asterisque* 74 (116) (2004), pp. 79–116.
- [KZ13] V Kaloshin and K Zhang. “A strong form of Arnold diffusion for two and half degrees of freedom”. In: *preprint* arXiv:1212.1150 (2013).
- [KZ14] V Kaloshin and K Zhang. “Arnold diffusion for three and half degrees of freedom”. In: *preprint* <http://www2.math.umd.edu/vkaloshi/papers/announce-three-and-half.pdf> (2014).
- [Loc92] P. Lochak. “Canonical perturbation theory via simultaneous approximation”. In: *Russian Mathematical Surveys* 47.6 (1992), pp. 57–133.
- [Mar12a] J.-P. Marco. “Generic hyperbolic properties of classical systems on the torus \mathbb{T}^2 ”. In: *preprint* (2012).
- [Mar12b] J.-P. Marco. “Generic hyperbolic properties of nearly integrable systems on \mathbb{A}^3 ”. In: *preprint* (2012).
- [Mat03] J. N. Mather. “Arnold diffusion. I. Announcement of results”. In: *Sovrem. Mat. Fundam. Napravl.* 2 (2003), 116–130 (electronic).
- [Mat08] J. N. Mather. “Arnold diffusion. II.” In: *preprint* (2008), pp. 1–185.
- [Mat10] J. N. Mather. “Order structure on action minimizing orbits”. In: *Symplectic topology and measure preserving dynamical systems*. Vol. 512. Contemp. Math. Providence, RI: Amer. Math. Soc., 2010, pp. 41–125.
- [Mat11] J. N. Mather. “Shortest curves associated to a degenerate Jacobi metric on \mathbb{T}^2 ”. In: *Progress in variational methods*. Vol. 7. Nankai Ser. Pure Appl. Math. Theoret. Phys. World Sci. Publ., Hackensack, NJ, 2011, pp. 126–168.
- [Mat93] J. N. Mather. “Variational construction of connecting orbits”. In: *Ann. Inst. Fourier (Grenoble)* 43.5 (1993), pp. 1349–1386.
- [Mañ97] R. Mañé. “Lagrangian flows: the dynamics of globally minimizing orbits”. In: *Bol. Soc. Brasil. Mat. (N.S.)* 28.2 (1997), pp. 141–153.
- [Moe96] R. Moeckel. “Transition tori in the five-body problem”. In: *Journal of Differential Equations* 129 (1996), pp. 290–314.

- [Sie89] C. L. Siegel. *Lectures on the geometry of numbers*. Notes by B. Friedman, Rewritten by Komaravolu Chandrasekharan with the assistance of Rudolf Suter, With a preface by Chandrasekharan. Springer-Verlag, Berlin, 1989, pp. x+160.
- [Sor10] A. Sorrentino. “Lecture notes on Mather’s theory for Lagrangian systems”. In: *Online lecture notes* arxiv.org/abs/1011.0590 (2010).
- [Tre04] D. Treschev. “Evolution of slow variables in a priori unstable Hamiltonian systems”. In: *Nonlinearity* 17.5 (2004), pp. 1803–1841.
- [Tre12] D. Treschev. “Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems”. In: *Nonlinearity* 25.9 (2012), pp. 2717–2757.
- [Tur14] D. Turaev. “Hyperbolic sets near homoclinic loops to a saddle for systems with a first integral”. In: *Regular and Chaotic Dynamics* to appear (2014).
- [Zhe10] Y Zheng. “Arnold diffusion for a priori unstable systems and a five-body problem.” In: *preprint* Penn State University (2010).