# On energy transferring in a periodic pendulum lattice with analytic weak couplings

HUANG Guan

**Abstract.** Consider a system of periodic pendulum lattice with analytic weak coupling:

$$\ddot{x}_i + \sin x_i = -\epsilon \sum_{j=i-2}^{i} \partial_{x_i} \beta_\alpha(x_j, x_{j+1}, x_{j+2}), \quad x_i = x_{i+N}, \quad i \in \mathbb{Z},$$

where $N \geqslant 3$ is an integer, $\epsilon > 0$ is a small parameter and the function $\beta_\alpha$ is an analytic function of a certain form. It is shown in this paper that for small enough $\epsilon$, the system admits motions such that the energy transfers between the pendulums in any predetermined order.

## 1. Introduction

### 1.1. Equation.

Let us consider a pendulum lattice with near couplings:

$$\ddot{x}_i + \sin x_i = -\epsilon \sum_{j=i-2}^{i} \partial_{x_i} \beta_\alpha(x_j, x_{j+1}, x_{j+2}), \quad i \in \mathbb{Z}, \tag{1.1}$$

where $0 < \epsilon \ll 1$ and

$$\beta_\alpha(x, y, z) = (1 + \cos x)(1 + \cos y)(1 + \cos z) f_\alpha(x) f_\alpha(y) f_\alpha(z), \quad x,\ y,\ z \in \mathbb{R}. \tag{1.2}$$

For $\alpha \in (0, 1)$, we define the function $f_\alpha$ to be a real analytic function on $\mathbb{R}$ satisfying the following conditions:

1. It is $2\pi$-periodic.
2. $0 \leqslant f_\alpha(z) \leqslant 1$, $\forall z \in \mathbb{R}$.
3. $f_\alpha(z) \geqslant 1 - \alpha$, if $|z - \mathbb{Z}2\pi| \leqslant \alpha$, and $f_\alpha(z) \leqslant \alpha^2$, if $|z - \mathbb{Z}2\pi| \geqslant 2\alpha$.

The explicit form of $f_\alpha$ is not important and we will choose $\alpha$ to be a small but fixed number, see Lemma 2.1 below.

Let us impose a periodic condition to system (1.1):

$$x_{i+N} = x_i, \quad \forall i \in \mathbb{Z}, \tag{1.3}$$

where $N \geqslant 3$ is a fixed integer.

Thus, the system (1.1) and (1.3) can be written into a Hamiltonian system in $\mathbb{R}^{2N}$,

$$\dot{x}_i = y_i, \quad \dot{y}_i = -\partial_{x_i} H(X,Y), \quad i = 1, \dots, N,$$

where $X = (x_1 \dots, x_N) \in \mathbb{R}^N$, $Y = (y_1, \dots, y_N) \in \mathbb{R}^N$ and the Hamiltonian function $H(X,Y)$ (the total energy) is defined as

$$H(X,Y) = \sum_{i=1}^{N} \frac{y_i^2}{2} + V(x_i) + \epsilon\beta_\alpha(x_i, x_{i+1}, x_{i+2}),$$

with the pendulum potential $V(x) = -1 - \cos x$.

### 1.2. Result

For any solution $(x_1(t), \dots, x_N(t))$ of the system (1.1) and (1.3), let us denote the energy corresponding to each pendulum as

$$E_i(t) = \frac{\dot{x}_i^2(t)}{2} - \big(1 + \cos x_i(t)\big), \quad i = 1, \dots, N.$$

In this paper, the main interest is focused on the long time behaviors of the quantities $E_i$, $i = 1, \dots, N$. The system (1.1) is nearly integrable if $\epsilon$ is small enough. Due to the famous KAM theory, a large part (in the sense of the Lebesgue measure) of the phase space is foliated by the KAM tori. Particularly, for most of the initial data the energy $E_i(t)$ stays close to its initial value for all the time. Nevertheless, we will show that there exist certain motions of the system (1.1) such that the corresponding energy would slowly transfer along the chain of pendulums in any predetermined way, propagating to the right or to the left by any distances and for any moment the energy is almost concentrated in at most two pendulums. The precise statement is as follows.

Let us call a sequence

$$\cdots \sigma_{i-1}\sigma_i\sigma_{i+1} \cdots, \quad \sigma_i \in \{1, \dots, N\}, \quad i \in \mathbb{Z}$$

a regular path on $\mathbb{Z}$ if

$$|\sigma_i - \sigma_{i+1}| = 1, \quad \forall i \in \mathbb{Z}.$$

In the virtue of the periodic condition (1.3), we set $|1 - N| = 1$. The main result is the following theorem.

**Theorem 1.1.** *For any $r > 0$ and $N \geqslant 3$, there exists $\epsilon_0 > 0$ such that for any $0 < \epsilon < \epsilon_0$, and any sequence $\cdots \sigma_{-1}\sigma_0\sigma_1 \cdots \in \{1, \dots, N\}^{\mathbb{Z}}$ that is a regular path on $\mathbb{Z}$, there is a solution $\gamma(t)$ of the system (1.1) with total energy $H = 1$ and a sequence of time instances $\cdots t_{-1}t_0t_1 \cdots$ such that the energy of each pendulum satisfies*

$$|E_{\sigma_j}\big(\gamma(t_j)\big) - 1| \leqslant C\sqrt{\epsilon}, \quad |E_i\big(\gamma(t_j)\big)| \leqslant C\sqrt{\epsilon},\ i \neq \sigma_j, \tag{1.4}$$

*and*

$$0 < t_{j+1} - t_j < C\epsilon^{-4-r}, \tag{1.5}$$

*where $C$ is a universal constant.*

Particularly, if we choose the sequence to be a monotone,

$$\cdots 123 \cdots N 123 \cdots N \cdots ,$$

then the orbit we obtain may be viewed as a (very slow) "traveling wave" in in the pendulum lattice (1.1).

### 1.3. Strategy of the proof

To fix ideas, below we shall concentrate on the case of four pendulums ($N = 4$). The general cases are quite similar (see Appendix A). The proof is based on the variational methods initiated by Mather [22] and Bessi [3, 4]. Let us consider the Lagrangian

$$\mathcal{L}_\epsilon = \sum_{i=1}^{4} (\frac{\dot{x}_i^2}{2} + 1 + \cos x_i) - \epsilon \sum_{i=1}^{4} \beta_\alpha(x_i, x_{i+1}, x_{i+2}), \tag{1.6}$$

where we have adopted the periodic condition (1.3). Clearly, the corresponding Euler-Lagrange equation is exactly the same as the system (1.1). Generally, the orbits $\gamma$ are chosen to minimize the Lagrangian action $\int_\gamma \mathcal{L}_\epsilon dt$ under certain boundary conditions. In this work, we use a slightly different version of this approach. Instead of directly minimizing the Lagrangian action, using the Maupertuis principle, we will minimize the geodesic distance of the associated Jacobi metric on the energy surface, and construct the orbits as the geodesics of the Jacobi metric (so the orbits we get would have the prefixed total energy). These (long) geodesics are obtained by concatenating (short) geodesic segments that follow a well-arranged itinerary. This approach is sometimes referred as *the methods of broken geodesics*. Similar constructions were also carried out in [20, 5, 16, 15, 17, 25].

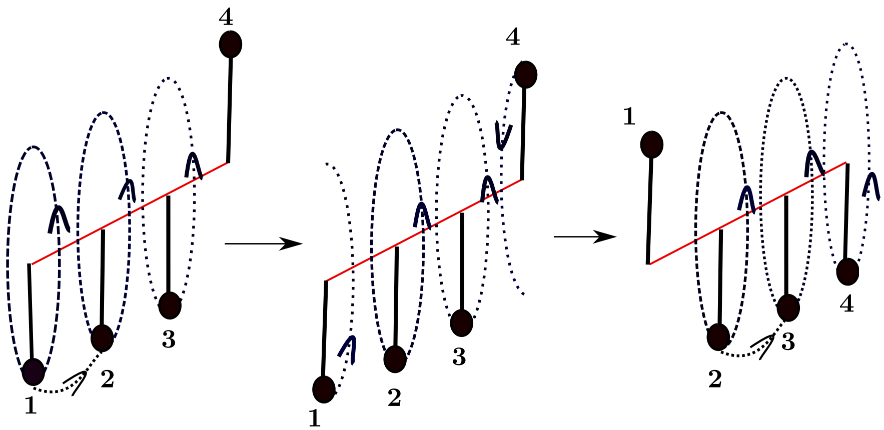


FIGURE 1. The stages of energy transferring and advancing.

Following [17], the itineraries that we construct have two types of stages: *energy transferring* and *advancing*. At the stage of energy transferring, e.g.

from the pendulum **1** to the pendulum **2**, only three pendulums, **1**, **2** and **3**, are "active", while the pendulum **4** keeps "sleeping" near the "top" ($x_4 = \pi$) throughout the whole stage (see the left of Figure 1). In the full period of this stage, there are many substages. For each substage, a small amount of energy transfers from **1** to **2**, while the pendulum **3** moves closely to a heteroclinic orbit of the system: its energy remains close to zero. At the end of the energy-transferring stage, the pendulum **2** takes up most of the energy, while the pendulums **1** and **3** have energy close to zero. There begins the stage of advancing. At this stage, the pendulum **4** slowly drops from the top to the "bottom" ($x_4 = 2\pi$) and the pendulum **1** slowly climbs up to the top, while the pendulum **2** rotates with speed $O(1)$ and the pendulum **3** moves along a heteroclinic orbit of the system (see the middle of Figure 1). After this stage, the three active pendulums would "advance" from **1**, **2** and **3** to **2**, **3** and **4**. Then there follows a new stage of energy transferring, e.g. from **2** to **3** (see the right of Figure 1). The detailed construction of these stages is presented in Paragraphs (I)-(III) in Section 6.

### 1.4. Further discussion

*Arnold diffusion.* In his celebrated paper [2], Arnold constructed a nearly integrable system with 2.5 degrees of freedom, showing that despite the fact that the KAM tori fill up most of the phase space of a nearly integrable system with many degrees of freedom ($> 2$), there may still exist orbits such that the action variables of the system exhibit a change of order 1 in a long time span. Now these phenomena of the large variations of the action variables of a nearly integrable system are called *Arnold diffusion.* Arnold conjectured that they should exist for generic (analytic or smooth) perturbations of an integrable system. Arnold diffusion has been a hot field of researches in the past decades. A survey (even a very short one) on the related results is out of the scope of the current paper. The reader is kindly referred to [24, 9, 8, 18, 6, 19, 7] and the references therein for several recent breakthroughs in the category of systems with finite smoothness and for historical remarks. However, it seems that the current sophisticated theories developed for Arnold diffusion, in the category of finitely many smooth systems, could not reach the scope of analytic nearly integrable systems. The result in this paper provides a good example of Arnold diffusion with $N$ degrees of freedom in the category of analytic systems. It captures some of the most important features of Arnold diffusion: Eg. the stages of energy transferring bear strong resemblances to the processes of passing single-resonant zones and the the stages of advancing are analogous to the crossing of double-resonant regions. The author believes that this result provides useful insights to the studies of Arnold diffusion in the analytic systems.

*Energy propagations in lattice systems.* Localization or delocalization are important physical phenomena observed in the context of wave propagation through disordered media, which was often modeled with lattice systems. When the disorder is significant, certain types of waves get trapped and wave propagation is extremely slow or absent. These phenomena were firstly systematically studied by Anderson ([1]), whom now these phenomena were named after. Nowadays, Anderson localization is extensively studied by mathematicians and physicists. Mathematically, it could be partially explained by the KAM theories and the Nekhoroshev estimates (see, e.g. [12]): these motions lie on the KAM tori, so no propagation occurs; or their rates of propagation are just too slow to be observed experimentally.

For the delocalization or energy propagation, rigorous results are rare (nonetheless, see [17][1]). In this direction, there are several essential questions asked by the physicists (see, e.g. [13]):

1. Is energy propagation possible when the interactions are weak?
2. If it is possible, what's the speed of the propagation? Is it polynomial or exponential (with respect to the intensity of the interactions)?
3. Does the speed depend on the range of interactions, the size of the lattice system, or both?

The result in the present paper gives an affirmative answer to the first question, at least for the particular lattice system (1.1), and the speed of propagation is polynomial (see (1.5)). Due to the special form of the interaction, given in (1.2), the range of interaction is fixed in the system (1.1), so no information of the third question is provided here. It would be interesting to investigate whether Theorem 1.1 still holds true for the system (1.1) with any other form of interactions. For other physical lattice models, see e.g. [13].

Perhaps, the most famous lattice lattice system is the FPU lattice, introduced by Fermi, Pasta, and Ulam in their work [11]. Most of the small-amplitude solutions of the FPU lattice do not exhibit energy propagation, due to the fact that near the equilibrium, the FPU system is nearly integrable (see, e.g. [14]). Proving the existence of energy propagation in the FPU lattice is still a wide open problem of great interest.

### 1.5. Plan of the paper

The paper is organized as follows: In Section 2, we introduce the Melnikoff function and choose the parameter $\alpha$ such that the Melnikoff function possess 'local-minimum'-like property at certain points. In Section 3, we study a modified Lagrangian (3.1), and show that its minimum actions under certain boundary conditions can be approximated by the Melnikoff function (up to a fixed constant). Mather's Barrier function for the modified Lagrangian (3.1) is studied in Section 4. We discuss the Maupertius principle in Section 5 and rewrite the results from Sections 3 and 4 in the context of the geodesic

[1] In [17], a result similar to Theorem 1.1 was proven for the system (1.1) with interaction function $\tilde{\beta}_\epsilon$ of the form,

$$\tilde{\beta}_\epsilon(\tilde{y}) = \epsilon^k \sum_{n\in\mathbb{Z}^3} \eta\Big(\frac{|\tilde{y}-2\pi n|}{\epsilon}\Big), \quad \tilde{y}\in\mathbb{R}^3,$$

where $k \geqslant 3$ is a fixed number and $\eta : \mathbb{R}_+ \to \mathbb{R}$ is a $C^\infty$ bump function such that $\eta(x) > 0$ if $|x| < 1$ and $\eta(x) = 0$ if $|x| \geqslant 1$. Notice that unlike the interaction $\beta_\alpha$ in (1.2), which is independent of the small parameter $\epsilon$ and analytic, the $C^{k-1}$-norm of $\tilde{\beta}_\epsilon$ tends to 0 as $\epsilon \to 0$, while the $C^{k+1}$-norm is unbounded as $\epsilon \to 0$.

distance corresponding to the associated Jacobi metric. Finally, Theorem 1.1 is proved in Section 6.

## 2. Melnikoff function

Let us set up some notations. For $-\infty \leqslant a < b \leqslant +\infty$, we denote $H^1([a,b],\mathbb{R}^n)$ the space of absolutely continuous functions with the $H^1$-norm

$$\|u\|_1^2 = \int_a^b \left(|u|^2 + |\dot{u}|^2\right) \mathrm{dt}.$$

For any two functions $f$, $g$ on $\mathbb{R}^n$ and $\lambda > 0$, we write

$$f = g \pm \lambda$$

if

$$-\lambda + g(x) \leqslant f(x) \leqslant g(x) + \lambda, \quad \forall x \in \mathbb{R}^n.$$

Since the potential of the system (1.1) is $2\pi$-periodic in all variables, we thus can take $\mathbb{T}^4$ as our configuration space, where

$$\mathbb{T} = [-\pi,\pi]/\{-\pi,\pi\}.$$

It is easy to see that the system (1.1), for any $\epsilon$, admits several families of 2-dimensional invariant tori, precisely, for any $h, \tilde{h} \geqslant 0$, we have

$$\mathcal{T}_i^2 = \begin{cases} \frac{\dot{x}_i^2}{2} - (1+\cos x_i) = h, \\ \frac{\dot{x}_{i+1}^2}{2} - (1+\cos x_{i+1}) = \tilde{h}, \\ x_{i+2} = x_{i+3} = \pi, \\ \dot{x}_{i+2} = \dot{x}_{i+3} = 0, \end{cases} \qquad i = 1,\ldots,4,$$

where we adopt the periodic condition $x_{i+4} = x_i$, $i = 1,\ldots,4$. These families are parametrized by $(h,\tilde{h})$ and belong to the energy surface $H = h + \tilde{h}$.

Recall that the pendulum of the Lagrangian $\frac{\dot{x}^2}{2} + (1+\cos x)$ has heteroclinic solution of the form

$$q_0(t) = -\pi + 4\arctan(e^t), \quad \dot{q}_0(t) = \frac{2}{\cosh(t)}.$$

Let us denote $\mathcal{A}^\pm(x)$, $x \in \mathbb{T}$, the action of the pendulum Lagrangian along this heteroclinic orbit,

$$\mathcal{A}^+(x) = \int_0^{+\infty} \left\{\frac{[\dot{q}_0(t_0+t)]^2}{2} + [1+\cos(q_0(t_0+t))]\right\}\mathrm{dt},$$

and

$$\mathcal{A}^-(x) = \int_{-\infty}^{0} \left\{\frac{[\dot{q}_0(t+t_0)]^2}{2} + [1+\cos(q_0(t+t_0))]\right\}\mathrm{dt},$$

where $q_0(t_0) = x$. Clearly, $\mathcal{A}^+(0) = \mathcal{A}^-(0)$ and we will simply denote $\mathcal{A}^+(0)$ as $\mathcal{A}_0$. Also, let us set

$$\mathcal{A}(x) = \min\{\mathcal{A}^+(x), \mathcal{A}^-(x)\}, \quad x \in \mathbb{T}. \tag{2.1}$$

For any $x \in \mathbb{R}$ and $h \geqslant 0$, let $q_h(x,t)$ be the solution of the free pendulum with energy $h$ and positive speed such that $q_h(x,0) = x$. See Figure 2.

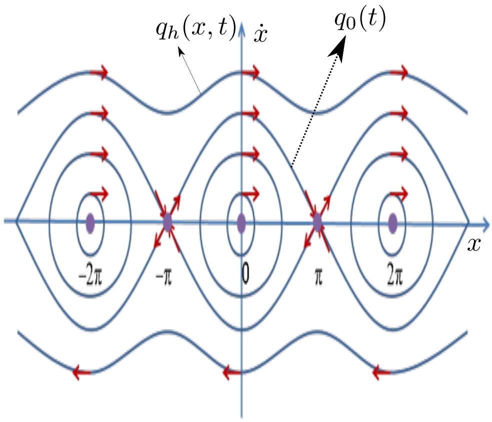


FIGURE 2. The phase portrait of a free pendulum.

In the case $h > 0$, for any $x, y \in \mathbb{R}$, there exists a unique $T \in \mathbb{R}$ such that $q_h(x,T) = y$. Let us define a metric on $\mathbb{R}$ by

$$\mathcal{S}^h(x,y) = \int_0^T \Big[\frac{\dot{q}_h^2(x,t)}{2} + 1 + \cos(q_h(x,t))\Big]\mathrm{d}t, \quad \mathrm{h} > 0. \tag{2.2}$$

Slightly abusing notations, we also denote

$$\mathcal{S}^0(x,y) = \mathcal{A}^+(x) + \mathcal{A}^-(y), \quad x,\ y \in \mathbb{T}. \tag{2.3}$$

For any $(x,y) \in \mathbb{R}^2$, and $h, \tilde{h} \geqslant 0$, we introduce the Melnikoff function $\mathcal{M}_\alpha(h,\tilde{h},x,y)$ as

$$\mathcal{M}_\alpha(h,\tilde{h},x,y) = -\int_{\mathbb{R}} \beta_\alpha\big(q_0(t), q_h(x,t), q_{\tilde{h}}(y,t)\big)\mathrm{d}t,$$

where the interaction $\beta_\alpha$ is defined in (1.2). The Melnikoff function can be seen as the integral of the perturbation along a heteroclinic orbit of the unperturbed system to the invariant torus $\mathcal{T}_i^2$ with the energy level $h + \tilde{h}$. Similarly let us define

$$\mathcal{M}_\alpha^+(h,\tilde{h},x,y) = -\int_0^{+\infty} \beta_\alpha(q_0(t), q_h(x,t), q_{\tilde{h}}(y,t))\mathrm{d}t,$$

and

$$\mathcal{M}_\alpha^-(h,\tilde{h},x,y) = -\int_{-\infty}^0 \beta_\alpha(q_0(t), q_h(x,t), q_{\tilde{h}}(y,t))\mathrm{d}t.$$

We will call them the one sided Melnikoff functions. It is easy to see that these functions are well defined, $2\pi$-periodic in $x, y$, and continuous with respect to all their arguments. Moreover, we have

$$\mathcal{M}_\alpha(h,\tilde{h},x,y)=\mathcal{M}_\alpha^+(h,\tilde{h},x,y)+\mathcal{M}_\alpha^-(h,\tilde{h},x,y).$$

We would like to choose a number $\alpha$ in the definition of the function $\beta_\alpha$ such that the Melnikoff function $\mathcal{M}_\alpha(h,\tilde{h},x,y)$, as a function of $(x,y)$, has 'local minimum'-like property at $(0,0)$. This is done in the following lemma.

**Lemma 2.1.** *For any $\delta\in(0,\frac{1}{2})$, there exist positive numbers $\alpha$ and $\kappa$ such that $\forall h,\tilde{h}\in[0,1]$, we have*

$$\inf\Big\{\mathcal{M}_\alpha(h,\tilde{h},x,y):\ (x,y)\in\partial\big([-\delta,\delta]\times[-\delta,\delta]\big)\Big\}-\mathcal{M}_\alpha(h,\tilde{h},0,0)>\kappa. \tag{2.4}$$

*Proof.* For $|t|<\frac{\alpha}{4}$, we have $|q_0(t)|,\ |q_h(0,t)|,\ |q_{\tilde{h}}(0,t)|\leqslant\alpha$, and

$$f_\alpha(q_0(t))f_\alpha(q_h(0,t))f_\alpha(q_{\tilde{h}}(0,t))\geqslant(1-\alpha)^3.$$

So for $\alpha$ small enough,

$$-\mathcal{M}_\alpha(h,\tilde{h},0,0)\geqslant\int_{|t|\leqslant\frac{\alpha}{4}}(1-\alpha)^3\mathrm{d}t\geqslant\frac{\alpha}{16}.$$

On the other side, if $|x|=\delta$ and $|y|\leqslant\delta$, we have that for $\alpha$ small enough,

$$0\leqslant f_\alpha(q_0(t))f_\alpha(q_h(x,t))\leqslant\alpha^2,\quad\forall t\in\mathbb{R},$$

implying that

$$-\mathcal{M}_\alpha(h,\tilde{h},q_h(x,t),q_{\tilde{h}}(y,t))\leqslant4\alpha^2\int_{\mathbb{R}}[1+\cos(q_0(t))]\mathrm{d}t,$$

Thus we have

$$-\mathcal{M}_\alpha(h,\tilde{h},x,y)\leqslant8\mathcal{A}_0\alpha^2,\quad\text{if}\quad|x|=\delta,\ |y|\leqslant\delta.$$

Analogously, for $|x|\leqslant\delta$ and $|y|=\delta$, we have

$$-\mathcal{M}_\alpha(h,\tilde{h},x,y)\leqslant8\mathcal{A}_0\alpha^2,\quad\text{if}\quad|x|\leqslant\delta,\ |y|=\delta.$$

This implies the assertion of the lemma. □

We now choose three numbers

$$(\delta,\alpha,\kappa),$$

such that the inequality (2.4) holds true and keep them *fixed* throughout the whole paper. From now on, we will drop the subindex $\alpha$ of the Melnikoff functions and denote them as $\mathcal{M}$ and $\mathcal{M}^\pm$.

By the continuity of the one sided Melnikoff functions, there exists $\lambda_0>0$ such that for any $h,h',\tilde{h},\tilde{h}'\in[0,1]$, if

$$|h-h'|+|\tilde{h}-\tilde{h}'|\leqslant\lambda_0,$$

then

$$\Big|\mathcal{M}^\pm(h,\tilde{h},x,y)-\mathcal{M}^\pm(h',\tilde{h}',x,y)\Big|<\frac{\kappa}{33},\quad\forall x,\ y\in\mathbb{R}. \tag{2.5}$$

From the definition of functions $\mathcal{A}^{\pm}$, we know that there exists $\bar{\delta} > 0$ such that

$$\mathcal{A}^{-}(-\pi + \bar{\delta}) = \mathcal{A}^{+}(\pi - \bar{\delta}) = \frac{\kappa}{400}. \tag{2.6}$$

We also keep $\lambda_0$ and $\bar{\delta}$ *fixed* in the rest of the paper.

## 3. Modified Lagrangian

For any $h_1, h_2 \in [0,1]$, we consider the following modified Lagrangian

$$\mathcal{L}_{\epsilon}^{h_1,h_2} = \sum_{i=1}^{2} \frac{1}{2}(\dot{x}_i - v_{h_i}(x_i))^2 + \sum_{i=3}^{4} \frac{\dot{x}_i^2}{2} + [1 + \cos x_i] - \epsilon \sum_{i=1}^{4} \beta_\alpha(x_i, x_{i+1}, x_{i+2}), \tag{3.1}$$

where we adopt the periodic condition $x_i = x_{i+4}$, $i = 1, \ldots, 4$, and

$$v_h(x) = \sqrt{2(h + 1 + \cos x)}. \tag{3.2}$$

Direct calculation shows that the Euler-Lagrange equation corresponds to the modified Lagrangian (3.1) is the same as the system (1.1).

**Lemma 3.1.** *For any $h_1, h_2 \in [0,1]$, let $(x_1(t), x_2(t), x_3(t), x_4(t))$ minimize*

$$\int_0^T \mathcal{L}_{\epsilon}^{h_1,h_2} \mathrm{dt},$$

*with the boundary conditions:*

$$\begin{cases} x_1(0) = x, & x_1(T) = q_{h_1}(x, T), \\ x_2(0) = y, & x_2(T) = q_{h_2}(y, T), \\ x_3(0) = 0, & x_3(T) = \pi, \\ x_4(0) = z, & x_4(T) = \pi, \end{cases} \tag{3.3}$$

*where $x, y \in \mathbb{R}$ are any real numbers and $|z - \pi| \leqslant \bar{\delta}$.*

*Then there exists $\epsilon_0 > 0$, which is independent of $h_1$ and $h_2$, such that for $0 < \epsilon \leqslant \epsilon_0$ and $T \geqslant \epsilon^{-1}$, we have*

$$\int_0^T \mathcal{L}_{\epsilon}^{h_1,h_2}(x_1, x_2, x_3, x_4)\mathrm{dt} = \mathcal{A}_0 + \mathcal{A}(z) + \epsilon \mathcal{M}^{+}(h_1, h_2, x, y) \pm \epsilon \frac{\kappa}{32}.$$

*Proof.* Without loss of generality, we assume $z < \pi$. Then

$$\mathcal{A}(z) = \mathcal{A}^{+}(z).$$

The key of the proof is to show that $(x_1(t), x_2(t), x_3(t), x_4(t))$ approximates well the orbit $(q_{h_1}(x,t), q_{h_2}(y,t), q_0(t), q_0(z,t))$ of the free pendulums for small enough $\epsilon$. Let us consider the curve $\gamma_T : [0, T] \to \mathbb{R}^4$ defined by

$$\gamma_T(t) = \begin{cases} \big(q_{h_1}(x,t), q_{h_2}(y,t), q_0(t), q_0(z,t)\big), & t \in [0, T-1], \\ \big(q_{h_1}(x,t), q_{h_2}(y,t), \bar{q}_0(t), \bar{q}_0(z,t)\big), & t \in [T-1, T], \end{cases}$$

where

$$\bar{q}_0(t) = q_0(T-1) + (\pi - q_0(T-1))(t - T + 1),$$

and

$$\bar{q}_0(z,t) = q_0(z,T-1) + (\pi - q_0(z,T-1))(t-T+1).$$

Clearly $\gamma_T$ satisfies the boundary conditions (3.3). Since $(x_1,x_2,x_3,x_4)(t)$ minimizes $\int_0^T \mathcal{L}_\epsilon^{h_1,h_2}\mathrm{dt}$, then if $\epsilon$ is small enough, we have

$$\begin{aligned}&\int_0^T \mathcal{L}_\epsilon^{h_1,h_2}(x_1(t),x_2(t),x_3(t),x_4(t))\mathrm{dt} \leqslant \int_0^T \mathcal{L}_\epsilon^{h_1,h_2}(\gamma_T(t))\mathrm{dt}\\ &\leqslant \mathcal{A}_0 + \mathcal{A}^+(z) + \epsilon\mathcal{M}^+(h_1,h_2,x,y) + 12\epsilon\mathcal{A}^+(z) + O(e^{-\frac{1}{2\epsilon}})\end{aligned}$$

Obviously, the terms in the right hand side are bounded as $\epsilon \to 0$. From the form of the modified Lagrangian $\mathcal{L}_\epsilon^{h_1,h_2}$, we thus have

$$\int_0^T \sum_{i=3}^4 \big(\frac{1}{2}\dot{x}_i^2(t) + 1 + \cos x_i(t)\big)\mathrm{dt} - \epsilon\int_0^T \sum_{i=1}^4 \beta_\alpha(x_i,x_{i+1},x_{i+2})\mathrm{dt} \leqslant C_1.$$

From the definition of the function $\beta_\alpha$ we have that

$$\sum_{i=1}^4 \beta_\alpha(x_i,x_{i+1},x_{i+2}) \leqslant C_2\sum_{i=3}^4 \big(1+\cos x_i\big).$$

Thus, for $\epsilon$ small enough,

$$\begin{aligned}&\int_0^T [\sum_{i=3}^4 \frac{\dot{x}_i^2}{2} + (1+\cos x_i)]\mathrm{dt} - \epsilon\int_0^T \sum_{i=1}^4 \beta_\alpha(x_i,x_{i+1},x_{i+2})\mathrm{dt}\\ &\geqslant \frac{1}{2}\int_0^T \sum_{i=3}^4 (1+\cos x_i + \frac{1}{2}\dot{x}_i^2)\mathrm{dt},\end{aligned}$$

which implies

$$\int_0^T \sum_{i=3}^4 (\frac{\dot{x}_i^2}{2} + 1 + \cos x_i)\mathrm{dt} \leqslant C_3. \tag{3.4}$$

Therefore,

$$\Big|\int_0^T \sum_{i=1}^4 \beta_\alpha(x_i,x_{i+1},x_{i+2})\mathrm{dt}\Big| \leqslant C_4.$$

That leads to

$$\Big|\int_0^T [\frac{\dot{x}_3^2}{2} + (1+\cos x_1)]\mathrm{dt} - \mathcal{A}_0\Big| \leqslant C\epsilon, \tag{3.5}$$

and

$$\Big|\int_0^T [\frac{\dot{x}_4^2}{2} + (1+\cos x_4)]\mathrm{dt} - \mathcal{A}^+(z)\Big| \leqslant C\epsilon, \tag{3.6}$$

By (3.5), we know that $x_3(t)$, as $\epsilon \to 0$, is a minimizing sequence of the pendulum Lagrangian in the class of orbits connecting 0 and $\pi$. So

$$x_3(t) \to q_0(t), \quad \text{in} \quad H^1(\mathbb{R}_+,\mathbb{T}), \quad \text{as} \quad \epsilon \to 0.$$

We set

$$E_1(t) = \frac{\dot{x}_1^2(t)}{2} - [1+\cos x_1(t)].$$

Since

$$\ddot{x}_1 = -\sin(x_1) - \epsilon\partial_{x_1}\sum_{i=1}^{3}\beta_\alpha(x_{i-2},x_{i-1},x_i),$$

with

$$\partial_a\beta_\alpha(a,b,c) = [-\sin a f_\alpha(a)+(1+\cos a)f'_\alpha(a)](1+\cos b)(1+\cos c)f_\alpha(b)f_\alpha(c), \quad \forall a,b,c\in\mathbb{R},$$

we have

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}E_1(t)\Big| = |\dot{x}_1\ddot{x}_1+\sin(x_1)\dot{x}_1| = \epsilon\Big|\dot{x}_1\cdot\partial_{x_1}\sum_{i=1}^{3}\beta_\alpha(x_{i-2},x_{i-1},x_i)\Big| \leqslant \epsilon C\sum_{i=3}^{4}(1+\cos x_i).$$

If $\epsilon$ is small enough, for any $0\leqslant t'<t''\leqslant T$, by (3.4), we have

$$E_2(t'')-E_2(t') = \int_{t'}^{t''}\frac{\mathrm{d}}{\mathrm{d}t}E_2(t)\mathrm{d}t \leqslant \epsilon C\int_{t'}^{t''}\sum_{i=3}^{4}(1+\cos(x_i(t))\mathrm{d}t \leqslant C'\epsilon. \quad (3.7)$$

Since $x_1$ covers the distance $q_{h_1}(x,T)-x$ in the time $T$, therefore there must exist some $t_0\in[0,T]$ such that $E_1(t_0)=h_1$. So

$$|E_1(t)-h_1|\leqslant C'\epsilon, \quad t\in[0,T].$$

This implies that for any given $R>0$,

$$\|x_1(t)-q_{h_1}(x,t)\|_{L^\infty([0,R])}\to 0 \quad \text{as} \quad \epsilon\to 0.$$

Similar relation is true for $x_2(t)$ and $q_{h_2}(y,t)$. Therefore we have

$$\Big|-\int_0^T\beta_\alpha(x_1,x_2,x_3)\mathrm{d}t-\mathcal{M}^+(h_1,h_2,x,y)\Big|\to 0 \quad \text{as} \quad \epsilon\to 0.$$

Due to (3.6) and (2.6), we know that for small enough $\epsilon$,

$$\Big|\int_0^T\sum_{i=2}^{4}\beta_\alpha(x_i,x_{i+1},x_{i+2})\mathrm{d}t\Big| \leqslant 12\mathcal{A}^+(z)+C\epsilon\leqslant\frac{\kappa}{33}.$$

It is obvious that

$$\int_0^T\mathcal{L}_\epsilon^{h_1,h_2}(x_1,x_2,x_3,x_4)\mathrm{d}t \geqslant \mathcal{A}_0+\mathcal{A}^+(z)-\epsilon\int_0^T\sum_{i=1}^{4}\beta_\alpha(x_i,x_{i+1},x_{i+2})\mathrm{d}t.$$

Thus follows the assertion of the lemma. □

Similarly, we obtain the following lemma.

**Lemma 3.2.** *For any $h_1,h_2\in[0,1]$, let $(x_1(t),x_2(t),x_3(t),x_4(t))$ minimize*

$$\int_{-T}^{0}\mathcal{L}_\epsilon^{h_1,h_2}\mathrm{d}t,$$

*with the boundary conditions:*

$$\begin{cases} x_1(0)=x, & x_1(-T)=q_{h_1}(x,-T),\\ x_2(0)=y, & x_2(-T)=q_{h_2}(y,-T),\\ x_3(0)=0, & x_3(-T)=-\pi,\\ x_4(0)=z, & x_4(-T)=-\pi,\end{cases}$$

*where $x$, $y \in \mathbb{R}$ are any real numbers and $|-\pi - z| \leqslant \bar{\delta}$. Then there exists $\epsilon_1 > 0$, which is independent of $h_1$ and $h_2$, such that for $0 < \epsilon \leqslant \epsilon_1$ and $T \geqslant \epsilon^{-1}$, we have*

$$\int_{-T}^{0} \mathcal{L}_{\epsilon}^{h_1,h_2}(x_1,x_2,x_3,x_4)\mathrm{dt} = \mathcal{A}_0 + \mathcal{A}(z) + \epsilon\mathcal{M}^{-}(h_1,h_2,x,y) \pm \epsilon\frac{\kappa}{32}.$$

Due to the symmetry of the system, we can interchange the positions of $x_3$ and $x_4$ and get the same results. It is not difficult to see that similar results also hold for the modified Lagrangians

$$\mathcal{L}_{\epsilon,i}^{h_1,h_2} = \sum_{j=i}^{i+1}\frac{1}{2}(\dot{x}_j - v_{h_j}(x_j))^2 + \sum_{j=i+2}^{i+3}\frac{\dot{x}_j^2}{2} + 1 + \cos x_j - \epsilon\sum_{j=1}^{4}\beta_\alpha(x_j,x_{j+1},x_{j+2}), \tag{3.8}$$

where $i = 1,\ldots,4$ and $h_j \in [0,1]$.

## 4. Mather's barrier function

Firstly, let us introduce some basic knowledge of the Mather theory.

Let $M$ be a compact connected Riemannian manifold, and let $L(\theta,v) : TM \to \mathbb{R}$ be a $C^2$ function with the following properties:

i) (Positive definite) $\partial_{vv}^2 L(\theta,v)$ as a bilinear form is strictly positive definite.

ii) (Superlinear growth on the fiber)

$$\lim_{\|v\|\to\infty}\frac{L(\theta,v)}{\|v\|} = \infty, \quad \forall\theta\in M,\ v\in T_\theta M.$$

We define

$$c[0] = -\inf_{\mu}\int L d\mu,$$

where the infimum is taken over all Borel probability measure $\mu$ on $TM$ that is invariant under the Euler-Lagrange flow of $L$. The number $c[0]$, sometimes called the Mañé critical value, has very deep meanings in dynamics and it plays a essential role in the Mather theory (and the weak KAM theory). The reader is kindly referred to e.g. [21, 22, 10, 23] for more information. All we need here is its definition.

For $T > 0$, let

$$\mathcal{B}_T(\theta_0,\theta_1) = \inf_{\gamma}\int_0^T L(\gamma,\dot{\gamma})\mathrm{dt},$$

where the infimum is taken over all absolutely continuous curves $\gamma : [0,T] \to M$ such that $\gamma(0) = \theta_0$, $\gamma(T) = \theta_1$. The following function is usually referred to as the Peierls barrier, defined by Mather [22]:

$$\mathcal{B}(\theta_0,\theta_1) = \liminf_{T\to\infty}(\mathcal{B}_T(\theta_0,\theta_1) + c[0]T).$$

The *liminf* is indeed a limit in the autonomous case (Corollary 6.3.3 in [10]). Nonetheless, we still use Mather's original definition, since it fits our purpose here.

We have the following statement:

**Proposition 4.1.** *([22] or Corollary 5.3.3 in [10]) The function $\mathcal{B}(\cdot,\cdot)$ is finite, Lipschitz, and satisfies the triangle inequality $\mathcal{B}(\theta_0,\theta_2)\leqslant\mathcal{B}(\theta_0,\theta_1)+\mathcal{B}(\theta_1,\theta_2)$.*

Now come back to our case in which

$$M=\mathbb{T}^2\times 2\mathbb{T}\times 2\mathbb{T}.$$

(We choose the double cover $2\mathbb{T}$ here so that we can distinguish between $\pi$ and $-\pi$ in the configuration space.) It is easy to see that the Lagrangian $\mathcal{L}_\epsilon^{h_1,h_2}$, introduced in (3.1), is well defined on $TM$ and satisfies the conditions i) and ii). From the special form of the Lagrangian $\mathcal{L}_\epsilon^{h_1,h_2}$, we know that for small enough $\epsilon$,

$$\mathcal{L}_\epsilon^{h_1,h_2}\geqslant 0.$$

Particularly, it vanishes on the set

$$\mathcal{T}_{h_1,h_2}=\{x_1,\dot{x}_1=v_{h_1}(x_1),x_2,\dot{x}_2=v_{h_2}(x_2),x_3=x_4=\pm\pi,\dot{x}_3=\dot{x}_4=0\},$$

where $v_h$ is defined in (3.2). Therefore, we have $c[0]=0$ and the infimum is attained by the invariant measure, which is supported on the torus $\mathcal{T}_{h_1,h_2}$.

We will denote

$$\mathcal{B}_{\epsilon,T}(\cdot,\cdot)\quad\text{and}\quad\mathcal{B}_\epsilon(\cdot,\cdot),$$

for the actions and the Peierls barrier of the Lagrangian $\mathcal{L}_\epsilon^{h_1,h_2}$. As a consequence,

$$\mathcal{B}_\epsilon(\theta_0,\theta_1)=\liminf_{T\to\infty}\mathcal{B}_{\epsilon,T}(\theta_0,\theta_1).$$

Moreover the sets

$$\mathbb{A}_\pi:=\{x_3=x_4=\pi\}\subset M\quad\text{and}\quad\mathbb{A}_{-\pi}=\{x_3=x_4=-\pi\},$$

have the property that the Peierls barrier vanishes in them.

**Lemma 4.2.** *Let $\theta_0,\theta_1\in\mathbb{A}_\pi\subset M$. Then $\mathcal{B}_\epsilon(\theta_0,\theta_1)=0$. The same result holds for points in $\mathbb{A}_{-\pi}$.*

*Proof.* Let $\theta_0=(x_1^0,x_2^0,\pi,\pi)$ and $\theta_1=(x_1^1,x_2^1,\pi,\pi)$. For every $T>0$, we consider

$$\gamma_T(t)=\big(q_{h_1}(x_1^0,t),q_{h_2}(x_2^0,t),\pi,\pi\big)+\Big(\frac{x_1^1-q_{h_1}(x_1^0,T)}{T}t,\frac{x_2^1-q_{h_2}(x_2^0,T)}{T}t,0,0\Big).$$

Clearly $\gamma_T(t)$ is absolutely continuous and $\gamma_T(0)=\theta_0$, $\gamma_T(T)=\theta_1$. We have

$$\int_{\gamma_T}\mathcal{L}_\epsilon^{h_1,h_2}\mathrm{dt}\leqslant\int_0^T\frac{C}{T^2}\mathrm{dt}=\frac{C}{T}\to 0\ \text{ as } T\to\infty.$$

Therefore $\mathcal{B}_\epsilon(\theta_0,\theta_1)=0$. □

For any point $\theta_0 \in \mathbb{A}_{-\pi}$, we consider function $\mathcal{B}_\epsilon(\theta_0, \theta)$ with

$$\theta = (x_1, x_2, 0, z) \in M,$$

where $|-\pi - z| \leqslant \bar{\delta}$. We will call this functions the "backward" barrier function. Roughly speaking, $\mathcal{B}_\epsilon(\theta_0, \theta)$ is the infimum of the actions among curves starting at $\theta$ and backward asymptotic to $\theta_0$.

**Proposition 4.3.** *With the above notations, for $\theta = (x_1, x_2, 0, z)$, we have the following statements:*

1. *The backward barrier function $B_\epsilon(\theta_0, \theta)$ is independent of the choice of $\theta_0$:*
$$\mathcal{B}_\epsilon(\theta_0, \theta) = \mathcal{B}_\epsilon(\theta_0', \theta), \quad \forall \theta_0' \in \mathbb{A}_{-\pi} \subset M.$$
2. $$\mathcal{B}_\epsilon(\theta_0, \theta) = \inf_{\gamma \in \mathcal{C}^-(\theta)} \liminf_{s \to -\infty} \int_s^0 \mathcal{L}_\epsilon^{h_1, h_2}(\gamma, \dot{\gamma}) \mathrm{dt},$$
*where $\mathcal{C}^-(\theta)$ denotes the set of absolutely continuous curves*
$$\gamma : (-\infty, 0] \to M, \quad t \mapsto (x_1(t), x_2(t), x_3(t), x_4(t)),$$
*such that $\gamma(0) = \theta$, $\lim_{s \to -\infty} x_3(t) = -\pi$ and $\lim_{s \to -\infty} x_4(t) = -\pi$.*
3. *There exists $\epsilon_2 > 0$ such that for any $0 < \epsilon \leqslant \epsilon_2$, we have*
$$\mathcal{B}_\epsilon(\theta_0, \theta) = \mathcal{A}_0 + \mathcal{A}(z) + \epsilon \mathcal{M}^-(h_1, h_2, x_1, x_2) \pm \epsilon \frac{\kappa}{31}.$$

*Proof.* We first show that the function $\mathcal{B}_\epsilon(\theta_0, \theta)$ is independent of the choice of $\theta_0$. Indeed, for any $\theta_0' \in \mathbb{A}_{-\pi}$, by Proposition 4.1 and Lemma 4.2, we have

$$\mathcal{B}_\epsilon(\theta_0', \theta) \leqslant \mathcal{B}_\epsilon(\theta_0', \theta_0) + \mathcal{B}_\epsilon(\theta_0, \theta) = \mathcal{B}_\epsilon(\theta_0, \theta).$$

In the same way, $\mathcal{B}_\epsilon(\theta_0, \theta) \leqslant \mathcal{B}_\epsilon(\theta_0', \theta)$. Therefore $\mathcal{B}_\epsilon(\theta_0, \theta)$ does not depend on the choice of $\theta_0$.

Now we begin the proof of statement (2). For $\gamma \in \mathcal{C}^-(\theta)$, let $t_n \to -\infty$ be a sequence such that

$$\lim_{n \to \infty} \int_{t_n}^0 \mathcal{L}_\epsilon^{h, \tilde{h}}(\gamma, \dot{\gamma}) \mathrm{dt} = \liminf_{s \to \infty} \int_s^0 \mathcal{L}_\epsilon^{h, \tilde{h}}(\gamma, \dot{\gamma}) \mathrm{dt}.$$

Passing to a subsequence, if necessary, we may assume that $\gamma(t_n)$ approaches $\bar{\theta} \in \{x_3 = x_4 = -\pi\}$ as $n \to \infty$. Let $d_n = ||\gamma_n(t_n) - \bar{\theta}||$, and define the curves $\xi_n : [t_n - d_n, 0] \to M$ by

$$\xi_n(t) = \begin{cases} \gamma(t), & t \in [t_n, 0] \\ \frac{\bar{\theta} - \gamma(t_n)}{d_n}(t_n - t) + \gamma(t_n), & t \in [t_n - d_n, t_n]. \end{cases}$$

Clearly, $\xi_n(0) = \theta$ and $\xi_n(t_n - d_n) = \bar{\theta}$. Therefore

$$\mathcal{B}_{\epsilon, -t_n + d_n}(\bar{\theta}, \theta) \leqslant \int_{t_n}^0 \mathcal{L}_\epsilon^{h_1, h_2}(\gamma, \dot{\gamma}) \mathrm{dt} + d_n K,$$

where $K = \sup_{\theta\in M,||v||\leqslant 1} \mathcal{L}_\epsilon^{h_1,h_2}(\theta, v)$. By taking the lower limit as $n \to \infty$, we obtain

$$\mathcal{B}_\epsilon(\theta_0,\theta) = \mathcal{B}_\epsilon(\bar\theta,\theta) \leqslant \liminf_{n\to\infty} \mathcal{B}_{\epsilon,-t_n+d_n}(\bar\theta,\theta) \leqslant \liminf_{s\to-\infty} \int_s^0 \mathcal{L}_\epsilon^{h_1,h_2}(\gamma,\dot\gamma)\mathrm{dt}.$$

It follows that

$$\mathcal{B}_\epsilon(\theta_0,\theta) \leqslant \inf_{\gamma\in\mathcal{C}^-(\theta)} \liminf_{s\to-\infty} \int_s^0 \mathcal{L}_\epsilon^{h_1,h_2}(\gamma,\dot\gamma)\mathrm{dt}.$$

For the opposite inequality, we have that there exists a sequence $T_n \to \infty$ and a family of curves $\gamma_n(t) : [-T_n, 0] \to M$ with $\gamma_n(0) = \theta$ and $\gamma_n(-T_n) = \theta_0 = (x_1^0, x_1^0, -\pi, -$ such that

$$\mathcal{B}_\epsilon(\theta_0,\theta) = \lim_{n\to\infty} \mathcal{B}_{\epsilon,T_n}(\theta_0,\theta) = \lim_{n\to\infty} \int_{-T_n}^0 \mathcal{L}_\epsilon^{h_1,h_2}(\gamma_n,\dot\gamma_n)\mathrm{dt}. \tag{4.1}$$

Let $\bar\gamma_n(t) : (-\infty, 0] \to M$ be such that

$$\bar\gamma_n(t) = \begin{cases} \gamma_n(t), & t\in[-T_n,0], \\ \big(q_{h_1}(x_1^0, t+T_n), q_{h_2}(x_2^0, t+T_n), -\pi, -\pi\big), & t\in(-\infty,-T_n). \end{cases}$$

Since $\mathcal{L}_\epsilon^{h,\tilde h}(\bar\gamma_n(t), \dot{\bar\gamma}_n(t)) = 0$ for $t \leqslant -T_n$, we have

$$\begin{aligned} \lim_{n\to\infty} \int_{-T_n}^0 \mathcal{L}_\epsilon^{h_1,h_2}(\gamma_n,\dot\gamma_n)\mathrm{dt} &= \lim_{n\to\infty} \liminf_{s\to-\infty} \int_s^0 \mathcal{L}_\epsilon^{h_1,h_2}(\bar\gamma_n,\dot{\bar\gamma}_n)\mathrm{dt} \\ &\geqslant \inf_{\gamma\in\mathcal{C}^-(\theta)} \liminf_{s\to-\infty} \int_s^0 \mathcal{L}_\epsilon^{h_1,h_2}(\gamma,\dot\gamma)\mathrm{dt}. \end{aligned}$$

This establishes statement (2).

For statement (3), let $\{T_n, n\in\mathbb{N}\}$ be as in (4.1). By the definition of $\mathcal{B}_{\epsilon,T}(\cdot,\cdot)$, we have that

$$\mathcal{B}_{\epsilon,T_n}(\theta_0,\theta) \leqslant \mathcal{B}_{\epsilon,\frac{T_n}{2}}(\theta_0,\theta_n) + \mathcal{B}_{\epsilon,\frac{T_n}{2}}(\theta_n,\theta),$$

and

$$\mathcal{B}_{\epsilon,2T_n}(\theta_n',\theta) \leqslant \mathcal{B}_{\epsilon,T_n}(\theta_n',\theta_0) + \mathcal{B}_{\epsilon,T_n}(\theta_0,\theta),$$

where

$$\theta_n = (q_{h_1}(x_1, -\frac{T_n}{2}), q_{h_2}(x_2, -\frac{T_n}{2}), -\pi, -\pi)$$

and

$$\theta_n' = (q_{h_1}(x_1, -2T_n), q_{h_2}(x_2, -2T_n), -\pi, -\pi).$$

By Lemmas 3.2 and 4.2, we have that for $\epsilon$ small enough,

$$B_\epsilon(\theta_0,\theta) = \mathcal{A}_0 + \mathcal{A}(z) + \epsilon\mathcal{M}^-(h_1,h_2,x_1,x_2) \pm \epsilon\frac{\kappa}{31}.$$

We thus prove all the assertions of the lemma. ∎

Similarly, for any $\theta_1 \in \mathbb{A}_\pi$ and

$$\theta' = (x_1', x_2', 0, z'),$$

with $|\pi - z| \leqslant \bar{\delta}$, we define the 'forward' barrier function $B_\epsilon(\theta', \theta_1)$, which is the infimum of the actions among curves staring at $\theta'$ and forward asymptotic to $\theta_1$. Mimicking the arguments in the proof of Proposition 4.3, we obtain the following statements.

**Proposition 4.4.** *With the above notations, we have the following:*

1. *The forward barrier function $B_\epsilon(\theta', \theta_1)$ is independent of the choice of $\theta_1$:*

$$\mathcal{B}_\epsilon(\theta', \theta_1) = \mathcal{B}_\epsilon(\theta', \theta_2), \quad \forall \theta_2 \in \mathbb{A}_\pi \subset M.$$

2. 

$$\mathcal{B}_\epsilon(\theta', \theta_1) = \inf_{\gamma \in \mathcal{C}^+(\theta')} \liminf_{s \to +\infty} \int_0^s \mathcal{L}_\epsilon^{h_1, h_2}(\gamma, \dot{\gamma}) \mathrm{d}t,$$

*where $\mathcal{C}^+(\theta')$ denotes the set of absolutely continuous curves*

$$\gamma : [0, +\infty) \to M, \quad t \mapsto (x_1(t), x_2(t), x_3(t), x_4(t)),$$

*such that $\gamma(0) = \theta$, $\lim_{s \to +\infty} x_3(t) = \pi$ and $\lim_{s \to \infty} x_4(t) = \pi$.*

3. *There exists $\epsilon_3 > 0$ such that for any $0 < \epsilon \leqslant \epsilon_3$, we have*

$$\mathcal{B}_\epsilon(\theta', \theta_1) = \mathcal{A}_0 + \mathcal{A}(z) + \epsilon \mathcal{M}^+(h_1, h_2, x_1', x_2') \pm \epsilon \frac{\kappa}{31}.$$

As in the Section 3, similar results of this section remain true for the modified Lagrangian $\mathcal{L}_{\epsilon,i}^{h_i, h_{i+1}}$, defined in (3.8).

From now on, for each $i = 1, \ldots, 4$ and $\theta \in \{x_{i+2} = 0\} \subset M$, we denote respectively

$$\mathcal{B}_\epsilon^+(i, h_i, h_{i+1}, \theta) \quad \text{and} \quad \mathcal{B}_\epsilon^-(i, h_i, h_{i+1}, \theta), \tag{4.2}$$

the "forward" and "backward" barrier functions for the Lagrangian $\mathcal{L}_{\epsilon,i}^{h_i, h_{i+1}}$, to indicate their dependence on $h_i$, $h_{i+1}$ and the positions of the pendulums. When the context is clear, we would simply write them as $\mathcal{B}_\epsilon^\pm(\theta)$.

## 5. The Maupertuis principle

In this section we describe the Maupertuis principle. Fix an energy level $H = 1$, the Maupertuis principle states that a geodesic $\gamma$ of the Riemannian metric

$$\mathrm{d}\rho_\epsilon = \sqrt{2} \sqrt{1 + \sum_{i=1}^4 (1 + \cos x_i) - \epsilon \sum_{i=1}^4 \beta_\alpha(x_i, x_{i+1}, x_{i+2})} \, \mathrm{d}s,$$

where $\mathrm{d}s$ denotes the usual Euclidean metric, is an orbit of the Euler-Lagrange system (1.1), up to a reparametrization. This special parametrization will be called the "time parametrization". If a geodesic $\gamma$ is parametrized by the time variable $t$, then we have that $\gamma(t) = (x_1, x_2, x_3, x_4)(t)$ solves the equation (1.1) and

$$\mathrm{d}\rho_\epsilon(\gamma(\mathrm{t})) = ||\dot{\gamma}||^2 \mathrm{dt},$$

with

$$\|\dot{\gamma}\| = \sqrt{2}\sqrt{1 + \sum_{i=1}^{4}(1+\cos x_i) - \epsilon\sum_{i=1}^{4}\beta_\alpha(x_i, x_{i+1}, x_{i+2})},$$

where $\dot{\gamma} = \frac{\mathrm{d}\gamma}{\mathrm{dt}}$ and $\|\cdot\|$ denote the standard Euclidean norm in $\mathbb{R}^4$.

In this section, we will unfold $x_1, x_2, x_3, x_4$ directions, i.e. $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. For any $\theta_0, \theta_1 \in \mathbb{R}^4$, we write

$$\rho_\epsilon(\theta_0, \theta_1) = \inf_{\gamma}\Big\{\int_\gamma \mathrm{d}\rho_\epsilon\Big\},$$

where the infimum is taken among all piecewise $C^1$ curves that connect $\theta_0$ and $\theta_1$. For small $\epsilon$, e.g. $\epsilon < \frac{1}{8}$, along a geodesic $\gamma$ parametrized by $t$, we have $1 \leqslant \|\dot{\gamma}(t)\| \leqslant 5$. Hence any geodesic can be extended indefinitely, implying the geodesic completeness of the Riemannian manifold. It follows that the infimum in the definition of $\rho_\epsilon(\theta_0, \theta_1)$ is realized by a minimizing geodesic. In the following lemma, we give the relation between the geodesics of the Riemannian metric $d\rho_\epsilon$ and the modified Lagrangian $\mathcal{L}_\epsilon^{h_1,h_2}$.

**Lemma 5.1.** *For any $\gamma : [0, T] \to \mathbb{R}^4$ that is a piecewise geodesic with the time parametrization, and any $h_1,\ h_2 \in [0, 1]$ such that $h_1 + h_2 = 1$, we have*

$$\int_\gamma \mathrm{d}\rho_\epsilon - v_{h_1}(x_1)\mathrm{dx}_1 - v_{h_2}(x_2)\mathrm{dx}_2 = \int_0^T \mathcal{L}_\epsilon^{h_1,h_2}(\gamma, \dot{\gamma})\mathrm{dt},$$

*where $\mathcal{L}_\epsilon^{h_1,h_2}$ is the modified Lagrangian defined in* (3.1) *and $v_h$ is defined in* (3.2)*.*

*Proof.* We have $\sum_{i=1}^4 \dot{x}_i^2 = H + \mathcal{L}_\epsilon = 1 + \mathcal{L}_\epsilon$. It follow that

$$\begin{aligned}
&\int_\gamma \mathrm{d}\rho_\epsilon - v_{h_1}(x_1)\mathrm{dx}_1 - v_{h_2}(x_2)\mathrm{dx}_2 \\
&= \int_0^T \Big[\|\dot{\gamma}\|^2 - v_{h_1}(x_1)\dot{x}_1 - v_{h_2}(x_2)\dot{x}_2\Big]\mathrm{dt} \\
&= \int_0^T \Big[\mathcal{L}_\epsilon(\gamma, \dot{\gamma}) + \Big(h_1 - v_{h_1}(x_1)\dot{x}_1\Big) + \Big(h_2 - v_{h_2}(x_3)\dot{x}_2\Big)\Big]\mathrm{dt} \\
&= \int_0^T \mathcal{L}_\epsilon^{h_1,h_2}(\gamma, \dot{\gamma})\mathrm{dt}.
\end{aligned}$$

Thus we prove the assertion of the lemma □

Before introducing the main results of this section, let us define the following metric on $\mathbb{R}^2$ by

$$\mathrm{d}p = \sqrt{2}\sqrt{1 + \sum_{i=1}^{2}(1+\cos x_i)}\ \mathrm{d}s,$$

and denote

$$\mathcal{P}(\tilde{\theta}_0, \tilde{\theta}_1) = \inf_{\tilde{\gamma}}\Big\{\int_{\tilde{\gamma}} \mathrm{d}p\Big\}, \qquad \tilde{\theta}_0,\ \tilde{\theta}_1 \in \mathbb{R}^2, \tag{5.1}$$

where the infimum is taken over all piecewise $C^1$ curves that connect $\tilde\theta_0$ and $\tilde\theta_1$. From the Maupertuis principle we know that the geodesics of the Riemannian metric $dp$ correspond to orbits of the system with two free pendulums and total energy 1. The following lemma is well known in literature, see e.g. Section 4 of [17].

**Lemma 5.2.** *(1) Let $\tilde\theta_i = (y_1^i, y_2^i) \in \mathbb{R}^2$, $i = 0, 1$. Assume there exist $k_0, k_1 \in \mathbb{Z}$ such that*

$$y_j^0 < (2k_j + 1)\pi < y_j^1 \quad j = 1, 2. \tag{5.2}$$

*There exists $K > 0$ such that if $D = \|\tilde\theta_0 - \tilde\theta_1\| \geqslant K$, then there exist unique $T$, $h$, $\tilde h > 0$ such that*

$$h + \tilde h = 1, \qquad q_h(y_1^0, T) = y_1^1, \quad \textit{and} \quad q_{\tilde h}(y_2^0, T) = y_2^1.$$

*Moreover, $\frac{D}{4} < T < D$ and*

$$\mathcal{P}(\tilde\theta_0, \tilde\theta_1) = \mathcal{S}^h(y_1^0, y_1^1) + \mathcal{S}^{\tilde h}(y_2^0, y_2^1),$$

*where $\mathcal{S}^h$ is defined in (2.2).*

*(2) There exists $C > 1$ such that if any two pairs of points $\tilde\theta_0^i$, $\tilde\theta_1^i$, $i = 1, 2$, in $\mathbb{R}^2$ satisfy condition (5.2) with*

$$\left|\tilde\theta_0^i - \tilde\theta_1^i\right| \geqslant C, \quad i = 1, 2,$$

*and*

$$\left|\mathbf{e}_2 - \mathbf{e}_1\right| \leqslant \frac{1}{C}, \quad \textit{where} \quad \mathbf{e}_i = \frac{\tilde\theta_1^i - \tilde\theta_0^i}{\left|\tilde\theta_1^i - \tilde\theta_0^i\right|}, \quad i = 1, 2,$$

*then*

$$\max\{|h_1 - h_2|,\ |\tilde h_1 - \tilde h_2|\} \leqslant \left|\mathbf{e}_2 - \mathbf{e}_1\right|,$$

*where $h_i$, $\tilde h_i$, $i = 1, 2$ are from the first part of the lemma.*

Now we can state the main technical lemmas. Pick any two points $\theta_0, \theta_2 \in \mathbb{R}^4$ on the sections $\{x_3 = 2i\pi\}$ and $\{x_3 = 2(i+1)\pi\}$ with $|x_4 - \mathbb{Z}2\pi| \leqslant \bar\delta$, we would like to estimate the geodesic distance $\rho_\epsilon(\theta_0, \theta_1)$, see Figure 3.

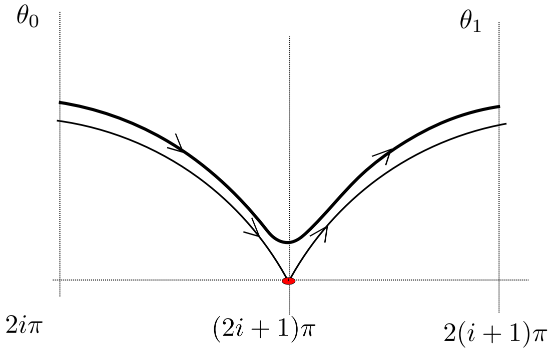


FIGURE 3. The geodesics between sections.

In the following lemma we will show that if $\|\theta_0-\theta_1\|$ is sufficiently large, the difference between geodesic distance $\rho_\epsilon$ and the pendulum distance $\mathcal{P}$ is approximate by the one-sided barrier functions $\mathcal{B}_\epsilon^\pm$, defined in (4.2).

**Lemma 5.3.** *There exists $\epsilon_4>0$ such that for any $0<\epsilon<\epsilon_4$, the following holds: Consider any $i\in\mathbb{Z}$, $\theta_j=(\tilde\theta_j, 2(i+j)\pi, z_j)\in\mathbb{R}^4$, $j=0,1$ with $\tilde\theta_j=(y_1^j,y_2^j)\in\mathbb{R}^2$ satisfying the condition (5.2), and*

$$|z_j-(2k+1)\pi|\leqslant\bar\delta,\quad j=0,1,$$

*for some $k\in\mathbb{Z}$. Then if $||\tilde\theta_0-\tilde\theta_1||\geqslant\frac{1}{\epsilon}$, then*

$$\rho_\epsilon(\theta_0,\theta_1)-\mathcal{P}(\tilde\theta_0,\tilde\theta_1)=2\mathcal{A}_0+\sum_{j=0}^{1}\mathcal{A}(z_j)+\epsilon\mathcal{M}^+(h_1,h_2,y_1^0,y_2^0)+\epsilon\mathcal{M}^-(h_1,h_2,y_1^1,y_2^1)\pm\epsilon\frac{\kappa}{14},$$

*where $h_1$, $h_2$ are determined by Lemma 5.2 for the pendulum distance $\mathcal{P}(\tilde\theta_0,\tilde\theta_1)$.*

*Proof.* It is sufficient to prove the statement for $i=k=0$. Without loss of generality, we assume $z_0<\pi<z_1$.

By Lemma 5.2, we have $T>\frac{1}{4\epsilon}$, $h_1>0$, $h_2>0$ such that

$$h_1+h_2=1,\ q_{h_1}(y_1^0,T)=y_1^1,\ q_{h_2}(y_2^0,T)=y_2^1,\ \text{and}\ \mathcal{P}(\tilde\theta_0,\tilde\theta_1)=\mathcal{S}^{h_1}(y_1^0,y_1^1)+\mathcal{S}^{h_2}(y_2^0,y_2^1).$$

We first prove the "$\leqslant$"-inequality. Let us define

$$x_3^+(t)=4\arctan e^t-\pi\quad\text{ and }\quad x_3^-(t)=4\arctan e^t+\pi.$$

Clearly $x_3^+(t)\to\pi$ as $t\to+\infty$ and $x_3^-(t)\to\pi$ as $t\to-\infty$. Let

$$x_1(t)=q_{h_1}(y_1^0,t),\quad x_2(t)=q_{h_2}(y_2^0,t),$$

and

$$x_4^+(t)=x_3^+(t+t_+),\quad x_4^-(t)=x_3^-(t+t_-),$$

where $x_3^+(t_+)=z_0$ and $x_3^-(t_-)=z_1$. Denote $\nu_m=|x_m^+(\frac{T}{2})-x_m^-(-\frac{T}{2})|$, $m=3,4$. It is easy to see that $\nu_4\leqslant\nu_3\leqslant 8e^{-\frac{T}{2}}$.

Let $\tilde T=T+\nu_3$. We will build a curve $\gamma_{\tilde T}$ as (see Figure 4)

$$\gamma_{\tilde T}(t)=\begin{cases}(x_1(t),x_2(t),x_3^+(t),x_4^+(t)), & t\in[0,T/2],\\ (x_1(\frac{T}{2}),x_2(\frac{T}{2}),\tilde x_3(t),\tilde x_4(t)), & t\in[\frac{T}{2},\frac{T}{2}+\nu_3],\\ (x_1(t-\nu_3),x_2(t-\nu_3),x_3^-(t-\tilde T),x_4^-(t-\tilde T),), & t\in(\frac{T}{2}+\nu_3,\tilde T]\end{cases}$$

where

$$\tilde x_m(t)=x_i^+(\frac{T}{2})+\min\{\frac{t-\frac{T}{2}}{\nu_m},1\}\cdot\Big(x_m^-(-\frac{T}{2})-x_m^+(\frac{T}{2})\Big),\quad m=3,4.$$

It follows that $\gamma_{\tilde T}:[0,\tilde T]\to M$ is a piecewise $C^1$ curve with $\gamma_{\tilde T}(0)=\theta_0$ and $\gamma_{\tilde T}(\tilde T)=\theta_1$. We have that

$$\rho_\epsilon(\theta_0,\theta_1)\leqslant\int_{\gamma_{\tilde T}}\mathrm{d}\rho_\epsilon\leqslant\int_{\gamma_{\tilde T}|[0,\frac{T}{2}]}\mathrm{d}\rho_\epsilon+\int_{\gamma_{\tilde T}|[\frac{T}{2}+\nu,\tilde T]}\mathrm{d}\rho_\epsilon+\nu_3:=\mathrm{I}+\mathrm{II}+\nu_3.$$

To simplify notations, we write $x_m^+(t)$ as $x_m(t)$, $m=3,4$, in the following arguments.

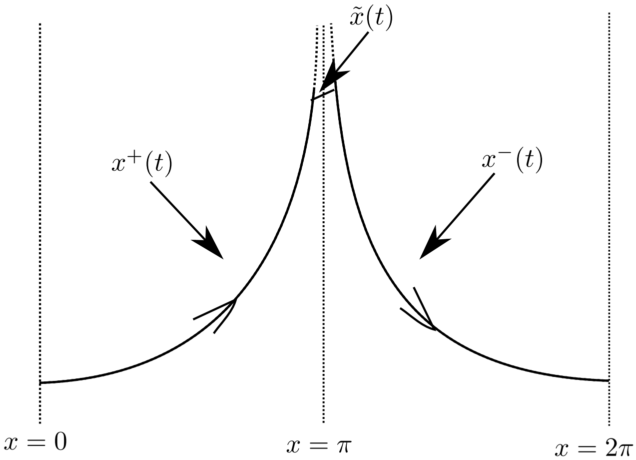


FIGURE 4. The construction of $\gamma_{\tilde{T}}$.

Since

$$\|\dot{\gamma}_{\tilde{T}}(t)\| = \sqrt{2}\sqrt{1+\sum_{i=1}^{4}[1+\cos x_i(t)]}, \quad t\in[0,\frac{T}{2}],$$

then

$$\mathrm{I} = 2\int_{\gamma_{\tilde{T}}\big|[0,\frac{T}{2}]}\left\{\sqrt{1+\sum_{i=1}^{4}[1+\cos x_i-\epsilon\beta_\alpha(x_i,x_{i+1},x_{i+2})]}\times\sqrt{1+\sum_{i=1}^{4}(1+\cos x_i)}\right\}\mathrm{dt}.$$

Using (2.6) and the inequality $\sqrt{a}\sqrt{a+b}\leqslant a+\frac{1}{2}b$, we have that

$$\begin{aligned}\mathrm{I} &\leqslant 2\int_0^{\frac{T}{2}}[1+\sum_{i=1}^{4}[1+\cos x_i(t)-\frac{1}{2}\epsilon\beta(x_i(t),x_{i+1}(t),x_{i+2}(t))]\mathrm{dt}\\ &\leqslant \mathcal{A}_0+\mathcal{A}(z_0)+\mathcal{S}^{h_1}(y_1^0,x_1(\frac{T}{2}))+\mathcal{S}^{h_2}(y_2^0,x_2(\frac{T}{2}))\\ &\quad+\epsilon\mathcal{M}^+(h_1,h_2,y_1^0,y_2^0)+\epsilon\frac{\kappa}{33}+O(e^{-\frac{T}{2}}).\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{II}\leqslant \mathcal{A}_0+\mathcal{A}(z_1)+\mathcal{S}^{h_1}(x_1(\frac{T}{2}),y_1^1)+\mathcal{S}^{h_2}(x_2(\frac{T}{2}),y_2^1)\\ +\epsilon\mathcal{M}^-(h_1,h_2,y_1^1,y_2^1)+\epsilon\frac{\kappa}{33}+O(e^{-\frac{T}{2}}).\end{aligned}$$

Since $T > \frac{1}{4\epsilon}$, for small enough $\epsilon$, we have $|O(e^{-\frac{T}{2}})| < \epsilon\frac{\kappa}{100}$. Thus we obtain that for $\epsilon$ small enough,

$$\rho_\epsilon(\theta_0,\theta_1)-\mathcal{P}(\tilde\theta_0,\tilde\theta_1) \leqslant 2\mathcal{A}_0+\sum_{j=0}^{1}\mathcal{A}(z_j)+\epsilon\mathcal{M}^+(h_1,h_2,y_1^0,y_2^0)+\epsilon\mathcal{M}^-(h_1,h_2,y_1^1,y_2^1)+\epsilon\frac{\kappa}{14}.$$

Now we continue to the proof of the "$\geqslant$"-part.

Let $\gamma = (x_1,x_2,x_3,x_4)(t) : [0,\bar T] \to \mathbb{R}^4$ be a minimizing geodesic for $\rho_\epsilon(\theta_0,\theta_1)$. Denote $\tau := \min\{\tau \in [0,\bar T] : x_3(\tau) = \pi\}$. See Figure 3.

We claim that there exists $C_0$ such that for small enough $\epsilon$, we have

$$|x_4(\tau)-\pi| = O(e^{-\frac{1}{C_0\epsilon}}). \tag{5.3}$$

Let us define

$$\xi^+(t) = \begin{cases}\gamma(t), & t\in[0,\tau],\\ (q_{h_1}(x_1(\tau),t-\tau), q_{h_2}(x_2(\tau),t-\tau), \pi, \hat x_4^+(t)), & t\in(\tau,+\infty),\end{cases}$$

and

$$\xi^-(t) = \begin{cases}\gamma(t+\bar T), & t\in[-(\bar T-\tau),0]\\ (q_{h_1}(x_1(\tau),t+\bar T-\tau), q_{h_2}(x_2(\tau),t+\bar T-\tau), \pi, \hat x_4^-(t)), & t\in(-\infty,-(\bar T-\tau)),\end{cases}$$

where

$$\hat x_4^+(t) = \begin{cases}x_4(\tau)+(t-\tau)(\pi-x_4(\tau)), & t\in[\tau,\tau+1]\\ \pi, & \text{else},\end{cases}$$

and

$$\hat x_4^-(t) = \begin{cases}x_4(\tau)+(t+\bar T-\tau)(x_4(\tau)-\pi), & t\in[-(\bar T-\tau)-1,-(\bar T-\tau)]\\ \pi, & \text{else}.\end{cases}$$

Clearly $\xi^+ \in \mathcal{C}^+(\theta_0)$ and $\xi^- \in \mathcal{C}^-(\theta_1)$. (See the defnitions in Propositions 4.3 and 4.4.) Moreover,

$$\mathcal{L}_\epsilon^{h_1,h_2}(\xi^+(t),\dot\xi^+(t)) = 0, \quad t \geqslant \tau+1 \quad \text{and} \quad \mathcal{L}_\epsilon^{h_1,h_2}(\xi^-(t),\dot\xi^-(t)) = 0, \quad t \leqslant -(\bar T-\tau)-1.$$

So by Proposition 4.3 and 4.4, we obtain

$$\begin{aligned}\mathcal{B}_\epsilon^+(h_1,h_2,\theta_0) &\leqslant \liminf_{t\to+\infty}\int_{\xi^+|[0,t]}\mathcal{L}_\epsilon^{h_1,h_2}(\xi^+(s),\dot\xi^+(s))\mathrm{d}s\\ &= \int_{\gamma|[0,\tau]} d\rho_\epsilon - v_{h_1}(x_1)\mathrm{dx}_1 - v_{h_2}(x_2)\mathrm{dx}_2 + O(e^{-\frac{1}{C_0\epsilon}}),\end{aligned}$$

and

$$\begin{aligned}\mathcal{B}_\epsilon^-(h_1,h_2,\theta_1) &\leqslant \liminf_{t\to-\infty}\int_{\xi^-|[t,0]}\mathcal{L}_\epsilon^{h_1,h_2}(\xi^-(t),\dot\xi^-(t))\mathrm{dt}\\ &= \int_{\gamma|[\tau,T]} d\rho_\epsilon - v_{h_1}(x_1)\mathrm{dx}_1 - v_{h_2}(x_2)\mathrm{dx}_2 + O(e^{-\frac{1}{C_0\epsilon}}),\end{aligned}$$

(here we have used (5.3)). So,

$$\begin{aligned}&\mathcal{B}_\epsilon^-(h_1,h_2,\theta_1)+\mathcal{B}_\epsilon^+(h_1,h_2,\theta_0)\\&\leqslant\int_{\gamma|[0,T]}d\rho_\epsilon-v_{h_1}(x_1)\mathrm{dx}_1-v_{h_2}(x_2)\mathrm{dx}_2+O(e^{-\frac{1}{C_0\epsilon}})\\&\leqslant\rho_\epsilon(\theta_0,\theta_1)-\mathcal{S}^{h_1}(y_1^0,y_1^1)-\mathcal{S}^{h_2}(y_2^0,y_2^1)+O(e^{-\frac{1}{C_0\epsilon}}).\end{aligned}$$

Therefore, for small enough $\epsilon$, we have

$$\rho_\epsilon(\theta_0,\theta_1)-\mathcal{P}(\tilde\theta_0,\tilde\theta_1)\geqslant 2\mathcal{A}_0+\sum_{j=0}^{1}\mathcal{A}(z_j)+\epsilon\mathcal{M}^+(h_1,h_2,y_1^0,y_2^0)+\epsilon\mathcal{M}^-(h_1,h_2,y_1^1,y_2^1)-\epsilon\frac{\kappa}{14}.$$

This implies the assertion of the lemma.

We still need to prove the claim (5.3). From the equation (1.1) we see that there exists $C>0$ such that

$$-(1+C\epsilon)\sin x_3\leqslant\ddot x_i\leqslant-(1-C\epsilon)\sin x_3\quad x_3\in[\frac{3\pi}{4},\pi],$$

and

$$-(1-C\epsilon)\sin x_3\leqslant\ddot x_3\leqslant-(1+C\epsilon)\sin x_3,\quad x_i\in[\pi,\frac{5\pi}{4}].$$

Let $v=\dot x_3(\tau)$. Clearly $v>0$ and $v\to 0$ as $\epsilon\to 0$. We have

$$\int_\pi^{\frac{5\pi}{4}}\frac{1}{\sqrt{v^2+2(1+C\epsilon)(1+\cos x)}}\mathrm{dx}\leqslant\bar T-\tau\leqslant 2\int_\pi^{\frac{5\pi}{4}}\frac{1}{\sqrt{v^2+2(1-C\epsilon)(1+\cos x)}}\mathrm{dx},$$

and

$$\int_{\frac{3\pi}{4}}^{\pi}\frac{1}{\sqrt{v^2+2(1+C\epsilon)(1+\cos x)}}\mathrm{dx}\leqslant\tau\leqslant 2\int_{\frac{3\pi}{4}}^{\pi}\frac{1}{\sqrt{v^2+2(1-C\epsilon)(1+\cos x)}}\mathrm{dx}.$$

So for $\epsilon$ small enough,

$$0<C_1\tau\leqslant\bar T-\tau\leqslant C_1'\tau.$$

This leads to

$$\bar T-\tau\geqslant\frac{1}{C_2\epsilon}\quad\text{ and }\quad\tau\geqslant\frac{1}{C_2\epsilon}.$$

Without loss of generality we assume $x_4(\tau)>\pi$. Therefore,

$$\bar T-\tau<\int_{x_4(\tau)}^{\frac{5\pi}{4}}\frac{1}{\sqrt{2(1-C\epsilon)(\cos x-\cos(x_4(\tau))}}\mathrm{dx}=O(\big|\log\big(x_4(\tau)-\pi\big)\big|).$$

Hence $|x_4(\tau)-\pi|=O(e^{-\frac{1}{C_0\epsilon}})$. This finishes the proof of the lemma. □

In the lemma above we deal with the "substages of energy transferring": the pendulum **3** is moving closely to a heteroclinic orbit of the system while the pendulum **4** remain close to the "top" ($x_4=\pi\in\mathbb{T}$). In the following lemma, we study the "stages of advancing": the pendulum **2** moves with speed $O(1)$, while the pendulum **3** move along a heteroclinic orbit of the system with the pendulum **1** climbing to the "top" and the pendulum **4** dropping from the "top" simultaneously.

**Lemma 5.4.** *There exists $\epsilon_5 > 0$ such that for $\epsilon \in (0, \epsilon_5]$, the following holds: Consider any $i, j, k \in \mathbb{Z}$, $\theta_0, \theta_1 \in \mathbb{R}^4$ with $\theta_0 = (x_1^0, x_2^0, 2i\pi, x_4^0)$ and $\theta_1 = (x_1^1, x_2^1, x_3^1, 2j\pi)$ satisfying*

$$\begin{cases} |x_1^0 - 2k\pi| \leqslant \delta, \\ |x_1^1 - (2k+1)\pi| \leqslant \bar{\delta}, \\ |x_3^1 - 2(i+1)\pi| \leqslant \delta, \\ |x_4^0 - (2j-1)\pi| \leqslant \bar{\delta}. \end{cases}$$

*Then if $D = x_2^1 - x_2^0 \geqslant \frac{1}{\epsilon}$, then*

$$\begin{aligned} &\rho_\epsilon(\theta_0, \theta_1) - \mathcal{S}^0(x_1^0, x_3^1) - \mathcal{S}^1(x_2^0, x_2^1) \\ &= 2\mathcal{A}_0 + \mathcal{A}(x_1^1) + \mathcal{A}(x_4^0) + \epsilon\mathcal{M}^+(0, 1, x_1^0, x_2^0) + \epsilon\mathcal{M}^-(0, 1, x_3^1, x_2^1) \pm \epsilon\frac{\kappa}{14}, \end{aligned}$$

*where $\mathcal{S}^1$ and $\mathcal{S}^0$ is defined in (2.2) and (2.3), respectively.*

*Proof.* It is sufficient to prove the statement for $i = j = k = 0$. Without loss of generality, let us assume $x_1^1 > \pi$ and $x_4^0 < -\pi$.

Let $x_2(t) = q_1(x_2^0, t)$. Then there exists unique $T > 0$ such that $x_2(T) = x_2^1$. Moreover $T > \frac{D}{4}$. Denote

$$x_1^+(t) = 4\arctan e^{t+t_1^+} - \pi, \quad x_1^-(t) = 4\arctan e^{t+t_1^-} + \pi,$$

$$x_3^+(t) = 4\arctan e^{t} - \pi, \quad x_3^-(t) = 4\arctan e^{t_3^-+t} + \pi,$$

and

$$x_4^+(t) = 4\arctan e^{t+t_4^+} - 3\pi, \quad x_4^-(t) = 4\arctan e^{t} - \pi.$$

where $t_1^\pm$, $t_3^-$ and $t_4^+$ are chosen such that

$$x_1^+(0) = x_1^0,\ x_1^-(0) = x_1^1,\ x_3^-(0) = x_3^1 \ \text{and}\ x_4^+(0) = x_4^0.$$

Let $\nu_i = \left|x_1^-(-\frac{T}{2}) - x_1^+(\frac{T}{2})\right|$, $i = 1, 3, 4$. Denote $\nu = \max\{\nu_1, \nu_3, \nu_4\}$. Set $\tilde{T} = T + \nu$. We construct a curve $\gamma_{\tilde{T}}(t) : [0, \tilde{T}] \to \mathbb{R}^4$ by

$$\gamma_{\tilde{T}}(t) = \begin{cases} (x_1^+(t), x_2(t), x_3^+(t), x_4^+(t)), & t \in [0, \frac{T}{2}], \\ (\hat{x}_1(t), x_2(\frac{T}{2}), \hat{x}_3(t), \hat{x}_4(t), & t \in (\frac{T}{2}, \frac{T}{2} + \nu], \\ (x_1^-(t - \tilde{T}), x_2(t - \nu), x_3^-(t - \tilde{T}), x_4^-(t - \tilde{T})), & t \in (\frac{T}{2} + \nu, \tilde{T}], \end{cases}$$

where

$$\hat{x}_i(t) = x_i^+(\frac{T}{2}) + \min\{\frac{t - T/2}{\nu_i}, 1\}\big(x_i^-(-\frac{T}{2}) - x_i^+(\frac{T}{2})\big), \quad i = 1, 3, 4.$$

Clearly the curve $\gamma_{\tilde{T}}$ is piecewise $C^1$ with $\gamma_{\tilde{T}}(0) = \theta_0$ and $\gamma_{\tilde{T}}(\tilde{T}) = \theta_1$. Therefore

$$\rho_\epsilon(\theta_0, \theta_1) \leqslant \int_{\gamma_{\tilde{T}}|[0,\tilde{T}]} \mathrm{d}\rho_\epsilon.$$

By similar argument in the proof of the Lemma 5.3 we obtain that for $\epsilon$ small enough,

$$\begin{aligned}&\rho_\epsilon(\theta_0,\theta_1)-\mathcal{S}^0(x_1^0,x_3^1)-\mathcal{S}^1(x_2^0,x_2^1)\\&\leqslant 2\mathcal{A}_0+\mathcal{A}(x_1^1)+\mathcal{A}(x_4^0)+\epsilon\mathcal{M}^+(0,1,x_1^0,x_2^0)+\epsilon\mathcal{M}^-(0,1,x_3^1,x_2^1)+\epsilon\frac{\kappa}{14},\end{aligned}$$

Let $\gamma(t)=(x_1(t),x_2(t),x_3(t),x_4(t)):[0,\bar{T}]\to\mathbb{R}^4$ be the minimizing geodesic of $\rho_\epsilon(\theta_0,\theta_1)$. Denote

$$\tau=\inf\{t\in[0,\bar{T}]:\ x_3(t)=\pi\}.$$

Argument as in Lemma 5.3 we have that there exists $C_0>1$ such that for small enough $\epsilon$,

$$|x_1(\tau)-\pi|=O(e^{-\frac{1}{C_0\epsilon}})\quad\text{and}\quad|-\pi-x_4(\tau)|=O(e^{-\frac{1}{C_0\epsilon}}).$$

We define

$$\xi^+(t)=\begin{cases}\gamma(t), & t\in[0,\tau],\\(\tilde{x}_1^+(t),q_1(x_2(\tau),t-\tau),\pi,\tilde{x}_4^+(t)), & t\in(\tau,+\infty),\end{cases}$$

and

$$\xi^-(t)=\begin{cases}\gamma(t+\bar{T}), & t\in[\tau,0],\\(\tilde{x}_1^-(t),q_1(x_2(\tau),t-T+\tau),\pi,\tilde{x}_4^-(t)), & t\in(-\infty,\tau)),\end{cases}$$

where for $i=1,4$,

$$\tilde{x}_i^+(t)=\begin{cases}x_i(\tau)+(t-\tau)(\pi-x_i(\tau)), & t\in[\tau,\tau+1]\\\pi, & \text{else},\end{cases}$$

and

$$\tilde{x}_i^-(t)=\begin{cases}x_i(\tau)+(t+\bar{T}-\tau)(x_i(\tau)-\pi), & t\in[-(\bar{T}-\tau)-1,-(\bar{T}-\tau)]\\\pi, & \text{else}.\end{cases}$$

Clearly $\xi^+(t)\in\mathcal{C}^+(\theta_0)$, defined for the modified Lagrangian $\mathcal{L}_{\epsilon,1}^{0,1}$ (see (3.8)), and $\xi^-(t)\in\mathcal{C}^-(\theta_1)$, for the modified Lagrangian $\mathcal{L}_{\epsilon,2}^{1,0}$. By the construction of $\xi^+$ we have

$$\mathcal{L}_{\epsilon,1}^{0,1}(\xi^+(t),\dot{\xi}^+(t))=0,\ t\geqslant\tau+1\quad\text{and}\quad\mathcal{L}_{\epsilon,2}^{1,0}(\xi^-(t),\dot{\xi}^-(t))=0,\ t\leqslant\tau-\bar{T}-1.$$

Then by similar arguments as in Lemma 5.3, we get that for sufficiently small $\epsilon$,

$$\begin{aligned}&\rho_\epsilon(\theta_0,\theta_1)-\mathcal{S}^0(x_1^0,x_3^1)-\mathcal{S}^1(x_2^0,x_2^1)\\&\geqslant 2\mathcal{A}_0+\mathcal{A}(x_1^1)+\mathcal{A}(x_4^0)+\epsilon\mathcal{M}^+(0,1,x_1^0,x_2^0)+\epsilon\mathcal{M}^-(0,1,x_3^1,x_2^1)-\epsilon\frac{\kappa}{14},\end{aligned}$$

This completes the proof of of the lemma. □

## 6. Construction of diffusion orbits

To fix ideas, in this section we will construct an orbit of the system (1.1) corresponding to the periodic monotone sequence, i.e.

$$\dots\sigma_{-1}\sigma_0\sigma_1\dots = \dots 123412341234\dots.$$

To construct orbits for general sequences will not pose any additional difficulty.

Fix any $r > 0$. Choose $K$ to be the smallest integer greater than $4\epsilon^{-2-r/2}$. Starting from the point $(0,0,0,\pi)$, we consider a sequence of sections in $\mathbb{R}^4$:

$$\underbrace{(\Sigma^1_{12,3},\Sigma^2_{12,3},\dots,\Sigma^K_{12,3})}_{1\to 2}\ \underbrace{(\Sigma^1_{23,4},\dots,\Sigma^K_{23,4})}_{2\to 3}, \tag{6.1}$$

where the sections are defined according to the rules explained below:

(I) The sections in the bracket "$1\to 2$" correspond to the stage of energy transferring that the energy slowly transfers from the pendulum $x_1$ to the pendulum $x_2$, while the pendulum $x_3$ moves along a heteroclinic orbit with the pendulum $x_4$ hanging near the top $x_4 = \pi$. The sections $\Sigma^i_{12,3}$, $i = 1,\dots,K$ are given by
   - i) $\Sigma^1_{12,3} := (0,0,0,\pi)$.
   - ii) $C^1_{r_1} = (0,0,0,\pi)$.
   - iii) $O_{12,3} = \{(x_1,x_2,x_3,x_4):\ |x_1|\leqslant\delta, |x_2|\leqslant\delta, x_3 = 0, |x_4|\leqslant\bar\delta\}\subset\mathbb{R}^4$.
   - iv) $C^1_{r_{i+1}} = C_{r_i} + 2\pi(K-i,i,1,0)$.
   - v) $\Sigma^{i+1}_{12,3} = C^1_{r_{i+1}} + O_{12,3}$, $i = 1,\dots,K-1$.

(II) The sections in the bracket "$2\to 3$" represent the stage of energy transferring that the energy transfers, bit by bit, from the pendulum $x_2$ to the pendulum $x_3$, while the pendulum $x_4$ moves closely to a heteroclinic orbit with the pendulum $x_1$ hanging near the top. The sections $\Sigma^i_{23,1}$ are given as
   - i) $C^2_{r_{i+1}} = C^2_{r_i} + (0,K-i,i,1)$, $i = 1,\dots,K-1$.
   - ii) $O_{23,4} = \{(x_1,x_2,x_3,x_4): |x_1|\leqslant\bar\delta, |x_2|\leqslant\delta, |x_3|\leqslant\delta, x_4 = 0\}$.
   - iii) $\Sigma^i_{23,4} = C^2_{r_i} + O_{23,4}$, $i = 1,\dots,K-1$.
   - iv) $\Sigma^K_{23,4} = C^2_{r_K}$.

(III) The sections $\Sigma^K_{12,3}$ and $\Sigma^1_{23,4}$ are connected by a stage of advancing:

$$C^2_{r_1} = C^1_{r_K} + 2\pi(\frac{1}{2},K,1,\frac{1}{2}).$$

Clearly, for any $i\in\{2,\dots,K-1\}$, $\theta_j\in\Sigma^j_{12,3}$, $j = i-1,i,i+1$, we have

$$|\mathbf{e}_j - \mathbf{e}_{j+1}|\leqslant\epsilon^{2+r/2},\quad \mathbf{e}_j = \frac{\theta_{j+1}-\theta_j}{|\theta_{j+1}-\theta_j|},\quad j = i-1,i,$$

and

$$|\mathbf{e}_{i-1} - \mathbf{e}'_{i-1}|\leqslant\epsilon^{2+r/2},\quad \mathbf{e}_{i-1} = \frac{\theta_{i-1}-\theta'_i}{|\theta_{i-1}-\theta'_i|},\quad \forall\theta'_i\in\Sigma^i_{12,3}.$$

The same estimates hold for the sections $\Sigma^i_{23,4}$, $i = 1, \dots, K-1$.

**Lemma 6.1.** *There exists $\epsilon_6 > 0$ such that for every $0 < \epsilon \leqslant \epsilon_6$, the following holds true:*

*(i) For each $i = 2, \dots, K-1$, any two points $\theta^j = (\tilde\theta^j, 2i\pi, x_4) \in \Sigma^i_{12,3}$, $j = 0, 1$, if we fix $\theta_k = (\tilde\theta_k, 2(i+k)\pi, x_4) \in \Sigma^{i+k}_{12,3}$, $k = \pm 1$, then*

$$\Big|\mathcal{P}(\tilde\theta_{-1}, \tilde\theta^0) + \mathcal{P}(\tilde\theta^0, \tilde\theta_{+1}) - \mathcal{P}(\tilde\theta_{-1}, \tilde\theta^1) - \mathcal{P}(\tilde\theta^1, \tilde\theta_{+1})\Big| \leqslant 4\epsilon^{1+r/4},$$

*where the pendulum distance $\mathcal{P}$ is defined in (5.1). The same relation holds true for points in the sections $\Sigma^i_{23,4}$, $i = 2, \dots, K-1$.*

*(ii) For any two points $\theta_j = (x_1^j, x_2^j, x_3, x_4) \in \Sigma^K_{12,3}$, $j = 0, 1$, if we fix $(\tilde\theta_1, \tilde x_3, \tilde x_4) \in \Sigma^{K-1}_{12,3}$ and $(x_1', x_2', x_3', x_4') \in \Sigma^1_{23,4}$, then*

$$\Big|\mathcal{P}(\tilde\theta_1, \tilde\theta^0) + \mathcal{S}^0(x_1^0, x_3') + \mathcal{S}^1(x_2^0, x_2') - \mathcal{P}(\tilde\theta_1, \tilde\theta^1) - \mathcal{S}^0(x_1^1, x_3') - \mathcal{S}^1(x_2^1, x_2')\Big| \leqslant 4\epsilon^{1+r/4},$$

*where $\tilde\theta^j = (x_1^j, x_2^j)$, $j = 0, 1$ and the metrics $\mathcal{S}^1$, $\mathcal{S}^0$ are defined in (2.2) and (2.3). The same result holds for points in $\Sigma^1_{23,4}$.*

For a proof of this lemma, see e.g. Section 2 of [17], or Lemma 2.1 of [4].

Now we set up our variational problem. Assume $\epsilon$ is small enough, i.e.

$$\epsilon < \min\Big\{\epsilon_0, \dots, \epsilon_6, (\frac{\kappa}{28})^{\frac{4}{r}}, (\frac{\lambda_0}{4})^{\frac{4}{r}}\Big\}.$$

Let

$$\mathbb{Y} = \{\Theta = (\theta_1, \dots, \theta_K, \theta^1, \dots, \theta^K) : \theta_i \in \Sigma^i_{12,3}, \theta^i \in \Sigma^i_{23,4}\}.$$

We define a function $\mathbb{F} : \mathbb{Y} \to \mathbb{R}$ by

$$\mathbb{F}(\theta_1, \dots, \theta_K, \theta^1, \dots, \theta^K) = \sum_{i=1}^{K-1} \rho_\epsilon(\theta_i, \theta_{i+1}) + \rho_\epsilon(\theta_K, \theta^1) + \sum_{i=1}^{K-1} \rho_\epsilon(\theta^i, \theta^{i+1}).$$

Since $ds \leqslant d\rho_\epsilon \leqslant 5ds$ as Riemannian metrics, the metric space defined by $d\rho_\epsilon$ is equivalent to the Euclidean metric space. It follows that the geodesic distance $\rho_\epsilon(\theta_0, \theta_1)$ is continuous in both $\theta_0$ and $\theta_1$. Hence $\mathbb{F}$ is also continuous. As $\mathbb{Y}$ is a compact manifold with boundary, the function $\mathbb{F}$ must attain its minimum in $\mathbb{Y}$. Let us pick one of the minimal points and denote it by

$$\bar\Theta = (\theta_1, \bar\theta_2, \dots, \bar\theta_K, \bar\theta^1, \dots, \bar\theta^{K-1}, \theta^K).$$

The following lemma is the key of our proof.

**Lemma 6.2.** *The point $\bar\Theta$ is in the interior of $\mathbb{Y}$.*

*Proof.* Assume the contrary that $\bar\Theta \in \partial\mathbb{Y}$. Then there exists $i$ such that $\bar\theta_i \in \partial\Sigma^i_{12,3}$ or $\bar\theta^i \in \partial\Sigma^i_{23,1}$.

**Case I**: *Energy transferring.* For any $i < K$. We write $\bar{\theta}_j = (x_1^j, x_2^j, x_3^j, x_4^j) \in \Sigma_{12,3}^j$, $j = i-1, i, i+1$. Then by Lemma 5.3, we have that

$$\begin{aligned}
&\rho_\epsilon(\bar{\theta}_{i-1}, \bar{\theta}_i) + \rho_\epsilon(\bar{\theta}_i, \bar{\theta}_{i+1}) - \mathcal{P}(\tilde{\theta}_{i-1}, \tilde{\theta}_i) - \mathcal{P}(\tilde{\theta}_i, \tilde{\theta}_{i+1}) \\
&= 4\mathcal{A}_0 + \mathcal{A}(x_4^{i-1}) + 2\mathcal{A}(x_4^i) + \mathcal{A}(x_4^{i+1}) + \epsilon\mathcal{M}^+(h_1^{i-1}, h_2^{i-1}, x_1^{i-1}, x_2^{i-1}) \\
&\quad + \epsilon\mathcal{M}^-(h_1^{i-1}, h_2^{i-1}, x_1^i, x_2^i) + \epsilon\mathcal{M}^+(h_1^i, h_2^i, x_1^i, x_2^i) \\
&\quad + \epsilon\mathcal{M}^-(h_1^i, h_2^i, x_1^{i+1}, x_2^{i+1}) \pm \epsilon\frac{\kappa}{7},
\end{aligned}$$

where $\tilde{\theta}_j = (x_1^j, x_2^j) \in \mathbb{R}^2$, $j = i-1, i, i_1$ and $h_1^j$, $h_2^j$, $j = i-1, i$ are determined by Lemma 5.2 for the pendulum distances $\mathcal{P}(\tilde{\theta}_j, \tilde{\theta}_{j+1})$, $j = i-1, i$, moreover,

$$|h_1^i - h_1^{i-1}| + |h_2^i - h_2^{i-1}| \leqslant \epsilon^{2+r/2}.$$

If $\bar{\theta}_i \in \partial\Sigma_{12,3}^i$, then there are two kinds of possibilities.

**Possibility 1**: $|x_1^i - 2k_1\pi| = \delta$ or $|x_2^i - 2k_2\pi| = \delta$. In this situation, let

$$\bar{\theta}_i' = (2k_1\pi, 2k_2\pi, x_3^i, x_4^i) \in \Sigma_{12,3}^i.$$

We will show that by moving $\bar{\theta}_i$ to the the point $\bar{\theta}_i'$ in the section $\Sigma_{12,3}^i$, the geodesic distance decreases,

$$\rho_\epsilon(\bar{\theta}_{i-1}, \bar{\theta}_i) + \rho_\epsilon(\bar{\theta}_i, \bar{\theta}_{i+1}) > \rho_\epsilon(\bar{\theta}_{i-1}, \bar{\theta}_i') + \rho_\epsilon(\bar{\theta}_i', \bar{\theta}_{i+1}),$$

which contradicts with the minimality of $\bar{\Theta}$. See Figure 5.

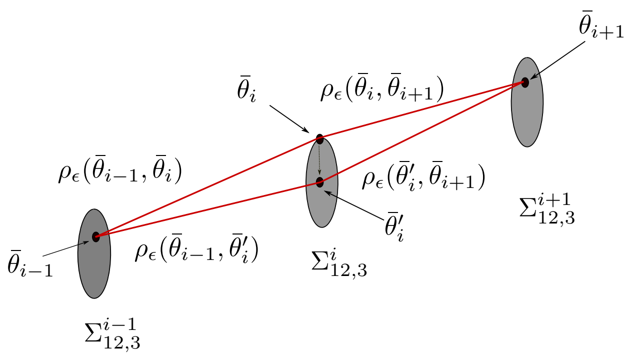


FIGURE 5. Moving the point from the boundary to the center.

Let $\tilde{\theta}_i' = (2k_1\pi, 2k_2\pi)$. Then, by Lemmas 5.3,

$$\begin{aligned}
&\rho_\epsilon(\bar{\theta}_{i-1}, \bar{\theta}_i') + \rho_\epsilon(\bar{\theta}_i', \bar{\theta}_{i+1}) - \mathcal{P}(\tilde{\theta}_{i-1}, \tilde{\theta}_i') - \mathcal{P}(\tilde{\theta}_i', \tilde{\theta}_{i+1}) \\
&= 4\mathcal{A}_0 + \mathcal{A}(x_4^{i-1}) + 2\mathcal{A}(x_4^i) + \mathcal{A}(x_4^{i+1}) + \epsilon\mathcal{M}^+(h_1^{i-1}, h_2^{i-1}, x_1^{i-1}, x_2^{i-1}) \\
&\quad + \epsilon\mathcal{M}^-(\tilde{h}_1^{i-1}, \tilde{h}_2^{i-1}, 0, 0) + \epsilon\mathcal{M}^+(\tilde{h}_1^i, \tilde{h}_2^i, 0, 0) + \epsilon\mathcal{M}^-(\tilde{h}_1^i, \tilde{h}_2^i, x_1^{i+1}, x_2^{i+1}) \pm \epsilon\frac{\kappa}{7},
\end{aligned}$$

where $\tilde{h}_1^j, \tilde{h}_2^j$, $j = i-1, i$ are determined by Lemma 5.2 for the pendulum distances $\mathcal{P}(\tilde{\theta}_{i-1}, \tilde{\theta}_i')$ and $\mathcal{P}(\tilde{\theta}_i', \tilde{\theta}_{i+1})$.

By Lemmas 6.1 and 5.2, we have

$$|\mathcal{P}(\tilde{\theta}_{i-1},\tilde{\theta}_i)+\mathcal{P}(\tilde{\theta}_i,\tilde{\theta}_{i+1})-\mathcal{P}(\tilde{\theta}_{i-1},\tilde{\theta}_i')-\mathcal{P}(\tilde{\theta}_i',\tilde{\theta}_{i+1})|\leqslant 4\epsilon^{1+r/4},$$

and

$$|h_1^{i-1}-\tilde{h}_1^{i-1}|+|h_2^i-\tilde{h}_2^i|\leqslant 2\epsilon^{2+r/2}.$$

So

$$\begin{aligned}&\rho_\epsilon(\bar{\theta}_{i-1},\bar{\theta}_i)+\rho_\epsilon(\bar{\theta}_i,\bar{\theta}_{i+1})-\rho_\epsilon(\bar{\theta}_{i-1},\bar{\theta}_i')-\rho_\epsilon(\bar{\theta}_i',\bar{\theta}_{i+1})\\&\geqslant \epsilon\mathcal{M}^+(h_1^{i-1},h_2^{i-1},x_1^{i-1},x_2^{i-1})+\epsilon\mathcal{M}^-(h_1^i,h_2^i,x_1^i,x_2^i)\\&\quad+\epsilon\mathcal{M}^+(h_1^i,h_2^i,x_1^i,x_2^i)+\epsilon\mathcal{M}^-(h_1^i,h_2^i,x_1^{i+1},x_2^{i+1})\\&\quad-\epsilon\mathcal{M}^+(\tilde{h}_1^{i-1},\tilde{h}_2^{i-1},x_1^{i-1},x_2^{i-1})+\epsilon\mathcal{M}^-(\tilde{h}_1^i,\tilde{h}_2^i,0,0)\\&\quad-\epsilon\mathcal{M}^+(\tilde{h}_1^i,\tilde{h}_2^i,0,0)+\epsilon\mathcal{M}^-(\tilde{h}_1^i,\tilde{h}_2^i,x_1^{i+1},x_2^{i+1})-4\epsilon^{1+r/4}\\&\geqslant \epsilon\mathcal{M}(h_1^i,h_2^i,x_1^i,x_2^i)-\epsilon\mathcal{M}(h_1^i,h_2^i,0,0)-\epsilon\frac{5\kappa}{7}\geqslant\epsilon\frac{2\kappa}{7}>0,\end{aligned}$$

where in the second inequality we have use the continuity property (2.5) of the one sided Melnikoff functions $\mathcal{M}^\pm$, while the third inequality we use Lemma 2.1.

**Possibility 2**: $|x_4-\pi|=\bar{\delta}$. Let us consider

$$\bar{\theta}_i''=(x_1^i,x_2^i,x_3^i,\pi).$$

Since $\mathcal{A}(\pi)=0$, then by Lemma 5.3, we have

$$\begin{aligned}&\rho_\epsilon(\bar{\theta}_{i-1},\bar{\theta}_i'')+\rho_\epsilon(\bar{\theta}_i'',\bar{\theta}_{i+1})-\mathcal{P}(\tilde{\theta}_{i-1},\tilde{\theta}_i)-\mathcal{P}(\tilde{\theta}_i,\tilde{\theta}_{i+1})\\&=4\mathcal{A}_0+\mathcal{A}(x_4^{i-1})+\mathcal{A}(x_4^{i+1})+\epsilon\mathcal{M}^+(h_1^{i-1},h_2^{i-1},x_1^{i-1},x_2^{i-1})\\&\quad+\epsilon\mathcal{M}^-(h_1^{i-1},h_2^{i-1},x_1^i,x_2^i)+\epsilon\mathcal{M}^+(h_1^i,h_2^i,x_1^i,x_2^i)\\&\quad+\epsilon\mathcal{M}^-(h_1^i,h_2^i,x_1^{i+1},x_2^{i+1})\pm\epsilon\frac{\kappa}{7}.\end{aligned}$$

Therefore

$$\rho_\epsilon(\bar{\theta}_{i-1},\bar{\theta}_i)+\rho_\epsilon(\bar{\theta}_i,\bar{\theta}_{i+1})-\rho_\epsilon(\bar{\theta}_{i-1},\bar{\theta}_i'')-\rho_\epsilon(\bar{\theta}_i'',\bar{\theta}_{i+1})=2\mathcal{A}(x_4)\pm\epsilon\frac{2\kappa}{7}>\frac{\kappa}{400}>0.$$

This excludes the possibility that $|x_4-\pi|=\bar{\delta}$. Hence $\bar{\theta}_i\notin\partial\Sigma_{12,3}^i$.

Similarly, we show that $\bar{\theta}^i\notin\partial\Sigma_{23,4}^i$, $i=2,\dots,K-1$.

**Case II**: *Advancing*. Let us write

$$\bar{\theta}_K=(x_1^K,x_2^K,x_3^K,x_4^K)\in\Sigma_{12,3}^K,\ \text{ and }\ \bar{\theta}^j=(\tilde{x}_1^j,\tilde{x}_2^j,\tilde{x}_3^j,\tilde{x}_4^j)\in\Sigma_{23,4}^j,\quad j=1,2.$$

Then by Lemmas 5.3 and 5.4 we have that

$$\begin{aligned}&\rho_\epsilon(\bar{\theta}_K,\bar{\theta}^1)+\rho_\epsilon(\bar{\theta}^1,\bar{\theta}^2)-\mathcal{S}^0(x_1^K,\tilde{x}_3^1)-\mathcal{S}^1(x_2^K,\tilde{x}_2^1)-\mathcal{P}(\tilde{\theta}^1,\tilde{\theta}^2)\\&=4\mathcal{A}_0+\mathcal{A}(x_4^K)+2\mathcal{A}(\tilde{x}_1^1)+\mathcal{A}(\tilde{x}_1^2)+\epsilon\mathcal{M}^+(0,1,x_1^K,x_2^K)\\&\quad+\epsilon\mathcal{M}^-(0,1,\tilde{x}_3^1,\tilde{x}_2^1)+\epsilon\mathcal{M}^-(\bar{h}_1,\bar{h}_2,\tilde{x}_3^1,\tilde{x}_2^1)\\&\quad+\epsilon\mathcal{M}^+(\bar{h}_1,\bar{h}_2,\tilde{x}_3^2,\tilde{x}_2^2)\pm\epsilon\frac{\kappa}{7},\end{aligned}$$

where $\tilde{\theta}^j = (\tilde{x}_3^j, \tilde{x}_2^j) \in \mathbb{R}^2$, $j = 1, 2$, and $\bar{h}_1, \bar{h}_2$ are determined by Lemma 5.2 for the pendulum distance $\mathcal{P}(\tilde{\theta}^1, \tilde{\theta}^2)$. Clearly, $|\bar{h}_1| + |1 - \bar{h}_2| \leqslant \epsilon^{2+r/2}$.

If $\bar{\theta}^1 \in \partial\Sigma_{23,4}^1$, then there are two possibilities.

**Possibility 1′**: $|x_2^1 - 2k_2'\pi| = \delta$ or $|x_3^1 - 2k_3'\pi| = \delta$. Let us consider

$$\bar{\theta}^{\sim 1} = (\tilde{x}_1^1, 2k_2'\pi, 2k_3'\pi, \tilde{x}_4^1) \in \Sigma_{23,4}^1.$$

Then by Lemmas 5.3 and 5.4, we have

$$\begin{aligned}
&\rho_\epsilon(\bar{\theta}_K, \bar{\theta}^{\sim 1}) + \rho_\epsilon(\bar{\theta}^{\sim 1}, \bar{\theta}^2) - \mathcal{S}^0(x_1^K, 2k_3'\pi) - \mathcal{S}^1(x_2^K, 2k_2'\pi) - \mathcal{P}(\tilde{\theta}^{\sim 1}, \tilde{\theta}^2) \\
&= 4\mathcal{A}_0 + \mathcal{A}(x_4^K) + 2\mathcal{A}(\tilde{x}_1^1) + \mathcal{A}(\tilde{x}_1^2) + \epsilon\mathcal{M}^+(0, 1, x_1^K, x_2^K) \\
&\quad + \epsilon\mathcal{M}^-(0, 1, 0, 0) + \epsilon\mathcal{M}^-(\bar{h}_1^\sim, \bar{h}_2^\sim, 0, 0) \\
&\quad + \epsilon\mathcal{M}^+(\bar{h}_1^\sim, \bar{h}_2^\sim, \tilde{x}_3^2, \tilde{x}_2^2) \pm \epsilon\frac{\kappa}{7},
\end{aligned}$$

where $\tilde{\theta}^{\sim 1} = (2k_3'\pi, 2k_2'\pi) \in \mathbb{R}^2$ and $\bar{h}_1^\sim$, $\bar{h}_2^\sim$ are determined by Lemma 5.2 for the pendulum distance $\mathcal{P}(\tilde{\theta}^{\sim 1}, \tilde{\theta}^2)$. It is easy to see that

$$|\bar{h}_1^\sim| + |1 - \bar{h}_2^\sim| \leqslant \epsilon^{2+r/2}.$$

Due to Lemma 6.1, we have

$$|\mathcal{S}^0(x_1^K, \tilde{x}_3^1) + \mathcal{S}^1(x_2^K, \tilde{x}_2^1) + \mathcal{P}(\tilde{\theta}^1, \tilde{\theta}^2) - \mathcal{S}^0(x_1^K, 2k_3'\pi) - \mathcal{S}^1(x_2^K, 2k_2'\pi) - \mathcal{P}(\tilde{\theta}^{\sim 1}, \tilde{\theta}^2)| \leqslant 4\epsilon^{1+r/4}.$$

Then by similar argument in Case I, we have that

$$\rho_\epsilon(\bar{\theta}_K, \bar{\theta}^1) + \rho_\epsilon(\bar{\theta}^1, \bar{\theta}^2) - \rho_\epsilon(\bar{\theta}_K, \bar{\theta}^{\sim 1}) - \rho_\epsilon(\bar{\theta}^{\sim 1}, \bar{\theta}^2) > 0,$$

which contradicts with the minimality of $\bar{\Theta}$. So this possibility is excluded.

**Possibility 2′**: $|\tilde{x}_1^1 - (2k_1' + 1)\pi| = \bar{\delta}$. Let us consider

$$\bar{\theta}_\sim^1 = ((2k_1' + 1)\pi, \tilde{x}_2^1, \tilde{x}_3^1, \tilde{x}_4^1) \in \Sigma_{23,4}^1.$$

Then by the same arguments in Case I, we obtain

$$\rho_\epsilon(\bar{\theta}_K, \bar{\theta}^1) + \rho_\epsilon(\bar{\theta}^1, \bar{\theta}^2) - \rho_\epsilon(\bar{\theta}_K, \bar{\theta}_\sim^1) - \rho_\epsilon(\bar{\theta}_\sim^1, \bar{\theta}^2) > 0.$$

This contradicts the minimality of $\bar{\Theta}$, which implies $\bar{\theta} \notin \partial\Sigma_{23,4}^1$.

Similarly, we show that $\bar{\theta}_K \notin \partial\Sigma_{12,3}^K$. Therefore $\bar{\Theta}$ is in the interior of $\mathbb{Y}$. □

When the minimum of $\mathbb{F}$ is attained in the interior of $\mathbb{Y}$, a standard approach in variational methods shows that the series of geodesics can be concatenated to a long geodesic of the Riemannian metric $d\rho_\epsilon$, which gives rise to a solution of the Euler-Lagrangian system (1.1).

Let $\gamma_l : [0, T_l] \to \mathbb{R}^4$, $l = 1, \ldots, 2K - 1$, be the minimizing geodesics, parametrized by time: For $l < K$, $\gamma_l$ minimizes $\rho_\epsilon(\bar{\theta}_l, \bar{\theta}_{l+1})$; for $l = K$, $\gamma_l$ minimizes $\rho_\epsilon(\bar{\theta}_K, \bar{\theta}^1)$; and $\gamma_l$ minimizes $\rho_\epsilon(\bar{\theta}^j, \bar{\theta}^{j+1})$ for $l > K$ with $j = l - K$. Let $T = \sum T_l$, we will show that $\gamma : [0, T] \to \mathbb{R}^4$ defined by

$$\gamma|[\sum_{l=0}^{j} T_l, \sum_{l=0}^{j+1} T_l] = \gamma_j|[0, T_j],$$

is a true geodesic. It is enough to prove that for any $l$, the concatenated curve $\gamma_l$ and $\gamma_{l+1}$ is a true geodesic. We will give detailed argument for the cases $l > K$. The other cases can be dealt with in the same way.

Write $\bar{\theta}^i = (\tilde{\theta}^k, \tilde{x}_4^i)$. Since $\bar{\theta}^i$ is in the interior of the section $\Sigma_{23,4}^i$, there exists a neighborhood $U$ of $\tilde{\theta}_i$ in $\mathbb{R}^3$ such that $U \times \{\tilde{x}_4^i\} \subset \Sigma_{23,4}^i$. Let us consider a cylindrical neighborhood of $\bar{\theta}^i$ given by $K = B_i \times \{|x_4 - \tilde{x}_4^i| \leqslant \nu\}$ where $\nu$ is small and $B_i$ is a small closed ball in $\mathbb{R}^3$ containing $\tilde{\theta}^i$. For $\nu$ and $B_i$ small enough, for any two points in $K$, the minimizing geodesic is unique in $K$ and contained in $U \times \{|x_4 - \tilde{x}_4^i| \leqslant \nu\}$. In fact we could take $\nu$ so small that $\gamma_i$ intersects $B_i \times \{x_4 = \tilde{x}_4^i - \nu\}$ at $y_0$, $\gamma_{i+1}$ intersects $B_i \times \{x_4 = \tilde{x}_4^i + \nu\}$ at $y_1$ and the minimizing geodesic connecting $y_0$ and $y_1$ crosses the section $\Sigma_{23,4}^i$ at the point $y'$. If the concatenated curve of $\gamma_i$ and $\gamma_{i+1}$ is not a geodesic, then $y' \neq \bar{\theta}_i$. See Figure 6.

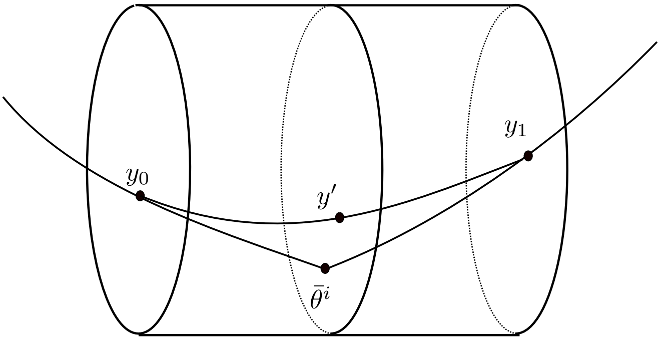


FIGURE 6. Concatenating geodesics.

Therefore,

$$\begin{aligned}
&\rho_\epsilon(\bar{\theta}^{i-1}, \bar{\theta}^i) + \rho_\epsilon(\bar{\theta}^i, \bar{\theta}^{i+1}) \\
&= \rho_\epsilon(\bar{\theta}^{i-1}, y_0) + \rho_\epsilon(y_0, \bar{\theta}^i) + \rho_\epsilon(\bar{\theta}^i, y_1) + \rho_\epsilon(y_1, \bar{\theta}^{i+1}) \\
&> \rho_\epsilon(\bar{\theta}^{i-1}, y_0) + \rho_\epsilon(y_0, y') + \rho_\epsilon(y', y_1) + \rho_\epsilon(y_1, \bar{\theta}^{i+1}) \\
&\geqslant \rho_\epsilon(\bar{\theta}^{i-1}, y') + \rho_\epsilon(y', \bar{\theta}^{i+1}).
\end{aligned}$$

This contradicts the minimality of $\bar{\Theta}$.

Now we have established that $\gamma$ is a real geodesic, hence it is an orbit of the Euler-Lagrangian system (1.1) when parametrized by time. It is not difficult to see that for each segment $\gamma_i$, for each $t_0 \in [0, T_i]$, we have at least one $j \in \{1, 2, 3, 4\}$ such that $E_j(t_0) \geqslant \frac{1}{4}$. By (3.7), $E_j(t) > \frac{1}{5}$, $\forall t \in [0, T_i]$, which implies $\dot{x}_j > \frac{1}{5}$. Therefore $T_i \leqslant 20\epsilon^{-2-r/2}$, and the total time

$$T = \sum_{i=1}^{2K-1} T_i \leqslant 160\epsilon^{-4-r}.$$

By similar argument as in Lemma 3.1, we have that for $i = 1, 2, 3$,

$$|E_i(t_i) - 1| \leqslant C\sqrt{\epsilon}, \quad |E_j(t_i)| \leqslant C\sqrt{\epsilon}, \quad j \neq i,$$

with $t_1 = 0$, $t_2 = \sum_{i=1}^{K} T_i$, and $t_3 = T$. Obviously,

$$|t_{i+1} - t_i| \leqslant 80\epsilon^{-4-r}, \quad i = 1, 2.$$

Clearly this construction can be carried out for any finite many sections of the types in (6.1). Thus for each $m \in \mathbb{N}$, and the sequence

$$\sigma_{-m} \dots \sigma_{-1}\sigma_0\sigma_1 \dots \sigma_m,$$

There exists an orbit $\gamma_m$ of the system such that there exists a sequence of time

$$t^m_{-m} < t^m_{-m+1} < \cdots < t^m_0 < t^m_1 < \cdots < t^m_m,$$

satisfying the following property:

$$|E_{\sigma_j}(t^m_j) - 1| \leqslant C\sqrt{\epsilon}, \quad |E_k(t^m_j)| \leqslant C\sqrt{\epsilon}, \quad k \neq \sigma_j.$$

Since the energy surface $H = 1$ is compact, so does the set $\{\gamma_m(t^m_0), \dot{\gamma}_m(t^m_0)\}_{m\in\mathbb{N}}$. Let $(\gamma^0_\infty, \dot{\gamma}^0_\infty)$ be one of its limiting points as $m \to \infty$. By Gronwall's lemma, we could see that the orbit $\gamma_\infty$ of the system with initial condition $(\gamma_\infty(0), \dot{\gamma}_\infty(0)) = (\gamma^0_\infty, \dot{\gamma}^0_\infty)$ satisfies the requirement of the Theorem 1.1. We thus prove the Theorem 1.1.

## Appendix A.

For arbitrary $4 \leqslant N \in \mathbb{N}$, let us consider the Lagrangian

$$\mathcal{L}^N_\epsilon = \sum_{i=1}^{N}(\frac{\dot{x}_i^2}{2} + 1 + \cos x_i) - \epsilon \sum_{i=1}^{N} \beta_\alpha(x_i, x_{i+1}, x_{i+2}),$$

Clearly, the corresponding Euler-Lagrange equation is the same as system (1.1). Now we choose $\bar{\delta}_N$ (similar to the $\bar{\delta}$ in (2.6)), such that

$$\mathcal{A}^-(-\pi + \bar{\delta}_N) = \mathcal{A}^+(\pi - \bar{\delta}_N) = \frac{\kappa}{400N}.$$

Then for the modified Lagrangians

$$\mathcal{L}^{h_k,h_{k+1}}_{\epsilon,k,N} = \sum_{i=k}^{k+1} \frac{1}{2}(\dot{x}_i - v_{h_i}(x_i))^2 + \sum_{i=k+2}^{k+N-1} \frac{\dot{x}_i^2}{2} + [1 + \cos x_i] - \epsilon \sum_{i=1}^{N} \beta_\alpha(x_i, x_{i+1}, x_{i+2}),$$

where $k \in \{1, \dots, N\}$ and we have imposed the periodic condition $x_{i+N} = x_i$, we have the following lemma which is paralleled to Lemma 3.1

**Lemma A.1.** *For any $k \in \{1, \dots, N\}$ and $h_k, h_{k+1} \in [0, 1]$, let $(x_1(t), \dots, x_N(t))$ minimizes*

$$\int_0^T \mathcal{L}^{h_k,h_{k+1}}_{\epsilon,k,N} \mathrm{dt},$$

*with the boundary conditions:*

$$\begin{cases} x_k(0) = x, & x_k(T) = q_{h_k}(x,T), \\ x_{k+1}(0) = y, & x_{k+1}(T) = q_{h_{k+1}}(y,T), \\ x_{k+2}(0) = 0, & x_{k+2}(T) = \pi, \\ x_i(0) = z_i, & x_i(T) = \pi, \quad i = k+3,\ldots,k+N-1, \end{cases}$$

*where* $x, y \in \mathbb{R}$ *are any real numbers and* $|z_i - \pi| \leqslant \bar{\delta}_N$.

*Then there exists* $\epsilon_0 > 0$, *which is independent of* $k$, $h_k$ *and* $h_{k+1}$, *such that for* $0 < \epsilon \leqslant \epsilon_0$ $T \geqslant \epsilon^{-1}$, *we have*

$$\int_0^T \mathcal{L}_{\epsilon,k,N}^{h_k,h_{k+1}}(x_1(t),\ldots,x_N(t))\mathrm{d}t = \mathcal{A}_0 + \sum_{i=k+3}^{k+N-1} \mathcal{A}(z_i) + \epsilon\mathcal{M}^+(h_1,h_2,x,y) \pm \epsilon\frac{\kappa}{32}.$$

Based on this fact, we could obtain results similar to those in Section 4 for the modified Lagrangians $\mathcal{L}_{\epsilon,k,N}^{h_k,h_{k+1}}$, $k = 1,\ldots,N$, and those in Section 5 for the corresponding Jacobi metric

$$\mathrm{d}\rho_{\epsilon,N} = \sqrt{2}\sqrt{1 + \sum_{i=1}^{N}(1+\cos x_i) - \epsilon\sum_{i=1}^{N}\beta_\alpha(x_i,x_{i+1},x_{i+2})}\,\mathrm{d}s,$$

where $\mathrm{d}s$ is the Euclidean metric in $\mathbb{R}^N$. Eventually, similar constructions as in Section 6 could be carried out and thus find the desired orbits of the system (1.1).

## Acknowledgements

The author thanks professor Vadim Kaloshin for proposing the project and for useful discussions. He also would like to thank the referee for valuable suggestions and comments.

## References

[1] P. W. Anderson. Absence of diffusion in certain random lattices. *Physical review*, 109(5):1492, 1958.

[2] V.I. Arnol'd. Instability of dynamical systems with many degrees of freedom. *Dokl.Akad. Nauk SSR*, 156:9–12, 1964.

[3] U. Bessi. An approach to Arnold's diffusion through the calculus of variations. *Nonlinear Analysis: Theory, Methods & Applications*, 26(6):1115–1135, 1996.

[4] U. Bessi. Arnold's diffusion with two resonances. *Journal of differential equations*, 137(2):211–239, 1997.

[5] S.V. Bolotin and P.H. Rabinowitz. A variational construction of chaotic trajectories for a reversible Hamiltonian system. *Journal of Differential Equations*, 148(2):364–387, 1998.

[6] C.Q. Cheng. Arnold diffusion in nearly integrable Hamiltonian systems. *preprint*, 2012.

[7] C.Q. Cheng and J.X. Xue. Arnold diffusion in nearly integrable Hamiltonian systems of arbitrary degrees of freedom. *preprint*, 2015.

[8] C.Q. Cheng and J. Yan. Arnold diffusion in Hamiltonian systems: a priori unstable case. *Journal of Differential Geometry*, 82(2):229–277, 2009.

[9] A. Delshams, R. de la Llave, and T.M. Seara. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. *Mem. Amer. Math. Soc*, 179(844), 2006.

[10] A. Fathi. Weak KAM Theorem in Lagrangian Dynamics, Seventh Preliminary Version. *preprint*, 2005.

[11] E. Fermi, J. Pasta, and S. Ulam. Studies of nonlinear problems. *Los Alamos Report LA-1940*, 978, 1955.

[12] J. Fröhlich, T. Spencer, and C. E. Wayne. Localization in disordered, nonlinear dynamical systems. *Journal of statistical physics*, 42(3-4):247–274, 1986.

[13] D. Hennig and G. P. Tsironis. Wave transmission in nonlinear lattices. *Physics Reports*, 307(5):333–432, 1999.

[14] A. Henrici and T. Kappeler. Results on normal forms for FPU chains. *Communications in Mathematical Physics*, 278(1):145–177, 2008.

[15] V. Kaloshin and M. Levi. An example of Arnold diffusion for near-integrable Hamiltonians. *Bulletin of the American Mathematical Society*, 45(3):409–427, 2008.

[16] V. Kaloshin and M. Levi. Geometry of Arnold diffusion. *SIAM review*, 50(4):702–720, 2008.

[17] V. Kaloshin, M. Levi, and M. Saprykina. Arnol'd diffusion in a pendulum lattice. *Communications on Pure and Applied Mathematics*, 67(5):748–775, 2014.

[18] V. Kaloshin and K. Zhang. A strong form of Arnold diffusion for two and a half degrees of freedom. *preprint*, 2012.

[19] V. Kaloshin and K. Zhang. Dynamics of the dominant Hamiltonian, with applications to Arnold diffusion. *preprint*, 2014.

[20] M. Levi. Shadowing property of geodesics in Hedlund's metric. *Ergodic Theory and Dynamical Systems*, 17(01):187–203, 1997.

[21] J. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Mathematische Zeitschrift*, 207(1):169–207, 1991.

[22] J. Mather. Variational construction of connecting orbits. *Annales de l'institut Fourier*, 43(5):1349–1386, 1993.

[23] A. Sorrentino. *Action-minimizing methods in Hamiltonian dynamics: an introduction to Aubry-Mather theory.* Princeton University Press, 2015.

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

[25] K. Zhang. Speed of Arnold diffusion for analytic Hamiltonian systems. *Inventiones mathematicae*, 186(2):255–290, 2011.

HUANG Guan
Department of Mathematics, University of Maryland at College Park, College Park, Maryland, USA