Bull. Soc. Math. France
146 (3), 2018, p. 517-574

DYNAMICS OF THE DOMINANT HAMILTONIAN

BY VADIM KALOSHIN & KE ZHANG

ABSTRACT. — It is well known that instabilities of nearly integrable Hamiltonian systems occur around resonances. Dynamics near resonances of these systems is well approximated by the associated averaged system, called *slow system*. Each resonance is defined by a basis (a collection of integer vectors). We introduce a class of resonances whose basis can be divided into two well separated groups and call them *dominant*. We prove that the associated slow system can be well approximated by a subsystem given by one of the groups, both in the sense of the vector field and weak KAM theory. As a corollary, we obtain perturbation results on normally hyperbolic invariant cylinders, and the Aubry/Mañe sets. This has applications in Arnold diffusion in arbitrary degrees of freedom.

RÉSUMÉ (*Dynamique de l'hamiltonien dominant*). — Il est bien connu que les instabilités des systèmes hamiltoniens presque intégrables interviennent au voisinage des résonances. La dynamique de ces systèmes près des résonances est bien approchée par les systèmes moyennés associés, appelés *systèmes lents*. Chaque résonance est définie par une base (une collection de vecteurs entiers). Nous introduisons une classe de résonances dont la base peut être divisée en deux groupes bien distincts, que nous appelons *dominantes*. Nous prouvons que le système lent associé peut être bien approché par un sous-système donné par l'un de ces deux groupes, à la fois comme champ de vecteurs et au sens de la théorie KAM faible. Comme corollaire, nous obtenons des résultats perturbatifs sur des cylindres invariants normalement hyperboliques, et sur

Texte reçu le 10 septembre 2016, accepté le 23 janvier 2017.

VADIM KALOSHIN, Department of Mathematics, University of Maryland at College Park, College Park, MD, USA • *E-mail :* `vadim.kaloshin@gmail.com`
KE ZHANG, Department of Mathematics, University of Toronto, Toronto, ON, Canada • *E-mail :* `kzhang@math.utoronto.edu`

Mathematical subject classification (2010). — 37J40, 37J50.

Key words and phrases. — Hamiltonian systems, resonant averaging, Mather theory, weak KAM theory, Arnold diffusion.

doi:10.24033/bsmf.2765

les ensembles d'Aubry/Mañé. Cela a des applications en diffusion d'Arnold pour un nombre arbitraire de degrés de liberté.

1. Introduction

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

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

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

$$D^{-1}\,\mathrm{Id} \leqslant \partial^2_{pp} H_0(p) \leqslant D\,\mathrm{Id}$$

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 [3], and asks ([1, 2, 4]) whether topological instability is "typical" in nearly integrable systems with $n \geqslant 2$ (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, $\mathrm{span}_{\mathbb{R}}\Lambda \cap \mathbb{Z}^{n+1} = \Lambda$.

We now consider the diffusion that occurs along a connected net Γ of $(n-1)$-resonances, which are curves in $\mathbb{R}^n$. The main difficulty in proving Arnold diffusion is in crossing the maximal (n-)resonances, which are intersections of Γ with a transveral 1-resonance manifold $\Gamma_{k'}$. A similar question is whether one can "switch" at the intersection of two resonant curves (see Figure 1.1).

For an n-resonance $\{p_0\} = \Gamma_\Lambda$, we assume that Λ is irreducible, and $\mathcal{B} = [k_1, \dots, k_n]$ is a basis over $\mathbb{Z}$. The averaging theory of H_ϵ near p_0 reduces to

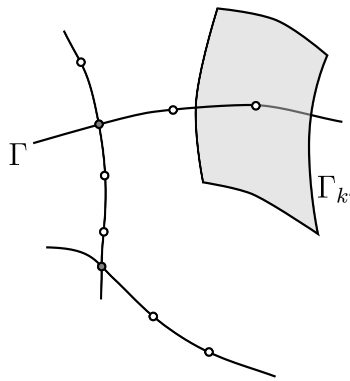

FIGURE 1.1. Diffusion path and essential resonances in $n = 3$. The hollow dots requires crossing, while the gray dots requires switching

the study of a particular *slow system* defined on $\mathbb{T}^n \times \mathbb{R}^n$, denoted $H^s_{p_0,\mathcal{B}}$. More precisely, in an $O(\sqrt{\varepsilon})$-neighborhood of p_0, the system H_ε admits the normal form (see [16], Appendix B)

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

where

$$\varphi_i = k_i \cdot (\theta, t),\ 1 \leqslant i \leqslant n, \quad (p - p_0)/\sqrt{\epsilon} = \bar{k}_1 I_1 + \cdots \bar{k}_n I_n.$$

Therefore, H_ϵ is conjugate to a fast periodic perturbation to $H^s_{p_0,\mathcal{B}}$. Note that our definition depend on the choice of basis $\mathcal{B}$. A basis free definition requires using a non-standard torus $\mathbb{T}^{n+1}/\omega(p_0)\mathbb{R}$ as the configuration space, and in this paper we choose to avoid this setting and fix a basis. Such averaged systems were studied in [24].

When $n = 2$, the slow system is a 2 degree 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 [23], [24], [25], [10], [16], [14], [17], [20], [21]). This is no longer the case when $n > 2$, which is a serious obstacle to proving Arnold diffusion in higher degrees of freedom. In [15] it is proposed that we can sidestep this difficulty by using *dimension reduction*: using existence of normally hyperbolic invariant cylinders (NHICs) 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.

In order to make this idea specific it is convenient to define the slow system for any p_0 and any d-resonance $d \leqslant n$. 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

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

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

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

$$U_{p_0,\mathcal{B}}(\varphi_1, \dots, \varphi_d) = -\sum_{l\in\mathbb{Z}^d} h_{l_1k_1+\cdots l_dk_d}(p_0)\, e^{2\pi i(l_1\varphi_1+\cdots+l_d\varphi_d)}. \tag{1.4}$$

The system $H^s_{p_0,\mathcal{B}}$ 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 nice space.

In the sequel we fix a rank $m < n$ lattice, called the *strong lattice*, and its basis $\mathcal{B} = [k_1, \dots, k_m]$. We say an irreducible lattice $\Lambda \supset \Lambda^{\mathrm{st}}$ of rank d is *dominated* by Λ^{st} if

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

where $|k| = \sup_i |k_i|$ is the sup-norm. Given the relation $\Lambda^{\mathrm{st}} \subset \Lambda$, we extend the basis $[k_1, \dots, k_m]$ of Λ^{st} to a basis $\mathcal{B} = [k_1, \dots, k_d]$ of Λ, such a basis is called *adapted*. Naturally, as $M(\Lambda|\Lambda^{\mathrm{st}}) \to \infty$, we have $|k_{m+1}|, \dots, |k_d| \to \infty$ for any adapted basis.

While we have fixed the basis $\mathcal{B}^{\mathrm{st}}$ of Λ^{st}, the system $H_{p_0,\mathcal{B}}$ strongly depends on the choice of the adapted basis. To get a meaningful result, we only consider particular bases that we call *κ-ordered*. Roughly speaking, given $\kappa > 1$, a basis $[k_1, \dots, k_d]$ is κ-ordered if k_i is, up to a factor of order κ, the vector of smallest norm in the set $\Lambda\setminus\mathrm{span}_{\mathbb{Z}}\{k_1, \dots, k_{i-1}\}$. The precise definition of this basis is given in Section 2.2. We will show that there exists κ depending only on $\mathcal{B}^{\mathrm{st}}$, such that any $\Lambda \supsetneq \Lambda^{\mathrm{st}}$ admits a κ-ordered basis.

After an ordered basis is chosen, we have two systems $H^s_{p_0,\mathcal{B}^{\mathrm{st}}}$ and $H^s_{p_0,\mathcal{B}}$, which we call the *strong system* and *slow system* respectively. When the lattices have a dominant structure (see (1.5)), the slow system $H^s_{p_0,\mathcal{B}}$ inherits considerable amount of information from the strong system. Indeed, let us denote

$$\begin{aligned} H^s_{p_0,\mathcal{B}^{\mathrm{st}}} &= K^{\mathrm{st}}(I_1, \dots, I_m) - U^{\mathrm{st}}(\varphi_1, \dots, \varphi_m), \\ H^s_{p_0,\mathcal{B}} &= K(I_1, \dots, I_d) - U(\varphi_1, \dots, \varphi_d), \end{aligned}$$

under (1.5) and we will show that

$$K^{\mathrm{st}}(I_1, \dots, I_m) = K(I_1, \dots, I_m, 0, \dots, 0), \quad \|U - U^{\mathrm{st}}\|_{C^2} \ll 1, \tag{1.6}$$

which indicates $H^s_{p_0,\mathcal{B}}$ can be approximated by an *extension* of $H^s_{p_0,\mathcal{B}^{\mathrm{st}}}$. The variables φ_i, I_i, $1 \leqslant i \leqslant m$ are called the *strong variables*, while φ_i, I_i, $m+1 \leqslant i \leqslant d$ are called the *weak variables*.

Recall that for each convex Hamiltonian H, we can associate a Lagrangian $L = L_H$, and the Euler-Lagrange flow is conjugate to the Hamiltonian flow. Denote by $X^{\rm st}_{\rm Lag}$ and $X^s_{\rm Lag}$ the Euler-Lagrange vector fields associated to the Hamiltonians $H^s_{p_0,\mathcal{B}^{\rm st}}$ and $H^s_{p_0,\mathcal{B}}$. Since the system for $X^{\rm st}_{\rm Lag}$ is only defined for the strong variables (φ_i, v_i), $1 \leqslant i \leqslant m$, we define a *trivial extension* of $X^{\rm st}_{\rm Lag}$ by setting $\dot\varphi_i = \dot v_i = 0$, $m+1 \leqslant i \leqslant d$.

We show that after performing a coordinate change[(1)] and rescaling transformation in the weak variables, the transformed vector field $X^s_{\rm Lag}$ converges to that of $X^{\rm st}_{\rm Lag}$ in some sense. In particular, if $X^{\rm st}_{\rm Lag}$ admits a normally hyperbolic invariant cylinder (NHIC), so does $X^s_{\rm Lag}$. In a separate direction, we also obtain a limit theorem on the weak KAM solutions by variational arguments. We now formulate our main results in loose language, leaving the precise version for the next section.

Main Result. — *Assume that $r > n + 2(d-m) + 4$. Given a fixed lattice $\Lambda^{\rm st}$ of rank m with a fixed basis $\mathcal{B}^{\rm st}$, there exist $\kappa > 1$ depending only on $\mathcal{B}^{\rm st}$, and the following hold. Each rank $d, m \leqslant d \leqslant n$ irreducible lattice $\Lambda \supset \Lambda^{\rm st}$ admits a κ-ordered basis, under which we have:*

1. *(Geometrical) As $M(\Lambda|\Lambda^{\rm st}) \to \infty$, the projection of $X^s_{\rm Lag}$ to the strong variables (φ_i, v_i), $1 \leqslant i \leqslant m$ converges to $X^{\rm st}_{\rm Lag}$ uniformly. Moreover, by introducing a coordinate change and rescaling affecting only the weak variables (φ_i, v_i), $m+1 \leqslant i \leqslant d$, the transformed vector field of $X^s_{\rm Lag}$ converges to a trivial extension of the vector field of $X^{\rm st}_{\rm Lag}$. As a corollary, we obtain that if $X^{\rm st}_{\rm Lag}$ admits an NHIC, then so does $X^s_{\rm Lag}$ for sufficiently large $M(\Lambda|\Lambda^{\rm st})$.*
2. *(Variational) If, in addition, we have $r > n + 4(d-m) + 4$, then as $M(\Lambda|\Lambda^{\rm st}) \to \infty$, the weak KAM solution of $H^s_{p_0,\mathcal{B}}$ (of properly chosen cohomology classes) converges uniformly to a trivial extension of a weak KAM solution of $H^s_{p_0,\mathcal{B}^{\rm st}}$, considered as functions on $\mathbb{R}^d$. We also obtain corollaries concerning the limits of Mañe, Aubry sets, rotation vector of minimal measure, and Peierl's barrier function. The precise definitions of these objects will be given later.*

The statement that $H^s_{p_0,\mathcal{B}^{\rm st}}$ approximates $H^s_{p_0,\mathcal{B}}$ is related to the classical concept of partial averaging (see for example [5]). The statement $\min_{k\in\Lambda\setminus\Lambda^{\rm st}} |k| \gg \max_{k\in\mathcal{B}^{\rm st}} |k|$ says that the resonances in $\Lambda^{\rm st}$ 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

1. The coordinate change we perform is known in analytic mechanics as the Routhian coordinates, which is an half-Lagrangian, half-Hamiltonian setting, see (2.8).

of the averaging independent of the small parameter ε. The theory is far from a simple corollary of (1.6), with the main difficulty coming from the fact that *as $M(\Lambda|\Lambda^{\mathrm{st}}) \to \infty$, the quadratic part of the system $H^s_{p_0,\mathcal{B}}$ becomes unbounded.*

In [24], John Mather developed a theory of (partial) averaging for the nearly integrable Lagrangian system, which is dual to our setting. Quantitative estimates on the action of minimizing orbits of the original system versus the slow system are obtained. Our variational result is related to [24], 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 for weak KAM solutions.

The formulation of the limit theorem in weak KAM solution requires special care. A natural candidate is Tonelli convergence (convergence of Lagrangian within the Tonelli family, see [7]). In our setup, $H^s_{p_0,\mathcal{B}^{\mathrm{st}}}$ and $H^s_{p_0,\mathcal{B}}$ are defined on different spaces, we need to consider the trivial extension of $H^s_{p_0,\mathcal{B}^{\mathrm{st}}}$ to a higher dimensional space. The extended Lagrangian is then *degenerate* and obviously not Tonelli. Moreover, the standard C^2 norm of the Lagrangian becomes unbounded in the limit process. We nevertheless obtain the convergence of weak KAM solutions.

While this paper is mainly motivated by Arnold diffusion, the paper is self-contained and do not relate to the actual diffusion problem. We hope our treatment of partial averaging and its variational aspects 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 a basis is handled in Section 3, and the estimates 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 . In Appendix A we prove Theorem 2.4 about existence of normally hyperbolic invariant cylinders stated in Section 2.4.

An earlier draft of the current paper is available on arxiv ([18]), which also includes a construction of the diffusion path such that the slow system at all strong resonances are dominant. We separated this construction from the current paper and it will appear in a future work.

2. Formulation of results

In order to state our results we establish an additional (filtrated) structure of the ambient lattice Λ relative to the strong lattice Λ^{st}:

$$\Lambda^{\mathrm{st}} = \Lambda_m \subset \Lambda_{m+1} \subset \cdots \subset \Lambda_d = \Lambda,$$

where each next lattice has rank by one exceeding the previous one, and associate to it a decomposition of the potential U into a filtrated sum (2.2).

2.1. The slow system and the choice of basis. — Recall that a slow system is

$$H^s_{p_0,\mathcal{B}}(\varphi, I) = K_{p_0,\mathcal{B}}(I) - U_{p_0,\mathcal{B}}(\varphi), \quad \varphi \in \mathbb{T}^d, I \in \mathbb{R}^d,$$

defined for a rank$-d$ irreducible lattice Λ with ordered basis $\mathcal{B} = [k_1, \dots, k_d]$, and a point $p_0 \in \mathbb{R}^n$. Let us denote

$$Z_{\mathcal{B}}(\varphi_1, \dots, \varphi_d, p) = \sum_{l \in \mathbb{Z}^d} h_{l_1 k_1 + \cdots l_d k_d}(p) e^{2\pi i (l_1 \varphi_1 + \cdots + l_d \varphi_d)}, \tag{2.1}$$

where $H_1(\theta, p, t) = \sum_{k \in \mathbb{Z}^{d+1}} h_k(p) e^{2\pi i k \cdot (\theta, t)}$, then (1.4) becomes

$$U_{p_0,\mathcal{B}}(\varphi) = -Z_{\mathcal{B}}(\varphi, p_0).$$

Let $\Lambda^{\mathrm{st}}, \mathcal{B}^{\mathrm{st}}$ be the strong lattice and basis, and consider an adapted basis $\mathcal{B}$ of an irreducible lattice $\Lambda \supset \Lambda^{\mathrm{st}}$ of rank d. We define the *filtration* $\Lambda^{\mathrm{st}} = \Lambda_m \subset \Lambda_{m+1} \subset \cdots \subset \Lambda_d = \Lambda$ associated to $\mathcal{B}$ by $\Lambda_i = \mathrm{span}_{\mathbb{Z}}\{k_1, \dots, k_i\}$, $m \leqslant i \leqslant d$. Then each Λ_i is irreducible of rank i.

Given $\kappa > 1$, $\mathcal{B}$ is called κ-ordered if:

1. For $m + 1 \leqslant i \leqslant n$, $|k_i| \leqslant \kappa M(\Lambda_i | \Lambda_{i-1})$.
2. For $m + 1 \leqslant i < j \leqslant n$, $|k_i| \leqslant \kappa(1 + |k_j|)$.

The following proposition, proved in Section 3, ensures existence of ordered bases.

Proposition 2.1. — *There is $\kappa = \kappa(\mathcal{B}^{\mathrm{st}}, d) > 0$ such that Λ admits a κ-ordered basis.*

We split the basis $\mathcal{B}$ into the strong and weak component, and introduce the following notations

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

Denote also

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

such naming convention will be kept for the whole paper.

Recall that a κ-ordered basis $\mathcal{B}$ comes with a filtration $\Lambda_m \subset \cdots \subset \Lambda_d$, with $\mathcal{B}_i = [k_1, \dots, k_i]$ being a basis of Λ_i. Let us define, for $m + 1 \leqslant i < d$,

$$U_{p_0,\mathcal{B}_{i-1},\mathcal{B}_i}(\varphi_1, \dots, \varphi_i) = U_{p_0,\mathcal{B}_i} - U_{p_0,\mathcal{B}_{i-1}},$$

As a result, we have

$$\begin{aligned} H^s_{p_0,\mathcal{B}} &= K_{p_0,\mathcal{B}} - U_{p_0,\mathcal{B}^{\mathrm{st}}} - U_{p_0,\mathcal{B}_m,\mathcal{B}_{m+1}} - \cdots - U_{\mathcal{B}_{d-1},\mathcal{B}_d} \\ &=: K_{p_0,\mathcal{B}} - U_{p_0,\mathcal{B}^{\mathrm{st}}} - U_{p_0,\mathcal{B}^{\mathrm{st}},\mathcal{B}^{\mathrm{wk}}}. \end{aligned} \tag{2.2}$$

Our first theorem gives estimates on $U_{p_0,\mathcal{B}_{i-1},\mathcal{B}_i}$ under a κ-ordered basis.

THEOREM 2.2. — *Let Λ^{st}, $\mathcal{B}^{\mathrm{st}}$ be the strong lattice and its basis, let $m < d \leqslant n$. Suppose H_1 is C^r with $r > n + 2(d-m) + 4$, and $\|H_1\|_{C^r} = 1$. Then there exists $\kappa = \kappa(\mathcal{B}^{\mathrm{st}}, n) > 1$ such that for each rank d irreducible lattice $\Lambda \supset \Lambda^{\mathrm{st}}$, there exists a κ-ordered basis $\mathcal{B}$ such that for $1 \leqslant i \leqslant d-m$, we have*

$$\|U^{\mathrm{wk}}_{p_0, \mathcal{B}_{i+m-1}, \mathcal{B}_{i+m}}\|_{C^2} \leqslant \kappa(1 + |k^{\mathrm{wk}}_i|)^{-r+n+2(d-m)+4}.$$

In particular, $\|U^{\mathrm{wk}}_{p_0, \mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}}\|_{C^2} \to 0$ as $M(\Lambda|\Lambda^{\mathrm{st}}) \to \infty$.

This theorem is proved in Section 3.

We will call any Hamiltonian that satisfy the conclusions of Theorem 2.2 a *dominant Hamiltonian.* In the next section, we define an abstract space of dominant Hamiltonians.

2.2. An abstract space of dominant Hamiltonians. — Define

$$\Omega^{m,d} := (\mathbb{Z}^{n+1})^d \times \mathbb{R}^n \times C^2(\mathbb{T}^m) \times C^2(\mathbb{T}^{m+1}) \times \cdots \times C^2(\mathbb{T}^d),$$
$$(\mathcal{B}^{\mathrm{st}} = [k^{\mathrm{st}}_1, \dots, k^{\mathrm{st}}_m], \mathcal{B}^{\mathrm{wk}} = [k^{\mathrm{wk}}_1, \dots, k^{\mathrm{wk}}_{d-m}], p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}} = (U^{\mathrm{wk}}_1, \dots, U^{\mathrm{wk}}_{d-m}))$$

and a mapping with

$$\mathcal{H}^s : \Omega^{m,d} \to C^2(\mathbb{T}^d \times \mathbb{R}^d),$$
$$\mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) = K_{p_0,\mathcal{B}}(I) - U^{\mathrm{st}}(\varphi_1, \dots, \varphi_m) - \sum_{i=1}^{d-m} U^{\mathrm{wk}}_i(\varphi_1, \dots, \varphi_{i+m}),$$

where $\mathcal{B} = [\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}]$ and $K_{p_0,\mathcal{B}}$ is defined by (1.3).

We equip $\Omega^{m,d}$ with the product topology, with a discrete topology on k^{wk}_i and the standard norms on other components. The map $\mathcal{H}^s$ is smooth in p_0, $U^{\mathrm{st}}, U^{\mathrm{wk}}_1, \dots, U^{\mathrm{wk}}_{d-m}$. Let $\Omega^{m,d}(\mathcal{B}^{\mathrm{st}})$ be the subset of $\Omega^{m,d}$ with fixed $\mathcal{B}^{\mathrm{st}}$.

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

1. For any $1 \leqslant i < j \leqslant d-m$, $|k^{\mathrm{wk}}_i| \leqslant \kappa\ (1 + |k^{\mathrm{wk}}_j|)$.
2. For each $1 \leqslant i \leqslant d-m$, $\|U^{\mathrm{wk}}_i\|_{C^2} \leqslant \kappa\,(1 + |k^{\mathrm{wk}}_i|)^{-q}$.

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

$$\mu(\mathcal{B}^{\mathrm{wk}}) = \min_{1 \leqslant i \leqslant d-m} |k^{\mathrm{wk}}_i|,$$

then in $\Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}})$, we have $\|U^{\mathrm{wk}}_i\| \leqslant \kappa\mu(\mathcal{B}^{\mathrm{wk}})^{-q}$, i.e., the weak potential $U^{\mathrm{wk}} := \sum_{i=1}^{d-m} U^{\mathrm{wk}}_i \to 0$ as $\mu(\mathcal{B}^{\mathrm{wk}}) \to \infty$.

We restate Theorem 2.2 using the new language:

THEOREM (Theorem 2.2 restated). — *Under the assumptions of Theorem 2.2, there exists a constant* $\kappa = \kappa(\mathcal{B}^{\mathrm{st}}, n) > 1$, *integer vectors* $\mathcal{B}^{\mathrm{wk}} = [k_1^{\mathrm{wk}}, \dots, k_{d-m}^{\mathrm{wk}}]$ *with* $\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}$ *forming an adapted basis, such that for* $q = r-n-2(d-m)-4 > 0$ *we have*

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

The strong Hamiltonian is defined by the mapping

$$\mathcal{H}^{\mathrm{st}} : (\mathbb{Z}^{n+1})^m \times \mathbb{R}^n \times C^2(\mathbb{T}^m) \to C^2(\mathbb{T}^m \times \mathbb{R}^m),$$
$$\mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p_0, U^{\mathrm{st}}) = K_{p_0,\mathcal{B}^{\mathrm{st}}}(I^{\mathrm{st}}) - U^{\mathrm{st}}(\varphi^{\mathrm{st}}).$$

We extend the definition to $\Omega^{m,d}$ by writing

$$\mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) = \mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p_0, U^{\mathrm{st}}).$$

We will prove all our limit theorems in the space $\Omega_{\kappa,q}^{m,d}(\mathcal{B}^{\mathrm{st}})$.

2.3. The rescaling limit. — We fix $\mathcal{B}^{\mathrm{st}}, \kappa > 1$ and $(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega_{\kappa,q}^{m,d}(\mathcal{B}^{\mathrm{st}})$. Denote

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

We write

$$\begin{aligned} H^s(\varphi, I) &= K(I) - U^{\mathrm{st}}(\varphi^{\mathrm{st}}) - U^{\mathrm{wk}}(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) \\ H^{\mathrm{st}}(\varphi^{\mathrm{st}}, I^{\mathrm{st}}) &= K^{\mathrm{st}}(I^{\mathrm{st}}) - U^{\mathrm{st}}(\varphi^{\mathrm{st}}) = K(I^{\mathrm{st}}, 0) - U^{\mathrm{st}}(\varphi^{\mathrm{st}}), \end{aligned} \tag{2.3}$$

where $U^{\mathrm{wk}} = \sum_{i=1}^{d-m} U_i^{\mathrm{wk}}$. We will keep using this notation throughout the paper.

As $\mu(\mathcal{B}^{\mathrm{wk}}) \to \infty$, we have $\|U^{\mathrm{wk}}\|_{C^2} \to 0$. However, $K(I^{\mathrm{st}}, I^{\mathrm{wk}})$ is not a small perturbation of $K(I^{\mathrm{st}}, 0)$, in fact, as $\mu(\mathcal{B}^{\mathrm{wk}}) \to \infty$, $K(I^{\mathrm{st}}, I^{\mathrm{wk}})$ becomes unbounded (see (2.5) below).

Let $Q_0(p) = \partial_{pp}^2 H_0(p)$ and $Q(p)$ be the $(n+1)\times(n+1)$ matrix $\begin{bmatrix} Q_0 & 0 \\ 0 & 0 \end{bmatrix}$, then

$$K(I) = Q(p_0)(k_1 I_1 + \cdots + k_d I_d) \cdot (k_1 I_1 + \cdots + k_d I_d).$$

We write

$$\partial_{II}^2 K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \quad A = \partial^2_{I^{\mathrm{st}} I^{\mathrm{st}}} K, B = \partial^2_{I^{\mathrm{st}} I^{\mathrm{wk}}} K, C = \partial^2_{I^{\mathrm{wk}} I^{\mathrm{wk}}} K, \tag{2.4}$$

then

$$(A)_{ij} = (k_i^{\mathrm{st}})^T Q k_j^{\mathrm{st}}, \quad (B)_{ij} = (k_i^{\mathrm{st}})^T Q k_j^{\mathrm{wk}}, \quad (C)_{ij} = (k_i^{\mathrm{wk}})^T Q k_j^{\mathrm{wk}}. \tag{2.5}$$

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

$$\begin{cases} \dot\varphi^{\mathrm{st}} = A I^{\mathrm{st}} + B I^{\mathrm{wk}}, & \dot I^{\mathrm{st}} = \partial_{\varphi^{\mathrm{st}}} U, \\ \dot\varphi^{\mathrm{wk}} = B^T I^{\mathrm{st}} + C I^{\mathrm{wk}}, & \dot I^{\mathrm{wk}} = \partial_{\varphi^{\mathrm{wk}}} U, \end{cases}$$

where $U = U^{\mathrm{st}} + U^{\mathrm{wk}}$. Then the Euler-Lagrangian vector field X^s_{Lag} is

$$\begin{cases} \dot\varphi^{\mathrm{st}} = v^{\mathrm{st}}, & \dot v^{\mathrm{st}} = A\partial_{\varphi^{\mathrm{st}}}U + B\partial_{\varphi^{\mathrm{wk}}}U, \\ \dot\varphi^{\mathrm{wk}} = v^{\mathrm{wk}}, & \dot v^{\mathrm{wk}} = B^T\partial_{\varphi^{\mathrm{st}}}U + C\partial_{\varphi^{\mathrm{wk}}}U, \end{cases} \tag{2.6}$$

which will be compared to the Euler-Lagrange vector field of H^{st}

$$\dot\varphi^{\mathrm{st}} = v^{\mathrm{st}}, \quad \dot v^{\mathrm{st}} = A\partial_{\varphi^{\mathrm{st}}}U^{\mathrm{st}}, \tag{2.7}$$

denoted X^{st}. To show that the projection of (2.6) converges to (2.7), we only need to show $\|B\partial_{\varphi^{\mathrm{wk}}}U\| \to 0$.

For convergence of weak variables, we will need a rescaling. It turns out that it is better to rescale the I^{wk} variable. Introduce the coordinate change

$$\Phi : (\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}}) \mapsto (\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, v^{\mathrm{wk}}), \quad v^{\mathrm{wk}} = B^TA^{-1}v^{\mathrm{st}} - \tilde C I^{\mathrm{wk}}, \tag{2.8}$$

where $\tilde C = C - B^TA^{-1}B$ which is symmetric and invertible.

This is a "half Lagrangian" setting in the sense that $(\varphi^{\mathrm{st}}, v^{\mathrm{st}})$ is remain the Lagrangian setup, while $(\varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$ is in the Hamiltonian format. This is known in analytic mechanics as the *Routhian coordinates.* Since the coordinate change is identity in $\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}$ let us compute the Jacobi matrix in the other two variables:

$$\frac{\partial(v^{\mathrm{st}}, v^{\mathrm{wk}})}{\partial(v^{\mathrm{st}}, I^{\mathrm{wk}})} = \begin{bmatrix} \mathrm{Id} & 0 \\ B^TA^{-1} & \tilde C \end{bmatrix}, \quad \frac{\partial(v^{\mathrm{st}}, I^{\mathrm{wk}})}{\partial(v^{\mathrm{st}}, v^{\mathrm{wk}})} = \begin{bmatrix} \mathrm{Id} & 0 \\ -\tilde C^{-1}B^TA^{-1} & \tilde C^{-1} \end{bmatrix}.$$

Thus, it is a diffeomorphism and the transformed vector field is

$$\begin{aligned} X^s &= (\Phi^{-1})_* X^s_{\mathrm{Lag}} = (D\Phi)^{-1} X^s_{\mathrm{Lag}} \circ \Phi \\ &= \begin{bmatrix} \mathrm{Id} & & & \\ & & \mathrm{Id} & \\ & & & \mathrm{Id} \\ & -\tilde C^{-1}B^TA^{-1} & & \tilde C^{-1} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} & v^{\mathrm{st}} & \\ A\partial_{\varphi^{\mathrm{st}}}U + B\partial_{\varphi^{\mathrm{wk}}}U & B^TA^{-1}v^{\mathrm{st}} - \tilde C I^{\mathrm{wk}} & B^T\partial_{\varphi^{\mathrm{st}}}U + C\partial_{\varphi^{\mathrm{wk}}}U \end{bmatrix} \\ &= \begin{bmatrix} & v^{\mathrm{st}} & \\ A\partial_{\varphi^{\mathrm{st}}}U + B\partial_{\varphi^{\mathrm{wk}}}U & B^TA^{-1}v^{\mathrm{st}} - \tilde C I^{\mathrm{wk}} & \partial_{\varphi^{\mathrm{wk}}}U \end{bmatrix} \end{aligned} \tag{2.9}$$

where the last line uses

$$\begin{aligned} &- \tilde C^{-1}B^TA^{-1}\left(A\partial_{\varphi^{\mathrm{st}}}U + B\partial_{\varphi^{\mathrm{wk}}}U\right) + \tilde C^{-1}\left(B^T\partial_{\varphi^{\mathrm{st}}}U + C\partial_{\varphi^{\mathrm{wk}}}U\right) \\ &= \tilde C^{-1}\left(-B^TA^{-1}B\partial_{\varphi^{\mathrm{wk}}}U + C\partial_{\varphi^{\mathrm{wk}}}U\right) = \partial_{\varphi^{\mathrm{wk}}}U \end{aligned}$$

by definition of $\tilde C$. We denote by $X^s(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$ the vector field of $(\Phi^{-1})_* X^s_{\mathrm{Lag}}$, lifted to the universal cover $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^{d-m}$.

Consider the trivial lift of the strong Lagrangian vector field X^{st}, defined on the universal cover

$$\begin{cases} \dot\varphi^{\mathrm{st}} = v^{\mathrm{st}}, & \dot v^{\mathrm{st}} = A\partial_{\varphi^{\mathrm{st}}}U, \\ \dot\varphi^{\mathrm{wk}} = 0, & \dot I^{\mathrm{wk}} = 0, \end{cases} \tag{2.10}$$

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

Given $1 \geqslant \sigma_1 \geqslant \cdots \geqslant \sigma_{d-m} > 0$, let $\Sigma = \mathrm{diag}\{\sigma_1, \ldots, \sigma_{d-m}\}$. We define a rescaling coordinate change $\Phi_\Sigma : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ by

$$\Phi_\Sigma : (\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}}) \mapsto (\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \Sigma^{-1}\varphi^{\mathrm{wk}}, \Sigma I^{\mathrm{wk}}). \tag{2.11}$$

The rescaled vector field for X^s is

$$\begin{aligned} \tilde X^s &:= (\Phi_\Sigma^{-1})_* X^s = (D\Phi_\Sigma)^{-1} X^s \circ \Phi_\Sigma, \\ \tilde X^s(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}}) &= (D\Phi_\Sigma)^{-1} X^s(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \Sigma^{-1}\varphi^{\mathrm{wk}}, \Sigma I^{\mathrm{wk}}), \end{aligned} \tag{2.12}$$

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

THEOREM 2.3. — *Fix $\mathcal{B}^{\mathrm{st}}$ and $\kappa > 1$. Assume that $q > 2$. Then there exists a constant $M = M(\mathcal{B}^{\mathrm{st}}, D, \kappa, q, d-m) > 1$, such that for $(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega_{\kappa,q}^{m,d}(\mathcal{B}^{\mathrm{st}})$, $H^s = \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$ and $H^{\mathrm{st}} = \mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p, U^{\mathrm{st}})$, the following hold.*

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

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

In particular, Theorem 2.3 implies that as $\mu(\mathcal{B}^{\mathrm{wk}}) \to \infty$, the vector field $\tilde X^s$ converges to X_L^{st} in the C^1 topology over compact sets. For applications, our version is more flexible as it is uniform over the whole space. Theorem 2.3 is proved in Section 4.1.

2.4. Persistence of normally hyperbolic invariant cylinders. — Our main application for Theorem 2.3 is to prove persistence of normally hyperbolic invariant cylinders (NHICs).

Let W be a manifold. For $R > 0$, let $B_R^l \subset \mathbb{R}^l$ denote the ball of radius R at the origin. A $2l$-cylinder Λ_R is defined by $\Lambda_R^l := \chi(\mathbb{T}^l \times B_R^l)$, where $\chi : \mathbb{T}^l \times B_R^l \to W$ is an embedding.

Let ϕ_t be a C^2 flow on W, and $\Lambda_R \subset W$ be a cylinder for some $R > 0$. We say that Λ_R^l is *normally hyperbolic (weakly) invariant cylinder* (NHWIC) if there exists $t_0 > 0$ such that the following hold.

- The vector field of ϕ_t is tangent to Λ_R^l at every $z \in \Lambda_R$.

- For each $z \in \Lambda^l_R$, there exists a splitting
$$T_z M = E^c(z) \oplus E^s(z) \oplus E^u(z), \qquad \text{where } E^c(z) = T_z\Lambda,$$
weakly invariant in the sense that
$$D\phi_{t_0}(z)E^\sigma(z) = E^\sigma(\phi_{t_0} z), \qquad \text{if } z,\, \phi_{t_0} z \in \Lambda \quad \text{and} \quad \sigma = c, s, u.$$
- There exist $0 < \alpha < \beta^{-1} < 1$ and a C^1 Riemannian metric g, called *an adapted metric, on a neighborhood of Λ^l_R such that whenever $z, \phi_{t_0} z \in \Lambda^l_R$,*
$$\|D\phi_{t_0}(z)|E^s\|, \quad \|(D\phi_{t_0}(z)|E^u)^{-1}\| < \alpha,$$
$$\|(D\phi_{t_0}(z)|E^c)^{-1}\|, \quad \|D\phi_{t_0}(z)|E^c\| < \beta,$$
where the norms are taken with respect to the metric g (see e.g., [13]*).*

The cylinder is called *normally hyperbolic (fully) invariant* (NHIC) if it satisfies the above conditions, and both Λ^l_R and $\partial\Lambda^l_R$ are invariant under ϕ_{t_0}. A more common definition of normally hyperbolic (fully) invariant cylinders assumes a spectral radius condition, but our definition is equivalent, see e.g., [9] Prop.5.2.2.

Moreover:

- If the parameters α, β satisfies the bunching condition $\alpha < \beta^2$, then the bundles E^s, E^u are C^1 smooth.
- When E^s, E^u are smooth, we can always choose the adapted metric g such that E^s, E^u and E^c are orthogonal.

Recall that $X^{\mathrm{st}}_{\mathrm{Lag}}$, X^s_{Lag} denotes the Lagrangian vector fields. Suppose $X^{\mathrm{st}}_{\mathrm{Lag}}$ admits a normally hyperbolic (fully) invariant cylinder Λ^{st}, we claim that X^s_{Lag} admits a normally hyperbolic weakly invariant cylinder diffeomorphic to $\Lambda^{\mathrm{st}} \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m})$.

Theorem 2.4. — *Consider a strong lattice $\mathcal{B}^{\mathrm{st}}$, a strong potential $U^{\mathrm{st}} \in C^2(\mathbb{T}^m)$, $\kappa > 0, a > 0$, and $q > 2$. Assume that for some $1 \leqslant l \leqslant m-1$, the Euler-Lagrange vector field $X^{\mathrm{st}}_{\mathrm{Lag}}$ of $H^{\mathrm{st}} = \mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p_0, U^{\mathrm{st}})$ admits a $2l$-dimensional NHIC $\Lambda^{\mathrm{st}}_{1+a} = \chi^{\mathrm{st}}(\mathbb{T}^l \times B^l_{1+a})$, given by the embedding $\mathbb{T}^l \times B^l_{1+a} \to \mathbb{T}^m \times \mathbb{R}^m$, with the parameters $0 < \alpha < \beta^2 < 1$.*

Then there exists an open set $V \supset \Lambda^{\mathrm{st}}_1$ such that for any $\delta > 0$, there exists $M > 0$, such that for any $(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}) \in \Omega^{\kappa,q}_{m,d}(\mathcal{B}^{\mathrm{st}})$ with $\mu(\mathcal{B}^{\mathrm{wk}}) > M$, $H^s = \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$, the following hold.

There exists a C^1 embedding
$$\eta^s = (\eta^{\mathrm{st}}, \eta^{\mathrm{wk}}) : (\mathbb{T}^l \times B^l) \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m}) \to (\mathbb{T}^m \times B^m) \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m}),$$
such that $\Lambda^s = \eta^s((\mathbb{T}^l \times B^l) \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m}))$ is a $2(l+d-m)$-dimensional NHWIC under X^s_{Lag}. Moreover,
$$\|\eta^{\mathrm{st}}(z^{\mathrm{st}}, z^{\mathrm{wk}}) - \chi^{\mathrm{st}}(z^{\mathrm{st}})\| < \delta, \quad \forall z^{\mathrm{st}} \in \mathbb{T}^l \times B^l,\ z^{\mathrm{wk}} \in \mathbb{T}^{d-m} \times \mathbb{R}^{d-m},$$

and any X^s_{Lag}-invariant set contained in $V \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m})$ is contained in Λ^s.

The assumption $\alpha < \beta^2$ is not necessary, and is assumed for simplicity of the proof. Nevertheless, the assumption is satisfied in our intended application and in most perturbative settings. The proof is presented in Appendix A.

2.5. 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

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

be a C^3 Hamiltonian satisfying the condition $D^{-1}\mathrm{Id} \leqslant \partial^2_{II} H(\varphi, I) \leqslant D\,\mathrm{Id}$. The associated Lagrangian $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ 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

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

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 $u : \mathbb{T}^d \to \mathbb{R}$ is called a (backward) 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 $\gamma : [0, t] \to \mathbb{T}^d$ is absolutely continuous. Weak KAM solutions exist and are Lipschitz (see [12], [7]).

The relation between cohomologies. We now turn to the weak KAM solutions of dominant Hamiltonians. Given

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

we write as before $H^s = \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$, $H^{\mathrm{st}} = \mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p, U^{\mathrm{st}})$ and recall the notations (2.3).

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

$$L^s(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}, v^{\mathrm{st}}, v^{\mathrm{wk}}) = \check{K}(v^{\mathrm{st}}, v^{\mathrm{wk}}) + U^{\mathrm{st}}(\varphi^{\mathrm{st}}) + U^{\mathrm{wk}}(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}),$$
$$L^{\mathrm{st}}(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) = \check{K}^{\mathrm{st}}(v^{\mathrm{st}}) + U^{\mathrm{st}}(\varphi^{\mathrm{st}}),$$

where $\check{K}$, $\check{K}^{\mathrm{st}}$ are quadratic functions with $(\partial^2_{vv}\check{K}) = (\partial^2_{II}K)^{-1}$ and $(\partial^2_{v^{\mathrm{st}}v^{\mathrm{st}}}\check{K}^{\mathrm{st}}) = (\partial^2_{I^{\mathrm{st}}I^{\mathrm{st}}}K)^{-1}$ as matrices.

Given $c = (c^{\mathrm{st}}, c^{\mathrm{wk}}) \in \mathbb{R}^m \times \mathbb{R}^{d-m} = \mathbb{R}^d$, we show that the weak KAM solution of $L^s - c \cdot v$ is related to the weak KAM solution of $L^{\mathrm{st}} - \bar{c} \cdot v^{\mathrm{st}}$, where $\bar{c}$ is defined as

$$\bar{c} = c^{\mathrm{st}} + A^{-1}Bc^{\mathrm{wk}},$$

with $\partial^2_{II}K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ as in (2.5). To understand this definition, note that $\bar{c}$ uniquely satisfies

$$\partial_{I^{\mathrm{st}}}K^{\mathrm{st}}(\bar{c}) = \partial_{I^{\mathrm{st}}}K(c^{\mathrm{st}}, c^{\mathrm{wk}}).$$

If we view $c, \bar{c}$ as the momentum variable, then their corresponding velocities have the same strong component.

Semi-continuity of weak KAM solutions. We now state our main variational results. For $\nu \in \mathbb{N}$, we consider a sequence of dominant Hamiltonians with $\mu(\mathcal{B}^{\mathrm{wk}}_\nu) \to \infty$, and cohomology classes c_ν such that the corresponding $\bar{c}_\nu$ converge. Then the weak KAM solutions has 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.

Fix $\mathcal{B}^{\mathrm{st}}$ and $\kappa > 1$. For $\nu \in \mathbb{N}$, consider

$$(\mathcal{B}^{\mathrm{wk}}_\nu, p_\nu, U^{\mathrm{st}}_\nu, \mathcal{U}^{\mathrm{wk}}_\nu) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}}), \quad c_\nu = (c^{\mathrm{st}}_\nu, c^{\mathrm{wk}}_\nu) \in \mathbb{R}^m \times \mathbb{R}^{d-m},$$

write

$$H^s_\nu = \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}_\nu, p_\nu, U^{\mathrm{st}}_\nu, \mathcal{U}^{\mathrm{wk}}_\nu), \quad L^s_\nu = L_{H^s_\nu},$$

and let u_ν be a weak KAM solution of $L^s_\nu - c_\nu \cdot v$.

Denote $\mathcal{B}_\nu = [\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}_\nu]$, $K_\nu = K_{p_\nu, \mathcal{B}_\nu}$, and

$$A_\nu = \partial^2_{I^{\mathrm{st}}I^{\mathrm{st}}}K_\nu, B_\nu = \partial^2_{I^{\mathrm{st}}I^{\mathrm{wk}}}K_\nu, C_\nu = \partial^2_{I^{\mathrm{wk}}I^{\mathrm{wk}}}K_\nu.$$

THEOREM 2.5. — *Given $p_0 \in \mathbb{R}^n$, $U^{\mathrm{st}}_0 \in C^2(\mathbb{T}^m)$ and $\bar{c} \in \mathbb{R}^m$, assume as $\nu \to \infty$:*

- *$\mu(B^{\mathrm{wk}}_\nu) \to \infty$, $p_\nu \to p_0$, $U^{\mathrm{st}}_\nu \to U^{\mathrm{st}}_0$.*
- *$c^{\mathrm{st}}_\nu + A^{-1}_\nu B_\nu c^{\mathrm{wk}}_\nu \to \bar{c}$.*

Then:

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

$$L_{\mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p_0, U^{\mathrm{st}}_0)} - \bar{c} \cdot v^{\mathrm{st}}.$$

The proof of Theorem 2.5 occupies Sections 4 and 5, with some technical statements deferred to Section 7.

REMARK. — *Theorem 2.2 implies that under an ordered basis, we can express a slow system as a dominant system with parameters κ, q, where $q = r - n - 2(d-m) - 4$. For Theorem 2.3 we need $q > 2$, and for Theorem 2.5 we need $q > 2(d-m)$. Notice that to apply our theorems to the slow system, we need $q > 2(d-m)$ (or, equiv., $r > n+4(d-m)+4$) and as stated in our main result. The additional requirement for Theorem 2.5 is due to the higher requirement of Proposition 5.3 (see also* (7.16)*).*

Using the point of view in [7], 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 a basis and averaging

In this section we prove Proposition 2.1 and Theorem 2.2.

3.1. The choice of a basis. — Recall that we have a fixed irreducible lattice $\Lambda^{\mathrm{st}} \subset \mathbb{Z}^{n+1}$ of rank $m < n$, and a fixed basis $\mathcal{B}^{\mathrm{st}} = \{k_1, \dots, k_m\}$ for Λ^{st}. For $\Lambda^{\mathrm{st}} \subset \Lambda$ irreducible of rank d, we first construct a filtration $\Lambda^{\mathrm{st}} = \Lambda_m \subset \cdots \subset \Lambda_d = \Lambda$, with each Λ_i containing the vector with the smallest norm in $\Lambda \backslash \Lambda_{i-1}$, $m+1 \leqslant i \leqslant n$.

Explicitly, we define $l_i = k_i$ for $1 \leqslant i \leqslant m$, and l_i with $i > m$ inductively using the following procedure. Suppose $l_1, \dots, l_i$ are defined, let

$$\Lambda_i = \mathrm{span}_{\mathbb{R}}\{l_1, \dots, l_i\} \cap \Lambda, \quad M_{i+1} = \min\{|k| : k \in \Lambda \backslash \Lambda_i\}.$$

We define l_{i+1} to be a vector reaching the minimum in the definition of M_{i+1}, i.e $|l_{i+1}| = M_{i+1}$. Then $\Lambda_{i+1} = \mathrm{span}_{\mathbb{R}}(\Lambda_i \cup l_{i+1}) \cap \Lambda$. This provides the filtration as needed.

We have

$$|l_i| = M_i,\ m < i \leqslant d, \quad |l_j| \leqslant |l_i|,\ m < j < i \leqslant d,$$

but $l_1, \dots, l_d$ may not form a basis. We turn them into a basis using the following *explicit* procedure (see [28]).

For each $i = 1, \dots, d$, define

$$c_i = \min\{s_i : \quad s_{i,1}l_1 + \cdots + s_{i,i-1}l_{i-1} + s_i l_i \in \Lambda,\ s_i \in \mathbb{R}^+,\ s_{i,j} \in \mathbb{R}^+ \cup \{0\}\}. \tag{3.1}$$

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

$$c_{i,i-1} = \min\{s_{i,i-1} : \quad s_{i,1}l_1 + \cdots + s_{i,i-1}l_{i-1} + c_i l_i,\ s_{i,j} \in \mathbb{R}^+ \cup \{0\}\}.$$

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

$$c_{i,j-1} = \min\{s_{i,j-1} : \\ s_{i,1}l_1 + \cdots + s_{i,j-1}l_{i-1} + c_{i,j}l_i + \cdots + c_{i,i-1}l_{i-1} + c_i l_i \in \Lambda, \\ s_{i,1}, \dots, s_{i,j-1} \in \mathbb{R}^+ \cup \{0\}\}.$$

Finally,

$$k_i = c_{i,1}l_1 + \cdots + c_{i,i-1}l_{i-1} + c_i l_i.$$

We have the following lemma from the geometry of numbers.

LEMMA 3.1 (see [28]). — *Let $\Lambda \subset \mathbb{Z}^{n+1}$ be a lattice of rank $d \leqslant n$ and let $l_1, \dots, l_d$ be any linearly independent set in Λ. Let*

$$k_i = c_{i,1}l_1 + \cdots + c_{i,i-1}l_{i-1} + c_i l_i, \quad i \leqslant d.$$

be defined using the procedure above. Then

1. *For each $1 \leqslant i \leqslant d$, $k_1, \dots, k_i$ form a basis of* $\mathrm{span}_{\mathbb{R}}\{l_1, \dots, l_i\} \cap \Lambda$ *over $\mathbb{Z}$. In particular, $k_1, \dots, k_d$ form a basis of Λ.*
2. *For $1 \leqslant i < d$ and $1 \leqslant j \leqslant i-1$, we have*

$$0 \leqslant c_{i,j} < 1, \quad 0 < c_i \leqslant 1.$$

3. *If for some m such that $1 \leqslant m \leqslant d$, $l_1, \dots, l_m$ already form a basis of* $\mathrm{span}_{\mathbb{R}}\{l_1, \dots, l_m\} \cap \Lambda$ *over $\mathbb{Z}$, then $k_1 = l_1, \dots, k_m = l_m$.*

Proof. — For the proof of item 1, we refer to [28], Theorem 18. Item 2 and 3 follow from definition and item 1 as we explain below.

For item 2, note that for any $k_i = c_{i,1}l_1 + \cdots + c_{i,i-1}l_{i-1} + c_i l_i \in \Lambda$, we can always subtract an integer from any $c_{i,j}$ or c_i and remain in Λ. If the estimates do not hold, we can get a contradiction by reducing $c_{i,j}$ or c_i.

For item 3, if $l_1, \dots, l_m$ is a basis (over $\mathbb{Z}$) of $\mathrm{span}_{\mathbb{R}}\{l_1, \dots, l_m\} \cap \Lambda$, then all coefficients of $k_i = c_{i,1}l_1 + \cdots + c_{i,i-1}l_{i-1} + c_i l_i \in \Lambda$ for $i \leqslant m$ must be integers. Then the constraints of item 2 implies $c_{i,j} = 0$ and $c_i = 1$, namely $k_i = l_i$. □

Proof of Proposition 2.1. — We choose the basis $k_1, \dots, k_d$ as described. Lemma 3.1 implies $k_i = l_i$ for $1 \leqslant i \leqslant m$. Using

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

we get

$$|k_i| \leqslant |l_1| + \cdots + |l_i| \leqslant |k_1| + \cdots + |k_m| + M_{m+1} + \cdots + M_i.$$

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

$$|k_i| \leqslant \bar{M} + (i-m)M_i \leqslant \bar{M} + (d-m)M_i.$$

Moreover, for $i < j$, we have

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

The proposition follows by taking $\chi = \bar{M} + (d-m)$. □

3.2. Estimating the weak potential. — In this section we prove Theorem 2.2. 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 $k_1, \dots, k_d$ be chosen as in Proposition 2.1. Recall that

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

then we have

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

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

Lemma 3.2 (c.f. [8], 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* $r \geqslant n + 4$. *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*

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

Since $\min_{k \in \Lambda_i \setminus \Lambda_{i-1}} |k| = M_i$, we apply Lemma 3.2 to $\Lambda_i \setminus \Lambda_{i-1}$ to get

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

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) so 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

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

and as a result $Z_{\mathcal{B}_i} - Z_{\mathcal{B}_{i-1}} = ([H_1]_{\Lambda_i} - [H_1]_{\Lambda_{i-1}}) \circ (P_i^T)^{-1}$. Using the Faa-di Bruno formula, we have

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

for some $c_n > 1$ depending on n. We apply the following lemma in linear algebra:

Lemma 3.3. — *Given $1 \leqslant s \leqslant n+1$, let $P = \begin{bmatrix} k_1 & \cdots & k_s \end{bmatrix}$ be an integer matrix with linearly independent columns. Then there exists $c_n > 1$ depending only on n such that*

$$\min_{\|v\|=1} \|Pv\| = \min_{\|v\|=1} (v^T P^T P v)^{\frac{1}{2}} = \|(P^T P)^{-1}\|^{-\frac{1}{2}} \geqslant c_n^{-1} |k_1|^{-1} \cdots |k_s|^{-1}.$$

In particular, if $s = n+1$, then $\|P^{-1}\| = \|(P^T)^{-1}\| \leqslant c_n |k_1| \cdots |k_{n+1}|$.

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

$$|b_{ij}| \leqslant \frac{1}{\det(P^T P)} \sum_{\sigma} \prod_{s \neq i} a_{s\sigma(s)},$$

where σ ranges over all one-to-one mappings from $\{1, \ldots, m\} \backslash \{i\}$ to $\{1, \ldots, m\} \backslash \{j\}$. Since P is a nonsingular integer matrix, we have $\det(P^T P) \geqslant 1$. Moreover, $a_{ij} = k_i^T k_j \leqslant n |k_i||k_j|$. Therefore,

$$|b_{ij}| \leqslant \sum_{\sigma} \prod_{s \neq i} |k_s||k_{\sigma(s)}| \leqslant c_n \,\big(\prod_{s \neq i} |k_s|\big)\,\big(\prod_{s \neq j} |k_s|\big),$$

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

$$\|(P^T P)^{-1}\| \leqslant c_n \sup_{i,j} |B_{ij}| \leqslant c_n \sup_{i,j} \big(\prod_{s \neq i} |k_s|\big)\,\big(\prod_{s \neq j} |k_s|\big) \leqslant c_n \big(\prod_{s=1}^{m} |k_s|\big)^2.$$

If $s = n+1$, then $\|P^{-1}\| = \|(P^T P)^{-1}\|^{\frac{1}{2}} = \|(P P^T)^{-1}\|^{\frac{1}{2}} = \|(P^T)^{-1}\|$. $\square$

Proof of Theorem 2.2. — Using Lemma 3.3, there exists a constant $c_n > 0$ depending only on n such that

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

We have $|k_1|, \ldots, |k_m| \leqslant \bar{M}$, $|\hat{k}_{i+1}| = \cdots = |\hat{k}_{n+1}| = 1$, and from Lemma 3.1, $|k_{m+1}|, \ldots, |k_i| \leqslant \chi M_i$. Therefore,

$$\|P_i^{-1}\| = \|(P_i^T)^{-1}\| \leqslant c_n \bar{M}^m \chi^{i-m} M_i^{i-m}.$$

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

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

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 2.3. Fix $\mathcal{B}^{\mathrm{st}}, \kappa > 1$ and let $(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}})$, we define H^{st}, H^s as before (see (2.3)). Recall from (2.4) that $\partial^2_{II} K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$, with

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

The vector field $X^s(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$, defined on the universal cover $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^{d-m}$, is obtained from the Euler-Lagrange vector field via the non-degenerate coordinate change $\tilde{C} I^{\mathrm{wk}} = B^T A^{-1} v^{\mathrm{st}} - v^{\mathrm{wk}}$ (see (2.8)). The vector field $X_L^{\mathrm{st}}(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$ is defined as a trivial extension of the Euler-Lagrange vector field of H^{st}, also defined on the universal cover. More explicitly (see (2.9), (2.10))

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

Given $1 \geqslant \sigma_1 \geqslant \cdots \geqslant \sigma_{d-m} > 0$, let $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_{d-m}\}$. The rescaling is $\Phi_\Sigma : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$, given by (2.11).

We denote by $\tilde{X}^s(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$ the rescaled X^s. Using (4.1), we have

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

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

$$D(\tilde{X}^s - X_L^{\mathrm{st}}) = (\Phi_\Sigma)^{-1} D X^s \circ \Phi_\Sigma - X_L^{\mathrm{st}} = \tag{4.3}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ A\partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{st}}} U^{\mathrm{wk}} + B \partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{wk}}} U^{\mathrm{wk}} & 0 & A\partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{wk}}} U^{\mathrm{wk}} + B \partial^2_{\varphi^{\mathrm{wk}}\varphi^{\mathrm{wk}}} U^{\mathrm{wk}} \Sigma^{-1} & 0 \\ 0 & \Sigma B^T A^{-1} & 0 & -\Sigma \tilde{C} \Sigma \\ \Sigma^{-1} \partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{wk}}} U^{\mathrm{wk}} & 0 & \Sigma^{-1} \partial^2_{\varphi^{\mathrm{wk}}\varphi^{\mathrm{wk}}} U^{\mathrm{wk}} \Sigma^{-1} & 0 \end{bmatrix}.$$

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

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

1. *For any $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant d-m$, $\|\partial_{\varphi_j^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0}, \|\partial^2_{\varphi_i^{\mathrm{st}}\varphi_j^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0} \leqslant M_1 |k_j^{\mathrm{wk}}|^{-q}$ for any $1 \leqslant i, j \leqslant d-m$, $\|\partial^2_{\varphi_i^{\mathrm{wk}}\varphi_j^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0} \leqslant M_1 \sup\{|k_i^{\mathrm{wk}}|^{-q}, |k_j^{\mathrm{wk}}|^{-q}\}$.*

2. $\|A\partial_{\varphi^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-q}\}$, $\|A\partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{st}}} U^{\mathrm{wk}}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-q}\}$.
3. $\|B\partial_{\varphi^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-(q-1)}\}$, $\|B\partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{wk}}} U\|_{C^0}$ $\leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-(q-1)}\}$.
4. $\|B\partial^2_{\varphi^{\mathrm{st}}\varphi^{\mathrm{wk}}} U\Sigma^{-1}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-\frac{2q-4}{3}}\}$.
5. $\|\Sigma^{-1}\partial^2_{\varphi^{\mathrm{wk}}\varphi^{\mathrm{wk}}} U\Sigma^{-1}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-\frac{q-2}{3}}\}$.
6. $\|\Sigma B^T A^{-1}\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-\frac{q-2}{3}}\}$.
7. $\|\Sigma\tilde{C}\Sigma\|_{C^0} \leqslant M_1 \sup_i\{|k_i^{\mathrm{wk}}|^{-\frac{2q-4}{3}}\}$.

We first prove Theorem 2.3 using our lemma.

Proof of Theorem 2.3. — Noting that $\Pi_{(\varphi^{\mathrm{st}},v^{\mathrm{st}})}(\tilde{X}^s - X_L^{\mathrm{st}})$ is the first and second line of (4.2), using item 2 and 3 of Lemma 4.1 we get

$$\|\Pi_{(\varphi^{\mathrm{st}},v^{\mathrm{st}})}(\tilde{X}^s - X_L^{\mathrm{st}})\| \leqslant M \sup_j\{|k_j^{\mathrm{wk}}|^{-(q-1)}\} \leqslant 2M_1\mu(\mathcal{B}^{\mathrm{wk}})^{-(q-1)},$$

where M_1 is from Lemma 4.1.

Since $D(\tilde{X}^s - X_L^{\mathrm{st}})$ 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

$$\|D\tilde{X}^s - DX_L^{\mathrm{st}}\| \leqslant M \sup_j\{|k_j^{\mathrm{wk}}|^{-\frac{q-2}{3}}\} \leqslant 2M_1\mu(\mathcal{B}^{\mathrm{wk}})^{-\frac{q-2}{3}}. \qquad \square$$

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

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

Item 1. — We have

$$\begin{aligned}\|\partial_{\varphi_i^{\mathrm{wk}}} U^{\mathrm{wk}}\|_{C^0} \leqslant \sum_{l=1}^{d-m} \|\partial_{\varphi_i^{\mathrm{wk}}} U_l^{\mathrm{wk}}\|_{C^0} &\leqslant \sum_{l\geqslant i} \|\partial_{\varphi_i^{\mathrm{wk}}} U_l^{\mathrm{wk}}\|_{C^0} \\ &\leqslant \kappa \sum_{l\geqslant i} |k_l^{\mathrm{wk}}|^{-q} \leqslant (d-m)\kappa^{q+1}|k_i^{\mathrm{wk}}|^{-q},\end{aligned}$$

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

$$\|\partial^2_{\varphi_i^{\mathrm{st}}\varphi_j^{\mathrm{wk}}} U^{\mathrm{wk}}\| \leqslant \sum_{l\geqslant j} \|U_l^{\mathrm{wk}}\|_{C^2} \leqslant (d-m)\kappa^{q+1}|k_j^{\mathrm{wk}}|^{-q},$$

$$\|\partial^2_{\varphi_i^{\mathrm{wk}}\varphi_j^{\mathrm{wk}}} U^{\mathrm{wk}}\| \leqslant \sum_{l\geqslant \sup\{i,j\}} \|U_l^{\mathrm{wk}}\|_{C^2} \leqslant (d-m)\kappa^{q+1}\sup\{|k_i^{\mathrm{wk}}|^{-q}, |k_j^{\mathrm{wk}}|^{-q}\}$$

the second and third estimate follows.

Item 2. — We have

$$|(A\partial_{\varphi^{\rm st}}U^{\rm wk})_i| = |\sum_l (k_l^{\rm st})^T Q k_i^{\rm st}\partial_{\varphi_i^{\rm st}}U^{\rm wk}| \leqslant m\bar{M}^2\|Q\|\|\partial_{\varphi_i^{\rm st}}U\|$$
$$\leqslant m(d-m)\bar{M}^2\|Q\|\kappa^{q+1}|k_i^{\rm wk}|^{-q},$$

where the last line is due to item 1. Similarly,

$$|(A\partial^2_{\varphi^{\rm st}\varphi^{\rm st}}U^{\rm wk})_{ij}| \leqslant \bar{M}^2\|Q\|\|\partial_{\varphi_i^{\rm st}\varphi_j^{\rm st}}U\| \leqslant (d-m)\bar{M}^2\|Q\|\kappa^{q+1}|k_j^{\rm wk}|^{-q}.$$

Since the vector or 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/vector entries.

Item 3. — Similar to item 2,

$$|(B\partial_{\varphi^{\rm wk}}U)_i| = |\sum_l (k_l^{\rm st})^T Q k_i^{\rm wk}\partial_{\varphi_i^{\rm wk}}U^{\rm wk}| \leqslant (d-m)\bar{M}\|Q\||k_i^{\rm wk}|\|\partial_{\varphi_i^{\rm wk}}U^{\rm wk}\|$$
$$\leqslant (d-m)\bar{M}\|Q\|(d-m)\kappa^{q+1}|k_i^{\rm wk}|^{-(q-1)},$$

while

$$|(B\partial^2_{\varphi^{\rm st}\varphi^{\rm wk}}U^{\rm wk})_{ij}| = |(k_i^{\rm st})^T Q k_j^{\rm wk}\partial^2_{\varphi_i^{\rm st}\varphi_j^{\rm wk}}U^{\rm wk}| \leqslant (d-m)\kappa^{q+1}\bar{M}\|Q\||k_j^{\rm wk}|^{-(q-1)}.$$

Item 4. — We have

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

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

Item 5. — Using item 1 and choice of σ_j, we have

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

Item 6. — We have

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

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

Item 7. — Recall $\tilde{C} = C - B^T A^{-1} B$. We have

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

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

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

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

$$\|\Sigma \tilde{C} \Sigma\| = \|\Sigma (C - B^T A^{-1} B) \Sigma\| \leqslant \|\Sigma C \Sigma\| + \|\Sigma B^T A^{-1} B \Sigma\| \leqslant 2 \|\Sigma C \Sigma\|.$$

Item 7 follows the previously obtained bound of $\|\Sigma C \Sigma\|$. □

4.2. The slow Lagrangian. — In this section, we derive the form of the slow Lagrangian in preparation for the variational part. We fix $\mathcal{B}^{\mathrm{st}}, \kappa > 1$ and $(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}})$. Recall the notations of H^{st}, H^s from (2.3).

We have

$$L^s(\varphi, v) = \check{K}(v) + U^{\mathrm{st}}(\varphi^{\mathrm{st}}) + U^{\mathrm{wk}}(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}), \quad L^{\mathrm{st}}(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) = \check{K}^{\mathrm{st}}(v^{\mathrm{st}}) + U^{\mathrm{st}}(\varphi^{\mathrm{st}}),$$

where $\partial^2_{vv} \check{K} = (\partial^2_{II} K)^{-1}$, $\partial^2_{v^{\mathrm{st}} v^{\mathrm{st}}} \check{K}^{\mathrm{st}} = (\partial^2_{I^{\mathrm{st}} I^{\mathrm{st}}} K)^{-1}$. Recall the notation

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

Lemma 4.2. — *With the above notations we have*

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

where

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

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

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

then

$$\begin{aligned} L^s(v, \varphi) - c \cdot v = L^{\mathrm{st}}&(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) - \bar{c} \cdot v^{\mathrm{st}} \\ &+ \frac{1}{2}(w^{\mathrm{wk}} - \tilde{C} c^{\mathrm{wk}}) \cdot \tilde{C}^{-1} (w^{\mathrm{wk}} - \tilde{C} c^{\mathrm{wk}}) \\ &- \frac{1}{2} c^{\mathrm{wk}} \cdot \tilde{C} c^{\mathrm{wk}} + U^{\mathrm{wk}}(\varphi^{\mathrm{wk}}, \varphi^{\mathrm{st}}). \end{aligned} \tag{4.6}$$

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

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

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

Then

$$\begin{aligned}
&\check{K}(v^{\mathrm{st}}, v^{\mathrm{wk}}) \\
&= \frac{1}{2} \begin{bmatrix} (v^{\mathrm{st}})^T (v^{\mathrm{wk}})^T \end{bmatrix} \left(\begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ \mathrm{Id} \end{bmatrix} \tilde{C}^{-1} \begin{bmatrix} -B^T A^{-1} & \mathrm{Id} \end{bmatrix} \right) \begin{bmatrix} v^{\mathrm{st}} \\ v^{\mathrm{wk}} \end{bmatrix} \\
&= \frac{1}{2} v^{\mathrm{st}} \cdot A^{-1} v^{\mathrm{st}} + \frac{1}{2} (v^{\mathrm{wk}} - B^T A^{-1} v^{\mathrm{st}}) \cdot \tilde{C}^{-1} (v^{\mathrm{wk}} - B^T A^{-1} v^{\mathrm{st}}) \\
&= \check{K}^{\mathrm{st}}(v^{\mathrm{st}}) + \frac{1}{2} (v^{\mathrm{wk}} - B^T A^{-1} v^{\mathrm{st}}) \cdot \tilde{C}^{-1} (v^{\mathrm{wk}} - B^T A^{-1} v^{\mathrm{st}}),
\end{aligned}$$

and (4.4) follows.

Moreover,

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

We obtain (4.6). □

The Euler-Lagrange flow of L^s satisfies the following estimates expressed in notations of Section 2.3.

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

1. *There exists a constant $M_1 = M_1(\mathcal{B}^{\mathrm{st}}, Q, \kappa, q)$ such that*
$$\|\ddot{\gamma}^{\mathrm{st}} - A\partial_{\varphi^{\mathrm{st}}} U^{\mathrm{st}}(\gamma^{\mathrm{st}})\|_{C^0} \leqslant M_1(\mu(\mathcal{B}^{\mathrm{wk}}))^{-(q-1)}.$$
2. *There exists a constant $M_2 = M_2(\mathcal{B}^{\mathrm{st}}, Q, \kappa, q, \|U^{\mathrm{st}}\|)$ such that*
$$\|\ddot{\gamma}^{\mathrm{st}}\|_{C^0} \leqslant M_2.$$

Proof. — We note that for L^s, we have

$$\ddot{\gamma}^{\mathrm{st}} = A\partial_{\varphi^{\mathrm{st}}} U + B\partial_{\varphi^{\mathrm{wk}}} U = A\partial_{\varphi^{\mathrm{st}}} U^{\mathrm{st}} + \left(A\partial_{\varphi^{\mathrm{st}}} U^{\mathrm{wk}} + B\partial_{\varphi^{\mathrm{wk}}} U^{\mathrm{wk}}\right)$$

using that U^{st} is independent of φ^{wk}. We now use item 2, 3 of Lemma 4.1, to get item 1.

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

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. In Sections 5.3–5.5, we prove Theorem 2.5 with some technical statements deferred to Section 7.

5.1. Weak KAM solutions of Tonelli Lagrangian.

— For an extensive exposition of the topic, we refer to [12].

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

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

The Lagrangians considered in this paper are Tonelli.

Minimizers. — An absolutely continuous curve $\gamma : [a,b] \to \mathbb{T}^d$ 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 $\xi : [a,b] \to \mathbb{T}^d$ 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 (cf. [12], 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 a lower bound of $b-a$ such that $\|\dot{\gamma}\| \leqslant D$ ([12] 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 [29], 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 ([19]).

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

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

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 always exists, and is invariant under the Euler-Lagrange flow (c.f [19, 6]). Hence this definition of the alpha function is equivalent to the one given in Section 2.5, where the minimization is over invariant probability measures.

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

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

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

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

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-Fenchel transform. — Define the Legendre-Fenchel transform associated to the beta function

$$\mathcal{LF}_\beta : H_1(\mathbb{T}^d, \mathbb{R}) \to \begin{array}{l}\text{the collection of nonempty,} \\ \text{compact convex subsets of } H^1(\mathbb{T}^d, \mathbb{R}),\end{array} \tag{5.1}$$

defined by

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

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

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

A piecewise C^1 curve $\gamma : I \to \mathbb{R}$ 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 $u : \mathbb{T}^d \to \mathbb{R}$ 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 $\gamma : (-\infty, 0] \to \mathbb{T}^d$ with $\gamma(0) = \varphi$.

This definition of the weak KAM solution is equivalent to the one given in Section 2.5 (see [12], 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 $h_L^T : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}$ 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 $h_L(\varphi, \psi) = \lim_{T\to\infty} h_L^T(\varphi, \psi)$. 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 [22]). Here we only introduce the projected version. Define the projected Aubry and Mañe sets as

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

The Mather set is $\tilde{\mathcal{M}}_L(c) = \overline{\bigcup_\mu \operatorname{supp}(\mu)}$ 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 relation, 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 2.5. 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 2.5.

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.

PROPOSITION 5.1. — *Fix $\mathcal{B}^{\rm st}, \kappa > 1$. For any $R > 0, q > 2$, there exists $M = M(\mathcal{B}^{\rm st}, D, R, \kappa, n)$ such that the following hold. For any*

$$(\mathcal{B}^{\rm wk}, p, U^{\rm st}, \mathcal{U}^{\rm wk}) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\rm st}) \cap \{\|U^{\rm st}\|_{C^2} \leqslant R\},$$

the Lagrangian $L^s = L_{\mathcal{H}^s(\mathcal{B}^{\rm st},\mathcal{B}^{\rm wk},p,U^{\rm st},\mathcal{U}^{\rm wk})}$, let $T \geqslant \frac{1}{2}$, $c = (c^{\rm st}, c^{\rm wk}) \in \mathbb{R}^m \times \mathbb{R}^{d-m}$ and $\gamma = (\gamma^{\rm st}, \gamma^{\rm wk}) : [0,T] \to \mathbb{T}^d$ be a minimizer of $L^s - c \cdot v$. Then for $\bar{c} = c^{\rm st} + A^{-1}Bc^{\rm wk}$, we have

$$\|\dot\gamma^{\rm st} - A\bar{c}\| \leqslant M.$$

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 $T \geqslant \frac{1}{2}$ and $c \in \mathbb{R}^d$, let $\gamma = (\gamma^{\rm st}, \gamma^{\rm wk}) : [0,T] \to \mathbb{T}^d$ be a minimizer for the Lagrangian $L^s - c \cdot v$. Then*

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

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

Proof. — Let $\gamma^{\rm st}_0 : [0,T] \to \mathbb{T}^m$ be such that

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

with $\zeta(0) = \gamma^{\rm st}(0)$, $\zeta(T) = \gamma^{\rm st}(T)$. Define $\gamma_0 = (\gamma^{\rm st}_0, \gamma^{\rm wk}_0) : [0,T] \to \mathbb{T}^d$, by

$$\gamma^{\rm wk}_0(t) = \gamma^{\rm wk}(t) - A^{-1}B\gamma^{\rm st}(t) + A^{-1}B\gamma^{\rm st}_0(t).$$

Note that

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

Using (4.6) and (5.4), we have

$$\begin{aligned}L^s-c\cdot v+\frac{1}{2}c^{\mathrm{wk}}\cdot\tilde C^{-1}c^{\mathrm{wk}}&=L^{\mathrm{st}}-\bar c\cdot v^{\mathrm{st}}\\&+\frac12(v^{\mathrm{wk}}-B^TA^{-1}v^{\mathrm{st}}-\tilde C^{\mathrm{wk}})\cdot\tilde C(v^{\mathrm{wk}}-B^TA^{-1}v^{\mathrm{st}}-\tilde C^{\mathrm{wk}})+U^{\mathrm{wk}}\end{aligned} \tag{5.3}$$

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

$$\int_0^T(L^s-c\cdot v)(\gamma,\dot\gamma)dt\leqslant\int_0^T(L^s-c\cdot v)(\gamma_0,\dot\gamma_0)dt.$$

By (5.3), we have

$$\begin{aligned}&\int_0^T(L^{\mathrm{st}}-\bar c\cdot v^{\mathrm{st}})(\gamma^{\mathrm{st}},\dot\gamma^{\mathrm{st}})dt+\int_0^T U^{\mathrm{wk}}(\gamma(t))dt\\&\qquad+\int_0^T\frac12(\dot\gamma^{\mathrm{wk}}-B^TA^{-1}\dot\gamma^{\mathrm{st}}-\tilde C^{\mathrm{wk}})\cdot\tilde C(\dot\gamma^{\mathrm{wk}}-B^TA^{-1}\dot\gamma^{\mathrm{st}}-\tilde C^{\mathrm{wk}})dt\\&\quad\leqslant\int_0^T(L^{\mathrm{st}}-\bar c\cdot v^{\mathrm{st}})(\gamma_0^{\mathrm{st}},\dot\gamma_0^{\mathrm{st}})dt+\int_0^T U^{\mathrm{wk}}(\gamma_0(t))dt\\&\qquad+\int_0^T\frac12(\dot\gamma_0^{\mathrm{wk}}-B^TA^{-1}\dot\gamma_0^{\mathrm{st}}-\tilde C^{\mathrm{wk}})\cdot\tilde C(\dot\gamma_0^{\mathrm{wk}}-B^TA^{-1}\dot\gamma_0^{\mathrm{st}}-\tilde C^{\mathrm{wk}})\,dt.\end{aligned}$$

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

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

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 $T\in[\frac12,1)$.

We first produce an upper bound for

$$\min_\zeta\int_0^T(L^{\mathrm{st}}-\bar c\cdot v^{\mathrm{st}}+\frac12\bar c\cdot A\bar c)(\zeta,\dot\zeta)dt.$$

By completing the squares as in Lemma 4.2, we have

$$L^{\mathrm{st}}-\bar c\cdot v^{\mathrm{st}}+\frac12\bar c\cdot A\bar c=\frac12(v^{\mathrm{st}}-A\bar c)\cdot A^{-1}(v^{\mathrm{st}}-A\bar c)+U^{\mathrm{st}}(\varphi^{\mathrm{st}}). \tag{5.4}$$

We then take

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

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

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

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

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

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

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

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

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

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

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

Then

$$\left\| \frac{1}{T}\int_0^T (\dot{\gamma}^{\mathrm{st}} - A\bar{c})\, dt \right\|^2 \leqslant \frac{1}{T^2}\int_0^T \|\dot{\gamma}^{\mathrm{st}} - A\bar{c}\|^2\, dt \leqslant 4M_1. \tag{5.5}$$

Moreover, from Lemma 4.3,

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

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

$$\left\| (\dot{\gamma}^{\mathrm{st}} - A\bar{c}) - \frac{1}{T}\int_0^T (\dot{\gamma}^{\mathrm{st}} - A\bar{c}) dt \right\|_{L^\infty} \leqslant \|\ddot{\gamma}^{\mathrm{st}}\|_{L^\infty} \leqslant QM_2.$$

Combine with (5.5) and we conclude the proof. □

5.3. Approximate Lipschitz property of weak KAM solutions. — The weak KAM solutions of the slow Hamiltonian is Lipschitz, however, it is not clear if the Lipschitz constant is bounded as $\mu(\mathcal{B}^{\mathrm{wk}}) \to \infty$. To get uniform estimates, we consider the following weaker notion.

DEFINITION. — *For $C, \delta > 0$, a function $u : \mathbb{R}^d \to \mathbb{R}$ is called (C, δ) approximately Lipschitz if*

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

For $u : \mathbb{T}^d \to \mathbb{R}$, 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 $\mathcal{B}^{\mathrm{st}}, \kappa > 1$. Assume that $q > 2(d-m)$. For $R > 0$, there exists a constant $M = M(\mathcal{B}^{\mathrm{st}}, D, \kappa, q, R) > 0$, such that for all*

$$(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}}) \cap \{\|U^{\mathrm{st}}\| \leqslant R\},$$

and

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

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

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

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

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

The proof of these statements are deferred to Section 7.

5.4. The alpha function and rotation vector estimate. — In this section we provide a few useful estimates in weak KAM theory and prove Theorem 2.5 using Propositions 5.3 and 5.4. Recall that the notations $c = (c^{\mathrm{st}}, c^{\mathrm{wk}}), \bar{c} = c^{\mathrm{st}} + A^{-1}Bc^{\mathrm{wk}}$.

Proposition 5.5. — *With these notations we have the following estimate:*

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

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

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

On the other hand, let μ^{st} be an ergodic minimal measure for $L^{\mathrm{st}} - \bar{c}\cdot v^{\mathrm{st}}$. For an L^{st}-Euler-Lagrange orbit $\varphi^{\mathrm{st}}(t)$ in the support of μ^{st}, and any $\varphi^{\mathrm{wk}}_0 \in \mathbb{T}^{d-m}$, define

$$\varphi^{\mathrm{wk}}(t) = \varphi^{\mathrm{wk}}_0 + B^T A^{-1}\varphi^{\mathrm{st}}(t) + \tilde{C}c^{\mathrm{wk}}t, \quad t \in \mathbb{R} \tag{5.7}$$

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

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

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

□

The following proposition establishes relations between rotation vectors 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 $(\rho^{\rm st}, \rho^{\rm wk})$ denote its rotation vector. Then*

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

and

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

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

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

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

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

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

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

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

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

which is the first conclusion.

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

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

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

5.5. Convergence of weak KAM solutions. — We now prove Theorem 2.5. Fix $\mathcal{B}^{\mathrm{st}}$ and $\kappa > 1$.

For $\nu \in \mathbb{N}$, let $(\mathcal{B}^{\mathrm{wk}}_\nu, p_\nu, U^{\mathrm{st}}_\nu, \mathcal{U}^{\mathrm{wk}}_\nu) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}})$ and $c_\nu = (c^{\mathrm{st}}_\nu, c^{\mathrm{wk}}_\nu)$ be a sequence satisfying the assumption of the theorem, namely, $\mu(\mathcal{B}^{\mathrm{wk}}_\nu) \to \infty$, $p_\nu \to p_0$, $U^{\mathrm{st}}_\nu \to U^{\mathrm{st}}_0$ in C^2, and $c^{\mathrm{st}}_\nu + A_\nu^{-1}B_\nu c^{\mathrm{wk}}_\nu \to \bar{c}$. Let us fix the notations

$$\begin{aligned} H^s_\nu &= \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}_\nu, p_\nu, U^{\mathrm{st}}_\nu, \mathcal{U}^{\mathrm{wk}}_\nu), \quad L^s_\nu = L_{H^s_\nu}, \quad \alpha_\nu = \alpha_{H^s_\nu}(c_\nu), \\ H^{\mathrm{st}}_\nu &= \mathcal{H}^{\mathrm{st}}(\mathcal{B}^{\mathrm{st}}, p_\nu, U^{\mathrm{st}}_\nu), \quad L^{\mathrm{st}}_\nu = L_{H^{\mathrm{st}}_\nu}. \end{aligned} \tag{5.9}$$

Item 1. — Let u_ν be the weak KAM solution to $L^s_\nu - c_\nu \cdot v$. We first show the sequence $\{u_\nu\}$ is equi-continuous.

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

$$|u_\nu(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) - u_\nu(\psi^{\mathrm{st}}, \psi^{\mathrm{wk}})| \leqslant M^*\|\varphi^{\mathrm{st}} - \psi^{\mathrm{st}}\| + \delta_\nu\|\varphi^{\mathrm{wk}} - \psi^{\mathrm{wk}}\| + 2\delta_\nu,$$

where $\delta_\nu = M^*(\mu(\mathcal{B}^{\mathrm{wk}}_\nu))^{-\frac{q}{2}-d+m}$.

Since $\delta_\nu \to 0$ as $\nu \to \infty$, for any $0 < \varepsilon < 1$ there exists $M > 0$ such that for all $\nu > M$, $3\delta_\nu < \frac{\varepsilon}{2}$. It follows that if $\|\varphi - \psi\| < \frac{\varepsilon}{2M^*} < 1$, then

$$|u_\nu(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) - u_\nu(\psi^{\mathrm{st}}, \psi^{\mathrm{wk}})| < \varepsilon.$$

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

$$|u_\nu(\varphi) - u(\psi)| < \varepsilon, \quad \text{ if } 1 \leqslant \nu \leqslant M, \|\varphi - \psi\| < \sigma.$$

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

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

$$\lim_{\nu\to\infty} \sup_{\varphi^{\mathrm{st}}} (\max_\nu u_\nu(\varphi^{\mathrm{st}}, \cdot) - \min_\nu u_\nu(\varphi^{\mathrm{st}}, \cdot)) \leqslant 2 \lim_{\nu\to\infty} \delta_\nu = 0,$$

therefore, u is independent of φ^{wk}.

Item 3. — From item 2, there exists $u^{\mathrm{st}}(\varphi^{\mathrm{st}}) = \lim_{\nu\to\infty} u_\nu(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}})$. We show u^{st} is a weak KAM solution of $L_0^{\mathrm{st}} - \bar{c}\cdot v^{\mathrm{st}}$. Recall the notations in (5.9), we have $L_\nu^{\mathrm{st}} \to L_0^{\mathrm{st}}$ in C^2.

We first show that u^{st} is dominated by $L_0^{\mathrm{st}} - \bar{c}\cdot v^{\mathrm{st}}$. Let $\xi^{\mathrm{st}} : [0,T] \to \mathbb{T}^m$ be an extremal curve of L_0^{st}. In the same way as (5.7) in the proof of Proposition 5.5, we define $\xi_\nu = (\xi_\nu^{\mathrm{st}}, \xi_\nu^{\mathrm{wk}}) : [a,b] \to \mathbb{T}^m$ such that $\xi_\nu^{\mathrm{st}}(a) = \xi^{\mathrm{st}}(a)$, $\xi_\nu^{\mathrm{st}}(b) = \xi^{\mathrm{st}}(b)$ and $\dot{\xi}_\nu^{\mathrm{st}} - B_\nu^T A_\nu^{-1}\dot{\xi} - \tilde{C}_\nu c_\nu^{\mathrm{wk}} = 0$. Since u_ν are dominated by $L_\nu^s - c_\nu \cdot v + \alpha_\nu$ (see (5.9)), we have

$$\begin{aligned} u_\nu(\xi_\nu(b)) - u_\nu(\xi_\nu(a)) &\leqslant \int_a^b (L_\nu^s - c_\nu \cdot v^s + \alpha_\nu(\xi_\nu, \dot{\xi}_\nu)dt \\ &= \int_a^b (L_\nu^{\mathrm{st}} - \bar{c}_\nu \cdot v^{\mathrm{st}})(\xi_\nu^{\mathrm{st}}, \dot{\xi}_\nu^{\mathrm{st}})dt \\ &\quad + \int_a^b (U_\nu^{\mathrm{wk}}(\xi_\nu) + \alpha_\nu - \frac{1}{2}\tilde{C}_\nu c_\nu^{\mathrm{wk}} \cdot c_\nu^{\mathrm{wk}})dt, \end{aligned}$$

where the equality is due to $\dot{\xi}_\nu^{\mathrm{st}} - B_\nu^T A_\nu^{-1}\dot{\xi}_\nu - \tilde{C} c_\nu^{\mathrm{wk}} = 0$. Using the fact that $\|U_\nu^{\mathrm{wk}}\|_{C^0} \to 0$, $L_\nu^{\mathrm{st}} \to L_0^{\mathrm{st}}$, and from Proposition 5.5, $\alpha_\nu - \frac{1}{2}\tilde{C}_\nu c_\nu^{\mathrm{wk}}\cdot c_\nu^{\mathrm{wk}} \to \alpha_{H^{\mathrm{st}}}(\bar{c})$ as $\nu \to \infty$, we get

$$u^{\mathrm{st}}(\xi^{\mathrm{st}}(b)) - u^{\mathrm{st}}(\xi^{\mathrm{st}}(a)) \leqslant \int_a^b (L_0^{\mathrm{st}} - \bar{c}\cdot v^{\mathrm{st}} + \alpha_{H^{\mathrm{st}}}(\bar{c}))dt. \tag{5.10}$$

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

Secondly, we show that for any $\varphi^{\mathrm{st}} \in \mathbb{T}^m$, there exists a $(u^{\mathrm{st}}, L_0^{\mathrm{st}}, \bar{c})$-calibrated curve $\gamma^{\mathrm{st}} : (-\infty, 0] \to \mathbb{T}^m$ with $\gamma^{\mathrm{st}}(0) = \varphi^{\mathrm{st}}$.

Because u_ν are weak KAM solutions of $L_\nu^s - c_\nu \cdot v$, for each ν there exists a $(u_\nu, L_\nu^s - c_\nu\cdot v, \alpha_\nu)$-calibrated curve $\gamma_\nu = (\gamma_\nu^{\mathrm{st}}, \gamma_\nu^{\mathrm{wk}}) : (-\infty, 0] \to \mathbb{T}^d$. By Proposition 5.1, all γ_ν^{st} are uniformly Lipschitz, so there exists a subsequence that converges in $C^1_{loc}((-\infty,0],\mathbb{T}^d)$. Assume without loss of generality that $\gamma_\nu^{\mathrm{st}} \to \gamma^{\mathrm{st}}$, since $\gamma_\nu = (\gamma_\nu^{\mathrm{st}}, \gamma_\nu^{\mathrm{wk}})$ is extremal for L_ν^s, we have

$$\ddot{\gamma}_\nu^{\mathrm{st}} = \frac{d}{dt}(A_\nu I^{\mathrm{st}} + B_\nu I^{\mathrm{wk}}) = A_\nu \partial_{\varphi^{\mathrm{st}}} U_\nu^{\mathrm{st}} + B_\nu \partial_{\varphi^{\mathrm{st}}} U_\nu^{\mathrm{wk}}.$$

By our assumption, as $i \to \infty$, $A_\nu \to A := \partial^2_{v^{\mathrm{st}}v^{\mathrm{st}}} L_0^{\mathrm{st}}$, and by Lemma 4.3 $\|B_\nu\|\|U_\nu^{\mathrm{wk}}\|_{C^2} \to 0$, we have

$$\ddot{\gamma}^{\mathrm{st}} = A\partial_{\varphi^{\mathrm{st}}} U^{\mathrm{st}}(\gamma^{\mathrm{st}}), \tag{5.11}$$

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

On the other hand, since γ_ν are $(u_\nu, L^s_\nu - c_\nu\cdot v, \alpha_\nu)$ calibrated, for any $[a,b]\subset(-\infty,0]$,

$$u_\nu(\gamma_\nu(b)) - u_\nu(\gamma_\nu(a)) = \int_a^b (L^s_\nu - c_\nu\cdot v^s + \alpha_\nu(\gamma_\nu,\dot\gamma_\nu))dt$$
$$\geqslant \int_a^b (L^{\mathrm{st}} - \bar c_\nu\cdot v^{\mathrm{st}})(\gamma_\nu^{\mathrm{st}},\dot\gamma_\nu^{\mathrm{st}})dt + \int_a^b (U^{\mathrm{wk}}_\nu + \alpha_{H^s_\nu}(c_\nu) - \frac{1}{2}\tilde C_\nu c^{\mathrm{wk}}_\nu\cdot c^{\mathrm{wk}}_\nu)(\gamma_\nu,\dot\gamma_\nu)dt.$$

Take the limit again to get

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

Because γ^{st} is an L^{st} extremal curve (see (5.11)), (5.10) hold for γ^{st}. Combining with the last displayed formula, (5.10) becomes an equality. Then γ^{st} is a calibrated curve for $L^{\mathrm{st}} - \bar c\cdot v^{\mathrm{st}} + \alpha_{H^{\mathrm{st}}}(\bar c)$, 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 $\mathcal{B}^{\mathrm{st}}$ and $\kappa > 1$. Assume that $(\mathcal{B}^{\mathrm{wk}}_\nu, p_\nu, U^{\mathrm{st}}_\nu, \mathcal{U}^{\mathrm{wk}}_\nu)$ satisfies the assumptions of Theorem 2.5. Denote H^s_ν, H^{st}_0, L^s_ν and L^{st}_0 as in the previous section (see (5.9)).*

1. *Any limit point of $\varphi_\nu\in\mathcal{N}_{H^s_\nu}(c_\nu)$ is contained in $\mathcal{N}_{H^{\mathrm{st}}_0}(\bar c)\times\mathbb{T}^{d-m}$.*
2. *If $\mathcal{A}_{H^{\mathrm{st}}_0}(\bar c)$ contains only finitely many static classes, then any limit point of $\varphi_\nu\in\mathcal{A}_{H^s_\nu}(c_\nu)$ is contained in $\mathcal{A}_{H^{\mathrm{st}}_0}(\bar c)\times\mathbb{T}^{d-m}$.*
3. *Assume that $\mathcal{A}_{H^{\mathrm{st}}}(\bar c)$ contains only one static class. Let $\varphi_\nu = (\varphi^{\mathrm{st}}_\nu, \varphi^{\mathrm{wk}}_\nu)\in\mathcal{A}_{H^s_\nu}(c_\nu)$ be such that $\varphi^{\mathrm{st}}_\nu\to\varphi^{\mathrm{st}}\in\mathcal{A}_{H^{\mathrm{st}}_0}(\bar c)$. Then for any $\psi = (\psi^{\mathrm{st}},\psi^{\mathrm{wk}})\in\mathbb{T}^d$,*
$$\lim_{\nu\to\infty} h_{L^s_\nu, c_\nu}(\varphi,\psi) = h_{L^{\mathrm{st}}_0,\bar c}(\varphi^{\mathrm{st}},\psi^{\mathrm{st}}).$$
4. *Let $(\rho^{\mathrm{st}}_\nu, \rho^{\mathrm{wk}}_\nu)$ be the rotation vector of any c_ν-minimal measure of L^s_ν. Then we have*
$$\lim_{\nu\to\infty}(\rho^{\mathrm{wk}}_\nu - B^T_\nu A^{-1}_\nu\rho^{\mathrm{st}}_\nu - \tilde C_\nu c^{\mathrm{wk}}_\nu) = 0,$$
and any accumulation point ρ of ρ^{st}_ν is contained in the set $\partial\alpha_{H^{\mathrm{st}}}(\bar c)$.

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 the barrier function. — We first state an alternate definition of the Aubry and Mañe sets due to Fathi (see also [6]). Let u be a weak KAM solution for the Lagrangian L. We define $\overline{\mathcal{G}}(L,u)$ 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 $\gamma : (-\infty, 0] \to \mathbb{T}^d$, with $(\varphi, v) = (\gamma(0), \dot\gamma(0))$. Let ϕ_t denote the Euler-Lagrange flow of L, then

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

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^s_ν, L^{st}_0, c_ν, $\bar c$ be as in the assumption. The strategy of the proof is similar to the one in [7].

Lemma 6.2. — *Let u_ν be a weak KAM solution of $L^s_\nu - c_\nu \cdot v$. Assume that $(\varphi_\nu, v_\nu) \in \tilde{\mathcal{I}}(L^s_\nu - c_\nu\cdot v, u_\nu)$ satisfies $(\varphi_\nu, v_\nu) \to (\varphi, v) = (\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}, v^{\mathrm{st}}, v^{\mathrm{wk}})$, and $u_\nu(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) \to u^{\mathrm{st}}(\varphi^{\mathrm{st}})$. Then*

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

Proof. — We first show that $(\varphi_\nu, v_\nu) \in \overline{\mathcal{G}}(L^s_\nu - c_\nu\cdot v, u_\nu)$ implies $(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) \in \overline{\mathcal{G}}(L^{\mathrm{st}} - \bar c\cdot v^{\mathrm{st}}, u^{\mathrm{st}})$. Indeed, there exists $\gamma_\nu : (-\infty, 0] \to \mathbb{T}^d$, each $(u_\nu, L^s_\nu - c_\nu\cdot v, \alpha_{L^s_\nu}(c_\nu))$-calibrated, with $(\gamma_\nu, \dot\gamma_\nu)(0) = (\varphi, v)$. We follow the same line as proof of item 3 in Theorem 2.5 (Section 5), then by restricting to a subsequence, γ^{st}_ν converges in $C^1_{loc}((-\infty,0],\mathbb{T}^d)$ to a $(u^{\mathrm{st}}, L^{\mathrm{st}}_0 - \bar c\cdot v^{\mathrm{st}}, \alpha_{H^{\mathrm{st}}_0}(\bar c))$-calibrated curve γ^{st}. In particular, $(\gamma^{\mathrm{st}}_\nu, \dot\gamma^{\mathrm{st}}_\nu) \to (\gamma^{\mathrm{st}}, \dot\gamma^{\mathrm{st}})$, which implies $(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) \in \overline{\mathcal{G}}(L^{\mathrm{st}} - \bar c\cdot v^{\mathrm{st}}, u^{\mathrm{st}})$.

Let ϕ^ν_t denote the Euler-Lagrange flow of L^s_ν, and ϕ^{st}_t the flow for L^{st}. Let π denote the projection to the strong components $(\varphi^{\mathrm{st}}, v^{\mathrm{st}})$, then from Lemma 4.3 $\pi\phi^\nu_t \to \phi^{\mathrm{st}}_t$ uniformly. As a result for a fixed $T > 0$ and $(\varphi_\nu, v_\nu) \in \tilde{\mathcal{I}}(L^{\mathrm{st}}_\nu - c_\nu\cdot v, u_\nu)$, we have

$$(\varphi_\nu, v_\nu) = (\varphi^{\mathrm{st}}_\nu, \varphi^{\mathrm{wk}}_\nu, v^{\mathrm{st}}_\nu, v^{\mathrm{wk}}_\nu) \in \phi^\nu_{-T}\left(\overline{\mathcal{G}}(L^s_\nu - c_\nu\cdot v, u_\nu)\right),$$

hence $(\varphi^{\mathrm{st}}_\nu, v^{\mathrm{st}}_\nu) \to (\varphi^{\mathrm{st}}, v^{\mathrm{st}}) \in \varphi^{\mathrm{st}}_{-T}\left(\overline{\mathcal{G}}(L^{\mathrm{st}} - \bar c\cdot v^{\mathrm{st}}, u^{\mathrm{st}})\right)$. Since $T > 0$ is arbitrary, we obtain $(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) \in \tilde{\mathcal{I}}(L^{\mathrm{st}}_0 - \bar c\cdot v^{\mathrm{st}}, u^{\mathrm{st}})$. □

Proof of Proposition 6.1, part I. — We first prove item 1.

Suppose $\tilde\varphi_\nu \in \tilde{\mathcal{N}}_{H^s_\nu}(c_\nu)$, then there exists weak KAM solutions u_ν of $L^s_\nu - c_\nu\cdot v$, such that $(\varphi_\nu, v_\nu) \in \tilde{\mathcal{I}}(L^s_\nu - c_\nu\cdot v, u_\nu)$. By Theorem 2.5, after restricting to a subsequence, we have $u_\nu(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) \to u^{\mathrm{st}}(\varphi^{\mathrm{st}})$. By Lemma 6.2, $(\varphi^{\mathrm{st}}_\nu, v^{\mathrm{st}}_\nu) \to (\varphi^{\mathrm{st}}, v^{\mathrm{st}})$ implies $(\varphi^{\mathrm{st}}, v^{\mathrm{st}}) \in \tilde{\mathcal{I}}(L^{\mathrm{st}}_0 - \bar c\cdot v^{\mathrm{st}}, u^{\mathrm{st}}) \subset \tilde{\mathcal{N}}_{H^{\mathrm{st}}_0}(\bar c)$.

For item 3, suppose $\varphi_\nu = (\varphi_\nu^{\mathrm{st}}, \varphi_\nu^{\mathrm{wk}}) \in \mathcal{A}_{H_\nu^s}(c_\nu)$ satisfies $\varphi_\nu^{\mathrm{st}} \to \varphi^{\mathrm{st}} \in \mathcal{A}_{H_0^{\mathrm{st}}}(\bar{c})$. Then $h_{L_\nu^s, c_\nu}(\varphi_\nu, \cdot)$ is a weak KAM solution of $L_\nu^s - c\cdot v$ (see [12], Theorem 5.3.6). By Theorem 2.5, by restricting to a subsequence, there exists a weak KAM solution u^{st} of $L_0^{\mathrm{st}} - \bar{c}\cdot v^{\mathrm{st}}$ such that

$$\lim_{\nu\to\infty} h_{L_\nu^s, c_\nu}(\varphi_\nu, \psi^{\mathrm{st}}, \psi^{\mathrm{wk}}) - h_{L_\nu^s, c_\nu}(\varphi_\nu, 0, 0) = u^{\mathrm{st}}(\psi^{\mathrm{st}}).$$

We may further assume that $h_{L_\nu^s, c_\nu}(\varphi_\nu, 0, 0) \to C \in \mathbb{R}$. Since $\mathcal{A}_{H_0^{\mathrm{st}}}(\bar{c})$ has only one static class, there exists a constant $C_1 > 0$ such that

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

Using the fact that $\varphi_\nu \in \mathcal{A}_{H_\nu^s}(c_\nu)$, we get $h_{L_\nu^s, c_\nu}(\varphi_\nu, \varphi_\nu) = 0$. Taking the limit,

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

Therefore,

$$\lim_{\nu\to\infty} h_{L_\nu^s, c_\nu}(\varphi_\nu^{\mathrm{st}}, \varphi_\nu^{\mathrm{wk}}, \psi^{\mathrm{st}}, \psi^{\mathrm{wk}}) = h_{L^{\mathrm{st}}, \bar{c}}(\varphi^{\mathrm{st}}, \psi^{\mathrm{st}}).$$

Item 4: Let $\rho_\nu = (\rho_\nu^{\mathrm{st}}, \rho_\nu^{\mathrm{wk}})$ be the rotation vector of minimal measures of $L_\nu^s - c_\nu \cdot v$, then from Proposition 5.6,

$$\lim_{\nu\to\infty} \rho_\nu^{\mathrm{wk}} - B_\nu^T A_\nu^{-1} \rho_\nu^{\mathrm{st}} - \tilde{C}_\nu c_\nu^{\mathrm{wk}} = 0.$$

Moreover, assume that $\rho_\nu^{\mathrm{st}} \to \rho^{\mathrm{st}} \in \mathbb{R}^m$, then by taking the limit in the second conclusion of Proposition 5.6, we get

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

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

6.2. Semi-continuity of the Aubry set. — Our strategy of the proof mostly follow [7].

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 $x_0, \ldots, x_N \in \mathcal{X}$ and $T_0, \ldots, T_{N-1} \geqslant T$, 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 [11]).

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 due to Bernard ([7]).

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* $x_\nu : [0, T_\nu] \to \mathcal{X}$ *is said to accumulate locally uniformly to* $(\mathcal{X}, \phi_t)$ *if for any sequence* $S_\nu \in [0, T_\nu]$*, the curves* $x_\nu(t + S_\nu)$ *has a subsequence which converges uniformly on compact sets to a trajectory of* ϕ_t*.*

LEMMA 6.4 ([7]). — *Suppose* $x_\nu : [0, T_\nu] \to \mathcal{X}$ *accumulates locally uniformly to* (X, ϕ_t)*,* $x_\nu(0) \to x$ *and* $x_\nu(T_\nu) \to y$*, then* $x \mathfrak{C}_X y$*.*

Proof of Proposition 6.1, part II. — We prove item 2. Let $\varphi_\nu = (\varphi_\nu^{\mathrm{st}}, \varphi_\nu^{\mathrm{wk}}) \in \mathcal{A}_{H_\nu^s}(c_\nu)$ and $\varphi_\nu^{\mathrm{st}} \to \varphi^{\mathrm{st}}$, we show that $\varphi^{\mathrm{st}} \in \mathcal{A}_{H_0^{\mathrm{st}}}(\bar{c})$. According to Proposition 6.3, $\varphi_\nu \mathfrak{C}_u \varphi_\nu$. Let $\tilde{\varphi}_\nu$ be the unique point in $\tilde{\mathcal{A}}_{H_\nu^s}(c_\nu)$ projecting to φ_ν, then there exists weak KAM solutions u_ν of $L_\nu^s - c_\nu \cdot v$, such that $\tilde{\varphi}_\nu \mathfrak{C} \tilde{\varphi}_\nu$ in $\overline{\mathcal{G}}(L_\nu^s - c_\nu \cdot v, u_\nu)$. Fix $\varepsilon_\nu \to 0$ and $M_\nu \to \infty$, then for each ν, there exists

$$T_{\nu,1} < \cdots < T_{\nu,N_\nu}, \quad T_{\nu,j+1} - T_{\nu,j} > M_\nu,$$

and a piecewise C^1 curve $\gamma_\nu = (\gamma_\nu^{\mathrm{st}}, \gamma_\nu^{\mathrm{wk}}) : [0, T_\nu] \to \mathbb{T}^d$, satisfying

1. $\gamma_\nu | (T_{\nu,j}, T_{\nu,j+1})$ satisfies the Euler-Lagrange equation of L_ν^s;
2. $d\big(((\gamma_\nu(T_{\nu,j}-), \dot{\gamma}_\nu(T_{\nu,j}-)),\ ((\gamma_\nu(T_{\nu,j}+), \dot{\gamma}_\nu(T_{\nu,j}+))\big) < \varepsilon_\nu$.

Using Lemma 4.3, the projection of the Euler-Lagrange flow of L_ν^s to $(\varphi^{\mathrm{st}}, v^{\mathrm{st}})$ 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 $(\gamma_\nu^{\mathrm{st}}, \dot{\gamma}_\nu^{\mathrm{st}})$ accumulates locally uniformly to

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

where ϕ_t^{st} is the Euler-Lagrange flow of L_0^{st}. Therefore, $\varphi_\nu^{\mathrm{st}} \to \varphi^{\mathrm{st}}$ impies $\varphi^{\mathrm{st}} \mathfrak{C}_u \varphi^{\mathrm{st}}$. Using Proposition 6.3 again, we get $\varphi^{\mathrm{st}} \in \mathcal{A}_{H_0^{\mathrm{st}}}(\bar{c})$. □

7. Technical estimates on weak KAM solutions

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

7.1. Approximate Lipschitz property in the strong component. — In this section we prove Proposition 5.4 using Proposition 5.3. Proposition 5.3 is proved 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 $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ be a Tonelli Hamiltonian, $T \geqslant 1$, and $\gamma : [0,T] \to \mathbb{T}^d$ be an extremal curve. Then for any $1 \leqslant i \leqslant d$, $h > 0$, and a unit vector $f \in \mathbb{R}^d$,*

$$\int_0^T L(\gamma + \frac{th}{T}f, \dot\gamma + \frac{h}{T}f)dt - \int_0^T L(\gamma, \dot\gamma)dt$$
$$\leqslant (\partial_v L(\gamma(T), \dot\gamma(T)) \cdot f)h + \left(\|f \cdot (\partial^2_{vv}L)f\| \frac{1}{T} + \|f \cdot (\partial^2_{\varphi v}L)f\| + T\|f \cdot (\partial^2_{\varphi\varphi}L)f\| \right) h^2.$$

Proof. — We compute

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

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. □

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

Lemma 7.2. — *For $C, \delta > 0$, assume that all $\varphi \in \mathbb{T}^d$ the function $u : \mathbb{T}^d \to \mathbb{R}$ satisfies the following condition: there exists $l \in \mathbb{R}^d$ such that*

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

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

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

$$0 = u(\varphi + e_i) - u(\varphi) \leqslant -|l_i| + C + \delta,$$

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

$$u(\varphi + y) - u(\varphi) \leqslant (\sqrt{d}(C+\delta) + C\|y\|)\|y\| + \delta < 2\sqrt{d}(C+\delta)\|y\| + \delta. \quad \square$$

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

(7.1)

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

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

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

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

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

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

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

   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

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

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

$$\dot\xi_0^{\mathrm{wk}} - B^T A^{-1} \xi^{\mathrm{st}} - \tilde C c^{\mathrm{wk}} = 0.$$

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

$$\begin{aligned} u(\varphi^{\mathrm{st}}+y, \xi^{\mathrm{wk}}(0)) &\leqslant u(\gamma(-T)) + \int_{-T}^{0} (L^s - c\cdot v + \alpha_{H^s}(c))(\xi,\dot\xi)dt \\ &= u(\gamma(-T)) + \int_{-T}^{0} (L^{\mathrm{st}} - \bar c\cdot v^{\mathrm{st}})(\xi^{\mathrm{st}},\dot\xi^{\mathrm{st}})dt \\ &\quad + \int_{-T}^{0} \frac12 (\dot\xi^{\mathrm{wk}} - B^T A^{-1}\dot\xi^{\mathrm{st}} - \tilde C c^{\mathrm{wk}}) \\ &\qquad \cdot \tilde C^{-1}(\dot\xi^{\mathrm{wk}} - B^T A^{-1}\dot\xi^{\mathrm{st}} - \tilde C c^{\mathrm{wk}})dt \\ &\quad + (\alpha_{H^s}(c) - \frac12 c^{\mathrm{wk}}\cdot \tilde C^{-1}c^{\mathrm{wk}})T + \int_{-T}^{0} U^{\mathrm{wk}}(\xi)dt \end{aligned}$$

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

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

From (7.2) we get

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

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

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

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

We now invoke Proposition 5.3 to get

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

where $\delta = M_1^*\mu(\mathcal{B}^{\mathrm{wk}})^{-\frac q2 - d + m}$ for some $M_1^* = M_1^*(\mathcal{B}^{\mathrm{st}}, Q, \kappa, q, R)$. Combine all the estimates, we get

$$u(\varphi^{\mathrm{st}}+y, \varphi^{\mathrm{wk}}) - u(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) \leqslant l\cdot y + (\|A^{-1}\| + \|U^{\mathrm{st}}\|_{C^2})\|y\|^2 + 2\delta + 2\|U^{\mathrm{wk}}\|^0_C.$$

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

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

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

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

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

7.2. Filtrated decomposition of the slow Lagrangian. — For the proof of Proposition 5.3, we need a filtrated decomposition of the Lagrangian L^s which treat all φ_i^{wk}, $1 \leqslant i \leqslant d-m$ 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*

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

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

Proof. — The proof is direct calculation. □

We write $H^s(\varphi, I) = K(I) - U(\varphi) = K(I) - U^{\mathrm{st}}(\varphi^{\mathrm{st}}) - U^{\mathrm{wk}}(\varphi)$ and $S = \partial^2_{II} K$. We describe a coordinate change block diagonalizing $\partial^2_{II} K$. Write S in the following block form

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

and for each $1 \leqslant i \leqslant d-m-1$, further decompose each X_{i+1} as

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

Note that in this notation, $X_1 = \partial^2_{I^{\mathrm{st}} I^{\mathrm{st}}} K = A$ (see (2.4)).

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

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

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

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

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

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

Let

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

then recursive computation yields

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

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

$$\lfloor v \rfloor_0 = v^{\mathrm{st}}, \quad \lfloor v \rfloor_i = (v^{\mathrm{st}}, v_1^{\mathrm{wk}}, \dots, v_i^{\mathrm{wk}}),\ 1 \leqslant i \leqslant d-m. \tag{7.6}$$

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*

$$w = \begin{bmatrix} w^{\mathrm{st}} \\ w_1^{\mathrm{wk}} \\ \vdots \\ w_{d-m}^{\mathrm{wk}} \end{bmatrix} = \begin{bmatrix} v^{\mathrm{st}} \\ v_1^{\mathrm{wk}} - y_1^T X_1^{-1} \lfloor v \rfloor_0 \\ \vdots \ v_{d-m}^{\mathrm{wk}} - y_{d-m}^T X_{d-m}^{-1} \lfloor v \rfloor_{d-m-1} \end{bmatrix}, \tag{7.7}$$

and

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

where A, B are defined in (2.4). *Then we have*

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

REMARK. — *This is a finer version of Lemma 4.2. In particular, the strong component $L^s - \eta^{\mathrm{st}}\cdot v^{\mathrm{st}}$ is identical to the $L^s - \bar c\cdot v^{\mathrm{st}}$, 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 $\eta^{\mathrm{st}} = c^{\mathrm{st}} + A^{-1}Bc^{\mathrm{wk}}$, we compute

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

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

We now prove (7.8). We have

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

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

$$\frac{1}{2}\tilde z_i^{-1}(w_i^{\mathrm{wk}})^2 - \eta_i^{\mathrm{wk}}w_i^{\mathrm{wk}} = \frac{1}{2}\tilde z_i^{-1}(w_i^{\mathrm{wk}} - \tilde z_i\eta_i^{\mathrm{wk}})^2 - \frac{1}{2}\tilde z_i(\eta_i^{\mathrm{wk}})^2, \quad 1 \leqslant i \leqslant d-m,$$

and (7.8) follows. $\square$

We derive some useful estimates.

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

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

the following hold.

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

Proof. — For item 1, note that for each $j \geqslant i$, $|k_i^{\mathrm{wk}}| \leqslant \kappa|k_j^{\mathrm{wk}}|$, hence

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

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

For item 2, inverting (7.3) we get

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

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

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

Moreover, using the definition (see (1.3))

$$S = \partial_{II}^2 K = \left[k_1^{\mathrm{st}} \cdots k_m^{\mathrm{st}}\, k_1^{\mathrm{wk}} \cdots k_{d-m}^{\mathrm{wk}}\right]^T Q \left[k_1^{\mathrm{st}} \cdots k_m^{\mathrm{st}}\, k_1^{\mathrm{wk}} \cdots k_{d-m}^{\mathrm{wk}}\right],$$

we have

$$\begin{aligned} X_{i+1} &= \left[k_1^{\mathrm{st}} \cdots k_m^{\mathrm{st}}\, k_1^{\mathrm{wk}} \cdots k_i^{\mathrm{wk}}\right]^T Q \left[k_1^{\mathrm{st}} \cdots k_m^{\mathrm{st}}\, k_1^{\mathrm{wk}} \cdots k_i^{\mathrm{wk}}\right] \\ &= \left[\bar{k}_1^{\mathrm{st}} \cdots \bar{k}_m^{\mathrm{st}}\, \bar{k}_1^{\mathrm{wk}} \cdots \bar{k}_i^{\mathrm{wk}}\right]^T Q_0 \left[\bar{k}_1^{\mathrm{st}} \cdots \bar{k}_m^{\mathrm{st}}\, \bar{k}_1^{\mathrm{wk}} \cdots \bar{k}_i^{\mathrm{wk}}\right] =: \bar{P}^T Q_0 \bar{P}, \end{aligned}$$

where $\bar{k}$ is the first n components of k. We have assumed $Q_0 \geqslant D^{-1}\mathrm{Id}$ for $D > 1$. By Lemma 3.3, there exists a constant $c_n > 1$ depending only on n such that

$$\begin{aligned} \|X_{i+1}^{-1}\| &= (\min_{\|v\|=1} v^T X_{i+1} v)^{-1} = (\min_{\|v\|=1} v^T \bar{P} Q_0 \bar{P})^{-1} \leqslant D\|\bar{P}^{-1}\|^2 \\ &\leqslant D c_n |k_1^{\mathrm{st}}|^2 \cdots |k_m^{\mathrm{st}}|^2 |k_1^{\mathrm{wk}}|^2 \cdots |k_i^{\mathrm{wk}}|^2 \leqslant D c_n \bar{M}^m \kappa^{i-1} |k_i^{\mathrm{wk}}|^{2i}, \end{aligned}$$

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

7.3. Approximate Lipschitz property in the weak component. — In this section we prove Proposition 5.3. We fix $(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{st}}) \in \Omega_{\kappa,q}^{m,d} \cap \{\|U^{\mathrm{st}}\|_{C^2} \leqslant R\}$, and write $L^s = L_{\mathcal{H}^s}(\mathcal{B}^{\mathrm{wk}}, p, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{st}})$.

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

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

then

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

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

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

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

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

Proof. — First we have

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

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 $\gamma : (-\infty, 0] \to \mathbb{T}^d$ 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,

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

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

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

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

$$u(\varphi + he_i^{\mathrm{wk}}) - u(\varphi) \leqslant lh + \frac{1}{2}\delta_{d-m}h^2.$$

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

$$|u(\varphi + he_i^{\mathrm{wk}}) - u(\varphi)| \leqslant (\frac{1}{2}\delta_{d-m} + \frac{1}{2}\delta_{d-m}h)h \leqslant \delta_{d-m}h.$$

This is the Lipschitz estimate. □

We now state the inductive step.

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

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

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 $\gamma : (-\infty, 0] \to \mathbb{T}^d$ 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 $\xi : [-T, 0] \to \mathbb{T}^d$ satisfies

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

then

(7.11)

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

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

$$\zeta(-T) = \lfloor \gamma \rfloor_i(-T), \quad \zeta(0) = \lfloor \gamma \rfloor_i(0),$$

where $\lfloor \cdot \rfloor_i$ 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,

$$\lfloor \xi \rfloor_i(t) = \zeta(t) + \frac{th}{T} e_i^{\mathrm{wk}}. \tag{7.12}$$

2. We define the other components inductively. For $i < j \leqslant d-m$, suppose $\lfloor\xi\rfloor_{j-1}(t) = (\xi^{\mathrm{st}}, \xi_1^{\mathrm{wk}}, \dots, \xi_{j-1}^{\mathrm{wk}})(t)$ has been defined. We define
$$\xi_j^{\mathrm{wk}}(t) = \gamma_j^{\mathrm{wk}}(t) + y_j^T X_j^{-1}\lfloor\xi\rfloor_{j-1}(t) - y_j^T X_j^{-1}\lfloor\gamma\rfloor_{j-1}(t).$$

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

$$\begin{cases} \xi_j^{\mathrm{wk}}(-T) = \gamma_j^{\mathrm{wk}}(-T), \\ \dot\xi_j^{\mathrm{wk}} - y_j^T X_j^{-1}\lfloor\dot\xi\rfloor_{j-1} = \dot\gamma_j^{\mathrm{wk}} - y_j^T X_j^{-1}\lfloor\dot\gamma\rfloor_{j-1}. \end{cases} \tag{7.13}$$

We define $\chi = \xi(0) - \varphi - he_i^{\mathrm{wk}}$, and note that from (7.12),
$$\lfloor\chi\rfloor_i = \lfloor\xi\rfloor_i(0) - \lfloor\gamma\rfloor_i(0) - he_i^{\mathrm{wk}} = 0.$$

Action comparison. — We now compute

$$\begin{aligned}
&\int_{-T}^0 (L^s - c\cdot v)(\xi,\dot\xi)dt - \int_{-T}^0 (L^s - c\cdot v)(\gamma,\dot\gamma)dt \\
&\quad= \int_{-T}^0 L^s_{c,i}(\lfloor\xi\rfloor_i, \lfloor\dot\xi\rfloor_i)dt - \int_{-T}^0 L^s_{c,i}(\lfloor\gamma\rfloor_i, \lfloor\dot\gamma\rfloor_i)dt \\
&\qquad+ \sum_{j=i+1}^{d-m} \int_{-T}^0 \big(U_j^{\mathrm{wk}}(\xi(t)) - U_j^{\mathrm{wk}}(\gamma(t))\big)\, dt \\
&\qquad+ \frac12 \sum_{j=i+1}^{d-m} \tilde z_j^{-1} \int_{-T}^0 \big((\xi_j^{\mathrm{wk}} - y_j^T X_j^{-1}\lfloor\dot\xi\rfloor_{j-1} - \tilde z_j \eta_j^{\mathrm{wk}})^2 \\
&\qquad\quad - (\gamma_j^{\mathrm{wk}} - y_j^T X_j^{-1}\lfloor\dot\gamma\rfloor_{j-1} - \tilde z_j\eta_j^{\mathrm{wk}})^2\big) \\
&\quad\leqslant \int_{-T}^0 L^s_{c,i}(\lfloor\xi\rfloor_i, \lfloor\dot\xi\rfloor_i)dt - \int_{-T}^0 L^s_{c,i}(\lfloor\gamma\rfloor_i, \lfloor\dot\gamma\rfloor_i)dt + 2T\sum_{j=i+1}^{d-m}\|U_j^{\mathrm{wk}}\|_{C^0}.
\end{aligned} \tag{7.14}$$

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 $\|U_j^{\mathrm{wk}}\|_{C^0}$.

We now have

$$\begin{aligned}
&\int_{-T}^0 L^s_{c,i}(\lfloor\xi\rfloor_i, \lfloor\dot\xi\rfloor_i)dt - \int_{-T}^0 L^s_{c,i}(\lfloor\gamma\rfloor_i, \lfloor\dot\gamma\rfloor_i)dt \\
&\qquad= \int_{-T}^0 L^s_{c,i}(\lfloor\xi\rfloor_i, \lfloor\dot\xi\rfloor_i)dt - \int_{-T}^0 L^s_{c,i}(\zeta,\dot\zeta)dt \\
&\qquad\quad+ \int_{-T}^0 L^s_{c,i}(\zeta,\dot\zeta)dt - \int_{-T}^0 L^s_{c,i}(\lfloor\gamma\rfloor_i, \lfloor\dot\gamma\rfloor_i)dt \\
&\qquad\leqslant \int_{-T}^0 L^s_{c,i}(\lfloor\xi\rfloor_i, \lfloor\dot\xi\rfloor_i)dt - \int_{-T}^0 L^s_{c,i}(\zeta,\dot\zeta)dt,
\end{aligned}$$

noting that ζ is minimizing for $L^s_{c,i}$.

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

$$\int_{-T}^{0} L^s_{c,i}(\lfloor\xi\rceil_i, \lfloor\dot{\xi}\rceil_i)dt - \int_{-T}^{0} L^s_{c,i}(\zeta,\dot{\zeta})dt \leqslant l\cdot h + \left(\frac{1}{T}\tilde{z}_i^{-1} + T\|\sum_{j=i}^{d-m} U_j^{\mathrm{wk}}\|_{C^2}\right)h^2,$$

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

(7.15)

$$\int_{-T}^{0} L^s_{c,i}(\lfloor\xi\rceil_i, \lfloor\dot{\xi}\rceil_i)dt - \int_{-T}^{0} L^s_{c,i}(\zeta,\dot{\zeta})dt \leqslant l\cdot h + \sigma_i h^2, \sigma_i = \left(\tilde{z}_i^{-1}\sum_{j=i}^{d-m}\|U_j^{\mathrm{wk}}\|_{C^2}\right)^{\frac{1}{2}}.$$

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

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

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

$$u(\varphi + he_i^{\mathrm{wk}} + \chi) - u(\varphi) \leqslant l\cdot h + \sigma_i h^2 + \sigma_i.$$

Since $\lfloor\chi\rceil_i = 0$, using the inductive assumption,

$$|u(\varphi + he_i^{\mathrm{wk}} + \chi) - u(\varphi + he_i^{\mathrm{wk}})| \leqslant 2\sum_{j=i+1}^{d-m}\delta_j.$$

Therefore

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

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}. □

Proof of Proposition 5.3. — We have shown by induction that for all $1 \leqslant i \leqslant d-m$, 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 $1 \leqslant i \leqslant d-m$

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

Then $\delta_{d-m} = 2\sigma_{d-m} \leqslant M^*|k^{\mathrm{wk}}_{d-m}|^{-\frac{q}{2}+d-m}$. For each $1 \leqslant i \leqslant d-m$, we have
(7.16)
$$\delta_i = \sqrt{d}(6\sigma_i + 4\sum_{j=i+1}^{d-m}\delta_i) \leqslant (6\sqrt{d})^{i-m}\sum_{j=i}^{d-m}\sigma_i \leqslant M^*(6\sqrt{d})^{i-m}|k^{\mathrm{wk}}_i|^{-\frac{q}{2}+d-m}.$$

For any $\varphi^{\mathrm{wk}}, \psi^{\mathrm{wk}} \in \mathbb{T}^{d-m}$ and $\varphi^{\mathrm{st}} \in \mathbb{T}^m$,
$$|u(\varphi^{\mathrm{st}}, \varphi^{\mathrm{wk}}) - u(\varphi^{\mathrm{st}}, \psi^{\mathrm{wk}})| \leqslant \sum_{i=1}^{d-m}\delta_i|\varphi^{\mathrm{wk}}_i - \psi^{\mathrm{wk}}_i| + \sum_{i=1}^{d-m}\delta_i$$

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

Appendix A. Normally hyperbolic invariant manifolds

In this section we state a version of the center manifold theorem and prove Theorem 2.4. While the central manifold theorem is classical, we need an version whose center direction is a non-compact set equipped with a Riemannian metric. This is done in the first two subsections. In the last subsection, we perform a reduction on our system H^s from (2.3) to apply the central manifold theorem.

A.1. Normally hyperbolic invariant manifolds via isolation block. — We state an abstract theorem on existence of normally hyperbolic invariant manifolds for a smooth map F, based on construction of Conley's isolating block (see McGehee, [26]).

We introduce a set of notations. We have three components $x \in \mathbb{R}^s, y \in \mathbb{R}^u, z \in \Omega^c \subset \mathbb{R}^c$, where Ω^c is a (possibly unbounded) convex set. We assume that Ω^c admits a C^1 complete Riemannian metric g. We also consider a Riemannian metric on the product space $W = \mathbb{R}^s \times \mathbb{R}^u \times \Omega^c$ by taking the tensor product of g and Ω^c, and the standard Euclidean metric on $\mathbb{R}^s, \mathbb{R}^u$.

Fix some $r > 0$ and let $D^s \subset \mathbb{R}^s$ and $D^u \subset \mathbb{R}^u$ be *closed* balls of radius r at the origin in $\mathbb{R}^s$ and $\mathbb{R}^u$ (r is considered fixed and we omit the dependence). Denote $D^{\mathrm{sc}} = D^s \times \Omega^c$, $D^{uc} = D^u \times \Omega^c$, and $D = D^{\mathrm{sc}} \times D^u$.

Consider a C^1 smooth map
$$F : D = D^s \times D^u \times \Omega^c \to \mathbb{R}^s \times \mathbb{R}^u \times \Omega^c,$$
we state a set of conditions guaranteeing the set
$$W^{\mathrm{sc}}(F) = \{Z \in D : F^k(Z) \in D \text{ for all } k > 0\},$$
called the center-stable manifold, is a graph i.e.,
$$W^{\mathrm{sc}}(F) = \{(X, Y) \in D^{\mathrm{sc}} \times D^u : w^{\mathrm{sc}}(X) = Y\}$$
for a C^1 function w^{sc}.

[C1] $\pi_{sc}F(D^{\mathrm{sc}} \times D^u) \subset D^{\mathrm{sc}}$.

[C2] F maps $D^{\mathrm{sc}} \times \partial D^u$ into $D^{\mathrm{sc}} \times \mathbb{R}^u \backslash D^u$ and is a homotopy equivalence.

The first two conditions guarantee a *topological isolating block*: F stretches $D^{\mathrm{sc}} \times B^u$ along the unstable component D^u and is a weak contraction along the center-stable component D^{sc}.

Now we state the *cone conditions*. For some $\mu > 0$

$$C^u_\mu(Z) = \{v = (v^c, v^s, v^u) \in T_Z D : \mu\|v^u\|^2 \geqslant \|v^c\|^2 + \|v^s\|^2\}.$$

Note that

$$(C^u_\mu(Z))^c = \{v = (v^c, v^s, v^u) \in T_Z D : \mu^{-1}(\|v^c\|^2 + \|v^s\|^2) \geqslant \|v^u\|^2\} =: C^{\mathrm{sc}}_{\mu^{-1}}(Z).$$

Let us also define

$$K^\mu_u(x_1, y_1, z_1) = \{(x_2, y_2, z_2) : \mu\|y_2 - y_1\|^2 \geqslant \|x_2 - x_1\|^2 + \mathrm{dist}(z_1, z_2)^2\},$$

where the distance is induced by the Riemannian metric g.

We assume there exist $\mu > 1$ and $\chi > 1$ with the property that for any $Z_1, Z_2 \in D$ such that $Z_2 \in K^\mu_u(Z_1)$ we have

[C3] $F(Z_2) \in K^u_\mu(F(Z_1))$.

[C4] $\|\pi_u(F(Z_2) - F(Z_1))\| \geqslant \chi\|\pi_u(Z_2 - Z_1)\|$.

Proposition A.1. — *(Lipschitz center-stable manifold theorem) Suppose F satisfies conditions [C1-C4], then $W^{\mathrm{sc}}(F)$ is given by the graph of a Lipschitz function*

$$W^{\mathrm{sc}}(F) = \{(x, y, z) \in D : w^{\mathrm{sc}}(x, z) = y\}.$$

Moreover, for if $W^{\mathrm{sc}}(F)$ is C^1, we have

$$T_Z W^{\mathrm{sc}}(F) \in C^{\mathrm{sc}}_{\mu^{-1}}(Z).$$

In order to obtain the center-unstable manifold, consider the involution $I : (x, y, z) \mapsto (y, x, z)$ and assume $\mathrm{inv}(F) = I \circ F^{-1} \circ I^{-1}$ satisfies the same conditions.

Theorem A.2. — *Assume that F and* $\mathrm{inv}(F)$ *satisfy the conditions [C1-C4], there exists a C^1 function $w^c : M \to D$ such that*

$$W^c(F) := W^{\mathrm{sc}}(F) \cap W^{uc}(F) = \{(x, y, z) \in D : w^c(z) = (x, y)\}.$$

Proof. — Proposition A.1 implies the existence of Lipschitz functions $w^{uc} : D^{uc} \to D$ and $w^{\mathrm{sc}} : D^{\mathrm{sc}} \to D$, with

$$W^{\mathrm{sc}}(F) = \{x = w^{\mathrm{sc}}(y, z)\}, \quad W^{uc}(F) = W^{\mathrm{sc}}(\mathrm{inv}(F)) = \{y = w^{uc}(x, z)\}.$$

Then standard arguments (see Theorem 5.18 in [27]) implies these functions are C^1. The fact that $\mu > 1$ and

$$T_Z W^{\mathrm{sc}}(F) \in C^{\mathrm{sc}}_{\mu^{-1}}(Z), \quad T_Z W^u(F) \in C^{\mathrm{sc}}_{\mu^{-1}}(Z)$$

implies $W^{\mathrm{sc}}(F)$ and $W^{uc}(F)$ intersect transversally, and $W^{\mathrm{sc}}(F) \cap W^{uc}(F)$ is a graph over the center component M. □

A.2. Existence of Lipschitz invariant manifolds. — We prove Proposition A.1. Let $\mathcal{V}$ be a collection of sets $\Gamma \subset D$ satisfying the following conditions:
(a) $\pi_u \Gamma = D^u$,
(b) $Z_2 \in K^u_\mu(Z_1)$ for all $Z_1, Z_2 \in \Gamma$, where π_u is the projection to the unstable component.

These conditions ensures $\pi_u : \Gamma \to D^u$ is one-to-one and onto, therefore, Γ is a graph over D^u. Moreover, condition (b) further implies that the graph is Lipschitz. In particular, each $\Gamma \in \mathcal{V}$ is a topological disk.

Lemma A.3. — *Let $\Gamma \in \mathcal{V}$, then $F(\Gamma) \cap D \in \mathcal{V}$.*

Proof. — By [C4] for any Z_1 and Z_2 we have that $F(Z_2)$ belongs to the cone $K^u_{F(Z_1)}$ of $F(Z_1)$. Thus, it suffices to show that $D^u \subset \pi_u(F(\Gamma) \cap D)$. The proof is by contradiction. Suppose there is $Z_* \in B^u$ such that $Z_* \notin \pi_u(F(\Gamma))$.

We have the following commutative diagram

$$\begin{array}{ccc} \partial\Gamma & \overset{i_1}{\hookrightarrow} & \Gamma \\ \downarrow \pi_u \circ F & & \downarrow \pi_u \circ F. \\ \mathbb{R}^u \backslash D^u & \overset{i_2}{\hookrightarrow} & \mathbb{R}^u \backslash \{Z_*\} \end{array} \tag{A.1}$$

From [C2] and using the fact that B^s and Ω^c are contractible, $\pi_u \circ F|\Gamma$ is a homotopy equivalence. Note that i_2 is a homotopy equivalences, and $\pi_u \circ F|\Gamma$ is a homeomorphism onto its image. Let h and g be the homotopy inverses of $\pi_u \circ F|\partial\Gamma$ and i_2, then $h \circ g \circ (\pi_u \circ F)$ defines a homotopy inverse of i_1. As a result Γ is homotopic to $\partial\Gamma$, this is a contradiction. □

Proposition A.1 follows from the following statement.

Proposition A.4. — *The mapping $\pi_{sc} : W^{\mathrm{sc}}(F) \to D^{\mathrm{sc}}$ is one-to-one and onto, therefore, it is the graph of a function w^{sc}. Moreover, w^{sc} is Lipschitz and*

$$T_Z W^{\mathrm{sc}}(F) \in (C^u_\mu(Z))^c = C^{\mathrm{sc}}_{\mu^{-1}}(Z), \quad Z \in W^{\mathrm{sc}}(F).$$

Proof. — For each $X \in D^{\mathrm{sc}}$, we define $\Gamma_X = (\pi_{sc})^{-1}X$, clearly $\Gamma_X \in \mathcal{V}$. We first show $\Gamma_X \cap W^{\mathrm{sc}}(F)$ is nonempty and consists of a single point. Assume first that $\Gamma_X \cap W^{\mathrm{sc}}(F)$ is empty. Then by definition of $W^{\mathrm{sc}}(F)$, there is $n \in \mathbb{N}$ such that $F^n(\Gamma_X) \cap D = \varnothing$. However, by Lemma A.3, $\bigcap_{i=1}^n F^i(\Gamma_X) \cap D \in \mathcal{V}$ is always nonempty, a contradiction. We now consider two points $Z_1, Z_2 \in W^{\mathrm{sc}}(F)$ with $\pi_u Z_1 = \pi_u Z_2$. Note that $F^k(Z_1), F^k(Z_2) \in D$ for all $k \geqslant 0$, and $Z_2 \in K^u_\mu(Z_1)$, by [C4] we have

$$2 \geqslant \|\pi_u(F^k(Z_1) - F^k(Z_2))\| \geqslant \chi^k \|\pi_u(Z_1 - Z_2)\|$$

for all k, which implies $Z_1 = Z_2$.

The last argument actually shows $Z_2 \notin K^u_\mu(Z_1)$ for all $Z_1, Z_2 \in W^{\mathrm{sc}}(F)$. For any $\epsilon > 0$, for $Z_1 = (X_1, Y_1), Z_2 = (X_2, Y_2) \in W^{\mathrm{sc}}(F)$ with $\mathrm{dist}(X_1, X_2)$ small, we have $\|Y_1 - Y_2\| \leqslant \mu^{-\frac{1}{2}} \mathrm{dist}(X_1, X_2)$. This implies both the Lipschitz and the cone properties in our proposition. □

A.3. NHIC for the dominant system. — We prove Theorem 2.4 in this section. First, an overview of notations.

1. The strong Hamiltonian is $H^{\mathrm{st}} = \mathcal{H}^{\mathrm{st}}(p_0, \mathcal{B}^{\mathrm{st}}, U^{\mathrm{st}})$ defined on $\mathbb{T}^m \times \mathbb{R}^m$, and its associated Lagrangian vector field is X^{st} (see (2.7)). We denote the time-1-map of X^{st} by G_0^{st} and lift it to the universal cover $\mathbb{R}^m \times \mathbb{R}^m$ without changing its name.
2. The vector field X^{st} is extended trivially to $(\mathbb{T}^m \times \mathbb{R}^m) \times (\mathbb{T}^{d-m} \times \mathbb{R}^{d-m})$ (see (2.10)). The time-1-map is denoted G_0, and we have $G_0 = G_0^{\mathrm{st}} \times \mathrm{Id}$. We will also lift it to the universal cover with the same name.
3. The slow Hamiltonian is $H^s = \mathcal{H}^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$, and consider its Lagrangian vector field X^s_{Lag}.

   We apply a coordinate change $(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, v^{\mathrm{wk}}) = \Phi(\varphi^{\mathrm{st}}, v^{\mathrm{st}}, \varphi^{\mathrm{wk}}, I^{\mathrm{wk}})$ as in (2.8), and a rescaling Φ_Σ as defined in (2.12). The new vector field is denoted $\tilde{X}^s = (\Phi_\Sigma^{-1})_*(\Phi^{-1})_* X^s_{\mathrm{Lag}}$ (see (2.9), (2.12)). We denote its time-1-map G, which is considered a map on the Euclidean space $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^{d-m}$.
4. Below we also use notations from Section 2.4.

By Theorem 2.3, we have:

Corollary A.5. — *Assume that $(\mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}}) \in \Omega^{m,d}_{\kappa,q}(\mathcal{B}^{\mathrm{st}})$, then for any $\delta_1 > 0$, there exists $M > 0$ such that for all $(\mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$ with $\mu(\mathcal{B}^{\mathrm{wk}}) > M$, uniformly on $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^{d-m}$, we have*

$$\|\Pi_{(\varphi^{\mathrm{st}}, v^{\mathrm{st}})}(G - G_0)\| < \delta_1, \quad \|DG - DG_0\| < \delta_1.$$

By assumption, the Hamiltonian flow H^{st}, (and hence G_0^{st}) admits an NHIC $\chi^{\mathrm{st}}(\mathbb{T}^l \times B^l_{1+a})$ with exponent α, β, where χ^{st} is an embedding. We use local coordinates in a tubular neighborhood to simplify the setting. (See also the left block of (A.3) below)

Lemma A.6. — *There exist a tubular neighborhood $N(\Lambda^{\mathrm{st}}_a) \subset \mathbb{T}^m \times \mathbb{R}^m$ of Λ^{st}_a and a C^1 diffeomorphism*

$$h^{\mathrm{st}} : B^{m-l}_1 \times B^l_1 \times (\mathbb{T}^l \times B^l_{1+a}) \to N(\Lambda^{\mathrm{st}}_a)$$

such that:

1. *$h^{\mathrm{st}}(0, 0, z) = \chi^{\mathrm{st}}(z)$, in particular, $h^{\mathrm{st}}(\mathcal{C}^{\mathrm{st}}_a) := h^{\mathrm{st}}(\{0\} \times \{0\} \times (\mathbb{T}^l \times B^l_{1+a})) = \Lambda^{\mathrm{st}}_a$.*
2. *For the map $F_0^{\mathrm{st}} := (h^{\mathrm{st}}) \circ G_0^{\mathrm{st}} \circ (h^{\mathrm{st}})^{-1}$:*
   (a) *$\mathcal{C}^{\mathrm{st}}_a$ is an NHIC for F^{st}_0 with the same exponents α, β.*

(b) *The associated stable/unstable bundles take the form*

$$E^s = \mathbb{R}^l \times \{0\} \times \{0\}, \quad E^u = \{0\} \times \mathbb{R}^l \times \{0\}.$$

In particular, DF_0^{st} is a block diagonal matrix in the blocks corresponding to the three components.

(c) *Let g_0 denote the Euclidean metric. Then there exists a Riemannian metric g on $\mathbb{T}^l \times B_{1+a}^l$ such that the tensor metric $g_0 \otimes g_0 \otimes g$ on $B_1^l \times B_1^l \times (\mathbb{T}^l \times B_{1+a}^l)$ is an adapted metric for the NHIC $\mathcal{C}_a^{\mathrm{st}}$.*

Proof. — We use the bundles E^u, E^s, and the parametrization χ^{st} of Λ_a^{st} to build a coordinate system for the normal bundle to Λ_a^{st}, which is diffeomorphic to the tubular neighborhood. We then pull back the adapted metric of Λ_a^{st} using this map to $\mathcal{C}_a^{\mathrm{st}}$. □

Denote $\Omega^{\mathrm{wk}} = \mathbb{R}^{d-m} \times \mathbb{R}^{d-m}$ and consider the trivial extension

$$h : B_1^l \times B_1^l \times ((\mathbb{T}^l \times B_{1+a}^l) \times \Omega^{\mathrm{wk}}) \to N(\Lambda_a^{\mathrm{st}}) \times \Omega^{\mathrm{wk}}$$

by $h(x, y, (z^{\mathrm{st}}, z^{\mathrm{wk}})) = (h^{\mathrm{st}}(x, y, z^{\mathrm{st}}), z^{\mathrm{wk}})$. Define the following maps

$$F_0 = h^{-1} \circ G_0 \circ h = (F_0^{\mathrm{st}}, \mathrm{Id}), \quad F = h^{-1} \circ G \circ h. \tag{A.2}$$

See diagram below, where "i" denote standard embeddings, $\Omega_a^{\mathrm{st}} = \mathbb{T}^l \times B_{1+a}^l$ and $\circlearrowleft \cdot$ denotes the (unperturbed) map of the given space.

(A.3)

$$\begin{tikzcd}
\Lambda_a^{\mathrm{st}} \arrow[r, "i"] & N(\Lambda_a^{\mathrm{st}}) \subset (\mathbb{T}^m \times \mathbb{R}^m) \circlearrowleft G_0^{\mathrm{st}} \arrow[r, "i"] & (\mathbb{T}^m \times \mathbb{R}^m) \times \Omega^{\mathrm{wk}} \circlearrowleft G_0 \\
\Omega_a^{\mathrm{st}} \arrow[u, "\chi^{\mathrm{st}}"] \arrow[r, "i"] & B_1^{m-l} \times B_1^{m-l} \times \Omega_a^{\mathrm{st}} \circlearrowleft F_0^{\mathrm{st}} \arrow[u, "h^{\mathrm{st}}"] \arrow[r, "i"] & B_1^{m-l} \times B_1^{m-l} \times (\Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}}) \circlearrowleft F_0. \arrow[u, "h"]
\end{tikzcd}$$

Finally, to apply Theorem A.2, we lift Ω_a^{st} to the universal cover $\mathbb{R}^l \times B_{1+a}^l$, and the maps F_0, F to the covering space without changing their names, namely

$$F, F_0 : B_1^l \times B_1^l \times (\Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}}) \circlearrowleft .$$

$\Omega_{1+a}^{\mathrm{st}} \times \Omega^{\mathrm{wk}}$ is our center component and is denoted by Ω. While the maps are defined on unbounded regions, we keep in mind that $F_0 = (F_0^{\mathrm{st}}, \mathrm{Id})$ where F_0^{st} is defined on a compact set $B_1^l \times B_1^l \times (\mathbb{T}^l \times B_{1+a}^l)$.

We still need one reduction to apply Theorem A.2. Recall that $\Omega_a^{\mathrm{st}} = \mathbb{R}^l \times B_{1+a}^l$. Write $F_0 = (F_0^x, F_0^y, F_0^z)$, define

$$L(x, y, z) = (D_x F_0^x(0, 0, z) \cdot x, D_y F_0^y(0, 0, y) \cdot y, F_0^z(0, 0, z)) \tag{A.4}$$

this is the linearized map at $(0, 0, z)$ (we used $F_0(0, 0, z) = (0, 0, F_0^z)$, and DF_0 is block diagonal from Lemma A.6). since $F_0 = (F_0^{\mathrm{st}}, \mathrm{Id})$ and F_0^{st} is defined over a compact set, we obtain as $r \to 0$,

$$\|L - F_0\| = o(r), \quad \|DL - DF_0\| = o(1) \text{ on } B_r^l \times B_r^l \times (\Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}}). \tag{A.5}$$

Moreover, since F_0 preserves $\{0\} \times \{0\} \times \partial\Omega$, we get

$$L(B_1^l \times B_1^l \times \partial\Omega) \subset \mathbb{R}^l \times \mathbb{R}^l \times \partial\Omega. \tag{A.6}$$

Namely, the linearized map L preserves the boundary of the center component. Finally, we modify the map F so that it also fixes the center boundary. Let ρ be a standard mollifier satisfying

$$\begin{cases} \rho(x, y, (z^{\mathrm{st}}, z^{\mathrm{wk}})) = \rho(z^{\mathrm{st}}) = 0 & z^{\mathrm{st}} \in \Omega_0^{\mathrm{st}} \\ \rho(x, y, (z^{\mathrm{st}}, z^{\mathrm{wk}})) = \rho(z^{\mathrm{st}}) = 1 & z^{\mathrm{st}} \in \Omega_a^{\mathrm{st}} \backslash \Omega_{a/2}^{\mathrm{st}}. \end{cases}$$

Let

$$\tilde{F} = F(1-\rho) + L\rho, \tag{A.7}$$

we have:

Lemma A.7. — *For any $\mu > 1$, $\epsilon > 0$, and $r_0 > 0$, there exist $\delta_1 > 0$ and $0 < r < r_0$ such that if G and G_0 satisfy*

$$\|\Pi_{(\varphi^{\mathrm{st}}, v^{\mathrm{st}})}(G - G_0)\| < \delta_1, \quad \|DG - DG_0\| < \delta_1,$$

then the map $\tilde{F}$, defined by (A.2), (A.4), *and* (A.7), *satisfies conditions* [C1]–[C4] *with the parameters μ and $\chi = \alpha^{-1} - \epsilon$ on $B_r^l \times B_r^l \times \Omega$. The same hold for the map* $\mathrm{inv}(\tilde{F})$.

Proof. — First of all, from Lemma A.6, $DF_0(0,0,z) = \mathrm{diag}\{D_x F_0^x, D_y F_0^y, D_z F_0^z\}$ with

$$\|D_x F_0^x\|, \|(D_y F_0^y)^{-1}\|^{-1} \leqslant \alpha, \quad \|D_z F_0^z\|, \|(D_z F_0^z)^{-1}\| \leqslant \beta^{-1}. \tag{A.8}$$

Recall that $F_0 = (F_0^{\mathrm{st}}, \mathrm{Id})$ where F_0^{st} is defined over a compact set. Therefore, for sufficiently small $r > 0$, we have

$$\|\Pi_x F_0(x,y,z)\| \leqslant (\alpha + \epsilon)\|x\|, \quad \|\Pi_y DF_0(x,y,z)\| \geqslant (\alpha+\epsilon)^{-1}\|y\|,$$

hence

$$\Pi_x F_0(B_r^l \times B_r^l \times \Omega) \subset B_{(\alpha+\epsilon)r}^l, \quad \|\Pi_y F_0(B_r^l \times \partial B_r^l \times \Omega)\| \geqslant (\alpha+\epsilon)^{-1} r.$$

Since $\|\tilde{F} - F_0\| \leqslant \|(1-\rho)(F - F_0)\| + \|\rho(L - F_0)\|$,

$$\|\Pi_{(x,y)}(F - F_0)\| \leqslant C\|\Pi_{(\varphi^{\mathrm{st}}, v^{\mathrm{st}})}(G - G_0)\| \leqslant C\delta_1,$$

and $\|L - F_0\| = o(r)$, by choosing δ_1, r small enough, we get

$$\Pi_x \tilde{F}(B_r^l \times B_r^l \times \Omega) \subset B_r^l, \quad \|\Pi_y \tilde{F}(B_r^l \times \partial B_r^l \times \Omega)\| > r.$$

The first half of the above formula, combined with (A.6), gives [C1], and the second half gives [C2].

We now prove the cone conditions [C3] and [C4]. We first show the map $\tilde{F}$ is well approximated by the linearized map $DF_0(0,0,z)$. Given any $\epsilon > 0$, we use Corollary A.5 to choose δ_1 so small such that

$$\|D(F-F_0)\| + \|\Pi_{z^{\mathrm{st}}}(F-F_0)\|\|d\rho\| \leqslant C\|D(G-G_0)\| + \|\Pi_{\varphi^{\mathrm{st}}, v^{\mathrm{st}}}(G-G_0)\|\|d\rho\| < \epsilon/2$$

By (A.5), we can choose r_1 such that for $0 < r < r_1$,

$$\|D(F_0 - L)\| + \|F_0 - L\|\|d\rho\| < \epsilon/2$$

on $B_r^l \times B_r^l \times (\Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}})$. Then from $\tilde{F} = F + (L - F)\rho$, and the fact that ρ depends only on z^{st} gives

$$\begin{aligned}&\|D\tilde{F}(x,y,z) - DL(x,y,z)\| \leqslant \|DF - DL\| + \|\Pi_{z^{\mathrm{st}}}(F - L)\|\|d\rho\| \\ &\leqslant \|D(F - F_0)\| + \|D(F_0 - L)\| + (\|\Pi_{z^{\mathrm{st}}}(F - F_0)\| + \|\Pi_{z^{\mathrm{st}}}(F_0 - L)\|)\|d\rho\| < \epsilon.\end{aligned}$$

Consider $(x_1, y_1, z_1), (x_2, y_2, z_2) \in B_r^l \times B_r^l \times \Omega$, denote $(\Delta x, \Delta y, \Delta z) = (x_2, y_2, z_2) - (x_1, y_1, z_1)$ and $d = \|\Delta x\| + \|\Delta y\| + \mathrm{dist}(z_1, z_2)$. For d small enough

$$\begin{aligned}&\|\tilde{F}(x_2, y_2, z_2) - \tilde{F}(x_1, y_1, z_1) - DF_0(0,0,z)(\Delta x, \Delta y, \Delta z)\| \\ &\quad = \|L(x_2, y_2, z_2) - L(x_1, y_1, z_1) + (\tilde{F} - L)(x_2, y_2, z_2) - (\tilde{F} - L)(x_1, y_1, z_1) \\ &\qquad - DF_0(0,0,z_1)(\Delta x, \Delta y, \Delta z)\| \\ &\quad \leqslant \|D(\tilde{F} - L)(x_1, y_1, z_1)\| d \leqslant \epsilon d.\end{aligned}$$

To prove [C3], we first show the linear map preserves the unstable cone. For any $\mu > 1$ and $(v_x, v_y, v_z) \in T_{(x_1,y_1,z_1)} B_r^l \times B_r^l \times \Omega$ with $\mu\|v^y\|^2 \geqslant \|v^z\|^2 + \|v^x\|^2$, let $(v_x', v_y', v_z') = DF_0(0,0,z)(v_x, v_y, v_z)$, we have

$$\begin{aligned}\mu\|v_y'\|^2 &\geqslant \mu\alpha^{-1}\|v_y\|^2 \geqslant \alpha^{-1}(\|v_x\|^2 + \|v_z\|^2) \\ &\geqslant \alpha^{-1}(\|v_x'\|^2 + \beta\|v_z'\|^2) \geqslant \frac{\beta}{\alpha}(\|v_x'\|^2 + \|v_z'\|^2).\end{aligned}$$

In other words, for any $\mu > 1$, we have $DF_0(0,0,z)C_\mu^u \subset C_{\alpha\mu/\beta}^u$.

Coming to the non-linear map $\tilde{F}$, for $(x_1, y_1, z_1), (x_2, y_2, z_2) \in B_r^l \times B_r^l \times \Omega$, let $(x_i', y_i', z_i') = \tilde{F}(x_i, y_i, z_i)$, and $(\Delta x, \Delta y, \Delta z)$, $(\Delta x', \Delta y', \Delta z')$ be the corresponding difference. If $(x_2, y_2, z_2) \in K_\mu^u(x_1, y_1, z_1)$, then $\mu\|\Delta y\|^2 \geqslant \|\Delta x\|^2 + \mathrm{dist}(z_1, z_2)^2$. In particular, $\mathrm{dist}(z_1, z_2)^2 \leqslant \mu\|\Delta y\|^2 \leqslant \mu r^2$. When r is small enough

$$\|(\Delta x', \Delta y', \Delta z') - DF_0(0,0,z_1)(\Delta x, \Delta y, \Delta z)\| \leqslant \epsilon d.$$

Furthermore, assume r is so small that $(1-\epsilon)\mathrm{dist}(z, z') \leqslant \|\Delta z\|_{(0,0,z_1)} \leqslant (1+\epsilon)\mathrm{dist}(z, z')$, where $\|\Delta z\|_{(0,0,z_1)}$ is measured using the local Riemannian metric. We drop the subscript from now on. Using the linear calculation, there exists a uniform constant $C > 1$ such that

$$\begin{aligned}\mu\|\Delta y'\|^2 &\geqslant \frac{\beta}{\alpha}(\|\Delta x'\|^2 + \|\Delta z'\|^2) - C\epsilon^2 d^2 \\ &\geqslant \frac{\beta}{\alpha}(\|\Delta x'\|^2 + \|\Delta z'\|^2) - C(1+\mu)\epsilon^2\|\Delta y\|^2 \\ &\geqslant (1-\epsilon)\frac{\beta}{\alpha}(\|\Delta x'\|^2 + \mathrm{dist}(z_1', z_2')^2) - C^2(1+\mu)\epsilon^2\|\Delta y'\|^2,\end{aligned}$$

noting that $\|D\tilde{F}^{-1}\|, \|D\tilde{F}\|$ are uniformly bounded. When ϵ is small enough, we get $\mu\|\Delta y\|^2 \geqslant \|\Delta x\|^2 + \mathrm{dist}(z_1', z_2')^2$. Thus, [C3] is proven.

[C4] follows directly from $\|(\Delta x', \Delta y', \Delta z') - DF_0(0,0,z_1)(\Delta x, \Delta y, \Delta z)\| \leqslant \epsilon d$ and (A.8). The proof for $\mathrm{inv}(\tilde{F})$ is identical and is omitted. □

Proof of Theorem 2.4. — For any $\delta > 0$, we choose $0 < r < \delta/C$ and $\mu > 1$, where C is a constant specified later. Apply Lemma A.7, there exists $M > 0$ such that whenever $\mu(\mathcal{B}^{\mathrm{wk}}) > M$, the map $\tilde{F}$ associated to $H^s(\mathcal{B}^{\mathrm{st}}, \mathcal{B}^{\mathrm{wk}}, p_0, U^{\mathrm{st}}, \mathcal{U}^{\mathrm{wk}})$ satisfies [C1]–[C4] on $B_r^l \times B_r^l \times (\Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}})$. As a result, we obtain a function $w^c : \Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}} \to B_r^{m-l} \times B_r^{m-l}$ such that

$$W^c = \mathrm{Graph}(w^c) = \{(x, y, (z^{\mathrm{st}}, z^{\mathrm{wk}})) : (x,y) = w^c(z^{\mathrm{st}}, z^{\mathrm{wk}})\}$$

is invariant under $\tilde{F}$, and is the maximally invariant set on $B_r^{m-l} \times B_r^{m-l} \times \Omega_a^{\mathrm{st}} \times \Omega^{\mathrm{wk}}$. Since $\tilde{F} = F$ on whenever $z^{\mathrm{st}} \in \Omega_{a/2}^{\mathrm{st}}$, any F invariant set in $U_1 := B_r^{m-l} \times B_r^{m-1} \times \Omega_{a/2}^{\mathrm{st}} \times \Omega^{\mathrm{wk}}$ is also $\tilde{F}$ invariant and hence is contained in W^c.

We now consider the map

$$\zeta : \Omega_0^{\mathrm{st}} \times \Omega^{\mathrm{wk}} \to (\mathbb{R}^m \times \mathbb{R}^m) \times (\mathbb{R}^{d-m} \times \mathbb{R}^{d-m}),$$
$$(z^{\mathrm{st}}, z^{\mathrm{wk}}) \mapsto h(w^c(z^{\mathrm{st}}, z^{\mathrm{wk}}), z^{\mathrm{st}}, z^{\mathrm{wk}}),$$

then $\zeta(\Omega_0^{\mathrm{st}} \times \Omega^{\mathrm{wk}})$ is weakly invariant for the vector field $\tilde{X}^s$, and maximally invariant on the set $U_2 := h(B_r^{m-l} \times B_r^{m-1} \times \Omega_0^{\mathrm{st}} \times \Omega^{\mathrm{wk}})$.

Finally, inverting the coordinate changes, we obtain for

$$\eta^s = \Phi \circ \Phi_\Sigma \circ \zeta, \quad \eta^s : \Omega_0^{\mathrm{st}} \times \Omega^{\mathrm{wk}} \to \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{d-m} \times \mathbb{R}^{d-m},$$

$\eta^s(\Omega_0^{\mathrm{st}} \times \Omega^{\mathrm{wk}})$ is weakly invariant for $X^s_{\mathrm{Lag}} = (\Phi)_*(\Phi_\Sigma)_*\tilde{X}^s$, and maximally invariant on $U_3 = \Phi \circ \Phi_\Sigma(U_2)$.

Since h is identity in the weak component

$$U_2 = h^{\mathrm{st}}(B_r^{m-l} \times B_r^{m-1} \times \Omega_0^{\mathrm{st}}) \times \Omega^{\mathrm{wk}}.$$

The coordinate changes Φ and Φ_Σ does not change $(\varphi^{\mathrm{st}}, v^{\mathrm{st}})$, therefore,

$$U_3 = \Phi \circ \Phi_\Sigma(V_2) = h^{\mathrm{st}}(B_r^{m-l} \times B_r^{m-1} \times \Omega_0^{\mathrm{st}}) \times \Omega^{\mathrm{wk}} =: V \times \Omega^{\mathrm{wk}}.$$

Moreover, using the fact that $h^{\mathrm{st}}|_{\{0\}\times\{0\}\times\Omega_0^{\mathrm{st}}} = \chi^{\mathrm{st}}|_{\Omega_0^{\mathrm{st}}}$, and $\|w^c\|_{C^0} < r$ for small enough r we have for some $C > 0$, uniformly over all z^{wk}

$$\|h^{\mathrm{st}}(w^c(z^{\mathrm{st}}, z^{\mathrm{wk}}), z^{\mathrm{st}}) - \chi^{\mathrm{st}}(z^{\mathrm{st}})\| \leqslant Cr.$$

Since $h(x, y, z^{\mathrm{st}}, z^{\mathrm{wk}}) = (h^{\mathrm{st}}(x, y, z^{\mathrm{st}}), z^{\mathrm{wk}})$, we get $\|\Pi_{\varphi^{\mathrm{st}}, v^{\mathrm{st}}}\zeta - \chi^{\mathrm{st}}\|_{C^0} < Cr$, where we abuse notation by writing $\chi^{\mathrm{st}}(z^{\mathrm{st}}, z^{\mathrm{wk}}) = z^{\mathrm{st}}(z^{\mathrm{st}})$. Finally, since Φ, Φ_Σ are identity in $\varphi^{\mathrm{st}}, v^{\mathrm{st}}$, we have

$$\|\Pi_{\varphi^{\mathrm{st}}, v^{\mathrm{st}}}\Phi \circ \Phi_\Sigma \circ \zeta - \chi^{\mathrm{st}}\|_{C^0} = \|\Pi_{\varphi^{\mathrm{st}}, v^{\mathrm{st}}}\zeta - \chi^{\mathrm{st}}\|_{C^0} < Cr.$$

We obtain $\|\Pi_{\varphi^{\mathrm{st}}, v^{\mathrm{st}}}\eta^s - \chi^{\mathrm{st}}\| < \delta$. □

Acknowledgments. — The first author acknowledges NSF for partial support grant DMS-5237860. The second author is partially supported by the NSERC DISCOVERY grant, reference number 436169-2013. The authors would like to thank John Mather, Marcel Guardia, and Abed Bounemoura for useful conversations. We also thank a referee for suggesting several improvements of the paper.

BIBLIOGRAPHY

[1] V. I. Arnold – "Small denominators and problems of stability of motion in classical and celestial mechanics", *Uspehi Mat. Nauk* **18** (1963), p. 91–192.

[2] ———, "Instabilities in dynamical systems with several degrees of freedom", *Sov Math Dokl* **5** (1964), p. 581–585.

[3] ———, "Instability of dynamical systems with many degrees of freedom", *Dokl. Akad. Nauk SSSR* **156** (1964), p. 9–12.

[4] ———, "Mathematical problems in classical physics", in *Trends and perspectives in applied mathematics*, Appl. Math. Sci., vol. 100, Springer, 1994, p. 1–20.

[5] V. I. Arnold, V. V. Kozlov & A. I. Neishtadt – *Mathematical aspects of classical and celestial mechanics*, third éd., Encyclopaedia of Math. Sciences, vol. 3, Springer, 2006.

[6] P. Bernard – "Young measures, superposition and transport", *Indiana Univ. Math. J.* **57** (2008), p. 247–275.

[7] ———, "On the Conley decomposition of Mather sets", *Rev. Mat. Iberoam.* **26** (2010), p. 115–132.

[8] P. Bernard, V. Kaloshin & K. Zhang – "Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders", preprint arXiv:1112.2773, 3000effacer.

[9] M. Brin & G. Stuck – "Introduction to dynamical systems", in *Introduction to Dynamical systems*, Cambridge Univ. Press, 2002.

[10] C.-Q. Cheng – "Arnold diffusion in nearly integrable systems", preprint arXiv:1503.04153, 3000effacer.

[11] C. Conley – "The gradient structure of a flow. I", *Ergodic Theory Dynam. Systems* **8*** (1988), p. 11–26.

[12] A. Fathi – "Weak kam theorem in lagrangian dynamics, 10th preliminary version", book preprint, 2008.

[13] N. Gourmelon – "Adapted metrics for dominated splittings", *Ergodic Theory and Dynamical Systems* **27** (2007), p. 1839–1849.

[14] M. Guardia & V. Kaloshin – "Orbits of nearly integrable systems accumulating to kam tori", preprint arXiv:1412.7088, 3000effacer.

[15] V. Kaloshin & K. Zhang – "Arnold diffusion for three and half degrees of freedom", preprint http://www2.math.umd.edu/~vkaloshi/papers/announce-three-and-half.pdf, 2014.

[16] ______, "A strong form of Arnold diffusion for two and half degrees of freedom", preprint arXiv:1212.1150, 3000effacer.

[17] V. Kaloshin, J. N. Mather & E. Valdinoci – "Instability of totally elliptic points of symplectic maps in dimension 4", *Astérisque* **74** (2004), p. 79–116.

[18] V. Kaloshin & K. Zhang – "Dynamics of the dominant hamiltonian, with applications to Arnold diffusion", preprint arXiv:1410.1844, 3000effacer.

[19] R. Mañé – "Lagrangian flows: the dynamics of globally minimizing orbits", *Bol. Soc. Brasil. Mat. (N.S.)* **28** (1997), p. 141–153.

[20] J.-P. Marco – "Generic hyperbolic properties of classical systems on the torus $\mathbb{T}^2$", preprint, 2012.

[21] ______, "Generic hyperbolic properties of nearly integrable systems on $\mathbb{A}^3$", preprint, 2012.

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

[23] ______, "Arnold diffusion. I. Announcement of results", *Sovrem. Mat. Fundam. Napravl.* **2** (2003), p. 116–130.

[24] J. N. Mather – "Arnold diffusion. II", preprint, 2008.

[25] J. N. Mather – "Shortest curves associated to a degenerate Jacobi metric on $\mathbb{T}^2$", in *Progress in variational methods*, Nankai Ser. Pure Appl. Math. Theoret. Phys., vol. 7, World Sci. Publ., Hackensack, NJ, 2011, p. 126–168.

[26] R. McGehee – "The stable manifold theorem via an isolating block", in *Symposium on Ordinary Differential Equations (Univ. Minnesota, Minneapolis, Minn., 1972; dedicated to Hugh L. Turrittin)*, Lecture Notes in Math., vol. 312, Springer, 1973, p. 135–144.

[27] M. Shub – *Global stability of dynamical systems*, Lecture Notes in Math., vol. 583, Springer, 1987.

[28] C. L. Siegel – *Lectures on the geometry of numbers*, Springer, 1989.

[29] A. Sorrentino – "Lecture notes on Mather's theory for lagrangian systems", preprint arXiv:1011.0590, 3000effacer.