Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Functional Analysis 229 (2005) 1–61
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

On diffusion in high-dimensional Hamiltonian systems

Jean Bourgain[a,*], Vadim Kaloshin[b]

[a]*Institute for Advanced Study, Princeton, NJ 08540, USA*
[b]*California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 30 September 2004
Communicated by J. Bourgain
Available online 13 December 2004

Abstract

The purpose of this paper is to construct examples of diffusion for ε-Hamiltonian perturbations of completely integrable Hamiltonian systems in $2d$-dimensional phase space, with d large.

In the first part of the paper, simple and explicit examples are constructed illustrating absence of 'long-time' stability for size ε Hamiltonian perturbations of quasi-convex integrable systems already when the dimension $2d$ of phase space becomes as large as $\log \frac{1}{\varepsilon}$. We first produce the example in Gevrey class and then a real analytic one, with some additional work.

In the second part, we consider again ε-Hamiltonian perturbations of completely integrable Hamiltonian system in $2d$-dimensional space with ε-small but not too small, $|\varepsilon| > \exp(-d)$, with d the number of degrees of freedom assumed large. It is shown that for a class of analytic time-periodic perturbations, there exist linearly diffusing trajectories. The underlying idea for both examples is similar and consists in coupling a fixed degree of freedom with a large number of them. The procedure and analytical details are however significantly different. As mentioned, the construction in Part I is totally elementary while Part II is more involved, relying in particular on the theory of normally hyperbolic invariant manifolds, methods of generating functions, Aubry–Mather theory, and Mather's variational methods.

Part I is due to Bourgain and Part II due to Kaloshin.

* Corresponding author.
E-mail address: bourgain@math.ias.edu (J. Bourgain).

0022-1236/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2004.09.006

Part I: an example of diffusion for Hamiltonian perturbations of integrable systems in high dimension

1. Introduction

Consider a real analytic Hamiltonian, expressed in action-angle variables, of the form

$$H(I,\theta) = h(I) + \varepsilon f(I,\theta), \tag{1.1}$$

where $(I,\theta) \in \mathbb{R}^d \times \mathbb{T}^d$ and h, f are real analytic.

The corresponding equations of motion are

$$\begin{cases} \dot{I}_j = -\frac{\partial H}{\partial \theta_j} = -\varepsilon \frac{\partial f}{\partial \theta_j} \\ \dot{\theta}_j = \frac{\partial H}{\partial I_j} = \frac{\partial h}{\partial I_j} + \varepsilon \frac{\partial f}{\partial I_j} \end{cases} \quad (1 \leqslant j \leqslant d). \tag{1.2}$$

Thus H is an ε-perturbation of the integrable Hamiltonian $h(I)$, which we assume moreover to satisfy a strict convexity or quasi-convexity property. Recall that quasi-convexity means that

$$\langle D^2 h(I) v, v \rangle \geqslant c|v|^2 \tag{1.3}$$

required only to hold for vectors v orthogonal to $\nabla h(I)$.

A typical example of a quasi-convex h is

$$h(I) = I_1^2 + \cdots + I_d^2 + I_{d+1}.$$

The interest of the weaker quasi-convexity assumption is that it allows non-autonomous perturbations with a time-periodic dependence of a strictly convex Hamiltonian. Thus

$$H(I,\theta,t) = \sum_{j=1}^{d} I_j^2 + \varepsilon f(I,\theta,t) \tag{1.4}$$

with f 1-periodic in t, may be put in the quasi-convex format, considering an extra pair of action-angle variables (I_{d+1}, θ_{d+1}) and putting

$$\begin{aligned} &\mathcal{H}(I_1,\ldots,I_d,I_{d+1};\theta_1,\ldots,\theta_d,\theta_{d+1}) \\ &\quad = \sum_{j=1}^{d} I_j^2 + I_{d+1} + \varepsilon f(I_1,\ldots,I_d;\theta_1,\cdots,\theta_d,\theta_{d+1}). \end{aligned} \tag{1.5}$$

Recall the classical Nekhoroshev stability theorem (in the analytic category) stating that

$$|I(t) - I(0)| < \varepsilon^b \tag{1.6}$$

for times t satisfying

$$|t| < \exp\left(c\left(\frac{1}{\varepsilon}\right)^a\right), \tag{1.7}$$

where a, b, c are constants depending on d (one may take $a = b = \frac{1}{2d}$, cf. [L-N,Po]).

Moreover, as proved in [M-S], analogs of this stability result remain valid in the Gevrey classes as well, with very similar conclusions (see [M-S] for the detailed statement).

Our original motivation is the implementation of the Nekhoroshev stability mechanism in the context of Hamiltonian PDEs, obviously requiring indeed a better understanding of the role of the phase space dimension when $d \to \infty$. In fact, what matters is the dimension of the unperturbed torus (or, more precisely, its compactness properties), for which we try to obtain a long-time stability property (the torus may not be of full dimension). Research along this line has been pursued by several authors during recent years (cf. [Ba,Bo] to cite just a few references). Roughly speaking, no results are obtained from this approach, unless the dimension d of the torus satisfies

$$d \ll \log\frac{1}{\varepsilon} \quad (\varepsilon = \text{perturbation}) \tag{1.8}$$

(or the torus is subject to additional compactness assumptions, which somehow permit us to make good finite-dimensional approximations).

One should observe that (1.8) may be rather restrictive. For instance, in the $n+1$ body problem, $d = 3n$ and condition on the perturbation (caused by mutual attraction between planets) become extremely restrictive even for relatively small n.

In the other direction, counter examples illustrating the sharpness of the positive results about dependence of the exponents a and b in (1.6), (1.7) on the dimension d, were constructed in [M-S]. Basically, it is shown that $a \sim \frac{1}{d}$ is optimal (at least in the Gevrey class setting). This fact again suggests that one cannot go beyond (1.8) (it does not read immediately from the statement in [M-S] because of dimensional dependence of certain other factors).

Our purpose here is to describe a very simple example of the form (1.4) in dimension $d \sim \log\frac{1}{\varepsilon}$, that exhibits instability

$$|I_1(t) - I_1(0)| > c > 0 \tag{1.9}$$

for $t \sim \frac{1}{\varepsilon}$ (which is the 'obvious' stability time). Again f is time periodic and in Gevrey class $G^\alpha (\alpha > 1)$. The example is closely related to a construction in [Bo] (with f time independent but $d \sim \frac{1}{\varepsilon}$). It turns out that with some extra work, we may achieve a time-periodic analytic perturbation ($d \sim \log \frac{1}{\varepsilon}$) but with instability time $T_\varepsilon \sim \frac{1}{\varepsilon^{1+\tau}}$ (where $\tau > 0$ may be taken arbitrarily small).

For simplicity, we first treat the Gevrey version and put the main idea in evidence. It bares in fact some resemblance with the method in [M-S]. However, the coupling with a 'kicked pendulum' in [M-S] is replaced here by a shift on a high-dimensional torus; this permits us to create well-spaced long orbits that are basically exploited the same way as the 'irregular distribution' of certain periodic orbits of the kicked pendulum. It should also be pointed out that the mechanism described here may easily be realized as perturbations of 'classical Hamiltonian models' from physics or PDE, in particular the non-linear lattice Schrödinger operator with short-range interactions from [F-S-W] or the one-dimensional NLS on $\mathbb{T}$ (cf. [Bo]).

In some sense, it does permit us to prove a certain topological 'genericity' of the diffusion phenomenon in infinite-dimensional Hamiltonian systems (within the smooth category). But this will not be elaborated on here.

2. The example in Gevrey class

The format is (1.4)

$$H(I, \theta, t) = \tfrac{1}{2}(I_0^2 + I_1^2 + \cdots + I_d^2) + \varepsilon f(\theta, t), \tag{2.1}$$

where

$$\begin{aligned} f(\theta, t) = \cos\theta_0 \sum_{k=1}^{K} \sigma_k \Bigg\{ & \varphi^2\left(\frac{1}{d}\sum_{j=1}^{d} \cos(\theta_j - k\lambda_j)\right) \\ & -2v(t)\,\varphi\left(\frac{1}{d}\sum_{j=1}^{d}\cos(\theta_j - k\lambda_j)\right)\Bigg\}. \end{aligned} \tag{2.2}$$

Here

$$d \sim \log\frac{1}{\varepsilon} \quad \text{and} \quad K \sim \frac{1}{\varepsilon}. \tag{2.3}$$

Take φ to satisfy

$$\begin{cases} \varphi(u) = 0 \ \text{ if } \ |u| < 10^{-2}, \\ \varphi(u) = 1 \ \text{ if } \ |u| > \frac{1}{2}. \end{cases} \tag{2.4}$$

Thus φ cannot be real analytic but may be taken in any Gevrey class $G^\alpha(\mathbb{R})$ for $\alpha > 1$.

The function $v(t)$ will be a 1-periodic Gevrey-class function on $\mathbb{R}$ satisfying

$$v(t) = \psi(t) \text{ for } |t| \leqslant \tfrac{1}{2}, \tag{2.5}$$

where

$$\psi(t) = \varphi\left(\frac{1}{d}\sum_{j=1}^{d} \cos \lambda_j t\right). \tag{2.6}$$

The frequencies $\bar{\lambda} = (\lambda_j)_{1 \leqslant j \leqslant d}$ are taken in a bounded interval in $\mathbb{R}$, chosen randomly according to some probability distribution, as described below. We will in particular insure that for any given $\theta = (\theta_1, \ldots, \theta_j) \in \mathbb{T}^d$, only a bounded number of terms in (2.2) do not vanish. This will permit us to claim the Gevrey bound on f.

Finally, the $\sigma_k (1 \leqslant k \leqslant K)$ are numbers in $[-1, 1]$ which should be considered as parameters. They will eventually be obtained from a Schauder fixpoint argument, in order to fulfill a certain resonance.

(I) *Choice of the frequencies*: We basically proceed as in [Bo] (see the last section).

Consider a probability density μ on $\mathbb{R}$, $\operatorname{supp} \mu$ bounded, satisfying

$$|\hat{\mu}(x)| < (1 + 10^4 |x|)^{-50}. \tag{2.7}$$

Choose the λ_j independently according to μ. Denote v the d-fold product measure $dv = \mu(d\lambda_1) \otimes \cdots \otimes \mu(d\lambda_d)$. Fix $1 \leqslant R \leqslant e^d$ and estimate using (2.7)

$$\begin{aligned} &\int \sup_{\frac{R}{10^2} \leqslant |x| \leqslant R^2} \left|\sum_{j=1}^{d} e^{i\lambda_j x}\right| v(d\bar{\lambda}) \\ &\leqslant \int \sup_{|x| \leqslant R^2} \left|\sum_{j=1}^{d} \left[e^{i\lambda_j x} - \int e^{i\lambda x} \mu(d\lambda)\right]\right| dv \\ &+ (100R)^{-50} d. \end{aligned} \tag{2.8}$$

For fixed x, the terms $e^{i\lambda_j x} - \int e^{i\lambda x}\mu(d\lambda) \quad (1 \leqslant j \leqslant d)$ are independent mean zero variables in λ wrt dv. So by the standard bounds (in the Gaussian setting referred to as Dudley's majoration), (2.8) is bounded by

$$C \int_0^\infty \sqrt{\log \mathcal{N}(\rho)}\, d\rho, \tag{2.9}$$

where $\mathcal{N}(\rho)$ stands for the metric entropy numbers of the set $[|x| \leqslant R^2]$ (i.e. the minimum number of balls of radius ρ needed to cover the set), for the pseudo-distance

$$d(x,x') = \sup_{\bar\lambda \in \operatorname{supp} \nu} \left[\sum_{j=1}^{d} \left| \left(e^{i\lambda_j x} - \int e^{i\lambda x} \mu(d\lambda) \right) - \left(e^{i\lambda_j x'} - \int e^{i\lambda x'} \mu(d\lambda) \right) \right|^2 \right]^{1/2}$$

$$< C d^{1/2} (|x - x'| \wedge 1).$$

Hence $\mathcal{N}(\rho) < C d^{1/2} R^2 \rho^{-1}$ and

$$(2.9) < C \int_0^{\sqrt{d}} \sqrt{\log\left(1 + \frac{d^{1/2} R^2}{\rho}\right)}\, d\rho \sim (\log R)^{1/2} \sqrt{d}. \tag{2.10}$$

Therefore we get

$$\int \sup_{\frac{R}{10^2} \leqslant |x| \leqslant R^2} \left| \sum_{j=1}^{d} e^{i\lambda_j x} \right| d\nu < C\sqrt{d}\sqrt{\log R} + (100R)^{-50}\, d. \tag{2.11}$$

Taking R of the form $e^{-2^{-s}d}$ $(s = 0, 1, 2, \ldots, \log d)$, one easily derives from (2.11)

Lemma 2.12. *A random choice* $\bar\lambda = (\lambda_j)_{1 \leqslant j \leqslant d}$ *as above satisfies the inequality*

$$\left| \sum_{j=1}^{d} e^{i\lambda_j x} \right| < 10^{-50} d + C d^{3/4} (\log(1 + |x|))^{1/4}$$

for all $|x| > 10^{-2}$.

We assume K in (2.2) to satisfy

$$\log K < c_0 d \tag{2.13}$$

with c_0 small enough to ensure that by (2.12)

$$\left| \sum_{j=1}^{d} e^{i\lambda_j x} \right| < 10^{-50} d \quad \text{if} \quad 10^{-2} < |x| < e^{c_0 d}. \tag{2.14}$$

Returning to (2.2), let us verify the claim that with this choice of $\bar\lambda = (\lambda_j)$, for any given $(\theta_1, \ldots, \theta_d) \in \mathbb{T}^d$, there are at most 10^5 non-vanishing terms in the k-sum. Let thus $I \subset \{1, \ldots, K\}$ with $\varphi\left(\frac{1}{d}\sum_{j=1}^{d} \cos(\theta_j - k\lambda_j)\right) \neq 0$ for $k \in I$.

By definition (2.4) of φ, it follows that

$$\left|\sum_{j=1}^{d} \cos(\theta_j - k\lambda_j)\right| \geqslant 10^{-2} d \quad \text{for } k \in I$$

and we may clearly assume

$$\sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) \geqslant 10^{-2} d \quad \text{for } k \in I. \tag{2.15}$$

Denoting the vectors $\xi = (e^{i\theta_j})_{1 \leqslant j \leqslant d}$ and $\xi_k = (e^{ik\lambda_j})_{1 \leqslant j \leqslant d}$ for $k \in I$, by (2.15)

$$10^{-2} d|I| < \sum_{k \in I} \operatorname{Re} \langle \xi, \xi_k \rangle \leqslant \sqrt{d} \left| \sum_{k \in I} \xi_k \right|.$$

Hence

$$|I|d + 2 \sum_{k < k'} \operatorname{Re} \langle \xi_k, \xi_{k'} \rangle \geqslant 10^{-4} d|I|^2$$

and since, by (2.14), $|\langle \xi_k, \xi_{k'} \rangle| = \left| \sum_{j=1}^{d} e^{i\lambda_j (k-k')} \right| < 10^{-50} d$ for $k \neq k'$,

$$d|I| + 10^{-50} d|I|^2 > 10^{-4} d|I|^2$$

and

$$|I| < 2 \cdot 10^4.$$

This proves the claim.

The function $\psi = \psi(t)$ defined in (2.6) is Gevrey $G^{\alpha} (\alpha > 1)$ and satisfies

$$\psi(t) = 0 \quad \text{for} \quad 10^{-2} < |t| < K$$

(the first statement follows from the composition properties of Gevrey functions; see for instance the Appendix in [M-S]).

Take v to be the 1-periodic function on $\mathbb{R}$ satisfying

$$v \Big|_{[-\frac{1}{2}, \frac{1}{2}]} = \psi \Big|_{[-\frac{1}{2}, \frac{1}{2}]}.$$

Obviously $v \in G^{\alpha}(\mathbb{T})$, $\alpha > 1$. By construction

$$\varphi'(\ ) = \varphi\left(\frac{1}{d}\sum_{j=1}^{d}\cos\lambda_j(t-k)\right) = 0 \quad \text{if} \quad |t|, |k| \leqslant K \quad \text{and} \quad |t-k| \geqslant 10^{-2} \tag{2.16}$$

and

$$\varphi\left(\frac{1}{d}\sum_{j=1}^{d}\cos\lambda_j(t-k)\right) = v(t-k) = v(t) \quad \text{if} \quad |t-k| \leqslant \frac{1}{2}. \tag{2.17}$$

(II) *Equations of motion and choice of* $\sigma_k (1 \leqslant k \leqslant K)$: We restrict time to

$$0 \leqslant t \leqslant K \quad \text{where } K = \left[\frac{\lambda}{10\varepsilon}\right] \tag{2.18}$$

(γ is a small constant).

Returning to (2.1), (2.2) the equations of motion are

$$\begin{cases} \dot{\theta} = I_j \quad (0 \leqslant j \leqslant d), & (2.19) \\ \dot{I}_0 = \varepsilon \sin\theta_0 \sum_{k=1}^{K} \sigma_k \left\{ \varphi^2\left(\frac{1}{d}\sum_{j=1}^{d}\cos(\theta_j - k\lambda_j)\right) \right. & \\ \qquad \left. -2v(t)\varphi\left(\frac{1}{d}\sum_{j=1}^{d}\cos(\theta_j - k\lambda_j)\right)\right\}, & (2.20) \\ \dot{I}_j = \frac{2\varepsilon}{d}\cos\theta_0 \sum_{k=1}^{K}\sigma_k\varphi'\left(\frac{1}{d}\sum_{j=1}^{d}\cos(\theta_{j'} - k\lambda_{j'})\right), & \\ \sin(\theta_j - k\lambda_j)\left\{\varphi\left(\frac{1}{d}\sum_{j'=1}^{d}\cos(\theta_{j'} - k\lambda_{j'})\right) - v(t)\right\} \quad (1 \leqslant j \leqslant d). & (2.21) \end{cases}$$

Take as initial conditions

$$\begin{cases} \theta_j(0) = 0 \quad (0 \leqslant j \leqslant d), & (2.22) \\ I_0(0) = \gamma \quad (\gamma \text{ as above}), & (2.23) \\ I_j(0) = \lambda_j \quad (1 \leqslant j \leqslant d). & (2.24) \end{cases}$$

First, we show that for $j = 1, \ldots, d$, the action variable I_j remains preserved for $0 \leqslant t \leqslant K$.

This will be a consequence to the definition of v and H.

Denote for $1 \leqslant j \leqslant d$

$$\delta_j = \delta_j(t) = \max_{0 \leqslant t' \leqslant t} |I_j(t') - \lambda_j|.$$

From (2.19)

$$|\theta_j(t) - \lambda_j t| \leqslant \delta_j t.$$

From (2.21), for $1 \leqslant j \leqslant d$

$$|\dot{I}_j| \leqslant \frac{2\varepsilon}{d} \sum_{k=1}^{K} \left| \varphi' \left(\frac{1}{d} \sum_{j'=1}^{d} \cos(\theta_{j'} - k\lambda_{j'}) \right) \right| \left| \varphi \left(\frac{1}{d} \sum_{j'=1}^{d} \cos(\theta_{j'} - k\lambda_{j'}) \right) - v(t) \right|$$

$$\leqslant \frac{2\varepsilon}{d} \sum_{k=1}^{K} \left| \varphi' \left(\frac{1}{d} \sum_{j'=1}^{d} \cos \lambda_{j'}(t-k) \right) \right| \left| \varphi \left(\frac{1}{d} \sum_{j'=1}^{d} \cos \lambda_{j'}(t-k) \right) - v(t) \right| \quad (2.25)$$

$$+ C \frac{\varepsilon K}{d^2} \sum_{j'=1}^{d} |\theta_{j'} - t\lambda_{j'}|. \quad (2.26)$$

Consider first (2.25). Recalling (2.16), (2.17), $\varphi' \left(\frac{1}{d} \sum_{j'=1}^{d} \cos \lambda_{j'}(t-k) \right) = 0$ if $|t - k| > 10^{-2}$ and $v(t) = \varphi \left(\frac{1}{d} \sum_{j'=1}^{d} \cos \lambda_{j'}(t-k) \right)$ if $|t-k| \leqslant \frac{1}{2}$. Hence (2.25) =0.

Since (2.26) $< C \frac{\varepsilon K t}{d^2} \sum_{j=1}^{d} \delta_j$, it follows that

$$\sum_{j=1}^{d} \delta_j(t) \leqslant C \frac{\varepsilon K}{d} t^2 \sum_{j=1}^{d} \delta_j(t).$$

Therefore, for t small enough, $\sum_{j=1}^{d} \delta_j(t) = 0$. Consequently, his holds for all $t \leqslant K$. Since $I_j(t) = \lambda_j$ for $1 \leqslant j \leqslant d$, $\theta_j(t) = \lambda_j t$ and

$$\varphi^2 \left(\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) \right) = \varphi \left(\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) \right) v(t).$$

Hence (2.20) becomes

$$\dot{I}_0 = -\varepsilon \sin \theta_0 \sum_{k=1}^{K} \sigma_k \varphi^2 \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j(t-k) \right). \quad (2.27)$$

Again by (2.16)

$$\sum_{k=1}^{K} \left| \varphi \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j(t-k) \right) \right| \leqslant 1$$

so that certainly $|\dot{I}_0| \leqslant \varepsilon$ and by (2.23), (2.18), (2.19)

$$|I_0(t) - \gamma| \leqslant \varepsilon |t| \leqslant \varepsilon K < \gamma/10,$$

$$|\theta_0(t + \Delta t) - \theta_0(t) - \gamma \Delta t| < \frac{\gamma}{10} |\Delta t|. \tag{2.28}$$

Next, integrate (2.27) between $0 \leqslant t \leqslant K$

$$\begin{aligned} I_0(K) - I_0(0) &= -\varepsilon \sum_{k=1}^{K} \sigma_k \int_{|\tau|<10^{-2}} \sin \theta_0(k+\tau) \varphi^2 \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j \tau \right) d\tau \\ &\overset{(2.28)}{=} -\varepsilon \left[\int_{|\tau|<10^{-2}} \varphi^2 \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j \tau \right) d\tau \right] \\ &\quad \times \left[\sum_{k=1}^{K} \sigma_k \sin \theta_0(k) \right] + 0(\varepsilon K \gamma). \end{aligned} \tag{2.29}$$

We now explain how $\bar{\sigma} = (\sigma_k)_{1 \leqslant k \leqslant K}$ is chosen. Summarizing the preceeding, the Hamiltonian H in (2.1), (2.2) depends on $\bar{\sigma} \in [-1, 1]^K$ and so does therefore $(\theta_0(k))_{1 \leqslant k \leqslant K}$ (since the initial conditions are fixed). The map

$$\bar{\sigma} \mapsto (\sin \theta_0(k))_{1 \leqslant k \leqslant K}$$

is obviously a continuous transformation of $[-1, 1]^K$. Invoking Schauder's fixpoint theorem, we may therefore find $\bar{\sigma} = (\sigma_k)$ such that

$$\sigma_k = \sin \theta_0(k) \quad (1 \leqslant k \leqslant K). \tag{2.30}$$

Eq. (2.29) implies then

$$|I_0(K) - I_0(0)| > c\varepsilon \sum_{k=1}^{K} \sin^2 \theta_0(k) - C\varepsilon\gamma K. \tag{2.31}$$

Using (2.28), it is easily seen that $\sum_{k=1}^{K} \sin^2 \theta_0(k) \sim \frac{K}{2}$. Letting γ be small enough and recalling (2.26), we get

$$\left| I_0\left(\frac{\gamma}{10\varepsilon}\right) - I_0(0) \right| > c\varepsilon K \sim \gamma. \tag{2.32}$$

Consequently, in a time interval $[0, T]$, $T \sim \frac{1}{\varepsilon}$, the action-variable I_0 undergoes a drift of 0(1). Recalling (2.3), (2.13), the phase dimension $2d \sim \log \frac{1}{\varepsilon}$. The perturbation is $\varepsilon f(\theta, t)$, where $\|f\|_{G^\alpha(\mathbb{T}^{d+1})} < C \quad (\alpha > 1 \text{ fixed})$.

3. An analytic example

Since non-zero analytic functions cannot vanish on an interval, approximations arise that create some complications. We redefine $f = f(\theta, t)$ in (2.2) as

$$f(\theta, t) = \cos\theta_0 \sum_{k=1}^{K} \sigma_k \left\{ \varphi^3 \left(\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) \right) - 3\varphi^2 (\cdots)\, v(t) + 3\varphi (\cdots)\, v(t)^2 \right\}. \tag{3.1}$$

The motion equations for the action variables then become

$$\dot{I}_0 = -\varepsilon \frac{\partial f}{\partial \theta_0} = \varepsilon \sin\theta_0 \sum_{k=1}^{K} \sigma_k \left\{ \left[\varphi \left(\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) \right) - v(t) \right]^3 + v(t)^3 \right\} \tag{3.2}$$

and for $j = 1, \ldots, d$

$$\begin{aligned} \dot{I}_j &= -\varepsilon \frac{\partial f}{\partial \theta_j} \\ &= \frac{3\varepsilon}{d} \cos\theta_0 \sum_{k=1}^{K} \sigma_k \varphi' \left(\frac{1}{d} \sum_{j'=1}^{d} \cos(\theta_{j'} - k\lambda_{j'}) \right) \sin(\theta_j - k\lambda_j) [\varphi(\cdots) - v(t)]^2 \end{aligned} \tag{3.3}$$

replacing (2.20), (2.21). The point of the modification is the $[\varphi(\cdots) - v(t)]^2$ factor in (3.3) that will create 'near vanishing' up to first order for $\bar{\theta} = \bar{\lambda} t$, where $\bar{\theta} = (\theta_1, \ldots, \theta_d)$.

Of course, φ and v in (3.1) need to be chosen as real analytic.

We first produce an appropriate analytic substitute for φ, cf. (2.4).

Lemma 3.4. *Fix $m \in \mathbb{Z}_+$ and consider the polynomial*

$$\varphi(u) = \sqrt{m} 4^m \int_{\frac{1}{2}}^{1} (su^2)^m (1 - su^2)^m \, ds. \tag{3.5}$$

The following properties hold:

$$|\varphi(z)| < 24^m \text{ for } z \in \mathbb{C}, |z| \leqslant \sqrt{2}, \tag{3.6}$$

$$|\varphi(z)| < 8^m |z|^{2m} \text{ for } z \in \mathbb{C}, |z| \leqslant 1, \tag{3.7}$$

$$0 \leqslant \varphi(u) \leqslant 10 \quad \text{if } u \in [-1, 1], \tag{3.8}$$

$$\varphi(u) > \frac{1}{10} \quad \text{if } u \in [-1, 1], |u| \geqslant \frac{1}{\sqrt{2}}. \tag{3.9}$$

Proof. Eqs. (3.6), (3.7) are immediate from (3.5).

For $u \in [-1, 1]$, write

$$\varphi(u) = \sqrt{m} \int_{\frac{1}{2}}^{1} (1 - (1 - 2su^2)^2)^m \, ds \leqslant \sqrt{m} \int_{\frac{1}{2}}^{1} e^{-m(1-2su^2)^2} \, ds < \frac{1}{u^2},$$

which together with (3.7) implies (3.8).

Also

$$\begin{aligned} \varphi(u) &> \sqrt{m} \int_{\frac{1}{2}}^{1} e^{-2m(1-2su^2)^2} \, ds \\ &= \frac{1}{2u^2} \int_{\sqrt{m}(1-2u^2)}^{\sqrt{m}(1-u^2)} e^{-2y^2} \, dy > \frac{1}{10} \quad \text{if } \frac{1}{\sqrt{2}} \leqslant |u| \leqslant 1. \end{aligned}$$

This proves Lemma 3.4. □

Define $m \in \mathbb{Z}_+$ by

$$\varepsilon = 10^{-10m}. \tag{3.10}$$

Take the frequencies $(\lambda_j)_{1 \leqslant j \leqslant d}$ as before and define $\psi = \psi(t)$ by (2.6). Thus, by (3.6), ψ admits an analytic extension to some neighborhood of the real axis, on which it is bounded by 24^m. Furthermore, it follows from (2.14), (3.7) that

$$|\psi(t)| < \left(\frac{8}{10^{100}}\right)^m < 10^{-99m} \quad \text{if} \quad 10^{-2} \leqslant |t| \leqslant K \tag{3.11}$$

and similar if in (2.6) we replace φ by φ' or φ''.

Next, we introduce the 1-periodic function v. Define

$$v(t) = \sum_{j \in \mathbb{Z}} \psi(t + j) e^{-\varepsilon^4 (t+j)}, \tag{3.12}$$

which is 1-periodic.

If $|t| \leqslant \frac{1}{2}$, it follows from (3.10), (3.11)

$$|v(t) - \psi(t)| \leqslant \varepsilon^4 + \sum_{j \neq 0} |\psi(t+j)| e^{-\varepsilon^4 (t+j)}$$
$$< \varepsilon^4 + 10^{-99m} \varepsilon^{-5} + \varepsilon^{-4} e^{-\frac{1}{\varepsilon}} < 2\varepsilon^4. \tag{3.13}$$

Assume, cf. (2.14)

$$e^{c_0 d} > \varepsilon^{-10}. \tag{3.14}$$

Consider the analytic extensions of ψ and v to $|\mathrm{Im}\, z| < \rho$ (ρ a sufficiently small constant) which we estimate in more detail.

Write according to (2.6), (3.12)

$$\psi(z) = \varphi\left(\frac{1}{d}\sum_{j=1}^{d} \cos \lambda_j z\right), \tag{3.15}$$

$$v(z) = \sum_{j \in \mathbb{Z}} \psi(z+j) e^{-\varepsilon^4 (z+j)}. \tag{3.16}$$

Write $z = t + iy$, $|y| < \rho$. Clearly

$$\frac{1}{d}\sum_{j=1}^{d} \cos \lambda_j z = \frac{1}{d}\sum_{j=1}^{d} \cos \lambda_j t + 0(\rho) \tag{3.17}$$

(since the λ_j are bounded by some constant).

Therefore, assuming $10^{-2} \leqslant |t| < e^{c_0 d}$, (2.14) implies (for ρ small enough)

$$\left|\frac{1}{d}\sum_{j=1}^{d} \cos \lambda_j z\right| < 10^{-50} + 0(\rho) < 10^{-49}$$

and by (3.7).

$$|\psi(z)| < (8 \cdot 10^{-98})^m < 10^{-97m}. \tag{3.18}$$

Using (3.18) for $10^{-2} < |\mathrm{Re}\, z| < \varepsilon^{-10}$, estimate from (3.16), (3.10), (3.6)

$$|v(\tau)| \leqslant |\psi(z)| + 10^{-97m} \varepsilon^{-5} + \varepsilon^{-4} e^{-\frac{1}{\varepsilon}} 24^m < 25^m \quad (|\mathrm{Im}\, z| < \rho). \tag{3.19}$$

Returning to the function f in (3.1) and its analytic extension, estimate

$$\sum_{k=1}^{K} \left| \varphi \left(\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j + iy_j - k\lambda_j) \right) \right|, \tag{3.20}$$

where $|y_j| \leqslant \rho$ $(1 \leqslant j \leqslant d)$.

Since

$$\frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j + iy_j - k\lambda_j) = \frac{1}{d} \sum_{j=1}^{d} \cos(\theta_j - k\lambda_j) + 0(\rho), \tag{3.21}$$

the same calculation performed in Section 2, based on (2.14), shows that (3.21) is bounded by 10^{-24}, except for at most 10^{50} values of k. It follows then from (3.6), (3.7) that

$$(3.20) < 10^{50} \cdot 24^m + K(8 \cdot 10^{-48})^m < 25^m. \tag{3.22}$$

Consequently, by (3.19), (3.22), the function $f(\theta, t)$ on $\mathbb{T}^{d+1} \times \mathbb{T}$ admits an analytic extension $\tilde{f}$ to a ρ-neighborhood, satisfying the bound

$$|\tilde{f}| < 25^{3m}. \tag{3.23}$$

Recalling (3.10), the perturbation $\varepsilon \tilde{f}$ satisfies

$$\varepsilon \|\tilde{f}\| < \varepsilon 10^{5m} < \varepsilon^{1/2}. \tag{3.24}$$

The preceeding may easily be modified to give a bound $\varepsilon^{1-\tau}$ ($\tau > 0$ arbitrary) as well (but the width ρ of the analyticity region will decrease).

Consider the equation of motion (2.19), (3.2), (3.3).

Estimate again for $1 \leqslant j \leqslant d$

$$\delta_j = \max_{0 \leqslant t \leqslant k} |I_j(t) - \lambda_j|. \tag{3.25}$$

From (2.19)

$$|\theta_j(t) - \lambda_j t| < \delta_j |t| < \delta_j K. \tag{3.26}$$

The bounds (2.25)–(2.26) are too crude. Performing a Taylor expansion up to order 2 for $\partial_{\theta_j} f$ at $\bar{\theta} = \bar{\lambda} t$, we get

$$\partial_{\theta_j} f(\theta_0, \bar{\theta}, t) = \partial_{\theta_j} f(\theta_0, \bar{\lambda} t, t) \tag{3.27}$$

$$+ \sum_{j'=1}^{d} \partial^2_{\theta_j \theta_{j'}} f(\theta_0, \bar{\lambda} t, t)(\theta_{j'} - \lambda_{j'} t) \tag{3.28}$$

$$+ 0\left(\left| \frac{\partial^2}{\partial \bar{\theta}^2} \partial_{\theta_j} f \right| \ |\bar{\theta} - \bar{\lambda} t|^2 \right). \tag{3.29}$$

From (3.3)

$$|(3.27)| \lesssim \frac{1}{d} \sum_{k=1}^{K} \left| \varphi' \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j (t-k) \right) \right| \psi(t-k) - v(t)|^2. \tag{3.30}$$

For $|t-k| \leqslant \frac{1}{2}$ (3.13) applies and $|\psi(t-k) - v(t)| \lesssim \varepsilon^4$.

If $|t-k| > \frac{1}{2}$, (3.9) gives

$$(|\varphi| + |\varphi'| + |\varphi''|) \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j (t-k) \right) < 10^{-99m}.$$

Therefore

$$|(3.27)| \lesssim \frac{\varepsilon^8}{d} + \frac{1}{d} K 10^{-99m} < \frac{\varepsilon^8}{d} + \frac{1}{d} \frac{1}{\varepsilon} \varepsilon^{\frac{99}{10}} < 2 \frac{\varepsilon^8}{d}. \tag{3.31}$$

Estimating from (3.3) the next derivative

$$\begin{aligned} &\left| \partial^2_{\theta_j \theta_{j'}} f(\theta_0, \bar{\lambda} t, t) \right| \\ &\quad \leqslant \sum_{k=1}^{k} \left\{ \left[\frac{1}{d} \left| \varphi' \left(\frac{1}{d} \sum_{j=1}^{d} \cos \lambda_j (t-k) \right) \right| + \frac{1}{d^2} |\varphi''(\cdots)| \right] |\psi(t-k) - v(t)|^2 \right. \\ &\qquad \left. + \frac{1}{d^2} |\varphi'(\cdots)|^2 |\psi(t-k) - v(t) \right\} \\ &\quad < \frac{\varepsilon^4}{d^2} + \frac{K}{d} 10^{-99m} \lesssim \frac{\varepsilon^4}{d^2}. \end{aligned} \tag{3.32}$$

Hence, from (3.26), (3.32)

$$|(3.28)| < \frac{\varepsilon^4}{d^2} K \sum_{j=1}^{d} \delta_j. \tag{3.33}$$

Since we already established the bound (3.23) on the analytic extension $\tilde{f}$ of f, write

$$|(3.29)| \lesssim \rho^{-3} 25^{3m} K^2 \left(\sum \delta_j\right)^2 < 10^{5m} K^2 \left(\sum \delta_j\right)^2. \tag{3.34}$$

Collecting estimates (3.31), (3.33), (3.34), (3.3) implies

$$\begin{aligned} \delta_j &\leqslant \varepsilon K \left(\frac{\varepsilon^8}{d} + \frac{\varepsilon^4}{d^2} K \left(\sum_{j \geqslant 1} \delta_j \right) + 10^{5m} K^2 \left(\sum \delta_j\right)^2 \right), \\ \sum_{j \geqslant 1} \delta_j &< \varepsilon^8 + \varepsilon^3 \left(\sum_{j \geqslant 1} \delta_j \right) + 10^{5m} \varepsilon^{-2} \left(\sum \delta_j\right)^2, \\ \sum_{j \geqslant 1} \delta_j &\lesssim \varepsilon^8 \end{aligned} \tag{3.35}$$

and again from (3.26)

$$|\theta_j(t) - \lambda_j t| < \varepsilon^7 \quad (1 \leqslant j \leqslant d). \tag{3.36}$$

From (3.36)

$$\varphi \left(\frac{1}{d} \sum_{j=1}^{d} \cos\left(\theta_j(t) - k\lambda_j\right) \right) = \psi(t-k) + 0(\varepsilon^7). \tag{3.37}$$

Substitution of (3.37) in (3.2) gives

$$\begin{aligned} \dot{I}_0 = \varepsilon \sin \theta_0 \sum_{k=1}^{K} \sigma_k \Bigg\{ & \varphi^3 \left(\frac{1}{d} \sum_{j \geqslant 1} \cos(\theta_j - k\lambda_j) \right) \\ & -3\varphi^2(\cdots) v(t) + 3\varphi(\cdots) v(t)^2 \Bigg\} \end{aligned}$$

$$= \varepsilon \sin\theta_0 \sum_{k=1}^{K} \sigma_k \left[\psi(t-k)^3 - 3\psi(t-k)^2 v(t) + 3\psi(t-k)v(t)^2\right] + 0(\varepsilon K \varepsilon^7)$$

$$\overset{(3.11),(3.13)}{=} \varepsilon \sin\theta_0 \sum_{k=1}^{K} \sigma_k \psi(t-k)^3 + 0(\varepsilon K \varepsilon^7 + \varepsilon K \varepsilon^4)$$

$$= \varepsilon \sin\theta_0(t) \sum_{k=1}^{K} \sigma_k \psi(t-k)^3 + 0(\varepsilon^4). \tag{3.38}$$

The remainder of the argument is the same. Thus from (3.38)

$$\begin{aligned} I_0(K) - I_0(0) &= \varepsilon \sum_{k=1}^{K} \sigma_k \int_{|t|<K} \sin\theta_0(k+\tau)\psi(\tau)^3 d\tau + 0(\varepsilon^4 K) \\ &\overset{(3.11)}{=} \varepsilon \sum_{k=1}^{K} \sigma_k \int_{|\tau|<10^{-2}} \sin\theta_0(k+\tau)\psi(\tau)^3\, d\tau + 0(\varepsilon^3) \\ &= \varepsilon \left(\sum_{k=1}^{K} \sigma_k \sin\theta_0(k)\right) \left(\int_{|\tau|<10^{-2}} \psi(\tau)^3\, d\tau\right) \\ &\quad + 0(\varepsilon^3 + \varepsilon K \gamma). \end{aligned} \tag{3.39}$$

Again using a Schauder fixpoint argument, we ensure that $\sigma_k = \sin\theta_0(k)$ $(1 \leqslant k \leqslant K)$.

From (3.9), (3.39) gives the minoration

$$\begin{aligned} |I_0(K) - I_0(0)| &> \varepsilon \left[\int_{|\tau|<10^{-2}} \psi(\tau)^3 d\tau\right] \left[\sum_{k=1}^{k} \sin^2\, \theta_0(k)\right] + 0(\varepsilon K \gamma) \\ &> c\varepsilon K(1 - C\gamma) > c'\gamma. \end{aligned}$$

Part II: diffusion for Hamiltonian perturbations of integrable systems in high dimensions using Mather's variational methods

4. Statement of the results

In this section we state several results about the existence of diffusing trajectories. Let $h(I)$ be a strictly convex Hamiltonian as above.[1] There will be two types of results: One is for *autonomous* and the other is for *time-periodic* real-analytic perturbations respectively. The latter could be also treated as autonomous perturbations of quasi-convex $h(I)$'s. According to the general principle a convex autonomous system close

[1] We could relax this condition (see (4.3)).

to an integrable behaves on a bounded time interval as if it is a quasi-convex system (see Section 4.1). This reduces proofs of autonomous convex results to time-periodic convex ones. We start by stating results for time-periodic perturbations and then state their analogs for autonomous perturbations.

The first, probably main, result is about existence of trajectories diffusing linearly fast for all time. The second result says that once we fix direction in action space (say I_0), then there are diffusing trajectories approximately performing any prescribed ahead motion along I_0 linearly fast in time. This is similar to a result in [MS2]. Even though constants of exponential are not explicitly computed our result holds not only for Gevrey (as in [M-S,MS2]), but also for analytic perturbations. Other important extensions are wider class of unperturbed integrable Hamiltonians $h(I)$'s and construction of diffusion in the case of autonomous perturbations (Theorem 4.3). After our results for time-periodic perturbations are stated, we formulate results in the autonomous convex case. Not to confuse with notations of the previous part of the paper below we consider a time-periodic real analytic Hamiltonian, expressed in action-angle variables, of the form

$$H(\theta, I, t) = h(I) + \varepsilon f(\theta, I_0, t), \tag{4.1}$$

where $\theta = (\theta_0, \theta_1, \dots, \theta_d) \in \mathbb{T}^{d+1}$, $I = (I_0, I_1, \dots, I_d) \in \mathbb{R} \times B^d_{10} \subset \mathbb{R}^{d+1}$, $t \in \mathbb{T}$, and B^d_{10} is the ball of radius 10 around the origin with the norm $\|I\| = \max_j |I_j|$. Assume that the integrable Hamiltonian part $h(I)$ is strictly convex. Consider shift normalization. Convexity implies that $\nabla h : \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ is one to one on its image so shifting I and value of h if necessary we could assume that

$$h(0) = 0 \quad \text{and} \quad \nabla h(0) = 0. \tag{4.2}$$

Let $\Pi_1 : (I_0, \dots, I_d) \to (I_1, \dots, I_d)$ be the natural projection. Assume that h satisfies the following hypothesis: for some $D > 1$

$$D^{-1} < \partial^2_{I_0} h(I_0, I_1, \dots, I_d) < D \quad \text{for each } (I_0, \dots, I_d) \in \mathbb{R} \times B^d_{10}. \tag{4.3}$$

$$B^d_1 \subset \pi_1 \nabla h|_{\{I_0\} \times B^d_{10}} \quad \text{for any } I_0 \in \mathbb{R}.$$

Certainly this condition holds true for $h(I_0, \dots, I_d) = \sum_{j=0}^d I_j^2$.

Let W_ρ be the set of real analytic functions in variables (θ, I_0, t) that are 1-periodic in $(\theta, t) = (\theta_0, \theta_1, \dots, \theta_d, t)$ and bounded on the strip of size $|\Im\ \theta| = (\sum_{j=0}^d |\Im\ \theta_j|^2)^{1/2} \leqslant \rho$, $|\Im\ I_0| \leqslant \rho$, $|\Im\ t| \leqslant \rho$. Consider the standard maximum norm

$$\|f\|_{\rho,R} \equiv \sup_{|\Im\theta, \Im I_0, \Im t| \leqslant \rho,\ |\Re I_0| < R} |f(\theta, I_0, t)|.$$

Theorem 4.1. *Suppose h satisfies* (4.3). *Then for Hamiltonian* (4.1) *there are positive constants c and ρ independent of d and ε so that for $|\varepsilon| > \exp(-d)$ there is a real* *analytic* 1*-periodic in* (θ, t) *function* $f(\theta, I_0, t)$ *depending on* h *with the property*: *for any* $R > 0$ *we have*

$$\|f(\theta, I_0, t)\|_{\rho,R} < 8\pi D(R+1)^2$$

the Hamiltonian $H(\theta, I, t) = h(I) + \varepsilon f(\theta, I_0, t)$ *has an orbit* $\{(\theta, I)(t)\}_{t\in\mathbb{R}}$ *of the Hamiltonian equation* (1.2) *such that*

$$|I(t) - I(0)| \geqslant |I_0(t) - I_0(0)| \geqslant \frac{\varepsilon c}{d}|t| - 1 \quad \textit{for all time}.$$

Remark 1. In condition (4.3, second line) outer radius 10 is not important and can be chosen arbitrary, whereas inner radius 1 has to be at least 1.

Remark 2. In the case of a C^∞ (or Gevrey) perturbation $\varepsilon f(\theta, I_0, t)$ we have certain simplifications. First, the constant c above can be chosen 1. Second, the proof is significantly simpler (all sections starting from Section 6 could be omitted). Third, a diffusing trajectory we construct satisfies $(I_1, \ldots, I_d)(t) = \Lambda$ for some $\Lambda \in \mathbb{Q}^d$ for all time t. In particular, we do not need the Aubry–Mather theory and Mather's variational method.

It seems possible that similar results can be obtained by Mather variational approach [Ma4]. However, it might require substantial efforts.

As we highlighted, the above diffusing trajectories might have a much more flexible behavior than just a linear drift. For example, fix a direction, say I_0, in action space. There are trajectories which exhibit ahead any given random walk along I_0 with time step $K \approx d/\varepsilon$.

Theorem 4.2. *With the notations and set up of Theorem* 4.1 *for any set of numbers* $\{I^k\}_{k\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ *there is a trajectory* $\{(\theta, I)(t)\}_{t\in\mathbb{R}}$ *of the Hamiltonian equation* (1.2) *and an increasing sequence of moments of time* $\{t_k\}_{k\in\mathbb{Z}}$ *with the property*

$$|I_0(t_k) - I^k| < 1 \quad \textit{and} \quad t_k - t_{k-1} < \frac{d(|I^k - I^{k-1}| + 1)}{c\varepsilon} \quad \forall k \in \mathbb{Z}.$$

4.1. Autonomous case

We consider an autonomous real analytic Hamiltonian, expressed in action-angle variables, of the form

$$\tilde{H}(\theta, I) = \tilde{h}(I) + \varepsilon \tilde{f}(\theta, I_0), \tag{4.4}$$

where $\theta = (\theta_0, \theta_1, \ldots, \theta_{d+1}) \in \mathbb{T}^{d+2}$, $I = (I_0, I_1, \ldots, I_{d+1}) \in B_{10}^{d+2} \subset \mathbb{R}^{d+2}$. Again B_r^{d+2} denotes the ball of radius r around the origin. Assume that the integrable Hamiltonian part $\tilde{h}(I)$ is convex, superlinear, and satisfies the same hypothesis (4.3) as h.

Let $\tilde{\pi}_1 : (I_0, \dots, I_{d+1}) \to (I_1, \dots, I_d)$ be the natural projection. Moreover, assume that for some $D' > 1$, e.g. $D' = D$, we have

$$(D')^{-1} < \partial^2_{I_{d+1}} h(I_0, \dots, I_{d+1}) < D' \quad \text{for each } (I_0, \dots, I_{d+1}) \in B^{d+2}_{10}. \tag{4.5}$$

$$B^d_1 \subset \pi_1 \nabla h|_{\{I_0\} \times B^d_{10}} \times \{I_{d+1}\} \quad \text{for any } I_0, I_{d+1} \in \mathbb{R}$$

Certainly this condition holds true for $h(I_0, \dots, I_{d+1}) = \sum_{j=0}^{d+1} I_j^2$.

Since $\tilde{h}$ is real analytic, by Sard's Lemma almost every value of $\tilde{h}$ is regular. Restrict to a compact part of the space $\mathbb{T}^{d+2} \times B^{d+2}_{10}$. Then if E is regular for $\tilde{h} : B^{d+2}_{10} \to \mathbb{R}$ and ε is small we have that E is also regular for $\tilde{H} : \mathbb{T}^{d+2} \times B^{d+2}_{10} \to \mathbb{R}$. By theorem on implicit function if E is a regular value of $\tilde{H}$, then the energy hypersurface $L_E = \tilde{H}^{-1}(E)$ is an analytic manifold, possibly with a boundary. Moreover, for all nearby E's this is also true. We choose positive E, δ and $I(0) \in L_E$ so that we have $|I_{d+1}(0)| > \delta$ and $|\partial\tilde{h}(I(0))/\partial I_{d+1}| > \delta$.

Theorem 4.3. *Suppose $\tilde{h}$ satisfies hypotheses* (4.5). *Fix positive E and δ so that $E > 2\delta^2/D^2$. There are positive constants c and ρ independent of d and ε such that if* $|\varepsilon| > \exp(-d)$, *then there exists a real analytic* 1*-periodic in θ perturbation $\varepsilon \tilde{f}(\theta, I_0)$ depending on h with the property*:

$$\| f(\theta, I_0) \|_{\rho, R} < 8\pi D'(R+1)^2$$

and such that there is an orbit $\{(\theta, I)(t)\}_{t \in \mathbb{R}}$ of the Hamiltonian equation (1.2) *with the Hamiltonian $H(\theta, I) = \tilde{H}(\theta, I) = \tilde{h}(I) + \varepsilon \tilde{f}(\theta, I_0)$ and initial conditions*

$$\tilde{H}((\theta, I)(0)) = E, \quad |I_{d+1}(0)| > \delta, \quad \textit{and} \quad \left| \frac{\partial \tilde{h}(I(0))}{\partial I_{d+1}} \right| > \delta, \tag{4.6}$$

then for any $t > 0$ such that the lower bounds on $|I_{d+1}(\tau)|$ and $\left| \frac{\partial \tilde{H}((\theta, I)(\tau))}{\partial I_{d+1}} \right|$ hold for each $\tau \in [0, t]$ we have

$$|I(t) - I(0)| \geqslant |I_0(t) - I_0(0)| \geqslant \delta \left(\frac{\varepsilon c t}{d} - 1 \right).$$

Remark. Role of radii 1 and 10 in (4.5, second line) is the same as in (4.3, second line) (see Remark 1). Namely, 1 should be fixed and 10 could vary.

Remark 3. Let us verify (4.6) in the case of C^∞-perturbations $\varepsilon f(\theta, I_0)$. In the analytic case similar arguments work. In the C^∞-case a diffusing orbit that we construct satisfies $(I_1, \ldots, I_d)(t) \equiv \Lambda$ for some $\Lambda \in \mathbb{Q}^d$ and all time t. Start verification of (4.6) with the first one. Since $\tilde{H}$ is autonomous, for all time $t \in \mathbb{R}$ we have

$$\tilde{H}\left((\theta, I_0, \Lambda, I_{d+1})(t)\right) = \tilde{h}\left((I_0, \Lambda, I_{d+1})(t)\right) + \varepsilon \tilde{f}\left((\theta, I_0)(t)\right).$$

Recall that $\tilde{h}$ is convex in both I_0 and I_{d+1} and satisfies bound (4.5). Conservation of energy provides a restriction on behavior of $(I_0, I_{d+1})(t)$.

To verify the second in (4.6) note that application of theorem on implicit function shows that for any $t = t > \dfrac{2d}{c\varepsilon}$ by choosing initial conditions appropriately (and this is possible as the proof shows) we can assure that for all $\tau \in [0, t]$ we have $|I_0(\tau) - I_0(0)| > |I_{d+1}(\tau) - I_{d+1}(0)|$. Thus, if $I_{d+1}(0)$ is large enough, lower bound on $|I_{d+1}(\tau)|$ also holds on $[0, t]$.

The third condition in (4.6) follows from shift normalization (4.2), the second condition, and bound (4.5).

In the case of a C^∞ (or Gevrey) perturbation $\varepsilon f(\theta, I_0)$ the constant c above can be chosen 1 and the proof is simpler.

In view of the reduction below, one can state an analog of Theorem 4.2 on a bounded time interval, which is left to the reader.

Now we reduce Theorem 4.3 to Theorem 4.1. The theorem on implicit function condition $|\partial \tilde{h}(I)/\partial I_{d+1}|_{I=(I_0,\ldots,I_{d+1}),\ |I_{d+1}|>\delta}| > \delta > 0$ implies existence of an analytic function $I_{d+1} = I_E(\theta_0, \ldots, \theta_d, \theta_{d+1}, I_0, \ldots, I_d)$ such that

$$\tilde{H}(\theta_0, \ldots, \theta_d, \theta_{d+1}, I_0, \ldots, I_d, I_E(\theta, I_0, \ldots, I_d)) \equiv E.$$

Trajectories of the Hamiltonian $H(\theta_0, \ldots, \theta_d, I_0, \ldots, I_d, t)$ given below coincide with projected trajectories of $\tilde{H}(\theta_0, \ldots, \theta_{d+1}, I_0, \ldots, I_d, I_E(\theta, I_0, \ldots, I_d))$ given above with t being the projection of θ_{d+1}. Another description is the following: we omit I_{d+1} and reparametrize time so that period is 1 and speed of θ_{d+1} becomes identically 1. Here is the formula for the differential of H we need

$$\begin{aligned} dH(\theta_0, \ldots, \theta_d, I_0, \ldots, I_d, \theta_{d+1}) \\ = \frac{\partial H}{\partial \theta_0} d\theta_0 + \cdots + \frac{\partial H}{\partial \theta_d} d\theta_d \\ + \frac{\partial H}{\partial I_0} dI_0 + \cdots + \frac{\partial H}{\partial I_d} dI_d + \frac{\partial H}{\partial \theta_{d+1}} d\theta_{d+1} \end{aligned}$$

$$= \left(\frac{\partial \tilde{H}}{\partial I_{d+1}}\right)^{-1} \left(\frac{\partial \tilde{H}}{\partial \theta_0} d\theta_0 + \cdots + \frac{\partial \tilde{H}}{\partial \theta_d} d\theta_d \right.$$

$$\left. + \frac{\partial \tilde{H}}{\partial I_0} dI_0 + \cdots + \frac{\partial \tilde{H}}{\partial I_d} dI_d \ + \frac{\partial \tilde{H}}{\partial \theta_{d+1}} d\theta_{d+1}\right). \tag{4.7}$$

Put $H(0) = 0$. Direct verification shows that H is a well-defined function on $\mathbb{T}^{d+1} \times \mathbb{R}^{d+1} \times \mathbb{T}$. This can be done as long as we could apply theorem on implicit function to express I_{d+1} in terms of the other variables and velocity of θ_{d+1} bounded away from zero. The fact that trajectories of $\tilde{H}$ coincide with projected trajectories of H follows from the form of Hamiltonian equations (1.2). Note also that the time reparametrization proposed above has ratio of time intervals for $\tilde{H}$ and H bounded by $\min |I_{d+1}|$ or its inverse.

Bring constructed H to the form (4.1). Fix E as above. By formula (4.7)

$$dH = (\partial_{I_{d+1}} \tilde{H})^{-1} \ d\tilde{H} = (\partial_{I_{d+1}} \tilde{h})^{-1} \ (d\tilde{h} + \varepsilon \ d\tilde{f}).$$

By theorem on implicit function there is an analytic function $I_E(I_0, \Lambda) = I_{E,\Lambda}(I_0)$, $\Lambda \in B_1^d$ such that $h_\Lambda(I_0, \Lambda, I_{E,\Lambda}(I_0)) = E$. Using (4.7) with $\tilde{H}(I) = \tilde{h}(I)$ express the function $H(I) = h_E(I)$. Direct calculation shows that if $\tilde{h}$ satisfies conditions (4.2) and (4.5) for some $\tilde{D}$, then $h_E(I)$ also satisfies these conditions with $D^{\text{new}} = D/\delta$ and defined above $H(\theta, I)$ can be written in the form

$$H(\theta_0, \ldots, \theta_d, I_0, \ldots, I_d, \theta_{d+1}) = h_E(I_0, \ldots, I_d) + \varepsilon f_E(\theta_0, \ldots, \theta_d, I_0, \theta_{d+1}).$$

Moreover, upper bounds on norms of f_E above can be given in terms of upper bounds of norms of $\tilde{f}$ and the other constants. This shows that Theorem 4.1 implies Theorem 4.3.

The rest of the paper is organized as follows. In the next section we describe three main steps of the proof of Theorem 4.1. It turns out that for C^∞-perturbations we need only the first two. These two steps use theory of normally hyperbolic manifolds and the method of generating functions. We realize these steps in Sections 5.1 and 5.2, respectively. The final third step requires involved arguments using Aubry–Mather theory and Mather's variational method. This step has an independent interest, since to the best of our knowledge it is the first detailed exposition of Mather's variational method of changing Lagrangians for twist maps. Mention also that there exist partially complete notes [Ma3] devoted to mechanical systems on $\mathbb{T}^2$.

5. Construction of "diffusing" perturbation

The ideas of the construction are the following.

Step 1: *Construction of a normally hyperbolic invariant cylinder* $\mathbb{A}_\Lambda \subset \mathbb{T}^d \times \mathbb{R} \times \mathbb{T}$. First, we choose a long straight closed line S_Λ in $\mathbb{T}^d \times \mathbb{T}$ of length $K \approx d/\varepsilon$. Then we construct a C^∞-smooth perturbation of $(2d+3)$-dimensional Hamiltonian system (4.1) in the phase space $\mathbb{T}^{d+1} \times \mathbb{R}^{d+1} \times \mathbb{T}$, which is time 1 periodic and has a hyperbolic invariant three-dimensional cylinder $\mathbb{A}_\Lambda \simeq \mathbb{T} \times \mathbb{R} \times \mathbb{T}_* \ni \{(\theta_0, I_0, \theta_*)\}$ embedded into the phase space. The first two components of $\mathbb{A}_\Lambda$ are the first pair of action-angle variables $(\theta_0, I_0) \in \mathbb{T} \times \mathbb{R} = \mathbb{A}_0$ and the last coordinate is a coordinate on the closed line S_Λ. With unit velocity period along S_Λ is K. Then the time-K-map restricted to this cylinder is an exact symplectic map of the cylinder $\Phi : \mathbb{A}_0 \to \mathbb{A}_0$. Usually Φ is called *a Poincare return map*. We construct $\mathbb{A}_\Lambda$ so that in its neighborhood dynamics in transversal directions is sufficiently hyperbolic. Namely, the Poincare map Φ either expands or contracts along transversal to $\mathbb{A}_0$ directions strongly enough. The classical result of Sacker–Fenichel (see Section 10) says $\mathbb{A}_\Lambda$ persists under small perturbations of underlying differential equations, $\mathbb{A}_\Lambda$ just gets slightly deformed. We shall prove Theorem 4.1 for C^∞-perturbations first and then use persistence of $\mathbb{A}_\Lambda$ to approximate C^∞-perturbations by analytic ones.

Step 2: *Ballistic trajectories of twist maps without invariant curves.* It turns out that one can find a C^∞-perturbation $\varepsilon f(\theta, I_0, t)$ of the restricted system of period $K \approx d/\varepsilon \gg 1$ so that the Poincare map Φ can be chosen more or less arbitrary by adjusting the perturbation $\varepsilon f(\theta, I_0, t)$. At this point we reduce investigation to the study of exact symplectic maps of $\mathbb{A}$. We shall consider a subclass of those called *exact area-preserving monotone twist* (EAPT) maps or simply *twist* maps of $\mathbb{A}$. For this class of maps we have Aubry–Mather theory and powerful variational methods due to Mather [Ma1,Ma3] available.

Now we need to construct twist maps with a ballistic trajectory to prove Theorem 4.1 or a "random walk" trajectory to prove Theorem 4.2, respectively. We start by working out the construction for the unperturbed Hamiltonian $h(I) = \sum_{j=0}^d I_j^2$ of the standard type. Then using *generation functions technique* we construct a C^∞-perturbation $\varepsilon f(\theta, I_0, t)$ such that the Poincare map Φ has the following standard form:

$$\Phi^* : (\theta_0, I_0) \to \left(\theta_0 + K\ I_0 + \sin 2\pi\theta_0, I_0 + \frac{1}{K}\ \sin 2\pi\theta_0\right). \tag{5.1}$$

Note that $(\Phi^*)^n(\frac{1}{4}, 0) = (\frac{1}{4}, \frac{n}{K})$ for all n and its generating function is $h^*(\theta, \theta') = \frac{1}{2K}(\theta - \theta')^2 - \frac{1}{2\pi K}\cos 2\pi\theta$. This remark suggested to author by Mather and Neishtadt. However, for a general convex $h(I)$ and an analytic perturbation such an exact form of Φ is problematic. For a general convex $h(I)$ satisfying (4.3) we shall need to modify the form (5.1) and follow the same strategy as for $h(I) = \sum_{j=0}^d I_j^2$. To extend the proof to treat analytic perturbations we approximate the above C^∞-perturbation $\varepsilon f(\theta, I_0, t)$ with an analytic one $\varepsilon f'(\theta, I_0, t)$. Persistence of $\mathbb{A}_\Lambda$ guarantee it survives and dynamics of deformed $\mathbb{A}'_\Lambda$ is close to the one on $\mathbb{A}_\Lambda$ (see Section 10). In particular, it means that a twist map Φ', the Poincare map of $\mathbb{A}'_\Lambda$, has no invariant curves. The simplified key result at the next step is:

Step 3: *If a* C^1 *twist map is* C^1*-close to the twist map* (5.1), *then there exist ballistic trajectories* $\Phi^n(\theta^0, I^0) = (\theta^n, I^n)$, *i.e. such that* $c_1 n < I^n < c_1 n + c_2$ *for some* $c_1,\ c_2 > 0$ *and all* $n \in \mathbb{Z}$. To construct those ballistic and "random walk" orbits we also apply Mather's variational approach [Ma3].[2] More exactly, we use his method of changing Lagrangians. This approach analyzes a family barrier function which is associated with the corresponding twist maps. We also use this as an opportunity to describe this powerful approach, since it is not available in the literature.

5.1. Construction of a normally hyperbolic invariant three-dimensional cylinder in the $(2d+3)$*-dimensional phase space*

The idea of the construction is as follows. Choose a large positive integer $K \approx de^d > \frac{d}{\varepsilon}$ and a rational direction $\Lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{Q}^d \subset \mathbb{R}^d$ such that $K\Lambda \in \mathbb{Z}^d$ and, therefore,

$$S_\Lambda = \{\Lambda t \ (\mathrm{mod}\ 1) \in \mathbb{T}^d :\ 0 \leqslant t \leqslant K\}$$

is a closed curve in $\mathbb{T}^d$. Choose also Λ so that S_Λ is reasonably equidistributed in $\mathbb{T}^d$. Consider dynamics projected onto coordinates: $(\theta_1, \dots, \theta_d, I_1, \dots, I_d, t)$.

For the initial values $(\theta_1, \dots, \theta_d) = 0,\ (I_1, \dots, I_d) = \Lambda$ the unperturbed Hamiltonian $h(I)$ has a periodic orbit $S_\Lambda \times \{\Lambda\} \subset \mathbb{T}^d \times \mathbb{R}^d$. We design a perturbation $\varepsilon f(\theta, I_0, t)$ so that the perturbed Hamiltonian H has the same property. Namely, it has the family of trajectories $\{(\theta, I)(t)\}_{t\in\mathbb{R}}$ with the property: independently of values of (θ_0, I_0) we have

$$\begin{aligned} &\text{if } (I_1(0), \dots, I_d(0)) = \Lambda \text{ and } (\theta_1(0), \dots, \theta_d(0)) = 0, \text{ then} \\ &(I_1(t), \dots, I_d(t)) = \Lambda \text{ and } (\theta_1(t), \dots, \theta_d(t)) \in S_\Lambda \text{ for all time.} \end{aligned} \tag{5.2}$$

As the result, we obtain that dynamics of the first symplectic pair $\{(\theta_0, I_0)(t)\}$ under the above condition is described by a time-periodic system of large period $K = |S_\Lambda|/|\Lambda| > \frac{d}{\varepsilon}$. Moreover, the system is ε-close to integrable (see Eq. (5.5)). However, *the large period* allows the Poincare map for the period being *far from integrable*.

The standard remark is the time-K-self-map of the cylinder $(\theta_0, I_0) \in \mathbb{A}_0$, denoted by $\Phi : \mathbb{A}_0 \to \mathbb{A}_0$, is an exact symplectic map. Here construction of a required perturbation:

Lemma 5.1. *Let* d *be a large positive integer. There are positive* C *and* δ *independent of* d *such that one can find* $\Lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{Q}^d \cap B^d_{10} \subset \mathbb{R}^d$ *and* $1 \leqslant |\lambda_j| \leqslant 2,\ 1 \leqslant j \leqslant d$ *with the property*: $S_\Lambda \subset \mathbb{T}^d$ *is a closed curve of length* $Cd^{3/2}\exp(d) < L < 2Cd^{3/2}\exp(d)$ *so that its* δ*-neighborhood is non-self-intersecting. Moreover, scaling* Λ *allows to choose it in a ball around the origin of any radius.*

[2] First fundamental step of constructing diffusing orbits using variational method with constrains was done in [Ma1].

Proof. Let $e_1, \dots, e_d$ be the corresponding basis of coordinates $(\theta_1, \dots, \theta_d)$. For any pair of rational vectors v and u denote $L_{u,v}$ (resp. $T^2_{u,v}$) the plane span by them in $\mathbb{R}^d$ (resp. its natural embedding in $\mathbb{T}^d$). Note that $T^2_{u,v}$ is diffeomorphic to the 2-torus. We shall choose a rational Λ according to the following inductive algorithm.

Let the first two components λ_1 and λ_2 be rational and such that the closed line directed along (λ_1, λ_2) is 2δ-dense, but not δ-dense in $T^2_{e_1,e_2}$. It is easy to find such (λ_1, λ_2). Denote $\Lambda_2 = (\lambda_1, \lambda_2)$ and the corresponding vector u_2. By the construction length of this line is between $(2\delta)^{-1}$ and δ^{-1}.

Let λ_3 be rational such that the line directed along $u_2 + \lambda_3 e_3$ is 2δ-dense, but not δ-dense in $T^2_{u_2,e_3}$. Denote $\Lambda_3 = (\lambda_1, \lambda_2, \lambda_3)$ and the corresponding vector u_3. By the construction length of this line is between $(2\delta)^{-2}$ and δ^{-2} and so on.

After $d-1$ steps we get such a λ_d that the line directed along $u_{d-1} + \lambda_d e_d$ is 2δ-dense, but not δ-dense in $T^2_{u_{d-1},e_d}$. Denote $\Lambda = (\lambda_1, \dots, \lambda_d)$ and the corresponding vector u_d. By the construction length of this line, which we denote S_Λ, is between $(2\delta)^{1-d}$ and δ^{1-d} and *its δ-neighborhood is non-self-intersecting in $\mathbb{T}^d$*. Following the same procedure one could choose S_Λ of required length. This completes the proof. □

We fix Λ and the corresponding periodic orbit $S_\Lambda \subset \mathbb{T}^d$ satisfying non-self-intersection property of the above lemma. Now we construct a C^∞ time 1-periodic perturbation $\varepsilon f(\theta, I_0, t)$ such that the invariance property (5.2) holds true. Construction consists of two stages in the case of C^∞ perturbations and of three stages in the case of real analytic perturbations.

Change the basis on $\mathbb{T}^d$ so that the first coordinate vector becomes collinear Λ. Choose the other components of the basis in an arbitrary way. Denote new coordinates on $\mathbb{T}^d$ by $(\dot\theta_1^\Lambda, \dots, \theta_d^\Lambda)$. In the new coordinates S_Λ is given by $(\theta_2^\Lambda, \dots, \theta_d^\Lambda) = 0$ and motion is given by $\dot\theta_1^\Lambda = |\Lambda|$.

Step 1: *Keeping dynamics along S_Λ invariant.* Consider a C^∞-smooth 1-periodic function of the form

$$\begin{aligned} f\varepsilon(\theta_0, \theta_1, \dots, \theta_d, I_0, t) &= f\varepsilon(\theta_0, \theta_1^\Lambda, \dots, \theta_d^\Lambda, I_0, t) \\ &= f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ \ f_2(\theta_2^\Lambda, \dots, \theta_d^\Lambda, t) \end{aligned} \tag{5.3}$$

with the following properties: Fix a small number τ, e.g. $\tau = 10^{-1}$. Suppose for all time $\theta_0 \in \mathbb{T},\ I_0 \in \mathbb{R},\ \theta_1^\Lambda \in S_\Lambda, t \in \mathbb{T}$ we have that $f_1(\theta_0, I_0, \theta_1^\Lambda, t)$ is strictly positive for $t \in [4\tau, 1-4\tau] \cup [0, \tau] \cup [1-\tau, 1]$ and $(f_1(\theta_0, I_0, \theta_1^\Lambda, t) - 1)$ is strictly positive for $t \in [0, 0.5\tau] \cup [1-0.5\tau, 1]$. Also suppose that the following conditions hold:

$$\begin{aligned} &f_1(\theta_0, I_0, \theta_1^\Lambda, t) \equiv 0, \quad t \ (\mathrm{mod}\ 1) \in [\tau, 4\tau] \cup [1-4\tau, 1-\tau], \\ &\partial_{\theta_1^\Lambda} f_1(\theta_0, I_0, \theta_1^\Lambda, t) \equiv 0, \quad t \ (\mathrm{mod}\ 1) \in [0, \tau] \cup [1-\tau, 1], \\ &f_2(0, t) \equiv 0, \quad t \ (\mathrm{mod}\ 1) \in [3\tau, 1-3\tau], \\ &f_2(0, t) \equiv 1, \quad t \ (\mathrm{mod}\ 1) \in [0, 2\tau] \cup [1-2\tau, 1]. \end{aligned} \tag{5.4}$$

Moreover, for each I_0 the function $f(\theta_0\theta_1^\Lambda, \ldots, \theta_d^\Lambda, I_0, t)$ has a non-degenerate local maximum at any point of the form $(|\Lambda|\theta_1^\Lambda - t, \ldots, \theta_d^\Lambda) = 0$. At this local maxima consider the Hessian matrix $\partial^2_{\theta_i^\Lambda\theta_j^\Lambda} f_2(0,t)$. Since maxima is non-degenerate, all eigenvalues of the Hessian are strictly negative. Denote by λ_t the maximal eigenvalue and $\lambda^* = \max_{0\leqslant t\leqslant 1} \lambda_t < 0$.

The Hamiltonian equations have the form

$$\begin{cases} \dot\theta_0 = \partial_{I_0} h(I) + \varepsilon\partial_{I_0} f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ f_2(\theta_2^\Lambda, \ldots, \theta_d^\Lambda, t), \\ \dot\theta_j = \partial_{I_j} h(I), \quad 1\leqslant j\leqslant d, \\ \dot I_0 = -\varepsilon\partial_{\theta_0} f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ f_2(\theta_2^\Lambda, \ldots, \theta_d^\Lambda, t), \\ \dot I_1^\Lambda = -\varepsilon\partial_{\theta_1^\Lambda} f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ f_2(\theta_2^\Lambda, \ldots, \theta_d^\Lambda, t), \\ \dot I_j^\Lambda = -\varepsilon f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ \partial_{\theta_j^\Lambda} f_2(\theta_2^\Lambda, \ldots, \theta_d^\Lambda, t), \quad 2\leqslant j\leqslant d. \end{cases} \tag{5.5}$$

Check that by the construction for all time the right-hand side of last d equations is identically zero. Start with the last $(d-1)$ equations. For all $t \pmod 1 \notin [2\tau, 3\tau]\cup[1-3\tau, 1-2\tau]$ the function f_2 is a constant and all its spacial partial derivatives vanish. In the remaining part of the time interval f_1 vanishes. Now check the equation for $\dot I_1^\Lambda$. Either $\partial_{\theta_1^\Lambda} f_1(\cdot, I_0, t)$ or $f_2(\cdot, t)$ vanish on the whole period in t by the choice above.

Choose speed of motion on S_Λ so that period is integer and, therefore, on the natural lift $\tilde S_\Lambda \subset \mathbb{T}^d \times \mathbb{T}$ motion is periodic. Now we make sure that the cylinder $\mathbb{A}_\Lambda$ is *normally hyperbolic*. Recall $\mathbb{A}_\Lambda = \mathbb{A}_0 \times S''_\Lambda \subset \mathbb{T}\times\mathbb{R}\times\mathbb{T}^d\times\mathbb{T}\times\mathbb{R}^d$ is an invariant three-dimensional cylinder. We would like to prove it is normally hyperbolic. It suffices to prove the following fact. Fix any point on $(\theta_0, I_0)\times\{0,\Lambda\}\in\mathbb{A}_0\times\mathbb{T}$ as an initial value and consider that the projected time-K-map Poincare map

$$F_{(\theta_0,I_0)} : (\theta_2^\Lambda, \ldots, \theta_d^\Lambda, I_2^\Lambda, \ldots, I_d^\Lambda)(0) \to (\theta_2^\Lambda, \ldots, \theta_d^\Lambda, I_2^\Lambda, \ldots, I_d^\Lambda)(K) \tag{5.6}$$

of the flow (5.5) in a neighborhood of S_Λ. Namely, we integrate properties of the flow (5.5) when initial conditions are close to $\mathbb{A}_0$. The standard fact about Poincare return maps of a Hamiltonian system says that eigenvalues of the linearization of $F_{(\theta_0,I_0)}$ of a symplectic map come in pairs: $\mu_j,\ \mu_j^{-1}$. It follows from existence of invariant non-degenerate 2-form $\sum d\theta_j^\Lambda\wedge dI_j^\Lambda$. Therefore, to prove that $F_{(\theta_0,I_0)}$ is hyperbolic, it is enough to present $(d-1)$ eigenvalues whose absolute value are different from 1 and estimate how much different.

By the construction for $t \pmod 1 \in [\tau, 1-\tau]$ the perturbation vanishes and dynamics is integrable. Outside of this time interval, f has a controllably non-degenerate maximum. More exactly, in a neighborhood of $(|\Lambda|\theta_1^\Lambda - t, \ldots, \theta_d^\Lambda) = 0$ the maximum value occurs on this set and the Hessian in transversal directions in strictly negative. It is coherent with conditions (5.4). Since maxima eigenvalues of the Hessian $\partial^2_{\theta_i\theta_j} f_2$ are strictly negative, $\dot I_j^\Lambda \geqslant \varepsilon\lambda^*\|(\theta_2^\Lambda, \ldots, \theta_d^\Lambda)\|$ for $|t-n| < 0.5\tau$. This shows that absolute values of eigenvalues of the linearization of $F_{(\theta_0,I_0)}$ if exceed 1 are bounded from below by $\exp(\varepsilon\tau\lambda^* K)$. Choose $K = Cd\exp(d) \geqslant Cd/\varepsilon$ with large enough C. Then we get $\exp(C\tau\lambda^* d) \geqslant (K+1)^2$.

5.2. Construction of a desired twist map on the cylinder $\mathbb{A}_\Lambda$ with ballistic trajectories

Before we state an appropriate suspension lemma we need to modify the form of map (5.1) we would like to construct. We start by construction in the case $h(I) = \sum_{j=0}^d I_j^2$. Later we extend it to the case of general convex h satisfying (4.3).

Consider a sequence of standard EAPT maps:

$$\begin{aligned}\Phi_j : (\theta_0, I_0) &\mapsto (\theta_0', I_0')\\ &= \left(\theta_0 + I_0', I_0 + \frac{2}{K}\sin 2\pi\left\{\theta_0 - \frac{j(j-1)}{2K}\right\}\right),\end{aligned} \tag{5.7}$$

where $1 \leqslant j \leqslant K$. Note that each of these maps is an EAPT map. Denote

$$\tilde{\Phi} = \Phi_K \circ \Phi_{K-1} \circ \cdots \circ \Phi_1. \tag{5.8}$$

Since this is a composition of analytic EAPT maps, $\tilde{\Phi}$ is an analytic EAPT map too. Note that

$$\Phi_j\left(\frac{1}{4} + \frac{j(j-1)}{K}, \frac{2(j-1)}{K}\right) = \left(\frac{1}{4} + \frac{(j+1)j}{K}, \frac{2j}{K}\right)$$

for each $j = 1, \ldots, K$. Therefore, $\tilde{\Phi}\left(\frac{1}{4}, 0\right) = \left(\frac{1}{4}, 2\right)$ and

$$\tilde{\Phi}^m\left(\frac{1}{4}, 0\right) = \left(\frac{1}{4}, 2m\right) \quad \text{for each} \quad m \in \mathbb{Z}.$$

Denote C^d-norm of a function f by $\|f\|_{C^d}$, i.e. supremum of f and all its partial derivatives of order up to d in the corresponding periodic variables.

Proposition 5.2. *Assume $h(I) = \sum_{j=0}^d I_j^2$. With the notations above there are C^∞ smooth functions $f_1(\theta_0, I_0, \theta_1^\Lambda, t)$ and $f_2(\theta_2^\Lambda, \ldots, \theta_d^\Lambda, t)$ satisfying conditions* (5.3)–(5.4) *such that the corresponding Poincare return map $\tilde{\Phi}(\theta_0, I_0)$ has the form* (5.8). *The perturbation $\varepsilon f(\theta_0, \theta_1, \ldots, \theta_d, I_0, t)$ of $h(I)$ in* (4.1), *given by* (5.3) *via the product of functions f_1 and f_2. Moreover, with respect to the original coordinate system for any fixed I_0 we have*

$$\begin{aligned}8D\pi 8^d(\delta\tau^2)^{-d}\exp(-d)(I_0+1)^2 &\leqslant \tilde{C}^2(\delta\tau^2)^{-d}\exp(-d)(I_0^2 + \tilde{C}),\\ 8D\pi 4^d\tau^{-4d}\exp(-d)(I_0+1)^2 &\leqslant \tilde{C}\tau^{-2d}(I_0^2+\tilde{C}), \quad 2^d\delta^{-d} \leqslant \tilde{C}\delta^{-d}\end{aligned} \tag{5.9}$$

where $D = 2$.

Corollary 5.3. *With the notations above and notions introduced in Section* 10 *the three-dimensional cylinder* $\mathbb{A}_\Lambda \subset \mathbb{T}^{d+1} \times \mathbb{R}^{d+1} \times \mathbb{T}$ *is* 1*-normally hyperbolic with respect to Hamiltonian flow of* H.

Proof of the Corollary. Consider the Poincare time-K-map of Eq. (5.5) in a neighborhood of the invariant cylinder $\mathbb{A}_0$. Since Φ has the form (5.1), maximal expansion of Φ restricted on the cylinder is $(K+1)$. While minimal expansion along transversal directions is $\exp(C\tau\lambda^* d) \geqslant (K+1)^2$. By the remark above each expanding direction has a conjugate contracting direction so that exponents of corresponding expansion and contraction are inverse to one another. □

Proof of the Proposition. Note that the conditions (5.4) by the construction imply (5.2) and, therefore, $(\theta_2^\Lambda, \ldots, \theta_d^\Lambda) = 0$ holds for all time. Also when $\dot{I}_0$ is non-zero, then $f_2 \equiv 1$. Therefore, we need only determine the function f_1 from Eq. (5.5) or, equivalently, its Hamiltonian $\{H(\cdot, t)\}_{t \in \mathbb{T}}$ near $\mathbb{A}_\Lambda$. We shall use the standard suspension method based on generating functions (see e.g. [Arn,Go]). It suffices to determine time-s-map $\Phi^s(\theta_0, I_0)$ for each $0 \leqslant s \leqslant K$. Each time-$s$-map, restricted to $\mathbb{A}_0$, is a symplectic map of the cylinder $\mathbb{A}_0$ and can be given by a generating function. More exactly, given a function $h(\tilde{\theta}_0, \tilde{\theta}_0')$ of two variables $\tilde{\theta}_0, \tilde{\theta}_0' \in \mathbb{R}$ one can define an exact symplectic map of the cylinder $\mathbb{A} = \mathbb{T} \times \mathbb{R}$ using Eq. (6.2). So we shall define a C^∞ smooth 1-parameter family of generating functions $\{h_s(\tilde{\theta}_0, \tilde{\theta}_0')\}_{0 \leqslant s \leqslant K}$ so that for each $0 \leqslant s \leqslant K$ it defines time-s-map $\Phi^s(\theta_0, I_0)$.

First we construct each map from set (5.7) whose composition is (5.8). Each of them will be chosen from a particular family of generating functions $\{h_s(\tilde{\theta}_0, \tilde{\theta}_0', \alpha)\}_{0 \leqslant \alpha < 1}$. This family is defined as follows. Pick a value $\alpha \in [0, 1)$ and define a C^∞-smooth function $\rho : [0, 1] \to \mathbb{R}_+$ as follows $\rho(1) = 2/K$ and $\rho'(s) = 0$ for all $t \pmod 1 \in [0, 1 - 0.5\tau]$, where K is the same as in (5.7). Then we have the family

$$h_s(\tilde{\theta}_0, \tilde{\theta}_0', \alpha) = \frac{1}{2s}(\tilde{\theta}_0 - \tilde{\theta}_0')^2 - \frac{\rho(s)}{2\pi s} \cos 2\pi(\tilde{\theta}_0 - \alpha). \tag{5.10}$$

Direct calculation shows that the corresponding family of C^∞ EAPT maps $\{\Phi^s\}_s$ is

$$\begin{cases} \tilde{\theta}_0' = \tilde{\theta}_0 + sI_0 + \rho(s) \sin 2\pi(\tilde{\theta}_0 - \alpha), \\ I_0' = I_0 + \frac{\rho(s)}{s} \sin 2\pi(\tilde{\theta}_0 - \alpha). \end{cases} \tag{5.11}$$

To extract from (5.11) the Hamiltonian $\{H'(\theta_0, I_0, s, t)\}_{0 \leqslant s \leqslant K}$, $\theta_1^\Lambda = s$ or, equivalently, of $\varepsilon f_1(\theta_0, I_0, s, t)$, we use the standard fact from symplectic geometry (see e.g. [Mc-S, Proposition 9.18]; [Go, Theorem 58.9]). First of all note that for every $0 \leqslant s \leqslant 1$ map (5.11) generates a Hamiltonian isotopy $\{\Phi^s\}_{s=0}^1$. This family generates the time-periodic Hamiltonian function $H'(\theta_0, I_0, s)$ as follows.

According to the standard formula in symplectic geometry the corresponding Hamiltonian can be expressed as follows. Denote by $A_s(\theta_0, I_0)$ action along the trajectory starting at (θ_0, I_0) of length s in time,[3] by $\{X_s\}_{s=0}^1$ the vector field generated by the isotopy $\{\Phi^s\}_{s=0}^1$, and $i_{X_s}\alpha$ the interior derivative of a one form α along X_s. Then

$$H'(\theta_0, I_0, s) = i_{X_s}(I_0 d\theta_0) - \left((\Phi^s)^{-1}\right)^* \frac{d}{ds}(A_s). \tag{5.12}$$

We need to compute this expression in order to see what class of perturbations $\varepsilon f(\theta, I_0, t)$ in (4.1) realizes a Poincare return map of the required form.

Calculation of the action functional $A_s(\theta_0, I_0)$: Denote the lift $A_s : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, $A_s(\theta_0, I_0)$ to $A_s : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $A_s(\tilde\theta_0, I_0)$ by the same A_s for brevity. By construction the isotopy $\{\Phi^s\}_s$ gives the family of the generating functions $\{h_s\}_s$ given by (5.10).[4] By one of definitions $h_s(\tilde\theta_0, \tilde\theta_0', \alpha)$ is the minimal action to get from $\tilde\theta_0$ to $\tilde\theta_0'$ in time s. The form of the EAPT map (5.11) gives that the initial condition $(\tilde\theta_0, I_0)$ mapped into $(\tilde\theta_0', I_0')$ for some I_0' has to be

$$\tilde\theta_0' - \tilde\theta_0 = sI_0 + \rho(s)\sin 2\pi(\tilde\theta_0 - \alpha).$$

Therefore,

$$\begin{aligned} h_s(\tilde\theta_0, \tilde\theta_0', \alpha) &= A_s(\tilde\theta_0, I_0) \\ &= \frac{\left(sI_0 + \rho(s)\sin 2\pi(\tilde\theta_0 - \alpha)\right)^2}{2s} - \frac{\rho(s)}{2\pi s}\cos 2\pi(\tilde\theta_0 - \alpha), \end{aligned}$$

where $\tilde\theta_0'$ is θ-coordinate of the position of the trajectory starting at $(\tilde\theta_0, I_0)$ at time s.

Calculation of the vector field $X_s(\tilde\theta_0, I_0)$: It is an exercise to show that

$$X_s(\tilde\theta_0, I_0) = \frac{d\Phi^s}{ds}\left((\Phi^s)^{-1}(\tilde\theta_0, I_0)\right).$$

Inversion of (5.11) gives that $(\Phi^s)^{-1}$

$$\begin{cases} \tilde\theta_0' = \tilde\theta_0 - sI_0, \\ I_0' = I_0 - \frac{\rho(s)}{s}\sin 2\pi(\tilde\theta_0 - \alpha - sI_0). \end{cases} \tag{5.13}$$

[3] We shall relate it to the generating function h_s below.

[4] An isotopy is a smooth 1-parameter family of diffeomorphisms.

Differentiation of (5.11) gives

$$\frac{d\Phi^s}{ds}(\tilde\theta_0, I_0) = \left(I_0 + \rho'(s)\sin 2\pi(\tilde\theta_0 - \alpha), \left(\frac{\rho(s)}{s}\right)' \sin 2\pi(\tilde\theta_0 - \alpha)\right).$$

Substitute

$$X_s(\tilde\theta_0, I_0) = \begin{pmatrix} I_0 + (\rho'(s) - \rho(s)/s)\ \sin 2\pi(\tilde\theta_0 - \alpha - sI_0) \\ \left(\frac{\rho(s)}{s}\right)' \sin 2\pi(\tilde\theta_0 - \alpha - sI_0) \end{pmatrix}.$$

It implies

$$i_{X_s}(I_0 d\tilde\theta_0) = I_0(I_0 + (\rho'(s) - \rho(s)/s)\sin 2\pi(\tilde\theta_0 - \alpha - sI_0)).$$

Calculate differential of $A_s(\tilde\theta_0, I_0)$ given above

$$\begin{aligned}\frac{d}{ds}A_s(\tilde\theta_0, I_0) = &\left(\frac{I_0^2}{2} + \rho'(s) I_0 \sin 2\pi(\tilde\theta_0 - \alpha)\right.\\ &\left. + \frac{\rho(s)\rho'(s)}{2s}\sin^2 2\pi(\tilde\theta_0 - \alpha) + \left(\frac{\rho(s)}{2\pi s}\right)' \cos 2\pi(\tilde\theta_0 - \alpha)\right) ds\\ &+ (sI_0 + \rho(s)\sin 2\pi(\tilde\theta_0 - \alpha))\, dI_0 + \left(2\pi\rho(s) I_0 \cos 2\pi(\tilde\theta_0 - \alpha) + \right.\\ &\left.\frac{\pi\rho^2(s)}{2s}\sin 4\pi(\tilde\theta_0 - \alpha) + \frac{\rho(s)}{s}\sin 2\pi(\tilde\theta_0 - \alpha)\right) d\tilde\theta_0,\end{aligned} \tag{5.14}$$

where differentials are given by $d\tilde\theta_0 = (I_0 + \rho'(s)\sin 2\pi(\tilde\theta_0 - \alpha))\,ds$ and $dI_0 = (\rho(s)/s)'\ \sin 2\pi(\tilde\theta_0 - \alpha)\,ds$. Applying $((\Phi^s)^{-1})^*$ to (5.14) and substituting dI_0 and $d\tilde\theta_0$ and applying $(\Phi^s)^{-1}$ we get

$$\begin{aligned}H'(\theta_0, I_0, s) = &\frac{I_0^2}{2} - 2\pi\rho(s) I_0^2 \cos 2\pi(\theta_0 - \alpha - sI_0) + \\ &I_0 g_1(s, \theta_0 - \alpha - sI_0) + g_2(s, \theta_0 - \alpha - sI_0),\end{aligned} \tag{5.15}$$

As straightforward calculation shows g_1 and g_2 are uniformly bounded by $(2\pi + 2)/K$ and $4/K$ respectively. It implies that $2\pi I_0^2 + (2\pi+2)I_0 + 4 < 4\pi(I_0+1)^2$, where g_1 and g_2 are 1-periodic in both variables. By the construction the time-1-map corresponding to H^* has the form (5.13). It is the same as (5.7) with $\alpha = \dfrac{j(j-1)}{K}$. To include the constructed time-1-maps into the time-K-map we recall that the Hamiltonian H

depends on $\theta_1^\Lambda = s$ which has period K. So we introduce the function $\alpha\colon [0, K] \to \mathbb{R}$ as follows:

$$\alpha(j+\tau) = \frac{j(j-1)}{2K} \quad \text{for each integer } j \text{ with } 1 \leqslant j \leqslant K$$
$$\text{and } \alpha'(s) = 0 \text{ if } \ s \,(\mathrm{mod}\,1) \ \notin [0.5-\tau, 0.5+\tau]. \tag{5.16}$$

Now we get the following Hamiltonian:

$$H(\theta_0, I_0, s, t) = \frac{I_0^2}{2} - 2\pi\rho(s) I_0^2 \cos 2\pi(\theta_0 - \alpha(s) - tI_0) +$$
$$I_0 g_1(s, \theta_0 - \alpha(s) - tI_0) + g_2(s, \theta_0 - \alpha(s) - tI_0), \tag{5.17}$$

where g_1 and g_2 as above. This implies 1-periodicity in θ_0 and K-periodicity in $s = \theta_1^\Lambda$. To see 1-periodicity in t choose time parametrization by s of θ_1^Λ. Then when α is fixed we have $g_1, \ g_2$ are period 1 in the first variable s, while the second is fixed.

In order to estimate C^d-norm of the corresponding perturbations εf it suffices to bound C^d norm of εf_1 with respect to the original coordinate system $(\theta_0, \theta_1, \ldots, \theta_d)$. In this coordinate system f_1 is periodic in θ_1^Λ of period K. Note that f_1 can be expressed in terms of the functions ρ and α. It is chosen so that it has period K and its C^d-norm is bounded by $\sim \tau^{-d}$, because it makes "jumps" of order $2/K$ over intervals of size $\sim \tau$. Similar estimate holds for α, which makes "jumps" of order of 1 over intervals of size $\sim \tau$. As one can see from expression (5.14) derivatives of other components of H are bounded say by $(2\pi)^d$. Note also that C^d-norm of f_2 can be bounded by $\sim \delta^{-d}$ because by Lemma 5.1 we have that $\mathbb{T}_\Lambda^{d-1}$ has to contain a non-self-intersecting δ-ball and f_2 is of order 1. Combining these remarks we complete the proof of the proposition. □

Now we need to take care of the general case. Suppose $h(I)$ is strictly convex. Let $h_0(I_0) = h(I, \Lambda)$, where $\Lambda \in \mathbb{R}^{d-1}$ satisfies hypothesis of Lemma 5.1 and h—hypothesis (4.3). First we build the analog of standard EAPT maps (5.7) and (5.8). Define a sequence of numbers: let j be integer $1 \leqslant j \leqslant K$

$$\begin{aligned} \alpha(j, 2m) &= \alpha(0, 2m) + \sum_{s=0}^{j-1} h_0'\left(2m + \tfrac{2s}{K}\right) + \frac{2j}{K} \,(\mathrm{mod}\,1), \\ \alpha(K, 2m) &= \alpha(0, 2m+2) \quad \text{for each } m \in \mathbb{Z} \ \ \& \ \ \alpha(0,0) = 0. \end{aligned} \tag{5.18}$$

Define collection of C^∞ function $\{\alpha(j, I)\}_{0 \leqslant j \leqslant K}$ on $\mathbb{R}$.

Define an analog of sequence of EAPT maps (5.7) as follows:

$$\Phi_j^h\colon (\theta_0, I_0) \to (\theta_0', I_0')$$

$$= \left(\theta_0 + h_0'(I_0) + \frac{2}{K}\sin 2\pi\left\{\theta_0 - \alpha\left(j, I_0 - \frac{2(j-1)}{K}\right)\right\},\right.$$
$$\left. I_0 + \frac{2}{K}\sin 2\pi\left\{\theta_0 - \alpha\left(j, I_0 - \frac{2(j-1)}{K}\right)\right\}\right). \tag{5.19}$$

By analogy with (5.16) extend these functions to one function defined on $[0, K] \times \mathbb{R}$ so that $\partial_s \alpha(s, I_0) = 0$ if $s \pmod 1 \notin [0.5 - \tau, 0.5 + \tau]$. By analogy with (5.8) denote

$$\tilde{\Phi}^h = \Phi^h_K \circ \Phi^h_{K-1} \circ \cdots \circ \Phi^h_1. \tag{5.20}$$

Note that by the construction

$$\Phi^h_j\left(\frac{1}{4} + \alpha(j-1, 2m), 2m + \frac{2(j-1)}{K}\right)$$
$$= \left(\frac{1}{4} + \alpha(j, 2m), 2m + \frac{2j}{K}\right)$$

for each $j = 1, \ldots, K$. Therefore, $\tilde{\Phi}^h\left(\frac{1}{4}, 0\right) = \left(\frac{1}{4}, 2\right)$ and

$$\left(\tilde{\Phi}^h\right)^m\left(\frac{1}{4}, 0\right) = \left(\frac{1}{4}, 2m\right) \quad \text{for each} \quad m \in \mathbb{Z}. \tag{5.21}$$

This implies linear growth of I_0 in time.

We shall not reproduce all the calculations from the proof of Proposition 5.2, we just indicate the changes one has to make to treat the general convex $h_0(I_0)$.

Denote

$$l_0(v) = \sup_{I_0} \{v I_0 - h_0(I_0)\} \tag{5.22}$$

the Legendre transform of $h_0(I_0)$. Define instead of (5.10)

$$h_s(\tilde{\theta}_0, \tilde{\theta}_0', \alpha) = l_0\left(\frac{\tilde{\theta}_0' - \tilde{\theta}_0}{s}\right) - \frac{\rho(s)}{2\pi s}\cos 2\pi(\tilde{\theta}_0 - \alpha). \tag{5.23}$$

It defines a substitute family $\{\Phi^s\}_{0 \leqslant s \leqslant 1}$ for (5.11)

$$\begin{cases} \tilde{\theta}_0' = \tilde{\theta}_0 + s h_0'(I_0) + \rho(s)\sin 2\pi(\tilde{\theta}_0 - \alpha), \\ I_0' = I_0 + \dfrac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0 - \alpha). \end{cases} \tag{5.24}$$

Since $\tilde{\theta}_0' - \tilde{\theta}_0 = sh_0'(I_0) + \rho(s)\sin 2\pi(\tilde{\theta}_0 - \alpha)$, to get from $\tilde{\theta}_0$ to $\tilde{\theta}_0'$ in time s we need I_0 satisfy the above equality. Therefore, action has the form

$$h_s(\tilde{\theta}_0, \tilde{\theta}_0', \alpha) = A_s(\tilde{\theta}_0, I_0) \\ = s\ l\left(h_0'(I_0) + \frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0 - \alpha)\right) - \frac{\rho(s)}{2\pi s}\cos 2\pi(\tilde{\theta}_0 - \alpha).$$

It is problematic to write explicitly an analog of (5.13), namely, inversion of Φ^s. So we semi-explicitly invert it, which suffices for our purposes. By Taylor-type formula

$$h_0'\left(I_0' - \frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0 - \alpha)\right) = h_0'(I_0') - h_0''(I_0^*)\,\frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0 - \alpha),$$

where $I_0^* \in [I_0' - \rho(s)/s, I_0' + \rho(s)/s]$. Recall that $\sup h_0''(I_0^*) \leqslant D$. Therefore, we have

$$\tilde{\theta}_0 = \tilde{\theta}_0' - sh_0'(I_0') - c_s(\alpha, \tilde{\theta}_0', I_0'),$$

where $c_s(\alpha, \tilde{\theta}_0', I_0')$ is an analytic 1-periodic in $\tilde{\theta}_0'$ function uniformly. By definition of $\rho(s)$ before (5.10) it is bounded by $\frac{D}{K}$. Thus, for $(\Phi^s)^{-1}$ we have

$$\begin{cases} \tilde{\theta}_0 = \tilde{\theta}_0' - sh_0'(I_0') - c_s(\alpha, \tilde{\theta}_0', I_0'), \\ I_0 = I_0' - \dfrac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0 - \alpha). \end{cases}$$

Calculate

$$\frac{d\Phi^s}{ds}(\tilde{\theta}_0, I_0) = \\ \left(h_0'(I_0) + \rho(s)\sin 2\pi(\tilde{\theta}_0 - \alpha),\ \left(\frac{\rho(s)}{s}\right)'\sin 2\pi(\tilde{\theta}_0 - \alpha)\right).$$

Apply $((\Phi^s)^{-1})^*$ and get

$$X_s(\tilde{\theta}_0, I_0) = ((\Phi^s)^{-1})^*\,\frac{d\Phi^s}{ds}(\tilde{\theta}_0, I_0) \\ = \left(h_0'\left(I_0 - \frac{\rho(s)}{s}\sin 2\pi\left\{\tilde{\theta}_0 - \alpha - sh_0'(I_0) - c_s(\alpha, \tilde{\theta}_0, I_0)\right\}\right) + \right.$$

$$\rho'(s)\sin 2\pi\left\{\tilde{\theta}_0-\alpha-sh_0'(I_0)-c_s(\alpha,\tilde{\theta}_0,I_0)\right\},$$
$$\left(\frac{\rho(s)}{s}\right)'\sin 2\pi\left\{\tilde{\theta}_0-\alpha-sh_0'(I_0)-c_s(\alpha,\tilde{\theta}_0,I_0)\right\}\Bigg)$$

Therefore,

$$\begin{aligned}&i_{X_s}(I_0 d\tilde{\theta}_0)\\&=I_0\left(h_0'\left(I_0-\frac{\rho(s)}{s}\sin 2\pi\left\{\tilde{\theta}_0-\alpha-sh_0'(I_0)-c_s(\alpha,\tilde{\theta}_0,I_0)\right\}\right)+\right.\\&\left.\rho'(s)\sin 2\pi\left\{\tilde{\theta}_0-\alpha-sh_0'(I_0)-c_s(\alpha,\tilde{\theta}_0,I_0)\right\}\right).\end{aligned}$$

Now we calculate the second term in (5.12)

$$\begin{aligned}\frac{d}{ds}A_s(\tilde{\theta}_0,I_0)=&\left[l\left(h_0'(I_0)+\frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0-\alpha)\right)\right.\\&+sl'\left(h_0'(I_0)+\rho'(s)\ \ \sin 2\pi(\tilde{\theta}_0-\alpha)\right)\left(\frac{\rho(s)}{2s}\right)'\sin 2\pi(\tilde{\theta}_0-\alpha)-\\&\left.\left(\frac{\rho(s)}{2\pi s}\right)'\cos 2\pi(\tilde{\theta}_0-\alpha)\right]ds+\\&sl'\left(h_0'(I_0)+\frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0-\alpha)\right)h_0''(I_0^*)\,dI_0+\\&\left(2\pi l'\left(h_0'(I_0)+\frac{\sin 2\pi(\tilde{\theta}_0-\alpha)}{s}\right)\cos 2\pi(\tilde{\theta}_0-\alpha)+\right.\\&\left.\frac{\rho(s)}{s}\sin 2\pi(\tilde{\theta}_0-\alpha)\right)\rho(s)\,d\tilde{\theta}_0,\end{aligned}$$

where differentials are given by $d\tilde{\theta}_0=(h_0'(I_0)+\rho'(s)\sin 2\pi(\tilde{\theta}_0-\alpha))\,ds$ and $dI_0=(\rho(s)/s)'\ \sin 2\pi(\tilde{\theta}_0-\alpha)\,ds$. Apply expressions for $((\Phi^s)^{-1})^*$, Taylor type formula for $h_0'(I_0)$, $l(v)$, $l'(v)$, and bounds on

$$D^{-1}<h_0''(I_0)<D\quad\Rightarrow\quad D^{-1}<l''(v)<D \tag{5.25}$$

to formula (5.12). Substituting expressions for dI_0 and $d\tilde{\theta}_0$ we get

$$\begin{aligned}H'(\theta_0,I_0,s,\alpha)=&\,h_0(I_0)-2\pi\rho(s)I_0h_0'(I_0)\cos 2\pi(\theta_0-\alpha-sI_0)+\\&g(I_0,s,\theta_0-\alpha-sh_0'(I_0)-c_s(\alpha,\tilde{\theta}_0,I_0)),\end{aligned}\tag{5.26}$$

where g is 1-periodic in the second and third variable and $|g(I_0, \varphi, s)| < 4\pi D(|I_0|+1)$. Notice that $I_0 h'(I_0) < 2D(|I_0|+1)^2$ for all $I_0 \in \mathbb{R}$. Periodicity of $c_s(\alpha, \tilde{\theta}_0, I_0)$ in $\tilde{\theta}_0$ of period 1 implies that g is 1-periodic both in $\tilde{\theta}_0$ and s. By the construction the time-1-map corresponding to H^* has the form (5.24). It is the same as (5.19) with $\alpha(s, I_0)$'s defined in (5.18). To include the time-1-maps we have constructed into the time-K-map recall that the Hamiltonian H we use dependence on $\theta_1^\Lambda = s$ which has period K. So we introduce the function $\alpha\colon [0, K] \to \mathbb{R}$ as in (5.18). This way we get a Hamiltonian of the form

$$\begin{aligned} H(\theta_0, I_0, s, t) = h_0(I_0) \\ -2\pi\rho(s) I_0 h_0'(I_0) \cos 2\pi(\theta_0 - \alpha(s, I_0) - tI_0) + \\ g(I_0, s, \theta_0 - \alpha(s) - tI_0 - c_s(\alpha(s, I_0), \tilde{\theta}_0, I_0)), \end{aligned} \tag{5.27}$$

where g is as above. This implies 1-periodicity in θ_0 and K-periodicity in $s = \theta_1^\Lambda$. To see 1-periodicity in t choose time parametrization by s of θ_1^Λ. Then when α is fixed we have g is period 1 in the variable s.

In order to estimate C^d-norm of the corresponding perturbations f it suffices to bound C^d norm of f_1 with respect to the original coordinate system $(\theta_0, \theta_1, \ldots, \theta_d)$. Here we proceed in the same way as in the end of proof of Proposition 5.2. This proves the following:

Proposition 5.4. *With the notations of Proposition* 5.2 *assume that* $\Lambda \in \mathbb{R}^{d-1}$ *satisfies conditions of Lemma* 5.1, *denote* $h_0(I_0) = h(I_0, \Lambda)$, *and assume*

$$D^{-1} < h_0''(I_0) < D \quad \textit{for some } D > 0 \textit{ and all } I_0 \in \mathbb{R}, \tag{5.28}$$

then there is a C^∞*-perturbation* $\varepsilon f(\theta, I_0, t)$ *satisfying* (5.9) *such that the corresponding Poincare map* $\tilde{\Phi}^h$ *has the form* (5.20).

Remark. Notice that D in (5.28) and in (4.3) can be choosen the same.

This completes the proof of Theorem 4.1 for C^∞-perturbations.

6. Linear diffusion for curveless twist maps

In this and three subsequent sections we prove a generalization of the statement we made in the beginning of the previous section about existence ballistic trajectories for curveless twist maps. First define the object of investigation of Aubry–Mather theory (see [Ba,M-F] for an exposition). This theory studies EAPT maps. These maps also naturally arise as Poincare return maps of Hamiltonian systems with two degrees of freedom. Let $\mathbb{A} = \mathbb{T} \times \mathbb{R}$ be an annulus, $\theta \in \mathbb{T}$, $I \in \mathbb{R}$, and $\pi_1(\theta, I) = \theta$ be the natural projection onto the first component, $r\geqslant 1$. A C^r-smooth EAPT map $\Phi\colon \mathbb{A}\to\mathbb{A}$ is a C^r-smooth map with the following properties:

- (*area-preservation*) it preserves a smooth non-degenerate area 2-form;
- (*twist*) let l_a be the image of a vertical line $a\times\mathbb{R}\subset\mathbb{A}$ under Φ, then $\pi\colon l_a\to\pi(l_a)\subset\mathbb{T}$ is a local diffeomorphism for every $a\in\mathbb{T}$;
- (*exact* or no "vertical drift") for any non-contractible curve γ on $\mathbb{A}$ area above γ below $\Phi(\gamma)$ and area above $\Phi(\gamma)$ below γ is the same.

It is easy to see that the last condition is independent of γ. An EAPT map can also be defined on an open set $U\subset\mathbb{A}$. In this case one usually extends Φ to the whole annulus so that Φ is C^r in $\mathbb{A}$.

By a result due to Moser [Mo] any EAPT map Φ of the annulus can be represented by a time-periodic Hamiltonian system on $H: T^*\mathbb{T}\times\mathbb{T}\simeq\mathbb{T}\times\mathbb{R}\times\mathbb{T}\to\mathbb{R}$ satisfying convexity condition: $H_{\mathrm{I_0I_0}}(\theta_0,I_0,t)>0$, i.e. Φ coincides with the time-1-map of the Hamiltonian flow associated to H. Thus, exact area-preserving twist maps are simply discretization of certain time-periodic Hamiltonian systems.

Theorem 6.1. *Let* $\Phi\colon\mathbb{A}\to\mathbb{A}$ *be a* C^1 EAPT *map of the annulus*, $(\theta,I)\in\mathbb{T}\times\mathbb{R}=\mathbb{A}$. *Suppose* Φ *is* C^1*-close to the* EAPT Φ^*, *given by* (5.20). *Then there are positive* c_1 *and* c_2 *and trajectories* $\{\Phi^n(\theta^0,I^0)=(\theta^n,I^n)\}_{n\in\mathbb{Z}}$ *such that*

$$c_1 n < I^n < c_1 n + c_2 \quad for\ all \quad n\in\mathbb{Z}. \tag{6.1}$$

The proof of this result utilizes Mather variational method of changing Lagrangians. This method is based on an analysis of the so-called barrier functions and is developed in [Ma3] (see also [Ma4] for the most recent treatment). To expose this method we need to introduce basic notions of *Aubry–Mather theory*, *Mather's and Peierls's barrier functions* and then describe the method. Our main contribution is linear estimates of diffusion time, which require detailed analysis of the method.

To prove existence of diffusing trajectories Xia [Xia] proposes slightly different variational approach also based on his way of defining barrier functions.

To fit the standard notations of Aubry–Mather theory in what follows for any $n\in\mathbb{Z}$ we denote (θ_n,I_n) a point on the first action-angle cylinder $\mathbb{A}$ *not in* nth *one*.

6.1. Aubry–Mather theory

Each C^1 smooth EAPT map can be described by a so-called *generating function* $h:\mathbb{R}^2\to\mathbb{R}^2$ in the following way: Let $\tilde\Phi\colon\mathbb{R}^2\to\mathbb{R}^2$ be a lift of Φ, given by $\tilde\Phi(\tilde\theta+1,\tilde\theta')=\tilde\Phi(\tilde\theta,\tilde\theta')+(1,0)$. Then $\tilde\Phi(\tilde\theta_0,I_0)=(\tilde\theta_1,I_1)$ can be implicitly defined by the following equations:

$$\begin{cases} I_0=-\partial_1 h(\tilde\theta_0,\tilde\theta_1),\\ I_1=\partial_2 h(\tilde\theta_0,\tilde\theta_1). \end{cases} \tag{6.2}$$

Here ∂_i, $i = 1, 2$ is the partial derivative with respect to the ith component, $h \in C^2$, $h(\tilde{\theta}_0 + 1, \tilde{\theta}_1 + 1) = h(\tilde{\theta}_0, \tilde{\theta}_1)$, and $\partial_2\partial_1 h \leqslant -\gamma < 0$ for some $\gamma > 0$. In our case it is important that EAPT maps $\tilde{\Phi}_j^h$ and Φ^h defined in (5.19)–(5.20) by assumption (5.28) have γ uniformly bounded from below. In turn (5.28) follows from (4.3).

The value of $h(\tilde{\theta}, \tilde{\theta}')$ equals minimal action to get from $\tilde{\theta}$ to $\tilde{\theta}'$ in time 1, where action arises from time-periodic Lagrangian system associated to Φ by Moser's theorem [Mo].

Aubry–Mather theory studies the orbit structure of EAPT maps by projecting orbits into their first components, which form *configuration space*. Consider the space of configurations $\mathbb{R}^{\mathbb{Z}} = \{\Theta |\ \Theta : \mathbb{Z} \to \mathbb{R}\}$—bi-infinite sequences of real numbers with product topology. Given a function $h : \mathbb{R}^2 \to \mathbb{R}$, extend h to arbitrary finite segments $(\tilde{\theta}_j, \dots, \tilde{\theta}_k)$, $j < k$, of configurations $\Theta \in \mathbb{R}^{\mathbb{Z}}$ by

$$h(\tilde{\theta}_j, \dots, \tilde{\theta}_k) = \sum_{i=j}^{k-1} h(\tilde{\theta}_i, \tilde{\theta}_{i+1}).$$

Say that segment is *minimal* or *action-minimizing* with respect to h if

$$h(\tilde{\theta}_j, \dots, \tilde{\theta}_k) \leqslant h(\tilde{\theta}_j^*, \dots, \tilde{\theta}_k^*)$$

for all $(\tilde{\theta}_j^*, \dots, \tilde{\theta}_k^*)$ with $\tilde{\theta}_j = \tilde{\theta}_j^*$ and $\tilde{\theta}_k = \tilde{\theta}_k^*$.

A configuration $\tilde{\theta} \in \mathbb{R}^{\mathbb{Z}}$ is called *minimal* or *action-minimizing* with respect to h if every finite segment of $\tilde{\theta}$ is minimal or action-minimizing with respect to h. The set of all action-minimizing trajectories is denoted by $\tilde{\Sigma} = \tilde{\Sigma}(h) \subset \mathbb{R}^{\mathbb{Z}}$.

A configuration $\tilde{\theta} \in \mathbb{R}^{\mathbb{Z}}$ is called *stationary* if

$$\partial_2 h(\tilde{\theta}_{k-1}, \tilde{\theta}_k) + \partial_1 h(\tilde{\theta}_k, \tilde{\theta}_{k+1}) = 0 \quad \text{for all} \quad k \in \mathbb{Z} \tag{6.3}$$

This equation is an analog of the Euler–Lagrange equation in this case. Indeed, this equation says that the sum $\sum_k h(\tilde{\theta}_k, \tilde{\theta}_{k+1})$ is extremized with respect to each $\tilde{\theta}_k$, because formal derivative of the sum with respect to each $\tilde{\theta}_k$ is zero. In particular, each minimal configuration is stationary. The set of all stationary trajectories will be denoted by $\widetilde{St} = \widetilde{St}(h) \subset \mathbb{R}^{\mathbb{Z}}$. We have $\tilde{\Sigma} \subset \widetilde{St}$.

Lemma 6.2. *Suppose h is a C^2 smooth function. Then there is one-to-one correspondence between stationary configurations and orbits of an* EAPT $\Phi : \mathbb{A} \to \mathbb{A}$, *given by the following relation: let* $0 \leqslant \tilde{\theta}_0 = \theta_0 < 1$, *then*

$$\begin{aligned} \{\tilde{\theta}_k\}_{k\in\mathbb{Z}}, &\longrightarrow \quad \Phi^k(\theta_0, I_0) = (\tilde{\theta}_k \bmod 1, \partial_2 h(\tilde{\theta}_{k-1}, \tilde{\theta}_k)), \\ \{\Phi^k(\theta_0, I_0)\}_{k\in\mathbb{Z}} &\longrightarrow \quad \tilde{\Phi}^k(\theta_0, I_0) = (\tilde{\theta}_k, I_k), \ \{\tilde{\theta}_k\}_{k\in\mathbb{Z}}. \end{aligned} \tag{6.4}$$

Proof is by direct calculation using (6.3).

Aubry graph of a configuration $\Theta = \{\tilde\theta_n\}_{n\in\mathbb{Z}}$ is a graph of piecewise linear function $\mathcal{A}(\Theta)$ equal to $\tilde\theta_n$ at each integer n and linearly interpolated in between. A configuration $\Theta \in \mathbb{R}^{\mathbb{Z}}$ has *rotation number* $\rho(\Theta)$ if the following limit exists $\rho(\Theta) = \lim_{n\to\pm\infty} \tilde\theta_n/n$. Similarly one-sided configuration $\Theta \in \mathbb{R}^{\mathbb{Z}_+}$ has rotation number if $\lim_{n\to+\infty} \tilde\theta_n/n$ exists. Action-minimizing configurations have the following properties.

Aubry–Mather Theorem. *Every minimal configuration $\Theta \in \tilde\Sigma$ has rotation number, i.e. $\rho(\Theta)$, exists and for every rotation number $\omega \in \mathbb{R}$ there is a minimal configuration $\Theta \in \tilde\Sigma$ with $\rho(\Theta) = \omega$.*

The union of the set of minimal configurations with the rotation number ω, denoted by $\widetilde\Sigma_\omega$, is called an *Aubry–Mather set* in the space of stationary configurations $\widetilde{St}$. To trace Aubry–Mather set on the cylinder $\mathbb{A}$ we define the projection $\pi^* : \widetilde{St} \to \mathbb{A}$ by $\pi^*(\Theta) = (\theta_0, I_0)$ and (θ_0, I_0) is as defined in (6.4). Then $\Sigma_\omega = \pi^*(\widetilde\Sigma_\omega) \subset \mathbb{A}$ is Aubry–Mather set on the cylinder.

Similar to the above, one can have one-sided minimizers. Fix $a \in \mathbb{R}$. A configuration $\Theta_a^\pm = \{\tilde\theta_n^a\}_{n\in\mathbb{Z}_\pm}$ is called one-sided minimizer starting at a if it is minimal for any $0 \leqslant j < k$ (resp. $j < k \leqslant 0$) and $\tilde\theta_0^a = a$. Aubry–Mather Theorem can be extended to one-sided minimizers.

Generalized Aubry Crossing Lemma (GAC Lemma). *For each $a \neq b \in \mathbb{R}$ and rotation number ω there are one-sided minimizers $\Theta_a^\pm$ and $\Theta_b^\pm$ of rotation number ω passing through a and b, respectively, and their Aurby graphs do not intersect. Moreover, each one-sided minimizer has rotation number and its limit set is contained in the Aubry–Mather set of the corresponding rotation number.*

Non-intersection property can be proven in the same way as Aubry Crossing Lemma (see e.g. [M-F, Section 9]). We postpone the rest of the proof until the next section.

Describe structure of Aubry–Mather sets. It turns out they can be only of a few topological types. Denote $\mathrm{Rec}(\Phi) = \{(\theta, I) \in \mathbb{A} : \ (\theta, I) \in \overline{\cup_{n\neq 0}\Phi^n(\theta, I)}\}$ the set of recurrent points of Φ.

Structure Theorem (irrational case $\omega \notin \mathbb{Q}$). *Aubry–Mather set $\Sigma_\omega \subset \mathbb{A}$ is*

- *either an invariant curve or*
- *if $\Sigma_\omega^{\mathrm{rec}}$ is the set of recurrent points in Σ_ω, then its projection $\pi_1(\Sigma_\omega^{\mathrm{rec}})$ is a Denjoy–Cantor set. Denjoy–Cantor set means that this is a Cantor set and dynamics of Φ, restricted to $\Sigma_\omega^{\mathrm{rec}}$, $\Phi|_{\Sigma_\omega^{\mathrm{rec}}} : \Sigma_\omega^{\mathrm{rec}} \to \Sigma_\omega^{\mathrm{rec}}$ is conjugated to a circle homeomorphism $\varphi_\omega : \mathbb{T} \to \mathbb{T}$ by the projection π_1 on the base, i.e. $\varphi_\omega \circ \pi_1 \equiv \pi_1 \circ \Phi$ on $\Sigma_\omega^{\mathrm{rec}}$. Generically $\Sigma_\omega^{\mathrm{rec}} = \Sigma_\omega$.*

Remark 4. Any circle homeomorphism is recurrent on the set of its recurrent points. Moreover, every recurrent trajectory is dense.

To formulate the Structure Theorem in the case of rational rotation number $\omega = p/q$, we introduce some sets. Let $\Sigma^{\text{per}}_{p/q}$ the set of action minimizing periodic points of period q and rotation number p/q. Two periodic points θ^- and θ^+ are adjacent elements of $\Sigma^{\text{per}}_{p/q}$ if projections $\pi_1(\theta^-)$ and $\pi_1(\theta^+)$ have a segment in $\mathbb{T}$ free from projection of all the other elements of $\pi_1(\Sigma^{\text{per}}_{p/q})$. For adjacent periodic points θ^- and θ^+ in $\Sigma^{\text{per}}_{p/q}$ let

$$\begin{aligned} \Sigma^+_{p/q}(\theta^-, \theta^+) &= \{\theta \in \Sigma_{p/q} : \ \theta \text{ is } \alpha \text{ (resp. } \omega\text{)–asymptotic to } \theta^- \text{ (resp. } \theta^+)\}, \\ \Sigma^-_{p/q}(\theta^-, \theta^+) &= \{\theta \in \Sigma_{p/q} : \ \theta \text{ is } \alpha \text{ (resp. } \omega\text{)–asymptotic to } \theta^+ \text{ (resp. } \theta^-)\}. \end{aligned} \tag{6.5}$$

Let $\Sigma^{\pm}_{p/q}$ be the union of $\Sigma^{\pm}_{p/q}(\theta^-, \theta^+)$ over all adjacent periodic joints θ^- and θ^+ in $\Sigma^{\text{per}}_{p/q}$.

Structure theorem (rational case $\omega = p/q \in \mathbb{Q}$). *The Aubry–Mather set $\Sigma_{p/q}$ is a disjoint union of $\Sigma^{\text{per}}_{p/q}$, $\Sigma^+_{p/q}$, and $\Sigma^-_{p/q}$. Moreover, $\Sigma^{\text{per}}_{p/q}$ is always non-empty and if $\Sigma^{\text{per}}_{p/q}$ is not a curve, then $\Sigma^-_{p/q}$ and $\Sigma^+_{p/q}$ are non-empty too.*

It turns out that position of Aubry–Mather sets on the cylinder has additional properties.

Graph Theorem. *For each $\omega \in \mathbb{R}$ the Aubry–Mather set $\Sigma_\omega \subset \mathbb{A}$ is a Lipschitz graph over the base $\mathbb{T}$, i.e. $\pi_1^{-1} : \pi_1 \Sigma_\omega \to \Sigma_\omega$ is Lipschitz with Lipschitz constant $L = L(\gamma)$ depending only on minimal angle of twisting γ.*

There is a different way of looking at Aubry–Mather theory by a special variational principle found by Mather (see e.g. [M-F, Section 6]). The way it is introduced below allows generalization to higher dimensional case [Ma2].

Denote by J_1 the set of maps $\varphi : \mathbb{R} \to \mathbb{R}$ which are increasing, i.e. if $s \leqslant t$, then $\varphi(s) \leqslant \varphi(t)$, and satisfy the periodicity condition $\varphi(t+1) = \varphi(t) + 1$. Let J denote J_1 modulo the following identifications: $\varphi \sim \psi$ if there exists $a \in \mathbb{R}$ such that $\varphi(t) = \psi(t+a)$ at all but at most countably many t.

For $\omega \in \mathbb{R}$ and $\varphi \in J_1$ we let

$$F_\omega(\varphi) = \int_a^{a+1} h(\varphi(t), \varphi(t+\omega))\, dt. \tag{6.6}$$

It turns out that

Theorem 6.3 (*see e.g. Mather–Forni [M-F]*). *There exists a minimizer φ_ω of F_ω in J_1 with the following properties*:

- *Let $\beta : \mathbb{R} \to \mathbb{R}$, $\beta(\omega) \mapsto F_\omega(\varphi_\omega)$. Then β is a convex continuous function called Mather's β-function.*[5]

[5] $\beta(\omega)$ is minimal average action along trajectories with rotation number ω.

- *For all $t \in \mathbb{T}$ we have that the Euler–Lagrange equation holds*:

$$\partial_2 h(\varphi_\omega(t-\omega\pm 0), \varphi_\omega(t\pm 0)) + \partial_1 h(\varphi_\omega(t\pm 0), \varphi_\omega(t+\omega\pm 0)) = 0.$$

- *Denote* $\eta^\pm(t) = -\partial_1 h(\varphi_\omega(t-\omega\pm 0), \varphi_\omega(t\pm 0)) = \partial_2 h(\varphi_\omega(t\pm 0), \varphi_\omega(t+\omega\pm 0))$. *The set*

$$\Sigma_\omega = \{(\varphi_\omega(t\pm 0), \eta^\pm(t)) : t \in \mathbb{T}\} \subset \mathbb{A}$$

is the same Aubry–Mather set as above.

Remark 5. Let $\Theta_\omega \in \widetilde{\Sigma}_\omega$ be a minimal configuration of rotation ω. Note that

$$\beta(\omega) = \frac{1}{n} \lim_{n\to\infty} \sum_{i=0}^{n-1} h(\tilde\theta_i, \tilde\theta_{i+1}).$$

Indeed, for rational ω it is obvious. Let ω be irrational. By the Structure Theorem, dynamics on Σ_ω is conjugate to a circle homeomorphism φ_ω with irrational rotation number ω. Every such homeomorphism has a unique invariant measure supported on recurrent trajectories $\mathrm{Rec}(\varphi_\omega)$ and every recurrent trajectory is dense in $\mathrm{Rec}(\varphi_\omega)$. Therefore, by Krylov–Bogolyubov construction average along every trajectory converges to a "space average". The latter one is unique by uniqueness of an invariant measure.

7. *c*-minimization and barriers

In this section we introduce Mather's and Peierls's barrier functions. These functions are the key tool to prove the existence of diffusing trajectories.

Consider a closed one form $\eta = c + df$ on $\mathbb{T}$, where f is a C^2 smooth function on $\mathbb{R}$ of period 1 and c is a real number. By definition cohomology class of η is $[\eta]_{\mathbb{T}} = c$. Denote by

$$\tilde h_\eta(\tilde\theta_0, \tilde\theta_1) = h(\tilde\theta_0, \tilde\theta_1) - (c(\tilde\theta_1 - \tilde\theta_0) + f(\tilde\theta_1) - f(\tilde\theta_0)). \tag{7.1}$$

It is easy to see that stationary configurations of $\tilde h_\eta$ and h are the same. Indeed, in any finite sums of $\tilde h_\eta$'s and h's

$$\sum_{i=j}^{k-1} \tilde h_\eta(\tilde\theta_i, \tilde\theta_{i+1}) = \left(\sum_{i=j}^{k-1} h(\tilde\theta_i, \tilde\theta_{i+1})\right) - [c(\tilde\theta_k - \tilde\theta_j) + f(\tilde\theta_k) - f(\tilde\theta_j)]$$

differs only at boundary terms. If end points are fixed, stationarity with respect to h implies stationarity with respect to $\tilde{h}_\eta$. What is important for our discussion is that minimizers of $\tilde{h}_\eta$ and h might be different.

Definition 7.1. The function $\alpha: \mathbb{R} \to \mathbb{R}$ given by

$$\alpha(c) = -\min_{\omega} \{\beta(\omega) - c\omega\} \tag{7.2}$$

is called *Mather's α-function*. Call the map

$$\mathcal{L}_\beta : \mathbb{R}_\omega \to \{\text{compact, convex, non-empty intervals in } \mathbb{R}_c\}, \tag{7.3}$$

defined by letting $\mathcal{L}_\beta$ be the set of $c \in \mathbb{R}_c$ for which the above inequality becomes equality, *Fenchel–Legendre transform*.

We keep the name Legendre transform for the classical transform which relates Hamiltonians and Lagrangians.

Denote

$$h_\eta(\tilde{\theta}_0, \tilde{\theta}_1) = \tilde{h}_\eta(\tilde{\theta}_0, \tilde{\theta}_1) - \alpha(c).$$

After we subtract a constant all minimal and stationary configurations stay minimal and stationary, respectively. Let

$$F_\eta(\varphi) = \inf_{\varphi} \int_a^{a+1} h_\eta(\varphi(t), \varphi(t+\omega))\, dt. \tag{7.4}$$

Remark 6. By definition of minimizer φ_ω of F_ω we get $F_\eta(\varphi_\omega) = \beta(\omega) - c\omega + \alpha(c)$. Using minimality in the definition of α-function one can prove that $F_\eta(\varphi_\omega) \geqslant 0$ and equals zero if and only if $c \in \mathcal{L}_\beta(\omega)$. In [Ma2, Section 6] Mather proved it in a more general case. This way we see that functional (7.4) of h_η attains its minimum on Aubry–Mather set Σ_ω, provided $c \in \mathcal{L}_\beta(\omega)$. By Structure Theorem we have that Aubry–Mather set could be either a Denjoy–Cantor set, or an invariant curve, or a union of periodic and heteroclinic orbits. This naturally implies that each of those invariant sets carries an invariant probability measure μ, called ω-minimal (resp. c-minimal) measure. For each $c \in \mathcal{L}_\beta(\omega)$ sometimes we redenote Σ_ω by Σ^c.

It turns out that generically the graph of Fenchel–Legendre transform is a devil staircase (see e.g. [Ba,Ma5]). At every irrational ω we have $\mathcal{L}_\beta(\omega)$ is one point and for every rational $\omega = p/q$ we have $\mathcal{L}_\beta(\omega)$ is a non-empty interval $[c^-, c^+] = [\beta'(p/q - 0), \beta'(p/q + 0)]$ (see [Ma2, Section 6]).

Introduce *Mather's barrier*

$$B_\eta(a) = \liminf_{n,n'\to\infty} \inf_{\tilde\theta_0=a} \sum_{i=-n}^{n-1} h_\eta(\tilde\theta_i, \tilde\theta_{i+1}). \tag{7.5}$$

It turns out that this limit exists and under some additional condition is independent of the function f defining η (see lemmas below). Peierls gave a somewhat different definition of a barrier function. We formulate a version of it below. The present definition is a slight modification of the following similar definition due to Mather [Ma2, Section 6, p. 1368], denoted by $B_c(\xi)$ there [6]

$$B_\eta^M(a) = \liminf_{n,\ n'\to+\infty} \inf_{\tilde\theta_0=a,\ \tilde\theta_n=\tilde\theta_{-n'}\ (\mathrm{mod}\ 1)} \sum_{i=-n'}^{n-1} h_\eta(\tilde\theta_i, \tilde\theta_{i+1}). \tag{7.6}$$

These definitions can be extended to convex Hamiltonian systems of arbitrary number of degrees of freedom. We shall prove that for EAPT maps these definitions are equivalent under an additional condition on f. Recall that we treat now the convex case of time-periodic Hamiltonian systems in 1 degrees of freedom what is usually called 1.5 degrees of freedom.

Before stating and proving properties of these barriers we need to introduce Peierls's barrier: For each rotation number ω we define a real-valued function P_ω so that it is identically zero on $\pi_1(\Sigma_\omega)$. Thus, it is enough to consider $a \in \mathbb{T} \setminus \pi_1(\Sigma_\omega)$. Since $\pi_1(\Sigma_\omega)$ is compact, the component of $\mathbb{T}\setminus\pi_1(\Sigma_\omega)$ which contains a is a segment, whose end points we denote a_- and a_+, $0\leqslant a < 1$. Since $a_\pm \in \pi_1(\Sigma_\omega)$, there are unique minimal configurations $\Theta^a_\pm = (\ldots, \tilde\theta^a_{i\pm}, \ldots)$ such that $\tilde\theta_{0\pm} = a_\pm$ respectively. Choose lifts $\tilde a, \tilde a_\pm, \tilde\theta_{i\pm}$ to the universal cover $\mathbb{R}$ such that $\tilde a_- < \tilde a < \tilde a_+ < \tilde a_- + 1$, $\tilde\theta^a_{0\pm} = \tilde a_\pm$ and such that the lifted $\Theta^a_\pm = (\ldots, \tilde\theta^a_{i\pm}, \ldots)$ are minimal configurations. *Peierls's barrier* is defined as follows:

$$P_\omega(a) = \min\left\{\sum_{i\in\mathbb{Z}} \left(h(\tilde\theta^a_i, \tilde\theta^a_{i+1}) - h(\tilde\theta^a_{i-}, \tilde\theta^a_{i+1-})\right)\right\}, \tag{7.7}$$

where h is the generating function under consideration. The minimum is taken over all $\Theta^a = (\ldots, \tilde\theta^a_i, \ldots) \in \mathbb{R}^{\mathbb{Z}}$ such that $\tilde\theta^a_{i-} \leqslant \tilde\theta_i \leqslant \tilde\theta^a_{i+}$ and $\tilde\theta_0 = a$, $0 \leqslant \tilde\theta_0 < 1$. The condition $\tilde\theta^a_{i-} \leqslant \tilde\theta_i \leqslant \tilde\theta^a_{i+}$ guarantees that the sum is absolutely convergent, since $\sum_i (\tilde\theta^a_{i+1-} - \tilde\theta^a_{i-}) \leqslant 1$ in the case ω is irrational. Note that if Θ^a_- is replaced by Θ^a_+, the above formula for $P_\omega(a)$ is still valid. Moreover, taking into account that minimum

[6] To get exactly the definition in [Ma2] one permutes indices $[-n', n] = [-n', 0] \cup [0, n]$ to $[0, n] \cup [-n', 0]$ below.

is realized on the union of one-sided minimizers the condition $\tilde{\theta}^a_{i-} \leqslant \tilde{\theta}^a_i \leqslant \tilde{\theta}^a_{i+}$ for all $i \in \mathbb{Z}$ corresponds to non-intersection property in GAC Lemma.

In the rational case there are three different barrier functions: $P_{p/q\pm}$ and $P_{p/q}$. Let $P_{p/q\pm}$ be defined as above with Σ_ω replaced by $\Sigma^{\pm}_{p/q}$. Arguments proving existence of the limits in the $p/q\pm$ case are similar. To define $P_{p/q}$ replace Σ_ω by $\Sigma^{\text{per}}_{p/q}$ and the infinite sum in (7.7) by the finite sum of q terms

$$P_{p/q}(a) = \min_{\tilde{\theta}_0 = \tilde{\theta}_q - p = a} \left\{ \sum_{i=0}^{q-1} \left(h(\tilde{\theta}^a_i, \tilde{\theta}^a_{i+1}) - h(\tilde{\theta}^a_{i-}, \tilde{\theta}^a_{i+1-}) \right) \right\} \tag{7.8}$$

with periodic boundary conditions. To distinguish p/q and $p/q\pm$ introduce rotation symbol ω^* is ω if $\omega \notin \mathbb{Q}$ and is either $p/q-$, or p/q, or $p/q+$ if $\omega \in \mathbb{Q}$. Then we have the family of barrier functions $\{P_{\omega^*}\}_{\omega^*}$. The following fact follows from the definition.

Corollary 7.2 (*Mather criterion of invariant curves; see e.g. Mather and Forni [M-F]*). *A C^1-smooth EAPT map $\tilde{\Phi}$ has an invariant curve Σ_ω of irrational* (*resp. rational*) *rotation number ω* (*resp. $\omega = p/q$*) *if and only if $P_\omega \equiv 0$* (*resp.* $\min\{P_{p/q-}, P_{p/q+}\} \equiv 0$).

Lemma 7.3. *Suppose $\eta = c + df$ and f is constant on $\pi_1(\Sigma_\omega)$. Then for barriers B_η and B^M_η, defined in* (7.5) *and* (7.6) *respectively, satisfy $B_\eta \equiv B^M_\eta$.*

Proof. We show that for a configuration Θ^a realizing the minimum of B_η one can find arbitrary large n' and n with $\tilde{\theta}_{-n'} - \tilde{\theta}_n \pmod 1$ being arbitrary small. This would imply that at an arbitrary small cost we could change that configuration and satisfy boundary condition in (7.6).

Since the only constrain in (7.5) is at the origin, the minimum is attained on the union of two one-sided minimizers. By GAC Lemma each of them has rotation number $\omega^\pm$ and accumulates the corresponding Aubry–Mather set $\Sigma_{\omega^\pm}$, respectively. By Remark 6 $\omega^\pm = \omega$ with $c \in \mathcal{L}_\beta(\omega)$. By GAC Lemma one-sided minimizers $\Theta^a_+ = (\tilde{\theta}^a_0, \ldots)$ and $\Theta^a_- = (\ldots, \tilde{\theta}^a_0)$ do not cross with minimizers of rotation number ω and approach them. If $\omega = p/q$ is rational, both $\Theta^a_\pm$ accumulate to the same boundary periodic point and, therefore, get arbitrary close. If ω is irrational, $\Theta^a_\pm$ accumulate to Σ_ω, and dynamics on Σ_ω is recurrent. Again $\tilde{\theta}^a_n$ and $\tilde{\theta}^a_{-n'} \pmod 1$ get arbitrary close. $\square$

It is also useful for the arguments below to define one-sided barriers:

$$\begin{aligned} B^+_\eta(a) &= f(a) + \liminf_{n \to +\infty} \inf_{\tilde{\theta}_0 = a} \sum_{i=0}^{n-1} h_\eta(\tilde{\theta}_i, \tilde{\theta}_{i+1}), \\ B^-_\eta(a) &= -f(a) + \liminf_{n \to +\infty} \inf_{\tilde{\theta}_0 = a} \sum_{i=-n}^{-1} h_\eta(\tilde{\theta}_i, \tilde{\theta}_{i+1}). \end{aligned} \tag{7.9}$$

Lemma 7.4. *Suppose $\eta = c + df$, $c \in \mathcal{L}_\beta(\omega)$ for some ω, and f is constant on $\pi_1(\Sigma_\omega)$. Then the limits (7.5) and (7.9) do exist and are achieved on the union one-sided minimizers of rotation number ω, respectively. Moreover, the functions $B_c^\pm(a) - B_\eta^\pm(a)$ are identically constant. In particular, the minima of B_η and $B_\eta^\pm$ are independent of choice of a closed one form η.*

The functions $B_c(a)$, $B_c^\pm(a)$, $B_\eta^\pm(a)$, and $B_\eta(a)$ are continuous in c and Lipschitz continuous in a. Moreover, for each c each of these functions takes its minimum on the projected Aubry–Mather set $\pi_1(\Sigma_\omega)$ as above.

Proof. We just proved that $B_\eta \equiv B_\eta^M$. Thus, instead of discovering bicycle we reduce the proof to the results of Mather [Ma2,Ma6]. In [Ma2, Propositions 7.1 and 7.2] Mather proved that when $\omega = \alpha'(c)$ is irrational we have $B_c^M = P_\omega$ and when $\omega = \alpha'(c)$ is rational and $\omega = p/q$ in the lowest terms we have $B_c^M = P_{p/q+}$ (resp. $B_c^M = P_{p/q-}$) for $c = \max\{c^* : \alpha'(c^*) = p/q\}$ (resp. $c = \min\{c^* : \alpha'(c^*) = p/q\}$).

Thus, Mather's barrier $B_c^M(a)$ equals Peierls's barrier $P_\omega(a)$ in all cases when $(c, \alpha(c))$ is an extremal point of the epigraph of the α-function. When $(c, \alpha(c))$ is not an extremal point of the epigraph, then it lies on a flat part of the graph of α. Let p/q be the slope of this flat part, expressed in lowest terms. Then, as it is shown in [Ma2, Proposition 7.2], B_c^M and $P_{p/q}$ have the same zero set $\Sigma_{p/q}^{\text{per}}$. However, they might not be equal.

In [Ma6] (see also [M-F, Theorem 18.2]) proved the following moduli continuity of barriers: Let p/q be rational in lowest terms, then

$$
\begin{aligned}
&|P_{\omega^*}(a) - P_{p/q}(a)| \leqslant C\gamma\left(\frac{1}{q} + |q\omega - p|\right), \\
&|P_{\omega^*}(a) - P_{p/q+}(a)| \leqslant C\gamma|q\omega - p| \quad \text{when } \omega \geqslant p/q+, \\
&|P_{\omega^*}(a) - P_{p/q-}(a)| \leqslant C\gamma|q\omega - p| \quad \text{when } \omega \leqslant p/q-,
\end{aligned}
\tag{7.10}
$$

where ω denote rotation number of the symbol ω^* and γ is the minimal angle of twisting (see [M-F, Section 7] for precise definition of γ denoted θ there). In particular, convexity of the α-function along with these estimates show continuity in ω at irrational ω's and implies continuity at c's with $\alpha'(c)$ being irrational. Now we prove continuity in c for $\alpha'(c)$ being rational.

In the rational case the Aubry–Mather set $\Sigma_{p/q}^{\text{per}}$ consists of action minimizing periodic points of period q and rotation number p/q. Pick one of them and lift it to the universal cover $\mathbb{R}$ and let $(\tilde\theta_0, \dots, \tilde\theta_{q-1}) \in \mathbb{R}^q$ be the lifted periodic block with $0 \leqslant \tilde\theta_0 < 1$. In this case by definition the β-function has the form

$$
\beta\left(\frac{p}{q}\right) = \frac{1}{q}\sum_{i=0}^{q-1} h(\tilde\theta_i, \tilde\theta_{i+1}).
$$

Therefore, for c's with the property $\alpha'(c) = p/q$ or, equivalently, $c \in [\beta'(p/q - 0), \beta'(p/q + 0)]$ we have

$$\alpha(c) = c\ \frac{p}{q} - \beta\left(\frac{p}{q}\right).$$

Note that the infimum in (7.5) is achieved on the union of two one-sided minimal configuration $\Theta_a^{c,+} = (\tilde{\theta}_0^a, \ldots)$, $\Theta_a^{c,-} = (\ldots, \tilde{\theta}_0^a)$, $\tilde{\theta}_0^a = a$. These configurations have to be one-sided minimizers, since the only restriction $\tilde{\theta}_0^a = a$ is in the middle of them. Moreover, the same configuration minimizes (7.5) for any $c' \in [\beta'(p/q - 0), \beta'(p/q + 0)]$. Indeed, the sets which minimize c-action and c'-action are the same. Application of (7.1), (7.5), and definitions of α and β-functions show that

$$\left|B_c^{\pm}(a) - B_{c'}^{\pm}(a)\right| < 2|c - c'|. \tag{7.11}$$

Lipschitz dependence on a follows from Lipschitz dependence on a of Peierls's barriers. The latter is proved in [Ma6, Lemma 6.3]. Moreover, Lipschitz constant is 2γ for any rotation number, where γ is the same angle of twisting as in (7.10). This completes the proof of Lemma 7.4. $\square$

Proof of Generalized Aubry Crossing Lemma (GAC Lemma). Non-intersection property has the same proof as Aubry Crossing Lemma (see e.g. [M-F, Section 9]). Therefore, it suffices to prove existence of one of one-sided minimizers passing through a and b, say Θ_a^+ and justify its properties. To prove the existence consider two cases: $a \in \Sigma_\omega$ and $a \notin \Sigma_\omega$.

In the first case we apply Aubry–Mather theorem and Θ_a^+ is just a positive part of minimizer from $\widetilde{\Sigma}_\omega$ passing through a. In the second case denote by $(a_-, a_+) \in \mathbb{T} \setminus \Sigma_\omega$ the maximal interval containing a. Consider the corresponding one-sided minimizers $\Theta_{a_-}^+ = \{\tilde{\theta}_{i-}^a\}_{i \in \mathbb{Z}_+}$ and $\Theta_{a_+}^+ = \{\tilde{\theta}_{i+}^a\}_{i \in \mathbb{Z}_+}$ as in definition (7.7). By non-intersection property if Θ_a^+ exists it should satisfy $\tilde{\theta}_{i-}^a \leqslant \tilde{\theta}_i^a \leqslant \tilde{\theta}_{i+}^a$. It implies that such a minimizer has rotation number ω. Existence now follows from existence of minimizer which realizes minimum in (7.7).

To prove that each one-sided minimizer has rotation number one could use the standard Aubry's Crossing Lemma and integer translations of one-sided Aubry graphs (see e.g. [M-F, Section 11, Trichotomy I]). In order to prove that limit of one-sided minimizer of rotation number ω belongs to Σ_ω consider two cases: rational and irrational. In both the above cases we have $\tilde{\theta}_{i-}^a \leqslant \tilde{\theta}_i^a \leqslant \tilde{\theta}_{i+}^a$ or non-intersection property for all $i \in \mathbb{Z}_+$.

In the irrational case, if Σ_ω is a curve there nothing to prove $a \in \Sigma_\omega$ and if it is a Denjoy–Cantor set, the intervals $(\tilde{\theta}_{i-}^a, \tilde{\theta}_{i+}^a) \pmod 1$, $i \in \mathbb{Z}$ belong to the complement $\mathbb{T} \setminus \Sigma_\omega$. Therefore, the sum of the is at most one and lengths have to tend to zero as the result $|\tilde{\theta}_i^a - \tilde{\theta}_{i-}^a| \to 0$.

In the rational case $\omega = p/q$ we have configurations $\Sigma^{\pm}_{p/q}$ and non-periodic one-side minimizers of rotation number p/q. The latter kind of minimizers have the following property: any one of them cannot cross at least one of minimizers from both $\Sigma^{+}_{p/q}$ and $\Sigma^{-}_{p/q}$, which follows from GAC lemma. □

Lemma 7.4 implies the following:

Corollary 7.5. *Aubry–Mather sets* $\{\Sigma_\omega\}_{\omega\in\mathbb{R}}$, $\Sigma_\omega \subset \mathbb{A}$ *depend continuously on* ω *with respect to Hausdorff distance.*

Proof. One way to prove the lemma is using the facts that Peierls's barrier P_ω vanishes on Σ_ω, its relation with Mather's barrier B_c, and continuity of the latter family. The other is to apply Theorems 11.2 and 11.3 from [M-F]. These two theorems roughly amount to the following. Every minimizer has a rotation number. Limit of minimizers is a minimizer with the limiting rotation number. □

Let I_0^j be given by $h_0'(I_0^j) = j/K$ for each $j \in \mathbb{Z}$. Since h_0 satisfies (5.28), which in turn follows from (4.3), we have that

$$D^{-1} < K|I_0^{j+1} - I_0^j| < D. \tag{7.12}$$

Define collection of restrictions of the EAPT map $\tilde{\Phi}^h$ to "integer" compact parts of the cylinder $\mathbb{A}$. Recall that L is a Lipschitz constant of $\pi^{-1}|_{\pi\Sigma_\omega} : \pi\Sigma_\omega \to \Sigma_\omega$ for the graphs of Aubry–Mather sets (see Graph Theorem Section 6.1). For an integer j define a compact annulus $\mathbb{A}_{j,L,D} = \mathbb{T} \times [I_0^j - L - 1, I_0^j + 1 + L + D]$ and a "local" EAPT map

$$\begin{aligned} &\tilde{\Phi}^h_j : \mathbb{A}_{j,L,D} \to \mathbb{A} \quad \tilde{\Phi}^h(\theta_0, I_0) = (\theta_0', I_0') \quad \text{where} \quad I_0 \in [-L, L+D] \\ &\text{and } (\theta_0', I_0') \ \text{ is given by } \ \tilde{\Phi}^h(\theta_0, I_0) = \tilde{\Phi}^h(\theta_0, I_0 + I_0^j) = (\theta_0', I_0' + I_0^j). \end{aligned} \tag{7.13}$$

Lemma 7.6. *For each integer* j *each Aubry–Mather set* Σ_ω, $\omega \in [j, j+1]$ *belongs to* $\mathbb{A}_{j,L,D}$.

Remark 7. Let $\omega \in [j, j+1]$ and $c \in \mathcal{L}_\beta(\omega)$. Similarly one can show that one-sided c-minimizers also belong to such a compact part of the cylinder with possibly different L. We only use the fact that such L exists.

Proof. First note that by (5.23) the generating function of $\tilde{\Phi}^h$ has the form

$$h(\tilde{\theta}, \tilde{\theta}') = l_0\left(\frac{\tilde{\theta}' - \tilde{\theta}}{K}\right) - \frac{\cos 2\pi\left\{\tilde{\theta} - \alpha\left(l_0'\left(\frac{\tilde{\theta}' - \tilde{\theta}}{K}\right)\right)\right\}}{2\pi K}. \tag{7.14}$$

Denote the second term by $r(\tilde{\theta}, \tilde{\theta}')$. Then we have bounds $\max |r(\tilde{\theta}, \tilde{\theta}')| \leqslant (2\pi K)^{-1}$ and $\max |\partial_{\tilde{\theta}} r(\tilde{\theta}, \tilde{\theta}'), \partial_{\tilde{\theta}'} r(\tilde{\theta}, \tilde{\theta}')| < DK^{-2}$.

To find Σ_j^{per} we need to minimize $\min_{\tilde{\theta}} h(\tilde{\theta}, \tilde{\theta} + j)$. By the form of $h(\tilde{\theta}, \tilde{\theta}')$ the minimum occurs at $\tilde{\theta}^*$ the maximum of $\cos 2\pi(\ldots)$. Therefore, $I' = \partial_2 h(\tilde{\theta}, \tilde{\theta}') = l_0'(j/K)/K = I_0^j$. So Σ_j^{per} is a fixed point with I_0-coordinate I_0^j and by Graph Theorem $\Sigma_j \subset \mathbb{A}_{j,L,D}$. Similarly, $\Sigma_{j+1} \subset \mathbb{A}_{j,L,D}$. It shows that for corresponding β-function we have $\beta'(j/K) = l_0'(j/K) = I_0^j$. By convexity we have $I_0^j < \beta'(\omega) < I_0^{j+1}$ for all $j/K < \omega < (j+1)/K$.

Now we show that $\Sigma_\omega \subset \mathbb{A}_{j,L,D}$ for any $j < \omega < j+1$. If at least one point from Σ_ω belongs to $A_{j,0,D}$, then by Graph Theorem all of Σ_ω in $\mathbb{A}_{j,L,D}$. The proof is by contradiction. Suppose $\Sigma_\omega \cap A_{j,0,D} = \emptyset$. Note that

$$\left| \partial_2 h(\tilde{\theta}, \tilde{\theta}') - \frac{1}{K} l_0' \left(\frac{\tilde{\theta}' - \tilde{\theta}}{K} \right) \right| \leqslant \frac{D}{K^2}.$$

If $\Theta = (\ldots, \tilde{\theta}_i, \ldots) \in \tilde{\Sigma}_\omega$, then it is also c-minimal for some $I_0^j < c < I_0^{j+1}$ of β-function. Compare c-action of this configuration and the straight line one $\Theta^\omega = (\ldots, \tilde{\theta}_i^\omega = \tilde{\theta}_0 + \omega i, \ldots)$. If $\Sigma_\omega \cap A_{j,0,D} = \emptyset$, then $\partial_2 h(\tilde{\theta}, \tilde{\theta}') \notin [I_0^j - 1, I_0^j + 1 + D]$. However,

$$|\partial_2 h(\tilde{\theta}_j, \tilde{\theta}_{j+1}) - c| \notin [-1 + DK^{-2}, 1 - DK^{-2}],$$

$$|\partial_2 h(\tilde{\theta}_j^\omega, \tilde{\theta}_{j+1}) - c| \in [-DK^{-2}, DK^{-2}].$$

Since K is large, relation of $\partial_2 h(\tilde{\theta}, \tilde{\theta}')$ and $l_0'\left(\frac{\tilde{\theta}' - \tilde{\theta}}{K}\right)$ shows that $\partial_2 h(\tilde{\theta}_j, \tilde{\theta}_{j+1}) - c$ has to be always of one side on $[-1 + DK^{-2}, 1 - DK^{-2}]$. This contradicts to c-minimality of Θ and proves the lemma. □

In the next three propositions we state and verify sufficient conditions for existence of ballistic trajectories:

Let $\Phi' : \mathbb{A} \to \mathbb{A}$ be a C^1 EAPT map, which is C^1-close to the EAPT map $\tilde{\Phi}^h : \mathbb{A} \to \mathbb{A}$, given by (5.20). Then for generating h_c' and barriers B_c', $P_{p/q\pm}'$ functions associated to the EAPT map Φ' the following claims hold:

Proposition 7.7. *There is $\Delta > 0$ such that*

$$\inf_{c \in \mathbb{R}} \sup_{a \in \mathbb{T}} B_c'(a) > \Delta, \qquad \inf_{p/q \in \mathbb{Q}} \sup_{a \in \mathbb{T}} \min\{P_{p/q+}'(a), P_{p/q-}'(a)\} > \Delta. \tag{7.15}$$

Remark 8. Let $p, q \in \mathbb{Z}$, $\mathcal{L}_\beta(p/q) = [c^-, c^+]$, $c^- \neq c^+$, and $c \in [c^-, c^+]$. Then

$$B'_c(a) \geqslant \min\{B'_{c^-}(a), B'_{c^+}(a)\} \quad \text{for any } a \in \mathbb{T}. \tag{7.16}$$

To see that, first, note that by (7.1) we have

$$h_c(\tilde\theta_0, \tilde\theta'_0) \in \left[\min h_{c^\pm}(\tilde\theta_0, \tilde\theta'_0), \max h_{c^\pm}(\tilde\theta_0, \tilde\theta'_0)\right]$$

and the same inclusion holds true for any finite or an infinite sum, which has a limit. Now recall that $c^\pm$, $c \in \mathcal{L}_\beta(p/q)$ implies that the values $B'_c(a)$ and $B'_{c^\pm}(a)$ are realized on a minimal configuration from one of the sets $\Sigma^\pm_{p/q}$. This proves the required bound (7.16).

Recall that by [Ma2, Propositions 7.1 and 7.2] we have $B'_{c^\pm} \equiv P'_{p/q\pm}$. Therefore, if (7.15) fails, then it happens for c's whose $(c, \alpha(c))$ is an extremal point of the epigraph of the α-function. By (7.10) it allows to approximate with c^* such that $c^* \in \mathcal{L}_\beta(\omega^*)$ with an irrational ω^* and B'_{c^*} does not satisfy (7.15).

Since a minimizing configuration has to satisfy admissibility conditions of Peierls' barrier (7.7), namely, each $\tilde\theta_j^{a,c}$ (mod 1) belongs to a maximal interval in $\mathbb{T} \setminus \Sigma^c$. Denote by $(\tilde\theta_{j-}^{a,c}, \tilde\theta_{j+}^{a,c})$ the end points of such an interval.

Two propositions below state that for any cohomology class $c \in \mathbb{R}$ and $a \in \mathbb{T}$ we have

Proposition 7.8. *With the notations above for any $\Delta > 0$ there is $N = N(\Delta) \in \mathbb{Z}_+$ such that for some n, n' such that $N \leqslant n, n' \leqslant 2N$ and subindex either $+$ or $-$ we have*

$$\begin{aligned} \left| \sum_{i \geqslant n} \left(h'_c(\tilde\theta_i^{a,c}, \tilde\theta_{i+1}^{a,c}) - h'_c(\tilde\theta_{i\pm}^{a,c}, \tilde\theta_{i+1\pm}^{a,c}) \right) \right| &\leqslant \frac{\Delta}{20}, \\ \left| \sum_{i \leqslant -n'} \left(h'_c(\tilde\theta_i^{a,c}, \tilde\theta_{i+1}^{a,c}) - h'_c(\tilde\theta_{i\pm}^{a,c}, \tilde\theta_{i+1\pm}^{a,c}) \right) \right| &\leqslant \frac{\Delta}{20}. \end{aligned} \tag{7.17}$$

Recall that a finite configuration $\{\tilde\theta_i\}_{i=0}^M$ is minimal subject to fixed end points $\tilde\theta_0 = a^-$ and $\tilde\theta_M = a^+$ if it minimizes $\sum_{i=0}^{M-1} h\left(\tilde\theta_i, \tilde\theta_{i+1}\right)$ over all configurations with $\tilde\theta_0 = a^-$ and $\tilde\theta_M = a^+$. For any cohomology class c let ω be such that $c = \mathcal{L}_\beta(\omega)$ we say that a finite minimal configuration $\{\tilde\theta_i\}_{i=0}^M$ with fixed end points is (c, M)*-admissible* or (c, M)*-admissible minimizer* if

$$|\tilde\theta_M - \tilde\theta_0 - M\omega| < 1. \tag{7.18}$$

Proposition 7.9. *Let ω be such that $c \in \mathcal{L}_\beta(\omega)$. Then for any $\rho > 0$, $N \in \mathbb{Z}_+$ there is $M_0 = M_0(\rho, N) < +\infty$ such that for any $M > M_0$ and any (c, M)-admissible minimizer $\{\tilde{\theta}_i^{a,c,N}\}_{i=0}^N$ there is a one-sided minimizer denoted $\{\tilde{\theta}_i^{a,c,+}\}_{i\in\mathbb{Z}}$ of rotation number ω starting at $\tilde{\theta}_0^{a,c,+} = a$ such that*

$$\max_{0\leqslant i\leqslant 2N} |\tilde{\theta}_i^{a,c,M} - \tilde{\theta}_{i+}^{a,c}| < \rho. \tag{7.19}$$

We shall prove all three propositions in Section 9.

Fix a sequence of strictly positive integers $\{\Delta n_i\}_{i\in\mathbb{Z}}$. Define another sequence of integers $\{n_i\}_{i\in\mathbb{Z}}$ as follows. Let $n_0 = 0$, $n_i = \Delta n_1 + \cdots + \Delta n_i$ for $i > 0$, and $n_i = -\Delta n_0 - \cdots - \Delta n_{i-1}$ for $i < 0$. Call $\{\tilde{\theta}_i^*\}_{i\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ a *connecting configuration* and $\mathcal{N} = \{n_i\}_{i\in\mathbb{Z}}$ *connecting times*. Let $\{\tilde{\theta}_{n_i} = \tilde{\theta}_i^*\}_{i\in\mathbb{Z}}$. For each integer i we choose $\tilde{\theta}_{n_i+1}, \ldots, \tilde{\theta}_{n_{i+1}-1} \subset \mathbb{R}$ so that $\{\tilde{\theta}_j\}_{n_i \leqslant j \leqslant n_{i+1}}$ is minimal subject to fixed end points $\tilde{\theta}_{n_i}$ and $\tilde{\theta}_{n_{i+1}}$. Call a configuration $\{\tilde{\theta}_i\}_{i\in\mathbb{Z}}$ obtained by such a minimization in between connecting times *minimizing the connecting configuration* $\{\tilde{\theta}_i^*\}_{i\in\mathbb{Z}}$. Fix a sequence of cohomology classes $\mathfrak{C} = \{c_i\}_{i\in\mathbb{Z}}$ and $\rho > 0$. Let $N = N(\Delta)$ and $M_0 = M_0(\rho, N)$ be the constants from Propositions 7.8 and 7.9, respectively. Call a configuration $(\mathfrak{C}, \mathcal{N}, \rho, \Delta)$-*admissible* if for each integer i the finite configuration $\{\tilde{\theta}_j\}_{j=n_i}^{n_{i+1}}$ is (c_i, n_i)-admissible minimizer, and $n_i > M_0$.

8. Mather method of changing "remote control" c

The idea of construction of ballistic (or simply diffusing) trajectories (6.1) is to find a variational problem with constrains so that its solution is a stationary configuration. More exactly, using notations above we shall construct sequences of:

- closed one forms $\eta_j = c_j + df_j$, where c_j's are increasing in j and $(c_{j+1} - c_j)$'s are sufficiently small and f_j are C^2 smooth, constant on Σ^{c_j}, and of uniformly bounded C^2-norm in j at the end;
- intervals $I(c_j) \subsetneq \mathbb{T}$ so that $\eta_j|_{I(c_j)} \equiv \eta_{j+1}|_{I(c_j)}$ for every $j \in \mathbb{Z}$. Denote by $\tilde{I}(c_j)$ the lift of $I(c_j)$ such that $\tilde{I}(c_j)$ either contained in $(0, 1)$ or contains zero;
- monotonically increasing moments of time $\cdots < n_{j-1} < n_j < n_{j+1} < \cdots$, $\Delta n_i = n_{i+1} - n_i$, and
- set of integer $\{m_j\}_{j\in\mathbb{Z}}$.

It turns out that it can be done so that the following *variational principle with constrains*:

$$\min \sum_{j\in\mathbb{Z}} \sum_{i=n_j}^{n_{j+1}-1} h_{\eta_j}(\tilde{\theta}_i, \tilde{\theta}_{i+1}), \tag{8.1}$$

provided

$$0 \leqslant \tilde{\theta}_0 < 1, \;\; \tilde{\theta}_{n_j} \in m_j + \tilde{I}(c_j) \quad \text{for all} \quad j \in \mathbb{Z} \tag{8.2}$$

has a solution $\Theta = (\ldots, \tilde{\theta}_i, \ldots)$ in the following sense: a solution or *a minimizer* of a variational principle (8.1) is a configuration Θ such that if $s < k$ and $\Theta' = (\ldots, \theta'_i, \ldots)$ is any sequence satisfying $\theta'_i = \tilde{\theta}_i$ for $n_{s-1} \leqslant i \leqslant n_s$ and $n_k \leqslant i \leqslant n_{k+1}$ and (8.2) for $s \leqslant j \leqslant k$, then

$$\sum_{j=s}^{k} \sum_{i=n_j}^{n_{j+1}-1} h_{\eta_j}(\tilde{\theta}_i, \tilde{\theta}_{i+1}) \leqslant \sum_{j=s}^{k} \sum_{i=n_j}^{n_{j+1}-1} h_{\eta_j}(\theta'_i, \theta'_{i+1}).$$

Show that if a minimizer Θ of a variational principle (8.1) *does not touch boundaries* of constrains, i.e. $\tilde{\theta}_{n_j} \in m_j + \operatorname{int} \tilde{I}(c_j)$, then it corresponds to a stationary configuration and, therefore, by Lemma 6.2 to a trajectory of Φ. Differentiate expression (8.1) with respect to $\tilde{\theta}_{n_j}$ for some j we get

$$\begin{aligned} &\partial_1 h(\tilde{\theta}_{n_j}, \tilde{\theta}_{n_j+1}) + (c_j + df(\tilde{\theta}_{n_j})) \\ &\quad + \partial_2 h(\tilde{\theta}_{n_j-1}, \tilde{\theta}_{n_j}) - (c_{j+1} + df(\tilde{\theta}_{n_j})) = 0. \end{aligned} \tag{8.3}$$

Equality $\eta_j|_{I(c_j)} \equiv \eta_{j+1}|_{I(c_j)}$ implies that the configuration Θ is stationary:

$$\partial_1 h(\tilde{\theta}_{n_j}, \tilde{\theta}_{n_j+1}) + \partial_2 h(\tilde{\theta}_{n_j-1}, \tilde{\theta}_{n_j}) = 0.$$

If a minimizer Θ does hit boundaries of constrains it could have "corners" and the corresponding configuration might not be stationary.

In order to get ballistic trajectories using this approach we prove that one for some $\delta > 0$ and an integer M one can choose $c_j = \delta\ j$ and $n_j \leqslant 2Mj,\;\; j \in \mathbb{Z}$ so that for a certain choice of the other parameters $\{I(c_j),\;\; \eta_j,\;\; m_j\}_{j \in \mathbb{Z}}$ the variational problem (8.1) has a solution $\tilde{\Theta}$ which does not touch boundaries of constrains. In other words, we need to prove that for a certain choice of parameters of the variational problem (8.1) has a solution with no corners.

Proof of no corners along with the choice of parameters of the variational problem (8.1): Let $\Delta > 0$ be a constant from Proposition 7.7, $N = N(\Delta) \in \mathbb{Z}_+$ be an integer from Proposition 7.8 and $\mathcal{L}'_\beta(\omega)$ be Fenchel-Legendre transform of Φ'. Choose $0 < \rho < (40C)^{-1}\Delta$, where C depends only on C^1-norm of Φ' and minimal angle of twisting γ. By Corollary 7.5 for some $\delta' = \delta'(\Delta) > 0$ and any pair of rotation numbers ω, ω' such that $|\omega - \omega'| < \delta'$ we have that $\operatorname{dist}(\Sigma_\omega, \Sigma_{\omega'}) < (20C)^{-1}\Delta$, where dist is Hausdorff distance between sets in $\mathbb{A}$. Using relation between Σ^c and Σ_ω given by Fenchel–Legendre transform (7.3) and Remark 5 we have $\mathcal{L}_\beta(\omega) = c$. By Theorem

6.3 β-function is convex, continuous. Therefore, Lemma 7.6 along with (7.12) imply that $\mathcal{L}'_\beta$ is equicontinuous. Namely, there is $\delta = \delta(\delta') > 0$ such that for any pair of cohomology classes c, c' such that $|c - c'| \leqslant \delta$ we have that $\mathrm{dist}(\Sigma^c, \Sigma^{c'}) < (20C)^{-1}\Delta$. Decrease ρ if necessary so that

$$\sup_{a\in\mathbb{T}} \sup_{|c-c'|<\delta} \max\{|B_c(a) - B_{c'}(a)|,\ |B^\pm_c(a) - B^\pm_{c'}(a)|\} \leqslant \frac{\Delta}{20}.$$

Existence of such ρ for c's on bounded interval it follows from continuity of barrier functions (see Lemma 7.4). To extend to unbounded c-interval notice that family (7.13) of restrictions of EAPT maps on compact c-intervals is C^2 bounded and, therefore, every sequence has a limiting point.

For each cohomology class $c \in H^1(\mathbb{T}, \mathbb{R})$ we define $I(c)$ as follows. By Proposition 7.7 there is $a_c \in \mathbb{T}$ such that $B_c(a_c) > \Delta$. Let (a_c^-, a_c^+) be the maximal interval in the complement of Σ^c. Since $B_c(a_c^\pm) = 0$, by continuity there are $\tilde{a}_c^\pm \in (a_c, a_c^\pm)$ such that $B_c(\tilde{a}_c^\pm) = \Delta/2$ and $B_c(a) > \Delta/2$ for any $a \in (\tilde{a}_c^-, \tilde{a}_c^+)$. Denote by $I(c)$ the interval $(\tilde{a}_c^-, \tilde{a}_c^+) \subset \mathbb{T}$ which does not contain a_c. Denote by $\tilde{I}(c)$ a lift of $I(c)$ to $\mathbb{R}$ such that $\tilde{I}(c)$ either contains zero or is contained in $[0, 1]$. By definition of $I(c)$ and Lipschitz continuity of $B_c(a)$ with respect to a we have that length $\min_c |I(c)|$ is bounded away from zero by $L^{-1}\Delta$, where L is Lipschitz constant of Φ'.

For each integer j choose an integer $d_j > 0$ such that fractional part $(M + d_j)\omega_j - [(M + d_j)\omega_j] < L^{-1}\Delta/2$. By the pigeon hole principle we can choose $d_j < 2L/\Delta$. Choose $M = M_0(\rho, N)$ and $n_j = n_{j-1} + M + i_j,\ \ j \in \mathbb{Z}$ with ρ, N as above and $M_0(\rho, N)$ from Proposition 7.8. Denote $\Delta n_j = n_j - n_{j-1}$.

Let $c_j = \delta j,\ \ j \in \mathbb{Z}$. Define $\{m_j\}_{j\in\mathbb{Z}}$ inductively in j as follows. For each integer j denote ω_j rotation number satisfying $c_j = \mathcal{L}_\beta(\omega_j)$. Let $m_0 = 0$. Choose an integer m_1 (resp. m_{-1}) so that $|m_1 - m_0 - M\omega_1| < L^{-1}\Delta/2$ (resp. $|m_0 - m_{-1} - M\omega_0| \leqslant L^{-1}\Delta/2$). Such an integer is uniquely defined. Choose m_2 (resp. m_{-2}) so that $|m_2 - m_1 - M\omega_2| \leqslant L^{-1}\Delta/2$ (resp. $|m_{-1} - m_{-2} - M\omega_{-1}| \leqslant L^{-1}\Delta/2$). Again m_2 and m_{-2} are uniquely defined and so on. We choose m_j's so that configurations satisfying condition (8.2) could be $(\mathcal{C}, \mathcal{N}, \rho, \Delta)$-*admissible*.

Consider the variational problem (8.1) with above parameters. It always has a solution $\Theta = \{\tilde{\theta}_i\}_{i\in\mathbb{Z}}$. The rest of the proof is by contradiction. Suppose for some j we have a possible corner, i.e. $\tilde{\theta}_{n_j}$ belongs to the boundary of $m_j + \tilde{I}(c_j)$. Below we outline the scheme of the arguments and leave minor technical details to the reader.

By the choice of m_j's both $\{\tilde{\theta}_i\}_{i=n_{j-1}}^{n_j}$ and $\{\tilde{\theta}_i\}_{i=n_j}^{n_{j+1}}$ are $(c_{j-1}, \Delta n_{j-1})$-admissible and $(c_j, \Delta n_j)$-admissible, respectively. By Proposition 7.9 there are one-sided minimizers $\{\tilde{\theta}_i^{a,c_{j-1}}\}_{i\leqslant 0}$ and $\{\tilde{\theta}_i^{a,c_j}\}_{i\geqslant 0}$ both starting at $\tilde{\theta}_{n_j} = \tilde{\theta}_0^{a,c_j,-} = \tilde{\theta}_0^{a,c_{j-1},-}$, respectively. Consider the maximal intervals in the complement of $\Sigma^{c_{j-1}}$ and Σ^{c_j} containing $\tilde{\theta}_{n_j}$. Denote $\tilde{\theta}_{n_j} = b_{j-1} = a_j$ and end points of these intervals by (b_{j-1}^-, b_{j-1}^+) and (a_j^-, a_j^+), respectively. By the choice of c_j's we have that $|a_j^\pm - b_{j-1}^\mp| < (20C)^{-1}\Delta$ up to permutation of $+$ and $-$. By GAC Lemma there are one-sided minimizers $\{\tilde{\theta}_{i\pm}^{b_{j-1}^\pm,c_{j-1}}\}_{i\leqslant 0}$ and

$\{\tilde\theta_{i\pm}^{a_j^\pm,c_j}\}_{i\geqslant 0}$ that do not intersect with the corresponding $\{\tilde\theta_i^{a,c_{j-1}}\}_{i\leqslant 0}$ and $\{\tilde\theta_i^{a,c_j}\}_{i\geqslant 0}$ and start at $a_j^\pm$ and $b_{j-1}^\pm$. By Proposition 7.8 for some $N\leqslant n_\pm\leqslant 2N$ we have that

$$\sum_{i=J}^{-1}\left(h'_{c_{j-1}}(\tilde\theta_i^{b_{j-1},c_{j-1}},\tilde\theta_{i+1}^{b_{j-1},c_{j-1}})-h'_{c_{j-1}}(\tilde\theta_{i-}^{b_{j-1}^-,c_{j-1}},\tilde\theta_{i+1-}^{b_{j-1}^-,c_j})\right)$$

for $J=-n_-$ and $J=+\infty$ the sums differ at most by $\Delta/20$ and

$$\sum_{i=0}^{J}\left(h'_{c_j}(\tilde\theta_i^{a_j,c_j},\tilde\theta_{i+1}^{a_j,c_j})-h'_{c_j}(\tilde\theta_{i-}^{a_j^-,c_j},\tilde\theta_{i+1-}^{a_j^-,c_j})\right)$$

for $J=n_+$ and $J=+\infty$ the sums differ at most by $\Delta/20$. Therefore, with an error of $\Delta/10$ to compare the sum along $\{\tilde\theta_{i-}^{b_{j-1}^-,c_{j-1}}\}_{i\leqslant 0}$ (resp. $\{\tilde\theta_{i-}^{a_j^-,c_j}\}_{i\geqslant 0}$) and $\{\tilde\theta_i^{b_{j-1},c_{j-1}}\}_{i\leqslant 0}$ (resp. $\{\tilde\theta_i^{a_j,c_j}\}_{i\geqslant 0}$) from 0 to n_- (resp. n_+) we could compare the semi-infinite sums from 0 to $-\infty$ (resp. $+\infty$). Semi-infinite sums by minimality represent values of one-sided barriers: $B^-_{c_{j-1}}(a_j)$ (resp. $B^+_{c_j}(a_j)$). By the choice of $|c_j-c_{j-1}|=\delta$ we have that $B_{c_j}(a)=B^+_{c_j}(a)+B^-_{c_j}(a)$ differs from $B^-_{c_{j-1}}(a)+B^+_{c_j}(a)$ by at most $\Delta/20$ for all $a\in\mathbb{T}$. Plug in $a=a_j,\ a_j^-$, and b_{j-1}^-. Recall that $|b_{j-1}^- - a_j^-|<(20C)^{-1}\Delta$. By Lemma = 7.4 $B_c^\pm(a)$ is Lipschitz in a with Lipschitz constant $L\leqslant C$. Therefore, $B^-_{c_j}(b^-_{j-1})+B^+_{c_j}(a_j^-)$ differs from $B_{c_j}(a_j^-)$ by at most $\Delta/20$. Now we need to compare $B_{c_j}(a_j^-)$ and $B_{c_j}(a_j)$. By definition of the barrier B_c the first is zero. By definition of $\tilde I(c_j)$ on the boundary of $=\tilde I(c_j)$ we have $B_{c_j}(a_j)>\Delta/2$. This contradicts minimality of the configuration $\{\tilde\theta_i\}_{i\in\mathbb{Z}}$ at $\tilde\theta_{n_j}$. This proves the existence of a solution to the variational problem (8.1) without corners, i.e. such a solution corresponds to a stationary configuration. Therefore, to a trajectory of the EAPT map Φ'. By the choice of c_j's and n_j's and Lemma 7.6 such a trajectory is ballistic, i.e. satisfies (6.1). This proves Theorem 6.1.

This, in turn, proves Theorem 4.1 for C^∞-perturbations. To prove Theorem 4.2 C^∞-perturbations one needs to change $\{c_j\}_{j\in\mathbb{Z}}$ in the setting of the variational principle so that c_j perform the random walk prescribed by $\{I^k\}_{k\in\mathbb{Z}}$.

To prove Theorems 4.1 and 4.2 for analytic perturbations we use persistency of the invariant cylinder and standard approximation arguments presented in Section 10.

9. Proofs of auxiliary propositions

Proof of Proposition 7.7. First we prove (7.15) for the unperturbed map $\tilde\Phi^h$ and then extend it to its C^1-small perturbations. The proof of (7.15) for $\tilde\Phi^h$ is by contradiction and consists of three steps.

Step 1: Assuming (7.15) fails construct a subsequential C^1-limit $\tilde{\Phi}_*$ of EAPT maps $\{\tilde{\Phi}_j^h\}_j$ and a limit ω_* of fractional parts of rotation numbers $\{\omega_j\}_j$ using the family of restricted EAPT maps (7.13).

Step 2: Show that if each element in the sequence of "local" EAPT maps $\tilde{\Phi}_j^h$ satisfies (5.21) and C^1-converges, then the limiting EAPT map $\tilde{\Phi}_*$ also fulfills (5.21). In particular, it implies that $\tilde{\Phi}_*$ has no invariant curves and, therefore, Mather's criterion of existence of invariant curves gives that for the family barrier functions $\{B_c^*(a)\}_c$ associated to $\tilde{\Phi}_*$ we have

$$\inf_{\omega\in[0,1]} \sup_{a\in\mathbb{T}} B_c^*(a) > \Delta' \quad \text{for some} \quad \Delta' > 0. \tag{9.1}$$

Step 3: Show lower semicontinuity of the left-hand side of the above quantity under C^1-small perturbations.

Step 1: Assume (7.15) fails, then for some sequence of c_j we have that $\sup_{a\in\mathbb{T}} B_{c_j}^*(a) \to 0$. Define ω_j so that $c_j \in \mathcal{L}_\beta(\omega_j)$ for each integer j. By Remark 8 the sequence $\{c_j\}_j$ can be chosen so that the corresponding ω_j's are irrational. Note that if $\Sigma_{\omega_j}^h \subset \mathbb{A}$ is an Aubry–Mather set of $\tilde{\Phi}^h$, then $\Sigma_{j,\{\omega_j\}}^h \subset \mathbb{A}_{L,D}$ is an Aubry–Mather set of rotation number $\{\omega_j\} = \omega_j - [\omega_j]$ of $\tilde{\Phi}_j^h$, as defined in (7.13). Choose a C^1-converging subsequence of $\{\tilde{\Phi}_{j_k}^h\}_k$ and of $\{\omega_{j_k}\}$. By compactness condition (5.28) and definitions (5.18)–(5.20) it is possible. Denote a limiting EAPT map by $\tilde{\Phi}_*^h$ and a limiting rotation number ω'.

Step 2: Since angular functions α and auxiliary EAPT maps $\{\tilde{\Phi}_j^h\}_j$ from (5.18)–(5.19) depend continuously on $h_0'(I_0)$, it implies that the limiting map $\tilde{\Phi}_*^h$ satisfies (5.18)–(5.20), where the corresponding α_*'s and $\tilde{\Phi}_{i,*}^h$ are C^1 limits of the corresponding α_j's and $\tilde{\Phi}_{i,j}^h$'s. It implies that $\tilde{\Phi}_*^h$ satisfies (5.21) and, therefore, does not have invariant curves. Absence of invariant curves, in a view of Mather's criterion of existence of invariant curves (see Corollary 7.2), implies that (9.1) for Peierls' barrier associated to $\tilde{\Phi}_*^h$ holds for some $\Delta > 0$.

Denote $\tilde{P}_{*,\omega^*}(a)$ the Peierls barrier of rotation symbol ω^* associated to the limiting map $\tilde{\Phi}_*^h$.

Step 3: We prove lower semicontinuity of

$$\inf_{\omega\in[0,1]} \sup_{a\in\mathbb{T}} \tilde{P}_{*,\omega^*}(a)$$

with respect to C^1-perturbations.

As we pointed out in Remark 8 it suffices to prove lower semicontinuity for irrational $\omega = \omega^*$'s. Since there is no invariant curves for $\tilde{\Phi}_*^h$, by Structure Theorem each $\Sigma_{*,\omega}$ is a Denjoy set.

Choose large N. Denote by Φ' an EAPT map C^1-close to $\tilde{\Phi}^h_*$ and by $\{P'_{\omega^*}(a)\}_\omega$ the corresponding family of barrier functions. Φ' is close enough to have generating functions of both maps satisfy

$$\sup_{|\tilde{\theta}'-\tilde{\theta}|\leqslant 1} |h'(\tilde{\theta},\tilde{\theta}') - \tilde{h}_*(\tilde{\theta}',\tilde{\theta})| < N^{-3}.$$

Suppose the above infimum for $P'_\omega(a)$ is achieved for an irrational ω'. By Lemma 7.4 there is a configuration $\Theta^{\omega'}_a$, which is the union of one-sided minimizers $\Theta^{\omega',+}_a = (\tilde{\theta}^a_0, \tilde{\theta}^{a,\omega'}_1, \ldots)$, $\Theta^{\omega',-}_a = (\ldots \tilde{\theta}^{a,\omega'}_{-1}, \tilde{\theta}^a_0)$, $\tilde{\theta}^a_0 = a$ minimizing (7.5) for P'_ω. Using the pigeon hole principle we shall decompose sum (7.5) into three parts

$$i \leqslant -n_-, \quad -n_- < i < n_+, \quad n_+ < i$$

so that

$$\left(|\tilde{\theta}^{a,\omega'}_{-n_-} - \tilde{\theta}^{a,\omega'}_{n_+} \ (\mathrm{mod}\ 1)| \right) < 1/N$$

and $N^2 \leqslant n_-, n_+ \leqslant 2N^2$. This would imply that the sum over the first and third part is close to $P'_{\omega'}(\tilde{\theta}^{a,\omega'}_{n_-})$, which is non-negative. Then comparing the sum over the second part $-n_- < i < n_+$ for generating functions h' and $\tilde{h}_*$ evaluated on the same configuration we get that

$$P'_{\omega'}(a) - \tilde{P}_{*,\omega'}(a) \geqslant -6N^{-1}.$$

By taking N large enough we could complete justification of Step 3.

Now we need to select n_- and n_+ with the above properties. By the Structure Theorem $\Sigma'_{\omega'}$ is either a Denjoy set or an invariant curve. Consider the first case. By GAC Lemma configurations admissible for minimization lie in the "holes" of the corresponding Denjoy set $\Sigma'_{\omega'}$ if there are "holes". Those holes have to shrink, since the sum of lengths of all of them is at most 1. This implies that Peierls' barrier minimizing configurations $\Theta^{\omega'}_a$ has to accumulate to $\Sigma'_{\omega'}$ both forward and backward in time. By Structure Theorem dynamics on $\Sigma'_{\omega'}$ is conjugated to circle dynamics and, therefore, is recurrent. Consider two cases: $a \in \Sigma'_{\omega'}$ and $a \notin \Sigma'_{\omega'}$.

Let $a \in \Sigma'_{\omega'}$. Denote by $H'_{\omega'}$ the conjugating map with a circle homeomorphism $\varphi_{\omega'}$ induced by projection from Structure Theorem (see Section 6.1). Consider the trajectory $\{H'_{\omega'}(\varphi^j_{\omega'}(a))\}_{j\in\mathbb{Z}}$ on the circle. By the pigeon hole principle and order preservation there are integers n' and n'' such that $N^2 \leqslant n', n'' \leqslant 2N^2$ and

$$\sum_{1\leqslant j\leqslant N} |H'_{\omega'} \circ \varphi^{-n'+j}_{\omega'}(a) - H'_{\omega'} \circ \varphi^{n''+j}_{\omega'}(a)| < \frac{1}{2N}$$

and intervals

$$[H'_{\omega'} \circ \varphi_{\omega'}^{-n'+j}(a), H'_{\omega'} \circ \varphi_{\omega'}^{n''+j}(a)]$$

are pairwise disjoint for $1 \leqslant j \leqslant N$. Conjugacy implies that corresponding intervals $\{[\varphi_{\omega'}^{-n'+j}(a), \varphi_{\omega'}^{n''+j}(a)]\}_{1 \leqslant j \leqslant N}$ are pairwise disjoint too. Again by the pigeon hole principle for some j^* we also have that

$$|\varphi_{\omega'}^{-n'+j^*}(a) - \varphi_{\omega'}^{n''+j^*}(a)| < 2N^{-2}.$$

This provides a choice of n_- and n_+ equal to n' and n'', respectively, and completes Step 3 for $a \in \Sigma'_{\omega'}$.

Consider $a \notin \Sigma'_{\omega'}$. Select the maximal interval $(a^-, a^+) \in \mathbb{T} \setminus \Sigma'_{\omega'}$ containing a. Go through the above procedure for $a = a^-$. Note also that $\varphi_{\omega'}$ maps maximal intervals in $\mathbb{T} \setminus \Sigma'_{\omega'}$ one into another and preserve their order on the circle. Therefore, the maximal intervals $\{[\varphi_{\omega'}^{-n'+j}(a^-), \varphi_{\omega'}^{-n'+j}(a^+)]\}_{1 \leqslant j \leqslant N}$ are pairwise disjoint too. By choice of admissible configurations for minimization in (7.7) we have that $\tilde{\theta}_n^{a,\omega'} \pmod 1 \in [\varphi_{\omega'}^n(a^-), \varphi_{\omega'}^n(a^+)]$. Again by the pigeon hole principle there is j^* such that

$$|\varphi_{\omega'}^{-n'+j^*}(a) - \varphi_{\omega'}^{n''+j^*}(a)| < N^{-1}.$$

This completes the proof of Step 3 and, therefore, of Proposition 7.7. □

Proof of Proposition 7.8. Fix $\Delta > 0$. Denote by ω rotation number such that $c \in \mathcal{L}'_\beta(\omega)$, where $\mathcal{L}'_\beta$ is Fenchel–Legendre transform associated with Φ'. There are two different cases: ω is irrational or rational p/q with "large" denominator $q > \Delta^{-1}$ and $\omega = p/q$ is rational with "small" denominator $q \leqslant \Delta^{-1}$.

The first case has the proof similar to the proof of Step 3 of Proposition 7.7. Namely, by definition of configurations admissible for minimization in (7.7) they have to lie in the "holes" of the corresponding Denjoy set $\Sigma'_{\omega'}$ if there are "holes". If rotation number is rational with "large" denominator or irrational by the pigeon hole principle one can find n such that $\tilde{\theta}_n^{a,c}$ belongs to a "hole" of size at most Δ. Then we have

$$\left| \sum_{i \geqslant n} h_c(\tilde{\theta}_i^{a,c}, \tilde{\theta}_{i+1}^{a,c}) - \sum_{i \geqslant n} h_c(\tilde{\theta}_{i\pm}^{a,c}, \tilde{\theta}_{i+1\pm}^{a,c}) \right| \leqslant C\Delta^2, \tag{9.2}$$

where C depends on C^1-norm of Φ' and minimal angle of twisting γ. Moreover, this estimate stays valid for one of subindices $\pm$ if $\tilde{\theta}_n^{a,c}$ is Δ-close to one of $\tilde{\theta}_{n\pm}^{a,c}$. Fix subindex—for brevity. To prove (9.2) note that by minimality the left-hand side is bounded by

$$h_c(\tilde\theta_n^{a,c},\tilde\theta_{n+1}^{a,c})+h_c(\tilde\theta_{n-}^{a,c},\tilde\theta_{n+1-}^{a,c})-(h_c(\tilde\theta_n^{a,c},\tilde\theta_{n+1-}^{a,c})+h_c(\tilde\theta_{n-}^{a,c},\tilde\theta_{n+1}^{a,c}))$$
$$=\int_{\tilde\theta_{n-}^{a,c}}^{\tilde\theta_n^{a,c}}\int_{\tilde\theta_{n+1-}^{a,c}}^{\tilde\theta_{n+1}^{a,c}}\partial_{12}h(\tilde\theta,\tilde\theta')\,d\tilde\theta\,d\tilde\theta'.$$

By Graph Theorem each Aubry–Mather set $\Sigma_\omega\subset\mathbb{A}$ is a Lipschitz graph over the base $\mathbb{T}$ and $\Phi':\mathbb{A}\to\mathbb{A}$ is a C^1 smooth map. Therefore, if $|\tilde\theta_{n+}^{a,c}-\tilde\theta_{n-}^{a,c}|<\Delta$ is bound on size of the "hole", then $|\tilde\theta_{n+1+}^{a,c}-\tilde\theta_{n+1-}^{a,c}|<\tilde C\Delta$ for some constant $\tilde C$ depending on Lipschitz constant and C^1-norm of Φ'. This proves (9.2).

Consider now the case of rational number p/q with small denominator $q\leqslant\Delta^{-1}$. First, note that (7.17) holds true for the standard EAPT Φ^*, given by (5.1). Indeed, we have periodicity in I_0. It gives periodicity in c and compactness in both θ_0 and I_0. Therefore, if (7.17) fails, i.e. for any N there is c_N failing it, then by periodicity in c there is a converging subsequence $c_{N_j}\to\tilde c$. But for $\tilde c$ there is $\tilde N$ such that (7.17) holds. This is a contradiction. Thus, there is $N^*\geqslant\tilde N$ such that (7.17) holds for the generating function associated to (5.1).

To prove (7.17) for Φ', which is C^1-close to $\tilde\Phi^h$, given by (5.20), we shall apply similar arguments. Recall that $\widetilde\Sigma_{p/q}$ and $\Sigma'_{p/q}$ denote Aubry–Mather sets of rotation number p/q of EAPT maps Φ' and $\tilde\Phi^h$, respectively.

We start by analyzing (7.17) for the unperturbed map $\tilde\Phi^h$ and $\widetilde\Sigma_{p/q}$. Recall that $c\in\mathcal{L}_\beta(p/q)$ by the standing assumption. Since barriers are continuous with respect to c, without loss of generality we could assume that $c\in\operatorname{int}\mathcal{L}_\beta(p/q)$. Fix a sufficiently small $\delta>0$. If $\tilde\theta_n^{a,c}$ belongs to δ-neighborhood of $\pi_1\widetilde\Sigma_{p/q}$, then by (9.2) we get that the left-hand side of ((7.17), line 1) is bounded by $C\delta^2$. Since δ is small, we have $C\delta^2<\Delta/20$ which gives the required estimate. For each cohomology class c by GAC Lemma there is N_c such that for any $n\geqslant N_c$ we have

$$\operatorname{dist}(\tilde\theta_n^{a,c},\widetilde\Sigma_{p/q})<\delta. \tag{9.3}$$

Now if we assume that $N_{c_j}\to\infty$ as $j\to\infty$ for some sequence $\{c_j\}_{j\geqslant0}$. From this sequence we could choose a converging subsequence in the following sense. Define ω_j so that $c_j\in\mathcal{L}_\beta(\omega_j)$. Proceed similarly to Step 1 of the proof of Proposition 7.7. Note that if $\widetilde\Sigma^h_{\omega_j}\subset\mathbb{A}$ is an Aubry–Mather set of $\tilde\Phi^h$, then $\Sigma^h_{j,\{\omega_j\}}\subset\mathbb{A}_{L,D}$ is an Aubry–Mather set of rotation number $\{\omega_j\}=\omega_j-[\omega_j]$ of $\tilde\Phi^h_j$, defined by (7.13). Choose a C^1-converging subsequence of $\{\tilde\Phi^h_{j_k}\}_k$ and of $\{\omega_{j_k}\}$. By compactness condition (5.28) and definitions (5.18)–(5.20) it is possible. Denote a limiting EAPT map by $\tilde\Phi^h_*$ and a limiting rotation number ω'. For ω', $\tilde\Phi^h_*$, and $c'\in\mathcal{L}'_\beta(\omega')$ there is $N'_{c'}$ such that

(9.3) holds. This is a contradiction. This shows that $\max_c N_c < \infty$. Denote its values by $\tilde{N}$.

To distinguish one-sided minimizers of Φ' and $\tilde{\Phi}^h$ denote them $\{\theta'^{a,c}_i\}_{i\geqslant 0}$ and $\{\tilde{\theta}^{a,c}_i\}_{i\geqslant 0}$, respectively, where a is the starting point and c is the cohomology class. Similarly for negative one-sided minimizers. Generating functions of Φ' and $\tilde{\Phi}^h$ are denoted by h' and h, respectively.

By Structure Theorem there are non-empty sets of periodic points $\widetilde{\Sigma}^{\text{per}}_{p/q} \subset \widetilde{\Sigma}_{p/q}$ and $(\Sigma'_{p/q})^{\text{per}} \subset \Sigma'_{p/q}$. Since we consider only rational p/q with small denominator $q \leqslant \Delta^{-1}$, we could choose Φ' close enough to $\widetilde{\Phi}^h$ so that

$$\sup_{a\in(\Sigma'_{p/q})^{\text{per}}} \operatorname{dist}(a, \widetilde{\Sigma}^{\text{per}}_{p/q}) < \frac{\delta}{4}.$$

Consider δ-neighbourhood of $\widetilde{\Sigma}^{\text{per}}_{p/q}$. Choose $N' \geqslant \tilde{N}$ divisible by q. Show that for any starting point $a \in \mathbb{T}$ and any one-sided c-minimizer $\{\theta'^{a,c}_i\}_{i\geqslant 0}$ for $n \geqslant N'$ we have

$$\operatorname{dist}(\theta'^{a,c}_n, \widetilde{\Sigma}^{\text{per}}_{p/q}) \leqslant 2\delta. \tag{9.4}$$

Suppose it fails. Consider $\{\tilde{\theta}^{a,c}_i\}_{i\geqslant 0}$ and compare $\widetilde{\Sigma}^n_c(\tilde{\Theta}) = \sum_{i=0}^{n-1} h'_c(\tilde{\theta}^{a,c}_i, \tilde{\theta}^{a,c}_{i+1})$ with $\Sigma'^n_c(\Theta') = \sum_{i=0}^{n-1} h'_c(\theta'^{a,c}_i, \theta'^{a,c}_{i+1})$. By GAC Lemma $\theta'^{a,c}_i \to (\Sigma'_{p/q})^{\text{per}}$ as $i \to +\infty$. Let (a^-, a^+) be the maximal interval in the complement of $(\Sigma'_{p/q})^{\text{per}}$ containing a. By Aubry Crossing Lemma $\theta'^{a,c}_{iq}$ monotonically tends to one of the end points $a^{\pm}$ as $i \to +\infty$. Since Φ' is C^1-close to $\tilde{\Phi}^h$, h' and h are C^2-close and, therefore, one-sided c-minimizers $\{\theta'^{a,c}_i\}_{i\geqslant 0}$ of h' are close to $\{\tilde{\theta}^{a,c}_i\}_{i\geqslant 0}$ of h for $0 \leqslant i \leqslant N'$. Therefore, we can make $|\theta'^{a,c}_{N'} - \tilde{\theta}^{a,c}_{N'}| < \delta$. Monotonicity shows that this is true for all $n \geqslant N'$. This proves (9.4).

We could have applied (9.2) to finish the proof now, but it is not possible to replace $\widetilde{\Sigma}^{\text{per}}_{p/q}$ with $\Sigma'_{p/q}{}^{\text{per}}$. In other words, distance between $\theta'^{a,c}_{N'}$ and $\Sigma'_{p/q}{}^{\text{per}}$ could be not close to zero even if N' is large. However, we are still able to show that (7.17) holds for any $n \geqslant N'$. Below we present a configuration $\{\hat{\theta}^{a,c}_i\}_{i\geqslant N'}$, $\hat{\theta}^{a,c}_{N'} = \theta'^{a,c}_{N'}$ satisfying (7.17). This would complete the proof.

Consider the first $nq = N'$ such that $\theta'^{a,c}_{nq}$ is 2δ-close to $\widetilde{\Sigma}^{\text{per}}_{p/q}$. Denote the closest point of $\widetilde{\Sigma}^{\text{per}}_{p/q}$ by $\tilde{a}$. The closest point of $\Sigma'_{p/q}{}^{\text{per}}$ according to our notations is one of $a^{\pm}$. Let it be a^+. Fix a large integer M. Subdivide the interval $(\tilde{a}, a^+)$ into M equal intervals. Denote end points of those intervals

$$a^+_0 = \tilde{a},\ a^+_1, \ldots, a^+_M = a^+.$$

Consider a configuration $\{\hat{\theta}^{a,c}_i\}_{i=0}^{(n+M)q}$, which is the union of minimal configurations connecting $a^+_0 = \theta'^{a,c}_{(n+1)q}$ $a^+_j = \hat{\theta}^{a,c}_{(n+j)q}$ and $a^+_{j+1} = \hat{\theta}^{a,c}_{(n+j+1)q}$. By construction

$\{\hat{\theta}^{a,c}_{(n+j)q}\}_{j=1}^{M+1}$ belong to $\widetilde{\Sigma}^{\mathrm{per}}_{p/q}$. Application of (9.2) gives

$$\left|\sum_{i=0}^{q-1} h_c(\hat{\theta}^{a,c}_{(n+j)q+i}\hat{\theta}^{a,c}_{(n+j)q+i+1})\right| < CM^{-2}.$$

So the sum of M terms is bounded by CM^{-1} and can be made arbitrary small. Now we choose h'_c close enough to h_c so that similar arguments work for the above sum with h_c replaced by h'_c. This shows that the left-hand side of (7.17) is small and completes the proof of Proposition 7.8. $\square$

Proof of Proposition 7.9. First, we fix a cohomology class $c \in \mathbb{R}$ and prove that (7.19) holds with c fixed and $M_0 = M_0(c)$ is replaced by $M_0^c = M_0(c, \rho, N)$. Then we show that $M_0(\rho, N) = \sup_{c\in\mathbb{R}} M_0(c, \rho, N) < +\infty$. Start with the first part.

Since any limit of minimizers is a minimizer, any converging subsequence of (c, M_i)-minimizers is a one-sided minimizer. Suppose for some $\rho_0 > 0$ condition (7.19) fails. Namely, there are $a \in \mathbb{R}$ and a sequence of (c, M_j)-minimizers $\{\theta_i^{a,c,M_j}\}_{i,j\geqslant 0}$ such that for any one-sided minimizer $\{\theta_i^{a,c,+}\}_{i\geqslant 0}$ starting at a for each j and some $0 < i_j < N$ we have

$$|\theta_{i_j}^{a,c,N_j} - \theta_{i_j}^{a,c,+}| > \rho. \tag{9.5}$$

Choosing a subsequence we get contradiction.

To get uniformity of $M_0(c, \rho, N)$ in c we use the sequence of restrictions (7.13) of the EAPT map Φ' to compact annuli. For each sequence satisfying (9.5) with varying c one can choose converging subsequence in the same way as in the proof of Proposition 7.7 Step 1. This gives required contradiction and proves Proposition 7.9. $\square$

10. Persistence of normally hyperbolic invariant manifolds

Recall that the *minimum norm* $m(L)$ of a linear transformation L is defined as

$$m(L) = \inf\{|Lx| : |x| = 1\}.$$

When L is invertible, $m(L) = \|L^{-1}\|^{-1}$. Let $\Phi : M \to M$ be a C^r smooth diffeomorphism of a smooth manifold M and $\mathbb{A}$ be a C^r smooth invariant submanifold of M. Let $T_{\mathbb{A}}M$ be the tangent bundle of M over $\mathbb{A}$. Suppose we have $d\Phi$-invariant splitting into three subspaces

$$T_{\mathbb{A}}M = W^u \bigoplus TV \bigoplus W^s,$$

i.e. for any $x \in M$ we have $d\Phi(x)W_x^s = W_{\Phi(x)}^s$ and $d\Phi(x)W_x^u = W_{\Phi(x)}^u$. Moreover, for some $C, \lambda > 1$ we have $|d\Phi^n(x)v| \geqslant C\lambda^n|v|$ for all $x \in \mathbb{A}$, all $v \in W_x^u$ (resp. W_x^s), and all $n \in \mathbb{Z}_+$ (resp. $n \in \mathbb{Z}_-$). Denote

$$d\Phi^s(x) = d\Phi(x)|_{W^s}, \quad d\Phi^u(x) = d\Phi(x)|_{W^u}, \quad d\Phi^c(x) = d\Phi(x)|_{T_x\mathbb{A}}.$$

Let $0 \leqslant k \leqslant r$. We say that Φ is *k-normally hyperbolic* at $\mathbb{A}$ iff there is a Riemann structure on *TM* such that for all $x \in \mathbb{A}$ we have

$$m(d\Phi_x^u) > \|d\Phi^c(x)\|^k, \quad m(d\Phi_x^c) > \|d\Phi^s(x)\|^k.$$

Let $T_{\mathbb{A}}$ be a tube neighborhood of $\mathbb{A}$ and $\pi_{\mathbb{A}} : \mathbb{T}_{\mathbb{A}} \to \mathbb{A}$ be the natural projection along normal to $\mathbb{A}$ directions.

Theorem 10.1 (*Sacker [Sa], Fenichel [Fe] and Hirsch et al. [H-P-S]*). *In the above notations assume that Φ is uniformly C^r, $r \geqslant 2$ in a neighborhood of $\mathbb{A}$. If Φ has a C^r smooth invariant manifold $\mathbb{A}$ and is k-normally hyperbolic at $\mathbb{A}$, then any C^r smooth Φ', which is C^r close to Φ, has the following property: Φ' has a C^k smooth invariant manifold $\mathbb{A}'$, which is C^k close to $\mathbb{A}$, and, therefore, the natural projection along normal to $\mathbb{A}$ directions $\pi|_{\mathbb{A}'} : \mathbb{A}' \to \mathbb{A}$ is a C^k diffeomorphism close to identity. Moreover, Φ' is k-hyperbolic at $\mathbb{A}'$ and $\pi \circ \Phi' \circ (\pi|_{\mathbb{A}'})^{-1} : \mathbb{A} \to \mathbb{A}$ is C^k close to Φ.*

11. Approximation by analytic Hamiltonians

Recall the standard result about approximation of smooth functions by analytic ones (see e.g. [S-Z]):

Lemma 11.1. *There is a family of convolution (or smoothing) operators of $C^0(\mathbb{R}^n)$ into the linear space of entire functions on $\mathbb{C}^n$*

$$S_r f(\eta) = r^{-n} \int_{\mathbb{R}^n} K(r^{-1}(\eta - \xi)) f(\xi)\, d\xi, \quad 0 < r \leqslant 1$$

such that for every $l > 0$ there exist a constant $c = c(l) > 0$ with properties: if $f \in C^l(\mathbb{R}^n)$, then for $|\alpha| \leqslant l$ and $0 < \rho \leqslant r$ we have

$$|\partial^\alpha S_r f - \partial^\alpha S_\rho f|_r \leqslant c|f|_{C^l} r^{l-|\alpha|}.$$

If f is periodic in some variables, then so are the approximating functions $S_r f$ in the same variables.

Choose $\rho > 0$ sufficiently small. Consider the perturbation

$$\varepsilon f(\theta, I_0, t) = f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ f_2(\theta_2^\Lambda, \dots, \theta_d^\Lambda, t)$$

from Proposition 5.2. We could approximate C^∞ functions f_1 and f_2 by analytic functions $f_1' = S_\rho f_1$ and $f_2' = S_\rho f_2$, respectively. Application of Theorem 10.1 gives that approximation f_2' satisfies $\|f_2' - f_2\|_{C^1} < c\|f_2\|_{C^d}\rho^{d-1}$. So choosing ρ small enough we get, say, $\exp(-3d)$-close C^1 approximation. Notice that for large $|I_0|$, say $|I_0| \approx R \gg 1$, one could rescale I_0 to $R\tilde{I}_0$, the unperturbed Hamiltonian $h(R\tilde{I}_0, I_1, \dots, I_d)$ to $R^2 h_R(\tilde{I}_0, I_1, \dots, I_d)$, and the I_0's part of the perturbation $f_1(\theta_0, R\tilde{I}_0, \theta_1^\Lambda, t)$ to $R^2 f_1^R(\theta_0, \tilde{I}_0, \theta_1^\Lambda, t)$ and still get perturbation of order of ε (see (5.17) and (5.27)). Standard calculation show that such a C^1-close approximation has a persistent 1-normally hyperbolic cylinder $\mathbb{A}_\Lambda \subset \mathbb{T}^{d+1} \times \mathbb{R}^{d+1} \times \mathbb{T}$ which is C^1-close to $\mathbb{A}_\Lambda$. If the time-K-map Φ of the corresponding Hamiltonian equation of

$$H(\theta, I, t) = h(I) + f_1(\theta_0, I_0, \theta_1^\Lambda, t)\ f_2(\theta_2^\Lambda, \dots, \theta_d^\Lambda, t),$$

then it has 1-hyperbolic invariant cylinder $\mathbb{A}'_\Lambda$ which is C^1-close to $\mathbb{A}_\Lambda$. By Theorem 10.1 the restriction of Φ' onto $\mathbb{A}$ of the form $\Phi = \pi \circ \Phi' \circ (\pi|_{\mathbb{A}'_\Lambda})^{-1} : \mathbb{A} \to \mathbb{A}$ is C^1 close to Φ^*. Therefore, by Theorem 6.1 the corresponding C^1 smooth EAPT $\Phi : \mathbb{A} \to \mathbb{A}$ has ballistic trajectories. To check that H can be approximated well enough by an analytic perturbation we apply Proposition 5.2 and Lemma 11.1. This proves Theorem 4.1. □

As we pointed out in Section 8 to prove Theorem 4.2 one needs to change $\{c_j\}_{j\in\mathbb{Z}}$ in the setting of the variational principle (8.1) so that c_j perform the random walk prescribed by $\{I^k\}_{k\in\mathbb{Z}}$.

Acknowledgments

V. Kaloshin expresses sincere gratitude to John Mather for teaching his theory and long-lasting support. The second author is partially supported by AIM and Sloan fellowships and NSF Grants.

References

[Arn] V. Arnold, Mathematical Methods in Classical Mechanics, Graduate Text in Mathematics, vol. 60, Springer, Berlin, 1989.

[Ba1] D. Bambusi, Long time stability of small amplitude solutions in nonlinear Schrödinger equations, CMP 189 (1997) 205–226.

[Ba2] D. Bambusi, Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations, Math. Z. 230 (1999) 345–387.

[Ba] V. Bangert, Mather sets for twist maps and geodesics on tori, Dynamics Reported 1 (1998) 1–56.

[Bo] J. Bourgain, On diffusion in high-dimensional Hamiltonian systems and PDE, J. Analyse Math. 80 (2000) 1–35.

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

[F-S-W] J. Fröhlich, T. Spencer, C. Wayne, Localization in disordered, nonlinear dynamical systems, J. Stat. Phys. 42 (3–4) (1986) 247–274.

[Go] Ch. Golé, Symplectic Twist Maps, Global Variational Techniques, Advanced Series in Nonlinear Dynamics, vol. 18, 2001.

[H-P-S] M. Hirsch, C. Pugh, M. Shub, Invariant Manifolds, Lecture Notes in Mathematics, vol. 583, Springer, Berlin, New York, 1977.

[L-N] P. Lochack, A. Neishtadt, Estimates in the theorem of N.N. Nekhorosheff for systems with quasi-convex Hamiltonian, Chaos 2 (1992) 495–499.

[M-S] J.-P. Marco, D. Sauzin, Stability and instability for Grevrey quasi-convex near-integrable Hamiltonian systems, Publ. Math. IHES 96 (2002) 199–275.

[MS2] J.-P. Marco, D. Sauzin, Wandering domains and random walks, 2002, preprint.

[Ma6] J. Mather, Modulus of Continuity for Peierls's Barrier, NATO Advances Scientific Institute Series C Mathematical Physics Sciences, vol. 209, Reidel, Dordrecht, 1987, pp. 177–202.

[Ma5] J. Mather, Differentiability of the minimal average action as a function of the rotation number, Bol. Soc. Brasil. Mat. (N.S.) 21 (1) (1990) 59–70.

[Ma1] J. Mather, Variational construction of orbits of twist diffeomorphisms, J. AMS 4 (2) (1991) 169 –207.

[Ma2] J. Mather, Variational construction of connecting orbits, Ann. Inst. Fourier (Grenoble) 43 (5) (1993) 1349–1386.

[Ma3] J. Mather, Variational construction of trajectories for time periodic Lagrangian systems on the two torus, unpublished, 1996.

[Ma4] J. Mather, Arnold Diffusion, I: Announcement of results, Kluwer Academic Plenum Publication Series, Dordrecht, Math. Sci., 2004, 20pp.

[M-F] J. Mather, G. Forni, Action-Minimizing Orbits in Hamiltonian Systems, Transition to Chaos in Classical and Quantum Mechanics (Montecatini Terme, 1991), Lecture Notes in Mathematics, vol. 1589, pp. 92–186.

[Mc-S] D. McDuff, D. Salamon, Introduction to Symplectic Topology, Oxford Mathematical Monographs, 1995.

[Mo] J. Moser, Monotone twist mappings and the calculus of variations, Ergodic Theory Dynamical Systems 6 (1986) 401–413.

[Po] J. Poschel, Nekhoroshew estimates for quasi-convex Hamiltonian systems, Math. Z. 213 (1993) 187–216.

[Sa] R. Sacker, A perturbation theorem for invariant manifolds and Holder continuity, J. Math. Mech. 18 (1969) 705–762.

[S-Z] D. Salamon, E. Zehnder, KAM theory in configuration space, Comment. Math. Helv. 64 (1) (1989) 84–132.

[Xia] Z. Xia, Arnold diffusion and instabilities in Hamiltonian dynamics, preprint, 2002, 34pp.