J. Eur. Math. Soc. 9, 905–956

DOI 10.4171/JEMS/683

Marian Gidea · Rafael de la Llave

Perturbations of geodesic flows by recurrent dynamics

To the memory of John N. Mather

Received April 22, 2014 and in revised form July 8, 2015

Abstract. We present a mechanism for Arnold diffusion based on intertwining homoclinic orbits to a normally hyperbolic invariant manifold, followed for long time intervals, with orbits of the dynamics restricted to that manifold, followed for short time intervals. The resulting trajectories are rather fast, and their construction is explicit, so they can be used in concrete applications.

The method to construct such orbits relies mainly on correctly aligned windows and does not require the use of delicate analytical methods such as averaging or KAM, nor the convexity assumptions of variational methods. It only requires control of the inner dynamics over short times, and so it allows easy verification of the hypotheses with finite precision numerical calculations.

As an illustration of this mechanism, we consider a geodesic flow on a compact manifold with a generic (Riemannian, Finsler or Lorentz) metric, subject to time-dependent perturbations via a generic potential. The genericity conditions are given explicitly. The perturbation is driven by a flow on an external manifold which satisfies some mild recurrence condition. Periodic and quasi-periodic perturbations are included, and there are many others. We show the existence of trajectories whose energy grows to infinity, as well as of trajectories that follow prescribed energy paths. The trajectories can be constructed to diffuse at speeds that match upper bounds.

While the study of perturbations of the geodesic flow is a classical problem, we use it mainly to showcase the main ingredients of the method in a simple way. Other results related to ours exist in the literature, but the trajectories we obtain are different from the previously constructed ones. Our approach is flexible enough to be widely applicable; some applications to the restricted three-body problem involving also numerical calculations are in progress.

Keywords. Mather acceleration theorem, Arnold diffusion, shadowing

1. Introduction

1.1. The Hamiltonian instability problem

The phenomenon of instability in nearly integrable Hamiltonian systems was discovered in [4], which described a fundamental mechanism based on whiskered tori and verified it in an example. Conceptually, this mechanism showed how to increase the energy of a

M. Gidea: Department of Mathematical Science, Yeshiva University, New York, NY 10018, USA; e-mail: Marian.Gidea@yu.edu

R. de la Llave: School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA; e-mail: rafael.delallave@math.gatech.edu

Mathematics Subject Classification (2010): Primary 37J40, 37C50, 37C29; Secondary 37B30

conservative mechanical system by using small, time-periodic perturbations. See Subsection 1.4 for a brief overview of the problem.

The main idea of this paper is that we can show Hamiltonian instability for rather general time-dependent perturbations. To establish instability we rely almost exclusively on homoclinic excursions to a normally hyperbolic invariant manifold (NHIM), through carefully choosing the times when to jump onto the homoclinics. By considering multiple homoclinic intersections we can always choose to jump onto the homoclinic that yields a larger energy growth. Between two consecutive jumps, we need to control the dynamics on the manifold only for short time periods. Hence, we do not need to analyze the dynamics on the manifold very much. We do not need to use either KAM theory or even first order averaging theory.

To illustrate the approach we study a model consisting of a geodesic flow perturbed by a time-dependent potential, sometimes known as the 'Mather acceleration theorem' [64]. As a departure from the original setting, we assume only some mild recurrence condition on the time dependence. The main assumptions can be verified in concrete examples by finite calculations. We do not need to assume that the perturbation is periodic or quasi-periodic as in previous works (e.g., [11, 51, 72]). We note that time-dependent perturbations that are not periodic or quasi-periodic are also considered in [43] through a different mechanism that involves first order averaging.

As already mentioned, the main contribution of this paper lies in the method, which has wider applicability. Indeed, we anticipate several other applications forthcoming [15, 46].

In spite of the simplicity of the tools used, we show that the method can produce very quantitative results, including optimal diffusion times, a result which has also been obtained by a variety of other methods for periodic or quasi-periodic systems (but, of course, leading to different orbits; see, for example, [72, 26, 43, 79]). Even if such a time estimate has been proved in some cases (and we do not consider it a novelty of this paper), we find it encouraging that one can use elementary tools to produce quantitative results even in systems with no averaging theory.

Remark 1.1. Note that for the models considered in this paper we cannot apply classical averaging theory in the extended system, as in [29, 30]; of course, averaging over fast variables, as in [43], can be done, but even this is not necessary for our mechanism. (Good expositions of classical averaging theory are in [60, 5].)

We do not know if the ergodic averaging theory [3] is enough to obtain geometric results.

Also, we do not know whether variational methods can be applied to general time-dependent systems. Some of the geometric features of minimizers well known for periodic perturbations have counter-examples for quasi-periodically forced systems [57]. Of course, variational methods work under the assumption that the system is convex (i.e., its Hessian matrix is positive definite), but this is not required by our method. In this paper, we do not assume that the metric defining the geodesic flow is Riemannian, and the arguments apply just as well to Finsler or Lorentz metrics.

Remark 1.2. Even in the particular cases of periodic and quasi-periodic potentials, the orbits constructed here seem to be different from previously constructed ones. The orbits we construct are most of the time performing homoclinic excursions and do not spend much time near the normally hyperbolic manifold. They are different from other orbits in other mechanisms where the energy gain happens near hyperbolic periodic orbits or where orbits spend a significant time near the NHIM.

1.2. Perturbed geodesic flow model

The explicit model that we treat in this paper is the following. We consider an (unperturbed) geodesic flow on the unit tangent bundle of a compact manifold. We assume that this geodesic flow has a hyperbolic periodic orbit whose stable and unstable manifolds intersect transversally. Via a potential, we couple the dynamics of the geodesic flow with that of an external flow on some other compact manifold. We assume that the external flow has a non-trivial uniformly recurrent point. We show that if the coupling between the geodesic flow and the external flow satisfies some non-degeneracy conditions, which we formulate explicitly, then the resulting system possesses some orbits along which the energy grows to infinity linearly in time; such growth rate is optimal. Also, we show that there exist orbits whose energy follows a prescribed energy path.

We point out that the explicit assumptions on the geodesic flow, on the external dynamics, and on the coupling, are very general. They hold generically and can be checked in concrete examples by a finite precision calculation. Also, the construction of orbits whose energy grows to infinity at an optimal rate, or of orbits whose energy behaves in some prescribed way, is explicit, a fact which can be potentially exploited in concrete applications, such as in celestial mechanics.

A particular class of systems that we refer to in this paper are quasi-periodic perturbations of the geodesic flow, which have been considered in the literature [30]. Nevertheless, the mechanisms we use are very different from those of the previous papers, and the orbits produced are different as well. This is relevant for applications, when one is not only interested in establishing the presence of the 'diffusion' phenomenon, but also in exploring various diffusion mechanisms which can feature different characteristics [77].

An intuition to keep in mind is that the unperturbed geodesic flow describes the dynamics of a mechanical system with a Maupertuis metric. Its energy is conserved. The external flow, which is governed by its own intrinsic dynamics, exerts a time-dependent perturbation on the mechanical system. For the perturbed system, the energy conservation law does not hold in general. If the external flow moves with some frequency, it seems conceivable that by choosing trajectories of the geodesic flow that are in resonance with that frequency one can obtain unbounded energy growth for the perturbed geodesic flow.

The striking conclusion of this paper is that one can achieve unbounded energy growth even when the external flow has no frequency of motion. The only assumption we use is recurrence. We can paraphrase this situation:

A little recurrence goes a long way.

1.3. Contents of the paper

The set-up and main results appear in Section 2. The proofs of the main results are based on a new mechanism combining geometric methods, perturbation theory, and topological techniques. Since all the constructions are rather explicit, we anticipate that this can be used in concrete applications, such as in celestial mechanics.

Section 3 describes the geometric method that is used in the paper. The perturbed system is first expressed as a slow and small perturbation of an integrable Hamiltonian, through some rescaling of the coordinates and of the time. Due to the assumptions on the geodesic flow, the unperturbed system possesses a normally hyperbolic invariant manifold whose stable and unstable manifolds intersect transversally, at various places. The transverse homoclinic intersections are used to define different 'scattering maps'. A scattering map is defined on a subset of the normally hyperbolic invariant manifold, and encodes information on the homoclinic trajectories. For all sufficiently small perturbations—which in our case amounts to all sufficiently large energies—the normally hyperbolic invariant manifold from the unperturbed system survives to the perturbed system, and its stable and unstable manifolds keep intersecting transversally. This allows us to continue the scattering maps from the unperturbed problem to the perturbed one. The effect of a scattering map on the energy of the system can be computed very explicitly using the fact that the energy is a slow variable.

In Section 4 we show that, by interspersing the 'outer' dynamics, given by a scattering map, with the 'inner' dynamics, given by the restriction of the flow to the normally hyperbolic invariant manifold, we can arrange that the energy changes by arbitrarily large quantities, and in particular grows to infinity.

At the first stage, we construct some elementary building blocks for the two-map dynamics, each block consisting of one application of a scattering map followed by a segment of a trajectory of the inner dynamics. We can compute explicitly the change of energy along such an elementary building block.

At the second stage, we show that, under appropriate conditions, we can construct sequences of elementary building blocks whose energy experiences the desired changes. We also show that these effects do not happen with one scattering map, but it is crucial to use at least two scattering maps. In particular, we show that we can arrange the choices of the scattering maps and of the corresponding blocks so that we can consistently increase the energy of the system, at an optimal rate.

These sequences of blocks determine pseudo-orbits, which are concatenations of orbits of the inner dynamics and of orbits of the outer dynamics. They do not immediately yield true orbits for the perturbed geodesic flow.

To prove the existence of true orbits, in Section 5 we apply the topological method of correctly aligned windows. Around the pseudo-orbits we construct windows that are correctly aligned, i.e., sequences of multi-dimensional rectangles that successively cross one another in a non-trivial way under the inner or the outer dynamics. A topological version of the Shadowing Lemma implies the existence of true orbits near those pseudo-orbits. We note that the main difficulty in carrying out the shadowing argument is that the underlying dynamical system is not hyperbolic, as there are neutral directions along

the normally hyperbolic invariant manifold. Hence the customary hyperbolic shadowing methods do not apply.

The orbits constructed in our paper are very different from the orbits previously considered in the literature. The orbits we consider stay very little time near the unperturbed periodic orbits between successive homoclinic excursions. This has several technical consequences: we do not need to rely on the KAM theorem, on Aubry–Mather theory, or on averaging, in order to control the behavior of the inner dynamics (the dynamics near the closed orbits between homoclinic excursions). Avoiding the use of averaging theory allows us to consider very general perturbations. Also, we do not need to study the geometric details on how the inner dynamics is affected by the perturbation. On the other hand, we do have to rely more heavily on the scattering map, and to develop some sophisticated shadowing mechanisms that are not based on hyperbolicity but rather on topological tools.

1.4. Some other related work

The Arnold mechanism of diffusion based on whiskered tori formed the basis of numerous subsequent developments, including incorporating variational methods, e.g., in [19, 9, 40]. A remarkable surge of interest in the problem was stimulated by the unpublished work [64], which led to several new approaches – variational, e.g., [64, 65, 66, 17, 18, 16], geometric, e.g., [11, 31, 51, 58, 43, 72, 26], or mixed, e.g., [7, 52, 53]. Instability mechanisms in multi-dimensional settings have been studied in [86, 12, 33]. Proofs for the Arnold conjecture for convex nearly integrable systems of two and a half and three degrees of freedom were announced in [16, 53]. Due to its importance to practical applications, the instability phenomenon has also been studied by physicists using heuristic and numerical methods, e.g., in [20, 84, 56, 77].

The geometric approach of [29, 30, 45, 47] does not only use whiskered tori but also normally hyperbolic invariant manifolds whose stable and unstable manifolds intersect (see, e.g., [29, 45, 47]). The method of [45] bypasses whiskered tori completely. The mechanism is based on interleaving homoclinic excursions (conveniently quantified by the scattering map) with the dynamics on the NHIM.

In the mechanisms discussed in the above papers [29, 30, 45, 47], there are two time scales. The homoclinic excursions are fast while the dynamics in the NHIM has slow components. If one assumes that the perturbations are periodic (or quasi-periodic), the inner dynamics can be controlled using the classical averaging method. Other methods based on normally hyperbolic manifolds, relying on the separatrix map rather than on the scattering map, are in [79, 73].

Results related to the ones considered in this paper were also proved in [10, 8, 25, 24, 62, 59, 79, 72, 73]. See also [11, 80, 81]. The existence of orbits whose energy grows in time at different rates, depending on the type of perturbation, was proved in [26, 79] for general Hamiltonian systems that slowly depend on time.

A result related to ours has been obtained in [43], where the time-dependent perturbation is also not assumed to be periodic or quasi-periodic, and not even recurrent. The main difference is that their mechanism to gain energy relies on following certain hyperbolic periodic orbits for a long time, while our mechanism does just the opposite, spending

as little time as possible near hyperbolic periodic orbits and jumping between such orbits most of the time. Implicitly, the mechanism in [43] involves averaging near periodic orbits while ours requires little control on the inner dynamics.

A general discussion on the speed of diffusion in a priori unstable Hamiltonian systems appears in [37]; in particular, that paper considers diffusion in the planar elliptic restricted three-body problem. The diffusion phenomenon in the case of the spatial circular restricted three-body problem is considered in [27, 28].

2. Set-up and main results

2.1. The geodesic flow

In this subsection we review some well known facts on the geodesic flow, and we set up the notation.

Let M be an n-dimensional C^r-smooth compact manifold and g be a C^r Riemannian metric on M, where $r \geq r_0$. (Some non-optimal values of r_0 are discussed in Subsection 5.7 at the end of the proofs of the main results.) The geodesic flow $\xi : TM \times \mathbb{R} \to TM$ is the flow on the tangent bundle TM (of dimension $2n$), defined by

$$\xi_t(x, v) = (\xi_{x,v}(t), (d\xi_{x,v}/dt)(t)),$$

where $\xi_{x,v}$ is the unique geodesic with $\xi_{x,v}(0) = x$ and $(d\xi_{x,v}/dt)(0) = v$. To a geodesic curve ξ on M corresponds a trajectory of the geodesic flow given by

$$t \mapsto (\xi(t), (d\xi/dt)(t)).$$

In this paper we do not distinguish between a geodesic curve and the corresponding trajectory of the geodesic flow. Also, we do not distinguish between a parametrized curve and its image.

Since the speed along a geodesic is constant, every geodesic can be reparametrized as a unit speed geodesic. This way we reduce the study of the geodesic flow to its restriction to the unit tangent bundle T^1M (of dimension $2n - 1$).

Using the standard identification between TM and T^*M given by g, we can interpret the geodesic flow as the Hamiltonian flow for the Hamiltonian $H_0 : T^*M \to \mathbb{R}$ given by

$$H_0(x, y) = \tfrac{1}{2} g_x(y, y). \tag{2.1}$$

On T^*M we consider the standard symplectic form $\omega = dy \wedge dx$.

Since H_0 is independent of time, it is a conserved quantity. The energy surfaces $\Sigma_E = \{H_0 = E\}$ are invariant under ξ_t. We will denote by $\xi_{E,t}$ the flow restricted to the energy surface Σ_E. The restriction $\xi_{1/2,t}$ of ξ_t to the energy manifold $\Sigma_{1/2} = \{H_0 = 1/2\}$ corresponds to the geodesic flow on T^1M.

It is clear that the flow restrictions to different energy surfaces are equivalent to one another and can be obtained just by a rescaling of time and momentum. The mapping $D_{\sqrt{2E}}(x, y) = (x, \sqrt{2E}\, y)$—dilation along the fibers of T^*M, which is a well defined

operation, because the fibers are linear spaces—gives a diffeomorphism between $\Sigma_{1/2}$ and Σ_E. Furthermore

$$\xi_{E,t} \circ D_{\sqrt{2E}} = D_{\sqrt{2E}} \circ \xi_{1/2,\sqrt{2E}\,t}, \qquad \xi_t \circ D_{\sqrt{2E}} = D_{\sqrt{2E}} \circ \xi_{\sqrt{2E}\,t}. \tag{2.2}$$

Explicitly, the trajectory of the geodesic flow $\xi_E = (\xi_E^x, \xi_E^y)$ at energy E is related to the trajectory of the geodesic flow $\xi_{1/2} = (\xi_{1/2}^x, \xi_{1/2}^y)$ at energy $1/2$ by the formula

$$(\xi_E^x(t), \xi_E^y(t)) = \big(\xi_{1/2}^x(\sqrt{2E}\cdot t), \sqrt{2E}\cdot \xi_{1/2}^y(\sqrt{2E}\cdot t)\big). \tag{2.3}$$

Below we will regard the geodesic flow as an unperturbed dynamical system, to which we will apply an external perturbation.

2.2. *Assumptions on the geodesic flow*

We make the following assumptions on the geodesic flow.

A1. There exists a closed, unit speed geodesic λ, which is a hyperbolic periodic orbit for the geodesic flow on T^1M.

A2. There exists a unit speed geodesic γ which is a transverse homoclinic orbit to λ for the geodesic flow on T^1M.

Note that condition *A1* means that λ has stable and unstable manifolds $W^s(\lambda)$ and $W^u(\lambda)$, of dimension n, in T^1M. The fact that a geodesic ξ is unit speed means that $g(d\xi/dt, d\xi/dt) = 1$. Without loss of generality, we can assume that the period of the geodesic λ from assumption *A1* is 1.

The transversality condition *A2* means that

$$T_z W^s(\lambda) + T_z W^u(\lambda) = T_z(T^1M) \quad \text{for all } z \in \gamma.$$

We note that there are many manifolds for which conditions *A1* and *A2* are satisfied for an abundance of Riemannian metrics; a partial review of the existing results is provided in [30]. As an example, any surface of genus 2 or higher, with any $C^{2+\delta}$ metric, $\delta > 0$, has hyperbolic geodesics with transverse homoclinic connections [54]. A very general result was obtained in [22], showing that on any closed manifold M with $\dim(M) \geq 2$ the set of C^∞ Riemannian metrics whose geodesic flow contains a non-trivial hyperbolic basic set is C^2-open and C^∞-dense. This implies that hypotheses *A1* and *A2* above hold for a C^2-open and C^∞-dense set of Riemannian metrics.

Note that the energy H_0 along both the hyperbolic orbit λ in *A1* and the homoclinic orbit γ in *A2* equals $1/2$, but similar orbits exist in all energy surfaces because of (2.3).

2.3. *Coupling the geodesic flow with an external dynamical system*

Next we will describe a class of perturbations of the geodesic flow for which we will show the existence of orbits with unbounded growth of energy over time, as well as of symbolic dynamics.

We first describe an external dynamical system which we couple with the geodesic flow through a time-dependent potential. Let $X : N \to TN$ be a C^1-smooth vector field on a compact d-dimensional manifold N. Let χ_t be the flow on N associated to X. Let $\theta_0 \in N$ and let $\chi_t(\theta_0)$ be an integral curve to X with the initial condition $\theta_0 \in N$.

Let

$$\mathcal{V} = \{V : M \times N \to \mathbb{R} \mid V \text{ is } C^1 \text{ in } (x, \theta) \in M \times N \text{ and } C^r \text{ in } x \in M\}.$$

This is a Banach space with the norm

$$\|V\|_{1;r} = \sum_{j=0,1} \sup_{(x,\theta)} \|D^j_{(x,\theta)} V(x, \theta)\| + \sum_{j=2,\dots,r} \sup_{(x,\theta)} \|D^j_x V(x, \theta)\|, \tag{2.4}$$

and we will refer to the induced topology on $\mathcal{V}$ as the $C^{1;r}$-*topology*.

We think of $V \in \mathcal{V}$ as a potential depending on the parameter θ evolving in N.

For every $V \in \mathcal{V}$ we consider a parameter-dependent, time-dependent Hamiltonian $H_{\theta_0} : T^*M \times \mathbb{R} \to \mathbb{R}$ given by

$$H_{\theta_0}(x, y, t) = H_0(x, y) + V(x, \chi_t(\theta_0)). \tag{2.5}$$

2.4. *Assumptions on the external dynamical system*

We now describe some conditions on the external dynamical system.

Given (N, χ), a point $\theta_0 \in N$ is said to be *uniformly recurrent* (or *syndetically recurrent*, or *almost periodic*) if for every open neighborhood U of θ_0 there exists $T > 0$ such that, every interval (a, b) with $b - a > T$ contains a time t with $\chi_t(\theta_0) \in U$. That is, a uniformly recurrent point is one which is recurrent with 'bounded return times'.

The flow χ on N is said to be *minimal* if every orbit is dense in N.

Since N is compact, if the flow is minimal then for every open set $O \subseteq N$ there exists $T \geq 0$, depending on O, such that $\bigcup_{t\in[0,T]} \chi_t(O) = N$.

Since N is compact, there always exists a point $\theta_0 \in N$ that is uniformly recurrent for (N, χ). This follows from the Poincaré Recurrence Theorem.

Moreover, if (N, χ) is minimal then every point $\theta \in N$ is uniformly recurrent.

Conversely, if $\theta_0 \in N$ is uniformly recurrent then its orbit closure is a minimal set.

If every point $\theta \in N$ is uniformly recurrent then N is the disjoint union of its minimal subsystems, in which case N is called *semisimple* (see [42]).

The following two alternative conditions will be used in the statements of the main results. *A3′* is clearly stronger than *A3* but leads to stronger results.

A3. The flow χ_t has a uniformly recurrent point $\theta_0 \in N$ with $X(\theta_0) \neq 0$.
A3′. The flow χ_t is minimal on N.

2.5. *Statement of the results*

The following result provides the existence of trajectories with unbounded growth of energy for the system (2.5).

Theorem 2.1. *Let g be a Riemannian metric on M satisfying conditions A1, A2, and let (N, χ) be an external dynamical system satisfying A3. Then there exist $\theta_0 \in N$, and $\mathcal{V}'$, open and dense in $\mathcal{V}$ with respect to the $C^{1;r}$-topology for $r \geq r_0$, such that, for every $V \in \mathcal{V}'$, the system (2.5) has a solution with $H_{\theta_0}(x(t), y(t), t) \geq At + B$ for some $A, B \in \mathbb{R}$ with $A > 0$, and for all t sufficiently large.*

For example, condition *A3* in Theorem 2.1 is automatically satisfied if $X(\theta) \neq 0$ for all θ. Note that in this case the only requirement on the flow χ on N is that it does not have any fixed points. This condition is necessary, since as we shall see in the argument in Subsection 4.4, we want the flow line $\chi_t(\theta_0)$ to leave some neighborhood of the point θ_0 before it returns again to that neighborhood.

Theorem 2.2. *Let g be a Riemannian metric on M satisfying conditions A1, A2, and let (N, χ) be an external dynamical system satisfying A3′. Then there exists a set $\mathcal{V}'$, open and dense in $\mathcal{V}$ with respect to the $C^{1;r}$-topology for $r \geq r_0$, such that, for every $V \in \mathcal{V}'$ and every $\theta_0 \in N$, the system (2.5) has a solution for which the energy $H_{\theta_0}(x(t), y(t), t)$ grows at least linearly to infinity as $t \to \infty$, i.e. $H_{\theta_0}(x(t), y(t), t) \geq At + B$ for some $A, B \in \mathbb{R}$ with $A > 0$, and for all t sufficiently large.*

Hypothesis *A3* in Theorem 2.1 requires choosing the parameter θ_0 to be a non-trivial uniformly recurrent point, and yields an unstable trajectory corresponding to that particular choice, while hypothesis *A3′* in Theorem 2.2 allows the parameter θ_0 to be arbitrary.

We also note that Theorem 2.2 remains valid under the weaker assumption that the flow χ_t on N is semisimple.

The linear growth rate $H_{\theta_0}(x(t), y(t), t) \approx t$ in Theorems 2.1 and 2.2 is optimal. Indeed, the energy $H_{\theta_0}(x(t), y(t), t)$ cannot grow in time faster than linearly, as we can easily show. By (2.5), we have

$$\frac{d}{dt} H_{\theta_0}(x(t), y(t), t) = \frac{\partial V}{\partial x}(x(t), \chi_t(\theta_0)) X(\chi_t(\theta_0)),$$

which is bounded due to the compactness of M and N.

In Subsection 4.3, we provide an explicit condition *A4* that ensures $V \in \mathcal{V}'$, as in Theorems 2.1 and 2.2. We emphasize here that this condition amounts to an explicit computation (4.14) that is verifiable in concrete systems. The potentials V satisfying this condition form an open and dense set in $\mathcal{V}$ relative to the $C^{1;r}$-topology.

Condition *A4* depends on a hyperbolic closed geodesic and on a pair of geometrically distinct homoclinic orbits associated to it. A generic geodesic flow has infinitely many homoclinic orbits associated to the same hyperbolic closed geodesic. Of course, for the main results we only need to verify the condition for just a single pair of homoclinic orbits.

Besides orbits whose energy grows unboundedly in time, there also exist orbits whose energy makes chaotic excursions, i.e., they follow any prescribed energy path. In other words, the system admits symbolic dynamics.

Theorem 2.3. *Assume the conditions of Theorem 2.1 hold, and let $V \in \mathcal{V}'$ with the set $\mathcal{V}'$ as in the statement of that theorem. Let E_* be sufficiently large. There exists $C > 0$, depending on the potential V, such that, for any differentiable function $\mathcal{E} : [0, \infty) \to [E_*, \infty)$ with $|\mathcal{E}'| \leq C$, there exist $\theta_0 \in N$, a time reparametrization $\mathcal{T} : [0, \infty) \to [0, \infty)$, a solution $(x(t), y(t))$ of the system (2.5), and a constant $D > 0$ such that*

$$|H_{\theta_0}(x(\mathcal{T}(t)), y(\mathcal{T}(t))) - \mathcal{E}(t)| \leq D\mathcal{E}(t)^{-1/2}.$$

Remark 2.4. A result similar to Theorem 2.2 appears in [43], where a more general class of Hamiltonian systems than that in the present paper is considered. While [43] does not require any recurrent property of the time-dependence, it requires some uniformity conditions on the frozen system, more precisely, that for all sufficiently large, fixed values of the time-dependent parameter, one is able to gain energy at a uniform rate. The difference is merely technical; we essentially consider a situation when one is able to ensure an energy growth at a uniform rate only locally, in some open subset of the parameter space, therefore we need recurrence to ensure that trajectories return to that subset infinitely often in a uniform way (see Subsection 4.3). Of course, if we choose to replace condition *A4* in Subsection 4.3 by a uniform, global condition, recurrence is no longer required.

Remark 2.5. It seems possible that Theorem 2.2 can be formulated in terms of a potential $V = V(x, t)$ that is an almost periodic function of time, rather than introducing the time-dependence via a trajectory of the auxiliary dynamical system (N, χ). However, Theorem 2.1 is formulated in terms of a local condition on the dynamics of the auxiliary dynamical system, while asking $V = V(x, t)$ being almost periodic is a global condition.

Remark 2.6. We note that Theorem 2.3 is similar in spirit to Theorem 1 in [14], which is formulated in the general class of slow-fast Hamiltonian systems; that paper shows that, under some assumptions, any continuous path in the slow phase space can be shadowed (see also the related paper [13]). One significant difference is that in Theorem 2.3 we shadow an energy path, not a path in the slow phase space; since we do not use averaging, we may not be able to shadow a specific path in the slow phase space. A key point of our paper is that we can drive the energy of the coupled system even without asking for much control on the inner dynamics.

2.6. Some examples of applications

The conditions for the flow on N in Theorem 2.1 are very general. What makes the examples presented below more surprising is that the flow on N can have a very simple orbit structure. The results include, as a particular case, quasi-periodic forcing, in which case we have the presence of the KAM and Nekhoroshev phenomena, which prevent, or delay, the onset of linear growth of energy, for a positive measure set of orbits. This is why we single out some examples to showcase Theorem 2.2.

For forcing systems with a rich orbit structure (e.g., horseshoes giving rise to symbolic dynamics), there are simpler arguments that show that one can get instability. Roughly, if the forcing system has essentially arbitrary orbits, we can choose initial conditions in

the forcing that lead to orbits that always increase the energy. Hence, for systems with complicated orbit structure there are many results of instability. A representative paper of this line of reasoning is [71]. Of course, even in the case when simpler mechanisms apply, the orbits constructed here are different, since we rely on the homoclinic excursions and not on riding the external forcing.

2.6.1. Perturbation of the geodesic flow by a quasi-periodic potential. We consider the particular case when $N = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, and the flow on N is a linear flow of rationally independent frequency vector ν, i.e. $\nu \cdot k \neq 0$ for all $k \in \mathbb{Z}^d$. Such a flow can be written as $\chi_t(\theta_0) = \theta_0 + \nu \cdot t$, for any initial point $\theta_0 \in \mathbb{T}^d$; for simplicity, we let $\theta_0 = 0$. The flow χ_t is minimal on $N = \mathbb{T}^d$.

The corresponding perturbed dynamical system is described by the time-dependent Hamiltonian $H : T^*M \times \mathbb{R} \to \mathbb{R}$ given by

$$H(x, y, t) = H_0(x, y) + V(x, \nu t). \tag{2.6}$$

As a consequence of Theorem 2.1, we find that, for a set of potentials $V : M \times \mathbb{T}^d \to \mathbb{R}$ that is generic relative to the $C^{1;r}$-topology, the system (2.6) has solutions $(x(t), y(t))$ for which the energy $H(x(t), y(t), t)$ grows linearly to infinity as $t \to \infty$. Note that the frequency ν is not required to be Diophantine, as in [30]. In addition to recovering the results from [30] under weaker conditions, we also obtain the existence of orbits whose energy grows at an optimal speed. Again, we point out that the orbits constructed in this paper are very different from those in [30], where the orbits stay near the closed geodesics for a long time.

When $d = 1$, the flow χ_t describes a periodic motion on $\mathbb{T}^1$, thus we obtain, as a particular case, the geodesic flow perturbed by a generic, periodic potential considered in, e.g., [64, 11, 29, 51].

2.6.2. Flows on Lie groups. In this section, we describe some other examples of external dynamical systems (N, χ_t) that can be used in Theorem 2.2. These examples are mild enough so that the mechanisms in [71] do not apply but nevertheless there is no averaging theory for them.

Note that we can interpret the example in Subsection 2.6.1 as a flow on a Lie group generated by a left-invariant vector field. Given a compact Lie group N, recall that a *left translation* on N is a map $L_g : N \to N$ given by $L_g(\theta) = g\theta$ for some $g \in N$. A vector field X on N is called *left-invariant* if X is invariant with respect to all left translations, i.e. $(L_g)_*(X) = X$ for all $g \in N$. Let χ_t be the flow of X on N. It is well known that each integral curve of χ_t is homeomorphic to the integral curve through the identity, that is, $\chi_t(\theta) = \theta \cdot \chi_t(e)$. The flow χ_t is minimal if and only if N is Abelian. However, every compact Abelian Lie group is a torus. Thus, our example in Subsection 2.6.1 has a natural interpretation as a left-invariant flow on a compact Abelian Lie group.

2.6.3. Horocycle flows. Let P be a compact connected surface with a Riemannian metric of negative curvature, and ξ_t be the geodesic flow restricted to the unit tangent

bundle T^1P. The *horocyle flow* is the unit speed flow on T^1P whose orbits (referred to as *horocycles*) are the strong stable manifolds $W^{ss}(z)$ of the geodesic flow,

$$W^{ss}(z) = \left\{ z' \in T^1M \;\middle|\; \lim_{t\to\infty} d(\xi_t(z'), \xi_t(z)) = 0 \right\}.$$

As a particular case, assume that P is a surface of constant negative curvature. The universal covering space is the Poincaré upper half-plane $\mathbb{H}$ with the Poincaré metric denoted ds. The geodesics in $\mathbb{H}$ are vertical lines and circles orthogonal to the real axis, and the horocycles are horizontal lines and circles tangent to the boundary. The orientation-preserving isometries of $\mathbb{H}$ are the linear fractional transformations, i.e., the elements of

$$\mathrm{PSL}(2,\mathbb{R}) = \left\{ z \mapsto \frac{az+b}{cz+d} \;\middle|\; ad - bc = 1 \right\}.$$

We can identify P as $\mathrm{PSL}(2,\mathbb{R})/\Gamma$ where Γ is a discrete cocompact subgroup[1] of $\mathrm{PSL}(2,\mathbb{R})$. The geodesic flow is given by

$$\xi_t(z) = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} z\Gamma$$

and the horocyle flow by

$$\chi_t(z) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} z\Gamma.$$

When the surface P has (variable) negative curvature, the universal Riemannian covering surface is the upper half-plane with metric f^2ds where f is a non-zero function. Then P can be regarded as the quotient of this space via a discrete cocompact group of isometries.

The geodesic flow on the unit tangent bundle is Anosov. Hence the unit tangent bundle is foliated by stable and unstable manifolds. When P is a surface, the stable and unstable manifolds are 1-dimensional. The horocycle flow is the motion at unit speed along the stable/unstable manifolds.

Hedlund [48] proved that the horocycle flow on a compact surface of constant negative curvature is minimal. Furstenberg [41] showed that it is uniquely ergodic (i.e., it admits a unique ergodic measure). B. Marcus [63] proved the same result for compact surfaces of variable negative curvature. Results on the minimality of the horocycle flow in higher dimensions were obtained by Eberlein (see e.g. [36]). We note that for surfaces the horocycle foliation is $C^{2-\varepsilon}$ [44] but in general not C^2. This is enough for our result, as argued in Subsection 5.7.

Hence, an interesting class of examples of external dynamical systems (N, χ) that can be used in Theorem 2.1 comes from horocycle flows on compact connected manifolds of negative curvature. We remark that horocycle flows do not determine perturbations of a fixed frequency (as in the quasi-periodic case).

[1] Γ is *cocompact* if $\mathrm{PSL}(2,\mathbb{R})/\Gamma$ is compact.

2.6.4. Flows on homogeneous spaces. Let G be a Lie group, Γ a discrete subgroup, $g : \mathbb{R} \to \Gamma$ a one-parameter subgroup, and $\chi : \mathbb{R} \times G/\Gamma \to G/\Gamma$ the G-induced flow, given by $\chi_t(z\Gamma) = (g(t)z)\Gamma$. Then G/Γ with the G-induced flow is a homogeneous space. An element $g \in G$ is *unipotent* if the adjoint[2] $\mathrm{Ad}(g)$ of g is a *unipotent matrix*, i.e., 1 is its only eigenvalue. If χ_t is unipotent for each t, then the flow χ is called a *unipotent flow*.

Ratner's classification theory [74] asserts that when Γ is cocompact and the flow χ is unipotent, if χ is uniquely ergodic then it is minimal. If $\mathrm{Vol}(G/\Gamma) < \infty$, where Vol denotes the Haar measure induced by χ, then the properties of minimality and unique ergodicity are equivalent.

Hence, another interesting class of examples of external dynamical systems (N, χ) that can be used in Theorem 2.1 are unipotent flows on homogeneous spaces of finite volume which are uniquely ergodic. Some specific examples appear in, e.g., [76, 68, 83].

2.6.5. Celestial mechanics and astrodynamics. We consider the Kepler problem, described by the Hamiltonian $H(y, x) = \frac{1}{2}|y|^2 - 1/|x|$. The solutions are conic sections or collision orbits. The regularized Keplerian motions for $H < 0$ can be lifted to trajectories of the geodesic flow on S^2. Let us consider this problem as a model for the motion of a satellite around the Earth. Following [23], the set of C^∞ Riemannian metrics on S^2 whose geodesic flow contains a non-trivial hyperbolic basic set is open and dense in the C^2-topology. Thus, we can choose a Riemannian metric that approximates the standard S^2-metric and for which there exists a hyperbolic closed geodesic with transverse homoclinic connection, satisfying conditions *A1*, *A2*. A closed geodesic will correspond to a closed orbit around the Earth. The motion of the Moon and the Sun can be regarded as a quasi-periodic forcing, as in *A3*. It seems possible that the method used to prove Theorem 2.2, which is constructive, can be adapted to this example in order to design explicit maneuvers on how to move, in specific ways, the satellite from one closed orbit to another around the Earth, or to increase the size and shape of a satellite orbit. The fact that an approximation to the metric on S^2 is used in this argument will result in some errors that can be corrected by low energy maneuvers. Satellite trajectory repositioning is very useful in astrodynamics: see, e.g., [70]. Some papers of related interest are [2, 1].

We should point out that in the above model we treat the satellite as an infinitesimal mass, that is, it moves under the gravitational influence of the other bodies (Earth, Sun, Moon) without exerting a gravitational influence on them. This situation is usually referred to in celestial mechanics as a *restricted problem*, and is an approximation of the full problem, when all bodies are considered to have positive masses and exert mutual gravitational influences. In a restricted problem, as is the case here, it is possible to achieve an energy growth of the infinitesimal mass while the total energy of the remaining bodies is conserved; of course, in the full problem, the energy growth of one mass is accompanied by an equal energy loss of the rest of the system. While idealized models, restricted problems are regarded as standard models in Hamiltonian dynamics, and they are successfully used in applications to astrodynamics [67].

[2] The *adjoint* $\mathrm{Ad}(g)$ of g is the derivative of the map $\Psi_g : G \to G$, $z \mapsto gzg^{-1}$, at e, i.e., $\mathrm{Ad}(g) = (d\Psi_g)_e : T_eG \to T_eG$, where d is the differential and T_eG is the tangent space at the origin e.

3. Geometric method

In this section we review briefly the geometric method to study perturbations of the geodesic flow developed in [29]. This method is based on the theory of normally hyperbolic invariant manifolds and on the scattering map. We will take the general set-up from [29] but we will make substantial modifications.

We consider the parameter-dependent Hamiltonian $H : T^*M \times N \to \mathbb{R}$ given by $H(x, y, \theta) = H_0(x, y) + V(x, \theta)$. The Hamilton equations and the parameter evolution equation are

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H_0}{\partial y}, \\ \frac{dy}{dt} &= -\frac{\partial H_0}{\partial x} - \frac{\partial V}{\partial x}, \\ \frac{d\theta}{dt} &= X(\theta). \end{aligned} \tag{3.1}$$

We write the corresponding flow $\psi : T^*M \times N \times \mathbb{R} \to T^*M \times N$ as

$$\psi_t(x, y, \theta) = (\xi_t(x, y, \theta), \chi_t(\theta)), \tag{3.2}$$

where χ_t is the flow defined by the vector field X on N.

3.1. Normal hyperbolicity

First we consider the unperturbed system, which is described by the Hamiltonian H_0 given by (2.1). Each energy manifold $\Sigma_E = \{(x, y) \mid H_0(x, y) = E\}$ is invariant under the geodesic flow ξ_t (now viewed as the Hamiltonian flow of H_0). We will denote by $\xi_{E,t} = (\xi^x_{E,t}, \xi^y_{E,t})$ a trajectory of the geodesic flow lying on Σ_E. By the rescaling property of the geodesic flow, assumptions *A1* and *A2* imply that for each energy level E there exists a periodic orbit λ_E that is hyperbolic in Σ_E, and a transverse homoclinic orbit γ_E to λ_E in Σ_E.

We consider a sufficiently large initial energy level $E_* \geq 0$ (to be specified later in the argument), and we define the 2-dimensional cylinder $\Lambda_0 = \bigcup_{E \geq E_*} \lambda_E$ in T^*M. Note that Λ_0 is a manifold with boundary but the flow is tangent to the boundary. This is a normally hyperbolic invariant manifold for the Hamiltonian flow on T^*M, whose stable and unstable manifolds are given by $W^s(\Lambda_0) = \bigcup_{E \geq E_*} W^s(\lambda_E)$ and $W^u(\Lambda_0) = \bigcup_{E \geq E_*} W^u(\lambda_E)$, respectively. The stable and unstable manifolds of Λ_0 intersect transversally along the 2-dimensional homoclinic manifold $\Gamma_0 = \bigcup_{E \geq E_*} \gamma_E$ in T^*M.

3.2. Scaled coordinates

We now rescale the coordinates (x, y) and the time t so that, for high energies, the flow corresponding to the rescaled Hamiltonian is a small and slow perturbation of the geodesic flow. For E_* sufficiently large we introduce a new parameter $\varepsilon = 1/\sqrt{E_*}$; we note that $E_* \to \infty$ if and only if $\varepsilon \to 0$.

The rescaled coordinates are (q, p) defined by $q = x$, $p = \varepsilon y$, and the rescaled time s is given by $s = t/\varepsilon$. The variable θ remains unchanged. The parameter-dependent

Hamiltonian in these new variables is

$$H_\varepsilon(q, p, \theta) = H_0(q, p) + \varepsilon^2 V(q, \theta) = \varepsilon^2 H(x, y, \theta). \tag{3.3}$$

The corresponding Hamilton equations and the parameter evolution equation are

$$\begin{aligned} \frac{dq}{ds} &= \frac{\partial H_0}{\partial p}, \\ \frac{dp}{ds} &= -\frac{\partial H_0}{\partial q} - \varepsilon^2 \frac{\partial V}{\partial q}, \\ \frac{d\theta}{ds} &= \varepsilon X(\theta). \end{aligned} \tag{3.4}$$

The corresponding flow ψ_s^ε is of the form $\psi_s^\varepsilon = (\xi_s^\varepsilon(q, p, \theta), \chi_s^\varepsilon(\theta))$.

When $\varepsilon \to 0$, the flow χ_s^ε approaches the constant flow id_θ on N, and the flow ξ_s^ε approaches the flow ξ_s in the C^{r-1}-topology. The limiting flow

$$\psi_s = (\xi_s, \mathrm{id}_\theta)$$

on the extended phase space $T^*M \times N$ has a normally hyperbolic invariant manifold $\tilde{\Lambda}_0 = \bigcup_{E \geq E_*} \lambda_E \times N$, since the exponential expansion rates of ξ_s are larger than those of id_θ. The stable and unstable manifolds of $\tilde{\Lambda}_0$ are given by $W^s(\tilde{\Lambda}_0) = \bigcup_{E \geq E_*} W^s(\lambda_E) \times N$ and $W^u(\tilde{\Lambda}_0) = \bigcup_{E \geq E_*} W^u(\lambda_E) \times N$, respectively. Obviously, $W^u(\tilde{\Lambda}_0)$ and $W^s(\tilde{\Lambda}_0)$ intersect transversally along the $(2 + d)$-dimensional homoclinic manifold $\tilde{\Gamma}_0 = \bigcup_{E \geq E_*} \gamma_E \times N$ in $T^*M \times N$.

Now we refer to the theory of normal hyperbolicity (see [50, 38, 6]) that shows that the invariant manifolds, which were identified in the limiting system, survive as locally invariant manifolds for the perturbed system. A very explicit proof of this can be found in [30]. The locally invariant manifolds are exactly invariant for a modified system constructed explicitly in [30]. The modifications are supported on $E \leq E_0$. When we refer to stable and unstable manifolds, we mean the stable and unstable manifolds of the extended system. Since we will be considering orbits whose energy stays large enough (in particular in regions where the extended system agrees with the original system), the orbits we construct will also be orbits of the original system.

Remark 3.1. When the system is differentiable enough and the perturbation is periodic or Diophantine, it is easy to remark [29] that the invariant manifold will contain KAM tori that act as boundaries, so that the manifold is in fact invariant. Of course, the argument in this paper does not require any differentiability and works for general perturbations.

Since ε enters both the size of the perturbation parameter and the time reparametrization, in order to apply the standard normally hyperbolicity theory one can rewrite the perturbed Hamiltonian as a two-parameter problem, with one parameter for the size of the perturbation and the other for the time change, and prove the persistence of the normally hyperbolic, locally invariant manifold from the unperturbed case to the perturbed case for all small enough sizes of the perturbation and uniformly in the time change parameter

(see [29, 30]). In the end, it follows that, for all sufficiently small ε (and implicitly for all sufficiently large E_*) the manifold $\tilde{\Lambda}_0$ can be continued to a normally hyperbolic locally invariant manifold $\tilde{\Lambda}_\varepsilon = \Lambda_\varepsilon \times N$ for the flow ψ_s^ε.

Moreover, the manifold $\tilde{\Lambda}_\varepsilon$ depends smoothly on the parameter ε, in the sense that there exists a C^{r-1}-smooth parametrization $\tilde{k}_\varepsilon : \tilde{\Lambda}_0 \to \tilde{\Lambda}_\varepsilon$ of $\tilde{\Lambda}_\varepsilon$, of the type $\tilde{k}_\varepsilon = (k_\varepsilon, \mathrm{id}_\theta)$, which depends C^{r-2}-smoothly on the parameter, such that $\tilde{k}_\varepsilon(\tilde{\Lambda}_0) = \tilde{\Lambda}_\varepsilon$ (and, in particular, $k_\varepsilon(\Lambda_0) = \Lambda_\varepsilon$).

In addition, there exist stable and unstable manifolds $W^u(\tilde{\Lambda}_\varepsilon)$ and $W^s(\tilde{\Lambda}_\varepsilon)$ that vary C^{r-1}-smoothly with ε. Similarly, there exist C^{r-1}-smooth parametrizations $\tilde{k}_\varepsilon^s : W^s(\tilde{\Lambda}_0) \to W^s(\tilde{\Lambda}_\varepsilon)$, $\tilde{k}_\varepsilon^u : W^u(\tilde{\Lambda}_0) \to W^u(\tilde{\Lambda}_\varepsilon)$ of the local stable and unstable manifolds of $\tilde{\Lambda}_\varepsilon$, which agree with $\tilde{k}_\varepsilon$ on $\tilde{\Lambda}_0$.

Since transversality is an open condition, $W^u(\tilde{\Lambda}_\varepsilon)$ and $W^s(\tilde{\Lambda}_\varepsilon)$ intersect transversally along a locally unique homoclinic manifold $\tilde{\Gamma}_\varepsilon = \Gamma_\varepsilon \times N$ for all ε sufficiently small (and so for all sufficiently large E_*).

3.3. Action-angle coordinates

For the unperturbed system, on the normally hyperbolic invariant manifold Λ_0 we can put a system of action-angle coordinates (J, ϕ), where the action coordinate is $J = \sqrt{2E}$, and the angle coordinate $\phi \in \mathbb{T}^1$ is symplectically conjugate to J, i.e., $dJ \wedge d\phi = (dy \wedge dx)_{|\Lambda_0}$. The unperturbed Hamiltonian is integrable on Λ_0, and it takes the following form in the action-angle coordinates:

$$H_0(\phi, J) = \tfrac{1}{2} J^2. \tag{3.5}$$

The unperturbed Hamiltonian flow on Λ_0 takes the form $J(t) = J$, $\phi(t) = \phi_0 + Jt$. Thus Λ_0 is foliated by invariant tori $\mathcal{T}_E = \{(J, \phi) \mid J = \sqrt{2E},\ \phi \in \mathbb{T}^1\}$ corresponding to each energy level $H_0 = E \geq E_*$.

For the perturbed system, the action-angle coordinate system (J, ϕ) on Λ_0 can be continued via k_ε to an action-angle coordinate system $(J_\varepsilon, \phi_\varepsilon)$ on Λ_ε. In these coordinates, the perturbed Hamiltonian function restricted to Λ_ε takes the form

$$H_\varepsilon(\phi_\varepsilon, J_\varepsilon, \theta) = \tfrac{1}{2} J_\varepsilon^2 + \varepsilon^2 V(\phi_\varepsilon, J_\varepsilon, \theta), \tag{3.6}$$

where $V(\phi_\varepsilon, J_\varepsilon, \theta)$ is obtained by expressing $V(\phi, J, \theta)$ in these coordinates. The Hamilton equations and the parameter evolution equation are

$$\begin{aligned} \frac{d\phi_\varepsilon}{ds} &= J_\varepsilon + \varepsilon^2 \frac{\partial V}{\partial J_\varepsilon}, \\ \frac{dJ_\varepsilon}{ds} &= -\varepsilon^2 \frac{\partial V}{\partial \phi_\varepsilon}, \\ \frac{d\theta}{ds} &= \varepsilon X(\theta). \end{aligned} \tag{3.7}$$

Note that, in general, the foliation of Λ_0 by invariant tori $\mathcal{T}_E$ in the unperturbed case does not survive to the perturbed case. For a perturbation driven by a general external

flow χ on N we cannot apply the KAM theorem since the perturbation affecting the tori is not necessarily periodic/quasi-periodic. For the same reason, we cannot apply averaging theory as in [45], since we cannot obtain level sets of the action that remain almost invariant for sufficiently long time.

3.4. The scattering map

A key tool to study the dynamics of a normally hyperbolic invariant manifold with a transverse homoclinic manifold is the scattering map [29]. The role of the scattering map is to relate the past asymptotic trajectory of an orbit in the homoclinic manifold to its future asymptotic trajectory. An extended study of the scattering map and its geometric properties can be found in [32].

Here we briefly recall the construction of the scattering map. We consider a flow on some manifold P. Let Λ be a normally hyperbolic invariant manifold for the flow, with the unstable and stable manifolds $W^u(\Lambda)$ and $W^s(\Lambda)$ intersecting transversally along a homoclinic manifold Γ. This means that $\Gamma \subseteq W^u(\Lambda) \cap W^s(\Lambda)$ and, for each $z \in \Gamma$,

$$T_z P = T_z W^u(\Lambda) + T_z W^s(\Lambda), \qquad T_z \Gamma = T_z W^u(\Lambda) \cap T_z W^s(\Lambda). \tag{3.8}$$

By the normal hyperbolicity of Λ, $W^u(\Lambda)$ is foliated by 1-dimensional fibers $W^u(z)$, $z \in \Lambda$, and $W^s(\Lambda)$ is foliated by 1-dimensional fibers $W^s(z)$, $z \in \Lambda$. For each $z \in W^u(\Lambda)$ there exists a unique $z^- \in \Lambda$ such that $z \in W^u(z^-)$, and for each $z \in W^s(\Lambda)$ there exists a unique $z^+ \in \Lambda$ such that $z \in W^s(z^+)$. We define the wave maps

$$\begin{aligned} &\Omega^+ : W^s(\Lambda) \to \Lambda, \qquad \Omega^+(z) = z^+, \\ &\Omega^- : W^u(\Lambda) \to \Lambda, \qquad \Omega^-(z) = z^-. \end{aligned} \tag{3.9}$$

These maps are differentiable.

To define the scattering map, we make the additional assumption that for each $z \in \Gamma$,

$$T_z W^s(\Lambda) = T_z W^s(z^+) \oplus T_z(\Gamma), \qquad T_z W^u(\Lambda) = T_z W^u(z^-) \oplus T_z(\Gamma). \tag{3.10}$$

By restricting Γ to some open subset of it if necessary, we can ensure that the restrictions of $\Omega^{\pm}$ to Γ are diffeomorphisms. We define the scattering map associated to Γ to be the diffeomorphism $S = \Omega^+ \circ (\Omega^-)^{-1}$ from $U^- := \Omega^-(\Gamma)$ in Λ to $U^+ := \Omega^+(\Gamma)$ in Λ.

As pointed out in [32], the scattering map can be defined in an analogous way in the case of time-dependent systems. In this case, one considers a flow associated to a skew product vector field $(Y(z,\theta), X(\theta))$ on a product manifold $P \times N$, where the skew product vector field is assumed to be close to an autonomous vector field, i.e.,

$$\|Y(z,\theta) - Y_0(z)\|_{C^r} \ll 1$$

for some vector field $Y_0(z)$ on P.

Assume that there exists a normally hyperbolic invariant manifold Λ_0 in P for the flow of $Y_0(z)$. Let S_0 be the scattering map associated to a homoclinic channel Γ_0. Assuming that the exponential rates of the flow on N are smaller than the exponential rates for the flow on P, it follows that $\Lambda_0 \times N$ is normally hyperbolic for the product flow of

$(Y_0(z), X(\theta))$, that $\Gamma_0 \times N$ is a homoclinic channel, and that the corresponding scattering map $\tilde{S}_0$ is a product of the form $\tilde{S}_0(z, \theta) = (S_0(z), \theta)$.

By the theory of normal hyperbolicity, there exists a normally hyperbolic invariant manifold close to $\Lambda_0 \times N$ in $P \times N$ for the flow of $(Y(z, \theta), X(\theta))$, provided $Y(z, \theta)$ is sufficiently close to $Y_0(z)$ in the sense described above. There also exists a homoclinic channel close to $\Gamma \times N$. The corresponding scattering map takes the skew product form $\tilde{S}(z, \theta) = (S(z, \theta), \theta)$.

3.5. The scattering map for the unperturbed geodesic flow

In the case of the unperturbed geodesic flow, described by the Hamiltonian H_0, condition (3.8) is satisfied in view of assumption *A2*, and (3.10) follows from the fact that Λ_0 is foliated by the invariant tori $\mathcal{T}_E$. The scattering map preserves each of these invariant tori and it only changes the phase ϕ along each torus by an amount a that is independent of the torus. The scattering map $S_0 : U_0^- \to U_0^+$ is expressed relative to the (J, ϕ)-coordinates on Λ_0 by

$$S_0(J^-, \phi^-) = (J^+, \phi^+) = (J^-, \phi^- + a), \tag{3.11}$$

where $J^- = J^+$ and the phase shift $a = \phi^+ - \phi^-$ is independent of the point $z^- = z^-(\phi^-, J^-)$ in U_0^-, as it only depends on the homoclinic manifold Γ_0.

Remarkably, in this case the scattering map can be globally defined as a continuous map on the whole of Λ_0, as it has no monodromy (see [32]). The continuation, however, is rather subtle because when the base point moves along a non-contractible closed curve in Λ, the point of homoclinic intersection changes. As already observed in [29], this causes that after perturbations, the global definition may be impossible. When we continue the scattering map along a non-contractible closed curve, we may end up with a different map.

3.6. The scattering map for the perturbed geodesic flow

We now describe the scattering map for the system (3.4). We stress that we use the notation $\tilde{}$ to denote the variables in the extended system. That is, $\tilde{}$ refers to adding the extra variable $\theta \in N$.

It is important to note that if the dynamics in N has a small growth rate—which happens in our case for small enough ε because the dynamics is given by the vector field εX—then we can apply the theory of scattering maps for non-autonomous systems recalled in Subsection 3.4.

As described in Subsection 3.2, $\tilde{\Lambda}_\varepsilon = \Lambda_\varepsilon \times N$ is a normally hyperbolic invariant manifold for the extended dynamics given by ψ_s^ε. Furthermore, $W^s(\tilde{\Lambda}_\varepsilon) = W^s(\Lambda_\varepsilon) \times N$ is the stable manifold of $\tilde{\Lambda}_\varepsilon$ under the extended dynamics. We also have $W^s(z, \theta) = W^s(z) \times \{\theta\}$, and analogously for the unstable manifold.

Hence, if Γ_ε is a homoclinic manifold satisfying (3.8) and (3.10), we see that $\tilde{\Gamma}_\varepsilon = \Gamma_\varepsilon \times N$ will also satisfy (3.8) and (3.10) in the extended system. The wave maps (3.9) associated to $\tilde{\Gamma}_\varepsilon$ in the extended system can be written as

$$\tilde{\Omega}_\varepsilon^\pm(z, \theta) = (\Omega_\varepsilon^\pm(z, \theta), \theta).$$

Therefore, if we can associate a scattering map to Γ_ε, we can also associate a scattering map to $\tilde{\Gamma}_\varepsilon$ and we have

$$\tilde{S}_\varepsilon(z,\theta) = (S_\varepsilon(z,\theta),\theta).$$

Note that the scattering map is the identity in the N component. For θ fixed, let $S_{\varepsilon,\theta}$ denote the map given by $S_{\varepsilon,\theta}(z) = S_\varepsilon(z,\theta)$.

We can think of the scattering map either as the mapping $\tilde{S}_\varepsilon$, or as a family of mappings $S_{\varepsilon,\theta}$, indexed by the parameter θ. As proved in [32], the mappings $\tilde{S}_\varepsilon$ and $S_{\varepsilon,\theta}$ are smooth and depend smoothly on parameters.

It is proved in [32] that $S_{\varepsilon,\theta}$ is symplectic as a mapping on a domain in Λ_ε if the flow is a time-dependent symplectic flow. (It is also proved in [32] that Λ_ε is a symplectic manifold.)

In our situation, one of the consequences of the smooth dependence on parameters of $\tilde{\Lambda}_\varepsilon$ and of its stable and unstable manifolds is that we can find regions $U_0^\pm \subset \Lambda_0$ which are independent of ε, for $0 < \varepsilon \leq \varepsilon_0 \ll 1$, such that $\tilde{k}_\varepsilon(U_0^\pm \times N) \subseteq \tilde{U}_\varepsilon^\pm$, where $\tilde{k}_\varepsilon$ is the parametrization of $\tilde{\Lambda}_\varepsilon$ described in Subsection 3.2.

Via the parametrization $\tilde{k}_\varepsilon$, we can consider the scattering map $S_{\varepsilon,\theta}$ as defined from U_0^- to U_0^+, for $\theta \in N$. That is, we can consider the scattering map as being defined between domains that are of product type and are independent of ε (of size of order $O(1)$).

4. Elementary building blocks for the dynamics

In this section we construct some elementary building blocks of the dynamics. Each building block is a pseudo-orbit determined by one application of the scattering map followed by the application of the inner dynamics for some time. The repeated construction of such elementary building blocks will produce a two-dynamics pseudo-orbit which intersperses the scattering map dynamics with the inner dynamics.

In Section 5.3, given a sequence of elementary building blocks, we will construct a sequence of windows which are correctly aligned by the dynamics. Afterwards, we will use the shadowing property of correctly aligned windows from Subsection 5.1 to deduce the existence of a true orbit following the sequence of elementary building blocks.

4.1. The effect of the scattering map on the scaled energy

The goal of this section is to compute the change of the energy H_ε by one of the scattering maps when $\varepsilon \in (0, \varepsilon_0)$, for ε_0 sufficiently small. Our goal will be to obtain estimates uniform in ε. The main observation is that the energy is a slow variable.

4.1.1. Preliminaries. We will be working with the scaled flow (3.1) in a fixed bounded range of scaled energies, which we choose to be $H_\varepsilon \in [1, 2]$. This determines a compact subset $H_\varepsilon^{-1}[1, 2] \cap \tilde{\Lambda}_\varepsilon$ in $\tilde{\Lambda}_\varepsilon$. Of course, obtaining estimates for all $\varepsilon \in (0, \varepsilon_0)$ corresponds to letting the physical variables (x, y) take values in a non-compact domain.

When we will say that some error term is bounded by a constant (or by $O(\varepsilon^a)$) it will mean uniformly in that compact set.

Denote by

$$\psi_s^\varepsilon(\tilde{z}_\varepsilon) \equiv (\xi_s^{\varepsilon,q}(\tilde{z}_\varepsilon), \xi_s^{\varepsilon,p}(\tilde{z}_\varepsilon), \chi_s^\varepsilon(\theta))$$

the trajectory of the scaled Hamilton equations (3.4) with initial condition $\tilde{z}_\varepsilon = (z_\varepsilon, \theta) \in T^*M \times N$. We have

$$\frac{d}{ds} H_\varepsilon(\psi_s^\varepsilon(\tilde{z}_\varepsilon)) = \varepsilon^3 \nabla_\theta V(\xi_s^{\varepsilon,q}(\tilde{z}_\varepsilon), \chi_s^\varepsilon(\theta)) \cdot X(\chi_s^\varepsilon(\theta)) \equiv \varepsilon^3 D_X V(\xi_s^{\varepsilon,q}(\tilde{z}_\varepsilon), \chi_s^\varepsilon(\theta)). \tag{4.1}$$

Here we regard $\frac{d}{ds} H_\varepsilon$ as a functional acting on the solution curves of the Hamiltonian equations (3.4) in the extended phase space.

We will compute the leading term in ε of the change in energy by a scattering map in the perturbed equation (see Subsection 3.6).

In the rest of the subsection, we will consider a fixed homoclinic intersection $\tilde{\Gamma}_\varepsilon$, and we will denote by $\tilde{S}_\varepsilon$ the corresponding scattering map.

The scattering map is defined from some domain to some range in $\tilde{\Lambda}_\varepsilon$. Since the manifold $\tilde{\Lambda}_\varepsilon$ depends on ε, it is convenient to reduce the scattering map to some manifold which is independent of ε. We use the fact that the manifold $\tilde{\Lambda}_\varepsilon$ is diffeomorphic to the manifold $\tilde{\Lambda}_0$ via the parametrization $\tilde{k}_\varepsilon$ mentioned in Subsection 3.2. It is important to note that $\tilde{k}_\varepsilon$ is ε-close to the identity relative to the C^r-topology for ε small. We will use this fact below, when we compute the leading term in the expansion with respect to ε of the change of energy by the scattering map; we will be able to approximate $\tilde{k}_\varepsilon$ by the identity and incur only error terms which are subdominant. Moreover, $\tilde{k}_\varepsilon$ can be chosen so that it is symplectic [30].

As Λ_0 is foliated by geodesics λ_E, $E \geq E_*$, we can parametrize Λ_0 by the map $(E, s) \mapsto \lambda_E(s)$, where $E \geq E_*$ and the time s is considered mod $(1/\sqrt{2E})$, i.e. $\lambda_E(s) = \lambda_E(s')$ if $s' - s \in (1/\sqrt{2E}) \cdot \mathbb{Z}$. Hence, we can parametrize Λ_ε by $(E, s) \mapsto k_\varepsilon(\lambda_E(s))$.

Thus, each point $z_\varepsilon \in \Lambda_\varepsilon$ can be written as $z_\varepsilon = k_\varepsilon(\lambda_E(s))$ for some unique E and s. Note that E and s depend on both the point $z_\varepsilon \in \Lambda_\varepsilon$ and the perturbation parameter ε. Therefore, each point $\tilde{z}_\varepsilon \in \tilde{\Lambda}_\varepsilon$ with $\tilde{z}_\varepsilon = (z_\varepsilon, \theta)$ can be written as $\tilde{z}_\varepsilon = \tilde{k}_\varepsilon(\tilde{\lambda}_E(s))$, where $\tilde{\lambda}_E(s) = (\lambda_E(s), \theta)$.

Via the parametrization $\tilde{k}_\varepsilon$, instead of $\tilde{S}_\varepsilon$ we consider the reduction $\tilde{S}_\varepsilon^o = \tilde{k}_\varepsilon^{-1} \circ \tilde{S}_\varepsilon \circ \tilde{k}_\varepsilon$: $\tilde{k}_\varepsilon^{-1}(\tilde{U}_\varepsilon^-) \to \tilde{k}_\varepsilon^{-1}(\tilde{U}_\varepsilon^+)$; note that the domain and codomain of $\tilde{S}_\varepsilon^o$ are subsets of $\tilde{\Lambda}_0$. We note that when $\varepsilon \to 0$ the scattering map $\tilde{S}_\varepsilon^o$ approaches the unperturbed scattering map S_0 in the C^{r-1}-topology.

Since on Λ_0 we consider two coordinate systems, (E, s) and (J, ϕ), we would like to make explicit the unperturbed scattering map in both coordinates. Given $S_0(z^-) = z^+$, in the action-angle coordinates, if $z^- = (J^-, \phi^-)$, $z^+ = (J^+, \phi^+)$, then $J^- = J^+$ and $\phi^- + a = \phi^+$, and in the energy-time coordinates, if $z^- = (E^-, s^-)$, $z^+ = (E^+, s^+)$, then $E^- = E^+$ and $s^- + a/\sqrt{2E} = s^+$.

4.1.2. The effect of the scattering map on the scaled energy. The first goal of this subsection is getting quantitative estimates on the change of scaled energy achieved by the

scattering map:

$$\Delta(\tilde{z}_\varepsilon^-) := H_\varepsilon(\tilde{z}_\varepsilon^+) - H_\varepsilon(\tilde{z}_\varepsilon^-) = H_\varepsilon(\tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-)) - H_\varepsilon(\tilde{z}_\varepsilon^-). \tag{4.2}$$

We will write each point $\tilde{z}_\varepsilon^-$ as $(k_\varepsilon(E,s),\theta)$ for some E, s, θ. We will express the leading term of the expansion of $\Delta(\tilde{z}_\varepsilon^-)$ with respect to ε in terms of the unperturbed system, and specifically in terms of the variables (E,s,θ).

Proposition 4.1. *Let $\tilde{z}_\varepsilon^- = (k_\varepsilon(\lambda_E(s)),\theta) \in \tilde{U}_\varepsilon^-$ and $\tilde{z}_\varepsilon^+ = \tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-)$. The change of the scaled energy H_ε by the scattering map from $\tilde{z}_\varepsilon^-$ to $\tilde{z}_\varepsilon^+$ is given by*

$$\Delta(\tilde{z}_\varepsilon^-) = \varepsilon^3 \Delta_1(E,s,\theta) + O(\varepsilon^4 |\ln \varepsilon|), \tag{4.3}$$

where the leading term Δ_1 in the expansion with respect to ε is given by

$$\begin{aligned}\Delta_1(E,s,\theta) = \lim_{T_\pm \to \pm\infty} \Bigg[& \int_{T_-}^{T_+} (D_X V)(\gamma_E^q(\sigma),\theta)\, d\sigma \\ & - \int_0^{T_+} (D_X V)\big(\lambda_E^q(\sigma + s + a/\sqrt{2E}),\theta\big)\, d\sigma - \int_{T_-}^{0} (D_X V)(\lambda_E^q(\sigma+s),\theta)\, d\sigma \Bigg].\end{aligned} \tag{4.4}$$

Proof. Let $\tilde{z}_\varepsilon = (\tilde{\Omega}^-)^{-1}(\tilde{z}_\varepsilon^-) = (\tilde{\Omega}^+)^{-1}(\tilde{z}_\varepsilon^+) \in \tilde{\Gamma}_\varepsilon$. We start by noting that the orbit starting at $\tilde{z}_\varepsilon$ is asymptotic in the future to the orbit of $\tilde{z}_\varepsilon^+$ and in the past to the orbit of $\tilde{z}_\varepsilon^-$. We can write

$$\begin{aligned}\Delta(\tilde{z}_\varepsilon^-) = \lim_{T_\pm \to \pm\infty} [& H_\varepsilon(\psi_{T_+}^\varepsilon(\tilde{z}_\varepsilon)) - H_\varepsilon(\psi_{T_-}^\varepsilon(\tilde{z}_\varepsilon)) \\ & - H_\varepsilon(\psi_{T_+}^\varepsilon(\tilde{z}_\varepsilon^+)) + H_\varepsilon(\tilde{z}_\varepsilon^+) + H_\varepsilon(\psi_{T_-}^\varepsilon(\tilde{z}_\varepsilon^-)) - H_\varepsilon(\tilde{z}_\varepsilon^-)].\end{aligned} \tag{4.5}$$

By the fundamental theorem of calculus, we have

$$\begin{aligned}\Delta(\tilde{z}_\varepsilon^-) &= \lim_{\substack{T_+ \to \infty \\ T_- \to -\infty}} \Bigg[\int_{T_-}^{T_+} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon))\, d\sigma \\ &\qquad - \int_0^{T_+} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^+))\, d\sigma - \int_{T_-}^{0} \left(\frac{dH_\varepsilon}{d\sigma}\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^-))\, d\sigma \Bigg] \\ &= \lim_{T_+ \to \infty} \int_0^{T_+} \left[\left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon)) - \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^+)) \right] d\sigma \\ &\quad + \lim_{T_- \to -\infty} \int_{T_-}^{0} \left[\left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon)) - \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^-)) \right] d\sigma.\end{aligned} \tag{4.6}$$

It is important to remark that the integrands in (4.6) converge exponentially fast. So do their derivatives of low order with respect to the initial conditions. It is shown in [32] that there is exponential convergence of the integrands in (4.6) for the derivatives of order lower than the ratio of the Lyapunov exponents in the stable and unstable directions and the Lyapunov exponents tangent to the manifold. In our case, since the directions tangent to the manifold have zero exponent for $\varepsilon = 0$, one can deduce that the number of derivatives of the integrands that converge exponentially fast is arbitrarily large for ε small. In the arguments presented in this paper, we will need just a moderate number of derivatives.

Our next goal is to prune (4.6) to extract a convenient expression for the leading term.

By the exponential convergence of the integrands, we see that for appropriately chosen constant $K > 0$, if we take $T_+ = -T_- = K|\ln\varepsilon|$, the integrals differ from the limit by no more than $O(\varepsilon^4)$. Therefore,

$$\Delta(\tilde{z}_\varepsilon^-) = \int_{-K|\ln\varepsilon|}^{K|\ln\varepsilon|} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon))\, d\sigma - \int_0^{K|\ln\varepsilon|} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^+))\, d\sigma$$
$$- \int_{-K|\ln\varepsilon|}^{0} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^-))\, d\sigma + O(\varepsilon^4). \tag{4.7}$$

We now express $\Delta(\tilde{z}_\varepsilon^-)$ in terms of the orbits of the unperturbed flow. By the smooth dependence on parameters of solutions of ordinary differential equations, for $|\sigma| \le K|\ln\varepsilon|$, we have that the orbits of the perturbed system are $O(\varepsilon|\ln\varepsilon|)$-close to the corresponding orbits of the unperturbed flow. Notice that, because $\varepsilon|\ln\varepsilon|$ is small, the separation between the orbits is still growing linearly and has not yet started to grow exponentially fast with time [78, estimate (3.5.4)].

More precisely, the orbits $\xi_\sigma^\varepsilon(\tilde{z}_\varepsilon^\pm)$ are $O(\varepsilon|\ln\varepsilon|)$-close to the orbit $(\lambda_E, \chi^\varepsilon)$ where λ_E is a closed geodesic, and the orbit $\xi_\sigma^\varepsilon(\tilde{z}_\varepsilon)$ is $O(\varepsilon|\ln\varepsilon|)$-close to $(\gamma_E, \chi^\varepsilon)$, where γ_E is a homoclinic orbit to λ_E for the geodesic flow. Also, in these intervals of time, the variable θ changes only by $O(\varepsilon|\ln\varepsilon|)$. Thus, for $|\sigma| \le K|\ln\varepsilon|$, we have

$$d(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon), \tilde{\gamma}_E(\sigma)) \le C\varepsilon|\ln\varepsilon|$$
$$d(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^-), \tilde{\lambda}_E(\sigma + s)) \le C\varepsilon|\ln\varepsilon|,$$
$$d\big(\psi_\sigma^\varepsilon(\tilde{z}_\varepsilon^+), \tilde{\lambda}_E(\sigma + s + a/\sqrt{2E})\big) \le C\varepsilon|\ln\varepsilon|,$$

for some constant $C > 0$, where $\tilde{\gamma}_E(\sigma) = (\gamma_E(\sigma), \theta)$ and $\tilde{\lambda}_E(\sigma) = (\lambda_E(\sigma), \theta)$.

Therefore, substituting in the integral in (4.7) the orbits of the geodesic flow instead of the orbits of the perturbed flow and keeping θ constant, we incur an error $O(\varepsilon^4|\ln\varepsilon|)$. Using also the notation $D_X V$ introduced in (4.1), we obtain

$$\begin{aligned}
\Delta(\tilde{z}_\varepsilon^-) &= \int_{-K|\ln\varepsilon|}^{K|\ln\varepsilon|} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\tilde{\gamma}_E(\sigma))\, d\sigma - \int_0^{K|\ln\varepsilon|} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\tilde{\lambda}_E(\sigma + s + a/\sqrt{2E}))\, d\sigma \\
&\quad - \int_{-K|\ln\varepsilon|}^{0} \left(\frac{d}{d\sigma} H_\varepsilon\right)(\tilde{\lambda}_E(\sigma + s))\, d\sigma + O(\varepsilon^4|\ln\varepsilon|) \\
&= \varepsilon^3 \int_{-K|\ln\varepsilon|}^{K|\ln\varepsilon|} (D_X V)(\gamma_E^q(\sigma), \theta)\, d\sigma \\
&\quad - \varepsilon^3 \int_0^{K|\ln\varepsilon|} (D_X V)(\lambda_E^q(\sigma + s + a/\sqrt{2E}), \theta)\, d\sigma \\
&\quad - \varepsilon^3 \int_{-K|\ln\varepsilon|}^{0} (D_X V)(\lambda_E^q(\sigma + s), \theta)\, d\sigma + O(\varepsilon^4|\ln\varepsilon|),
\end{aligned} \tag{4.8}$$

where λ_E^q, γ_E^q denote the q-components of λ_E, γ_E, respectively.

Finally, because of the exponentially fast convergence we can change the integral over $O(\ln \varepsilon)$ time interval as $t \to \infty$ incurring an error $O(\varepsilon^4)$, so we obtain:

$$\begin{aligned}
\Delta(\tilde{z}_\varepsilon^-) &= \varepsilon^3 \lim_{\substack{T_+\to\infty \\ T_-\to\infty}} \Bigg[\int_{T_-}^{T_+} (D_X V)(\gamma_E^q(\sigma), \theta)\, d\sigma \\
&\quad - \int_0^{T_+} (D_X V)(\lambda_E^q(\sigma + s + a/\sqrt{2E}), \theta)\, d\sigma - \int_{T_-}^{0} (D_X V)(\lambda_E^q(\sigma + s), \theta)\, d\sigma \Bigg] \\
&\quad + O(\varepsilon^4 |\ln \varepsilon|) \\
&= \varepsilon^3 \Bigg[\lim_{T_+\to\infty} \int_0^{T_+} [(D_X V)(\gamma_E^q(\sigma), \theta) - (D_X V)(\lambda_E^q(\sigma + s + a/\sqrt{2E}), \theta)]\, d\sigma \\
&\quad + \lim_{T_-\to-\infty} \int_{T_-}^{0} [(D_X V)(\gamma_E^q(\sigma), \theta) - (D_X V)(\lambda_E^q(\sigma + s), \theta)]\, d\sigma \Bigg] \\
&\quad + O(\varepsilon^4 |\ln \varepsilon|) \\
&= \varepsilon^3 \Delta_1(E, s, \theta) + O(\varepsilon^4 |\ln \varepsilon|). \qquad (4.9)
\end{aligned}$$

In the last line we have just defined $\Delta_1 = \Delta_1(E, s, \theta)$ as the leading term of $\Delta = \Delta(\tilde{z}_\varepsilon^-)$, where $\tilde{z}_\varepsilon^- = (k_\varepsilon(\lambda_E(s)), \theta) \in \tilde{U}_\varepsilon^-$. □

Notice that in the above argument we used the exponential convergence of the integrands to justify the change of the limits, which is often done in Melnikov theory.

It is important to realize that the expression for the leading term Δ_1 is in terms of the unperturbed trajectories and depends only on the perturbing potential. It can be considered as a global Melnikov function. In contrast with many standard treatments in which the Melnikov function is only defined for periodic or quasi-periodic orbits, (4.9) is well defined for all orbits in the domain of the scattering map independently of what is their dynamics. Also, note that the function Δ_1 can be viewed as an analogue of what was called the reduced Poincaré function in [29, 30].

4.1.3. The effect of the scattering map on the action-angle coordinates. Using action-angle coordinates, we write the geodesic $\lambda_E(s)$ as $\lambda_{J^2/2}(s)$ and the homoclinic $\gamma_E(s)$ as $\gamma_{J^2/2}(s)$, where $E = J^2/2$. By the rescaling property of the geodesic flow (2.3) we have $\lambda_{J^2/2}(s) = \lambda_1(Js)$ and $\gamma_{J^2/2}(s) = \gamma_1(Js)$. Note that for an energy $E = 1$ the corresponding action is $J = \sqrt{2}$. By a change of variable we obtain

$$\begin{aligned}
\Delta_1(J, \phi, \theta) &= \lim_{T_+\to\infty} \int_0^{T_+} [(D_X V)(\gamma_{J^2/2}^q(\sigma), \theta) - (D_X V)(\lambda_{J^2/2}^q(\sigma + \phi + a), \theta)]\, d\sigma \\
&\quad + \lim_{T_-\to-\infty} \int_{T_-}^{0} [(D_X V)(\gamma_{J^2/2}^q(\sigma), \theta) - (D_X V)(\lambda_{J^2/2}^q(\sigma + \phi), \theta)]\, d\sigma
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{T_+ \to \infty} \int_0^{T_+} [(D_X V)(\gamma_1^q(J\sigma), \theta) - (D_X V)(\lambda_1^q(J(\sigma + \phi + a)), \theta)]\, d\sigma \\
&\quad + \lim_{T_- \to -\infty} \int_{T_-}^{0} [(D_X V)(\gamma_1^q(J\sigma), \theta) - (D_X V)(\lambda_1^q(J(\sigma + \phi)), \theta)]\, d\sigma \\
&= \frac{1}{J} \Delta_1(\sqrt{2}, J\phi, \theta).
\end{aligned}
$$

Thus, we have the following rescaling property of Δ_1 relative to action-angle coordinates:

$$
\Delta_1(J, \phi, \theta) = \frac{1}{J} \Delta_1(\sqrt{2}, J\phi, \theta). \tag{4.10}
$$

4.2. Change of energy over an elementary building block of a pseudo-orbit

The goal of this section is to compute the change of energy over an elementary building block of a pseudo-orbit, which consists in applying the scattering map followed by applying the inner dynamics for some prescribed time. Later we will also formulate some conditions that ensure that the effect on the energy is non-trivial. Of course, this requires some choices (e.g., the time we decide to follow the inner dynamics), and, given the choices made, some non-degeneracy assumptions on the perturbations (e.g., a perturbation that vanishes identically will not produce any effect).

We first compute the energy change over an elementary building block obtained by starting at a point $\tilde{z}_\varepsilon^- = (z_\varepsilon, \theta)$, applying the scattering map, and then applying the inner dynamics for some time. We specify the time for which we apply the inner dynamics implicitly by requiring that the change of angle coordinate along the pseudo-orbit starting from $\tilde{z}_\varepsilon^-$ is some fixed number L, chosen sufficiently large, to be specified later. We denote such an elementary building block by $B(\tilde{z}_\varepsilon^-)$.

As specified before, we focus on a bounded energy range $H_\varepsilon \in [1, 2]$. From Subsection 4.1, the leading term in the energy change (4.9) depends on the effect of the perturbing potential on the unperturbed trajectories. Let $\tilde{z}_\varepsilon^- = (k_\varepsilon(\lambda_E(s)), \theta)$, with E in the energy range. Assume that the angle-action coordinates of $\lambda_E(s) \in \Lambda_0$ are (J, ϕ) with $J = \sqrt{2E}$.

By (3.11) the effect of the unperturbed scattering map on Λ_0, taking a point z^- to z^+, is to increase the angle coordinate ϕ by a. The scaled time s to follow the inner dynamics starting from z^+ and ending at a point of angle coordinate L is $(L - a)/\sqrt{2E}$. Hence, in terms of the unperturbed system, one follows the geodesic flow trajectory $\lambda_E(s)$ for the time interval $s \in [a/\sqrt{2E}, L/\sqrt{2E}]$.

In the perturbed system, we choose to follow the inner dynamics for the same time interval, which is independent of ε. Since the energy is a slow variable and we are considering only scaled times of order 1, the change of energy during the time spent along the inner dynamics can be computed, with a very small error, using the fundamental theorem of calculus.

This implies that the change of energy along an orbit segment starting at some point $\tilde{z}_\varepsilon^+ = (z_\varepsilon^+, \theta^+)$, where $\theta^+ = \chi^\varepsilon_{a/\sqrt{2E}}(\theta)$, and following it for a time interval

$s \in [a/\sqrt{2E}, L/\sqrt{2E}]$ is

$$\int_{a/\sqrt{2E}}^{L/\sqrt{2E}} \frac{d}{dt} H_\varepsilon(\xi_\sigma^\varepsilon(z_\varepsilon^+), \chi_\sigma^\varepsilon(\theta))\, d\sigma = \varepsilon^3 \int_{a/\sqrt{2E}}^{L/\sqrt{2E}} (D_X V)(\lambda_E^q(\sigma), \theta)\, d\sigma + O(\varepsilon^4 |\ln \varepsilon|).$$

Proposition 4.2. *Let $\tilde{z}_\varepsilon^- = (k_\varepsilon(\lambda_E(s), \theta)$, and $\tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-) = \tilde{z}_\varepsilon^+$. Consider an elementary building block consisting of one application of the scattering map $\tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-) = \tilde{z}_\varepsilon^+$, and a trajectory segment with initial point $\tilde{z}_\varepsilon^+$ following the inner dynamics for a time interval $s \in [a/\sqrt{2E}, L/\sqrt{2E}]$. The change $G(\tilde{z}_\varepsilon^-)$ of scaled energy H_ε over the building block is of the form*

$$G(\tilde{z}_\varepsilon^-) = \varepsilon^3 G_1(E, s, \theta) + O(\varepsilon^4 |\ln \varepsilon|), \tag{4.11}$$

where G_1, the leading term of G, is given by

$$\begin{aligned} G_1(E, s, \theta) = \lim_{\substack{T_+ \to \infty \\ T_- \to -\infty}} \Bigg[& \int_{T_-}^{T_+} (D_X V)(\gamma_E^q(s), \theta)\, d\sigma \\ & - \int_0^{T_+} (D_X V)(\lambda_E^q(\sigma + s + a/\sqrt{2E}), \theta)\, d\sigma - \int_{T_-}^{0} (D_X V)(\lambda_E^q(\sigma + s), \theta)\, d\sigma \\ & + \int_{a/\sqrt{2E}}^{L/\sqrt{2E}} (D_X V)(\lambda_E^q(\sigma), \theta)\, d\sigma \Bigg]. \end{aligned} \tag{4.12}$$

Proof. By the fundamental theorem of calculus, the change of energy along an orbit segment starting at $\tilde{z}_\varepsilon^+$ and following it for a time interval $s \in [a/\sqrt{2E}, L/\sqrt{2E}]$ is

$$\int_{a/\sqrt{2E}}^{L/\sqrt{2E}} \frac{d}{dt} H_\varepsilon(\xi_s^\varepsilon(z_\varepsilon^+), \chi_s^\varepsilon(\theta))\, ds = \varepsilon^3 \int_{a/\sqrt{2E}}^{L/\sqrt{2E}} (D_X V)(\lambda_E^q(s), \theta)\, ds + O(\varepsilon^4 |\ln \varepsilon|).$$

Combining this with (4.9), we see that, over an elementary building block, the energy has changed by

$$\begin{aligned} G(\tilde{z}_\varepsilon^-) &= \Delta(\tilde{z}_\varepsilon^-) + \int_{a/\sqrt{2E}}^{L/\sqrt{2E}} \frac{d}{dt} H_\varepsilon(\xi_s^\varepsilon(z_\varepsilon^+), \chi_s^\varepsilon(\theta))\, ds \\ &= \varepsilon^3 \left[\Delta_1(\tilde{z}_0^-) + \int_{a/\sqrt{2E}}^{L/\sqrt{2E}} (D_X V)(\lambda_E^q(s), \theta)\, ds \right] + O(\varepsilon^4 |\ln \varepsilon|). \end{aligned}$$

In conclusion,

$$G(\tilde{z}_\varepsilon^-) = \varepsilon^3 G_1(E, s, \theta) + O(\varepsilon^4 |\ln \varepsilon|). \qquad \square$$

Similarly to (4.10), we have the following rescaling property of G_1 relative to action-angle coordinates:

$$G_1(J, \phi, \theta) = \frac{1}{J} G_1(\sqrt{2}, J\phi, \theta). \tag{4.13}$$

We make two important remarks.

Remark 4.3. For the scattering map $\tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-) = \tilde{z}_\varepsilon^+$, there is no trajectory of the system asymptotic to $\tilde{z}_\varepsilon^-$ in the past and to $\tilde{z}_\varepsilon^+$ in the future. Rather, $\psi^\varepsilon_{T_-}(\tilde{z}_\varepsilon)$ will approach $\psi^\varepsilon_{T_-}(\tilde{z}_\varepsilon^-)$ as $T_- \to -\infty$, and $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon)$ will approach $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon^+)$ as $T_+ \to \infty$. Here $\tilde{z}_\varepsilon = (\tilde{\Omega}^-)^{-1}(\tilde{z}_\varepsilon^-) = (\tilde{\Omega}^+)^{-1}(\tilde{z}_\varepsilon^+)$. For $-T_- = T_+ = K|\ln\varepsilon|$, the change of energy $\Delta(\tilde{z}_\varepsilon^-)$ along the homoclinic trajectory from $\psi^\varepsilon_{T_-}(\tilde{z}_\varepsilon)$ to $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon)$ is the same as in (4.9), up to an error term which is subdominant.

Remark 4.4. Similarly, if instead of fixing the angle shift to be a constant value L, we allow choosing, for each elementary building block, a value of L which is constant plus an $O(\varepsilon)$-term, the change of energy $G(\tilde{z}_\varepsilon^-)$ is the same as in (4.11), up to an error term which is subdominant.

In Section 5.3 we will show that there is a trajectory of the dynamics that follows closely the pseudo-orbit consisting of the segment of the homoclinic trajectory from $\psi^\varepsilon_{T_-}(\tilde{z}_\varepsilon)$ to $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon)$, followed by a segment of the trajectory of the inner flow $(\psi^\varepsilon_s)_{|\tilde{\Lambda}_\varepsilon}$ with initial point $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon^+)$.

Based on these remarks, we will keep in mind that to an elementary building block we can associate the following objects:

- one application of the scattering map $\tilde{S}_\varepsilon(\tilde{z}_\varepsilon^-) = \tilde{z}_\varepsilon^+$, plus one segment of the trajectory of the inner flow $(\psi^\varepsilon_s)_{|\tilde{\Lambda}_\varepsilon}$ with initial point $\tilde{z}_\varepsilon^+$;
- a pseudo-orbit, consisting of a segment of a homoclinic orbit from $\psi^\varepsilon_{T_-}(\tilde{z}_\varepsilon)$ to $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon)$, followed by a segment of the trajectory of the inner flow $(\psi^\varepsilon_s)_{|\tilde{\Lambda}_\varepsilon}$ with initial point $\psi^\varepsilon_{T_+}(\tilde{z}_\varepsilon^+)$;
- a true orbit, which closely follows the pseudo-orbit described above.

The change of energy along either one of these objects is given by the estimate in Proposition 4.2, up to a subdominant error term.

4.3. Generic set of potentials

In this section we specify the set $\mathcal{V}'$ of potentials claimed in Theorem 2.1. The potentials $V \in \mathcal{V}'$ are required to satisfy a condition that ensures that one can achieve consistent energy growth by applying the scattering map followed by the inner dynamics as in Proposition 4.2.

A key observation is that since there exist a closed hyperbolic geodesic λ_E and a corresponding transverse homoclinic orbit γ_E, by the Birkhoff–Smale Homoclinic Orbit Theorem (see, e.g., [55]) there exist, in fact, at least two geometrically distinct homoclinic orbits to the same geodesic. The existence of at least two homoclinic orbits can also be deduced via variational methods. We shall denote a pair of such homoclinic orbits by γ_E^1, γ_E^2, and the associated scattering maps $\tilde{S}_\varepsilon^1$, $\tilde{S}_\varepsilon^2$. We denote the corresponding leading terms G_1 from Proposition 4.2 by G_1^1, G_1^2, respectively. The main idea is that, under some generic condition on the potential V, utilizing one of the homoclinic orbits to grow energy

will be more advantageous than utilizing the other, so one can select which one of the two homoclinic orbits to use in a way to ensure a net gain of energy over time.

To express the generic condition on the potentials $V \in \mathcal{V}'$, we write G_1^1, G_1^2 in action-angle coordinates. Indeed, the energy growth (4.11) along an elementary building block depends only on the initial point $\tilde{z}_\varepsilon^- = (k_\varepsilon(\lambda_{J^2/2}(\phi)), \theta) \in \tilde{\Lambda}_\varepsilon$ of the block, and the leading term G_1 in (4.12) depends on the corresponding angle-action coordinates and parameter value (J, ϕ, θ) corresponding to $\tilde{z}_\varepsilon^-$.

We define $\mathcal{V}'$ to be the set of potentials V for which the following non-degeneracy condition holds:

A4. Fix $J_0 = \sqrt{2}$ so the corresponding energy level of H_0 is $E_0 = 1$. Fix a non-trivial uniformly recurrent point $\theta_0 \in N$. For the geodesic flow at this energy level there exist two geometrically different homoclinic trajectories $\gamma_{E_0}^1, \gamma_{E_0}^2$ to the same closed geodesic λ_{E_0}, and a common domain $\tilde{U}$ for the corresponding scattering maps $\tilde{S}_\varepsilon^1, \tilde{S}_\varepsilon^2$ such that

$$\sup_{\phi_1} G_1^1(J_0, \phi_1, \theta_0) \neq \sup_{\phi_2} G_1^2(J_0, \phi_2, \theta_0), \tag{4.14}$$

where G_1^1, G_1^2 are the leading terms of the energy gain, given by (4.12).

In the above, it is understood that the angles ϕ_1, ϕ_2 are restricted to some closed intervals where $(\lambda_{J_0^2/2}(\phi_1), \theta_0) \in \tilde{U}$, $(\lambda_{J_0^2/2}(\phi_2), \theta_0) \in \tilde{U}$, where $\tilde{k}_\varepsilon(\tilde{U})$ is the domain $(\tilde{U}_\varepsilon^1)^-$ of $\tilde{S}_\varepsilon^1$ and also in the domain $(\tilde{U}_\varepsilon^2)^-$ of $\tilde{S}_\varepsilon^2$. Of course, a domain $\tilde{U}$ as in condition *A4* is not unique. The condition requires only the existence of at least one domain $\tilde{U}$ on which (4.14) holds.

We notice that because we have assumed that the closed geodesics in the unit tangent bundle have transverse homoclinic connections for the geodesic flow in the unit tangent bundle, we can define the projections along each of the points on the orbit. Similarly, we can lift for any value of the energy by the scaling invariance.

Therefore, we can always define locally two scattering maps and we can continue them. The only obstruction to defining a scattering map in an arbitrary domain is that the local continuation along a closed loop may have some monodromy. Each of the scattering maps can be defined on any domain that does not contain essential circles, i.e., non-contractible loops of the cylinder. Hence, the assumption in *A4* that the two scattering maps have a common domain is satisfied automatically.

Condition *A4* is a condition on V along trajectories of the unperturbed system, as shown by (4.12). This is an explicit condition, which can be verified for a given V along a given closed geodesic and a given pair of homoclinic orbits to that closed geodesic.

Assuming condition *A4*, suppose, without loss of generality, that

$$\sup_{\phi_1} G_1^1(J_0, \phi_1, \theta_0) > \sup_{\phi_2} G_1^2(J_0, \phi_1, \theta_0).$$

Then there exists $\delta > 0$, ϕ_* and an open neighborhood $\mathcal{P} \subseteq N$ of θ_0 such that

$$G_1^1(J_0, \phi_*, \theta) - G_1^2(J_0, \phi, \theta) \geq 2\delta \tag{4.15}$$

for all $\theta \in \mathcal{P}$ and all ϕ with $\tilde{k}_\varepsilon(\lambda_{E_0}(\phi), \theta_0)$ in the domain of $\tilde{S}^2_\varepsilon$. Moreover, we can choose the neighborhood $\mathcal{P}$ to be a flow box for the flow χ^ε on N, i.e., a homeomorphic copy in N of a d-dimensional open rectangle, of the form $\{\chi^\varepsilon_s(\theta) \mid \theta \in \Sigma,\ s \in (-\rho, \rho)\}$, where Σ is a $(d-1)$-dimensional open disk transverse to the flow χ^ε_s, and $\rho > 0$ is sufficiently small so that the flow χ^ε_s is transverse to each surface $\{\chi^\varepsilon_{s_0}(\theta) \mid \theta \in \Sigma\}$ for each $s_0 \in (-\rho, \rho)$.

Condition *A4* is formulated in terms of the geodesic and homoclinic orbits at some fixed energy level $E_0 = 1$, corresponding to $J_0 = \sqrt{2}$. By the rescaling property of the geodesic flow, and by the corresponding rescaling property (4.13), it follows that

$$G^1_1\left(J, \frac{1}{J}\phi_*, \theta\right) - G^2_1(J, \phi', \theta) \geq \frac{2\delta}{J} \tag{4.16}$$

for all $\theta \in \mathcal{P}$ and all $\phi' = \frac{1}{J}\phi$ with ϕ restricted as before. Since we restrict to the interval $H_\varepsilon \in [1, 2]$, we have $\sqrt{2} \leq J \leq 2$, hence

$$G^1_1\left(J, \frac{1}{J}\phi_*, \theta\right) - G^2_1(J, \phi', \theta) \geq \delta$$

for all $\theta \in \mathcal{P}$ and all ϕ' as before.

Lemma 4.5. *Given a geodesic flow, a closed geodesic λ and two geometrically distinct homoclinic orbits $\gamma^1_{E_0}$, $\gamma^2_{E_0}$, at the energy level $E_0 = 1$, the set $\mathcal{V}'$ of potentials V that satisfy assumption A4 is open and dense in the $C^{1;r}$-topology of the set of all potentials.*

Proof. To reach a contradiction we assume that, for some domain $\tilde{U}$, condition *A4* does not hold.

Let $\sup_{\phi_1} G^1_1(J_0, \phi_1, \theta_0) = G^1_1(J_0, \phi^*_1, \theta_0)$ for some ϕ^*_1, and $\sup_{\phi_2} G^2_1(J_0, \phi_2, \theta_0) = G^2_1(J_0, \phi^*_2, \theta_0)$ for some ϕ^*_2. Then we have

$$G^1_1(J_0, \phi^*_1, \theta_0) = G^2_1(J_0, \phi^*_2, \theta_0). \tag{4.17}$$

Since for fixed (J_0, ϕ^*_1) and (J_0, ϕ^*_2) the functions $G^1_1(J_0, \phi^*_1, \theta_0)$ and $G^2_1(J_0, \phi^*_2, \theta_0)$, respectively, considered as functionals of V, are continuous when the set $\mathcal{V}$ of potentials is given the C^0-topology, it is clear that (4.17) defines a C^0-closed set (intersection of closed sets), and therefore condition *A4* holds in a C^0-open set of potentials.

To complete the proof of Lemma 4.5, it suffices to show that, given a potential V that satisfies (4.17), there is an arbitrarily small perturbation of V, relative to the $C^{1;r}$-topology, which does not satisfy (4.17).

The construction is very clear. We note that G^j_1, $j \in \{1, 2\}$, is a sum of integrals over several trajectory segments of the geodesic flow: some are segments of closed geodesics in Λ_0, which are recurrent, and one of them is a segment of a homoclinic trajectory.

Because the two homoclinic orbits $\gamma^1_{E_0}, \gamma^2_{E_0}$ are geometrically different, we can find an $s_0 \in \mathbb{R}$ and a small enough ball $B \subset T^*M$ centered at $\gamma^1_{E_0}(s_0)$ in such a way that

$B \cap \lambda_{E_0} = B \cap \gamma^2_{E_0} = \emptyset$. Moreover, we can choose the ball B such that there exists a small interval I around s_0 such that $\gamma^1(\mathbb{R}) \cap B = \gamma^1(I)$. Choose a small ball B' in N centered at θ_0.

Now we choose a $C^{1;r}$ function $W : M \times N \to \mathbb{R}$ with support in $B \times B'$ so that $\nabla_\theta W(\gamma^1_{E_0}(s), \theta_0) \cdot X(\theta_0) \geq \rho > 0$ for all $s \in I$ and $\theta \in B'$, and for some $\rho > 0$. It follows from (4.4) that perturbing V to $V + W$ yields $G^1_1(J_0, \phi^*_1, \theta_0) \neq G^2_1(J_0, \phi^*_2, \theta_0)$. □

4.4. Gain of energy along sequences of elementary building blocks

We will estimate the gain of the scaled energy H_ε in time, along some suitably chosen sequences of elementary building blocks, for potentials $V \in \mathcal{V}'$. We will account for both the scaled time s and the physical time t.

Assume that V satisfies condition *A4* from the previous section. Then for given any J_0 with $\|J_0\| = \sqrt{2}$ there exist ϕ_* with $\tilde{k}_\varepsilon(\lambda_{E_0}(\phi_*), \theta_0) \in (\tilde{U}^1_\varepsilon)^-$ and a flow box $\mathcal{P} \subseteq N$, of size $O(1)$, with the property that if $\theta \in \mathcal{P}$ and $\tilde{k}_\varepsilon(\lambda_{E_0}(\phi), \theta_0) \in (\tilde{U}^2_\varepsilon)^-$, then

$$G^1_1(J_0, \phi_*, \theta) - G^2_1(J_0, \phi, \theta) \geq 2\delta.$$

From the rescaling property of the geodesic flow, it follows that for any $J \in [\sqrt{2}, 2]$ (corresponding to the energy range $E \in [1, 2]$ fixed in Subsection 4.2), there exists $\phi_*(J) = \phi/J_*$ such that

$$G^1_1(J, \phi_*(J), \theta) - G^2_1(J, \phi, \theta) \geq \delta \tag{4.18}$$

for all $\theta \in \mathcal{P}$ and all ϕ in the appropriate domain.

The flow χ^ε_s is slow, and so is its time-1 map. It takes a time $O(1/\varepsilon)$ to travel a distance $O(1)$. The flow box $\mathcal{P}$ was chosen of the form $\{\chi^\varepsilon_s(\theta) \mid \theta \in \Sigma, s \in (-\rho, \rho)\}$ with Σ as in Subsection 4.3.

Assume that $\theta_0 \in N$ is a uniformly recurrent point with $X(\theta_0) \neq 0$. Choosing the neighborhood $\mathcal{P}$ small enough ensures that the trajectory of θ_0 will successively leave $\mathcal{P}$ and return to $\mathcal{P}$. The lengths of the time intervals when the trajectory of θ_0 moves through $\mathcal{P}$ are uniformly bounded above and below, and because of the uniform recurrence hypothesis *A3*, so are the lengths of the time intervals when the trajectory of θ_0 moves through $N \setminus \mathcal{P}$. More precisely, there exist $0 < \tau_0 < \tau'_0$, independent of ε, such that the trajectory of θ_0 spends a scaled time between τ_0/ε and τ'_0/ε in $\mathcal{P}$, and there exist $0 < \tau_1 < \tau'_1$, independent of ε, such that the trajectory of θ_0 spends a scaled time between τ_1/ε and τ'_1/ε in $N \setminus \mathcal{P}$ between successive returns to $\mathcal{P}$.

Now we compute the growth of energy during a range of scaled time $\Delta s = 1/\varepsilon^2$. We follow a sequence of elementary building blocks of the type $B^1(J, \phi, \theta)$, $B^2(J, \phi, \theta)$, where the superscripts correspond to the two choices of homoclinic orbits/scattering maps γ^1_E, γ^2_E respectively, where the succession of blocks is chosen as follows. When $\theta \in \mathcal{P}$ we use blocks of the type $B^1(J, \phi_*(J), \theta)$, where $\phi_*(J)$ is defined as before. When $\theta \notin \mathrm{cl}(\mathcal{P})$ we use blocks of the type $B^2(J, \phi_0, \theta)$ for some ϕ_0 fixed. Thus, the sequence is composed of strings of the type $B^1(J, \phi_*(J), \theta)$, alternating with strings of the type $B^2(J, \phi_0, \theta)$; the proportion of time when we switch from $B^1(J, \phi_*(J), \theta)$ to $B^2(J, \phi_0, \theta)$ or vice versa is $O(\varepsilon)$.

We compute the growth of energy along such a sequence of building blocks spanning a range of scaled time of $1/\varepsilon^2$. The initial condition for the dynamics on N is the point θ_0 which is assumed to be uniformly recurrent. The J-coordinate along the pseudo-orbit takes the successive values $J(n)$, and the θ-coordinate along the pseudo-orbit takes the successive values $\theta(n)$; the ϕ-coordinate is maintained fixed $\phi = \phi_0$. We have

$$\begin{aligned}\Delta H_\varepsilon = \varepsilon^3\Big[&\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\in\mathcal{P},\,\theta((n+1)\varepsilon)\in\mathcal{P}}} G_1^1(J(n),\phi_*(J(n)),\theta(n))\\ &+\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\notin\mathcal{P},\,\theta((n+1)\varepsilon)\notin\mathcal{P}}} G_1^2(J(n),\phi_0,\theta(n))\\ &+\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\in\mathcal{P},\,\theta((n+1)\varepsilon)\notin\mathcal{P}}} G_1^2(J(n),\phi_0,\theta(n))\\ &+\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\notin\mathcal{P},\,\theta((n+1)\varepsilon)\in\mathcal{P}}} G_1^2(J(n),\phi_0,\theta(n))\Big] + O(\varepsilon^2|\ln\varepsilon|), \end{aligned} \tag{4.19}$$

where the above error term is due to the accumulation error term of $O(\varepsilon^4|\ln\varepsilon|)$ from (4.12) over $1/\varepsilon^2$ time steps. The terms corresponding to the times n when $\theta(n) \in \mathcal{P}$, $\theta((n+1)\varepsilon) \notin \mathcal{P}$ or $\theta(n) \notin \mathcal{P}$, $\theta((n+1)\varepsilon) \in \mathcal{P}$ are $O(\varepsilon \cdot 1/\varepsilon^2) = O(1/\varepsilon)$, and since G_1^1, G_1^2 are bounded, they contribute to a combined error term $O(\varepsilon^3 \cdot 1/\varepsilon) = O(\varepsilon^2)$, which is subdominant.

Thus we can write

$$\begin{aligned}\Delta H_\varepsilon \geq \varepsilon^3\Big[&\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\in\mathcal{P},\,\theta((n+1)\varepsilon)\in\mathcal{P}}} G_1^1(J(n),\phi_0,\theta(n))\\ &+\sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\notin\mathcal{P},\,\theta((n+1)\varepsilon)\notin\mathcal{P}}} G_1^2(J(n),\phi_0,\theta(n))\Big] + O(\varepsilon^2|\ln\varepsilon|). \end{aligned} \tag{4.20}$$

Now we rearrange the summation above to estimate the total gain of energy while $\theta \in \mathcal{P}$:

$$\begin{aligned}\Delta H_\varepsilon &\geq \varepsilon^3 \sum_{\substack{n\in[0,1/\varepsilon^2]\\ \theta(n)\in\mathcal{P},\,\theta((n+1)\varepsilon)\in\mathcal{P}}} [G_1^1(J(n),\phi_0,\theta(n)) - G_1^2(J(n),\phi_0,\theta(n))]\\ &\quad + \varepsilon^3 \sum_{n\in[0,1/\varepsilon^2]} G_1^2(J(n),\phi_0,\theta(n)) + O(\varepsilon^2|\ln\varepsilon|)\\ &\geq \varepsilon^3\delta\frac{\tau_0}{\varepsilon^2} + \varepsilon^3 \sum_{n\in[0,1/\varepsilon^2]} G_1^2(J(n),\phi_0,\theta(n))) + O(\varepsilon^2|\ln\varepsilon|)\\ &= \varepsilon\tau_0\delta + \varepsilon^3 \sum_{n\in[0,1/\varepsilon^2]} G_1^2(J(n),\phi_0,\theta(n)) + O(\varepsilon^2|\ln\varepsilon|). \end{aligned} \tag{4.21}$$

Now we treat the remaining summation as a Riemann sum of mesh 1 and we approximate it by a Riemann integral. Since $J(\sigma)$ changes by at most $O(\varepsilon^2)$ and $\theta(\sigma)$ changes by at most $O(\varepsilon)$ over each interval $[n, n+1]$, we have $\max_{\sigma\in[n,n+1]}\left|\frac{d}{d\sigma}G_1(J(\sigma),\phi,\theta(\sigma))\right| = O(\varepsilon)$ for all n. Hence the error in approximating the integral on a subdivision by a term of the Riemann sum is less than $O(\varepsilon)$, and since there are $1/\varepsilon^2$ terms in the Riemann sum, the error in approximating the integral by the Riemann sum is $O(1/\varepsilon)$. Taking into account the ε^3-leading factor we obtain

$$\varepsilon^3 \sum_{n\in[0,1/\varepsilon^2]} G_1^2(J(n),\phi_0,\theta(n)) \le \varepsilon^3 \int_0^{1/\varepsilon^2} G_1^2(J(\sigma),\phi_0,\theta(\sigma))\,d\sigma + O\left(\varepsilon^3\frac{1}{\varepsilon}\right)$$
$$\le \varepsilon^3[A_1^2(J(1/\varepsilon^2),\phi_0,\theta(1/\varepsilon^2)) - A_1^2(J(0),\phi_0,\theta(0))] + O(\varepsilon^2) = O(\varepsilon^2), \qquad (4.22)$$

where $G_1^2 = D_X A_1^2$, where the antiderivative A_1^2 of G_1^2 is bounded by the compactness of $(\Lambda_0 \times N) \cap H_0^{-1}[1,2]$. (Here we have used the fact that $\frac{d}{dt}A_1^2(J(t),\phi_0,\theta(t)) = \frac{dA_1^2}{dJ}\dot{J} + D_X A_1^2 = D_X A_1^2 + O(\varepsilon^2) = G_1^2 + O(\varepsilon^2)$.)

From (4.21) we conclude that, under the non-degeneracy assumption *A4*, the gain of energy following a string of elementary building blocks of the type $B^1(J,\phi_0,\theta)$, $B^2(J,\phi_0,\theta)$ over a range of scaled time $1/\varepsilon^2$ satisfies

$$\Delta H_\varepsilon \ge \varepsilon\tau_0\delta + O(\varepsilon^2|\ln\varepsilon|). \qquad (4.23)$$

Note that during this time interval we do not leave the scaled energy interval $E \in [1,2]$ fixed at the beginning of the argument.

Thus, during a time period of $1/\varepsilon^2$, moving along the pseudo-orbits corresponding to the elementary building blocks of the type $B^1(J,\phi_0,\theta)$, $B^2(J,\phi_0,\theta)$, in the specified order, corresponds to a scaled energy growth of $O(\varepsilon)$. Since $H_\varepsilon = \varepsilon^2 H_0$ and $\Delta s = \Delta t/\varepsilon$, we obtain a physical energy growth of $\Delta H = O(1/\varepsilon^2)$ during a physical time interval $\Delta t = O(1/\varepsilon^2)$, that is, a linear growth rate of the physical energy in physical time.

Remark 4.6. If we construct a sequence of elementary building blocks corresponding to a single homoclinic $\gamma_{E_0}^j$, $j \in \{1,2\}$, for a time of $O(1/\varepsilon)$, and if the variable ϕ_0 in the construction is fixed as above, a calculation as in (4.22) shows that the change of scaled energy is $O(\varepsilon^3)$. This implies that G_1^j cannot be always positive or always negative for all this time. Along any sequence of elementary building blocks for which the external flow χ^ε returns to a small neighborhood of its initial point θ_0, there will always be regions in $\tilde{\Lambda}_0$ where G_1^j is positive as well as regions where G_1^j is negative. For the same reason, we cannot have $G_1^1 > G_1^2$ for all time. Hence, besides a flow box $\mathcal{P} \subseteq N$ such that $G_1^1(J_0,\phi_0,\theta) - G_1^2(J_0,\phi_0,\theta) > 0$ for all $\theta \in \mathcal{P}$, there should also exist another flow box $\mathcal{P}'$ such that $G_1^2(J_0,\phi_0,\theta) - G_1^1(J_0,\phi_0,\theta) > 0$ for all $\theta \in \mathcal{P}'$.

4.5. Sequences of elementary building blocks achieving unbounded growth of energy

The above construction of pseudo-orbits can be continued for a time $O(1/\varepsilon^3)$ to achieve an energy growth corresponding to the whole interval $H_\varepsilon \in [1,2]$. Since $\varepsilon = 1/\sqrt{E_*}$, the corresponding growth of physical energy is $H = \varepsilon^{-2}H_\varepsilon \in [E_*, 2E_*]$. To grow the

physical energy to infinity, we repeat the procedure, re-initializing the process starting with $\varepsilon = 1/\sqrt{2E_*}$. For this new value of the small parameter ε in (3.3), growing the physical energy $H \in [2E_*, 4E_*]$ amounts to growing the scaled energy $H_\varepsilon \in [1, 2]$. So all the estimates made in this subsection remain valid and carry through. Thus, this construction of pseudo-orbits featuring energy growth can be repeated indefinitely.

In Section 5 we show the existence of true orbits 'shadowing' the pseudo-orbits constructed in this section.

4.6. Sequences of elementary building blocks achieving symbolic dynamics

In order to construct elementary building blocks along which the energy follows a prescribed path $\mathcal{E} : [0, \infty) \to \mathbb{R}$, we alternate elementary building blocks leading to energy growth with blocks leading to energy loss. Choosing the proportions of the energy growth and of the energy loss allows us to control the energy change. In particular, we can obtain rates close to zero by alternating blocks which gain energy with blocks which loose energy. The change of energy along an elementary building block is not more than ε^3 in the scaled variables, which corresponds to a change of energy of $E^{-1/2}$ in the physical variables. In this way, we can follow the prescribed energy path $\mathcal{E}$ up to $E^{-1/2}$. Then, once the sequence of elementary building blocks is constructed, in Section 5.3 we will construct a sequence of correctly aligned windows along this sequence of blocks, and apply the shadowing Theorem 5.3 to obtain an orbit that follows these windows.

Remark 4.7. We remark that in the present mechanism we do not achieve small rates of growth by staying near a KAM torus, as in [29, 30], rather the orbits we construct are performing homoclinic excursions most of the time. Thus, these orbits are very different from the previously constructed orbits.

5. Existence of orbits following sequences of elementary building blocks

In this section, we show that we can concatenate infinitely many elementary building blocks as above, and that there exists a true orbit that follows the pseudo-orbit underlying those blocks, thus achieving infinite energy growth. Since our system is not hyperbolic, the classical shadowing lemma for hyperbolic systems, saying that any pseudo-orbit can be 'shadowed' by a true orbit, does not apply. We will show that, nevertheless, the pseudo-orbits constructed in the previous section can be approximated by a true orbit. For this, we use a topological argument based on correctly aligned windows. This argument is constructive and robust, so it allows us to also estimate the energy growth rate along the resulting orbit.

5.1. Topological method

In this section we briefly review the topological method of correctly aligned windows, following [85, 45, 47]. Earlier versions of the method go back to [21, 35, 34].

A window is a triple consisting of a mapping, a set, and a partition of the boundary of that set: the mapping is a homeomorphism from a multi-dimensional rectangle in some

Euclidean space to a manifold, the set is the image of the multi-dimensional rectangle through the homeomorphism, and the partition divides the boundary of the set into an exit set and an entry set, which play a dynamical role.

Definition 5.1. An (n_1, n_2)-*window* in an m-dimensional manifold M, where $n_1 + n_2 = m$, is an ensemble $(W, W^{\text{exit}}, W^{\text{entry}}, c)$ consisting of:

(1) a homeomorphism $c : \text{dom}(c) \to \text{im}(c)$, where $\text{dom}(c)$ is an open neighborhood of $[0, 1]^{n_1} \times [0, 1]^{n_2} \subseteq \mathbb{R}^m$, and $\text{im}(c)$ is an open set in M,

(2) a homeomorphic copy $W := c([0, 1]^{n_1} \times [0, 1]^{n_2}) \subseteq \text{im}(c)$ of $[0, 1]^{n_1} \times [0, 1]^{n_2}$ in M,

(3) an 'exit set'
$$W^{\text{exit}} := c(\partial[0, 1]^{n_1} \times [0, 1]^{n_2})$$
and an 'entry set'
$$W^{\text{entry}} := c([0, 1]^{n_1} \times \partial[0, 1]^{n_2}).$$

We adopt the following notation: $W_c = c^{-1}(W) = [0, 1]^{n_1} \times [0, 1]^{n_2}$, $(W^{\text{exit}})_c = c^{-1}(W^{\text{exit}}) = \partial[0, 1]^{n_1} \times [0, 1]^{n_2}$, and $(W^{\text{entry}})_c = c^{-1}(W^{\text{entry}}) = [0, 1]^{n_1} \times \partial[0, 1]^{n_2}$. When the coordinate system c is evident from context, we suppress the subscript c from the notation.

Informally, two windows are correctly aligned under some map, provided that the image of the first window under the map crosses the second window all the way through and across its exit set. Below we present a version of the definition of correct alignment that is sufficient for the purpose of this paper. More details can be found in [85]. Given two windows $(W_1, W_1^{\text{exit}}, W_1^{\text{entry}}, c_1)$ and $(W_2, W_2^{\text{exit}}, W_2^{\text{entry}}, c_2)$ and a continuous map $f : M \to M$ with $f(\text{im}(c_1)) \subseteq \text{im}(c_2)$, we will denote $f_{c_1,c_2} = c_2^{-1} \circ f \circ c_1$.

Definition 5.2. The window W_1 is *correctly aligned* with the window W_2 under f if the following conditions are satisfied:

(1) When $n_1, n_2 \neq 0$, the conditions are:

(1.i) There exists a continuous homotopy $h : [0, 1] \times (W_1)_{c_1} \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ with
$$h_0 = f_{c_1,c_2},$$
$$h([0, 1], (W_1^{\text{exit}})_{c_1}) \cap (W_2)_{c_2} = \emptyset,$$
$$h([0, 1], (W_1)_{c_1}) \cap (W_2^{\text{entry}})_{c_2} = \emptyset.$$

(1.ii) There exists a linear map $A : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1}$ such that

(1.ii.a) $h_1(x, y) = (Ax, 0)$ for $x \in [0, 1]^{n_1}$ and $y \in [0, 1]^{n_2}$,
(1.ii.b) $A(\partial[0, 1]^{n_1}) \subset \mathbb{R}^{n_1} \setminus [0, 1]^{n_1}$.

(2) When $n_2 = 0$, the conditions are:

(2.i) $(W_1^{\text{exit}})_{c_1} = (\partial W_1)_{c_1}$, $(W_2^{\text{exit}})_{c_2} = (\partial W_2)_{c_2}$,
(2.ii) $(W_2)_{c_2} \subseteq \text{int}(f_{c_1,c_2}((W_1)_{c_1}))$,

(3) When $n_1 = 0$, the conditions are:

(3.i) $(W_1^{\text{exit}})_{c_1} = \emptyset$, $(W_2^{\text{exit}})_{c_2} = \emptyset$,
(3.ii) $\text{int}(W_2)_{c_2} \supseteq f_{c_1,c_2}((W_1)_{c_1})$.

The correct alignment of windows is robust, in the sense that if two windows are correctly aligned under a map, then they remain correctly aligned under a sufficiently small C^0-perturbation of the map. This allows one to verify the correct alignment of long, finite sequences of windows by breaking them into shorter, finite sequences of windows whose correct alignment can be easily controlled by perturbative arguments. This property will not be used in this paper.

Also, the correct alignment satisfies a natural product property. Given two windows and a map, if each window can be written as a product of window components, and if the components of the first window are correctly aligned with the corresponding components of the second window under the appropriate components of the map, then the first window is correctly aligned with the second window under the given map. We refer to [45, 47] for details.

The following result can be thought of as a topological version of the Shadowing Lemma. Note that this result does not assume that the system is hyperbolic.

Theorem 5.3 ([85]). *Assume that $\{W_i\}_{i\in\mathbb{Z}}$ is a bi-infinite sequence of (n_1, n_2)-windows in M, and $\{f_i\}_{i\in\mathbb{Z}}$ are continuous maps on M. If W_i is correctly aligned with W_{i+1} under f_i for every $i \in \mathbb{Z}$, then there exists a point $p \in W_0$ such that*

$$(f_i \circ \cdots \circ f_0)(p) \in W_{i+1} \quad \text{for all } i \in \mathbb{Z}.$$

Assume now that $\{W_i\}_{i\in\{0,\dots,d\}}$ is a finite sequence of (n_1, n_2)-windows in M, and $\{f_i\}_{i\in\{0,\dots,d\}}$ are continuous maps on M. If W_i is correctly aligned with W_{i+1} under f_i for every $i = 0, \dots, d-1$, and W_d is correctly aligned with W_0 under f_d, then there exists a point $p \in W_0$ such that

$$(f_d \circ \cdots \circ f_0)(p) = p.$$

A sequence $\{W_i\}_i$ of windows as above will be referred to as a *sequence of correctly aligned windows*. Note that the verification of the correct alignment of the sequence amounts to verifying correct alignment relations between successive pairs W_i, W_{i+1}. A consequence of this fact, which is important for our applications, is that the concatenation of finite sequences of correctly aligned windows is a finite sequence of correctly aligned windows. That is, if $W_0, \dots, W_k$ is a sequence of correctly aligned windows and $W_k, \dots, W_l$ is a sequence of correctly aligned windows, then so is $W_0, \dots, W_l$ (for simplification, we do not specify the mappings under which the correct alignment is realized).

In the context of this paper, the maps f_i from Theorem 5.3 will be different powers of the time-1 map associated to the flow.

We emphasize again the difference between hyperbolic dynamics and correct alignment of windows. To assert that an orbit is hyperbolic one needs to examine the expansion and contraction rates of the derivative of the map along the whole orbit. The concatenation of hyperbolic segments could fail to be hyperbolic (if the stable and unstable directions do not match). In contrast, to assert that a sequence of windows is correctly aligned, one only needs to verify that the image of one window under the map is correctly aligned with the next window in the sequence. Also, concatenations of finite sequences of correctly aligned windows are correctly aligned.

5.2. Reduction to a discrete dynamical system

We reduce the perturbed geodesic flow to a discrete dynamical system by considering the time-1 map of the flow ψ^ε—where the time refers to the rescaled time s in (3.3)—which we denote F_ε. Also, we denote by χ_1^ε the time-1 map (relative to the rescaled time s) associated to the flow χ^ε on N. Since $s = t/\varepsilon$, the time-1 map relative to the rescaled time s is the time-ε map relative to the physical time t.

The distinguished geometric objects for the perturbed geodesic flow, described in the earlier sections, give rise to similar objects for the discrete dynamical system. We start by listing these objects and summarizing their properties.

B1. The map $F_\varepsilon : T^*M \times N \to T^*M \times N$ is a C^{r-1}-diffeomorphism.

B2. The manifold $\tilde\Lambda_\varepsilon = \Lambda_\varepsilon \times N \subseteq T^*M \times N$ is a normally hyperbolic invariant manifold for F_ε, of dimension $d+2$; this manifold has stable and unstable manifolds $W^s(\tilde\Lambda_\varepsilon)$ and $W^u(\tilde\Lambda_\varepsilon)$, of dimension $d+n+1$.

B3. There exist exponential rates $0 < \lambda_- < \lambda_+ < \lambda_1 < 1 < \mu_1 < \mu_- < \mu_+$ such that $\lambda_- < \|DF_{\varepsilon|E^s_{\varepsilon,z}}\| < \lambda_+$, $\mu_- < \|DF_{\varepsilon|E^u_{\varepsilon,z}}\| < \mu_+$, $\lambda_1 < \|DF_{\varepsilon|E^c_{\varepsilon,z}}\| < \mu_1$, where E^s_ε and E^u_ε are the stable and unstable bundles in the decomposition $T_z(T^*M \times N) = T_z\tilde\Lambda_\varepsilon \oplus E^s_{\varepsilon,z} \oplus E^u_{\varepsilon,z}$. The above exponential rates can be chosen independently of ε.

B4. The stable and unstable manifolds $W^s(\tilde\Lambda_\varepsilon)$ and $W^u(\tilde\Lambda_\varepsilon)$ have a transverse intersection along a $(d+2)$-dimensional manifold $\tilde\Gamma_\varepsilon$.

B5. For the discrete dynamical system defined by F_ε there exist two scattering maps $\tilde S^j_\varepsilon : \tilde U^{j,-}_\varepsilon \to \tilde U^{j,+}_\varepsilon$ associated to two homoclinic channels $\tilde\Gamma^j_\varepsilon$, $j = 1, 2$, satisfying assumption *A4*. The set $\tilde U^{j,-}_\varepsilon$ is of size $O(1)$ in the sense that there exist open sets $U^{j,-} \subseteq \Lambda_0$ such that $\tilde k_\varepsilon(U^{j,-} \times N) \subseteq \tilde U^{j,-}_\varepsilon$ for $j = 1, 2$, and for all ε sufficiently small, where $\tilde k_\varepsilon : \tilde\Lambda_0 = \Lambda_0 \times N \to \tilde\Lambda_\varepsilon$ is the parametrization of $\tilde\Lambda_\varepsilon$ from Subsection 3.2.

B6. Consider the action-angle coordinates $(J_\varepsilon, \phi_\varepsilon)$ on Λ_ε. The restriction of F_ε to $\tilde\Lambda_\varepsilon$ has the form

$$F_\varepsilon(J_\varepsilon, \phi_\varepsilon, \theta) = \big(J_\varepsilon + O(\varepsilon^2), \phi_\varepsilon + J_\varepsilon + O(\varepsilon^2), \chi_1^\varepsilon(\theta)\big).$$

There exists $\tau > 0$, independent of ε, such that

$$\frac{\partial(\pi_{\phi_\varepsilon} \circ F_{\varepsilon|\Lambda_\varepsilon})}{\partial J_\varepsilon}(J_\varepsilon, \phi_\varepsilon) > \tau \quad \text{for all } (J_\varepsilon, \phi_\varepsilon, \theta) \in \tilde\Lambda_\varepsilon.$$

In particular $F_{\varepsilon|\Lambda_\varepsilon}$ is an integrable twist map in the variables $(J_\varepsilon, \phi_\varepsilon)$, up to order $O(\varepsilon^2)$, with the twist coefficient lower bounded by τ.

B7. Each scattering map $\tilde S^j_\varepsilon : \tilde U^{j,-}_\varepsilon \to \tilde U^{j,+}_\varepsilon$, associated to the homoclinic channel $\tilde\Gamma^j$, $j \in \{1, 2\}$, is of the form

$$\tilde S^j_\varepsilon(J^-_\varepsilon, \phi^-_\varepsilon, \theta^-) = (J^+_\varepsilon, \phi^+_\varepsilon, \theta^+),$$

where

$$\phi^+_\varepsilon = \phi^-_\varepsilon + a + O(\varepsilon^2), \quad |J^+_\varepsilon - J^-_\varepsilon| = O(\varepsilon^2), \quad \theta^+ = \theta^-.$$

B8. There exists a sequence of elementary building blocks, of the type $B^j(J, \phi, \theta)$, $j = 1, 2$, as in Subsection 4.4, along which the scaled energy H_ε grows by $O(\varepsilon)$ in a scaled time interval $\Delta s = O(1/\varepsilon^2)$. Each building block consists of one application of one of the scattering maps, $\tilde{S}^j_\varepsilon$, $j = 1, 2$, succeeded by an orbit of the inner dynamics followed for a scaled time of $O(1)$. The succession of elementary building blocks is chosen as in Subsection 4.4.

The general strategy to prove Theorems 2.1–2.3 is the following. We will choose a sequence of elementary building blocks as before, and will fix a two-dynamics pseudo-orbit determined by it. We will prove that there exists a true orbit that follows this pseudo-orbit. Also, we will estimate the time it takes for such an orbit to perform the trip.

As explained in Subsection 4.2, an elementary building block consists of a segment of a homoclinic orbit $\psi^\varepsilon_s(\tilde{z}_\varepsilon)$, $s \in [T_-, T_+]$, followed by a trajectory segment $(\psi^\varepsilon_s)_{\tilde{\Lambda}_\varepsilon}(\tilde{z}^+_\varepsilon)$ of the flow ψ^ε_s restricted to $\tilde{\Lambda}_\varepsilon$, during a time interval of order $O(1)$. Consider a sequence of $n_1 + 1$ successive elementary building blocks. For each of them, we denote by $\tilde{\gamma}^\varepsilon_i(s)$ the homoclinic segment of the elementary building block, and by $\tilde{\lambda}^\varepsilon_i(s)$ the trajectory segment of the flow ψ^ε_s restricted to $\tilde{\Lambda}_\varepsilon$ corresponding to that building block, where $i = 0, \dots, n_1$. Also, we denote by $z^\varepsilon_{i-1,i}$ the corresponding homoclinic point on $\tilde{\gamma}^\varepsilon_i$, with $i \in \{1, \dots, n_1\}$. We allow the homoclinic orbit segments under study to correspond to different homoclinic manifolds $\tilde{\Gamma}^j_\varepsilon$, with $j = 1, 2$, but we do not make this distinction in the notation (for typographical reasons), since all estimates are uniform.

Each trajectory segment that is a part of an elementary building block can be written as $\tilde{\gamma}^\varepsilon_i(s) = (\gamma^\varepsilon_i(s), \theta(s))$, and $\tilde{\lambda}^\varepsilon_i(s) = (\lambda^\varepsilon_i(s), \theta(s))$, respectively. Since a trajectory segment $\tilde{\lambda}^\varepsilon_i(s)$ is followed for a time $O(1)$, the action J_ε-coordinate along $\lambda^\varepsilon_i(s)$ stays constant up to an $O(\varepsilon^2)$ error. Thus, to the trajectory segment $\lambda^\varepsilon_i(s)$ we can associate a level set $\{J_\varepsilon = J^i_\varepsilon\}$ of the J_ε-coordinate in Λ_ε, which is almost invariant up to $O(\varepsilon^2)$.

For each $i \in \{0, \dots, n_1\}$, denote $\mathcal{L}^\varepsilon_i = \lambda^\varepsilon_i \times N$; this is a $(1+d)$-dimensional manifold in $T^*M \times N$. The stable and unstable manifolds of $\mathcal{L}^\varepsilon_i$ are $(n+d)$-dimensional manifolds $W^s(\mathcal{L}^\varepsilon_i) = \bigcup_{z \in \mathcal{L}^\varepsilon_i} W^s(z)$ and $W^u(\mathcal{L}^\varepsilon_i) = \bigcup_{z \in \mathcal{L}^\varepsilon_i} W^u(z)$, respectively, where the stable and unstable fibers of points are well defined due to the normal hyperbolicity of $\tilde{\Lambda}_\varepsilon$. Note that these sets are not, in general, invariant, as their fibers are not invariant; we have $F_\varepsilon(W^s(z)) \subseteq W^s(F_\varepsilon(z))$ and $F^{-1}_\varepsilon(W^u(z)) \subseteq W^u(F^{-1}_\varepsilon(z))$. By construction, for each $i \in \{1, \dots, n_1\}$, $W^u(\mathcal{L}^\varepsilon_{i-1})$ intersects $W^s(\mathcal{L}^\varepsilon_i)$ transversally along $\{z^\varepsilon_{i-1,i}\} \times N$.

To show the existence of an orbit that follows the pseudo-orbit determined by the given sequence of elementary building blocks, we will use the stable and unstable manifolds to build a chain of sets $\mathcal{L}^i_\varepsilon$, with $i = 0, \dots, n_1$, around which we will construct windows that are correctly aligned. The construction will be done according to the following sequence of steps, which we first describe informally below.

At the first step, for each $i = 1, \dots, n_1$, we construct a pair of windows W^-_{i-1} and W^+_i about the heteroclinic intersection $(\{z^\varepsilon_{i-1,i}\} \times N) \cap \tilde{\Gamma}^j_\varepsilon \subseteq W^u(\mathcal{L}^\varepsilon_{i-1}) \cap W^s(\mathcal{L}^\varepsilon_i)$ such that W^-_{i-1} is correctly aligned with W^+_i under the identity mapping. Here the homoclinic

intersection $\tilde{\Gamma}_\varepsilon^j$ is one of the two homoclinic intersections defined by the distinguished homoclinic orbits γ^j, $j \in \{1, 2\}$, described in assumption *A4*.

At the second step we iterate W_i^+ forward in time along $W^s(\mathcal{L}_i^\varepsilon)$ until its image is contained in some conveniently chosen neighborhood of $\mathcal{L}_i^\varepsilon$ in T^*M, and we construct a new window $\check{W}_i$ near $\mathcal{L}_i^\varepsilon$ such that W_i^+ is correctly aligned with $\check{W}_i$ under some positive iterate of F_ε. Also, we iterate W_{i-1}^- backwards in time along $W^u(\mathcal{L}_{i-1}^\varepsilon)$ until its image is contained in some conveniently chosen neighborhood of $\mathcal{L}_{i-1}^\varepsilon$, and we construct a new window $\hat{W}_{i-1}$ about $\mathcal{L}_{i-1}^\varepsilon$ such that $\hat{W}_{i-1}$ is correctly aligned with W_{i-1}^- under some positive iterate of F_ε. At this step, the neighborhoods of $\mathcal{L}_i^\varepsilon$ and $\mathcal{L}_{i-1}^\varepsilon$ are chosen so that the system of coordinates from Subsection 5.3.1, near the normally hyperbolic invariant manifold $\tilde{\Lambda}_\varepsilon$, is well defined in those neighborhoods. The Λ_ε-component of the window $\hat{W}_{i-1}$ is close to the level set $J_\varepsilon = J_\varepsilon^{i-1}$ of the action coordinate J_ε, and the Λ_ε-component of the window $\check{W}_i$ is close to the level set $J_\varepsilon = J_\varepsilon^i$ of the action coordinate J_ε on Λ_ε.

At the third step we align the window $\hat{W}_i$ corresponding to the heteroclinic connection $W^u(\mathcal{L}_{i-1}^\varepsilon) \cap W^s(\mathcal{L}_i^\varepsilon)$ to the window $\check{W}_i$ corresponding to the heteroclinic connection $W^u(\mathcal{L}_i^\varepsilon) \cap W^s(\mathcal{L}_{i+1}^\varepsilon)$. The Λ_ε-components of these two windows are close to the same level set $J_\varepsilon = J_\varepsilon^i$ of the action coordinate. The construction at the third step concatenates the sequence of correctly aligned windows constructed about one heteroclinic connection with the sequence of correctly aligned windows constructed about the next heteroclinic connection.

At the fourth step we concatenate short sequences of correctly aligned windows constructed as above, obtaining a long sequence of correctly aligned windows that follows the pseudo-orbit underlying the sequence of elementary building blocks that achieves the desired energy growth.

The conclusion is that, once the windows have been constructed, the shadowing result (Theorem 5.3) will provide the existence of true orbits that visit the windows in the prescribed order, hence these orbits will visit the prescribed level sets of the averaged action.

The construction of windows is similar to that in [45]. Therefore we will describe most steps of the construction succinctly. One step that is quite different is the third step, in which we align two windows $\check{W}_i$ and $\hat{W}_i$ about $\mathcal{L}_i^\varepsilon$ under some iterate of the map F_ε. The difference is that in [45] the map is a twist map, while in our case, the map is a twist map in one component and a time-1 map of some general flow in the other component. We will explain this step of the construction in more detail.

Now we explain how the above strategy is used to prove the statements from Theorems 2.1–2.3.

To prove that there exist trajectories along which the energy grows unboundedly we proceed as follows. We fix a potential $V \in \mathcal{V}'$, i.e., satisfying condition *A4*. Under the hypotheses of Theorem 2.1 we choose an initial condition $(J_\varepsilon^0, \phi_0, \theta_0)$, where θ_0 is chosen to be a non-trivial uniformly recurrent point, and ϕ_0 is fixed as in condition *A4*. Under the hypotheses of Theorem 2.2 we choose the same initial condition with an arbitrary θ_0. Starting with this initial condition, we construct a finite sequence of elementary building blocks as in Subsection 4.2, with the underlying pseudo-orbit

$$\tilde{\gamma}_0^\varepsilon, \tilde{\lambda}_0^\varepsilon, \ldots, \tilde{\gamma}_{n_1}^\varepsilon, \tilde{\lambda}_{n_1}^\varepsilon$$

for some $n_1 = O(1/\varepsilon^3)$, along which the scaled energy grows from $E = 1$ to $E = 2$ in a scaled time $O(1/\varepsilon^3)$.

This corresponds to a growth of the physical energy by $O(1/\varepsilon^2)$ in a physical time $O(1/\varepsilon^2)$. This physical energy growth is linear in the physical time. If the initial energy level of the physical energy is E_*, the physical energy at the end of this sequence of elementary building blocks is $2E_*$.

The windowing construction above provides a finite sequence of correctly aligned windows whose Λ_ε-components are close to these level sets. Any shadowing orbit obtained from Theorem 5.3 will result in an $O(1)$ growth of the scaled energy in a scaled time $O(1/\varepsilon^3)$.

Note that, as remarked before, segments of correctly aligned windows can be concatenated. Since the constructions are uniform for all $\varepsilon < \varepsilon_0$, we can construct infinitely many segments that make the energy grow to ∞, and concatenate them.

More precisely, to obtain orbits whose scaled energy grows to infinity we proceed as in Subsection 4.5. We reset ε to $\varepsilon = 1/\sqrt{2E_*}$ and we repeat the construction of a sequence of elementary building blocks whose scaled energy grows from $E = 1$ to $E = 2$ in a scaled time $O(1/\varepsilon^3)$. The energy growth rate is still linear, and at least as large as in the previous step. The physical energy grows from $2E_*$ to $4E_*$. Then we concatenate the segment of correctly aligned windows that grows the physical energy from E_* to $2E_*$, with the segment that grows the physical energy from $2E_*$ to $4E_*$. The concatenation of these segments of correctly aligned windows constructed is also a segment of correctly aligned windows. This construction of segments of elementary building blocks and corresponding segments of correctly aligned windows can be repeated indefinitely. The shadowing orbit that visits the resulting infinite sequence of correctly aligned windows yields an infinite physical energy growth at a linear rate with respect to the physical time. In this way we obtain the statement on the existence of orbits with unbounded energy growth from Theorems 2.1 and 2.2.

To obtain an orbit whose scaled energy follows a prescribed energy path, we note that a path $\mathcal{E} : [0, \infty) \to \mathbb{R}$ determines, via time-discretization, a sequence of J_ε-action level sets $\{\sqrt{2\mathcal{E}(k)}\}_{k\in\mathbb{N}}$ in Λ_ε. As in Subsection 4.6, we construct a sequence of elementary building blocks whose corresponding J_ε-values visit, in the prescribed order, the values $\sqrt{2\mathcal{E}(k)}$, $k \in \mathbb{Z}$. Since, in general, one cannot move from one value $\sqrt{2\mathcal{E}(k)}$ to the next $\sqrt{2\mathcal{E}(k+1)}$ via a single elementary building block, the construction yields a sequence $\{J_\varepsilon^i\}_{i\in\mathbb{Z}}$ of action level sets such that successive elementary building blocks correspond to successive values of J_ε^i, and there exists a subsequence $\{i_k\}_{k\in\mathbb{Z}}$ of $\mathbb{N}$ such that $J_\varepsilon^{i_k} = \sqrt{2\mathcal{E}(k)}$. Thus, we are in the same situation as described above, and we can proceed in the same way.

5.3. Construction of windows

In this subsection we will work out the details of the construction of correctly aligned windows described earlier.

We consider a finite sequence of elementary building blocks as in Subsection 4.2, with the corresponding pseudo-orbit

$$\tilde{\gamma}_0^\varepsilon, \tilde{\lambda}_0^\varepsilon, \ldots, \tilde{\gamma}_{n_1}^\varepsilon, \tilde{\lambda}_{n_1}^\varepsilon$$

for some $n_1 = O(1/\varepsilon^3)$. The corresponding action J level sets corresponding to the curves λ_i^ε in Λ_ε, $i = 0, \ldots, n_1$, are

$$J_\varepsilon^0, J_\varepsilon^1, \ldots, J_\varepsilon^{n_1}.$$

5.3.1. A system of coordinates near a normally hyperbolic invariant manifold. The construction of windows requires a coordinate systems relative to which the windows can be defined.

We now describe a general construction of a system of coordinates near a normally hyperbolic invariant manifold. Given a normally hyperbolic invariant manifold $\Lambda \subseteq M$ for $F : M \to M$, the bundle $N_x = E_x^u \oplus E_x^s$, $x \in \Lambda$, gives a concrete realization of the normal bundle N_Λ. That is, N_x is the complementary subspace to $T_x\Lambda$, i.e., $T_xM = T_x\Lambda \oplus N_x$ for all $x \in \Lambda$.

For any point p in a sufficiently small neighborhood $\mathcal{N}(\Lambda)$ of Λ in M, we can find unique $x \in \Lambda$, $s \in E_x^s$, and $u \in E_x^u$, with u, s small, such that $p = \exp_x(s + u)$. Therefore, it is natural to use (x, s, u) as a system of coordinates in $\mathcal{N}(\Lambda)$. We denote $h(x, s, u) = \exp_x(s + u)$, where $\exp : TM \to M$ is the exponential map.

In general, this system of coordinates is as smooth as the bundles E^s, E^u, whose regularity is limited by the regularity of the map F and by the ratios of the exponential rates.

In our case, because the exponential rate of F_ε on the manifold $\tilde{\Lambda}_\varepsilon$ is close to 0 for ε small enough, the corresponding maps $h_\varepsilon, h_\varepsilon^{-1}$ are C^r. We also note that if we express F_ε in terms of the coordinate mapping, we have, for $F_\varepsilon \in C^2$,

$$h_\varepsilon^{-1} \circ F_\varepsilon \circ h_\varepsilon(x, s, u) = \big(F_\varepsilon(x), DF_\varepsilon(x)s, DF_\varepsilon(x)u\big) + O(\delta^2),$$

where $\delta > 0$ is the size of the neighborhood $\mathcal{N}(\Lambda_\varepsilon)$. Furthermore

$$Dh_\varepsilon^{-1} \circ F_\varepsilon \circ h_\varepsilon^{-1}(x, s, u) = DF_\varepsilon(x) + O(\delta).$$

If we express F_ε relative to this system of coordinates, $\tilde{F}_\varepsilon = h_\varepsilon^{-1} \circ F_\varepsilon \circ h_\varepsilon$ is close to the tangent map

$$TF_\varepsilon(x, s, u) = (F_\varepsilon(x), DF_\varepsilon(x)s, DF_\varepsilon(x)u).$$

Indeed, we have

$$\|\tilde{F}_\varepsilon - TF_\varepsilon\|_{C^0} \le C\delta^2, \qquad \|\tilde{F}_\varepsilon - TF_\varepsilon\|_{C^1} \le C\delta,$$

for some $C > 0$, where δ is the size of the neighborhood $\mathcal{N}(\Lambda_\varepsilon)$. The constants δ and C can be chosen independent of ε.

In particular, by considering a sufficiently small neighborhood $\mathcal{N}(\Lambda_\varepsilon)$ we can ensure that the map F_ε is contracting in the s-components, expanding in the u-components, and more or less neutral in x.

Since we can describe the manifold $\tilde{\Lambda}_\varepsilon$ via the coordinates (J, ϕ, θ), the above construction provides us with a C^1-smooth coordinate system (J, ϕ, θ, s, u) in a neighborhood $\mathcal{N}(\tilde{\Lambda}_\varepsilon)$ of $\tilde{\Lambda}_\varepsilon$. Relative to this coordinate system the map F_ε can be approximated by a skew product of a map acting in the center directions of $\tilde{\Lambda}_\varepsilon$ and a map acting in the hyperbolic directions. To simplify notation, we will not make explicit the dependence on ε of the coordinate systems. Also, from this point on we will denote F_ε by F.

We will use this system of coordinates to construct windows near Λ_ε, and in particular about the sets $\mathcal{L}_i^\varepsilon$.

We will also need to construct windows about the heteroclinic intersections $(\{z_{i-1,i}^\varepsilon\} \times N) \cap \tilde{\Gamma}_\varepsilon$ of $W^u(\mathcal{L}_{i-1}^\varepsilon)$ with $W^s(\mathcal{L}_i^\varepsilon)$. For this, we propagate the above coordinate system from a neighborhood of the set $\mathcal{L}_{i-1}^\varepsilon$ in $\tilde{\Lambda}_\varepsilon$ along the unstable manifold $W^u(\mathcal{L}_{i-1}^\varepsilon)$, giving rise to a C^1-smooth coordinate system $(J^-, \phi^-, \theta, s^-, u^-)$ in a neighborhood of $(\{z_{i-1,i}^\varepsilon\} \times N) \cap \tilde{\Gamma}_\varepsilon$. Also, we propagate the same coordinate system from a neighborhood of the leaf $\mathcal{L}_i^\varepsilon$ along the stable manifold $W^s(\mathcal{L}_i^\varepsilon)$ to a neighborhood of the heteroclinic intersection $(\{z_{i-1,i}^\varepsilon\} \times N) \cap \tilde{\Gamma}_\varepsilon$, producing a C^1-smooth coordinate system $(J^+, \phi^+, \theta, s^+, u^+)$ in the neighborhood of $(\{z_{i-1,i}^\varepsilon\} \times N) \cap \tilde{\Gamma}_\varepsilon$.

The two coordinate systems differ by order $O(1)$, that is, if Φ denotes the coordinate change from one system to the other then $C_3^{-1} \leq \|D\Phi\| \leq C_3$ uniformly in a compact neighborhood of $\tilde{\Gamma}_\varepsilon$, for some $C_3 > 1$. Since we obtain two coordinate systems around $(\{z_{i-1,i}\} \times N) \cap \tilde{\Gamma}_\varepsilon$, we can construct windows and verify their correct alignment in either coordinate system. The sizes of the components of a window in one coordinate system will differ from the sizes in the other coordinate system by some multiplicative constants that are independent of ε and, by compactness, they can be chosen to be the same for all heteroclinic intersections.

5.3.2. Choice of constants. We choose some constants that will be used throughout the proof. Define

$$a^\pm = \sup_{z \in H^{-1}[1,2] \cap \tilde{\Gamma}_\varepsilon} d(z^\pm, z),$$

where the distance is measured along the stable or unstable fiber through x, respectively. Since $H^{-1}[1, 2] \cap \tilde{\Gamma}_\varepsilon$ and $H^{-1}[1, 2] \cap \tilde{\Lambda}_\varepsilon$ are compact, we have $0 < a^\pm < \infty$.

First we choose some constant $\varepsilon_1 > 0$, independent of ε, and then we choose positive constants $\alpha^-, \alpha^+, \check{\alpha}, \hat{\alpha}$ and $\beta^-, \beta^+, \check{\beta}, \hat{\beta}$, independent of ε, such that

$$3\varepsilon_1 < \alpha^- = \hat{\alpha} = \check{\alpha} < C_3^{-1}\alpha^+, \tag{5.1}$$

$$3\varepsilon_1 < \beta^+ = \hat{\beta} = \check{\beta} < C_3^{-1}\beta^-. \tag{5.2}$$

Second, we choose positive integers N, M sufficiently large so that

$$\lambda_+^N(a^+ + \alpha_+) < 2\varepsilon_1, \quad \check{\beta} + \varepsilon_1 < \mu_-^N\beta^+, \quad \mu_-^{-M}(\beta^- + a^-) < 2\varepsilon_1, \tag{5.3}$$

$$F^N(\tilde{\Gamma}_\varepsilon) \subseteq \mathcal{N}(\tilde{\Lambda}_\varepsilon), \quad F^{-M}(\tilde{\Gamma}_\varepsilon) \subseteq \mathcal{N}(\tilde{\Lambda}_\varepsilon), \tag{5.4}$$

where $\mathcal{N}(\tilde{\Lambda}_\varepsilon)$ is the neighborhood of $\tilde{\Lambda}_\varepsilon$ where the coordinate system described in Subsection 5.3.1 is defined. Note that there exist finite N, M as in (5.4) due to the definition of $\tilde{\Gamma}_\varepsilon$. Since the dynamical system (3.3) tends to a product system when $\varepsilon \to 0$, we can choose N, M to be independent of ε for all ε sufficiently small.

Third, we choose a positive integer K satisfying

$$2\check{\gamma} + \hat{\gamma} + 3\varepsilon_1 < K\tau\check{\delta}, \tag{5.5}$$

where τ is the twist constant from B6. Such a finite K exists for all small enough ε, since the inner map F is a twist map in the (J, ϕ)-variables.

Fourth, we choose positive constants δ^-, δ^+, $\check{\delta}$, $\hat{\delta}$, γ^-, γ^+, $\check{\gamma}$, $\hat{\gamma}$ such that

$$\check{\delta} < \check{\delta} + \varepsilon_1 < \delta^+ < C_3^{-1}\gamma^- < \gamma^- < \gamma^- + M\tau\delta^- + \varepsilon_1 < \hat{\gamma}, \tag{5.6}$$

$$\check{\delta} < \check{\delta} + \varepsilon_1 < \hat{\delta} < \hat{\delta} + \varepsilon_1 < \delta^- < C_3\delta^- < \gamma^+ < \gamma^+ + N\tau\delta^+ + \varepsilon_1 < \check{\gamma}. \tag{5.7}$$

Additionally, we want these constants to be sufficiently small. We will explain later in the argument how small they should be, but we emphasize here that the smallness condition is independent of ε and can be made precise from the beginning of the argument.

Fifth, we make ε even smaller if necessary, as described below. Note that the estimates on the inner map $F_{|\tilde{\Lambda}_\varepsilon}$ and on the outer map S_ε involve some error terms of order $O(\varepsilon^2)$, as in B6 and B7. Due to the compactness of $H^{-1}[1, 2] \cap \tilde{\Lambda}_\varepsilon$ and $H^{-1}[1, 2] \cap \tilde{\Gamma}_\varepsilon$, these error terms can be bounded from above by $C_4\varepsilon^2$ for some constant $C_4 > 0$ independent of ε. Now we choose ε sufficiently small so that

$$NC_4\varepsilon^2 < \varepsilon_1, \qquad MC_4\varepsilon^2 < \varepsilon_1, \qquad KC_4\varepsilon^2 < \varepsilon_1. \tag{5.8}$$

Thus, when we estimate the error terms when iterating the inner map or the outer map up to $\max\{M, N, K\}$ times, we will be able to conclude that the error terms are always less than ε_1.

From now on, ε is sufficiently small and fixed.

5.3.3. Step 1. Let us consider the heteroclinic intersection $W^u(\mathcal{L}^\varepsilon_{i-1})$ with $W^s(\mathcal{L}^\varepsilon_i)$ at $\{z^\varepsilon_{i-1,i}\} \times N$. As seen before, the normal hyperbolicity implies that there exist $z^{\varepsilon,-}_{i-1,i} \in \mathcal{L}^\varepsilon_{i-1}$ and $z^{\varepsilon,+}_{i-1,i} \in \mathcal{L}^\varepsilon_i$ such that $z^\varepsilon_{i-1,i} \in W^u(z^{\varepsilon,-}_{i-1,i}) \cap W^s(z^{\varepsilon,+}_{i-1,i})$. Let a^-_{i-1} be the distance between $z^\varepsilon_{i-1,i}$ and $z^{\varepsilon,-}_{i-1,i}$ measured along $W^u(\mathcal{L}^\varepsilon_{i-1})$, and let a^+_i be the distance between $z^\varepsilon_{i-1,i}$ and $z^{\varepsilon,+}_{i-1,i}$ measured along $W^s(\mathcal{L}^\varepsilon_i)$. We have $a^+_i < a^+$ and $a^-_i < a^-$.

At this step we construct a pair of windows W^-_{i-1}, W^+_i about $(\{z^\varepsilon_{i-1,i}\} \times N) \cap \tilde{\Gamma}_\varepsilon$ such that W^-_{i-1} is correctly aligned with W^+_i under the identity mapping.

We define the window W^-_{i-1} in the coordinates $(s^-, u^-, \phi^-, J^-, \theta)$ of the type

$$W^-_{i-1} = (R^{s^-}_{i-1} \times R^{u^-}_{i-1}) \times (R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}) \times R^\theta_{i-1},$$

where R^c denotes a rectangle in the coordinate c. We will think of W^-_{i-1} as a product of three window components: $R^{s^-}_{i-1} \times R^{u^-}_{i-1}$, corresponding to the hyperbolic coordinates,

$R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}$, corresponding to the angle-action coordinates, and R^{θ}_{i-1}, corresponding to the external system on N.

We define the exit set $(W^-)^{\text{exit}}_{i-1}$ of W^-_{i-1} by

$$\begin{aligned}(W^-)^{\text{exit}}_{i-1} = &(R^{s^-}_{i-1} \times \partial R^{u^-}_{i-1}) \times (R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}) \times R^{\theta}_{i-1} \\ &\cup (R^{s^-}_{i-1} \times R^{u^-}_{i-1}) \times (\partial R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}) \times R^{\theta}_{i-1} \\ &\cup (R^{s^-}_{i-1} \times R^{u^-}_{i-1}) \times (R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}) \times \partial R^{\theta}_{i-1}.\end{aligned}$$

This means that the exit directions of W^-_{i-1} correspond to the unstable direction u^- of the hyperbolic window $R^{s^-}_{i-1} \times R^{u^-}_{i-1}$, to the angle direction ϕ^- of the action-angle window $R^{\phi^-}_{i-1} \times R^{J^-}_{i-1}$, and to all directions of the external-state window R^{θ}_{i-1}. This definition of the exit set follows the construction of product of windows described in [45].

Similarly, we define the window W^+_i in the coordinates $(s^+, u^+, \phi^+, J^+, \theta)$ to be given by the product

$$W^+_i = (R^{s^+}_i \times R^{u^+}_i) \times (R^{\phi^+}_i \times R^{J^+}_i) \times R^{\theta}_i,$$

and its exit set $(W)^{\text{exit}}_i$ by

$$\begin{aligned}(W^+)^{\text{exit}}_i = &(R^{s^+}_i \times \partial R^{u^+}_i) \times (R^{\phi^+}_i \times R^{J^+}_i) \times R^{\theta}_i \\ &\cup (R^{s^+}_i \times R^{u^+}_i) \times (R^{\phi^+}_i \times \partial R^{J^+}_i) \times R^{\theta}_i \\ &\cup (R^{s^+}_i \times R^{u^+}_i) \times (R^{\phi^+}_i \times R^{J^+}_i) \times \partial R^{\theta}_i.\end{aligned}$$

The exit directions of W^+_i correspond to the direction u^+ of $R^{s^+}_i \times R^{u^+}_i$, to the direction J^+ of $R^{\phi^+}_i \times R^{J^+}_i$, and to all directions of R^{θ}_i. Note that the only difference in the exit directions of W^+_i from W^-_{i-1} is the switching from the direction ϕ^- to the direction J^+ in the angle-action components. Since we are not using any dynamics yet in constructing these windows, this switching in the exit direction of the action-angle components may seem arbitrary. The reason for this switching will become apparent later when we align windows by the inner map: when an action-angle component is iterated under the inner dynamics until it crosses another action-angle component, the twist property of the inner map causes the first window to have its action direction stretched across the second window along its angle direction. Therefore we will have a switching of the exit directions due to the alignment under the inner map. By having another switching of the exit direction at the heteroclinic intersection, we ensure the consistency of the correct alignment in the two-step process. This will be explained in greater detail at Step 3.

We set the sizes of the rectangular components of the windows as follows:

$$\|R^{s^-}_{i-1}\| = \alpha^-, \quad \|R^{u^-}_{i-1}\| = \beta^-, \quad \|R^{\phi^-}_{i-1}\| = \gamma^-, \quad \|R^{J^-}_{i-1}\| = \delta^-. \tag{5.9}$$

$$\|R^{s^+}_{i}\| = \alpha^+, \quad \|R^{u^+}_{i}\| = \beta^+, \quad \|R^{\phi^+}_{i}\| = \gamma^+, \quad \|R^{J^+}_{i}\| = \delta^+. \tag{5.10}$$

Above, by $\|R^c\|$ we mean the diameter of the rectangle R^c of the projection onto the coordinate c.

We want to ensure that W_{i-1}^- is correctly aligned with W_i^+ under the identity mapping. The two windows are defined in two different coordinate systems, so we need to use the coordinate change Φ to compare them relative to the same coordinate system. We therefore require that $\Phi(R_{i-1}^{s^-}) \subseteq \mathrm{int}(R_i^{s^+})$, $\mathrm{int}(\Phi(R_{i-1}^{u^-})) \supseteq R_i^{u^+}$, $\mathrm{int}(\Phi(R_{i-1}^{\phi^-})) \supseteq R_i^{J^+}$, $\Phi(R_{i-1}^{J^-}) \subseteq \mathrm{int}(R_i^{\phi^+})$, and $R_i^\theta \subseteq \mathrm{int}(\Phi(R_{i-1}^\theta))$. These conditions imply that W_{i-1}^- is correctly aligned with W_i^+ under the identity mapping. These conditions are consistent with the size specifications in (5.9) and (5.10), by the choice of the constants in (5.1)–(5.4).

5.3.4. Step 2. At this step we take a forward iterate of the window W_i^+ along the stable manifold $W^s(\mathcal{L}_i^\varepsilon)$, and we align its image with a window $\check{W}_i$ near $\tilde{\Lambda}_\varepsilon$. Similarly, we take a backward iterate of the window W_i^- along the unstable manifold $W^u(\mathcal{L}_{i-1}^\varepsilon)$, and we align its image with a window $\hat{W}_i$ near $\tilde{\Lambda}_\varepsilon$.

First, we consider the positive iterate $F^N(W_i^+)$ of W_i^+, with N as in Subsection 5.3.2. Since the window W_i^+ has been defined in the coordinate system $(s^+, u^+, \phi^+, J^+, \theta)$ near $\tilde{\Gamma}_\varepsilon$, which was obtained by propagating the system near $\tilde{\Lambda}_\varepsilon$ along the stable manifold, and N as in (5.4), the image $F^N(W_i^+)$ is naturally defined as a multi-dimensional rectangle in the coordinate system (s, u, ϕ, J, θ) defined in $\mathcal{N}(\mathcal{L}_i^\varepsilon)$. The dynamics in these coordinates is the skew product of the dynamics in the center directions with the dynamics in the hyperbolic directions.

The twist condition satisfied by F in the (ϕ, J) variables, described in B6, determines a sheering of the ϕ-direction by a quantity of τJ per iterate along each level set of the action variable J, modulo an error up to ε_1; also, the J-coordinate is preserved up to ε_1. Since $F^N(W_i^+)$ is in $\mathcal{N}(\tilde{\Lambda}_\varepsilon)$, there exists a rectangle $\check{R}_i^\phi \times \check{R}_i^{J^+}$ in the (ϕ, J)-coordinate such that the (ϕ, J)-component of W_i^+ is correctly aligned under F^N with $\check{R}_i^\phi \times \check{R}_i^{J^+}$. Since the size of the (ϕ, J) component of W_i^+ is $\gamma^- \times \delta^-$, we can choose $\check{R}_i^\phi \times \check{R}_i^{J^+}$ to be of size $\|\check{R}_i^\phi\| = \check{\gamma} > \gamma^+ + N\tau\delta^+ + \varepsilon_1$, and $\|\check{R}_i^{J^+}\| = \check{\delta} < \delta^+ - \varepsilon_1$. These inequalities are justified by (5.6) and (5.7).

In the θ-direction the map F acts as the time-ε map χ of the flow χ. To achieve correct alignment of the windows in the θ-variable, it is sufficient to choose $\check{R}_i^\theta$ to be a topological rectangle in the interior of $\chi^N(R_i^{\theta+})$.

The distance between the point $F^N(z_{i-1,i}^\varepsilon) \in F^N(W_i^+)$ and $F^N(z_{i-1,i}^{\varepsilon,+})$, measured along the stable manifold $W^s(\mathcal{L}_i^\varepsilon)$, is less than $\lambda_+^N a^+$.

The size of the projection of $F^N(W_i^+)$ onto the s-coordinate is at most $\lambda_+^N\alpha^+ + \varepsilon_1$. The size of the u-component of $F^N(W_i^+)$ is at least $\mu_-^N\beta^+ - \varepsilon_1$. By the choice of the coordinates in Subsection 5.3.1, the hyperbolic directions of $F^N(R_i^{s^+} \times R_i^{u^+})$ coincide with the hyperbolic directions in $\mathcal{N}(\tilde{\Lambda}_\varepsilon)$. Since by (5.1) and (5.3) we have $\check{\alpha} > \lambda^N(\alpha^+ + a^+) + \varepsilon_1$, and by (5.2) and (5.4) we have $\check{\beta} + \varepsilon_1 < \mu_-^N\beta^+$, we can construct a rectangle $\check{R}_i^s \times \check{R}_i^u$

in the hyperbolic variables, of size $\check{\alpha} \times \check{\beta}$, such that $\check{R}_i^{s^+} \times \check{R}_i^{u^+}$ is correctly aligned with $\check{R}_i^s \times \check{R}_i^u$ under F^N relative to the the hyperbolic variables.

Through these choices, we construct a new window $\check{W}_i$ about $z_{i-1,i}^{\varepsilon,+}$, given by

$$\check{W}_i = (\check{R}_i^s \times \check{R}_i^u) \times (\check{R}_i^\phi \times \check{R}_i^J) \times \check{R}_i^\theta$$

in the coordinates (s, u, ϕ, J, θ), such that $F^N(W_i^+)$ is correctly aligned with $\check{W}_i$ under the identity mapping, or equivalently W_i^+ is correctly aligned with $\check{W}_i$ under F^N. Its exit set $(\check{W})_i^{\text{exit}}$ is given by

$$\begin{aligned}(\check{W})_i^{\text{exit}} = &(\check{R}_i^s \times \partial \check{R}_i^u) \times (\check{R}_i^\phi \times \check{R}_i^J) \times \check{R}_i^\theta \\ &\cup (\check{R}_i^s \times \check{R}_i^u) \times (\check{R}_i^\phi \times \partial \check{R}_i^J) \times \check{R}_i^\theta \\ &\cup (\check{R}_i^s \times \check{R}_i^u) \times (\check{R}_i^\phi \times \check{R}_i^J) \times \partial \check{R}_i^\theta .\end{aligned}$$

We note that the correct alignment of W_i^+ with $\check{W}_i$ under F^N follows from the correct alignment of the window components, according to Definition 5.2, and the product property from [45]. The time it takes to achieve this alignment is N, which is independent of ε, and so is of order $O(1)$.

In a similar fashion, we construct a new window $\hat{W}_{i-1}$ near $\mathcal{L}_{J_{i-1}}^\varepsilon$ such that $F^M(\hat{W}_{i-1})$ is correctly aligned with W_{i-1}^- under the identity mapping, or equivalently $\hat{W}_{i-1}$ is correctly aligned with W_{i-1}^- under F^M for some M as in Subsection 5.3.2. For this purpose, we take a negative iterate $F^{-M}(W_{i-1}^-)$ of W_{i-1}^- and we construct a new window $\hat{W}_{i-1}$ about $z_{i-1,i}^{\varepsilon,-}$, of the type

$$\hat{W}_{i-1} = \hat{R}_{i-1}^s \times \hat{R}_{i-1}^u \times \hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J \times \hat{R}_{i-1}^\theta$$

in the coordinates (s, u, ϕ, J, θ), with the exit set given by

$$\begin{aligned}(\hat{W})_{i-1}^{\text{exit}} = &\hat{R}_{i-1}^s \times \partial \hat{R}_{i-1}^u \times \hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J \times \hat{R}_{i-1}^\theta \\ &\cup \hat{R}_{i-1}^s \times \hat{R}_{i-1}^u \times \partial \hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J \times \hat{R}_{i-1}^\theta \\ &\cup \hat{R}_{i-1}^s \times \hat{R}_{i-1}^u \times \hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J \times \partial \hat{R}_{i-1}^\theta .\end{aligned}$$

We choose the size of the window components to be $\|\hat{R}_{i-1}^s\| = \hat{\alpha}$, $\|\hat{R}_{i-1}^u\| = \hat{\beta}$, $\|\hat{R}_{i-1}^\phi\| = \hat{\gamma}$, and $\|\hat{R}_{i-1}^J\| = \hat{\delta}$. By the choice of the coordinates in Subsection 5.3.1, we can choose $\hat{R}_{i-1}^s \times \hat{R}_{i-1}^u$ so that it is correctly aligned with $R_{i-1}^{s-} \times R_{i-1}^{u-}$ under F^M. By condition B6 we can choose $\hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J$ so that it is correctly aligned with $R_{i-1}^{\phi-} \times R_{i-1}^{J-}$ under F^M. In the θ-variable we choose $\hat{R}_{i-1}^\theta$ so that $R_{i-1}^{\theta-} \subseteq \text{int}(\chi^M(\hat{R}_{i-1}^\theta))$.

We also require that $(\hat{R}_{i-1}^\phi \times \hat{R}_{i-1}^J) \times \hat{R}_{i-1}^\theta$ is contained, via the parametrization k described in Subsection 3.2, in $U^- \times N \subseteq k^{-1}(\tilde{U}^-)$.

Now we verify that the choice of the sizes of the window components is compatible with these correct alignment relations. Using (5.1), (5.2), (5.4), we can ensure that $\hat{\beta} >$

$\mu_-^{-M}(\beta^- + a^-) + \varepsilon_1$ and $\hat{\alpha} + \varepsilon_1 < \lambda_+^{-M}\alpha^-$. By (5.6) and (5.7) we can ensure that $\hat{\delta} + \varepsilon_1 < \delta^-$ and $\hat{\gamma} > \gamma^- + M\tau\delta^- + \varepsilon_1$, where the term $M\tau\delta^-$ represents the effect of the twist map on $R_{i-1}^{\phi-} \times R_{i-1}^{J-}$ under M negative iterates.

By the product property of correctly aligned windows, we find that $F^M(\hat{W}_{i-1})$ is correctly aligned with W_{i-1}^- under the identity mapping, or equivalently $\hat{W}_{i-1}$ is correctly aligned with W_i^- under F^M.

We note that to achieve the correct alignment of the windows in Step 2 we do not use the Lambda Lemma as in [39, 45]. The role of the Lambda Lemma in this paper is taken over by the choice of coordinates near $\tilde{\Lambda}_\varepsilon$ from Subsection 5.3.1, extended along the stable and unstable manifolds to produce convenient coordinate systems about $\tilde{\Gamma}_\varepsilon$.

5.4. Step 3

We consider the heteroclinic intersection $W^u(\mathcal{L}_{i-1}^\varepsilon) \cap W^s(\mathcal{L}_i^\varepsilon)$ and the heteroclinic intersection $W^u(\mathcal{L}_i^\varepsilon) \cap W^s(\mathcal{L}_{i+1}^\varepsilon)$.

By applying Step 2 for each heteroclinic connection we obtain a pair of windows $\hat{W}_i$ and $\check{W}_i$, both located about the set $\mathcal{L}_i^\varepsilon$.

At this step we want to get $\check{W}_i$ correctly aligned with $\hat{W}_i$ under some iterate F^K. We pointed out earlier that the dynamics in the coordinates (ϕ, J, θ, u, s) is the skew product of the dynamics in the center directions and the dynamics in the hyperbolic directions.

The first task is to ensure the correct alignment of the rectangle $\check{R}_i^u \times \check{R}_i^s$, of exit set $\partial\check{R}_i^u \times \check{R}_i^s$, with the rectangle $\hat{R}_i^u \times \hat{R}_i^s$, of exit set $\partial\hat{R}_i^u \times \hat{R}_i^s$. These rectangles are defined in the same coordinate system on $\mathcal{N}(\tilde{\Lambda}_\varepsilon)$. The correct alignment reduces to ensuring the inclusions $\mathrm{int}(F^K(\check{R}_i^u)) \supseteq \hat{R}_i^u$ and $F^{K_i}(\check{R}_i^s) \subseteq \mathrm{int}(\hat{R}_i^s)$. Since the unstable directions get uniformly expanded and the stable directions get uniformly contracted, it is sufficient for the correct alignment to have $\check{R}_i^u \times \check{R}_i^s$ and $\hat{R}_i^u \times \hat{R}_i^s$ of the same size $\check{\alpha} \times \check{\beta} = \hat{\alpha} \times \hat{\beta}$ (see (5.1) and (5.2)).

The second task is to correctly align $(\check{R}_i^\phi \times \check{R}_i^J) \times \check{R}_i^\theta$ with $(\hat{R}_i^\phi \times \hat{R}_i^J) \times \hat{R}_i^\theta$. The exit set of $\check{R}_i^\phi \times \check{R}_i^J$ is given by $\check{R}_i^\phi \times \partial\check{R}_i^J$, while the exit set of $\hat{R}_i^\phi \times \hat{R}_i^J$ is given by $\partial\hat{R}_i^\phi \times \hat{R}_i^J$. Under F^K, the rectangle $\check{R}_i^\phi \times \check{R}_i^J$ is transformed into a topological rectangle whose exit set components—'top' edge and 'bottom' edge—are shifted apart from one another by at least $K\tau\check{\delta} - \varepsilon_1$ in the ϕ-direction, and they are separated by a distance of at least $\check{\delta} - \varepsilon_1$ in the J-direction. In order for $F^K(\check{R}_i^\phi \times \check{R}_i^J)$ to be correctly aligned with $\hat{R}_i^\phi \times \hat{R}_i^J$ under the identity, the image $F^K(\check{R}_i^\phi \times \check{R}_i^J)$ should stretch across $\hat{R}_i^\phi \times \hat{R}_i^J$ in the direction of ϕ and its exit set components should come out through the exit set components—left edge and right edge—of $\hat{R}_i^\phi \times \hat{R}_i^J$. To ensure the stretching all the way of the first window across the second, we need $(K\tau\|\check{R}^J\| - \varepsilon_1) - 2(\|\check{R}_i^\phi\| + \varepsilon_1)$, representing the shearing in the ϕ direction of $F^K(\check{R}_i^\phi \times \check{R}_i^J)$ minus the size of the top and the bottom edge, to be bigger than $\|\hat{R}_i^\phi\|$, representing the 'width' of $\hat{R}_i^\phi \times \hat{R}_i^J$ in the ϕ-direction. Also, to ensure that the image of the first rectangle does not meet the entry set of the second rectangle, we need $\|\check{R}^J\| + \varepsilon_1$, representing an upper bound of the

'height' of $F^K(\check{R}_i^\phi \times \check{R}_i^J)$ in the J-direction, to be smaller than $\|\hat{R}^J\|$, representing the height of $\hat{R}_i^\phi \times \hat{R}_i^J$ in the J-direction. By construction $\|\check{R}_i^\phi\| = \check{\gamma}$, $\|\check{R}^J\| = \check{\delta}$ $\|\hat{R}_i^\phi\| = \hat{\gamma}$, and $\|\hat{R}^J\| = \hat{\delta}$. By (5.5) we have $K\check{\delta} > 2\check{\gamma} + \hat{\gamma} + 3\varepsilon_1$, and by (5.7) we have $\hat{\delta} > \check{\delta} + \varepsilon_1$.

We also need to ensure the correct alignment of the d-dimensional rectangle $\check{R}_i^\theta$, of exit set $\partial\check{R}_i^\theta$, with the d-dimensional rectangle $\hat{R}_i^\theta$, of exit set $\hat{R}_i^\theta$. The correct alignment condition for windows that have exit sets consisting of their whole boundary means that the image of the first window contains the second window in its interior. Thus, to make $\check{R}_i^\theta$ correctly aligned with $\hat{R}_i^\theta$ under F^K, we need to choose $\hat{R}_i^\theta$ so that $\hat{R}_i^\theta \subset \mathrm{int}(\chi_\varepsilon^K(\check{R}_i^\theta))$.

By the product property of correctly aligned windows, the outcome of this step is that $\check{W}_i$ is correctly aligned with $\hat{W}_i$ under F^K.

5.4.1. Step 4. To construct a long sequence of correctly aligned windows that visits all the sets $\{\mathcal{L}_i^\varepsilon\}_{i=0}^{n_1}$, we start by constructing a pair of correctly aligned windows W_0^- and W_1^+, by the heteroclinic intersection $(\{z_{0,1}\} \times N) \cap \tilde{\Gamma}_\varepsilon$ of $W^u(\mathcal{L}_0^\varepsilon)$ and $W^s(\mathcal{L}_1^\varepsilon)$, as in Step 1. Then, as in Step 2, we construct the window $\hat{W}_0$ near $\mathcal{L}_{J_0}^\varepsilon$ and $\check{W}_1$ near $\mathcal{L}_1^\varepsilon$. At this initial step, for the uniformity of the notation, we set $\check{W}_0 = \hat{W}_0$. Then, we continue the construction recursively, following Steps 1–3, until we arrive at a window $\check{W}_{n_1}$ near $\mathcal{L}_{n_1}$. Hence, we obtain the sequence of correctly aligned windows

$$\hat{W}_0, W_0, W_1, \check{W}_1, \ldots, \hat{W}_{n_1-1}, W_{n_1-1}, W_{n_1}, \check{W}_{n_1}$$

starting from $\mathcal{L}_0^\varepsilon$ and ending with $\mathcal{L}_{n_1}^\varepsilon$. We apply the Shadowing Lemma type Theorem 5.3, and conclude that there is an orbit $\{y_i\}_{i=0}^{n_1}$ with $y_i \in \check{W}_i$ for all i, and $y_{i+1} = F^{K+M+N}(y_i)$ for $i = 0, \ldots, n_1 - 1$.

5.5. *Proofs of Theorems 2.1 and 2.2*

For Theorem 2.1 the initial condition θ_0 is chosen to be a non-trivial uniformly recurrent point. For Theorem 2.2 the initial condition θ_0 can be any point in N. Then the construction from the previous section is performed. The outcome is an orbit $\{y_i\}_{i=0}^{n_1}$ that is $O(1)$ close to the pseudo-orbit underlying the sequence of elementary building blocks constructed in Subsection 4.4. We can estimate the time it takes for an orbit $\{y_i\}$ to move from a window $\check{W}_0$ to the window $\check{W}_{n_1}$. The pseudo-orbit from Subsection 4.4 achieves a growth of scaled energy $\Delta H_\varepsilon = O(\varepsilon)$ in a scaled time $\Delta s = O(1/\varepsilon^2)$. Each pseudo-orbit corresponding to a single elementary building block gives rise to a sequence of windows that are correctly aligned. The time required by the correct alignment of windows, and the corresponding time it takes for the orbit to move along those windows, is $O(1)$, with the constants being the same for all windows in the sequence. Thus, the scaled time it takes for the orbit $\{y_i\}_{i=0}^{n_1}$ to travel from y_0 to y_{n_1} is $O(1/\varepsilon^2)$, equal to the time along the sequence of elementary building blocks multiplied by some constant independent of ε, and hence on the energy. The gain of physical energy along this orbit is $\Delta H = O(1/\varepsilon)$, and the physical time spent is $\Delta t = O(1/\varepsilon)$. Thus the change in the physical energy is proportional to the time along the orbit, i.e. $\Delta t \approx \Delta H \approx O(1/\varepsilon)$, as claimed.

Since concatenations of correctly aligned windows are still correctly aligned, the above construction of correctly aligned windows can be continued for a time $O(1)$ to achieve an $O(1)$ change of the physical energy that covers the interval $H = \varepsilon^{-2} H_\varepsilon \in [E_*, 2E_*]$, where E_* was the choice of initial energy. To grow the physical energy to infinity, we re-initialize the process starting with $\varepsilon = 1/\sqrt{2E_*}$, and we construct a sequence of correctly aligned windows as before. Again, the fact that concatenations of correctly aligned windows are still correctly aligned yields a longer sequence of correctly aligned windows. The construction can be repeated inductively, yielding an infinite sequence of correctly aligned windows, and a corresponding shadowing orbit whose energy grows to infinity in time. The estimate on the time does not get any worse since moving up along the infinite chains corresponds to higher energy levels where the speed of diffusion is only growing faster. Thus the energy growth is linear with respect to time.

Remark 5.4. In the above construction we should remark that the windows are of size $O(1)$, while the distance from $\mathcal{L}_i^\varepsilon$ to $\mathcal{L}_{i+1}^\varepsilon$ is only of order $O(\varepsilon^3)$. Thus, for an orbit $\{y_i\}_{i=0}^{n_1}$ that goes from $\breve{W}_i$ about $\mathcal{L}_i^\varepsilon$ to $\breve{W}_{i+1}$ about $\mathcal{L}_{i+1}^\varepsilon$ the net energy change is undetermined. However, the method of correctly aligned windows does detect an orbit whose physical energy changes by $O(1)$, from $E = E_*$ to $E = 2E_*$, within a scaled time $O(1)$. Hence, the topological argument is incapable of detecting the detailed changes of the energy along the orbit; it can only detect significant changes.

We could get more control on the energy levels visited by choosing the windows smaller, but then we would get worse estimates on the time. This phenomenon resembles the 'energy-time' uncertainty principle of Heisenberg [49].

5.6. *Proof of Theorem 2.3*

The function $\mathcal{E}$ represents a prescribed energy path with upper bounded derivative; this condition is necessary as the energy of the perturbed system cannot grow faster than linearly. The function $\mathcal{T}$ represents a parametrization of the time. The argument is the same as for Theorem 2.1, provided that we choose the infinite sequence $\{J_\varepsilon^i\}_{i\in\mathbb{N}}$ of level sets to follow the energy path $\mathcal{E}$, as described in Subsection 5.2.

5.7. *Regularity*

We note that if we assume that the metric and the potential are C^r, we conclude that the flow is C^{r-1} and so is the map F. The theory of normal hyperbolicity requires at least C^1-differentiability.

Since we are using a derivative with respect to parameters and even estimate the remainders, we would like that $r - 1 \geq 2$.

We note that the C^1-differentiability of the flow χ on the external manifold N is sufficient for the argument, since the estimates do not involve any derivatives along the solution curves of χ.

Hence, in Section 2 it suffices to take $r_0 = 3$.

Acknowledgments. Part of this work has been done while M.G. and R.L. participated in the program Stability and Instability in Mechanical Systems, at the Centre de Recerca Matemàtica, Barcelona, Spain. Another part of this work has been done while M.G. was a member of and R.L. was a visitor at the IAS. Both authors are very grateful to these two institutions. M.G. would also like to thank Keith Burns and Amie Wilkinson for useful discussions. Both authors would like to thank the anonymous referees and Prof. K. Burns for their useful comments and suggestions that helped to improve this paper.

Research of M.G. was partially supported by NSF grants DMS 0635607 and DMS 1515851, and by The Fund for Math of the IAS.

Research of R.L. was partially supported by NSF grants, DMS 1162544 and DMS 1500943.

References

[1] Albers, P., Fish, J., Frauenfelder, U., Hofer, H., Koert, O.: Global surfaces of section in the planar restricted 3-body problem. Arch. Ration. Mech. Anal. **204**, 273–284 (2012) Zbl 1287.70006 MR 2898741

[2] Albers, P., Frauenfelder, U., van Koert, O., Paternain, G.: The contact geometry of the restricted 3-body problem. Comm. Pure Appl. Math. **65**, 229–263 (2012) Zbl 1300.70006 MR 2855545

[3] Anosov, D. V.: Averaging in systems of ordinary differential equations with rapidly oscillating solutions. Izv. Akad. Nauk SSSR Ser. Mat. **24**, 721–742 (1960) (in Russian) Zbl 0102.08003 MR 0126592

[4] Arnold, V. I.: Instability of dynamical systems with several degrees of freedom. Soviet Math. Dokl. **5**, 581–585 (1964) Zbl 0135.42602 MR 0163026

[5] Arnold, V., Kozlov, V., Neishtadt, A.: Dynamical Systems III. Encyclopaedia Math. Sci. 3, Springer, Berlin (1988) Zbl 0623.00023 MR 0923953

[6] Bates, P. W., Lu, K., Zeng, C.: Invariant foliations near normally hyperbolic invariant manifolds for semiflows. Trans. Amer. Math. Soc. **352**, 4641–4676 (2000) Zbl 0964.37018 MR 1675237

[7] Bernard, P., Kaloshin, V., Zhang, K.: Arnold diffusion in arbitrary degrees of freedom and crumpled 3-dimensional normally hyperbolic invariant cylinders. arXiv:1112.2773 (2011)

[8] Berti, M., Biasco, L., Bolle, P.: Drift in phase space: a new variational mechanism with optimal diffusion time. J. Math. Pures Appl. (9) **82**, 613–664 (2003) Zbl 1025.37037 MR 1996776

[9] Bessi, U.: An approach to Arnol'd's diffusion through the calculus of variations. Nonlinear Anal. **26**, 1115–1135 (1996) Zbl 0867.70013 MR 1375654

[10] Bessi, U., Chierchia, L., Valdinoci, E.: Upper bounds on Arnold diffusion times via Mather theory. J. Math. Pures Appl. (9) **80**, 105–129 (2001) Zbl 0986.37052 MR 1810511

[11] Bolotin, S., Treschev, D.: Unbounded growth of energy in non-autonomous Hamiltonian systems. Nonlinearity **12**, 365–388 (1999) Zbl 0989.37050 MR 1677779

[12] Bounemoura, A.: An example of instability in high-dimensional Hamiltonian systems. Int. Math. Res. Notices **2012**, 685–716 Zbl 1242.37038 MR 2885986

[13] Brännström, N., Gelfreich, V.: Drift of slow variables in slow-fast Hamiltonian systems. Phys. D **237**, 2913–2921 (2008) Zbl 1153.37338 MR 2514071

[14] Brännström, N., de Simone, E., Gelfreich, V.: Geometric shadowing in slowfast Hamiltonian systems. Nonlinearity **23**, 1169–1184 (2010) Zbl 1207.37015 MR 2630096

[15] Capiński, M. J., Gidea, M., de la Llave, R.: Arnold diffusion in the planar elliptic restricted three-body problem: mechanism and numerical simulation. Nonlinearity **30**, 329–360 (2017)

[16] Cheng, C.-Q.: Arnold diffusion in nearly integrable Hamiltonian systems. arXiv:1207.4016 (2012)

[17] Cheng, C.-Q., Yan, J.: Existence of diffusion orbits in a priori unstable Hamiltonian systems. J. Differential Geom. **67**, 457–517 (2004) Zbl 05033776 MR 2153027

[18] Cheng, C.-Q., Yan, J.: Arnold diffusion in Hamiltonian systems: a priori unstable case. J. Differential Geom. **82**, 229–277 (2009) Zbl 1179.37081 MR 2520793

[19] Chierchia, L., Gallavotti, G.: Drift and diffusion in phase space. Ann. Inst. H. Poincaré Phys. Théor. **60**, no. 1, 144 pp. (1994) Zbl 1010.37039 MR 1259103

[20] Chirikov, B. V., Lieberman, M. A., Shepelyansky, D. L., Vivaldi, F. M.: A theory of modulational diffusion. Phys. D **14**, 289–304 (1985) Zbl 0584.58033 MR 0793707

[21] Conley, C., Easton, R.: Isolated invariant sets and isolating blocks. Trans. Amer. Math. Soc. **158**, 35–61 (1971) Zbl 0223.58011 MR 0279830

[22] Contreras, G.: Geodesic flows with positive topological entropy, twist maps and hyperbolicity. Ann. of Math. 172 (2010), 761–808. Zbl 1204.37032 MR 2680482

[23] Contreras, G., Paternain, G.: Genericity of geodesic flows with positive topological entropy on S^2. J. Differential Geom. **61**, 1–49 (2002) Zbl 1048.53061 MR 1949783

[24] Cresson, J.: Symbolic dynamics and Arnold diffusion. J. Differential Equations **187**, 269–292 (2003) Zbl 1018.37038 MR 1949441

[25] Cresson, J., Guillet, C.: Periodic orbits and Arnold diffusion. Discrete Contin. Dynam. Systems **9**, 451–470 (2003) Zbl 1031.37054 MR 1952386

[26] Delshams, A., Gidea, M., de la Llave, R., Seara, T. M.: Geometric approaches to the problem of instability in Hamiltonian systems. An informal presentation. In: Hamiltonian Dynamical Systems and Applications, W. Craig (ed.), Springer, Dordrecht (2008) Zbl 1144.37022 MR 2446259

[27] Delshams, A., Gidea, M., Roldán, P.: Transition map and shadowing lemma for normally hyperbolic invariant manifolds. Discrete Contin. Dynam. Systems **33**, 1089–1112 (2013) Zbl 1276.37043

[28] Delshams, A., Gidea, M., Roldán, P.: Arnold's mechanism of diffusion in the spatial circular restricted three-body problem: a semi-analytical argument. Phys. D **334**, 29–48 (2016) MR 3545968

[29] Delshams, A., de la Llave, R., Seara, T. M.: A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of $\mathbb{T}^2$. Comm. Math. Phys. **209**, 353–392 (2000) Zbl 0952.70015 MR 1737988

[30] Delshams, A., de la Llave, R., Seara, T. M.: Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows. Adv. Math. **202**, 64–188 (2006) Zbl 1091.37018 MR 2218821

[31] Delshams, A., de la Llave, R., Seara, T. M.: A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. Mem. Amer. Math. Soc. **179**, no. 844, viii+141 pp. (2006) Zbl 1090.37044 MR 2184276

[32] Delshams, A., de la Llave, R., Seara, T. M.: Geometric properties of the scattering map of a normally hyperbolic invariant manifold. Adv. Math. **217**, 1096-1153 (2008) Zbl 1131.37032 MR 2383896

[33] Delshams, A., de la Llave, R., Seara, T. M.: Instability of high dimensional Hamiltonian Systems: Multiple resonances do not impede diffusion. Adv. Math. **294**, 689–755 (2016) Zbl 06567884 MR 3479576

[34] Easton, R.: Orbit structure near trajectories biasymptotic to invariant tori. In: Classical Mechanics and Dynamical Systems, R. Devaney and Z. Nitecki (eds.), Dekker, 55–67 (1981) Zbl 0502.70024 MR 0640118

[35] Easton, R., McGehee, R.: Homoclinic phenomena for orbits doubly asymptotic to an invariant three-sphere, Indiana Univ. Math. J. **28**, 211–240 (1979) Zbl 0435.58010 MR 0523100

[36] Eberlein, P.: Horocycle flows on certain surfaces without conjugate points. Trans. Amer. Math. Soc. **233**, 1–36 (1977) Zbl 0369.58020 MR 0516501

[37] Féjoz, J., Guàrdia, M., Kaloshin, V., Roldán, P.: Kirkwood gaps and diffusion along mean motion resonances in the restricted planar three-body problem. J. Eur. Math. Soc. **18**, 2315–2403 (2016) Zbl 06655296 MR 3551192

[38] Fenichel, N.: Asymptotic stability with rate conditions. Indiana Univ. Math. J. **23**, 1109–1137 (1973/74) Zbl 0284.58008 MR 0339276

[39] Fontich, E., Martín, P.: Differentiable invariant manifolds for partially hyperbolic tori and a lambda lemma. Nonlinearity **13**, 1561–1593 (2000) Zbl 0985.37065 MR 1781809

[40] Fontich, E., Martín, P.: Arnold diffusion in perturbations of analytic integrable Hamiltonian systems. Discrete Contin. Dynam. Systems **7**, 61–84 (2001) Zbl 1014.37037 MR 1806373

[41] Furstenberg, H.: The unique ergodicity of the horocycle flow. In: A. Beck (ed.), Recent Advances in Topological Dynamics, Lecture Notes in Math. 318, Springer, 95–115 (1973) Zbl 0256.58009 MR 0393339

[42] Furstenberg, H.: Poincaré recurrence and number theory. Bull. Amer. Math. Soc. **5**, 211–234 (1981) Zbl 0481.28013 MR 0628658

[43] Gelfreich, V., Turaev, D.: Unbounded energy growth in Hamiltonian systems with a slowly growing parameter. Comm. Math. Phys. **283**, 769–794 (2008) Zbl 1152.37023 MR 2434747

[44] Gerber, M., Wilkinson, A.: Hölder regularity of horocycle foliations. J. Differential Geom. **52**, 41–72 (1999) Zbl 1144.37309 MR 1743465

[45] Gidea, M., de la Llave, R.: Topological methods in the instability problem of Hamiltonian systems. Discrete Contin. Dynam. Systems Ser. A **14**, 295–328 (2006) Zbl 1098.37056 MR 2163534

[46] Gidea, M., de la Llave, R., Seara, T.: A general mechanism of diffusion in Hamiltonian systems: qualitative results. arXiv:1405.0866 (2014)

[47] Gidea, M., Robinson, C.: Shadowing orbits for transition chains of invariant tori alternating with Birkhoff zones of instability. Nonlinearity **20**, 1115–1143 (2007) Zbl 1116.37041 MR 2312385

[48] Hedlund, G.: Fuchsian groups and transitive horocycles. Duke Math. J. **2**, 530–542 (1936) Zbl 0015.10201 MR 1545946

[49] Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Z. Phys. **43**, 172–198 (1927)

[50] Hirsch, M. W., Pugh, C. C., Shub, M.: Invariant Manifolds. Lecture Notes in Math. 583, Springer (1977) Zbl 0355.58009 MR 0501173

[51] Kaloshin, V.: Geometric proofs of Mather's connecting and accelerating theorems. In: Topics in Dynamics and Ergodic Theory, S. Bezuglyi and S. Kolyada (eds.), London Math. Soc. Lecture Note Ser. 310, Cambridge Univ. Press, 81–106 (2003) Zbl 1043.37045 MR 2052276

[52] Kaloshin, V., Zhang, K.: Normally hyperbolic invariant manifolds near strong double resonance. arXiv:1202.1032 (2012)

[53] Kaloshin, V., Zhang, K.: A strong form of Arnold diffusion for two and a half degrees of freedom. arXiv:1212.1150 (2012)

[54] Katok, A.: Entropy and closed geodesics. Ergodic Theory Dynam. Systems **2**, 339–365 (1982) Zbl 0525.58027 MR 0721728

[55] Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. Encyclopedia Math. Appl. 54, Cambridge Univ. Press, Cambridge (1995) Zbl 0878.58020 MR 1326374

[56] Lieberman, M. A., Tennyson, J., Jeffrey, L.: Chaotic motion along resonance layers in near-integrable Hamiltonian systems with three or more degrees of freedom. In: Long-Time Prediction in Dynamics (Lakeway, TX, 1981), C. W. Horton and L. Reichl (eds.), Wiley, New York, 179–211 (1983) MR 0714724

[57] Lions, P.-L., Souganidis, P. E.: Correctors for the homogenization of Hamilton–Jacobi equations in the stationary ergodic setting. Comm. Pure Appl. Math. **56**, 1501–1524 (2003) Zbl 1050.35012 MR 1988897

[58] de la Llave, R.: Orbits of unbounded energy in perturbations of geodesic flows by periodic potentials. A simple construction. Preprint (2004)

[59] Lochak, P., Marco, J.-P.: Diffusion times and stability exponents for nearly integrable analytic systems. Cent. Eur. J. Math. **3**, 342–397 (2005) Zbl 1107.37049 MR 2152466

[60] Lochak, P., Meunier, C.: Multiphase Averaging for Classical Systems. With Applications to Adiabatic Theorems. Appl. Math. Sci. 72, Springer, New York (1988) Zbl 0668.34044 MR 0959890

[61] Marco, J.-P.: Transition le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques. Ann. Inst. H. Poincaré Phys. Théor. **64**, 205–252 (1996) Zbl 0854.70011 MR 1386217

[62] Marco, J.-P., Sauzin, D.: Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems. Publ. Math. Inst. Hautes Études Sci. **96**, 199–275 (2003) Zbl 1086.37031 MR 1986314

[63] Marcus, B.: Unique ergodicity of the horocycle flow: variable negative curvature case. Israel J. Math. **21**, 133–144 (1975) Zbl 0314.58013 MR 0407902

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

[65] Mather, J.: Arnold diffusion. I: Announcement of results. J. Math. Sci. (N.Y.) **124**, 5275–5289 (2004) Zbl 1069.37044 MR 2129140

[66] Mather, J.: Arnold diffusion by variational methods. In: Essays in Mathematics and its Applications, 271–285 (2012) Zbl 1350.37067 MR 2975591

[67] Meyer, K., Hall, G. R., Offin, D.: Introduction to Hamiltonian Dynamical Systems and the N-Body Problem. 2nd ed., Springer (2009) Zbl 1179.70002 MR 2468466

[68] Mozes, S., Weiss, B.: Minimality and unique ergodicity for subgroup actions. Ann. Inst. Fourier (Grenoble) **48**, 1533–1541 (1998) Zbl 0910.43010 MR 1662271

[69] Newhouse, S. E.: Quasi-elliptic periodic points in conservative dynamical systems. Amer. J. Math. **99**, 1061–1087 (1977) Zbl 0379.58011 MR 0455049

[70] Ocampo, C.: Trajectory analysis for the lunar flyby rescue of AsiaSat-3/HGS-1. Ann. New York Acad Sci. **1065**, 232–253 (2005)

[71] Ortega, R.: Unbounded motions in forced Newtonian equations. Ann. Mat. Pura Appl. **185**, S245–S257 (2006) Zbl 1103.34042 MR 2187762

[72] Piftankin, G. N.: Diffusion speed in the Mather problem. Nonlinearity **19**, 2617–2644 (2006) Zbl 1190.37067 MR 2267721

[73] Piftankin, G. N., Treschev, D.: Separatrix maps in Hamiltonian systems. Russian Math. Surveys **62**, 219–322 (2007) Zbl 1151.37050 MR 2352365

[74] Ratner, M.: Raghunathan's topological conjecture and distributions of unipotent flows. Duke Math. J. **63**, 235–280 (1991) Zbl 0733.22007 MR 1106945

[75] Sinai, Ya. G.: Geodesic flows on manifolds of negative curvature. In: Algorithms, Fractals, and Dynamics (Okayama/Kyoto, 1992), Plenum, New York, 201–215 (1995) Zbl 0860.58032 MR 1402492

[76] Starkov, A. N.: Minimal sets of homogeneous flows. Ergodic Theory Dynam. Systems **15**, 361–377 (1995) Zbl 1008.22004 MR 1332409

[77] Tennyson, J. L., Lieberman, M. A., Lichtenberg, A. J.: Diffusion in near integrable Hamiltonian systems with three degrees of freedom. On Nonlinear Dynamics and the Beam-Beam Interaction (New York, 1979), M. Month and J. C. Herrera (eds.), Amer. Inst. Phys., New York, 272–301 (1980) MR 0624988

[78] Thirring, W.: A Course in Mathematical Physics: I Classical Dynamical Systems. Springer, New York (1978) Zbl 0387.70001 MR 0507189

[79] Treschev, D.: Evolution of slow variables in a priori unstable Hamiltonian systems. Nonlinearity **17**, 1803–1841 (2004) Zbl 1075.37019 MR 2086152

[80] Treschev, D.: Multidimensional symplectic separatrix maps. J. Nonlinear Sci. **12**, 27–58 (2002) Zbl 1022.37041 MR 1888569

[81] Treschev, D.: Trajectories in a neighborhood of asymptotic surfaces of a priori unstable Hamiltonian systems. Nonlinearity **15**, 2033–2052 (2002) Zbl 1031.37052 MR 1938480

[82] Treschev, D.: Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems. Nonlinearity **25**, 2717–2757 (2012) Zbl 1266.37026 MR 2967122

[83] Weiss, B.: Unique ergodicity on compact homogeneous spaces. Proc. Amer. Math. Soc. **129**, 585–592 (2001) Zbl 0960.22007 MR 1800240

[84] Zaslavsky, G. M.: Chaos, fractional kinetics, and anomalous transport. Phys. Rep. **371**, 461–580 (2002) Zbl 0999.82053 MR 1937584

[85] Zgliczyński, P., Gidea, M.: Covering relations for multidimensional dynamical systems. J. Differential Equations **202**, 32–58 (2004) Zbl 1061.37013 MR 2060531

[86] Zhang, K.: Speed of Arnold diffusion for analytic Hamiltonian systems. Invent. Math. **186**, 255–290 (2011) Zbl 1251.37063 MR 2845619