VARIATIONAL CONSTRUCTION OF POSITIVE ENTROPY INVARIANT MEASURES OF LAGRANGIAN SYSTEMS AND ARNOLD DIFFUSION

SINIŠA SLIJEPČEVIĆ, ZAGREB

Abstract. We develop a variational method for constructing positive entropy invariant measures of Lagrangian systems without assuming transversal intersections of stable and unstable manifolds, and without restrictions to the size of non-integrable perturbations. We apply it to a family of two and a half degrees of freedom a-priori unstable Lagrangians, and show that if we assume that there is no topological obstruction to diffusion (precisely formulated in terms of topological non-degeneracy of minima of the Peierl's barrier function), then there exists a vast family of "horsheshoes", such as "shadowing" ergodic positive entropy measures having precisely any closed set of invariant tori in its support. Furthermore, we give bounds on the topological entropy and the "drift acceleration" in any part of a region of instability in terms of a certain extremal value of the Fréchet derivative of the action functional, generalizing the angle of splitting of separatrices. The method of construction is new, and relies on study of formally gradient dynamics of the action (coupled parabolic semilinear partial differential equations on unbounded domains). We apply recently developed techniques of precise control of the local evolution of energy (in this case the Lagrangian action), energy dissipation and flux. In Part II of the paper we will apply the theory to obtain sharp bounds for topological entropy and drift acceleration for the same class of equations in the case of small perturbations.

Keywords: Hamiltonian dynamics, Arnold diffusion, entropy, variational methods, instability, invariant sets, Lyapunov exponents

MSC2010: Primary 37J40, 37J45; Secondary: 37L45, 37L15, 34C28, 37A35, 37D25

1. Introduction

Consider a C^2, Tonelli Lagrangian $L : \mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T} \to \mathbb{R}$ (see [44] for definitions). The deep motivation for the paper is the Birkhoff ergodicity hypothesis on equivalence of space and time averages for Hamiltonian systems (see [3] for an overview and recent results). Related to that, it is important to know if the "size" (i.e. the natural measure) of the "chaotic", or "unstable" part of the phase space for a typical Hamiltonian is non-zero, which is essentially open even in the simplest case of area-preserving twist diffeomorphisms [25].

A more focused approach to investigate instability of Hamiltonian systems is to consider existence of orbits which "drift" in phase space, or in other words the existence of Arnold diffusion, following Arnold's construction [2] in the case of a weakly coupled rotator and pendulum with a weak periodic forcing. Typical considered questions are on existence of such orbits in specific examples [2, 4, 5, 11, 15], on genericity of existence of such orbits [7, 12, 13, 14, 27, 24, 34, 35, 45, 46], and on the fastest possible drift [8, 10, 48] (the references include only a small sample of the relevant results). Two typical approaches to construction of Arnold diffusion orbits are "geometric" and "variational" (see [6, 29] for an introduction and further references). The geometric approach essentially relies on finding a normally hyperbolic "scaffolding" of a perturbed integrable ("a-priori stable"), or integrable weakly coupled with an "a-priori unstable" (e.g. a pendulum, or a kicked pendulum) Hamiltonian. Futhermore, the geometric method typically relies on transversal intersections of stable and unstable manifolds of the "scaffolding" (e.g. the remaining KAM tori, or a normally hyperbolic cylinder), and construct an orbit which shadows it typically by an application of the implicit function theorem. On the other hand, the variational approach typically relies on minimizing the action under carefully constructed constraints. The variational approach is frequently complemented

Date: 21 June, 2016.

by leveraging the weak-KAM theory [17], or the description of the action-minimizing Mather's sets and invariant measures and their extensions [34, 44], which can result with insightful descriptions of the regions of instability [5, 13].

Our aim is to propose an alternative technique for construction of Arnold diffusion, enabling in addition to construction of individual orbits, also a construction of "shadowing" invariant measures in an a-priori specified region of the phase space. Specifically, we construct a rich family of positive entropy ergodic measures, and are able to estimate their metric entropy. As a result, we can relate the speed of drift in the phase space to the average of locally the largest Lyapunov exponents along a drift trajectory. We thus describe dynamics in a significant region of the phase space (though still most likely of the measure zero with respect to the natural measure of the manifold). Importantly, for our construction to hold, it suffices that there is no topological obstacle to diffusion, and we require no transversal intersections of stable and unstable manifolds. Even though our construction is essentially variational, it is precise enough to incorporate "geometric" information if available.

To introduce the method, we recall an alternative construction of shadowing orbits constructed by Mather [33] in the case of area-preserving twist diffeomorphisms on the cylinder, or equivalently of 1 1/2 degrees of freedom Tonelli Lagrangians on the torus. In [40], we considered formally gradient dynamics of the action (see Section 2 and the equations (2.1) for details) for that system. This is an extension of the variational approach, where we consider evolution of approximate orbits along the gradient of the action, until they "relax" to an equilibrium, which is by LaSalle principle the actual solution of the Euler-Lagrange equations. One of the novelties in [40] is that the evolution dynamics is considered on an unbounded domain, i.e. for all times $t \in \mathbb{R}$, when the dynamics is not gradient-like any more (see Remark 3.1 for details). The construction in [40] is simple: heteroclinic orbits are constructed whenever there is no obstruction to diffusion (in that case homotopically non-trivial invariant circles, or KAM-circles). We simply let any function $q(t)$ asymptotic to two Mather's sets at $t \to -\infty$, respectively $t \to \infty$, in the same region of instability evolve (or relax) along the formal gradient of the action, denoting the relaxation time by s. As there are no invariant KAM circles between these two Mather's sets, we show that the s-evolution of q must asymptotically stop after a finite distance (we show that otherwise there would be a KAM-circle in the limit set - a contradiction), so the "tails" must remain asymptotic to either Mather's sets. The configuration $\lim_{s\to\infty} q(s)$ is the required heteroclinic orbit crossing an arbitrarily large part of a region of instability.

The main tool in [40] - the order-preserving property of the dynamics - does not extend to higher dimensions. The techniques of study of formally gradient systems (called also extended gradient systems), introduced in [19], have recently matured enough [20, 21, 22], so that we can extend the approach to more degrees of freedom. We are now able to replace the monotonicity techniques by "energy methods", or specifically, by considering local interplay of energy, energy dissipation and energy flux, energy being in this case the Lagrangian action.

To explain it, we compare the approach with the variational method introduced by Bessi [11] in the Arnold's example. Bessi somewhat implicitly considered gradient dynamics of the action on a large, but still finite domain. He was then able to "control" evolution of an approximate shadowing orbit, by showing that the *total available action* along the entire constructed (finite, but very long) orbit is less than what is needed for every single section between two "jumps" (corresponding to one heteroclinic orbit in a diffusion "chain") to significantly move and "escape" from the desired region of the phase space. With this method it is difficult to construct orbits with "infinitely many jumps" (as the total "available action" is infinite), and the "control" decays proportionally to the length of the considered orbit.

Thierry Gallay and the author recently developed techniques establishing stability results for dissipative partial differential equations independent of the size of the domain. For example, in [21, 22], we established a-priori bounds for relaxation of unforced Navier-Stokes equations on a strip, independent of the domain size, thus holding for the equations on the unbounded domain. Applying and extending these ideas to formally gradient dynamics of the action, we are thus able to construct orbits of infinite length and invariant measures. We are also able to obtain sharp estimates on the drift acceleration, matching (the case of non-degenerate Melnikov function and small perturbation), or improving (the case of degenerate Melnikov function) the results obtained by the geometric or an alternative approach (further details will be reported in the Part II [43]).

After introducing the general method of constructing shadowing invariant measures, we apply it here to a family of 2 1/2-degrees of freedom a-priori unstable Lagrangians. We, however, believe that the approach can eventually be extended to more general Tonelli Lagrangians, as long as there is a rich family of partially hyperbolic Mather's sets (not necessarily invariant tori), and as long as there is no topological barrier to diffusion, expressed as a certain topological non-degeneracy of the Peierl's barrier function, or of weak KAM solutions.

1.1. **Statements of the main results.** We first develop new tools for construction of orbits and invariant measures for Lagrangians of the type $L(q,q_t,t)=\frac{1}{2}q_t^2+V(q,t)$ (in Remark 2.1 we explain how the tools can be applied to the entire class of Tonelli Lagrangians). We then apply the general theory to a family of a-priori unstable Lagrangians with 2 1/2 degrees of freedom, already considered in e.g. [2, 4, 8, 10, 45], given with

$$\begin{aligned} L(u,v,u_t,v_t,t) &= \frac{1}{2}u_t^2+\frac{1}{2}v_t^2+V(u,v,t), \\ V(u,v,t) &= \varepsilon\,(1-\cos u)\,(1-\mu f(u,v,t))\,, \end{aligned} \tag{1.1}$$

where $(u,v,u_t,v_t,t)\in\mathbb{T}^2\times\mathbb{R}^2\times\mathbb{T}$, $\mathbb{T}$ is parametrized with $[0,2\pi)$, $\varepsilon,\mu\geq 0$ are parameters, and f is 2π-periodic in all the variables. Our standing assumptions are as follows:

(A1): f is $C^{4+\gamma}(\mathbb{R}^3)$, $\gamma>0$ and $|f|\leq 1$, $|f_v|\leq 1$,

(A2): $0\leq 16\mu\leq\varepsilon\leq 1$.

Note that the bounds on f in (A1) are not an essential restriction, as we can always scale f and adjust μ for this to hold. The restricted range of parameters is also used mainly for convenience in the calculations. In any case, (A2) allows the case of "a-priori unstable" small perturbations in μ, as well as other physically relevant cases such as those considered in [39].

The main tool, but also an object of study, is the formally gradient dynamics associated to (1.1), given by the equations

$$u_s = u_{tt}-\partial_u V(u,v,t), \tag{1.2a}$$

$$v_s = v_{tt}-\partial_v V(u,v,t), \tag{1.2b}$$

$$u(0,t)=u^0(t).$$

The techniques we develop here can also be interpreted as new results on uniformly local stability of the equation (1.2) and similar equations on unbounded domains. We hope to make it more explicit in future research.

Consider solutions of the Euler-Lagrange flow induced by the Lagrangian (1.1)

$$u_{tt}=V_u=\varepsilon\sin u(1-\mu f(u,v,t)))+\varepsilon\mu(1-\cos u)f_u(u,v,t), \tag{1.3a}$$

$$v_{tt}=V_v=\varepsilon\mu(1-\cos u)f_v(u,v,t). \tag{1.3b}$$

Let ϕ be the non-autonomous flow induced by (1.3) on $\mathbb{T}^2\times\mathbb{R}^2\times\mathbb{T}$. We use the notation $(u,v,u_t,v_t,t)\in\mathbb{T}^2\times\mathbb{R}^2\times\mathbb{T}$, and always parametrize tori with $[0,2\pi)$. We denote by $\sigma=\phi^{2\pi}$ the time-2π map, and then σ is a diffeomorphism of $\mathbb{T}^2\times\mathbb{R}^2$. As in the classical Arnold's example [2], for each $\varepsilon,\mu\in\mathbb{R}$, and for each "speed of the rotator" $\omega\in\mathbb{R}$, the invariant tori $\mathbb{T}_\omega:=\{(0,v,0,\omega),v\in\mathbb{R}\}$ are σ-invariant (and the sets $(0,v,0,\omega,t)$, $(v,t)\in\mathbb{R}\times\mathbb{T}$) are ϕ-invariant), i.e. these invariant tori persist for all perturbations.

We define regions of instability, generalizing an analogous notion for area-preserving twist maps [33] as follows. Given $\omega\in\mathbb{R}$, let S_ω be the Peierl's barrier function (closely related to weak-KAM solutions of Hamilton-Jacobi equations [17, 44]), defined on $\mathbb{R}^2$ with

$$S_\omega(t_0,v_0)=\inf\left\{\int_{-\infty}^{\infty}L_\omega(q,q_t,t)dt,\ q=(u,v)\in H^1_{\rm loc}(\mathbb{R})^2,\ q(t_0)=(\pi,v_0),\ \lim_{t\to-\infty}u(t)=0,\ \lim_{t\to\infty}u(t)=2\pi\right\},$$

where $L_\omega(q, q_t, t) = u_t^2/2 + (v_t - \omega)^2/2 + V(u(t), v(t), t)$ is the adjusted Lagrangian. It is well-known, and recalled in Section 5, that S_ω is continuous and 2π-periodic in each variable, that for each (t_0, v_0) there exist $q \in H^1(\mathbb{R})^2$ for which the infimum is attained, and that for (t_0, v_0) which are critical points of S_ω, these q correspond to solutions of (1.3) homoclinic to $\mathbb{T}_\omega$. We say that $\omega \in \mathbb{R}$ is non-degenerate, if each connected component of its set of global minima of S_ω is bounded. We interpret it as "no topological obstacle to diffusion"[1]. The set of non-degenerate ω is open (see Remark 6.1), and we call each its connected component a region of instability. We frequently restrict attention to a closed subset $[\omega^-, \omega^+]$ of a single region of instability. We will require an additional, technical assumption that the bounded components of the global minima of S_ω are not too large, i.e. we assume the following:

(S1) For each $\omega \in [\omega^-, \omega^+]$ and for each global minimum (t_0, v_0) of S_ω, there exists a closed neighbourhood $\mathcal{N}$ of (t_0, v_0) such that for each $(t_1, v_1) \in \partial\mathcal{N}$,

$$S_\omega(t_1, v_1) - S_\omega(t_0, v_0) \geq 3\Delta_0, \tag{1.4}$$

where $\Delta_0 > 0$ is an uniform constant over $[\omega^-, \omega^+]$, and the diametar of $\mathcal{N}$ is not greater than R, such that

$$R\sqrt{\varepsilon} \leq 1/144. \tag{1.5}$$

Apart from the technical restriction[2] (1.5), the condition (S1) is equivalent to the definition of the region of instability (see Remark 6.1). Our definition of the region of instability is closely related to the results of Cheng[3] [12, 13, 14] and Bernard[4] [5, 6]. We note that the quantity $3\Delta_0$ in (1.4) is analogous to the quantity ΔW quantifying transport in the case of area-preserving twist maps[5] [32].

Importantly, we do not require any non-degeneracy of the minima of S_ω, for example we do not require that the second derivative of S_ω at these minima is positive definite. As perhaps noted first by Angenent [1], such non-degeneracy would be equivalent to requiring that the stable and unstable manifolds (the "whiskers") of $\mathbb{T}_\omega$ intersect transversally, and would lead to the "geometric" approach to diffusion phenomena ([29] and references therein). On the contrary, we allow the whiskers to intersect non-transversally, and construct possibly non-uniformly hyperbolic invariant sets. The main result is now a construction of a large number of "horsheshoes", i.e. of ergodic positive entropy measures roughly contained in an arbitrary part of a region of instability:

Theorem 1.1. *Assume $[\omega^-, \omega^+]$ satisfies (S1). Then for each closed subset $\mathcal{O}$ of $[\omega^-, \omega^+]$, there exist an ergodic, ϕ-invariant, positive entropy Borel probability measure μ on $\mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{R}$, such that*

$$\cup_{\omega \in \mathcal{O}} \mathbb{T}_\omega \subset \operatorname{supp} \mu, \quad \cup_{\omega \in \mathbb{R} - \mathcal{O}} \mathbb{T}_\omega \subset (\operatorname{supp} \mu)^c. \tag{1.6}$$

We call the invariant measures described by (1.6) the shadowing measures, as they indeed "shadow" an arbitrary closed set of invariant tori. A direct corollary of the Theorem 1.1 is that we can find a single, "dense" orbit whose closure contains the entire $\cup_{\omega \in [\omega_1, \omega_2]} \mathbb{T}_\omega$ (we choose $\mathcal{O} = [\omega^-, \omega^+]$, and by ergodicity we find an orbit which is dense in the support of the measure μ). We actually by our method also give a direct proof of existence of such and more elaborated shadowing orbits in Theorem 13.2. One can apply it to construct other complex structures in a vicinity of a chosen $\mathbb{T}_\omega$ for non-degenerate ω, such as an analogoue of the Mather's construction in the case of twist maps[6] [31].

Importantly, we are also able to obtain estimates of the metric entropy of constructed measures, topological entropy of σ and ϕ (by the variational principle for metric and topological entropy), and of the drift acceleration (or, using a less precise term, the speed of diffusion). Specifically, we can associate to each closed interval $[\omega^-, \omega^+]$ a value Δ_1, expressed as a certain extremal value of the norm of the Fréchet derivative of the action, and defined precisely in Proposition 10.1. The value Δ_1 can be understood as a generalization of the lower bound on the angle of splitting of separatrices[7]. Importantly, $\Delta_1 > 0$ whenever (S1) holds. Let $\varpi = \max\{|\omega^-|, |\omega^+|, 1\}$. To avoid repetition in the statements, we say that the topological entropy on $[\omega^-, \omega^+]$ is $O(h)$, if there exists an absolute constant $1 \geq c_0 > 0$ such that the the topological entropy of ϕ and σ restricted to an invariant subset of

$$\begin{aligned} &|u_t| \leq c_0 \varpi \sqrt{\varepsilon}, \\ &v_t \in [\omega^- - c_0 \varpi \sqrt{\varepsilon}, \omega^+ + c_0 \varpi \sqrt{\varepsilon}] \end{aligned} \tag{1.7}$$

is at least $c_0 h$. We say that the *drift acceleration* on $[\omega^-, \omega^+]$ is $O(d)$, if for each $\delta > 0$ there exists an absolute constant $1 \geq c(\delta) > 0$ (depending only on δ), a solution $q = (u, v)$ of (1.3) satisfying (1.7) for all $t \in \mathbb{R}$, and the times $t^- < t^+$ such that $|v_t(t^-) - \omega^-| \leq \delta$, $|v_t(t^+) - \omega^+| \leq \delta$, and such that

$$d \geq c(\delta) \frac{|\omega^+ - \omega^-|}{|t^+ - t^-|}.$$

Theorem 1.2. *Assume $\omega^- \leq \omega^+$ such that $[\omega^-, \omega^+]$ satisfies (S1), and let $\Delta_1 > 0$ be as is defined in Proposition 10.1. Then:*

(i) *The topological entropy of ϕ and σ on $[\omega^-, \omega^+]$ is*

$$O\left(\frac{\Delta_1}{\varpi^5 |\log \Delta_1|} \right),$$

(ii) *The drift acceleration on $[\omega^-, \omega^+]$ is*

$$O\left(\frac{\Delta_0 \Delta_1}{\varpi^6 (R \vee \mu) |\log \Delta_1|} \right).$$

Our estimates are consistent with upper bounds on the drift acceleration in the cases considered by Nekhoroshev[8] [37], as well as on upper bounds in [10] (see below).

The emerging picture of the Arnold diffusion is actually more subtle. We can set $\omega^- = \omega^+$, in Theorem 1.2, and find $\Delta_1 = \Delta_1(\omega)$ and an ergodic positive entropy measure μ_ω in a neighborhood of $\mathbb{T}_\omega$ with the locally maximal metric entropy as in Theorem 1.2, (ii). By the Margulis-Ruelle inequality, there is a positive Lyapunov exponent with respect to μ_ω, which is at least $\sim \Delta_1(\omega)/|\log \Delta_1(\omega)|$. The drift acceleration seems to be proportional to the integral of these Lyapunov exponents with respect to ω along a diffusion path. This picture somewhat explains, and provides tools to study the dynamics of various "time-scales", i.e. the observed diffusion in Hamiltonian systems with transport speed substantially varying in different regions of the phase space (see e.g. [9, 23] and references therein).

Consider now the case of small perturbations (i.e. small $\mu > 0$). Recall the definition of the Melnikov primitive $M_\omega : \mathbb{R}^2 \to \mathbb{R}$,

$$M_\omega(t_0, v_0) = -\varepsilon \int_{-\infty}^{\infty} (1 - \cos(u^\varepsilon(t - t_0))f(u^\varepsilon(t - t_0), v_0 + \omega(t - t_0), t)dt, \tag{1.8}$$

where $u^\varepsilon = 4 \operatorname{arctg} e^{\sqrt{\varepsilon}\, t}$ is the separatrix of the unperturbed system, i.e. the case $\mu = 0$ when (1.3) reduces to an uncoupled pendulum and a rotator. It is well-known that for sufficiently small μ and fixed ω, the oscillations of M_ω approximate well the oscillations of S_ω, the minima of M_ω approximate the minima of S_ω, and non-degeneracy of the minima of M_ω implies non-degeneracy of minima of S_ω (i.e. transversal intersection of "whiskers"). To get the conclusions of Theorems 1.1 and 1.2, it thus suffices to assume an analogue of (S1) for M_ω:

(S2) For each $\omega \in [\omega^-, \omega^+]$, and for each global minimum (t_0, v_0) of M_ω, there exists a closed neighbourhood $\mathcal{N}$ of (t_0, v_0) such that for each $(t_1, v_1) \in \partial \mathcal{N}$,

$$M_\omega(t_1, v_1) - M_\omega(t_0, v_0) \geq 4\tilde{\Delta}_0, \tag{1.9}$$

where $\tilde{\Delta}_0 > 0$ is an uniform constant over $[\omega^-, \omega^+]$, and the diametar of $\mathcal{N}$ is not greater than R, where R satisfies (1.5).

Theorem 1.3. *Assume that (S2) holds. Then there exists $\mu_0 > 0$, such that for for each $0 < \mu \leq \mu_0$, the conclusions of Theorem 1.1 and 1.2 hold, with $\Delta_0 = \mu \tilde{\Delta}_0$.*

In the Part II of the paper [43], we will explicitly estimate Δ_0, Δ_1, the topological entropy and drift acceleration for small μ under different assumptions on the Melnikov primitive. We will show that the approach seems to give optimal estimates as compared to known results for (1.1). For example, we will obtain the drift acceleration $O(\mu/|\log \mu|)$ and topological entropy $O(1/|\log \mu|)$ which is optimal[9], in the case of Melnikov primitive with non-degenerate minima and small μ. We will strengthen known results, and also show that for such fast drift acceleration, it suffices that $||D^2 M_\omega(t_0(\omega), v_0(\omega)||^{-1}$ (where $(t_0(\omega), v_0(\omega))$

minimize M_ω) is integrable along a diffusion path (equivalently, that the inverse of the splitting angles of separatrices is integrable), as long as (S2) holds. Furthermore, we will show that for small $\mu \leq \mu_0$, (1.5) is not needed (however, μ_0 is then inverse proportional to R), and will obtain new estimates for topological entropy and drift acceleration for M_ω with degenerate minima as a function of the leading term in the Taylor expansion of M_ω at the minimum.

Finally, we remark that all the results hold in the classical Arnold's example [2] with $f(u,v,t) = \cos v + \cos t$. In that case, the Melnikov primitive M_ω can be explicitly calculated [2, 11], and it satisfies (S2) with the regions of instability $(-\infty, 0)$ and $(0, \infty)$. We can thus obtain diffusion orbits and shadowing measures in the Arnold's example for $\varepsilon \leq 1$ and sufficiently small[10] $\mu > 0$, without restrictions to ω, as long as the sign of ω does not change. One can tailor the argument to also obtain accelerating orbits, i.e. orbits with unbounded ω in that case (and any case when f does not depend on u).

Remarks 1.1. (1) Consider the stable and unstable manifolds of $\mathbb{T}_\omega$ of the time-2π map σ. Then an unbounded family of global minima of S_ω corresponds to an unbounded, connected family of homoclinic orbits of σ, which can not be "crossed" by other orbits on 2-dimensional stable and unstable manifolds of σ. This would prevent drift and the complex dynamics we describe in Theorem 1.1 in that region of the phase space.

(2) The restriction on R is used only in the proof of Proposition 12.1 at the end of Section 12, to assist in the cases of topologically complex $\mathcal{N}$ in (S1). We will show in [43] that this is not needed in the case of small perturbations (sufficiently small μ). Alternatively, one can instead assume that $\mathcal{N}$ is convex, and bound the interval $[\omega^-, \omega^+]$ away from zero. Note that Bernard in [4] considered the same equation for sufficiently small μ with an assumption analogous to Theorem 1.3, but assuming $\mathcal{N}$ to be rectangular. Bernard used the method of Bessi [11] (see the earlier discussion on the Bessi method for a comparison), and constructed diffusion orbits of finite length in a restricted range of ω.

(3) The definition of the region of instability by Cheng and Yan [12, 13, 14] seems to be essentially equivalent to ours (they give it in a more general and abstract setting). Our understanding is that the method in these papers does not result with diffusion orbits of infinite length, thus does not imply existence of invariant measures, and that it does not immediately give estimates of the drift acceleration, topological entropy and Lyapunov exponents.

(4) Our results in Theorem 1.2 can be interpreted in the sense of Bernard's *forcing relation* of cohomology classes [5]: if ω and $\tilde{\omega}$ are in the same region of instability, then $(0,\omega), (0,\tilde{\omega}) \in H^1(\mathbb{T}^2, \mathbb{R})$ are related.

(5) Assume we take $\omega^- = \omega^+$, and find the largest $3\Delta_0$ such that (S1) holds. Then $3\Delta_0$ is the difference of actions of a minimizing and a "minimax" (in this case a "saddle") homoclinic orbit. In the case of area-preserving twist maps, one would analogously obtain exactly the quantity ΔW quantifying transport through gaps in Cantori [32].

(6) Mather in [31] constructed uncountably many minimal sets of twist maps with the same irrational angular rotation ω. One can adapt our construction in the proof of Theorem 1.2 to construct uncountably many minimal sets for irrational ω, supported on the set $\lim_{T\to\infty}(v(T) - v(-T))/2 = \omega$, by essentially constructing orbits shadowing a couple of orbits homoclinic to $\mathbb{T}_\omega$ (jumping "forward" and "backward") with the time between "jumps" in the set nL, n in a fixed subset of $\mathbb{N}$. We intend to provide details separately.

(7) We explain it by analogy. Consider a sufficiently smooth function $S : \mathbb{R}^n \to \mathbb{R}$ with a local minimum $x = 0$. If the minimum is nondegenerate and $(D^2S(0)x, x) \geq A||x||^2$, for small enough $\delta > 0$ on the level sets $S(x) = A\delta^2$ in a neighborhood of 0 we trivially have that $||DS(x)|| \geq 2A\delta + O(\delta^2)$. If, however, we merely assume that the set of local minima of S is bounded, we can by Morse-Sard theorem find level sets of S arbitrarily close to the set of minima, such that $||DS(x)|| \neq 0$ on that level set, and by compactness of level sets close enough to the bounded set of minima and smoothness of S, we can find a lower bound $||DS(x)|| \geq \Delta_1 > 0$ on any such level set. Now we take S_ω instead of S, acting on a suitable Banach space (see [1] or Section 10). If the stable and unstable manifolds of $\mathbb{T}_\omega$ intersect transversally, D^2S_ω is in a certain sense non-degenerate [1], the constant A can be interpreted as the angle of splitting, and $\Delta_1 \sim$ the lower bound on the norm of the Fréchet derivative of S_ω on a level set of S_ω is proportional to A (we actually take the square of the norm and find a level set which maximizes Δ_1). If, however, we only assume that the set of minima of S_ω is bounded, we analogously to the finite-dimensional case obtain $\Delta_1 > 0$ by an application of an infinite dimensional analogoue of the Morse-Sard theorem (see the proof of Proposition 10.1, and Part II for further discussion and examples [43]).

(8) In the cases considered by Nekhoroshev such as the Arnold's example [2], both Δ_0 and Δ_1 are exponentially small in ε, which is consistent with [37]. They are, however, polynomial in μ - see below.

(9) The "fast diffusion" (with respect to the perturbation μ) for sufficiently small μ has the drift acceleration $O(\mu/|\log\mu|)$, as conjectured by Lochak, and proved for a class of a-priori unstable systems similar to (1.1) (also allowing the dimension of the rotator variable v to be ≥ 1) with non-transveral intersection of whiskers by Berti, Biasco, Bolle, and Treschev [10, 45, 46]. In [10], it was established that this is under certain assumptions the largest possible drift acceleration. If we only assume (S2), then the drift acceleration is $O(\mu^2)$, as shown in [4, 11].

(10) One could extend the results in the Arnold's example to arbitrary μ (and other cases of (1.1)), by developing a computer-assisted proof verifying (S1) in the spirit of [39], by using all the a-priori bounds we develop here.

1.2. **The structure of the proof and notation.** In Sections 2-4, we consider the Lagrangian of the type $L(q, q_t, t) = \frac{1}{2}q_t^2 + V(q,t)$ on $\mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T}$, and develop general tools for constructing solutions and "shadowing" invariant measures of Euler-Lagrange equations. Specifically, in Section 2 we prove existence of solutions of the formally gradient dynamics of the action on unbounded domains, on function spaces large enough to contain solutions of Euler-Lagrange equations with merely a bounded momentum. We then in Section 3 show that invariant sets with respect to the formally gradient dynamics bounded in norm contain in its closure the Euler-Lagrange equations. The key tool is then developed in Section 4, where we extend these ideas to construction of shadowing invariant measures.

In Sections 5-12, we then focus on the a-priori unstable Lagrangian (1.1), and construct invariant sets of the formally gradient flow as required by the general setting. The construction is based on the following simple idea. Assume for the moment that ξ is an abstract continuous semiflow on a separable metric space $\mathcal{X}$, and let $\mathcal{A}$, $\tilde{\mathcal{B}}$ be subsets of $\mathcal{X}$. Assume they satisfy the following:

- **(B1):** $\tilde{\mathcal{B}}$ is $\mathcal{A}$-relatively $\xi-$invariant set. That means if $q(s_0) \in \tilde{\mathcal{B}}$, and if there exists $s_1 > s_0$ such that for all $s \in [s_0, s_1]$, $\xi^{s-s_0}(q) \in \mathcal{A}$, then for all $s \in [s_0, s_1]$, $\xi^{s-s_0}(q) \in \tilde{\mathcal{B}}$.
- **(B2):** There exists $\lambda > 0$ such that, if $q(s_0) \in \mathcal{A} \cap \tilde{\mathcal{B}}$, then for all $s \in [s_0, s_0 + \lambda]$, $q(s) \in \mathcal{A}$.

Lemma 1.4. *Assume $\mathcal{A}$, $\tilde{\mathcal{B}}$ are subsets of a separable metric space $\mathcal{X}$ satisfying (B1), (B2) with respect to a continuous semiflow ξ on $\mathcal{X}$. Then $\mathcal{B} = \mathcal{A} \cap \tilde{\mathcal{B}}$ is ξ-invariant.*

Proof. Assume the contrary and find a semi-orbit $q(s)$ of ξ, $s \geq s_0$, $q(s_0) \in \mathcal{A} \cap \tilde{\mathcal{B}}$ which violates the conclusion of the Lemma. Let

$$s_2 = \sup \left\{ s_1 \geq 0,\ q(s) \in \mathcal{A} \cap \tilde{\mathcal{B}} \text{ for all } s \in [0, s_1] \right\}.$$

Then if $s_3 = \max\{s_0, s_2 - \lambda/2\}$, by construction $q(s_3) \in \mathcal{A} \cap \tilde{\mathcal{B}}$. Now by (B2), for all $s \in [s_3, s_3 + \lambda]$, we have $q(s) \in \mathcal{A}$, and by (B1), for all $s \in [s_3, s_3 + \lambda]$ we obtain $q(s) \in \tilde{\mathcal{B}}$. But $s_3 + \lambda > s_2$, which is a contradiction. □

We construct the sets $\mathcal{A}, \tilde{\mathcal{B}}$ as follows. Let ξ be the "formally gradient semiflow" introduced in Section 2. We define the set $\mathcal{A}$ in Section 7 by very roughly fixing the times of "jumps" between invariant tori $\mathbb{T}_\omega$. The set $\tilde{\mathcal{B}}$ satisfying the conditions (B1), (B2) is then built using the "Russian doll" approach. We construct a decreasing sequence of sets $\mathcal{B}_1 \supset \mathcal{B}_2 \supset ... \supset \mathcal{B}_6$, showing inductively in each step that they are $\mathcal{A}$-relatively ξ-invariant. Finally, in $\mathcal{B}_6$ we establish sufficient control to also prove by an energy method the condition (B2), as required by Lemma 1.4. Specifically, in Sections 5, 6 we recall the known results on homoclinic, heteroclinic orbits and the Peierl's barrier, and prove a-priori bounds on minimizing homoclinics and heteroclinics. In Section 7, we fix a sequence of tori $\mathbb{T}_{\omega_k}$, $k \in \mathbb{Z}$, and construct a rough approximation of a shadowing orbit. In our method, it is not required that this approximation is very precise. We then in Sections 8 and 9 construct sets $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$, $\mathcal{B}_4$, by establishing L^∞-bounds, as well as weighted L^2-bounds on the first, second and third derivatives. The core of the argument is then in Sections 10 and 11, where we establish local control of the dynamics ξ between two "jumps" roughly independently of the behavior away from these jumps. The argument relies on precise control of the local "energy", "energy dissipation" and "energy flux" with respect to ξ, where "energy" is in this case the Lagrangian action. The approach is inspired by the results from [19, 20, 21, 22]. One of the novelties is the use of an infinite-dimensional version of the Morse-Sard Theorem, enabling us to establish lower > 0 bounds on the dissipation on certain "action" levels arbitrarily close to the minimal action along a heteroclinic orbit. We then show that these action levels can not be crossed by ξ, as the action dissipation is larger than the action flux through the boundary of the considered interval $t \in [\tilde{T}_k - L, \tilde{T}_k + L]$, where $\tilde{T}_k$ is the approximate time of a "jump" and L the minimal time between jumps, for L large enough.

We thus establish uniformly local control of the dynamics ξ, enabling us to construct invariant sets independently of the number of "jumps" between invariant tori. This allows the number of jumps to be infinite, and establishes "variational" control for all $t \in \mathbb{R}$. We complete the construction of an invariant set

$\mathcal{B}$, as a function of a given sequence of tori to be shadowed, in Section 12. The last step of the construction is somewhat subtle (the set $\mathcal{B}_6$), and uses in a topological way the existence of the semiflow ξ.

In Sections 13 and 14, we then focus on proving Theorems 1.1-1.3. Several technical results needed throughout the paper are given in Appendices A-D at the end of the paper.

Main notation. We denote by $\mathcal{X}$ the phase space on which we consider the formally gradient semiflow ξ, introduced in Section 2. The elements of $\mathcal{X}$ are always denoted by q, and in the case $N = 2$ consistently with $q = (u, v)$. By $\mathcal{E}$ we denote the set of equilibria of ξ, which by correspondences π_t and π specified in Section 3 are the solutions of the Euler-Lagrange equations. The fixed velocities v_t specifying the invariant tori $\mathbb{T}_\omega$ are denoted by ω. The sequence of "jump" times is given with $\tilde{T}_k$, and $T_k = \tilde{T}_k \mod 2\pi$, $T_k \in [0, 2\pi)$. An approximate shadowing orbit is denoted by q^0, and defined in Section 7. Various constants are fixed throughout the proof: the constants L (minimal length of time t between the "jumps"), M (maximal oscillations of the "rotator" variable v) depend on the particular choices of f, and the sequence of tori to be shadowed. The constants R (introduced in (S1)) and $\varpi = \max\{|\omega^-|, |\omega^+|, 1\}$ depend on the choice of f and the segment $[\omega^-, \omega^+]$ in a region of instability. We denote by $c_1, c_2, ...$ fixed absolute constants, though they may change within the proof when introduced. The symbol $g \ll h$ always means $g \leq c \cdot h$ for some absolute constant c. If the constant depends on ε or f, we write $g \ll_\varepsilon h$ or $g \ll_f h$. We use $g \wedge h$ instead of $\min\{g, h\}$ and $g \vee h$ instead of $\max\{g, h\}$. Given two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are σ-algebras, if μ is a probability measure on $(\Omega_1, \mathcal{F}_1)$ and $g : \Omega_1 \to \Omega_2$ measurable, we denote by $f^*\mu$ the pulled measure on $(\Omega_2, \mathcal{F}_2)$ given with $f^*\mu(\mathcal{C}) = \mu(f^{-1}(\mathcal{C}))$.

I: VARIATIONAL CONSTRUCTION OF ORBITS AND INVARIANT MEASURES

2. Existence of solutions and the function spaces

In this and the next two sections, we consider the Lagrangian $L(q, q_t, t) = \frac{1}{2}q_t^2 + V(q, t)$, $L : \mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T} \to \mathbb{R}$, where $\mathbb{T}$ is always parametrized with $[0, 2\pi)$, and V is C^2, 2π-periodic in all coordinates. Here we prove existence of solutions of the formally gradient dynamics of the action, given with:

$$q_s = q_{tt} - \frac{\partial}{\partial q}V(q, t). \tag{2.1}$$

We consider (2.1) on the Banach space $\mathcal{X}$ of all $q = (u, v) \in H^2_{\text{loc}}(\mathbb{R})^N$, such that $q_t \in H^1_{\text{ul}}(\mathbb{R})^N$ (in Appendix A we recall the definition of the uniformly local spaces $H^k_{\text{ul}}(\mathbb{R})^N$, the associated norms and their properties). The norm on $\mathcal{X}$ is given with

$$||q||_{\mathcal{X}_{\text{ul}}} = \left(q(0)^2 + ||q_t||^2_{H^1_{\text{ul}}(\mathbb{R})^N}\right)^{1/2}.$$

We will frequently require an alternative, weaker, localized topology on $\mathcal{X}$, induced by the $H^1_{\text{loc}}(\mathbb{R})^N$ topology on $\mathcal{X}$. We use the notation $\mathcal{X}_{\text{ul}}$ and $\mathcal{X}_{\text{loc}}$ respectively to distinguish the topologies on the same set $\mathcal{X}$. Now $\mathcal{X}_{\text{loc}}$ is a normed (but not complete) space with the norm

$$||q||_{\mathcal{X}_{\text{loc}}} = \left(\int_{-\infty}^{\infty} e^{-|t|}\left(q^2(t) + q_t^2(t)\right)dt\right)^{1/2}.$$

Denote by $\varphi^y q(t) = q(y + t)$ the translation (corresponding to the time evolution of the Euler-Lagrange equations, as discussed in detail in the next section). By definition of uniformly local spaces, φ is a continuous flow on $\mathcal{X}_{\text{ul}}$, and by definition of the topology, also on $\mathcal{X}_{\text{loc}}$. The existence of solutions of (2.1) is given with:

Theorem 2.1. *Assume $q^0 \in \mathcal{X}$ at s_0 is the initial condition. Then:*

(i) There exists unique $q(s) \in \mathcal{X}$ for all $s \in [s_0, \infty)$, $q(s_0) = q^0$, so that

$$q - q^0 \in C^0([s_0, \infty), H^2_{ul}(\mathbb{R})^N) \cap C^1((s_0, \infty), H^2_{ul}(\mathbb{R})^N),$$

and so that for all $s > s_0$, q is a solution of (2.1).

(ii) The system (2.1) generates a continuous semiflow ξ on $\mathcal{X}_{loc}$ and $\mathcal{X}_{ul}$.

(iii) The semiflow ξ and the flow φ commute.
(iv) If $V \in H^k(\mathbb{T}^{N+1})$, $k \geq 1$, then for all $s > s_0$ we have that $q(s) \in H^k_{ul}(\mathbb{R})^N$.

The proof of Theorem 2.1 follows the standard approach for semilinear parabolic equations [26], only on a larger space than usual, and is given in the Appendix A.

Remark 2.1. All the results of the Part I hold if we consider a more general, C^2 Tonelli Lagrangian $L(q, q_t, t)$ on $\mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T}$ (see e.g. [34, 44] for background and definitions). In that case, instead of (2.1), we consider

$$q_s = q_{tt} + \left(\frac{\partial^2}{\partial q_t^2} L(q, q_t, t) \right)^{-1} \left(\frac{\partial^2}{\partial q \partial q_t} L(q, q_t, t) + \frac{\partial^2}{\partial t \partial q_t} L(q, q_t, t) - \frac{\partial}{\partial q} L(q, q_t, t) \right). \tag{2.2}$$

For example, stationary points of (2.2) are indeed solutions of the Euler-Lagrange equations. One can in particular verify that (2.2) on $\mathcal{X}$ is an extended gradient system in the sense of [19, 20], and that the proofs of Theorems 3.1 and 4.4 can be generalized to hold. We develop the theory in a simpler case for clarity of the introduced ideas.

3. Existence of Euler-Lagrange orbits in invariant sets

Consider equilibria (or stationary solutions) of (2.1), i.e. the solutions of

$$q_{tt} = \frac{\partial}{\partial q} V(q, t). \tag{3.1}$$

We denote by $\mathcal{E}$ the set of all $q \in \mathcal{X}$ satisfying (3.1). Let $\pi_t : \mathcal{E} \to \mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T}$, $\pi_t(q) = (q(t) \mod 2\pi, q_t(t), t \mod 2\pi)$. Then by the continuous dependence on initial conditions of (3.1), π_t is continuous in both of the topologies $\mathcal{X}_{\text{ul}}$, $\mathcal{X}_{\text{loc}}$ induced on $\mathcal{E}$. Furthermore, we have that for any $t_1, t_2 \in \mathbb{R}$, $\pi_{t_1+t_2} \circ \varphi^{t_1} = \phi^{t_2} \circ \pi_{t_1}$, i.e. $\mathcal{E}$ correspond to the solutions of (2.1), and the t-translation on $\mathcal{E}$ corresponds to the t-evolution of a solution of (3.1). Similarly, if $\pi : \mathcal{E} \to \mathbb{T}^N \times \mathbb{R}^N$ is the projection in the zero-coordinate, $\pi(q) = (q(0) \mod 2\pi, q_t(0))$, and $S = \varphi^{2\pi}$ is the t-translation for one period, then $\pi \circ S = \sigma \circ \pi$.

The space $\mathcal{X}$ is large enough so that the projection of $\mathcal{E}$ to $\mathbb{T}^N \times \mathbb{R}^N$ contains a rich set of orbits. Specifically, we show in Lemma 15.5 in the Appendix A that if a solution of (3.1) satisfies $||q_t||_{L^\infty(\mathbb{R})^N} < \infty$, then $q \in \mathcal{X}$, thus $q \in \mathcal{E}$.

The result of this section is that, to construct elements of $\mathcal{E}$, i.e. to find solutions of (3.1), it suffices to find an invariant set of (2.1):

Theorem 3.1. *Assume $\mathcal{B}$ is a non-empty, ξ-invariant subset of $\mathcal{X}$, bounded in the $\mathcal{X}_{ul}$-norm. Then there is a $q \in \mathcal{E}$ in the closure of $\mathcal{B}$ in $\mathcal{X}_{loc}$.*

Remark 3.1. We comment why Theorem 3.1 is not entirely straightforward. The system (2.1) belongs to a class of extended gradient systems (or formally gradient systems), introduced in a general setting in [19, 20]. These are dynamical systems which, when considered on bounded domains, are gradient-like; but on an unbounded domain may behave differently. For example, for the system of equations $q_{tt} = \Delta q - \partial_q V(q)$, if we require that q decays fast enough at ∞, the system is gradient-like, and ω-limit sets (considered in a sufficiently weak topology so that orbits uniformly bounded in norm are relatively compact) by LaSalle principle consist of equilibria. If q is merely bounded, and we consider solutions $q : \mathbb{R}^M \to \mathbb{R}^N$, then for $M = 1, 2$, ω-limit sets may contain non-equilibria, but always contain at least one equilibrium, and for $M \geq 3$, there are examples of ω-limit sets without equilibria at all [19, 20]. As in (2.1), $M = 1$ (the dimension of the variable t), Theorem 3.1 holds. We adapt the proof from [19, 20] to our setting.

In the first lemma below, we establish a bound on the action dissipation, then we establish relative compactness of the required set, and finally construct $q \in \mathcal{E}$ by a variational argument.

Lemma 3.2. *Assume $q(s) \in \mathcal{B}$, $s \geq s_0$ is an orbit of ξ. There exists an absolute constant $c_1 > 0$ and a sequence of relaxation times s_n, $n \in \mathbb{Z}$, so that*

$$\int_{-n}^{n} q_s(t, s_n)^2 dt \leq \frac{c_1}{n} \left(||q(s_0)||^2_{\mathcal{X}_{ul}} + 1 \right).$$

Proof. Let $\delta > 0$, and denote by E_δ, D_δ the weighted action and action dissipation,

$$E_\delta(q) = \int_{-\infty}^{\infty} e^{-\delta|t|} L(q(t), q_t(t), t)dt, \qquad D_\delta(q) = \int_{-\infty}^{\infty} e^{-\delta|t|} q_s(t)^2 dt.$$

It is straightforward to verify that E_δ, D_δ are well-defined on $\mathcal{X}$ (the integrals are absolutely integrable). Furthermore, we can differentiate with respect to s, by calculating on a dense subset and then extending the final result to the entire $\mathcal{X}$ by continuity. We partially integrate and apply the Young's inequality in the second line below:

$$\begin{aligned}\frac{d}{ds}E_\delta(q(s)) &= \int_{-\infty}^{\infty} e^{-\delta|t|}\left(q_t q_{ts} + \frac{\partial}{\partial q}V(q,t)q_s\right)dt \leq \delta \int_{-\infty}^{\infty} e^{-\delta|t|}|q_t q_s|dt - \int_{-\infty}^{\infty} e^{-\delta|t|} q_s^2 dt \\ &\leq \frac{\delta^2}{2}\int_{-\infty}^{\infty} e^{-\delta|t|} q_t^2 dt + \frac{1}{2}D_\delta(q(s)) - D_\delta(q(s)) \\ &\leq \delta^2 E_\delta(q(s)) - \frac{1}{2}D_\delta(q(s)).\end{aligned}$$

Now by the Gronwall Lemma, integrating it over $[s_0, s_0 + 1/\delta^2]$, we have

$$e^{-1}E_\delta(q(s_0 + 1/\delta^2)) + \frac{1}{2}\int_{s_0}^{s_0+1/\delta^2} e^{-(s-s_0)\delta^2} D_\delta(q(s))ds \leq E_\delta(q(s_0)). \tag{3.2}$$

It is easy to see that $E_\delta(q)$ can be bounded by $O\left(\left(||q_t||^2_{L^2_{\mathrm{ul}}(\mathbb{R})^2} + 1\right)/\delta\right)$, thus by definition of the $\mathcal{X}_{\mathrm{ul}}$-norm,

$$E_\delta(q) \ll \frac{1}{\delta}\left(||q||^2_{\mathcal{X}_{\mathrm{ul}}} + 1\right).$$

Also by definition, $L(q, q_t, t) \geq 0$, thus $E_\delta(q) \geq 0$. Inserting it in (3.2) we otain

$$\int_{s_0}^{s_0+1/\delta^2} D_\delta(q(s))ds \ll \frac{1}{\delta}\left(||q(s_0)||^2_{\mathcal{X}_{\mathrm{ul}}} + 1\right),$$

thus by definition of D_δ,

$$\int_{s_0}^{s_0+1/\delta^2}\left(\int_{-1/\delta}^{1/\delta} q_s(t)^2 dt\right)ds \ll \frac{1}{\delta}\left(||q(s_0)||^2_{\mathcal{X}_{\mathrm{ul}}} + 1\right). \tag{3.3}$$

Now set $\delta = 1/n$. From (3.3) it follows immediately that there exists the required s_n, $s_0 \leq s_n \leq s_0 + n^2$, so that the claim holds with c_1 being the absolute constant in (3.3). □

Lemma 3.3. *If $\mathcal{B}$ is bounded in $\mathcal{X}_{ul}$, then it is relatively compact in the closure $\bar{\mathcal{X}}_{loc}$ of $\mathcal{X}_{loc}$ in $H^1_{loc}(\mathbb{R})^N$.*

Proof. It is easy to check from the definition of the $\mathcal{X}_{\mathrm{ul}}$-norm, that boundedness of $\mathcal{B}$ in $\mathcal{X}_{ul}$ implies boundedness of $q|_{[-n,n]}$ in $H^2([-n,n])^N$, uniformly for $q \in \mathcal{B}$, for any $n > 0$. Thus for any sequence $q^{(j)}$ in $\mathcal{B}$, by compact embedding we can find a subsequence (again denoted by $q^{(j)}$) so that $q^{(j)}|_{[-n,n]}$ converges in $H^1([-n,n])^N$; and by diagonalization a further subsequence converging in $H^1_{\mathrm{loc}}(\mathbb{R})^N$ (which induces the $\mathcal{X}_{\mathrm{loc}}$ topology by definition). □

Lemma 3.4. *Assume $s_n \to \infty$ as $n \to \infty$ is a sequence of times such that*

$$\lim_{n\to\infty}\int_{-n}^{n} q_s(t, s_n)^2 dt \to 0.$$

Then any limit point of $q(s_n)$ in $\bar{\mathcal{X}}_{loc}$ is in $\mathcal{E}$.

Proof. Fix $m \in \mathbb{Z}$, and choose a test function $g \in H_0^1([-m,m])^2$ (that is, vanishing at $t=-m,m$). Now by partial integration and Cauchy-Schwartz,

$$\begin{aligned}\int_{-m}^{m}\left(q_t(t,s_n)g_t(t)+\frac{\partial V}{\partial q}(q(t,s_n),t)g(t)\right)dt &= \int_{-m}^{m}(-q_s(t,s_n)g(t))\,dt \\ &\leq \left(\int_{-m}^{m}q_s(t,s_n)^2dt\right)^{1/2}||g||_{L^2([-m,m])^2}.\end{aligned}$$

Now if $q(s_n,.)$ converges to some q^0 in $\bar{\mathcal{X}}_{\text{loc}}$, their restrictions to $[-m,m]$ converge in $H^1([-m,m])^2$. We deduce that

$$\int_{-m}^{m}\left(q_t^0(t)g_t(t)+\frac{\partial V}{\partial q}(q^0(t),t)g(t)dt\right)dt=0.$$

As it holds for an arbitrary test function g, we conclude that the variation of the action at q^0 is 0, so q^0 is a solution of (3.1). By construction, $q_t \in L^2_{\text{ul}}(\mathbb{R})^N$, thus by Lemma 15.5 we have that $q \in \mathcal{E}$. □

Theorem 3.1 follows by combining Lemmas 3.2, 3.3 and 3.4.

4. Construction of shadowing invariant measures

In Section 3, we showed how to construct a solution of (3.1), given an invariant set with respect to the semiflow ξ. In this section we develop a measure-theoretical analogue to that. In the first subsection, we propose an abstract notion of a shadowing measure and derive its properties. In the second subsection, we prove existence of such measures, if a certain sub-algebra of Borel sets with certain invariance property with respect to the semiflow ξ is given.

4.1. **Shadowing of invariant measures in an abstract setting.** We propose an abstract definition of a shadowing invariant measure as follows. In this subsection we will always consider a measurable space $(\Omega,\mathcal{F})$, where Ω is a compact metric space and $\mathcal{F}$ the Borel σ-algebra. Let S be a homeomorphism on Ω and μ a S-invariant probability measure on $(\Omega,\mathcal{F})$. Recall that μ is a factor of a S-invariant probability measure ν on the same space $(\Omega,\mathcal{F})$, if there exist two Borel-measurable sets $\mathcal{M}_1$, $\mathcal{M}_2$ such that $\mu(\mathcal{M}_1)=1$, $\nu(\mathcal{M}_2)=1$, and a measurable map $\theta:\mathcal{M}_2\to\mathcal{M}_1$, such that $\theta\circ S|_{\mathcal{M}_2}=S\circ\theta|_{\mathcal{M}_1}$, and such that θ pulls the measure ν into μ, i.e. for any set $\mathcal{D}\in\mathcal{F}$, $\nu(\theta^{-1}(\mathcal{D}))=\mu(\mathcal{D}))$ (where we extended θ to a measurable function on the entire Ω in an arbitrary way).

Definition 4.1. Let $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$. We say that a S-invariant Borel-probability measure ν $\mathcal{G}$-shadows a S-invariant probability measure μ on $(\Omega,\mathcal{F})$, if μ is a factor of ν, and if for each $\mathcal{D}\in\mathcal{G}$, we have $\mu(\mathcal{D})=\nu(\mathcal{D})$.

We now in several lemmas show relation of the shadowing property to the support of a measure, ergodicity and entropy. To control certain topological properties of the shadowing measure, we introduce the notion of the conditional support of a probability measure μ with respect to a σ-subalgebra of Borel sets $\mathcal{G}$. We denote it by $\text{supp}(\mu|\mathcal{G})$, and define it as the set of all $x\in\Omega$ such that there exists a sequence of closed sets $\mathcal{D}_j\in\mathcal{G}$, $j\in\mathbb{N}$, $\mu(\mathcal{D}_j)>0$ such that $\cap_{j\in\mathbb{N}}\mathcal{D}_j=\{x\}$. Furthermore, let $\text{supp}^c(\mu|\mathcal{G})$ be the complement-conditional support, defined as the set of all $x\in\Omega$ for which there exists an open $\mathcal{D}\in\mathcal{G}$ such that $\mu(\mathcal{D})=0$ and $x\in\mathcal{D}$. Clearly, if $\mathcal{G}=\mathcal{F}$, we have $\text{supp}(\mu|\mathcal{F})=\text{supp}(\mu)$, and $\text{supp}^c(\mu|\mathcal{F})=\text{supp}(\mu)^c$. In general, it is easy to deduce from the definition of the support of a measure that we have

$$\text{supp}(\mu|\mathcal{G})\subseteq\text{supp}(\mu)\subseteq\text{supp}^c(\mu|\mathcal{G})^c. \tag{4.1}$$

The following Lemma follows directly from the definitions:

Lemma 4.1. *Assume that ν $\mathcal{G}$-shadows μ. Then* $\text{supp}(\nu|\mathcal{G})=\text{supp}(\mu|\mathcal{G})$ *and* $\text{supp}^c(\nu|\mathcal{G})=\text{supp}^c(\mu|\mathcal{G})$.

The relation of shadowing to ergodicity is important and somewhat more involved:

Lemma 4.2. *Assume that ν $\mathcal{G}$-shadows μ, that μ is S-ergodic, and that $\mathcal{G}$ satisfies the following: for each $\mathcal{D}\in\mathcal{G}$, $\theta^{-1}(\mathcal{M}_1\cap\mathcal{D})\subset\mathcal{D}$. Then almost every measure in the ergodic decomposition of ν $\mathcal{G}$-shadows μ.*

Proof. Consider the ergodic decomposition of ν, i.e. a Borel-probability measure χ on the compact, metrizable space of probability measures $\mathcal{M}(\Omega)$ (equipped with the weak*-topology), such that χ-a.e. measure is S-invariant and ergodic, and such that the usual representation formula for ν in terms of χ holds [47]. Then it is straightforward to check by verifying the definition of the ergodic decomposition [47] that $(\theta^*)^*\chi$ is the ergodic decomposition of μ, where $(\theta^*)^*$ is the double pull defined in a natural way. However, the ergodic decomposition is unique, and as μ is ergodic, $(\theta^*)^*\chi$ must be concetrated on μ. That means that for χ-a.e. measure $\tilde{\nu}$ (i.e. almost every measure in the ergodic decomposition of ν), we have $\theta^*(\tilde{\nu}) = \mu$. By construction, μ is then a factor of $\tilde{\nu}$.

It remains to show the shadowing property. As $\mu(\mathcal{M}_1) = 1$, we have

$$\tilde{\nu}(\mathcal{D}) \geq \tilde{\nu}(\theta^{-1}(\mathcal{M}_1 \cap \mathcal{D})) = \mu(\mathcal{M}_1 \cap \mathcal{D}) = \mu(\mathcal{D}), \tag{4.2}$$

and analogously $\tilde{\nu}(\mathcal{D}^c) \geq \mu(\mathcal{D}^c)$. However, $1 = \mu(\mathcal{D}) + \mu(\mathcal{D}^c) = \tilde{\nu}(\mathcal{D}) + \tilde{\nu}(\mathcal{D}^c) = 1$. We conclude that the equality in (4.2) must hold. □

Finally, we establish relation of shadowing to the metric (or Kolmogorov-Sinai) entropy $h_\mu(S)$ of a measure μ.

Lemma 4.3. *If ν $\mathcal{G}$-shadows μ, then $h_\nu(S) \geq h_\mu(S)$.*

Proof. This holds, as entropy is always non-increasing under factor maps and μ is a factor of ν [47]. □

4.2. **Variational construction of shadowing measures.** In this subsection $\mathcal{X}$ will always be equipped with the topology $\mathcal{X}_{\text{loc}}$. Prior to the variational construction of measures, we introduce the required spaces and projections. Recall the projections $\pi : \mathcal{X} \to \mathbb{T}^N \times \mathbb{R}^N$, given with $\pi(q) = (q \mod 2\pi, q_t)$. Let $\hat{\mathcal{X}}$ be the quotient space induced by the relation of equivalence: $q \sim \tilde{q}$ whenever there is $k \in \mathbb{Z}$, such that $q - \tilde{q} = 2k\pi$, and with the induced topology. Let $\iota : \mathcal{X} \to \hat{\mathcal{X}}$ be the canonical projection, and let $\hat{\xi}$, $\hat{\varphi}$ be the induced semi-flow ξ and flow φ on $\hat{\mathcal{X}}$. By (1.2) and by definition, $\hat{\xi}$ and $\hat{\varphi}$ are well-defined. If $\hat{\mathcal{E}} = \iota(\mathcal{E})$, $S = \varphi^{2\pi}$, $\hat{S} = \hat{\varphi}^{2\pi}$ are the 2π-shifts in the variable t, and $\hat{\pi} : \hat{\mathcal{X}} \to \mathbb{T}^N \times \mathbb{R}^N$ is defined with $\hat{\pi}(\hat{q}) = (\hat{q}, \hat{q}_t)$, then the following commutative diagrams hold:

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{i} & \hat{\mathcal{X}} \\ \downarrow{\scriptstyle \xi} & & \downarrow{\scriptstyle \hat{\xi}} \\ \mathcal{X} & \xrightarrow{i} & \hat{\mathcal{X}} \end{array} \qquad \begin{array}{ccccc} \mathcal{E} & \xrightarrow{i} & \hat{\mathcal{E}} & \xrightarrow{\hat{\pi}} & \mathbb{T}^N \times \mathbb{R}^N \\ \downarrow{\scriptstyle S} & & \downarrow{\scriptstyle \hat{S}} & & \downarrow{\scriptstyle \sigma} \\ \mathcal{E} & \xrightarrow{i} & \hat{\mathcal{E}} & \xrightarrow{\hat{\pi}} & \mathbb{T}^N \times \mathbb{R}^N \end{array}$$

By the continuous dependence on initial conditions of (1.3), $\pi|_{\mathcal{E}}$ and $\hat{\pi}|_{\hat{\mathcal{E}}}$ are continuous. As for notation, we will always denote the functions on the quotient set $\hat{\mathcal{X}}$ by $\hat{.}$. To simplify the notation, the subsets and elements of $\mathcal{X}$ and $\hat{\mathcal{X}}$ will be denoted by the same symbol, as the meaning will always be clear from the context.

We now focus on constructing ϕ-, or equivalently σ-invariant measures of (1.3) (we always implicitly assume that the measures are Borel probability measures). We denote by $\mathcal{M}(\mathcal{X})$, $\mathcal{M}(\hat{\mathcal{X}})$ and $\mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N)$ the spaces of S-, $\hat{S}$-, respectively σ-invariant measures on these spaces, equipped with the weak*-topology. Analogously we define $\mathcal{M}(\mathcal{E})$, $\mathcal{M}(\hat{\mathcal{E}})$. We always denote by $.^*$ the functions, flows and semi-flows pulled to these spaces of measures. By all the commutative relations established so far, it is straightforward to check that the objects below are well-defined, and that the following commutative diagrams hold:

$$\begin{array}{ccc} \mathcal{M}(\mathcal{X}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{X}}) \\ \downarrow{\scriptstyle \xi^*} & & \downarrow{\scriptstyle \hat{\xi}^*} \\ \mathcal{M}(\mathcal{X}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{X}}) \end{array} \qquad \begin{array}{ccccc} \mathcal{M}(\mathcal{E}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{E}}) & \xrightarrow{\hat{\pi}^*} & \mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N) \\ \downarrow{\scriptstyle S^*} & & \downarrow{\scriptstyle \hat{S}^*} & & \downarrow{\scriptstyle \sigma^*} \\ \mathcal{M}(\mathcal{E}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{E}}) & \xrightarrow{\hat{\pi}^*} & \mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N) \end{array}$$

(by definition, S^*, $\hat{S}^*$ and σ^* are identities). Thus constructing invariant measures of (1.3), i.e. elements of $\mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N)$, is equivalent to finding required objects in $\mathcal{M}(\hat{\mathcal{E}})$, i.e. fixed points of $\hat{\xi}^*$ on $\mathcal{M}(\hat{\mathcal{X}})$, or fixed points of ξ^* on $\mathcal{M}(\mathcal{X})$. The approach to constructing such measures is as follows: we will construct an element $\mu \in \mathcal{M}(\hat{\mathcal{X}})$ (typically not supported on $\hat{\mathcal{E}}$), e.g. by embedding a Bernoulli shift. We will then find an element of $\nu \in \mathcal{M}(\hat{\mathcal{E}})$ which shadows μ, as an element of the ω-limit set of μ with respect to $\hat{\xi}^*$. To achieve the shadowing property, given a fixed $\mu \in \mathcal{M}(\hat{\mathcal{X}})$, we will require that a σ-subalgebra $\mathcal{G}$ of the σ-algebra of Borel sets on $\hat{\mathcal{X}}$ satisfies the following conditions:

- (M1) **The separation property.** There exists a Borel-measurable set $\mathcal{M}_1 \subset \hat{\mathcal{X}}$ such that $\mu(\mathcal{M}_1) = 1$, and such that $\{\mathcal{D} \cap \mathcal{M}_1,\ \mathcal{D} \in \mathcal{G}\}$ generates all Borel-measurable sets on $\mathcal{M}_1$. Specifically, for each $q \in \mathcal{M}_1$, there exists $\mathcal{D}_q \in \mathcal{G}$ such that if $q, \tilde{q} \in \mathcal{M}_1$, $q \neq \tilde{q}$, then $\mathcal{D}_q \cap \mathcal{D}_{\tilde{q}} = \emptyset$. Furthermore, for any $q \in \mathcal{M}_1$, $\mathcal{D}_{\hat{S}(q)} = \hat{S}(\mathcal{D}_q)$.
- (M2) **The ξ-invariance.** For each $q \in \mathcal{M}_1$ and each $\mathcal{D} \in \mathcal{G}$, if $q \in \mathcal{D}$, then for all $s \geq 0$, $\hat{\xi}^s(q) \in \mathcal{D}$.
- (M3) **Measurability.** If $\mathcal{M}_2 = \cup_{q \in \mathcal{M}_1} \mathcal{D}_q$, then the map $\hat{\theta} : \mathcal{M}_2 \to \mathcal{M}_1$ given with $\hat{\theta}(\mathcal{D}_q) = q$ is Borel-measurable. Specifically, $\mathcal{M}_2$ is Borel-measurable.
- (M4) **The closed-sets property.** There exists a family $\mathcal{D}_i \in \mathcal{G}$ of closed sets, $i \in \mathcal{I}$, such that $\mathcal{G}$ is generated by this family (i.e. $\mathcal{G}$ is the smallest σ-algebra containing all $(\mathcal{D}_i)_{i \in \mathcal{I}}$). Furthermore, for each $i_1 \in \mathcal{I}$ there exists a sequence $i_n \in \mathcal{I}$, $n \in \mathbb{N}$ such that $\mathcal{D}_{i_n}$ are pairwise disjoint, and such that $\mu(\cup_{n=1}^{\infty} \mathcal{D}_{i_n}) = 1$.

In applications, (M1), (M3) and (M4) will follow relatively easily from the construction of μ, and the focus will be on ensuring the ξ-invariance of the constructed σ-algebra $\mathcal{G}$. An important tool in the construction of the shadowing measure, already suggested in [41], is that ξ^* and $\hat{\xi}^*$ are gradient-like semiflows with the Lyapunov function (given below for $\hat{\xi}^*$)

$$\hat{\mathcal{L}}^*(\mu) = \int_{\hat{\mathcal{X}}} \int_0^{2\pi} L(q(t), q_t(t), t) dt \, d\mu(q) \tag{4.3}$$

(see Lemma 4.5 below). We fix $\mu \in \mathcal{M}(\hat{\mathcal{X}})$ and denote by $\mu(s) = \hat{\xi}^*(\mu, s)$ for $s \geq 0$, i.e. $\mu(0) = \mu$, and $\mu(s)$ is the pulled measure μ with respect to the map $\hat{\xi}^s$.

Theorem 4.4. Variational construction of shadowing measures. *Assume $\mu \in \mathcal{M}(\hat{\mathcal{X}})$ such that $||q_t(s)||_{H^1_{ul}(\mathbb{R})^N}$ is bounded on the support of $\mu(s)$, uniformly in $s \geq 0$. Assume $\mathcal{G}$ is a σ-subalgebra of Borel sets on $\hat{\mathcal{X}}$ satisfying (M1)-(M4). Then there exists $\nu \in \mathcal{M}(\hat{\mathcal{E}})$ which $\mathcal{G}$-shadows μ.*

Furthermore, if μ is $\hat{S}$-ergodic, we can choose ν to be $\hat{S}$-ergodic.

To prove the theorem, we first construct a measure $\nu \in \mathcal{M}(\hat{\mathcal{E}})$ in two Lemmas, and then show that it is indeed the shadowing measure by using (M1)-(M4).

Lemma 4.5. *The function $s \to \hat{\mathcal{L}}^*(\mu(s))$ is strictly decreasing, unless $\mu(0) \in \mathcal{M}(\hat{\mathcal{E}})$, in which case it is constant. Furthermore,*

$$\frac{d}{ds}\hat{\mathcal{L}}^*(\mu(s)) = -\int_{\tilde{\mathcal{X}}} \int_0^{2\pi} q_s(t)^2 dt \, d\mu(s)(q). \tag{4.4}$$

Proof. We first note that by the uniform bound on $||q_t(s)||_{H^1_{\mathrm{ul}}(\mathbb{R})^N}$, we have that $\hat{\mathcal{L}}^*(\mu(s)) < \infty$ for all $s \geq 0$. By the smoothening property Theorem 2.1, (iv), for any $s > 0$ and any $q \in \operatorname{supp} \mu(s)$, such q is smooth enough so that we can differentiate as follows:

$$\begin{aligned}
\frac{d}{ds} \int_0^{2\pi} L(q, q_t, t) dt &= \int_0^{2\pi} \left(\frac{d}{ds} L(q, q_t, t) \right) dt = \int_0^{2\pi} \left(q_t q_{ts} + \frac{\partial}{\partial q} V(q,t) q_s \right) dt ds \\
&= \int_0^{2\pi} \left(-q_{tt} q_s + \frac{\partial}{\partial q} V(q,t) q_s \right) dt + q_t(2\pi) q_s(2\pi) - q_t(0) q_s(0) \\
&= -\int_0^{2\pi} q_s^2 dt + q_t(2\pi) q_s(2\pi) - q_t(0) q_s(0).
\end{aligned}$$

Now for $0 < s_0 < s_1$, we have

$$\int_0^{2\pi} L(q(s_0), q_t(s_0), t)dt - \int_0^{2\pi} L(q(s_1), q_t(s_1), t) = -\int_{s_0}^{s_1}\int_0^{2\pi} q_s(s,t)^2 dtds$$
$$+ \int_{s_0}^{s_1} \left(q_t(s,2\pi)q_s(s,2\pi) - q_t(s,0)q_s(s,0)\right) ds. \tag{4.5}$$

By the dominated convergence theorem, we can extend (4.5) also to $0 \le s_0 < s_1$. By the assumptions we have for any $q \in \operatorname{supp}\mu$,

$$\int_{\tilde{\mathcal{X}}}\int_0^{s_1}\int_0^{2\pi} |q_t(s,t)q_s(s,t)|dtdsd\mu(q) \ll \int_{\tilde{\mathcal{X}}}\int_0^{s_1}\left(\int_0^{2\pi}(q_t(s,t)^2 + q_{tt}(s,t)^2 dt\right)^{1/2} dsd\mu(q)$$
$$\ll \int_{\tilde{\mathcal{X}}}\int_0^{s_1} ||q_t(s)||_{H^1_{\mathrm{ul}}(\mathbb{R})^N} dsd\mu(q) \ll A\cdot s_1,$$

where A is the uniform bound $||q_t(s)||_{H^1_{\mathrm{ul}}(\mathbb{R})^N}$. Thus without loss of generality, we can assume that the function $\int_0^{s_1} q_t(s,0)q_s(s,0)ds$ is absolutely integrable with respect to μ (otherwise we choose some other $T \in [0,2\pi)$ instead of $T = 0$ and repeat the argument over the interval $[T, T+2\pi]$). By the $\hat{S}$-invariance of μ, we now have for any s_0, $0 \le s_0 < s_1$,

$$\int_{\tilde{\mathcal{X}}}\int_{s_0}^{s_1} q_t(s,0)q_s(s,0)dsd\mu(q) = \int_{\tilde{\mathcal{X}}}\int_{s_0}^{s_1} q_t(s,2\pi)q_s(s,2\pi)dsd\mu(q).$$

Integrating (4.5) with respect to μ, we now get for any $0 \le s_0 < s_1$,

$$\hat{\mathcal{L}}^*(\mu(s_0)) - \hat{\mathcal{L}}^*(\mu(s_1)) = -\int_{\tilde{\mathcal{X}}}\int_{s_0}^{s_1}\int_0^{2\pi} q_s(s,t)^2 dtdsd\mu(q).$$

By the Fubini theorem, we can swap integrals over ds and $d\mu$, which completes the proof. □

Lemma 4.6. *There exists $\nu \in \mathcal{M}(\hat{\mathcal{E}})$ which is a weak*-limit of a subsequence of $\mu(s)$, $s \ge 0$.*

Proof. As $\mathcal{X}$ equipped with the localized topology is not complete, to establish compactness required for the construction of ν, we need to consider its closure in $H^1_{\mathrm{loc}}(\mathbb{R})^N$, denoted by $\mathcal{Y}$. Let $\hat{\mathcal{Y}}$ be the quotient set with the same relation of equivalence $\sim$ and the induced topology, and $\hat{\mathcal{X}} \hookrightarrow \hat{\mathcal{Y}}$ the natural embedding. It is straightforward to check that the closure of the set of all q satisfying $||q_t||_{H^1(\mathbb{R})} \le A$ is compact in $\hat{\mathcal{Y}}$ (we choose representatives in $\mathcal{Y}$ such that $q(0) \in [0,2\pi]^N$ and find a convergent subsequence by diagonalization). Thus by the Banach-Alaoglu theorem, $\mu(s)$, $s \ge 0$ has a convergent subsequence $\mu(s_n)$ which converges to some measure ν on $\hat{\mathcal{Y}}$ in the weak* topology induced by the induced $H^1_{\mathrm{loc}}(\mathbb{R})^N$ topology.

It suffices to show that $\nu \in \mathcal{M}(\hat{\mathcal{E}})$. Choose $h \in H^1(\mathbb{R})^N$ with compact support, say in $[-2n\pi, 2n\pi]$, $n \in \mathbb{N}$. Then by the $\hat{S}$-invariance of μ in the third row below, we obtain

$$\int_{\tilde{\mathcal{X}}} |\partial L(q,q_t,t)h|d\mu(s_n)(q) = \int_{\tilde{\mathcal{X}}}\left|\int_{-2n\pi}^{2n\pi}\left(q_t(s_n,t)h_t(t) + DV(q(s_n,t),t)h(t)\right)dt\right| d\mu(q)$$
$$\le ||h||_{L^2(\mathbb{R})^N}\left(\int_{\tilde{\mathcal{X}}}\int_{-2n\pi}^{2n\pi} q_s^2(s_n,t)dtd\mu(q)\right)^{1/2}$$
$$= (2n+1)^{-1/2}||h||_{L^2(\mathbb{R})^N}\left(\int_{\tilde{\mathcal{X}}}\int_0^{2\pi} q_s^2(s_n,t)dtd\mu(q)\right)^{1/2},$$

which by Lemma 4.5 converges to zero. We thus have that for any $h \in H^1(\mathbb{R})^N$ with compact support, $\int_{\tilde{\mathcal{Y}}} |\partial L(q,q_t,t)h|d\nu(q) = 0$. By choosing such a countable, dense set of h, and by continuity, we conclude that ν is supported on the solutions of (1.3). By construction, ν is supported on q such that $q_t \in L^2_{\mathrm{ul}}(\mathbb{R})^N$, thus by Lemma 15.5, $\nu \in \mathcal{M}(\hat{\mathcal{E}})$. □

Proof of Theorem 4.4. Take ν constructed in Lemma 4.6. We apply results from subsection 4.1 with $\Omega = \mathcal{Y}$, $\mathcal{Y}$ as in Lemma 4.6, thus compact and metrizable. As ν is a weak*-limit of $\hat{S}$-invariant measures and $\hat{S}$ is continuous, ν is $\hat{S}$-invariant. It suffices to show that ν shadows μ. We take as the factor the function $\hat{\theta} : \mathcal{M}_2 \to \mathcal{M}_1$ from the property (M3). Let $s_k > 0$ be the sequence from Lemma 4.6 such that ν is the weak* limit of $\mu(s_k)$.

First we show that for each $\mathcal{D} \in \mathcal{G}$, $\nu(\mathcal{D}) = \mu(\mathcal{D})$. It suffices to show it for the generator $\mathcal{D}_i$, $i \in \mathcal{I}$, from (M4). Choose $i_1 \in \mathcal{I}$, and find $i_n \in \mathcal{I}$, $n \in \mathbb{N}$ so that (M4) holds. Now by (M2) and (M3), we have $\mu(\mathcal{D}_{i_n}) = \mu(\mathcal{D}_{i_n} \cap \mathcal{M}_1) \leq \mu(s_k)(\mathcal{D}_{i_n})$ for all $n \in \mathbb{N}$. However, by (M4) and σ-aditivity of μ and $\mu(s_n)$, we get

$$1 = \sum_{n=1}^{\infty} \mu(\mathcal{D}_{i_n}) \leq \sum_{n=1}^{\infty} \mu(s_k)(\mathcal{D}_{i_n}) = \mu(s_k)(\cup_{n=1}^{\infty} \mathcal{D}_{i_n}) \leq 1,$$

thus equality must hold in all the terms. As ν is the weak*-limit of $\mu(s_k)$ and $\mathcal{D}_{i_n}$ are closed, we have that for all $n \in \mathbb{N}$, $\nu(\mathcal{D}_{i_n}) \geq \limsup_{k \to \infty} \mu(s_k)(\mathcal{D}_{i_n}) = \mu(\mathcal{D}_{i_n})$. As $\mathcal{D}_{i_n}$, we analogously as above have

$$1 = \sum_{n=1}^{\infty} \mu(\mathcal{D}_{i_n}) \leq \sum_{n=1}^{\infty} \nu(\mathcal{D}_{i_n}) = \nu(\cup_{n=1}^{\infty} \mathcal{D}_{i_n}) \leq 1,$$

thus again equality must hold in all the terms.

By (M1) and the definition of $\hat{\theta}$, $\hat{\theta}$ and $\hat{S}$ commute. As $\nu(\mathcal{D}) = \mu(\mathcal{D}) = \mu(\mathcal{D} \cap \mathcal{M}_1)$, and $\mathcal{D} \cap \mathcal{M}_1$ generate all Borel-measurable sets on $\mathcal{M}_1$, to show that $\hat{\theta}$ is measure-preserving, it suffices to show that for all $i \in \mathcal{I}$,

$$\nu(\mathcal{D}_i) = \nu(\hat{\theta}^{-1}(\mathcal{D}_i \cap \mathcal{M}_1)). \tag{4.6}$$

Choose $i_1 \in \mathcal{I}$, and find a sequence $i_n \in \mathcal{I}$ so that (M4) holds. By definition of $\hat{\theta}$ in (M3), we have that $\mathcal{D}_{i_n} \subset \hat{\theta}^{-1}(\mathcal{D}_{i_n} \cap \mathcal{M}_1)$, thus $\nu(\mathcal{D}_{i_n}) \leq \nu(\hat{\theta}^{-1}(\mathcal{D}_{i_n} \cap \mathcal{M}_1))$. By (M1) and (M3), the sets $\hat{\theta}^{-1}(\mathcal{D}_{i_n} \cap \mathcal{M}_1)$, $n \in \mathbb{N}$ are pairwise disjoint. Analogously as above, from all of this and $\sum_{n=1}^{\infty} \nu(\mathcal{D}_{i_n}) = 1$ we conclude that (4.6) must hold.

If μ is $\hat{S}$-ergodic, we can find a $\hat{S}$-ergodic ν by Lemma 4.2. □

II: INVARIANT SETS IN THE A-PRIORI UNSTABLE CASE

5. The homoclinic orbits

As of this section, we focus on the a-priori unstable case with the Lagrangian (1.1) and $N = 2$. In this section we recall the key properties of the Peierl's barrier function and stable and unstable manifolds of the invariant tori $\mathbb{T}_\omega$. The results of this section are standard (see [17, 44] and references therein). As we were unable to find in the literature the a-priori bounds we require later, we give self-contained proofs.

For a fixed $\omega \in \mathbb{R}$, let $S_\omega^-, S_\omega^+ : \mathbb{R}^2 \to \infty$ be the Peierl's barrier functions defined with

$$S_\omega^-(t_0, v_0) = \inf \left\{ \int_{-\infty}^{t_0} L_\omega(q(t), q_t(t), t) dt,\ q = (u, v) \in H^1_{\text{loc}}((-\infty, t_0])^2, q(t_0) = (\pi, v_0), \lim_{t \to -\infty} u(t) = 0 \right\},$$

$$S_\omega^+(t_0, v_0) = \inf \left\{ \int_{t_0}^{\infty} L_\omega(q(t), q_t(t), t) dt,\ q = (u, v) \in H^1_{\text{loc}}([t_0, \infty))^2, q(t_0) = (\pi, v_0), \lim_{t \to \infty} u(t) = 2\pi \right\}.$$

The functions for which the minima $S_\omega^-(t_0, v_0)$, $S_\omega^+(t_0, v_0)$ are attained are the solutions of (1.3) and lie on unstable, respectively stable manifolds of $\mathbb{T}_\omega$ (see Proposition 5.4, (i) below). We call them one-sided (left-, respectively right-hand) sided minimizers at (ω, t_0, v_0).

We first obtain a-priori bounds on S_ω^-, S_ω^+. We then introduce the notion and construct specific super- and sub-solutions of (1.2) required in this section and later, and finally construct one-sided minimizers and prove explicit a-priori bounds. In particular, we prove that there exists an absolute constant $c_2 > 0$ so that, if $q^- = (u^-, v^-) : (-\infty, t_0] \to \mathbb{R}^2$, $q^+ = (u^+, v^+) : [t_0, \infty) \to \mathbb{R}^2$ are any one-sided minimizers, then

$$|u^-(t)| \leq c_2 e^{-\frac{1}{2}\sqrt{\varepsilon}|t - t_0|}, \text{ for all } t \leq t_0, \qquad |u^+(t) - 2\pi| \leq c_2 e^{-\frac{1}{2}\sqrt{\varepsilon}|t - t_0|}, \text{ for all } t \geq t_0. \tag{5.1}$$

Lastly, we prove continuity of S^-, S^+ in ω, t_0, v_0 and estimate the Lipschitz constant in ω.

Lemma 5.1. *For all* $\omega, t_0, v_0 \in \mathbb{R}$*, we have*

$$4\sqrt{\varepsilon(1-\mu)} \le S^-_\omega(t_0, v_0), S^+_\omega(t_0, v_0) \le 4\sqrt{\varepsilon(1+\mu)}. \tag{5.2}$$

Proof. By definition,

$$\int_{t_0}^{\infty} L_\omega(q, q_t, t)dt \le \int_{t_0}^{\infty} \left(\frac{1}{2}u_t^2 + \frac{1}{2}(v_t - c)^2 + (\varepsilon(1+\mu))(1-\cos u(t))\right) dt. \tag{5.3}$$

It is well-known [2, 11] that the right-hand side of (5.3) attains minimum for the separatrix solution of the pendulum equation $u^0(t) = 4\operatorname{arctg} e^{\sqrt{\varepsilon(1+\mu)}\,(t-t_0)}$, $v^0(t) = \omega(t-t_0) + v_0$, and that the value of the integral on the right-hand side of (5.3) is then by direct calculation $4\sqrt{\varepsilon(1+\mu)}$. Analogously we deduce that for any $q \in H^1_{\text{loc}}((t_0, \infty))^2$,

$$4\sqrt{\varepsilon(1-\mu)} \le \int_{t_0}^{\infty} \left(\frac{1}{2}u_t^2 + \frac{1}{2}(v_t - c)^2 + (\varepsilon(1-\mu))(1-\cos u(t))\right) dt \le \int_{t_0}^{\infty} L_\omega(q, q_t, t)dt,$$

which completes the proof for $S^+_\omega(t_0, v_0)$. The bounds for $S^-_\omega(t_0, v_0)$ are analogous. □

In order to obtain a-priori bounds on one-sided minimizers, we require the notion of super-, respectively sub-solutions of (1.2a) or (1.2b). We say that $q = (u, v)$ is a super-solution of (1.2a) on $U = (t_0, t_1) \times (s_0, s_1]$, where $-\infty \le t_0 < t_1 \le \infty$, $-\infty \le s_0 < s_1 \le \infty$, if it is continuous on $\bar{U}$ and for any $(s, t) \in U$, $u_s - u_{tt} + \partial_u V(u, v, t) \le 0$. Analogously we say that $q = (u, v)$ is a sub-solution of (1.2a) on U, if $u_s - u_{tt} + \partial_u V(u, v, t) \ge 0$. We say that q is a strict super-, respectively sub-solution, if strict inequalities hold.

We say that a $z : \mathbb{R} \to \mathbb{R}$ is a stationary super-solution on $I = (t_0, t_1)$, $-\infty \le t_0 < t_1 \le \infty$, if it is continuous on $[t_0, t_1]$ and C^2 on (t_0, t_1), and such that for any $v \in C^2(\mathbb{R}^2)$, and for any $t \in (t_0, t_1)$,

$$z_{tt} - \partial_u V(z, v, t) \le 0.$$

We see that then for any $v \in C^2(\mathbb{R})$, (z, v) is a super-solution of (1.2a) on $(t_0, t_1) \times \mathbb{R}$, where z is considered as a fixed function in s. Analogously we define the notions of strict stationary super-solutions, sub-solutions, and analogous notions for (1.2b).

The following Lemma is a special case of the parabolic maximum principle [16].

Lemma 5.2. *Assume z is a strict stationary super-solution of (1.2a) on (t_0, t_1), $-\infty \le t_0 < t_1 \le \infty$, assume $q = (u, v) \in \mathcal{E}$, and let $u(t) \le z(t)$ for all $t \in (t_0, t_1)$. Then for all $t \in (t_0, t_1)$, $u(t) < z(t)$.*

Analogous statements hold for strict stationary sub-solutions.

Proof. Assume the contrary and find $t_2 \in (t_0, t_1)$ such that $u(t_2) = z(t_2)$. Direct calculation yields that

$$u_{tt}(t_2) - \partial_u V(u(t_2), v(t_2), t_2) \le z_{tt}(t_2) - \partial_u V(z(t_2), v(t_2), t_2) < 0,$$

which is in contradiction to $q \in \mathcal{E}$. □

Lemma 5.3. *There exist $z^- : (-\infty, 3/(4\sqrt{\varepsilon})] \to \mathbb{R}$ and $z^+ : [-3/(4\sqrt{\varepsilon}), \infty) \to \mathbb{R}$, depending only on ε, μ, satisfying for all t in the domain of definition:*

(i) $0 < z^-(t) < 3\pi/2$, $\pi/2 < z^+(t) < 2\pi$, both are continuous and C^2 in the interior of the domain,

(ii) $z^-(0) = z^+(0) = \pi$,

(iii) z^- is a strict stationary super-solution on $(-\infty, 3/(4\sqrt{\varepsilon}))$ of (1.2a), and z^+ is a strict stationary sub-solution (1.2a) on $(-3/(4\sqrt{\varepsilon}), \infty)$. Furthermore, for any constant $T \ge 0$, $z^-(t+T)$ and $z^+(t-T)$ are strict stationary super-, resp. sub-solutions in the interior of their domain of definition.

(iv) z^-, z^+ are strictly increasing and we have

$$\sqrt{\varepsilon}/2 < z^-_t(t), z^+_t(t) \quad \text{for all } t \in (-1/(4\sqrt{\varepsilon}), 1/(4\sqrt{\varepsilon})). \tag{5.4}$$

(v) For all $t \in [1/(4\sqrt{\varepsilon}), 3/(4\sqrt{\varepsilon})]$, $z^-(t) \le \pi + 1/4$ and $z^+(-t) \ge \pi - 1/4$,

(vi) There exists an absolute constant $c_2 > 0$ such that for all t in the domains of definition,

$$|z^-(t)| \le c_2 e^{-\frac{1}{2}\sqrt{\varepsilon}|t|}, \qquad |z^+(t) - 2\pi| \le c_2 e^{-\frac{1}{2}\sqrt{\varepsilon}|t|}. \tag{5.5}$$

(vii) There exists an absolute constant $c_3 > 0$ such that

$$|z^-(t) - u^{(\varepsilon)}(t)| \le c_3\sqrt{\varepsilon\mu},\ t \le 0, \qquad |z^+(t) - u^{(\varepsilon)}(t)| \le c_3\sqrt{\varepsilon\mu},\ t \ge 0, \tag{5.6}$$

where $u^{(\varepsilon)}(t) = 4\,\mathrm{arctg}\, e^{\sqrt{\varepsilon}t}$ is the separatrix solution in the case $\mu = 0$.

An explicit construction of z^-, z^+ and the proof of Lemma 5.3 is given in the Appendix B.

Proposition 5.4. *Let $(\omega, t_0, v_0) \in \mathbb{R}^3$. Then there exist one-sided minimizers $q^- : (\infty, t_0] \to \mathbb{R}$, $q^+ : [t_0, \infty) \to \mathbb{R}$, for which $S^-_\omega(t_0, v_0)$, $S^+_\omega(t_0, v_0)$ attain their minimal value.*

Furthermore, any such one-sided minimizers $q^- = (u^-, v^-)$, $q^+ = (u^+, v^+)$ at (ω, t_0, v_0) satisfy the following:

(i) They are C^4 and solutions of the Euler-Lagrange equations on $(-\infty, t_0)$, (t_0, ∞) respectively,

(ii) For all $0 \le T \le 3/(4\sqrt{\varepsilon})$,

$$0 < u^-(t) \le z^-(t - t_0 + T) \qquad \text{for all } t \le t_0, \tag{5.7}$$

$$z^+(t - t_0 - T) \le u^+(t) < 2\pi \qquad \text{for all } t \ge t_0. \tag{5.8}$$

The proof is in the Appendix B. (Existence and (i) are a consequence of the Tonelli theorem [33, Appendix 1], and the a-priori bounds follow from Lemma 5.2 applied to z^-, z^+ constructed in Lemma 5.3.)

Combining (5.5), (5.7) and (5.8) with $T = 0$ we get:

Corollary 5.5. *Any one-sided minimizers q^-, q^+ at (ω, t_0, v_0) satisfy (5.1).*

We finally deduce the Lipschitz constant for S^- and S^+ in the variable ω.

Corollary 5.6. *The functions S^-, S^+ are continuous in t_0, v_0, ω. Furthermore, there exists an absolute constant $c_4 \ge 1$ such that for any $(t_0, v_0) \in \mathbb{R}^2$,*

$$|S^-_\omega(t_0, v_0) - S^-_{\tilde\omega}(t_0, v_0)| \le c_4\mu|\tilde\omega - \omega|, \qquad |S^+_\omega(t_0, v_0) - S^+_{\tilde\omega}(t_0, v_0)| \le c_4\mu|\tilde\omega - \omega|. \tag{5.9}$$

Proof. We fix first (t_0, v_0) and show that the Lipschitz constant of $S^+_\omega(t_0, v_0)$ in ω. Choose $\omega, \tilde\omega \in \mathbb{R}$, and let $q = (u, v)$ be a right-hand sided minimizer constructed in Proposition 5.4 at (ω, t_0, v_0) respectively. Define

$$\tilde q(t) = (\tilde u(t), \tilde v(t)) := (u(t), v(t) + (\tilde\omega - \omega)(t - t_0)),$$

defined for $t \in [t_0, \infty)$. By definition of $q, \tilde q$, by applying $1 - \cos u \le (u - 2\pi)^2/2$, (5.1) and the standing assumption (A2), we get

$$\begin{aligned}
S^+_{\tilde\omega}(t_0, v_0) &\le \int_{t_0}^{\infty} L_{\tilde\omega}(\tilde q(t), \tilde q_t(t), t)dt = \int_{t_0}^{\infty} \left(\frac{1}{2}u_t^2 + \frac{1}{2}(v_t - \omega)^2 + V(\tilde u, \tilde v, t)\right) dt \\
&= S^+_\omega(t_0, v_0) + \int_{t_0}^{\infty} (V(\tilde u, \tilde v, t) - V(u, v, t))dt \\
&\le S^+_\omega(t_0, v_0) + \varepsilon\mu \int_{t_0}^{\infty} \{1 - \cos(u(t))\} \left\{ |\sup_{a \in [v(t), \tilde v(t)]} f_v(u(t), a, t)||\tilde\omega - \omega|(t - t_0) \right\} dt \\
&\le S^+_\omega(t_0, v_0) + c_4\varepsilon\mu \int_{t_0}^{\infty} e^{-\sqrt{\varepsilon}(t - t_0)}|\tilde\omega - \omega|(t - t_0)dt \le S^+_\omega(t_0, v_0) + c_4\mu|\tilde\omega - \omega|,
\end{aligned}$$

where c_4 (chosen to be ≥ 1) is an absolute constant. The other inequalities in (5.9) are proved analogously. Continuity in t_0, v_0 is follows similarly from the definitions of S^-, S^+. ☐

6. The heteroclinic orbits and the region of instability

We discuss first the notion of a region of instability defined in the Introduction. We then recall the fact that, if $\omega, \tilde\omega$ are sufficiently close and in the same region of instability, then there exists a heteroclinic orbit connecting the tori $\mathbb{T}_\omega$ and $\mathbb{T}_{\tilde\omega}$. We also establish a-priori bounds on heteroclinic orbits, and show that the set of heteroclinic orbits is compact in $H^2_{\mathrm{loc}}(\mathbb{R})^2$. As we were unable to find in the literature the a-priori bounds we need later, we give self-contained proofs.

We can write the function S_ω defined in the Introduction as $S_\omega(t, v) = S^-_\omega(t, v) + S^+_\omega(t, v)$.

Definition 6.1. The region of instability is a connected component of the set of non-degenerate $\omega \in \mathbb{R}$, where ω is non-degenerate if every connected component of the set of global minima of S_ω in $\mathbb{R}^2$ is bounded.

One can easily show as a consequence of Corollary 5.6 that a region of instability is open. We do not require it here, as we take (S1) as the standing assumption. Thus for every global minimum (t_0, v_0) of S_ω we can find a closed, bounded set $\mathcal{N}(t_0, v_0) \subset \mathbb{R}^2$, $(t_0, v_0) \in \mathcal{N}(t_0, v_0)$, such that for each $(t_1, v_1) \in \partial\mathcal{N}(t_0, v_0)$, (1.4) holds, and such that there is a constant $R \geq \sup\{|x - y|,\ x, y \in \mathcal{N}(t_0, v_0)\}$ satisfying (1.5).

Remark 6.1. By continuity and periodicity of S_ω, if $[\omega^-, \omega^+]$ is a segment in a region of instability, we can always find $\Delta_0 > 0$, $R > 0$ satisfying (1.4), uniform over $\omega \in [\omega^-, \omega^+]$, and uniform over (t_0, v_0) which are global minimizers of S_ω. The proof is analogous to the argument used in Lemma 6.6, and omitted as not needed in the following.

Let $[\omega^-, \omega^+]$ be an interval in the same region of instability satisfying (S1), and let $\varpi = \max\{\omega^-, \omega^+, 1\}$. We fix Δ_0, R associated to $[\omega^-, \omega^+]$ from now on. We define the action $\mathcal{L}_{\omega,\tilde{\omega}} : H^1_{\text{loc}}(\mathbb{R})^2 \to \mathbb{R} \cup \{\infty\}$ and the minimal action $\Sigma_{\omega,\tilde{\omega}} : \mathbb{R}^2 \to \mathbb{R}$ along a trajectory of the heteroclinic orbit as:

$$\mathcal{L}_{\omega,\tilde{\omega}}(q) = \int_{-\infty}^{0} L_\omega(q, q_t, t)dt + \int_{0}^{\infty} L_{\tilde{\omega}}(q, q_t, t)dt + (\tilde{\omega} - \omega)v(0), \tag{6.1}$$

$$\Sigma_{\omega,\tilde{\omega}}(t_0, v_0) := S^-_\omega(t_0, v_0) + S^+_{\tilde{\omega}}(t_0, v_0) + (\tilde{\omega} - \omega)v_0 + \frac{1}{2}(\omega^2 - \tilde{\omega}^2)t_0. \tag{6.2}$$

The main result of the section is:

Proposition 6.1. *Assume that* $\omega, \tilde{\omega} \in [\omega^-, \omega^+]$ *satisfy*

$$|\omega - \tilde{\omega}| \leq \frac{\Delta_0}{4c_4(R \vee \mu) \cdot \varpi}. \tag{6.3}$$

(i) There exist $q = (u, v) \in \mathcal{E}$ *and* $(t_0, v_0) \in [0, 2\pi)^2$ *such that* $q(t_0) = (\pi, v_0)$*, and such that* $q|_{t \leq t_0}$ *and* $q|_{t \geq t_0}$ *are one-sided minimizers at* (ω, t_0, v_0)*, respectively* $(\tilde{\omega}, t_0, v_0)$*.*

(ii) Furthermore, there exists a closed, bounded set $\mathcal{N}_q \subset \mathbb{R}^2$ *containing* (t_0, v_0)*, of radius at most* R*, such that for any* $(t_1, v_1) \in \partial\mathcal{N}_q$*,*

$$\Sigma_{\omega,\tilde{\omega}}(t_1, v_1) - \Sigma_{\omega,\tilde{\omega}}(t_0, v_0) \geq 2\Delta_0 > 0. \tag{6.4}$$

We denote by $\mathcal{H}$ the set of all $q \in \mathcal{E}$ satisfying (i), (ii) in Proposition 6.1 for some $\omega, \tilde{\omega} \in [\omega^-, \omega^+]$ satisfying (6.3). We say that such $q \in \mathcal{H}$ is a heteroclinic minimizer connecting $\omega, \tilde{\omega}$. Within this section, denote by $(t_\omega, v_\omega) \in [0, 2\pi)^2$ a minimizer of S_ω, fixed if non-unique, and by $\mathcal{N}_\omega = \mathcal{N}(t_\omega, v_\omega)$.

Lemma 6.2. *If (6.3) holds, then* $\Sigma_{\omega,\tilde{\omega}}$ *attains a local minimum* (t_0, v_0) *in the interior of* $\mathcal{N}_\omega$*, such that for any* $(t_1, v_1) \in \partial\mathcal{N}_\omega$*, (6.4) holds.*

Proof. Note first that $\Sigma_{\omega,\tilde{\omega}}$ is continuous by the definition and Corollary 5.6. Thus by the definition and compactness of $\mathcal{N}_\omega$, it suffices to show that for some (t_2, v_2) in the interior of $\mathcal{N}_\omega$, and any $(t_1, v_1) \in \partial\mathcal{N}_\omega$, $\Sigma_{\omega,\tilde{\omega}}(t_1, v_1) - \Sigma_{\omega,\tilde{\omega}}(t_2, v_2) \geq 2\Delta_0 > 0$. Let $(t_2, v_2) = (t_\omega, v_\omega)$. Then by definition, because of $|t_1 - t_\omega| \leq R$, $|v_1 - v_\omega| \leq R$, $c_4 \geq 1$ and (6.3) we obtain

$$\begin{aligned}\Sigma_{\omega,\tilde{\omega}}(t_1, v_1) - \Sigma_{\omega,\tilde{\omega}}(t_\omega, v_\omega) &\geq S^-_\omega(t_1, v_1) - S^-_\omega(t_\omega, v_\omega) + S^+_{\tilde{\omega}}(t_1, v_1) - S^+_{\tilde{\omega}}(t_\omega, v_\omega) \\ &\quad - |\tilde{\omega} - \omega||v_1 - v_\omega| - \varpi|\tilde{\omega} - \omega||t_1 - t_\omega| \\ &\geq S^-_\omega(t_1, v_1) - S^-_\omega(t_\omega, v_\omega) + S^+_{\tilde{\omega}}(t_1, v_1) - S^+_{\tilde{\omega}}(t_\omega, v_\omega) - \Delta_0/2. \end{aligned} \tag{6.5}$$

From (5.9) and $\varpi \geq 1$ we deduce that

$$S^+_{\tilde{\omega}}(t_1, v_1) - S^+_{\tilde{\omega}}(t_\omega, v_\omega) \geq S^+_\omega(t_1, v_1) - S^+_\omega(t_\omega, v_\omega) - \Delta_0/2,$$

which combined with (6.5) and (1.4) gives

$$\Sigma_{\omega,\tilde{\omega}}(t_1, v_1) - \Sigma_{\omega,\tilde{\omega}}(t_\omega, v_\omega) \geq S_\omega(t_1, v_1) - S_\omega(t_\omega, v_\omega) - \Delta_0 \geq 2\Delta_0.$$

□

Lemma 6.3. *Assume (t_0, v_0) is a local minimum of $\Sigma_{\omega,\tilde{\omega}}$, and let*

$$q(t) = \begin{cases} q^-(t) & t \le t_0, \\ q^+(t) & t \ge t_0, \end{cases} \tag{6.6}$$

where q^- is the left-hand sided minimizer at (ω, t_0, v_0), and q^+ the right-sided minimizers at $(\tilde{\omega}, t_0, v_0)$. We then have that q is a solution of (1.3).

Proof. Let q and (t_0, v_0) be as in the statement of the Lemma. We first show that, if $\tilde{q} = (\tilde{u}, \tilde{v}) \in H^1(\mathbb{R})^2$ such that for some $(t_1, v_1) \in \mathbb{R}^2$, $\tilde{q}(t_1) = (\pi, v_1)$, then

$$\mathcal{L}_{\omega,\tilde{\omega}}(\tilde{q}) - \mathcal{L}_{\omega,\tilde{\omega}}(q) \ge \Sigma_{\omega,\tilde{\omega}}(t_1, v_1) - \Sigma_{\omega,\tilde{\omega}}(t_0, v_0). \tag{6.7}$$

Indeed, by definitions and the partial integration,

$$\begin{aligned} \mathcal{L}_{\omega,\tilde{\omega}}(\tilde{q}) &= \int_{-\infty}^{t_1} L_\omega(\tilde{q}(t), \tilde{q}_t(t), t)dt + \int_{t_1}^{\infty} L_{\tilde{\omega}}(\tilde{q}(t), \tilde{q}_t(t), t)dt + (\tilde{\omega} - \omega)v_1 + \frac{1}{2}(\omega^2 - \tilde{\omega}^2)t_1 \\ &\ge S^-_\omega(t_1, v_1) + S^+_{\tilde{\omega}}(t_1, v_1) + (\tilde{\omega} - \omega)v_1 + \frac{1}{2}(\omega^2 - \tilde{\omega}^2)t_1 = \Sigma_{\omega,\tilde{\omega}}(t_1, v_1). \end{aligned}$$

By the definition of q, we obtain an equality in an analogous calculation for q, thus $\mathcal{L}_{\omega,\tilde{\omega}}(q) = \Sigma_{\omega,\omega*}(t_0, v_0)$. This gives (6.7).

We now claim that for any $h \in H^1(\mathbb{R})^2$, $\partial\mathcal{L}(q)h = \lim_{\delta\to 0} \frac{1}{\delta}(\mathcal{L}(q + \delta h) - \mathcal{L}(q))$ is equal to 0. It suffices to show that for any $h = (u^h, v^h) \in H^1(\mathbb{R})^2$ and sufficiently small $\delta > 0$, $\mathcal{L}(q + \delta h) \ge \mathcal{L}(q)$, and that $\partial\mathcal{L}(q)h$ exists. Consider an open neighborhood U of (t_0, v_0) in $\mathbb{R}^2$ such that $\Sigma_{\omega,\tilde{\omega}}|_U \ge \Sigma_{\omega,\tilde{\omega}}(t_0, v_0)$.

We show first that for any $h \in H^1(\mathbb{R})^2$ we can find $\delta_0 > 0$ small enough, such that for any $0 \le \delta \le \delta_0$, there exists $(t_1, v_1) \in U$ such that $(q + \delta h)(t_1) = (\pi, v_1)$. Let $\tilde{q} = (\tilde{u}, \tilde{v}) = q + \delta h$. Indeed, we can find t_1 sufficiently close to t_0 such that $\tilde{u}(t_1) = \pi$ for δ small enough, because of (5.7), (5.8) with $T = 0$ and the fact that z^-, z^+ are strictly increasing at $t = 0$. We find $v_1 = \tilde{v}(t_1)$ sufficiently close to v_0 by finding δ_0 small enough so that all the terms in

$$|\tilde{v}(t_1) - v_0| \le \int_{t_0}^{t_1} |\tilde{v}_t|dt + |\tilde{v}(t_0) - v_0| \le \sqrt{2}|t_1 - t_0|^{1/2} \left(\int_{t_0}^{t_1} (v_t^2(t) + \delta^2 (v^h)_t^2)dt \right)^{1/2} + \delta|v^h(t_0)|$$

are small enough. Now combining it with (6.7) and the fact that the right-hand side in (6.7) is ≥ 0 on U, we obtain $\mathcal{L}(q + \delta h) \ge \mathcal{L}(q)$ for $\delta \le \delta_0$. Now, a straightforward calculation and the fact from Proposition 5.4 that q solves (1.3) for all t except perhaps $t = t_0$ yields

$$\partial L(q)h = (u_t^-(t_0) - u_t^+(t_0), v_t^-(t_0) - v_t^+(t_0)) \cdot h(t_0),$$

and $\partial L(q)h$ exists, thus $\partial L(q)h = 0$. As h is arbitrary, we get that q is C^1 at t_0. By the uniqueness of solutions of the Euler-Lagrange equations, q is a solution of (1.3) also at $t = t_1$. □

Proposition 6.1 now follows from Lemmas 6.2 and 6.3, with $\mathcal{N}_q := \mathcal{N}_\omega$.

Lemma 6.4. *If $q \in \mathcal{H}$, then there exists a unique $(t_0, v_0) \in [0, 2\pi]$ such that $q(t_0) = (\pi, v_0)$.*

Proof. Existence follows from the definition of $\mathcal{H}$, and uniqueness from (5.8) and the properties of z^-, z^+ proved in Lemma 5.3. □

Lemma 6.5. *Let $q = (u, v) \in \mathcal{H}$ connecting $\omega, \tilde{\omega}$, such that $q(t_0) = (\pi, v_0)$. Then there exists an absolute constant $c_5 > 0$ such that for all $t \in \mathbb{R}$,*

$$|u(t) - 2\pi \mathbf{1}_{[t_0,\infty)}(t)| \le c_5 e^{-\frac{1}{2}\sqrt{\varepsilon}|t - t_0|}, \tag{6.8}$$

$$|u_t(t)| \le c_5 \sqrt{\varepsilon} e^{-\frac{1}{2}\sqrt{\varepsilon}|t - t_0|}, \tag{6.9}$$

$$|v(t) - v_0 - \omega(t - t_0)| \le c_5\mu, \qquad |v_t(t) - \omega| \le c_5\sqrt{\varepsilon}\mu\, e^{-\sqrt{\varepsilon}|t - t_0|}, \qquad t \le t_0, \tag{6.10}$$

$$|v(t) - v_0 - \tilde{\omega}(t - t_0)| \le c_5\mu, \qquad |v_t(t) - \tilde{\omega}| \le c_5\sqrt{\varepsilon}\mu\, e^{-\sqrt{\varepsilon}|t - t_0|}, \qquad t \ge t_0, \tag{6.11}$$

$$|u_{tt}(t)| \leq c_5\varepsilon\, e^{-\frac{1}{2}\sqrt{\varepsilon}|t-t_0|}, \qquad |v_{tt}(t)| \leq c_5\varepsilon\mu\, e^{-\sqrt{\varepsilon}|t-t_0|}, \tag{6.12}$$

$$|u_{ttt}(t)| \leq c_5\varepsilon\varpi\, e^{-\frac{1}{2}\sqrt{\varepsilon}|t-t_0|}, \qquad |v_{ttt}(t)| \leq c_5\varepsilon\mu\varpi\, e^{-\frac{1}{2}\sqrt{\varepsilon}|t-t_0|}. \tag{6.13}$$

Proof. The absolute constant c_5 may change from line to line in the proof. The relation (6.8) follows from the definition of q and (5.1). Now, by using $\sin x \leq |x - 2k\pi|$ for $k = 0, 1$, the fact that u is a solution of (1.3), and finally using (6.8), we see that

$$|u_{tt}(t)| \leq \varepsilon(1 - \cos u(t) + |\sin u(t)|) \ll \varepsilon\, e^{-\frac{1}{2}\sqrt{\varepsilon}|t-t_0|},$$

which is the left-hand side of (6.12). By integrating it over $[t, \infty)$ for $t \geq t_0$, alternatively over $(-\infty, t]$ for $t \leq t_0$, and using $\lim_{|t|\to\infty} u_t = 0$, we get (6.9). Analogously, as $v(t)$ is a solution of (1.3), by using $\cos x \leq (x - 2k\pi)^2/2$ for $k = 0, 1$ and (6.8), we obtain

$$|v_{tt}(t)| \leq \varepsilon\mu(1 - \cos u(t)) \ll \varepsilon\mu\, e^{-\sqrt{\varepsilon}|t-t_0|},$$

which is the right-hand side of of (6.12). As $\lim_{t\to-\infty} v_t(t) = \omega$ and $\lim_{t\to\infty} v_t(t) = \tilde{\omega}$, by integrating it over $(-\infty, t]$ for $t \leq t_0$, respectively over $[t, \infty)$ for $t \geq t_0$, we obtain the right-hand sides of (6.10) and (6.11). We use $v(t_0) = v_0$, integrate the right-hand side of (6.10) over $[t, t_0]$, respectively the right-hand side of (6.11) over $[t_0, t]$, and obtain the left-hand sides of (6.10) and (6.11). Finally, to bound the third derivatives, by careful differentiation, while using uniform bounds on f and its derivatives, and as $\mu \leq 1$ and $\varpi \geq 1$, we obtain

$$\begin{aligned}|u_{ttt}| = |D_t V_u(u(t), v(t), t)| &\ll \varepsilon\mu|u_t| + \varepsilon|u_t| + \varepsilon\mu(1 - \cos u + |\sin u|)|(v_t| + \varepsilon\mu(1 - \cos u + |\sin u|) \\ &\ll \varepsilon|u_t| + \varepsilon|v_t - \omega\mathbf{1}_{(-\infty,t_0)}(t) - \tilde{\omega}\mathbf{1}_{[t_0,\infty)}(t)| + \varepsilon|(1 - \cos u + |\sin u|)|\varpi.\end{aligned}$$

By inserting the bounds (6.8), (6.9) and the right-hand sides of (6.10), (6.11), we obtain the left-hand side of (6.13). Similarly we get

$$|v_{ttt}| = |D_t V_v(u(t), v(t), t)| \ll \varepsilon\mu|u_t| + \varepsilon\mu|v_t - \omega\mathbf{1}_{(-\infty,t_0)}(t) - \tilde{\omega}\mathbf{1}_{[t_0,\infty)}(t)| + \varepsilon\mu\varpi(1 - \cos u),$$

which analogously as above implies the right-hand side of (6.13). □

Lemma 6.6. *The set $\mathcal{H}$ is compact in $H^2_{loc}(\mathbb{R})^2$. Furthermore, for each $q \in \mathcal{H}$ we have that $q \in \mathcal{E}$ and $q_t \in H^2_{ul}(\mathbb{R})^2$.*

Proof. It is straightforward to observe that the closure in $H^2_{\mathrm{loc}}(\mathbb{R})^2$ of all q satisfying (6.8)-(6.13) and $(t_0, v_0) \in [0, 2\pi]^2$ is compact. Thus it suffices to show that $\mathcal{H}$ is closed in $H^2_{\mathrm{loc}}(\mathbb{R})^2$. Assume $q_n \in \mathcal{H}$ connecting ω_n and $\tilde{\omega}_n$, $q_n(t_n) = (\pi, v_n)$, is a sequence converging to $q \in H^2_{\mathrm{loc}}(\mathbb{R})^2$. We first show that $q \in \mathcal{E}$. By construction, q is a solution of (1.3), and by Lemma 6.5 and the construction we easily show that $q_t \in L^\infty(\mathbb{R})^2$. Now $q \in \mathcal{E}$ follows from Lemma 15.5. We see that q must satisfy the condition (i) from the definition of $\mathcal{H}$, as the sequences $\int_{-\infty}^{t_n} L_{\omega_n}(q_n, (q_n)_t, t)dt$, $\int_{t_n}^{\infty} L_{\tilde{\omega}_n}(q_n, (q_n)_t, t)dt$ are convergent by (6.8)-(6.11) and the Lebesgue dominated convergence theorem, and as S^+, S^- are continuous.

To show (ii), note that $\mathcal{N}_{q_n}$ is a family of compact sets with a bounded union, thus we can find a convergent subsequence converging to a set $\mathcal{N}_q$ in the Hausdorff topology. By the construction, the sequence (t_n, v_n) converges to some $(t_0, v_0) \in \mathcal{N}_q$ such that $q(t_0) = (\pi, v_0)$. If $(\tilde{t}_0, \tilde{v}_0) \in \partial\mathcal{N}_q$, it is a limit of a subsequence of $(\tilde{t}_{n_k}, \tilde{v}_{n_k})$ lying on the boundaries of the convergent sub-sequence of $\partial\mathcal{N}_{q_n}$. The relation (6.4) now follows by the continuity of $(\omega, \tilde{\omega}, t, v) \mapsto \Sigma_{\omega,\tilde{\omega}}(t, v)$, established by the definition and Corollary 5.6. □

7. An approximate shadowing orbit

In this section we define an approximate shadowing orbit q^0 which can be understood as a suitable initial condition for (1.2). Furthermore, we define the set $\mathcal{A}$ from Lemma 1.4, and introduce the constants L, L_k, $k \in \mathbb{Z}$ (the time between the jumps) and M (the magnitude of oscillations of v with respect to (1.2b)) to be optimized later, as a scaffolding for the proofs. Finally we show that $q^0 \in \mathcal{A}$, and evaluate bounds on q^0 needed later. We will eventually see that essentially the only role of q^0 in the proofs is to show that the constructed sets $\mathcal{A}$, $\mathcal{B}$ are not empty.

Fix a closed subset of a region of instability $[\omega^-,\omega^+]$, with the uniform constants Δ_0, R as in (S1) and ϖ as in Introduction. Assume ω_k, $k\in\mathbb{Z}$ is a sequence in $[\omega^-,\omega^+]$ such that for all $k\in\mathbb{Z}$, (6.3) holds. The constant $4L$ will be the minimal time between two "jumps". Let $\tilde{L}_k$ be the approximate time of the "jumps", satisfying $\tilde{L}_k\equiv 0 \mod 2\pi$ and $\tilde{L}_{k+1}-\tilde{L}_k\geq 4L+2\pi$. Let $q_k=(u_k,v_k)\in\mathcal{H}$ and (T_k,V_k) be such that $q_k(T_k)=(\pi,V_k)$, and let $\mathcal{N}_{q_k}\subset\mathbb{R}^2$ be the sets associated to q_k as in the definition of $\mathcal{H}$. Note that if $(t_1,v_1)\in\partial\mathcal{N}_{q_k}$, and if $\tilde{q}=(\tilde{u},\tilde{v})\in H^1_{\rm loc}(\mathbb{R})^2$ such that $\tilde{q}(t_1)=(\pi,v_1)$, $\lim_{t\to-\infty}\tilde{u}(t)=0$, $\lim_{t\to\infty}\tilde{u}(t)=2\pi$, then by (6.4) and (6.7), we have

$$\mathcal{L}_{\omega,\omega^*}(\tilde{q})-\mathcal{L}_{\omega,\omega^*}(q)\geq 2\Delta_0>0. \tag{7.1}$$

Also, by the definition of Δ_0 and (5.2), we can easily deduce (using $\mu\leq 1/16$ by (A2)) the useful bound

$$\Delta_0\leq 9\sqrt{\varepsilon}\mu. \tag{7.2}$$

We construct the required parameters, functions and sets inductively in $|k|$ as follows: $\tilde{T}_0=T_0$, $\tilde{V}_0=V_0$, $\tilde{q}_0=(\tilde{u}_0,\tilde{v}_0):=q_0$, and

$$\begin{aligned}
\tilde{T}_k&=T_k \mod 2\pi,\quad \text{so that } -\pi<\tilde{T}_k-\tilde{L}_k\leq\pi,\\
\tilde{q}_k(t)&=(u_k(t-\tilde{T}_k+T_k)+2k\pi,\ v_k(t-\tilde{T}_k+T_k)+\tilde{V}_k-V_k),\\
\tilde{V}_k&=V_k \mod 2\pi \quad \text{so that } -\pi<\tilde{v}_{k-1}(\tilde{T}_k)-\tilde{V}_k\leq\pi \text{ for } k\geq 1,\\
\tilde{V}_k&=V_k \mod 2\pi \quad \text{so that } -\pi<\tilde{v}_{k+1}(\tilde{T}_k)-\tilde{V}_k\leq\pi \text{ for } k\leq -1,\\
\tilde{\mathcal{N}}_k&=\mathcal{N}_{q_k}+(\tilde{T}_k-T_k,\tilde{V}_k-V_k),\\
L_k&=\tilde{T}_{k+1}-\tilde{T}_k,
\end{aligned}$$

where we always use the notation $\tilde{q}_k=(\tilde{u}_k,\tilde{v}_k)$. We now require "smoothening" functions φ^-, φ^+, defined over an arbitrary interval $[a,b]$, $a<b$:

$$\begin{aligned}
\varphi^-_{a,b}(t)&=\begin{cases}1 & t\leq a,\\ \frac{\exp(-(b-a)/(t-a))}{\exp(-(b-a)/(t-a))+\exp(-(b-a)/(b-t))} & t\in[a,b],\\ 0 & t\geq b,\end{cases}\\
\varphi^+_{a,b}(t)&=1-\varphi^-_{a,b}(t).
\end{aligned} \tag{7.3}$$

By definition φ^-,φ^+ are C^∞, with values in $[0,1]$, and with uniformly bounded derivatives

$$(\varphi^-_{a,b})^{(k)}(t),(\varphi^+_{a,b})^{(k)}(t)=O_k\left(\frac{1}{|b-a|^k}\right), \tag{7.4}$$

where the implicit constant depends only on k. Let

$$q^0(t)=\varphi^-_{\tilde{T}_{k-1}+L,\tilde{T}_k-L}(t)\tilde{q}_{k-1}(t)+\varphi^+_{\tilde{T}_{k-1}+L,\tilde{T}_k-L}(t)\tilde{q}_k(t)\quad \text{for all } t\in[\tilde{T}_{k-1},\tilde{T}_k]. \tag{7.5}$$

Remark 7.1. Assume we fix a segment $[\omega^-,\omega^+]$ in a region of instability, and that for each $\omega,\tilde{\omega}\in[\omega^-,\omega^+]$ satisfying (6.3) we chose a single $q\in\mathcal{H}$ (as such q is not necessarily unique). Then q^0 is uniquely defined by the choice of L, $(\tilde{L}_k)_{k\in\mathbb{Z}}$, $(\omega_k)_{k\in\mathbb{Z}}$ (uniqueness of $T_k,V_k,\tilde{T}_k,\tilde{V}_k$ follows from Lemma 6.4). In the proofs of the main theorems, we thus use the notation $q^0(L,(\tilde{L}_k)_{k\in\mathbb{Z}},(\omega_k)_{k\in\mathbb{Z}})$. We fix q^0 for now and do not use such notation until Section 13.

Finally, let M be a constant chosen later, so that

$$M\geq\sup_{k\in\mathbb{Z}}\{|\tilde{v}_{k-1}(\tilde{T}_k)|-\tilde{V}_k|,|\tilde{v}_k(\tilde{T}_{k-1})|-\tilde{V}_{k-1}|\}+(\varpi+1)\mu. \tag{7.6}$$

The set $\mathcal{A}$ is defined as the set of all $q = (u, v) \in H^3_{\rm loc}(\mathbb{R})^2 \cap \mathcal{X}$ such that $q_t \in H^2_{\rm ul}(\mathbb{R})^2$, and such that for all $k \in \mathbb{Z}$,

$$|u(\tilde{T}_k) - (2k+1)\pi| \le \frac{1}{3}, \tag{7.7}$$

$$|v(\tilde{T}_k) - \tilde{V}_k| \le M. \tag{7.8}$$

Lemma 7.1. *We have that $q^0 \in \mathcal{A}$.*

Proof. The smoothness of q^0 and $q^0_t \in H^2_{\rm ul}(\mathbb{R})^2$ follow from the construction, Lemma 6.5 and Remark 15.1. By definition, $u(\tilde{T}_k) = (2k+1)\pi$ and $v(\tilde{T}_k) = \tilde{V}_k$, which trivially implies (7.7), (7.8). □

Let $k(t) = j$ for $t \in (T_{j-1}, T_j]$ and $\|t\| = \min\{t - T_{k(t)-1}, T_{k(t)} - t\} = \min\{|t - \tilde{T}_k|,\ k \in \mathbb{Z}\}$. Furthermore, let

$$\tilde{\omega}_k = \frac{\tilde{V}_k - \tilde{V}_{k-1}}{\tilde{T}_k - \tilde{T}_{k-1}}.$$

Lemma 7.2. *There exist an absolute constant $c_6 \ge 1$ so that any $q^0 = (u^0, v^0)$ given by (7.5) satisfies:*

$$|u^0(t) - 2k(t)\pi| \le c_6 e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|}, \tag{7.9}$$

$$|v^0(t) - V_{k(t)-1} - \tilde{\omega}_{k(t)}(t)(t - T_{k(t)-1})| \le c_6(1 \wedge M), \tag{7.10}$$

$$|u^0_t(t)| \le c_6 \left(\sqrt{\varepsilon}\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L_{k(t)}} \right), \qquad |v^0_t(t) - \omega_{k(t)}| \le c_6 \left(\sqrt{\varepsilon}\mu\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L_{k(t)}} \right), \tag{7.11}$$

$$|u^0_{tt}(t)| \le c_6 \left(\varepsilon\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L^2_{k(t)}} \right), \qquad |v^0_{tt}(t)| \le c_6 \left(\varepsilon\mu\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L^2_{k(t)}} \right), \tag{7.12}$$

and finally

$$|u^0_{ttt}(t)| \le c_6 \varpi \left(\varepsilon\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L^3_{k(t)}} \right), \tag{7.13a}$$

$$|v^0_{ttt}(t)| \le c_6 \varpi \left(\varepsilon\mu\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|} + \frac{1}{L^3_{k(t)}} \right). \tag{7.13b}$$

Proof. We write (7.5) in an abbreviated form $q^0 = \varphi^- \tilde{q}_{k-1} + \varphi^+ \tilde{q}_k$ for $t \in [\tilde{T}_{k-1}, \tilde{T}_k]$. Then

$$u^0 = \varphi^-(\tilde{u}_{k-1} - 2k\pi) + \varphi^+(\tilde{u}_k - 2k\pi) + 2k\pi, \tag{7.14}$$

$$v^0 = \varphi^-(\tilde{v}_{k-1} - \tilde{V}_{k-1} - \tilde{\omega}_k(t - \tilde{T}_{k-1})) + \varphi^+(\tilde{u}_k - \tilde{V}_{k-1} - \tilde{\omega}_k(t - \tilde{T}_{k-1})) + \tilde{V}_{k-1} + \tilde{\omega}_k(t - \tilde{T}_{k-1}). \tag{7.15}$$

To obtain (7.9), the left-hand side of (7.11), and (7.12), (7.13), it suffices to differentiate (7.14), (7.15) and insert (6.8)-(6.13) as required and (7.4).

From the left-hand sides of (6.10), (6.11) and the definition of $\tilde{v}_k$, $\tilde{T}_k$, $\tilde{V}_k$, we easily obtain that for all $t \in [T_{k-1}, T_k]$,

$$|\tilde{v}_{k-1}(t) - \tilde{V}_{k-1} - \omega_k(t - \tilde{T}_{k-1})| \ll \mu, \tag{7.16a}$$

$$|\tilde{v}_k(t) - \tilde{V}_k - \omega_k(t - \tilde{T}_k)| \ll \mu. \tag{7.16b}$$

By inserting $t = \tilde{T}_k$ in (7.16a), we get

$$\left| \omega_k - \frac{\tilde{v}_{k-1}(\tilde{T}_k) - \tilde{V}_{k-1}}{\tilde{T}_k - \tilde{T}_{k-1}} \right| \ll \frac{\mu}{\tilde{T}_k - \tilde{T}_{k-1}}.$$

By definition and (7.6) we know that $|\tilde{V}_k - \tilde{v}_{k-1}(\tilde{T}_k)| \ll 1 \wedge M$. As also $\mu \leq 1 \wedge M$, we have

$$|\omega_k - \tilde{\omega}_k| \leq \left|\omega_k - \frac{\tilde{v}_{k-1}(\tilde{T}_k) - \tilde{V}_{k-1}}{\tilde{T}_k - \tilde{T}_{k-1}}\right| + \left|\frac{\tilde{V}_k - \tilde{v}_{k-1}(\tilde{T}_k)}{\tilde{T}_k - \tilde{T}_{k-1}}\right| \ll \frac{1 \wedge M}{\tilde{T}_k - \tilde{T}_{k-1}}. \tag{7.17}$$

Combining it with (7.16), and using $\tilde{V}_{k-1} + \tilde{\omega}_k(t - \tilde{T}_{k-1}) = \tilde{V}_k + \tilde{\omega}_k(t - \tilde{T}_k)$ and $\mu \leq 1 \wedge M$, we obtain

$$|\tilde{v}_{k-1}(t) - \tilde{V}_{k-1} - \tilde{\omega}_k(t - \tilde{T}_{k-1})| \ll 1 \wedge M, \tag{7.18a}$$

$$|\tilde{v}_k(t) - \tilde{V}_{k-1} - \tilde{\omega}_k(t - \tilde{T}_{k-1})| \ll 1 \wedge M. \tag{7.18b}$$

Now (7.10) follows from (7.15) and (7.18). The right-hand side of (7.11) is obtained easily by differentiating (7.15), using the right-hand sides of (6.10), (6.11) and finally (7.17). □

8. Invariant sets with L^∞ bounds

We now construct $\mathcal{A}$-relatively ξ-invariant sets with respect to the dynamics (1.2) satisfying a-priori L^∞ bounds. More specifically, we construct $\mathcal{B}_1$ such that any $q = (u, v) \in \mathcal{B}_1$ satisfies for all $t \in \mathbb{R}$

$$|u(t) - 2k(t)\pi| \leq c_7 e^{-\frac{1}{2}\sqrt{\varepsilon}\,\|t\|}, \tag{8.1}$$

$$|v(t) - v^0(t)| \leq c_8 M, \tag{8.2}$$

where $c_7, c_8 > 0$ are absolute constants. We first define the set $\mathcal{B}_1$, then show that it is $\mathcal{A}$-relatively ξ-invariant, that $q^0 \in \mathcal{B}_1$, and finally we deduce (8.1) and (8.2).

Let $v_k^-, v_k^+ : [\tilde{T}_k, \tilde{T}_{k+1}] \to \mathbb{R}$ be the unique C^2 functions satisfying

$$v_k^-(\tilde{T}_k) = \tilde{V}_k + c_6 M, \qquad v_k^+(\tilde{T}_k) = \tilde{V}_k - c_6 M,$$
$$v_k^-(\tilde{T}_{k+1}) = \tilde{V}_{k+1} + c_6 M, \qquad v_k^+(\tilde{T}_{k+1}) = \tilde{V}_{k+1} - c_6 M,$$

and

$$-(v_k^-)_{tt}(t) = (v_k^+)_{tt}(t) = c_2^2 e^2\, \varepsilon\mu\, e^{-\sqrt{\varepsilon}\,\|t\|}.$$

We define $\mathcal{B}_1$ to be the set of all $q \in \mathcal{A}$ satisfying

$$z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2k\pi \leq u(t) \leq z^-(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2(k+1)\pi, \qquad t \in [\tilde{T}_k, \tilde{T}_{k+1}], \tag{8.3}$$

$$v_k^+(t) \leq v(t) \leq v_k^-(t), \qquad t \in [\tilde{T}_k, \tilde{T}_{k+1}]. \tag{8.4}$$

Lemma 8.1. *The set $\mathcal{B}_1$ is $\mathcal{A}$-relatively ξ-invariant.*

Proof. We apply twice the parabolic maximum principle [16, Sec. 7, Theorem 12]. Assume that $q(s_0) \in \mathcal{B}_1$, and that for all $s \in [s_0, s_1]$, $q(s) \in \mathcal{A}$. We have already shown in Lemma 5.3, (iii), that $z^+(. - \tilde{T}_k - 3/(4\sqrt{\varepsilon}))$ is a strict stationary sub-solution, and $z^-(. - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon}))$ a strict stationary super-solution of (1.2a) on $(\tilde{T}_k, \tilde{T}_{k+1})$. The assumptions and Lemma 5.2,(v) imply that

$$z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2(k-1)\pi \leq u(t, s) \leq z^-(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2k\pi$$

holds on the parabolic boundary

$$(t, s) \in \{[\tilde{T}_k, s], s \in [s_0, s_1]\} \cup \{[t, s_0], t \in [\tilde{T}_k, \tilde{T}_{k+1}]\} \cup \{[\tilde{T}_{k+1}, s], s \in [s_0, s_1]\}, \tag{8.5}$$

thus by the parabolic maximum principle, (8.3) holds for all $s \in [s_0, s_1]$ and all $k \in \mathbb{Z}$.

Consider now the bounds on z^-, z^+. By Lemma 5.2,(vi), we that for $t \in [\tilde{T}_k, \tilde{T}_{k+1}]$,

$$|z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) - 2\pi| \leq c_2 e^{-\frac{1}{2}\sqrt{\varepsilon}|t - \tilde{T}_k| + \frac{3}{8}} \leq c_2 e\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|}.$$

We obtain analogous bounds on $|z^-(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon}))|$. We deduce that whenever $u(t)$ satisfies (8.3), then (8.1) holds for some absolute $c_7 = c_2 e$.

We now show that whenever $u(t,s)$ satisfies (8.1), v^- is a super-solution, and v^+ a sub-solution of (1.2b) on $(\tilde{T}_k, \tilde{T}_{k+1})$. Consider v^-. By using the definition of $(v_k^-)_{tt}$, the relation $(1-\cos x) \leq (x - 2k(t)\pi)^2/2$, (8.1) with $c_7 = c_2 e$ and the standing assumption (A1), we get that for $t \in (\tilde{T}_k, \tilde{T}_{k+1})$,

$$\begin{aligned}(v_k^-)_{tt} - V_v(u, v^-, t) &= -c_2^2 e^2\, \varepsilon\mu e^{-\sqrt{\varepsilon}\,\|t\|} + \varepsilon\mu(1-\cos u(t))|f_v(u, v_k^-, t)| \\ &\leq -c_2^2 e^2\, \varepsilon\mu e^{-\sqrt{\varepsilon}\,\|t\|} + c_2^2 e^2\, \varepsilon\mu e^{-\sqrt{\varepsilon}\,\|t\|} \leq 0.\end{aligned}$$

Analogously we get that $(v_k^+)_{tt} - V_v(u, v_k^+, t) \geq 0$. Furthermore, by the definition of $\mathcal{A}$, we have that for all $k \in \mathbb{Z}$, $v_k^+(\tilde{T}_k) \leq v(\tilde{T}_k) \leq v_k^-(\tilde{T}_k)$ and $v_k^+(\tilde{T}_{k+1}) \leq v(\tilde{T}_{k+1}) \leq v_k^-(\tilde{T}_{k+1})$. It suffices now to apply the parabolic maximum principle to (1.2b) for all $k \in \mathbb{Z}$, with the same parabolic boundary (8.5). □

Lemma 8.2. *We have that $q^0 \in \mathcal{B}_1$.*

Proof. Use the notation $\tilde{q}_k = (\tilde{u}_k, \tilde{v}_k)$ and $\tilde{q}_{k+1} = (\tilde{u}_{k+1}, \tilde{v}_{k+1})$. From Proposition 5.4, we see that for $t \in [\tilde{T}_k, \tilde{T}_{k+1}]$,

$$\begin{aligned}z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2k\pi \leq \tilde{u}_k(t) \leq 2(k+1)\pi \leq z^-(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2(k+1)\pi, \\ z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2k\pi \leq 2(k+1)\pi \leq \tilde{u}_{k+1}(t) \leq z^-(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2(k+1)\pi.\end{aligned}$$

As u^0 is a convex combination of $\tilde{u}_k(t)$, $\tilde{u}_{k+1}(t)$ on $[\tilde{T}_k, \tilde{T}_{k+1}]$, (8.3) holds for $u = u^0$.

The relation (8.4) for $v = v^0$ follows from (7.10) and the definitions of v_k^-, v_k^+ (as v_k^- is concave and v_k^+ is convex). □

Lemma 8.3. *The relations (8.1), (8.2) hold for all $q \in \mathcal{B}_1$.*

Proof. The relation (8.1) has already been established in the proof of Lemma 8.1. As $v(t), v^0(t)$ satisfy (8.4), we have that

$$|v(t) - v^0(t)| \leq \sup\{|v_k^+(t) - v_k^-(t)|, k \in \mathbb{Z}, t \in [\tilde{T}_k, \tilde{T}_{k+1}]\}. \tag{8.6}$$

To establish a bound on $|v_k^+(t) - v_k^-(t)|$, we introduce $w(t) = v_k^-(t) - \tilde{V}_k - (\tilde{V}_{k+1} - \tilde{V}_k)(t - \tilde{T}_k) - M$. As $w(\tilde{T}_k) = w(\tilde{T}_{k+1}) = 0$, by symmetry $w_t(\tilde{T}) = 0$, where $\tilde{T} = (\tilde{T}_k + \tilde{T}_{k+1})/2$. Consider $T \in [\tilde{T}, \tilde{T}_{k+1}]$.

$$\begin{aligned}|w_t(T)| &= c_2^2 e^2\, \varepsilon\mu \int_{\tilde{T}}^{T} \exp(-\sqrt{\varepsilon}\,(\tilde{T}_k - t))dt \leq 8e^2\sqrt{\varepsilon}\,\mu \exp(-\sqrt{\varepsilon}(\tilde{T}_k - T)), \\ |w(T) - w(\tilde{T})| &\leq c_2^2 e^2\, \sqrt{\varepsilon}\,\mu \int_{\tilde{T}}^{T} \exp(-\sqrt{\varepsilon}(\tilde{T}_k - t))dt \leq c_2^2 e^2 \mu.\end{aligned}$$

As $w(T)$ is decreasing on $[\tilde{T}, \tilde{T}_{k+1}]$, we get that $|w(t)| \leq c_2^2 e^2 \mu$. By analogy, the same holds on $[\tilde{T}_k, \tilde{T}]$. We see that $|v_k^-(t) - \tilde{V}_k - (\tilde{V}_{k+1} - \tilde{V}_k)(t - \tilde{T}_k)| \leq M + c_2^2 e^2 \mu$. Analogously, $|v_k^+(t) - \tilde{V}_k - (\tilde{V}_{k+1} - \tilde{V}_k)(t - \tilde{T}_k)| \leq M + c_2^2 e^2 \mu$, thus

$$|v_k^+(t) - v_k^-(t)| \leq 2M + 2c_2^2 e^2\, \mu, \text{ for all } t \in [\tilde{T}_k, \tilde{T}_{k+1}],$$

which is by (7.6) $\leq c_8 M$, with $c_8 = 2c_2^2 e^2 + 2$. It suffices to insert this in (8.6). □

9. Bounds on the derivatives

Here we show that there is a $\mathcal{A}$-relatively ξ-invariant set such that the norms of the first, second and third order derivatives of u, v all behave as $O_\varepsilon(\log(\|t\|)/\|t\|)$. Let

$$\lambda(\tau) = \frac{\sqrt{\varepsilon}}{4} \wedge \frac{8\log\|\tau\|}{\|\tau\|}, \tag{9.1}$$

where $\wedge$ is the minimum. In this section we use the weighted L^2 norm

$$||w||^2_{L^2_\tau(\mathbb{R})} := \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|} w^2(t)dt.$$

Let $\mathcal{B}_2$ be the set of all $q \in \mathcal{B}_1$ such that

$$||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})^2} \leq c_9\, \lambda(\tau), \tag{9.2a}$$

$$||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})^2} \leq c_9(M^2+1)\, \lambda(\tau), \tag{9.2b}$$

for all $\tau \in \mathbb{R}$. Let $\mathcal{B}_3$ be the set of all $q \in \mathcal{B}_2$ such that for all $\tau \in \mathbb{R}$,

$$\varepsilon||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})^2} + ||u_{tt}||^2_{L^2_\tau(\mathbb{R})^2} \leq c_{10}(M^2+\varpi^2)\varepsilon\, \lambda(\tau), \tag{9.3a}$$

$$\varepsilon||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})^2} + ||v_{tt}||^2_{L^2_\tau(\mathbb{R})^2} \leq c_{10}(M^2+\varpi^2)\varepsilon\, \lambda(\tau). \tag{9.3b}$$

Finally, let $\mathcal{B}_4$ be the set of all $q \in \mathcal{B}_3$ such that for all $\tau \in \mathbb{R}$,

$$\varepsilon||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})^2} + ||u_{tt}||^2_{L^2_\tau(\mathbb{R})^2} + ||u_{ttt}||^2_{L^2_\tau(\mathbb{R})^2} \leq c_{11}(M^4+\varpi^4)\varepsilon\, \lambda(\tau), \tag{9.4a}$$

$$\varepsilon||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})^2} + ||v_{tt}||^2_{L^2_\tau(\mathbb{R})^2} + ||v_{ttt}||^2_{L^2_\tau(\mathbb{R})^2} \leq c_{11}(M^4+\varpi^4)\varepsilon\, \lambda(\tau). \tag{9.4b}$$

Proposition 9.1. *There exist absolute constants c_9, c_{10} and c_{11} so that the sets $\mathcal{B}_2$, $\mathcal{B}_3$ and $\mathcal{B}_4$ are $\mathcal{A}$-relatively ξ-invariant, and such that $q^0 \in \mathcal{B}_4$.*

The proof of Proposition is routine but technical, and as such postponed to the Appendix C. In essence, by differentiating the weighted integral versions of (1.2a), (1.2b), we obtain a differential inequality which by the Gronwall's lemma implies invariance of the sets as required. An important step is use of a variant of the Poincaré inequality (Lemma 17.4) which relies on the L^∞ bounds obtained in the previous section. We do the procedure iteratively for the three sets.

The main implication needed in the following is that we can for each k approximate $q - \tilde{q}_k$ close to $\tilde{T}_k$ with a "well-behaved" h vanishing at $\pm\infty$.

Lemma 9.2. *Assume that $q \in \mathcal{B}_4$. Then there exists an absolute constant $c_{12} > 0$ such that for each $k \in \mathbb{Z}$ there exist $\tilde{h} = (\tilde{u}^h, \tilde{v}^h) \in H^3_{loc}(\mathbb{R})^2$ satisfying the following:*

(i) *For all $t \in [\tilde{T}_k - L, \tilde{T}_k + L]$, $\tilde{h}(t) = q(t) - \tilde{q}_k(t)$,*
(ii) *For $t \geq \tilde{T}_k + L(1 + 1/\log L)$ and for $t \leq \tilde{T}_k - L(1+1/\log L)$ we have $\tilde{h}(t) = 0$,*
(iii) *For all $t \in \mathbb{R}$, $|\tilde{u}^h(t)| \leq c_{12} e^{-\frac{1}{2}\sqrt{\varepsilon}|t-\tilde{T}_k|}$,*
(iv) *For all $t \in [\tilde{T}_k - 2\pi, \tilde{T}_k + 2\pi]$, we have that $|\tilde{v}^h(t)| \leq c_8 M$,*
(v) *For all $T \geq 0$,*

$$||\tilde{h}_t||^2_{H^2((-\infty,\tilde{T}_k - T])^2} + ||\tilde{h}_t||^2_{H^2([\tilde{T}_k+T,\infty))^2} \leq c_{12}(M^4+\varpi^4)\left(\frac{\log^2 T}{T} \wedge \frac{\sqrt{\varepsilon}|\log\varepsilon|}{8}\right).$$

(vi) *Specifically, for $T = L$, we have*

$$||\tilde{h}_t||^2_{H^2((-\infty,\tilde{T}_k - L])^2} + ||\tilde{h}_t||^2_{H^2([\tilde{T}_k+L,\infty))^2} \leq c_{12}(M^4+\varpi^4)\frac{\log L}{L}.$$

We will use the following simple Lemma:

Lemma 9.3. *Assume $y_0 \geq 4$ and let $y_{j+1} = y_j\,(1 + 1/\log y_j)$. Then for some absolute implicit constant,*

$$\sum_{j=0}^{\infty} \frac{\log y_j}{y_j} \ll \frac{\log^2 y_0}{y_0}.$$

Proof. We first show inductively in $j = 0, 1, ...$ that $y_j \geq e^{\frac{1}{2}\sqrt{x+j}}$, where x is chosen so that $y_0 = e^{\frac{1}{2}\sqrt{x}}$, i.e. $x = 4\log^2 y_0$. Indeed, by the Mean Value Theorem, there is a real number z, $j \leq z \leq j+1$, so that

$$e^{\frac{1}{2}\sqrt{x+j+1}} = e^{\frac{1}{2}\sqrt{x+j}} + \frac{e^{\frac{1}{2}\sqrt{x+z}}}{4\sqrt{x+z}} \leq e^{\frac{1}{2}\sqrt{x+j}} + \frac{e^{1/2}}{8}\cdot\frac{e^{\frac{1}{2}\sqrt{x+j}}}{\frac{1}{2}\sqrt{x+j}} \leq y_j + \frac{y_j}{\log y_j} = y_{j+1}.$$

Now as $\log y/y$ is decreasing for $y \geq 4$,

$$\sum_{j=0}^{\infty}\frac{\log y_j}{y_j} \leq \frac{\log y_0}{y_0} + \int_x^\infty \frac{1}{2}\sqrt{z}e^{-\frac{1}{2}\sqrt{z}}dz \ll \frac{\log y_0}{y_0} + xe^{-\frac{1}{2}\sqrt{x}} = \frac{\log y_0}{y_0} + 4\frac{\log^2 y_0}{y_0} \ll \frac{\log^2 y_0}{y_0}.$$

□

Proof of Lemma 9.2. Let

$$\tilde{h}(t) = \begin{cases} \varphi^{+}_{\tilde{T}_k - L(1+1/\log L), \tilde{T}_k - L}(t) \cdot (q(t) - \tilde{q}_k(t)), & t \le \tilde{T}_k, \\ \varphi^{-}_{\tilde{T}_k + L, \tilde{T}_k + L(1+1/\log L)}(t) \cdot (q(t) - \tilde{q}_k(t)), & t \ge \tilde{T}_k. \end{cases}$$

Now, (i) and (ii) follow from the definition of the smoothening functions (7.3), and (iii) from (6.8), the definition of $\tilde{q}_k$ and (8.1). We claim that for each $T \ge 4/\sqrt{\varepsilon}$,

$$||\tilde{h}_t||^2_{H^2([\tilde{T}_k + T, \tilde{T}_k + T(1+1/\log T)])^2} \ll (M^4 + \varpi^4) \frac{\log T}{T}. \tag{9.5}$$

For $T \le L$, this follows directly from (9.4) with $\tau = \tilde{T}_k + T$, where the bounds on $\tilde{q}_k$ follow from its definition (as it is a translate of $q_k \in \mathcal{H}$), the bounds on q_k from (6.10)-(6.13), and finally by using (7.4). To show (iv), we use (i), and the fact that for all $t \in [\tilde{T}_k - 2\pi, \tilde{T}_k + 2\pi]$, $q^0(t) = \tilde{q}^k(t)$, thus by (8.2)

$$|\tilde{v}^h(t)| = |v(t) - \tilde{v}_k(t)| = |v(t) - v^0(t)| \le c_8 M.$$

As by (ii), $h(t)$ vanishes for $t \ge L(1 + 1/\log L)$, the claim holds for $T \ge L(1 + 1/\log L)$. For $T \in [L, L(1 + 1/\log L)]$, (9.5) similarly follows from the case $T = L$. Analogously we obtain for such T,

$$||\tilde{h}_t||^2_{H^2([\tilde{T}_k - T(1+1/\log T), \tilde{T}_k - T])^2} \ll (M^4 + \varpi^4) \frac{\log T}{T}, \tag{9.6}$$

and

$$||\tilde{h}_t||^2_{H^2([-4/\sqrt{\varepsilon}, 0])^2} \ll (M^4 + \varpi^4)\sqrt{\varepsilon}, \quad ||\tilde{h}_t||^2_{H^2([0, 4/\sqrt{\varepsilon}])^2} \ll (M^4 + \varpi^4)\sqrt{\varepsilon}. \tag{9.7}$$

Now (vi) follows from (9.5) and (9.6) with $T = L$, and again by noting that by (ii), $\tilde{h}(t)$ vanishes for $t \le \tilde{T}_k - L(1 + 1/\log L)$ and $t \ge \tilde{T}_k + L(1 + 1/\log L)$. We obtain (v) as follows: in the case $T \ge 4/\sqrt{\varepsilon}$, we combine (9.5) and (9.6) while inserting a sequence of $y_0 = T$, $y_j = y_{j-1}(1 + 1/\log y_j)$ instead of T, and applying Lemma 9.3. If $T \le 4/\sqrt{\varepsilon}$, we add another term in the that estimate by using (9.7). □

10. Lower bound on the action dissipation

In this section we develop a lower bound for the dissipation of the action with respect to the dynamics (1.2).

We now fix the constant M with

$$M = 2\pi + 2(\varpi + 1)(R + \mu) + 6R^{1/2}\varepsilon^{1/4}. \tag{10.1}$$

(clearly M satisfies (7.6) as required). Let $\mathcal{C}$ be the closure in $H^2_{\text{loc}}(\mathbb{R})^2$ of the set of all $h = (u^h, v^h) \in H^3_{\text{loc}}(\mathbb{R})^2$ satisfiying for all $t \in \mathbb{R}$ and all $T \ge 0$,

$$|u^h(t)| \le c_{12}\, e^{-\frac{1}{2}\sqrt{\varepsilon}|t| + 2\pi}, \tag{10.2}$$

$$|v^h(0)| \le c_8 M, \tag{10.3}$$

$$||h_t||^2_{H^2((-\infty, -T])^2} + ||h_t||^2_{H^2([T, \infty))^2} \le 2c_{12}(M^4 + \varpi^4)\left(\frac{\log^2 T}{T} \wedge \frac{\sqrt{\varepsilon}|\log \varepsilon|}{8}\right). \tag{10.4}$$

Consider for $(q, h) \in \mathcal{H} \times \mathcal{C}$, $q = (u, v)$, $h = (u^h, v^h)$,

$$E_q(h) = \int_{-\infty}^{0} L_{\omega^-(q)}(q + h, q_t + h_t, t)dt + \int_0^\infty L_{\omega^+(q)}(q + h, q_t + h_t, t) + (\omega^+(q) - \omega^-(q))(v(0) + v^h(0)),$$

$$D_q(h) = \int_{-\infty}^{\infty} (q + h)_s^2 dt,$$

where we take $\omega^-(q) = \lim_{t\to-\infty} v_t$ and $\omega^+(q) = \lim_{t\to\infty} v_t$, and q_s is evaluated by inserting (1.2). We establish an uniform lower bound on the action dissipation D_q on a certain level of action:

Proposition 10.1. *There exists a constant $\Delta_1 > 0$, $0 \leq \Delta_1 \leq \Delta_0/2$, depending on the region of instability $[\omega^-, \omega^+]$, R, Δ_0 and f, and constants $0 \leq \Delta_0(q) \leq \Delta_0/2$ defined for all $q \in \mathcal{H}$, so that for all $(q,h) \in \mathcal{H}\times\mathcal{C}$, if*

$$|E_q(h) - E_q(0) - \Delta_0(q)| \leq \Delta_1,$$

then

$$D_q(h) \geq \Delta_1.$$

To obtain Δ_1 we introduce for any $q \in \mathcal{H}$:

$$\Delta_1(q,e) = \inf\{D_q(h),\ \ h \in \mathcal{C},\ E_q(h) = E_q(0) + e\}, \tag{10.5}$$

$$\Delta_1(q) = \sup_{e\in[0,\Delta_0]} \Delta_1(q,e). \tag{10.6}$$

We prove the Proposition in several steps. First we recall an infinite-dimensional version of the Morse-Sard theorem, which will enable us to deduce that $\Delta_1(q) > 0$ for all $q \in \mathcal{H}$. We then in several lemmas establish various continuity and lower semi-continuity properties, which combined with compactness of $\mathcal{H}$, $\mathcal{C}$ enables us to complete the proof.

Recall first the Pohožaev infinite-dimensional version of the Morse-Sard theorem. Consider a real functional E on a real, separable, reflexive Banach space $\mathcal{Y}$. We say that E is Fredholm, if it is C^2 (in the sense of Fréchet derivatives), and the dimension of Ker $D^2E(h)$, $D^2E(h) : \mathcal{Y} \to \mathcal{Y}^*$ is finite dimensional for any $h \in \mathcal{Y}$. (Equivalently, D^2E is Fredholm, as in this case it suffices to check finite dimensionality of the kernel [38].) A critical value of E is any value $e \in \mathbb{R}$ for which there exists $h \in \mathcal{Y}$ so that $E(h) = e$ and $DE(h) = 0$.

Lemma 10.2. Morse-Sard-Pohožaev [38]. *Assume that $E : \mathcal{Y} \to \mathbb{R}$ is a real, C^k functional defined on a real, separable, reflexive Banach space $\mathcal{Y}$. Assume that* $\dim($*Ker* $D^2E(h)) \leq m < \infty$ *for any $h \in \mathcal{Y}$, and let $k \geq \max\{m, 2\}$. Then the set of critical values of E has Lebesgue measure 0.*

We will apply Lemma 10.2 to the functionals $h \mapsto E_q(h)$ for $q \in \mathcal{H}$. Let $\mathcal{Y}$ be the set of all $q = (u,v) \in H^2_{\text{loc}}(\mathbb{R})^2$ such that $||q||_{\mathcal{Y}} < \infty$, where

$$||q||_{\mathcal{Y}} = \left(\int_{-\infty}^{\infty} e^{\frac{1}{4}\sqrt{\varepsilon}\,|t|} u(t)^2 dt + |v(0)|^2 + ||u_t||^2_{H^1(\mathbb{R})} + ||v_t||^2_{H^1(\mathbb{R})}\right)^{1/2}. \tag{10.7}$$

The space $\mathcal{Y}$ is a Hilbert space, as the norm (10.7) is induced by a scalar product defined in a straight-forward way. Thus $\mathcal{Y}$ is separable and reflexive.

Let us establish compactness of $\mathcal{H}$ and $\mathcal{C}$ and continuity of D_q, E_q. Recall that assumed topology on $\mathcal{H}$ and $\mathcal{C}$ is induced by $H^2_{\text{loc}}(\mathbb{R})^2$. It is straightforward to verify that on $\mathcal{C}$ it coincides with the topology induced by the $\mathcal{Y}$-norm.

Lemma 10.3. *(i) The sets $\mathcal{H}$ and $\mathcal{C}$ are compact,*

(ii) The functions $(q,h) \mapsto E_q(h), D_q(h)$ are well-defined and continuous on $\mathcal{H} \times \mathcal{C}$.

Proof. We have showed compactness of $\mathcal{H}$ in Lemma 6.6. Compactness of $\mathcal{C}$ follows directly from the definition, the compact embedding theorem applied to h restricted to any bounded closed interval, and a diagonalization argument. The claim (ii) follows easily from the definitions of E_q, D_q, the uniform bounds on $q \in \mathcal{H}$ in Lemma 6.5, the definition of $\mathcal{C}$ and the assumed localized topologies on the sets $\mathcal{H}$ and $\mathcal{C}$. □

Lemma 10.4. *For any $q \in \mathcal{H}$,*

(i) $E_q : \mathcal{Y} \to \mathbb{R}$ is C^4,

(ii) E_q is Fredholm and the dimension of Ker D^2E_q is at most 4,

(iii) If $DE_q(h) \neq 0$, then $D_q(h) > 0$.

Proof. Let $q = (u, v)$, $h = (u^h, v^h) \in \mathcal{Y}$, and $g^{(1)}, g^{(2)}, g^{(3)}, g^{(4)} \in \mathcal{Y}$, $g^{(j)} = (u^{(j)}, v^{(j)})$. We will show that the Fréchet derivatives of E_q are for given with

$$DE_q(h)g^{(1)} = \int_{-\infty}^{0} (v_t - \omega + v_t^h)v_t^{(1)}dt + \int_0^\infty (v_t - \tilde{\omega} + v_t^h)v_t^{(1)}dt + (\tilde{\omega} - \omega)v^{(1)}(0)$$
$$+ \int_{-\infty}^{\infty} \left\{(u_t + u_t^h)u_t^{(1)} + D_{u,v}V(q(t) + h(t), t)g^{(1)}(t)\right\} dt, \tag{10.8}$$

$$D^2E_q(h)(g^{(1)}, g^{(2)}) = \int_{-\infty}^{\infty} \left\{g_t^{(1)}g_t^{(2)} + D_{u,v}^2 V(q(t) + h(t), t)(g^{(1)}(t), g^{(2)}(t))\right\} dt, \tag{10.9}$$

$$D^kE_q(h)(g^{(1)}, ..., g^{(k)}) = \int_{-\infty}^{\infty} \left\{D_{u,v}^k V(q(t) + h(t), t)(g^{(1)}(t), ..., g^{(k)}(t))\right\} dt, \tag{10.10}$$

where $k = 3, 4$. We first show that the integrals on the right-hand sides are finite. Consider the terms containing u_t, u_t^h, v_t, v_t^h in (10.8). They are absolutely integrable by Cauchy-Schwartz, as $||u_t^h||_{L^2(\mathbb{R})}^2 < \infty$, $||v_t^h||_{L^2(\mathbb{R})}^2 < \infty$ by the definition of $\mathcal{Y}$, and by Lemma 6.5 applied to $u_t, |v_t - \omega|, |v_t - \tilde{\omega}|$. Analogously the term $g_t^{(1)}g_t^{(2)}$ in (10.9) is absolutely integrable. Thus it suffices to show that for any integers $i, j \geq 0$, $i + j = k$, $k = 1, 2, 3, 4$, the integral

$$X := \int_{-\infty}^{\infty} \left\{\partial_u^i \partial_v^j V(q(t) + h(t), t)u^1(t)...u^i(t)v^{i+1}(t)...v^k(t)\right\} dt \tag{10.11}$$

is absolutely integrable. It is straightforward to check that for $g^{(j)} = (u^{(j)}, v^{(j)}) \in \mathcal{Y}$,

$$|u^{(j)}(t)| \leq ||u^{(j)}||_{L^\infty(\mathbb{R})} \ll ||u^{(j)}||_{H^1(\mathbb{R})} \leq ||g^{(j)}||_{\mathcal{Y}},$$
$$|v^{(j)}(t)| \leq |v^{(j)}(0)| + \int_0^t |v_t^{(j)}(\tau)|d\tau \ll (1 + |t|^{1/2})||g^{(j)}||_{\mathcal{Y}}.$$

If $i \geq 1$, we thus have by applying uniform bounds on derivatives of V,

$$X \ll_f ||g^{(2)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}} \int_{-\infty}^{\infty} |u^{(1)}(t)|(1 + |t|^{1/2})^{k-1}dt$$
$$\ll_f ||g^{(2)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}} \left(\int_{-\infty}^{\infty} e^{-\frac{1}{4}\sqrt{\varepsilon}|t|}(1 + |t|^{k-1})dt\right)^{1/2} \left(e^{\frac{1}{4}\sqrt{\varepsilon}|t|}|u^{(1)}(t)|^2dt\right)^{1/2}$$
$$\ll_{f,\varepsilon} ||g^{(1)}||_{\mathcal{Y}}||g^{(2)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}}. \tag{10.12}$$

Analogously, in the case $i = 0$,

$$X \ll ||g^{(1)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}} \int_{-\infty}^{\infty} |\partial_v^k V(q(t) + h(t), t)|(1 + |t|^{1/2})^k dt$$
$$\ll_f ||g^{(1)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}} \int_{-\infty}^{\infty} \left(1 - \cos(u(t) + u^h(t))\right)(1 + |t|^{1/2})^k dt$$
$$\ll_f ||g^{(1)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}} \int_{-\infty}^{\infty} \left(e^{-\frac{1}{2}\sqrt{\varepsilon}(t-t_0)} + u^h(t)^2\right)(1 + |t|^{1/2})^k dt,$$

where in the last row we applied $1 - \cos(u(t) + u^h(t)) \ll (u(t) - 2\pi\mathbf{1}_{[t_0,\infty)}(t))^2 + u^h(t)^2$ and then (6.8). Analogously to (10.12) we establish bound on the remaining terms containing u^h and get

$$X \ll_{f,\varepsilon} ||g^{(1)}||_{\mathcal{Y}}...||g^{(k)}||_{\mathcal{Y}}(1 + ||h||_{\mathcal{Y}}^2). \tag{10.13}$$

We now show that (10.8) is indeed the Fréchet derivative of E_q; the proof for higher order derivatives is analogous (we use Hölder continuity of fourth derivatives assumed in (A1) for D^4E_q). By the Mean Value theorem and analogously as when evaluating (10.12), (10.13), we obtain

$$\begin{aligned}|E_q(h+g^{(1)}) - E_q(h) - DE_q(h)g^{(1)}| &\leq \int_{-\infty}^{\infty} \left\{ \frac{1}{2}(g_t^{(1)})^2 + |D^2_{u,v}V(q(t)+h(t),t)||g^{(1)}|^2 \right\} dt \\ &\ll_{f,\varepsilon} (1+||h||^2_{\mathcal{Y}})||g^{(1)}||^2_{\mathcal{Y}},\end{aligned}$$

which by the definition of the Fréchet derivative gives the claim.

To show (ii), we observe that the kernel of $D^2E_q(h)$ is by partial integration the set of all $g^{(1)} \in \mathcal{Y}$ such that $-g_{tt}^{(1)} + D^2_{u,v}V(q(t)+h(t),t)g^{(1)}(t) = 0$ for all $t \in \mathbb{R}$. This is a system of four linear ordinary differential equations (a linearization of (1.3)), so the space of its solutions in $\mathcal{Y}$ is at most four-dimensional.

By partial integration and (1.2), we can write (10.8) as $DE_q(h)g^{(1)} = \int_{-\infty}^{\infty} -(q+h)_s g^{(1)} dt$, which is clearly $\equiv 0$ if and only if $(q+h)_s \equiv 0$, which implies (iii). □

We are now ready to apply the Morse-Sard-Pohožaev Lemma.

Lemma 10.5. *For any $q \in \mathcal{H}$, there exists e, $0 \leq e \leq \Delta_0$ such that $\Delta_1(q,e) > 0$.*

Proof. Because of Lemma 10.4, (i) and (ii), we can apply the Morse-Sard-Pohožaev Lemma 10.2 to the functional E_q, and find any level set $E_q(h) = e$, $0 \leq e \leq \Delta_0$, so that for any $h \in \mathcal{Y}$, $E_q(h) = E_q(0) + e$ implies that $DE_q(h) \neq 0$. By Lemma 10.4, (iii), for any $h \in \mathcal{Y}$, $D_q(h) > 0$.

By Lemma 10.3, (i) and (ii), the level set $\{E_q(h) = E_q(0)+e\} \cap \mathcal{C}$ is a compact subset of $\mathcal{Y}$. By continuity of D_q, we can bound $D_q(h)$ away from zero on that level set, which by definition of $\Delta_1(q,e)$ completes the proof. □

Lemma 10.6. *The function $(q,e) \mapsto \Delta_1(q,e)$ is lower semi-continuous on $\mathcal{H} \times [0,\Delta_0]$.*

Proof. Choose a sequence $(q_n,e_n) \in \mathcal{H} \times [0,\Delta_0]$ converging to (q,e). As the level sets $\{E_q(h) = E_q(0)+e\} \cap \mathcal{C}$ are compact, we can find $h_n \in \mathcal{C}$, $n \in \mathbb{N}$ so that $E_{q_n}(h_n) = E_{q_n}(0) + e_n$ and $D_{q_n}(h_n) = \Delta_1(q_n,e_n)$.

Find a subsequence k_n so that $\liminf_{n\to\infty} \Delta_1(q_n,e_n) = \lim_{n\to\infty} \Delta_1(q_{k_n},e_{k_n})$, and a further subsequence (denoted again by k_n) so that h_{k_n} is convergent in $\mathcal{C}$. Let $h = \lim_{n\to\infty} h_{k_n}$. Now because of continuity of $(q,h) \mapsto D_q(h)$ by Lemma 10.3, (ii), we have

$$\liminf_{n\to\infty} \Delta_1(q_n,e_n) = \lim_{n\to\infty} \Delta_1(q_{k_n},e_{k_n}) = \lim_{n\to\infty} D_{q_n}(h_n) = D_q(h).$$

Again by Lemma 10.3, (ii), $E_q(h) = E_q(0) + e$, thus by definition $D_q(h) \geq \Delta_1(h,e)$, which completes the proof. □

Proof of Proposition 10.1. We prove it by using several times the well-known properties of lower semi-continuous functions. As $q \mapsto \Delta_1(q)$ is by definition a supremum of a family of lower semi-continuous functions, it is lower-semi continuous. By Lemma 10.5, $\Delta_1(q) > 0$ for every $q \in \mathcal{H}$. As lower semi-continuous functions on a compact set attain a minimum, and $\mathcal{H}$ is compact, there exists $\tilde{\Delta}_1 > 0$ so that $\Delta_1(q) \geq \tilde{\Delta}_1$ for all $q \in \mathcal{H}$.

Now consider the set $\{\Delta_1(q,e) > \tilde{\Delta}_1/2\} \subset \mathcal{H} \times [0,\Delta_0/2]$. By lower semi-continuity of $(q,e) \mapsto \Delta_1(q,e)$, it is an open subset of $\mathcal{H} \times [0,\Delta_0/2]$. We can find its open cover consisting of sets $U_i \times B(e_i,r_i)$, $i \in \mathcal{I}$, where $U_i \subset \mathcal{H}$ is open in $\mathcal{H}$, and $B(e_i,r_i)$ are open balls in $\mathbb{R}$. By choice of $\tilde{\Delta}_1$, the set $\{\Delta_1(q,e) > \tilde{\Delta}_1/2\}$ projects in the first coordinate to the entire $\mathcal{H}$, so U_i, $i \in \mathcal{I}$ is an open cover of $\mathcal{H}$. By compactness, we find its finite subcover. Denote it by $U_1,...,U_n$, and its associated open balls by $B(e_1,r_1),...,B(e_n,r_n)$.

We now set $\Delta_1 = \min\{\tilde{\Delta}_1/2, r_1,...,r_n\}$, and set for $q \in U_j$, $\Delta_0(q) = e_j$ (we choose any j if q is in more than one U_j). It is straightforward to check that this completes the proof. □

11. Local upper bounds on the action

This section contains the core of the argument, as we show that the action within two "intersections", or more precisely in the segment $[\tilde{T}_k - L, \tilde{T}_k + L]$, can not increase more than an arbitrarily small constant, proportional to $\log L/L$. This will complete our method of control of the dynamics. The proof relies on an action-balance law, stating that the change of action with respect to (1.2) is equal to the action dissipation and action flux. Let $\tilde{E}_k, \tilde{D}_k, \tilde{F}_k : \mathcal{X} \to \mathbb{R}$ be the truncated action, action dissipation and action flux near $\tilde{T}_k$, defined for $q = (u, v)$ as

$$\begin{aligned}
\tilde{E}_k(q) &= \int_{\tilde{T}_k - L}^{\tilde{T}_k} L_{\omega_k}(q, q_t, t)dt + \int_{\tilde{T}_k}^{\tilde{T}_k + L} L_{\omega_{k+1}}(q, q_t, t)dt + (\omega_{k+1} - \omega_k)v(\tilde{T}_k), \\
\tilde{D}_k(q) &= \int_{\tilde{T}_k - L}^{\tilde{T}_k + L} \left\{ u_s^2 + v_s^2 \right\} dt, \\
\tilde{F}_k(q) &= u_t(\tilde{T}_k + L)u_s(\tilde{T}_k + L) + (v_t(\tilde{T}_k + L) - \omega_{k+1})v_s(\tilde{T}_k + L) \\
&\quad - u_t(\tilde{T}_k - L)u_s(\tilde{T}_k - L) - (v_t(\tilde{T}_k - L) - \omega_k)v_s(\tilde{T}_k - L).
\end{aligned}$$

Let $\mathcal{B}_5$ be the set of all $q \in \mathcal{B}_4$ so that for all $k \in \mathbb{Z}$,

$$\tilde{E}_k(q) \le \tilde{E}_k(\tilde{q}_k) + \Delta_0(q_k), \tag{11.1}$$

where $\Delta_0(\tilde{q}_k)$, Δ_1 are as constructed in Proposition 10.1.

Proposition 11.1. *There exists an absolute constant $c_{13} > 0$ such that, if*

$$L \ge c_{13}(\varpi^5 + M^5)\frac{|\log \Delta_1|}{\Delta_1}, \tag{11.2}$$

then $\mathcal{B}_5$ is $\mathcal{A}$-relatively ξ-invariant. Furthermore, $q^0 \in \mathcal{B}_5$.

(Recall that M in (11.2) is given by (10.1).) To prove it, we first establish the action-balance law (11.3), then find upper bounds on the action flux, lower bounds on the action dissipation on the energy level $\tilde{E}_k(q) = \tilde{E}_k(\tilde{q}_k) + \Delta_0(\tilde{q}_k)$, and then complete the proof. The constant c_{13} may change throughout the section.

Lemma 11.2. *For any $q \in \mathcal{X}$,*

$$\frac{d}{ds}\tilde{E}_k(q) = -\tilde{D}_k(q) + \tilde{F}_k(q). \tag{11.3}$$

Proof. The proof is straightforward, by differentiating $\tilde{E}_k(q)$, partial integration and inserting (1.2). □

Lemma 11.3. *There exists an absolute constant $c_{13} > 0$ such that if (11.2) holds, then for any $q \in \mathcal{B}_4$ and any $k \in \mathbb{Z}$, $|\tilde{F}_k(q)| \le \Delta_1/4$.*

Proof. By (7.11), (9.3a), the definition of $\lambda(\tau)$, using that $\|\tau\| \le L_{k(\tau)}$ and $\varepsilon \le 1$, we obtain

$$\begin{aligned}
|u_t(\tau)| &\ll ||u_t||_{H^1([\tau,\tau+1])} \ll (M + \varpi)\lambda(\tau)^{1/2} + ||u_t^0||_{L^2[\tau,\tau+1]} \\
&\ll (M + \varpi)\left(\frac{\log ||\tau||}{\|\tau\|}\right)^{1/2},
\end{aligned} \tag{11.4}$$

and analogously

$$|v_t(\tau) - \omega_{k(\tau)}| \ll (M + \varpi)\left(\frac{\log ||\tau||}{\|\tau\|}\right)^{1/2}. \tag{11.5}$$

Because of (17.12) and (17.13) in the Appendix C, we get for all $q \in \mathcal{B}_4 \subset \mathcal{B}_1$, $|V_u| \ll e^{-\sqrt{\varepsilon}\|t\|/2} \ll (\log\|t\|/\|t\|)^{1/2}$, $|V_v| \ll e^{-\sqrt{\varepsilon}\|t\|} \ll (\log\|t\|/\|t\|)^{1/2}$. Using this and (9.4a), (9.4b) applied to $|u_{tt}|$, $|v_{tt}|$ analogously as above, we get

$$|u_s| \le |u_{tt}| + |V_u| \ll (M^2 + \varpi^2)\left(\frac{\log\|\tau\|}{\|\tau\|}\right)^{1/2}, \tag{11.6}$$

$$|v_s| \le |v_{tt}| + |V_v| \ll (M^2 + \varpi^2)\left(\frac{\log\|\tau\|}{\|\tau\|}\right)^{1/2}. \tag{11.7}$$

As by definition, for $\tau = \tilde{T}_k \pm L$, $\|\tau\| = L$, we deduce that for any $q \in \mathcal{B}_4$,

$$\tilde{F}_k(q) \le c_{13}(M^3 + \varpi^3)\frac{\log L}{L} \tag{11.8}$$

for some absolute constant $c_{13} > 0$. Now it is straightforward to check that

$$L \gg (\varpi^5 + M^5)\frac{|\log \Delta_1|}{\Delta_1} \gg (\varpi^3 + M^3)\log(\varpi^3 + M^3)\frac{|\log \Delta_1|}{\Delta_1}$$

suffices for the right-hand side of (11.8) to be $\ll \Delta_1$, which completes the proof for a large enough absolute constant c_{13}. □

Lemma 11.4. *There exists an absolute constant $c_{13} > 0$ such that if (11.2) holds, then for any $q = (u, v) \in \mathcal{B}_4$ and any $k \in \mathbb{Z}$, there exists $h \in \mathcal{C}$, so that*

$$|\tilde{E}_k(q) - E_{q_k}(h) - T_k(\omega_k^2 - \omega_{k+1}^2)/2 - (\omega_{k+1} - \omega_k)(\tilde{V}_k - V_k)| \le \Delta_1/2, \tag{11.9}$$

$$|\tilde{D}_k(q) - D_{q_k}(h)| \le \Delta_1/2. \tag{11.10}$$

Furthermore, for all $k \in \mathbb{Z}$,

$$|\tilde{E}_k(\tilde{q}_k) - E_{q_k}(0) - T_k(\omega_k^2 - \omega_{k+1}^2)/2 - (\omega_{k+1} - \omega_k)(\tilde{V}_k - V_k)| \le \Delta_1/2. \tag{11.11}$$

Proof. Fix $k \in \mathbb{Z}$. We define $\tilde{h}$ as in Lemma 9.2, and let $h(t) = \tilde{h}(t + \tilde{T}_k - T_k)$. We first show that $h \in \mathcal{C}$. As by definition $|T_k| \le 2\pi$, by inserting the definition of h in the bounds in Lemma 9.2, we easily see that (10.2) and (10.3) follow from Lemma 9.2, (iii) and (iv). Analogously we show that (10.4) follows from Lemma 9.2, (vi).

Denote by $X_k = T_k(\omega_k^2 - \omega_{k+1}^2)/2 + (\omega_{k+1} - \omega_k)(\tilde{V}_k - V_k)$ (a constant independent of q). By definitions, the partial integration to change the range of integration in the second line, and substitution $t \to t + \tilde{T}_k - T_k$ in the third line, and finally by using Lemma 9.2, (i), we obtain

$$\begin{aligned}
E_{q_k}(h) &= \int_{-\infty}^{0} L_{\omega_k}(q_k + h, (q_k)_t + h_t, t)dt + \int_0^\infty L_{\omega_{k+1}}(q_k + h, (q_k)_t + h_t, t) + (\omega_{k+1} - \omega_k)(v_k(0) + v^h(0)) \\
&= \int_{-\infty}^{T_k} L_{\omega_k}(q_k + h, (q_k)_t + h_t, t)dt + \int_{T_k}^\infty L_{\omega_{k+1}}(q_k + h, (q_k)_t + h_t, t) \\
&\qquad + (\omega_{k+1} - \omega_k)(V_k + v^h(T_k)) + (\omega_k^2 - \omega_{k+1}^2)T_k/2 \\
&= \int_{-\infty}^{\tilde{T}_k} L_{\omega_k}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t)dt + \int_{\tilde{T}_k}^\infty L_{\omega_{k+1}}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t)dt + (\omega_{k+1} - \omega_k)v(\tilde{T}_k) + X_k \\
&= \tilde{E}_k(q) + \int_{-\infty}^{\tilde{T}_k - L} L_{\omega_k}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t)dt + \int_{\tilde{T}_k + L}^\infty L_{\omega_{k+1}}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t)dt + X_k. \qquad (11.12)
\end{aligned}$$

By the definition of V, $\tilde{q}_k$, and then (6.8) and Lemma (9.2), (iii), we obtain

$$\begin{aligned}
\int_{-\infty}^{\tilde{T}_k - L} V(\tilde{q}_k + \tilde{h}, t)dt &\le \varepsilon \int_{-\infty}^{\tilde{T}_k - L} (1 - \cos(\tilde{q}_k + \tilde{h}))\|t\| \ll \varepsilon \int_{-\infty}^{T_k - L} q_k^2(t)dt + \varepsilon \int_{-\infty}^{\tilde{T}_k - L} \tilde{h}_k^2(t)dt \\
&\ll \varepsilon \int_{-\infty}^{T_k - L} e^{-\frac{1}{2}\sqrt{\varepsilon}|t - T_k|} + \varepsilon \int_{-\infty}^{\tilde{T}_k - L} e^{-\frac{1}{2}\sqrt{\varepsilon}|t - \tilde{T}_k|} \ll \sqrt{\varepsilon}e^{-\frac{1}{2}\sqrt{\varepsilon}L} \ll \frac{\log L}{L}.
\end{aligned}$$

Inserting it in the definition of L_{ω_k} and combining with (6.9), (6.10) and Lemma (9.2), (vi), we get

$$\int_{-\infty}^{\tilde{T}_k-L} L_{\omega_k}(\tilde{q}_k+\tilde{h},(\tilde{q}_k)_t+\tilde{h}_t,t)dt \ll ||(\tilde{u}_k)_t||^2_{L^2((-\infty,\tilde{T}_k-L])}+||(\tilde{v}_k)_t-\omega_k||^2_{L^2((-\infty,\tilde{T}_k-L])}$$
$$+||\tilde{h}_t||^2_{L^2([\tilde{T}_k-L(1+1/L),\tilde{T}_k-L])^2}+\frac{\log L}{L}$$
$$\ll (M^4+\varpi^4)\frac{\log L}{L}.$$

By proving an analogous statement for the second integral in (11.12) and combining all the relations above, we see that

$$|\tilde{E}_k(q)-E_{q_k}(h)-X_k| \ll (M^4+\varpi^4)\frac{\log L}{L}.$$

We complete the claim analogously as in Lemma 11.3. The proof for (11.10) is analogous, as we also by Lemmas 6.5 and 9.2 control the second derivatives of $\tilde{q}_k$ and $\tilde{h}$. We get (11.11) by inserting $h=0$ in (11.9). □

Proof of Proposition 11.1. Let c_{13} be the larger of the constants in Lemmas 11.2 and 11.4. Let for some $k\in\mathbb{Z}$, s_1 be the supremum of all the times s such that (11.1) holds. Then by continuity,

$$\tilde{E}_k(q)=\tilde{E}_k(\tilde{q}_k)+\Delta_0(q_k). \tag{11.13}$$

Combining (11.9), (11.11) and (11.13), we obtain $|E_{q_k}(h)-E_{q_k}(0)-\Delta_0(q_k)|\le\Delta_1$. By Proposition 10.1, we thus obtain $D_{q_k}(h)\ge\Delta_1$, so by (11.10), $\tilde{D}_k(q)\ge\Delta_1/2$. Inserting this and the bound from Lemma 11.3 in the action balance law (11.3), we get

$$\frac{d}{ds}\tilde{E}_k(q)\le-\Delta_1/2+\Delta_1/4\le-\Delta_1/4<0,$$

which is in contradiction with the assumption. To show $q^0\in\mathcal{B}_5$, we recall that q^0 and $\tilde{q}_k$ coincide on $[\tilde{T}_k-L,\tilde{T}_k+L]$. As we always have $\Delta_0(q_k)>0$, we conclude that $\tilde{E}_k(q^0)=\tilde{E}_k(\tilde{q}_k)<\tilde{E}_k(\tilde{q}_k)+\Delta_0(q_k)$, so $q^0\in\mathcal{B}_5$ by definition. □

12. Completion of construction of an invariant set

In this section we complete the construction of an invariant set $\mathcal{B}$ with respect to (1.2) by finally applying Lemma 1.4.

Proposition 12.1. *Assume ω_k is a sequence in a closed subset $[\omega^-,\omega^+]$ satisfying (6.3), and that L satisfies*

$$L\ge c_{14}\varpi^5\frac{|\log\Delta_1|}{\Delta_1} \tag{12.1}$$

for some large enough absolute constant c_{14}. Then there exists a ξ-invariant set $\mathcal{B}\subset\mathcal{B}_5\subset\mathcal{X}$, such that $q^0\in\mathcal{B}$.

Remark 12.1. Assume that as in the Remark 7.1 we fix a segment $[\omega^-,\omega^+]$ in a region of instability, and that for each $\omega,\tilde{\omega}\in[\omega^-,\omega^+]$ satisfying (6.3) we chose a single $q\in\mathcal{H}$ (as such q is not necessarily unique). We also fix $\mathcal{N}_q$ associated to such q as in the definition of $\mathcal{H}$. We will see that the set $\mathcal{B}$ is then uniquely defined by the choice of L, $(\tilde{L}_k)_{k\in\mathbb{Z}}$ and $(\omega_k)_{k\in\mathbb{Z}}$, and satisfies all the relations in the definitions of $\mathcal{B}_1$-$\mathcal{B}_5$. In the proofs of the main theorems, we thus use the notation $\mathcal{B}(L,(\tilde{L}_k)_{k\in\mathbb{Z}},(\omega_k)_{k\in\mathbb{Z}})$.

We say that q intersects the set $\mathcal{N}_k\subset\mathbb{R}^2$ at k, if there exists $(t,v)\in\mathcal{N}_k$ such that $q(t)=((2k+1)\pi,v)$. Analogously we define the notion of q intersecting $\partial\mathcal{N}_k$ at k. We first in two lemmas establish that $q\in\mathcal{B}_5$ can not intersect $\partial\mathcal{N}_k$ at k, then define $\mathcal{B}_6$ and prove its ξ-relative invariance, and finally complete the proof of Proposition 12.1.

Lemma 12.2. *There exists an absolute constant c_{15} such that, if*

$$L \geq c_{15} \frac{|\log \mu|}{\sqrt{\varepsilon}}, \tag{12.2}$$

then the following holds: for each $q \in \mathcal{B}_5$ intersecting $\partial \mathcal{N}_k$ at k, we have

$$\tilde{E}_k(q) \geq \tilde{E}_k(\tilde{q}_k) + \Delta_0. \tag{12.3}$$

Proof. To prove (12.3), we will approximate $q(t - T_k + \tilde{T}_k) - (2k\pi, \tilde{V}_k - V_k)$ with a $\tilde{q} = (\tilde{u}, \tilde{v}) \in H^1_{\text{loc}}$ so that we can apply (7.1). Indeed, let $\tilde{q}(t) = q(t - T_k + \tilde{T}_k) - (2\pi k, \tilde{V}_k - V_k)$ for $t \in [T_k - L, T_k + L]$. We define $\tilde{v}(t)$ uniquely for all $t \in \mathbb{R}$ by $\tilde{v}_t = \omega_k$ for $t < T_k - L$, $\tilde{v}_t = \omega_{k+1}$ for $t > T_k + L$, $\tilde{v}$ continuous. Let

$$\tilde{u}(t) = a \cdot e^{-\frac{1}{2}\sqrt{\varepsilon}|t - T_k|}, \ t \leq T_k - L, \qquad \tilde{u}(t) = 2\pi - b \cdot e^{-\frac{1}{2}\sqrt{\varepsilon}|t - T_k|}, \ t \geq T_k + L,$$

where the constants a, b are uniquely chosen so that $\tilde{u}$ is continuous. By construction, (8.1) and the definition of V, we have

$$\begin{aligned} |\tilde{E}_k(q) - \mathcal{L}_{\omega_k, \omega_{k+1}}(\tilde{q})| &= \int_{-\infty}^{T_k - L} L_{\omega_k}(\tilde{q}, \tilde{q}_t, t) dt + \int_{T_k + L}^{\infty} L_{\omega_{k+1}}(\tilde{q}, \tilde{q}_t, t) dt \\ &\ll \varepsilon \int_{-\infty}^{T_k - L} e^{-\frac{1}{2}\sqrt{\varepsilon}|t - T_k|} dt + \varepsilon \int_{T_k + L}^{\infty} e^{-\frac{1}{2}\sqrt{\varepsilon}|t - T_k|} dt \ll \sqrt{\varepsilon} e^{-\frac{1}{2}\sqrt{\varepsilon} L} \leq e^{-\frac{1}{2}\sqrt{\varepsilon} L}. \end{aligned} \tag{12.4}$$

Similarly, we obtain

$$|\tilde{E}_k(\tilde{q}_k) - \mathcal{L}_{\omega_k, \omega_{k+1}}(q_k)| \ll e^{-\frac{1}{2}\sqrt{\varepsilon} L}. \tag{12.5}$$

By (5.2) and the definition of Δ_0, we have $\Delta_0 \ll \sqrt{\varepsilon}\, \mu \leq \mu$. Thus we can choose $L \gg |\log \mu| / \sqrt{\varepsilon} \gg |\log \Delta_0| / \sqrt{\varepsilon}$ for a large enough absolute constant, so that the left-hand sides od (12.4) and (12.5) are bounded by $\Delta_0/2$. Combining it with (7.1), we complete the proof. ☐

Lemma 12.3. *If $q \in \mathcal{B}_5$, then for all $k \in \mathbb{Z}$, q can not intersect $\partial \mathcal{N}_k$ at k.*

Proof. As $q \in \mathcal{B}_5$, (11.1) holds. If q would intersect $\partial \mathcal{N}_k$ at k, this would contradict (12.3) and $\Delta_0(\tilde{q}_k) \leq \Delta_0/2$ established in Proposition 10.1. ☐

We now define the number of times q intersects of $\mathcal{N}_k$ at k. Let Υ_k (as a function of $q = (u, v) \in \mathcal{X}$) be the set of all $t \in \mathbb{R}$ satifying

$$\Upsilon_k = \{\tilde{T}_{k-1} \cup \tilde{T}_{k+1}\} \cup \{t \in [\tilde{T}_{k-1}, \tilde{T}_{k+1}], (t, v(t)) \in \partial B_k\}.$$

By the definition of $\mathcal{A}$, $u(\tilde{T}_{k-1}) < (2k+1)\pi$ and $u(\tilde{T}_{k+1}) > (2k+1)\pi$. Now by Lemma (12.3), for any $t \in \Upsilon_k$, $u(t) \neq (2k+1)\pi$. Let $\sim_k$ be a relation of equivalence on Υ_k defined with $t_1 \sim_k t_2$ whenever for all $t_3 \in \Upsilon_k$ such that $t_1 \leq t_3 \leq t_2$, we have that $u(t_1) - (2k+1)\pi$, $u(t_2) - (2k+1)\pi$ and $u(t_3) - (2k+1)\pi$ have the same sign. Let $\tilde{\Upsilon}_k = \Upsilon_k / \sim_k$ with the induced topology. As by assumptions, Υ_k is a closed subset of a compact set, Υ_k is compact. By definition, continuity of q and compactness of $\partial \mathcal{N}_k$, we see that $\tilde{\Upsilon}_k$ is totally disconnected and compact, thus finite, and $|\tilde{\Upsilon}_k| \geq 2$. Consider $|\tilde{\Upsilon}_k| - 1$ intervals $(t_j, t_{j+1}) \in \Upsilon_k^c$, where $t_j, t_{j+1} \in \Upsilon_k$ and $t_j \not\sim_k t_{j+1}$. We say that q intersects $\mathcal{N}_k$ at k exactly m-number of times, if m is the number of such intervals, for which q intersects $\mathcal{N}_k$ at k for some $(t, v) \in \mathcal{N}_k$ such that $t \in (t_j, t_{j+1})$.

Let $\mathcal{B}_6$ be the set of all $q \in \mathcal{B}_5$, such that for each $k \in \mathbb{Z}$, q intersects $\mathcal{N}_k$ at k odd number of times (i.e. that m in the definition of the number of intersections is odd).

Lemma 12.4. *The set $\mathcal{B}_6$ is $\mathcal{A}$-relatively ξ-invariant, and $q^0 \in \mathcal{B}_6$.*

The proof relies on somewhat subtle topological considerations, and is postponed to the Appendix D. We note that the proof uses in a fundamental way the existence of a continuous semiflow which solves (1.2).

Proof of Proposition 12.1. Assume that L satisfies (11.2) and (12.2), and that M is given as in (10.1). Let $\tilde{\mathcal{B}} = \mathcal{B}_6$ and $\mathcal{B} = \tilde{\mathcal{B}} \cap \mathcal{A}$. It suffices to show that $q^0 \in \mathcal{B}$, and that the conditions (B1), (B2) hold, as the claim will then hold by Lemma 1.4. We have already shown (B1) and $q^0 \in \mathcal{B}$ in Lemma 12.4. We now show (B2). The smoothness requirement in the definition of $\mathcal{A}$ follows from Theorem 2.1, (iv). We now show that (7.7) and (7.8) hold.

Let $q \in \mathcal{B}_6$, fix $k \in \mathcal{B}_6$, and find $(t, v(t)) \in \tilde{\mathcal{N}}_k$ such that $u(t) = (2k+1)\pi$ (this exists by the definition of $\mathcal{B}_6$). By definition of $\tilde{\mathcal{N}}_k$, $\tilde{T}_k$ and $\tilde{V}_k$ we have that $|t - \tilde{T}_k| \le R$, $|v(t) - \tilde{V}_k| \le R$. Without loss of generality, let $t \le \tilde{T}_k$, $v(t) \le \tilde{T}_k$ (the other cases are analogous). Then by the definition of Δ_0 and (5.2),

$$\int_t^{\tilde{T}_k} \left(u_t^2 + (v_t - \omega_k)^2\right) dt \le \max_{(t_0,v_0)\in\mathbb{R}^2} S_\omega(t_0, v_0) \le 8\sqrt{\varepsilon(1+\mu)} \le 9\sqrt{\varepsilon}.$$

We thus have

$$|u(\tilde{T}_k) - (2k+1)\pi| \le \int_t^{\tilde{T}_k} |u_t| dt \le R^{1/2} \left(\int_t^{\tilde{T}_k} u_t^2 dt\right)^{1/2} \le 3R^{1/2}\varepsilon^{1/4},$$

$$\begin{aligned} |v(\tilde{T}_k) - \tilde{V}_k| &\le R + \int_t^{\tilde{T}_k} |v_t| dt \le R(1+|\varpi|) + R^{1/2}\left(\int_t^{\tilde{T}_k} (v_t - \omega_k)^2 dt\right)^{1/2} \\ &\le R(1+|\varpi|) + 3R^{1/2}\varepsilon^{1/4}. \end{aligned}$$

By the assumptions (1.5) and (10.1), we thus have $|u(\tilde{T}_k) - (2k+1)\pi| \le 1/4$, $|v(\tilde{T}_k) - \tilde{V}_k| \le M/2$. Let $q(s_0) = q$, and consider the solution $q(s)$ of (1.2) with the initial condition $q(s_0)$ at $s = s_0$. By (11.7) and the definition of M in (10.1), we have $|u_s| \ll_{\varepsilon,\varpi} 1$, $|v_s| \ll_{\varepsilon,\varpi} 1$, thus the upper bound on $|u_s(t)|$, $|v_s(t)|$ is independent of the choice of $t = \tilde{T}_k$ and $q = (u, v) \in \mathcal{B}_6$. We conclude that for any $\tilde{\lambda} > 0$, there exists $\lambda > 0$, independent of the choice of $q(s_0) \in \mathcal{B}_6$, such that for all $s \in [s_0, s_0+\lambda]$, we have $|u(\tilde{T}_k) - (2k+1)\pi| \le 1/4 + \tilde{\lambda}$, $|v(\tilde{T}_k) - \tilde{V}_k| \le M/2 + \tilde{\lambda}$. By the definition of $\mathcal{A}$, this proves (B2). We conclude that $\mathcal{B} := \mathcal{B}_6 \cap \mathcal{A}$ is indeed ξ-invariant.

It remains to show that (12.1) suffices for (11.2) and (12.2) to hold. By definition of M in (10.1), $c(\varpi) \ll \varpi^5$, thus (11.2) is satisfied. As by definition of Δ_1 in Proposition 10.1, (7.2) and (A2), we know that $\Delta_1 \le \Delta_0/2 \le 4\sqrt{\varepsilon}\mu \le \sqrt{\varepsilon}$, we deduce $|\log\mu|/\sqrt{\varepsilon} \ll |\log\sqrt{\varepsilon}|/\sqrt{\varepsilon} \le |\log\Delta_1|/\Delta_1$, which was required. ☐

III: PROOFS OF THE MAIN THEOREMS

13. Proofs of the shadowing Theorem, Theorems 1.1 and 1.2

We first prove a "classical" shadowing theorem, showing existence of a solution of (1.3) shadowing an arbitrary sequence of tori $\mathbb{T}_\omega$, and then Theorems 1.1 and 1.2. Prior to all of it, we establish useful a-priori bounds on the derivatives of $q \in \mathcal{B}$. The constant c_0 may change from line to line within the section.

Lemma 13.1. *Assume $q \in \mathcal{B}$. Then for all $t \in \mathbb{R}$ and some absolute $c_0 > 0$,*

$$|u_t(t)| \le c_0\varpi\left(\sqrt{\varepsilon} \wedge \left(\frac{\log\|t\|}{\|t\|}\right)^{1/2}\right), \qquad |v_t(t) - \omega_{k(t)}| \le c_0\varpi\left(\sqrt{\varepsilon} \wedge \left(\frac{\log\|t\|}{\|t\|}\right)^{1/2}\right). \tag{13.1}$$

Specifically, (1.7) holds.

Proof. We first establish the following fact: assume $w \in H^1([\tau, \tau + 4/\sqrt{\varepsilon}])$ and $A > 0$ a constant, such that

$$||w||^2_{L^2([\tau,\tau+4/\sqrt{\varepsilon}])} \le A\sqrt{\varepsilon}, \qquad ||w_t||^2_{L^2([\tau,\tau+4/\sqrt{\varepsilon}])} \le A\varepsilon^{3/2}. \tag{13.2}$$

Then we have $||w||_{L^\infty([\tau,\tau+4/\sqrt{\varepsilon}])} \ll A^{1/2}\sqrt{\varepsilon}$ (for some absolute implicit constant). Indeed, by substitution $\tilde{w}(t) = w(\tau + 4t/\sqrt{\varepsilon})$, we obtain by direct calculation $||\tilde{w}||_{H^1([0,1])} \ll A^{1/2}\sqrt{\varepsilon}$, which implies $||\tilde{w}||_{L^\infty([0,1])} \ll A^{1/2}\sqrt{\varepsilon}$, thus the claim.

By (7.11), for any $\tau \in \mathbb{R}$ we have that $||u_t^0||^2_{L^2([\tau,\tau+4/\sqrt{\varepsilon}])} \ll \sqrt{\varepsilon}$ and $||v_t^0 - \omega_{k(.)}||^2_{L^2([\tau,\tau+4/\sqrt{\varepsilon}])} \ll \sqrt{\varepsilon}$. Inserting that in (9.3a) and (9.3b), and in the view of the definition (10.1) of M and that $\lambda(\tau) \leq \sqrt{\varepsilon}/4$, we get that for any $\tau \in \mathbb{R}$ the relations (13.2) hold with $w = u_t$, and also with $w = v_t - \omega_{k(.)}$, with $A = c_0\varpi^2$, for some absolute c_0. This yields the $c_0\varpi\sqrt{\varepsilon}$ upper bound in (13.1). The other bound has already been established (by an analogous argument) in (11.4) and (11.5). □

Theorem 13.2. *Let $[\omega^-, \omega^+]$ be such that (S1) holds. Let ω_k, $k \in \mathbb{Z}$ be a sequence in $[\omega^-, \omega^+]$, and choose arbitrarily small $\delta_k > 0$, $k \in \mathbb{Z}$. Then there exists $q = (u, v) \in \mathcal{E}$ and a sequence of times t_k such that for all $k \in \mathbb{Z}$*

$$|u(t_k) - 2k\pi| < \delta_k, \quad |u_t| < \delta_k, \quad |v_t(t_k) - \omega_k| < \delta_k. \tag{13.3}$$

Proof. We set L large enough so that (12.1) holds. Choose $\tilde{L}_k$ so that $\tilde{L}_{k+1} - \tilde{L}_k \geq 4L \vee c_{16}\varpi^2|\log\delta_k|^2/\delta_k^2 + 2\pi$ for some $c_{16} > 0$ to be determined later, and then $L_k \geq 4L \vee c_{16}\,\varpi^2|\log\delta_k|^2/\delta_k^2$. Let $q^0(L, (\tilde{L}_k)_{k\in\mathbb{Z}}, (\omega_k)_{k\in\mathbb{Z}})$ as in Remark 7.1, and let $\mathcal{B} = \mathcal{B}(L, (\tilde{L}_k)_{k\in\mathbb{Z}}, (\omega_k)_{k\in\mathbb{Z}})$ as in Remark 12.1. Then by Proposition 12.1, $\mathcal{B}$ is ξ-invariant, non-empty, and by Theorem 3.1, there is a $q \in \mathcal{E}$ which is also in the closure of $\mathcal{B}$ in $\mathcal{X}_{\text{loc}}$. Let $t_k = (\tilde{T}_{k-1} + \tilde{T}_k)/2) = \tilde{T}_{k-1} + L_k/2$. By (8.1) (the bound on u) and (13.1) (the bounds on u_t, v_t), inserting $\tau = t_k$, thus $||\tau|| = L_k/2$, it follows that we can choose an absolute constant c_{16} large enough so that (13.3) holds. □

We prove first Theorem 1.2 as the construction is simpler and illustrative, and then Theorem 1.1.

Proof of Theorem 1.2, (i). Let $L \equiv 0 \mod 2\pi$ be such that $c_{17}\frac{|\log\mu|}{\sqrt{\varepsilon}} \leq L \leq c_{17}\frac{|\log\mu|}{\sqrt{\varepsilon}} + 2\pi$ for some $c_{17} \geq c_{15}$, i.e. such that (12.1) holds, for some absolute c_{17} to be determined later.

To prove (i), we will embed the standard Bernoulli shift in $\hat{\mathcal{X}}$ in such a way that the Theorem 4.4 can be applied by using the construction of ξ-invariant sets in Proposition 12.1. Let $(\Omega_0, \mathcal{F}_0, \mu_0, s)$ be the standard Bernoulli shift, where Ω_0 is the set of all $\chi = (\chi_j)_{j\in\mathbb{Z}}$, $\chi_j \in \{0,1\}$, $\mathcal{F}_0$ is the σ-algebra on Ω_0 induced by finite cylinders, μ_0 is the product measure, where 0 and 1 have the same probability 1/2, and $s : \Omega_0 \to \Omega_0$ is the right shift. The mapping $\iota_0 : \Omega_0 \to \mathcal{X}$ and the induced map $\hat{\iota}_0 = \iota \circ \iota_0 : \Omega_0 \to \hat{\mathcal{X}}$ is defined as follows: we associate to each $\chi \in \Omega_0$ a sequence $\omega_k(\chi)$ and $\tilde{L}_k(\chi)$, $k \in \mathbb{Z}$. We then define $q_\chi = q^0(L, (\tilde{L}_k(\chi))_{k\in\mathbb{Z}}, (\omega_k(\chi))_{k\in\mathbb{Z}})$ as in Remark 7.1, and the ξ-invariant sets $\mathcal{B}_\chi$ containing q_χ, $\mathcal{B}_\chi = \mathcal{B}(L, (\tilde{L}_k(\chi))_{k\in\mathbb{Z}}, (\omega_k(\chi))_{k\in\mathbb{Z}})$ as in Remark 12.1. Let $\hat{q}_\chi$, $\hat{\mathcal{B}}_\chi$ be their embeddings in the quotient set $\hat{\mathcal{X}}$.

Choose arbitrary $\omega_k = \omega \in [\omega^-, \omega^+]$ fixed for all $k \in \mathbb{Z}$. We set $\tilde{L}_k$ so that q_χ has a "jump" at $T = 4nL$ if and only if $\chi_n = 0$. More precisely, let $\tilde{\Omega}_0$ be the subset of Ω_0 of χ with infinitely many 0 (clearly measurable and $\mu_0(\tilde{\Omega}_0) = 1$), and for $\chi \in \tilde{\Omega}_0$, let $k_j(\chi)$ be the increasing sequence of integers of "positions" of zeros in χ, uniquely defined by the requirement $k_0(\chi) \leq 0$, $k_1(\chi) \geq 1$. We now set $\tilde{L}_j(\chi) = 4L \cdot k_j(\chi)$, and complete the definition of $\hat{q}_\chi$, $\mathcal{B}_\chi$. By the definition of the induced localized topology on $\hat{\mathcal{X}}$, $\hat{\iota_0} : \tilde{\Omega}_0 \to \hat{\mathcal{X}}$ is continuous (assuming the product topology on Ω_0, $\tilde{\Omega}_0$), thus measurable. Also note that by construction, $\hat{q}_{s(\chi)} = \hat{S}^{4L/2\pi}(\hat{q}_\chi)$, $\hat{\mathcal{B}}_{s(\chi)} = \hat{S}^{4L/2\pi}(\hat{\mathcal{B}}_\chi)$. Let $\tilde{\mu} = (\hat{\iota}_0)^*\mu_0$ be the pulled measure, and let $\tilde{\mathcal{M}}_1 = \hat{\iota}_0(\tilde{\Omega}_0)$.

As the entropy $h_{\mu_0}(s) = \log 2$ [47], by construction we have $h_{\tilde{\mu}}(\hat{S}^{4L/2\pi}) = \log 2$. Finally we define a measure μ "to be shadowed" by

$$\mu = \sum_{n=0}^{4L/2\pi-1} (\hat{S}^n)^*\tilde{\mu}, \qquad \mathcal{M}_1 = \bigcup_{n=0}^{4L/2\pi-1} \hat{S}^n(\tilde{\mathcal{M}}_1).$$

By construction, μ is $\hat{S}$-invariant. By [47], Theorem 4.13, (i), $h_\mu(\hat{S}) = 2\pi\log 2/(4L)$.

To apply Theorem 4.4 and establish existence of a shadowing measure ν, we need to construct a σ-subalgebra $\mathcal{G}$. First we define the sets $\mathcal{D}_i \subset \hat{\mathcal{X}}$, $i \in \mathcal{I}$, where $\mathcal{I} = \{(j,k,n),\ j \in \{0,1\},\ k \in \{0,1,...,4L/2\pi - 1\},\ n \in \mathbb{Z}\}$, as the set of all $q = (u,v) \in \hat{\mathcal{X}}$ satisfying the following conditions

$$u(4nL + 2k\pi) \in \begin{cases} [\pi - 1, \pi + 1] \mod 2\pi & j = 0, \\ [-1, 1] \mod 2\pi & j = 1. \end{cases} \tag{13.4}$$

Now let $\mathcal{G}$ be the σ-subalgebra generated by $\mathcal{D}_i$ $i \in \mathcal{I}$. As each $q \in \mathcal{M}_1$ can be represented as $\hat{S}^{k_0} q_\chi$ for some $k_0 = 0, ..., 4L/2\pi - 1$, $\chi \in \tilde{\Omega}_0$, we define

$$\mathcal{D}_q = \bigcap_{n=-\infty}^{\infty} \left\{ \mathcal{D}_{\chi_n, k_0, n} \bigcap_{\substack{k=0 \\ k \neq k_0}}^{4L/2\pi - 1} \mathcal{D}_{0,k,n} \right\}, \qquad \mathcal{B}_q = \hat{S}^k \mathcal{B}_\chi.$$

By construction we can choose an absolute constant $c_{17} \geq c_{15}$ large enough so that $\mathcal{B}_q \subset \mathcal{D}_q$ for all $q \in \mathcal{M}_1$. The reason is as follows: (i) in the case $j = 0$ in (13.4), because of the relations (7.7), (8.3) and the properties of z^+, z^- (as there is some $\tilde{T}_m$ such that $|\tilde{T}_m - (4nL + 2k\pi)| \leq 2\pi$), and (ii) in the case $j = 1$, because of (8.3) and the exponentially fast decay of z^+, z^- towards $0 \mod 2\pi$. Thus by Proposition 12.1, for all $q \in \mathcal{M}_1$ and all $s \geq 0$, $\hat{\xi}^s(q) \in \mathcal{D}_q$. By construction we have that if $q \in \mathcal{D} \cap \mathcal{M}_1$ and $\mathcal{D} \in \mathcal{G}$, then $\mathcal{D}_q \subset \mathcal{D}$. This completes the proof of the condition (M2). By construction, (M1), (M3) hold. To verify (M4), it suffices to note that for fixed k, n, $\mathcal{D}_{0,k,n}$ and $\mathcal{D}_{1,k,n}$ are disjoint and cover the entire $\mathcal{M}_1$. As this completes the proof of conditions (M1)-(M4), we can apply Theorem 4.4 and obtain a $\nu \in \mathcal{M}(\hat{\mathcal{E}})$ which shadows μ. By Lemma 4.3 we get $h_\nu(\hat{S}) \geq h_\mu(\hat{S}) \sim 1/L$. The claim now follows from (12.1) and the variational principle for the topological and metric entropy [47]. □

Proof of Theorem 1.2, (ii). Choose $L \equiv 0 \mod 2\pi$ satisfying $c_{15}\frac{|\log \mu|}{\sqrt{\varepsilon}} \leq L \leq c_{15}\frac{|\log \mu|}{\sqrt{\varepsilon}} + 2\pi$, i.e. such that (12.1) holds. Choose an increasing sequence $\omega_1 = \omega^-$, ..., $\omega_n = \omega^+$, such that $\omega_k - \omega_{k-1} \leq \Delta_0/(4c_4(R \vee \mu)\varpi)$ for $k = 2, ..., n$ (as required by (6.3)), with equality for all except possibly $k = n$. Choose ω_k for $k \notin \{1, ..., n\}$ in an arbitrary way as long as $\omega_k \in [\omega^-, \omega^+]$ and as long as (6.3) holds. Let $\tilde{L}_{k+1} - \tilde{L}_k = 4L + 2\pi$ for all k except $k = 1$ and $k = n$ for which we set

$$\tilde{L}_{k+1} - \tilde{L}_k = c_0 \varpi^2 \frac{|\log \delta|^2}{\delta^2} + 4L + 2\pi, \tag{13.5}$$

for c_0 large enough to be chosen later. By Proposition 12.1, $\mathcal{B} = \mathcal{B}(L, (\tilde{L}_k(\chi))_{k\in\mathbb{Z}}, (\omega_k(\chi))_{k\in\mathbb{Z}})$ is ξ-invariant, and by Theorem 3.1, there is a $q \in \mathcal{E}$ which lies in the closure of $\mathcal{B}$ in $\mathcal{X}_{\text{loc}}$.

Let $t^- = (\tilde{T}_1 + \tilde{T}_2)/2 = \tilde{T}_1 + L_1/2$, and $t^+ = (\tilde{T}_n + \tilde{T}_{n+1})/2 = \tilde{T}_n + L_n/2$. Now by (13.1), as $||t^-|| = L_1/2$, $||t^+|| = L_n/2$, as $L_1, L_n \geq c_0 \varpi^2 \frac{|\log \delta|^2}{\delta^2}$, and finally as $\omega_{k(t^-)} = \omega_1 = \omega^-$ and $\omega_{k(t^+)} = \omega_n = \omega^+$; we can choose c_0 in (13.5) to be a a large enough absolute constant such that $|v_t(t^-) - \omega^-| \leq \delta$, $|v_t(t^+) - \omega^+| \leq \delta$. Now as $n \sim \varpi(R \vee \mu)(\omega^+ - \omega^-)/\Delta_0$, and by (12.1) and (13.5),

$$\begin{aligned} |t^+ - t^-| &\leq \sum_{k=1}^n L_k \ll \varpi^2 \frac{|\log \delta|^2}{\delta^2} + nL \ll \varpi^2 \frac{|\log \delta|^2}{\delta^2} + \varpi^6 \frac{|\log \Delta_1|(R \vee \mu)}{\Delta_0 \Delta_1}(\omega^+ - \omega^-) \\ &\ll \varpi^6 \tilde{c}(\delta) \frac{|\log \Delta_1|(R \vee \mu)}{\Delta_0 \Delta_1}(\omega^+ - \omega^-), \end{aligned}$$

where we can set $\tilde{c}(\delta) = |\log \delta|^2/\delta^2$ as $nL \gg \varpi^2$. The proof is completed. □

Proof of Theorem 1.1. We first complete the proof in the case when the entire $\mathcal{O} \subseteq [\tilde{\omega}^-, \tilde{\omega}^+] \subseteq [\omega^-, \omega^+]$, such that $\tilde{\omega}^+ - \tilde{\omega}^- \leq \Delta_0/(4c_4(R \vee \mu) \cdot \varpi)$ (i.e. we can by Proposition 6.1 connect any two ω, $\tilde{\omega}$ in $\mathcal{O}$ with a heteroclinic orbit with only one "jump"). The only modification as compared to the proof of Theorem 1.2 is now the choice of the sequence ω_k. Let $\{\tilde{\omega}_j,\ j \in \mathcal{P}\}$, $\mathcal{P}$ finite or countable, be a dense subset of $\mathcal{O}$. Let $p : \mathbb{N} \to \mathcal{P}$ be any function such that for each $j \in \mathcal{P}$, $p^{-1}(j)$ is infinite (constructed e.g. by a diagonalization procedure). By keeping the notation as in the proof of Theorem 1.2, (i), for a given χ we define $\omega_j = \tilde{\omega}_{p(k_{j+1} - k_j)}$ (where by definition $k_{j+1} - k_j - 1$ is the number of consecutive "ones" between the zeroes at positions j and $j+1$ in a chosen $\chi \in \tilde{\Omega}_0$). The rest of the construction is analogous to the proof of Theorem 1.1, by which we obtain $\nu \in \mathcal{M}(\hat{\mathcal{E}})$, which is supported on the closure of the union of ξ-invariant sets $\mathcal{B}(L, (\tilde{L}_k)_{k\in\mathbb{Z}}, (\omega_k(\chi))_{k\in\mathbb{Z}})$ (considered as subsets of $\hat{\mathcal{X}}$), for $\chi \in \Omega_0$. It is easy to check by construction and (13.1) that supp $\hat{\pi}^*\nu$ intersects $\mathbb{T}_\omega$ for every $\omega \in \mathcal{O}$ (which implies the first part of (1.6)), and that for every $\omega \notin \mathcal{O}$, supp $\hat{\pi}^*\nu \cap \mathbb{T}_\omega = \emptyset$.

Consider now the general case, and let n be an integer such that

$$n \geq \frac{4c_4(R \vee \mu) \cdot \varpi}{\Delta_0}(\omega^+ - \omega^-)$$

(i.e. n is the minimal number of "jumps" required by the Proposition 12.1 and (6.3) to cross the entire $[\omega^-, \omega^+]$). Consider the subshift of finite type $(\Omega_1, \mathcal{F}_1, \mu_1, s)$ ($(\Omega_1, \mathcal{F}_1)$ a subspace of $(\Omega_0, \mathcal{F}_0)$), μ_1 the s-invariant probability measure, where we "allow" only sequences with no less than n consecutive zeros. The only change in the construction above is that we associate to each consecutive sequence of $m \geq n$ zeros in a chosen $\chi \in \Omega_1$ a sequence ω_1, ω_2, ..., ω_m in an arbitrary, s-invariant way, such that (6.3) holds for any $\omega = \omega_k$, $\tilde{\omega} = \omega_{k+1}$, $k \in \mathbb{Z}$ (this is possible by the choice of n). The rest of the proof is analogous. □

Remark 13.1. We could have sharpened the definition of the subalgebra $\mathcal{G}$ in the proof of Theorem 1.1, to be able to use the notion of the conditional support from Section 4. More specifically, we could obtain that $\hat{\pi}(\text{supp}(\nu|\mathcal{G}))$ intersects all $\mathbb{T}_\omega$, $\omega \in \mathcal{O}$, and that $\cup_{\omega \in \mathbb{R}-\mathcal{O}} \subset \hat{\pi}(\text{supp}^c(\nu|\mathcal{G}))$. This would, however, unnecessarily complicate the proof.

14. Proof of Theorem 1.3

In this section we first state an analogue of Lemma 5.3 which will enable us more precise control for small μ, then evaluate the error term when approximating S_ω with M_ω, and finally we prove Theorem 1.3.

Lemma 14.1. *There exist a constant $\tilde{T} > 0$ and functions $\tilde{z}^- : [-\tilde{T}, 0] \to \mathbb{R}$ and $\tilde{z}^+ : [0, \tilde{T}] \to \mathbb{R}$, depending only on ε, μ, satisfying for all t in the domain of definition:*

(i) $\tilde{z}^-$, $\tilde{z}^+$ are continuous and C^2 in the interior of the domain,

(ii) $\tilde{z}^-(0) = \tilde{z}^+(0) = \pi$, $\tilde{z}^-(-\tilde{T}) = -\sqrt{\varepsilon\mu}$, $\tilde{z}^+(\tilde{T}) = 2\pi + \sqrt{\varepsilon\mu}$,

(iii) $\tilde{z}^-$ is a strict stationary sub-solution on $(-\tilde{T}, 0)$ of (1.2a), and $\tilde{z}^+$ is a strict stationary super-solution (1.2a) on $(0, \tilde{T})$. Furthermore, for any constant $T \in \mathbb{R}$, $\tilde{z}^-(t+T)$ and $\tilde{z}^+(t+T)$ are strict stationary sub-, resp. super-solutions in the interior of their domain,

(iv) $\tilde{z}^-$, $\tilde{z}^+$ are strictly increasing,

(v) There exists an absolute constant $c_{18} > 0$ such that

$$|\tilde{z}^-(t) - u^{(\varepsilon)}(t)| \leq c_{18}\sqrt{\varepsilon\mu},\ t \in [-\tilde{T}, 0], \quad |\tilde{z}^+(t) - u^{(\varepsilon)}(t)| \leq c_{18}\sqrt{\varepsilon\mu},\ t \in [0, \tilde{T}], \tag{14.1}$$

where $u^{(\varepsilon)}(t) = 4 \operatorname{arctg} e^{\sqrt{\varepsilon}t}$ is the separatrix solution in the case $\mu = 0$.

The proof is a relatively straightforward modification of the proof of Lemma 5.3, thus also done in the Appendix B.

Lemma 14.2. *(i) There exists an absolute constant $c_{19} > 0$ such that if $q = (u, v)$ is a two-sided minimizer at $(\omega, t_0, v_0) \in \mathbb{R}^3$, then for all $t \in \mathbb{R}$,*

$$|u(t) - u^{(\varepsilon)}(t - t_0)| \leq c_{19}\sqrt{\varepsilon\mu}, \tag{14.2}$$

(ii) For all $(t_0, v_0), (t_1, v_1) \in \mathbb{R}^2$, and all $\omega \in \mathbb{R}$,

$$|S_\omega(t_1, v_1) - S_\omega(t_0, v_0) - \mu(M_\omega(t_1, v_1) - M_\omega(t_0, v_0))| = O_f(\varepsilon\mu^{3/2}). \tag{14.3}$$

Proof. We first note that analogously to the proof of Proposition 5.4, (ii) we can show that for all $t \in [t_0 - \tilde{T}, t_0]$, we have $u(t) \geq \tilde{z}^-(t - t_0)$, thus by Proposition 5.4, (ii), $\tilde{z}^-(t - t_0) \leq u(t) \leq z^-(t - t_0)$. We deduce that

$$|u(t) - u^{(\varepsilon)}(t - t_0)| \leq |\tilde{z}^-(t - t_0) - u^{(\varepsilon)}(t - t_0)| + |z^-(t - t_0) - u^{(\varepsilon)}(t - t_0)|$$

which implies (14.2) by and (5.6) and (14.1). The case $t < -\tilde{T} + t_0$ follows from $\tilde{z}^-(\tilde{T}) = -\sqrt{\varepsilon\mu}$ and $0 \leq u(t) \leq z^-(t - t_0)$, $\tilde{z}^-$ increasing. The proof for $t \geq t_0$ is analogous, which completes (i).

Let $q = (u, v) \in H^1_{\rm loc}(\mathbb{R})^2$ be a two-sided minimizer at (ω, t_0, v_0), and let $\tilde{q}(t) = (u(t+t_0-t_1), v(t+t_0-t_1)+ v_1 - v_0)$ (i.e. we translate q in t, v so that $\tilde{q}(t_1) = (\pi, v_1)$). As by definition, $\int_{-\infty}^{\infty} L_\omega(\tilde{q}, \tilde{q}_t, t)dt \geq S_\omega(t_1, v_1)$, we by straightforward calculation have

$$S_\omega(t_1, v_1) - S_\omega(t_0, v_0) - \mu(M_\omega(t_1, v_1) - M_\omega(t_0, v_0)) \leq 2\varepsilon\mu \left| \int_{-\infty}^{\infty} (1-\cos u) f(u, v, t)dt \right.$$

$$\left. - \int_{-\infty}^{\infty} (1 - \cos(u^\varepsilon(t-t_0)))\, f(u^\varepsilon(t-t_0), v_0 + \omega(t-t_0), t)dt \right|. \tag{14.4}$$

By the Mean Value Theorem, (6.9), by the left-sides of (6.10), (6.11) (which also hold for two-sided minimizers which are not necessarily homoclinics) and (14.2), the absolute value of the difference of the integrands in (14.4) is $\ll_f e^{-\frac{1}{2}\sqrt{\varepsilon}|t-t_0|}\sqrt{\varepsilon\mu}$, thus (14.4) is $\ll_f \varepsilon\mu^{3/2}$. The other inequality in (14.4) is obtained analogously, by starting with a two-sided minimizer at (ω, t_1, v_1). ☐

Proof of Theorem 1.3. This follows from (14.3), as we choose μ_0 to be small enough, so that for $\mu \leq \mu_0$, the $O_f(\varepsilon\mu^{3/2})$ term is $\leq \mu\tilde{\Delta}_0$. Then for all $0 < \mu \leq \mu_0$, (S1) holds with $\Delta_0 = \mu\tilde{\Delta}_0$. ☐

IV: APPENDICES

15. Appendix A: The function spaces and existence of solutions

This section is dedicated to the proof of Theorem 2.1. We first recall definition of the required function spaces (see [19] and references therein for details), then prove the theorem in four separate lemmas: on local existence and uniqueness, global existence, continuous dependence on initial conditions and regularity of solutions.

Let $\varphi^y(q)(t) = q(t+y)$ be the translation, $y \in \mathbb{R}$. The uniformly local norms and spaces are given with:

$$\begin{aligned}
||q||_{L^2_{\rm ul}(\mathbb{R})^N} &= \sup_{y\in\mathbb{R}} \left(\int_{\mathbb{R}} e^{-|t+y|} q(t)^2 \right)^{1/2}, \\
L^2_{\rm ul}(\mathbb{R})^N &= \left\{ q \in L^2_{\rm loc}(\mathbb{R})^N,\ ||q||_{L^2_{\rm ul}(\mathbb{R})^N} < \infty,\ \lim_{y\to 0} ||\varphi^y q - q||_{L^2_{\rm ul}(\mathbb{R})^N} = 0 \right\}, \\
H^k_{\rm ul}(\mathbb{R})^N &= \left\{ u \in L^2_{\rm ul}(\mathbb{R})^N \mid \partial_t^j u \in L^2_{\rm ul}(\mathbb{R})^N \text{ for all } j \leq k \right\}, \\
||q||_{H^k_{\rm ul}(\mathbb{R})^N} &= \left(\sum_{j=0}^{k} ||\partial_t^j q||^2_{L^2_{\rm ul}(\mathbb{R})^N} \right)^{1/2}.
\end{aligned}$$

Remark 15.1. For our purposes it suffices to note that if $q : \mathbb{R}^N \to \mathbb{R}$ is Lipschitz continuous and bounded in $L^\infty(\mathbb{R})^N$, then it is in $L^2_{\rm ul}(\mathbb{R})^N$. Specifically, the Lipschitz continuity implies that $\lim_{y\to 0} ||\varphi^y q - q||_{L^2_{\rm ul}(\mathbb{R})^N} = 0$ holds, i.e. that $\varphi^y(q)$ is continuous in $y \in \mathbb{R}$.

Denote by $A : H^2_{\rm loc}(\mathbb{R})^N \to L^2_{\rm loc}(\mathbb{R})^N$ the linear operator $Aq = -q_{tt}$. The system (2.1) can then be written in a compact form

$$q_s = -Aq + F(q), \tag{15.1}$$

$F(q)(t) = \partial V(q(t), t)/\partial q$. Fix an initial condition $q^0 \in \mathcal{X}$. We substitute $\tilde{q} = q - q^0$, and consider

$$\begin{aligned}
\tilde{q}_s &= -\tilde{A}\tilde{q} + \tilde{F}(\tilde{q}), \\
\tilde{q}(0) &= 0,
\end{aligned} \tag{15.2}$$

where $\tilde{F}(\tilde{q}) = F(\tilde{q} + q^0) - Aq^0$, and $\tilde{A}$ is the restriction of A to $H^2_{\rm ul}(\mathbb{R})^N$, $\tilde{A} : H^2_{\rm ul}(\mathbb{R})^N \to L^2_{\rm ul}(\mathbb{R})^N$. It is straightforward to check that $\tilde{F}$ is well-defined as $\tilde{F} : H^1_{\rm ul}(\mathbb{R})^N \to L^2_{\rm ul}(\mathbb{R})^N$, and uniformly Lipschitz (for a fixed q^0) on the entire domain. Without loss of generality, we assume that the initial condition is given at $s_0 = 0$, and fix q^0 throughout the proofs.

Lemma 15.1. *For some $S > 0$ small enough, there exists an unique solution $\tilde{q}$ of (15.2) on $(0, S)$,*

$$\tilde{q} \in C^0([0,S)\,, H^1_{ul}(\mathbb{R})^N) \cap C^1((0,S)\,, H^1_{ul}(\mathbb{R})^N) \cap C^0((0,S)\,, H^2_{ul}(\mathbb{R})^N).$$

Proof. We follow [26, Chap. 3]. First note that $\tilde{A}$ generates an analytic semigroup $S(s) = \exp(-s\tilde{A})$ of bounded linear operators in $L^2_{\text{ul}}(\mathbb{R})^N$, which can for example be verified by using the explicit expression of the heat kernel (see e.g. [19] for details). As $\tilde{F}$ is uniformly, thus locally Lipschitz in q, and constant in s, the claim follows from [26, Theorem 3.3.3] with (using the notation from [26]) $X = L^2_{\text{ul}}(\mathbb{R})^N$, $D(\tilde{A}) = H^2_{\text{ul}}(\mathbb{R})^N$, $\alpha = 1/2$, $X^{1/2} = H^1_{\text{ul}}(\mathbb{R})^N$. □

We now require the well-known fact [26] that $\tilde{q}$ is a (”classical”) solution of (15.2) on $(0, S)$ if and only if it is a mild solution, i.e. if for any $0 < s_1 \leq S$, $\tilde{q}$ satisfies the integral equation

$$\tilde{q}(s_1) = \int_0^{s_1} e^{-\tilde{A}(s_1 - s)} \tilde{F}(\tilde{q}(s)) ds. \tag{15.3}$$

Lemma 15.2. *The solution $\tilde{q}$ of (15.2) exists on $(0, \infty)$.*

Proof. As $\tilde{F}(\tilde{q})$ is uniformly bounded in $L^2_{\text{ul}}(\mathbb{R})^N$ by some constant A, we have that if the solution of (15.2) exists on $(0, S)$, then for all $s \in (0, S)$, $||\tilde{q}(s)||_{H^2_{\text{ul}}(\mathbb{R})^N} \leq e^S \cdot A$. The claim now follows from [26, Corollary 3.3.5], as ”blow-up” is not possible. □

Lemma 15.3. *The solution of (15.1) is continuous with respect to initial conditions in both $\mathcal{X}_{ul}$ and $\mathcal{X}_{loc}$.*

Proof. We substitute back q instead of $\tilde{q}$ in (15.3), and obtain

$$q(s_1) = q^0 + \int_0^{s_1} e^{-\tilde{A}(s_1 - s)} \left(F(q(s)) - Aq^0 \right) ds. \tag{15.4}$$

Consider a sequence of initial conditions $q^{0,(n)}$ converging in either $\mathcal{X}_{\text{ul}}$ or $\mathcal{X}_{\text{loc}}$ norm to q^0, and consider associated solutions $q^{(n)}(s)$, $q(s)$. Continuity is then shown by bounding the difference of the right-hand sides of (15.4), for $q^{(n)}(s_1)$ and $q(s_1)$ for n large enough and $s_1 > 0$ small enough, in either $\mathcal{X}_{\text{ul}}$- or $\mathcal{X}_{\text{loc}}$-norm (see also [26, Corollary 3.4.1] for details). □

Lemma 15.4. *If $V \in H^k(\mathbb{T}^{N+1})$, $k \geq 2$, then for all $s > 0$, if q is the solution of (15.1), then $q(s) \in H^k_{ul}(\mathbb{R})^N$.*

Proof. We prove it inductively. In the proofs of Lemmas 15.1, 15.2, we already established the case $k = 2$. Consider the case $k \geq 3$, and assume the claim holds for a given $k-1$. Let $r = d^{k-2}q/dt^{k-2}$. By the inductive assumption, $r(s) \in H^1_{\text{ul}}(\mathbb{R})^N$ for all $s > 0$. For arbitrarily small $\delta > 0$, consider the system of equations

$$\begin{aligned} r_s &= -\tilde{A}r + F^{(k-2)}(s), \\ r(\delta) &= \frac{d^{k-2}}{dt^{k-2}} q(\delta) \end{aligned} \tag{15.5}$$

where $F^{(k-2)}(s) = \frac{d^{k-2}}{dt^{k-2}} \frac{\partial}{\partial q} V(q(s,.),.)$ is a fixed function. One can verify by using the inductive assumption $q \in H^{k-1}_{\text{ul}}(\mathbb{R})^N$, the assumed regularity of V and the embedding properties of the uniformly local spaces [19], that $F^{(k-2)}(s) \in L^2_{\text{ul}}(\mathbb{R})^N$ for all $s \in (0, \infty)$, and that it is uniformly bounded in $L^2_{\text{ul}}(\mathbb{R})^N$ on $(\delta, S]$ for any $S > \delta$. Now by repeating the argument of existence and uniqueness of solutions as in Lemma 15.1, we deduce that for any $s \in (\delta, \infty)$, $r(s)$ is a solution of (15.5), thus $r(s) \in H^2_{\text{ul}}(\mathbb{R})^N$ and $q(s) \in H^k_{\text{ul}}(\mathbb{R})^N$. As $\delta > 0$ is arbitrarily small, the claim is proved. □

Theorem 2.1 now follows from Lemmas 15.1, 15.2, 15.3 and 15.4.

We close the section with a frequently required result that the solutions of (1.3) we construct are indeed in $\mathcal{X}$.

Lemma 15.5. *If q is a solution of (3.1) such that either $q_t \in L^\infty(\mathbb{R})^N$ or $q_t \in L^2_{ul}(\mathbb{R})^N$, then $q \in \mathcal{X}$ (and by definition, $q \in \mathcal{E}$).*

Proof. First we note that as $\partial V(q,t)/\partial q$ is C^1 and periodic, q_{tt} is continuous and in $L^\infty(\mathbb{R})^N$. We easily deduce that also in the case $q_t \in L^2_{\text{ul}}(\mathbb{R})^N$, we have $q_t \in L^\infty(\mathbb{R})^N$. The Mean Value Theorem shows that the Lipschitz constant for q_{tt} is bounded with $||\partial^2 V/\partial q^2||_{L^\infty(\mathbb{R})^N}||q_t||_{L^\infty(\mathbb{R})^N} + ||\partial^2 V/\partial q\partial t||_{L^\infty(\mathbb{R})^N}$, which is finite as V is C^2 and periodic in all the variables. By Remark 15.1, we now have $q_t \in L^2_{\text{ul}}(\mathbb{R})^N$, $q_{tt} \in L^2_{\text{ul}}(\mathbb{R})^N$, thus $q \in \mathcal{X}$. □

16. Appendix B: A-priori bounds on one-sided minimizers

This Appendix is dedicated to the proofs of Lemma 5.3 and Proposition 5.4 in Section 5 and Lemma 14.1 in Section 14, i.e. the construction of one-sided minimizers in Section 5 and calculation of a-priori bounds.

16.1. **Proofs from Section 5.**

Lemma 16.1. *Define* $\tilde{w}(t) = 4\,\text{arctg}\exp(\sqrt{\varepsilon(1-2\mu^{1/2})}\,t)$*, and*

$$w(t) = \begin{cases} \tilde{w}(t) & t \geq t_1 \\ \tilde{w}(t) + \varepsilon^{3/2}\mu^{1/2}(t_1 - t)^3 & t \leq t_1, \end{cases}$$

where t_1 *is chosen so that* $\tilde{w}(t) = \pi + 2\mu^{1/2}$*. Then*

(i) $0 < t_1 \leq 2\sqrt{\mu/\varepsilon}$*,*

(ii) w *is is* C^2*, for all* $t \in [-1/(2\sqrt{\varepsilon}), \infty)$*,* $0 < w_t < 2\sqrt{\varepsilon}$*, and for all* $t \in [-1/(2\sqrt{\varepsilon}), 1/(2\sqrt{\varepsilon})]$*,*

$$\sqrt{\varepsilon}/2 \leq w_t, \tag{16.1}$$

(iii) There is a unique t_0 *satisfying* $w(t_0) = \pi$ *on* $[-1/\sqrt{\varepsilon}, \infty)$*, and it satisfies* $0 > t_0 \geq -\mu^{1/2}$*,*

(iv) w *is a strict stationary sub-solution of (1.2a) on* $[-1/\sqrt{\varepsilon}, \infty)$*. Furthermore, for any* $T \geq 0$*,* $w(t-T)$ *is a strict stationary sub-solution of (1.2a) on* $[-1/\sqrt{\varepsilon} + T, \infty)$*.*

Proof. (i) As $\tilde{w}$ is strictly increasing and $\tilde{w}(0) = \pi$, clearly t_1 is unique and $t_1 > 0$. To show $t_1 \leq 2\sqrt{\mu/\varepsilon}$, it suffices to show that $\tilde{w}(t^*) > \pi + 2\mu^{1/2}$ for $t^* = 2\sqrt{\mu/\varepsilon}$. This is straightforward by the mean-value theorem and by noting that the derivative of $4\,\text{arctg}\,e^t$ is $\geq 9/5$ on $[0, \sqrt{2}/4]$.

(ii) This follows by elementary calculation, applying (i) and the standing assumption $\sqrt{\mu} \leq 1/4$.

(iii) As $w(0) > \pi$, it suffices to show that $\tilde{w}(-\mu^{1/2}/2) > \pi$. For $t^* = -\mu^{1/2}/2$, $|t^* - t_1| \leq \sqrt{\mu}/\sqrt{\varepsilon}$, thus then value of the polynomial in the definition of w is $\leq 27\mu^2$. It is elementary to show that $\tilde{w}(-\mu^{1/2}/2) < \pi - 27\mu^2$, applying (A2).

(iv) It is easy to see that it suffices to show that

$$\tilde{\mathcal{F}}(w) := w_{tt} - \varepsilon \sin w(t) - \varepsilon\mu|\sin w(t)| - \varepsilon\mu(1 - \cos w(t)) > 0$$

for $t \in [-1/\sqrt{\varepsilon}, \infty)$. Note that $\tilde{w}_{tt} = (\varepsilon(1 - 2\mu^{1/2}))\sin\tilde{w}(t)$. For $t \geq t_1$, $\sin\tilde{w}(t) = \sin w(t)$ by definition. For $t \in [-1/\sqrt{\varepsilon}, t_1]$, by the strict monotonicity of both $w(t), \tilde{w}(t)$ we see that $\pi/2 \leq \tilde{w}(t) \leq w(t) \leq 3\pi/2$, thus $\sin\tilde{w}(t) \geq \sin w(t)$. In both cases we thus get

$$\tilde{\mathcal{F}}(w) = (\varepsilon(1 - 2\mu^{1/2}))\sin\tilde{w}(t) + 6\varepsilon^{3/2}\mu^{1/2}((t_1 - t) \vee 0) - \varepsilon\sin w(t) - \varepsilon\mu|\sin w(t)| - \varepsilon\mu(1 - \cos w(t))$$

$$\geq 6\varepsilon^{3/2}\mu^{1/2}((t_1 - t) \vee 0) - 2\varepsilon\mu^{1/2}\sin w(t) - \varepsilon\mu|\sin w(t)| - \varepsilon\mu(1 - \cos w(t)). \tag{16.2}$$

Denote the expression (16.2) by $\mathcal{F}(w(t))$. It suffices to show now that $\mathcal{F}(w(t)) > 0$.

Consider first the case $w(t) \geq \pi + 2\mu^{1/2}$, which is equivalent to $t \geq t_1$. As in this case, $\sin w < 0$, and always $\mu < \mu^{1/2}$, we have

$$\mathcal{F}(w) > \varepsilon\mu^{1/2}|\sin w| - \varepsilon\mu(1 - \cos w). \tag{16.3}$$

For any $w \in (0, \pi)$, the inequality

$$|\sin w| \geq \frac{1}{2}|w - \pi|(1 - \cos w). \tag{16.4}$$

holds. Inserting this and $|w - \pi| \geq 2\mu^{1/2}$ in (16.3) we get $\mathcal{F}(w) > 0$.

Let $w \in [\pi, \pi + 2\mu^{1/2}]$, or equivalently $t \in [t_0, t_1]$. As $\mathcal{F}(w(t_1)) > 0$, it suffices to show that $\mathcal{F}(w(t)) - \mathcal{F}(w(t_1)) \geq 0$. Calculating we get

$$\mathcal{F}(w(t)) - \mathcal{F}(w(t_1)) \geq 6\varepsilon^{3/2}\mu^{1/2}|t-t_1| - (2\varepsilon\mu^{1/2}+\varepsilon\mu)|\sin w(t) - \sin w(t_1)| - \varepsilon\mu|\cos w(t) - \cos w(t_1)|. \tag{16.5}$$

Using the mean-value theorem, (16.1) and $\mu^{1/2} \leq 1/2$, we get

$$(2\varepsilon\mu^{1/2} + \varepsilon\mu)|\sin w(t) - \sin w(t_1)| \leq \frac{5}{2}\varepsilon\mu^{1/2} \cdot 2\varepsilon^{1/2}|t - t_1| = 5\varepsilon^{3/2}\mu^{1/2}|t-t_1|,$$
$$\varepsilon\mu|\cos w(t) - \cos w(t_1)| \leq \frac{1}{2}\varepsilon\mu^{1/2} \cdot 2\varepsilon^{1/2}|t - t_1| \leq \varepsilon^{3/2}\mu^{1/2}|t-t_1|.$$

We combine it with (16.5) to get $\mathcal{F}(w(t)) - \mathcal{F}(w(t_1)) \geq 0$.

We now also know that $\mathcal{F}(w(t_0)) > 0$. In the case $t \in [-1/\sqrt{\varepsilon}, t_0]$ which is equivalent to $\sin w(t) \in [\pi/2, \pi]$, it suffices to show that $\mathcal{F}(w(t)) - \mathcal{F}(w(t_0))$. We again obtain that $\mathcal{F}(w(t)) - \mathcal{F}(w(t_0))$ is equal to the right-hand side of (16.5) with t_0 instead of t_1. The rest of the proof is analogous to the previous case. □

Proof of Lemma 5.3. We take w constructed in Lemma 16.1 and set $z^+ = w(t - t_0)$, $z^- = 2\pi - w(-t + t_0)$. The claims (i)-(iv) are now straightforward, (v) can be easily checked by direct calculation, and (vi) follows from the definition and the relations $2\pi - 4\operatorname{arctg} x < 4/x$, $4\operatorname{arctg}(1/x) = 2\pi - 4\operatorname{arctg} x$, holding for all $x > 0$. By the Mean Value Theorem, by noting that $|d(\operatorname{arctg} e^x)/dt| \leq 4e^{-|x|}$, the construction and bounds on t_0, t_1, we easily get that for $t \geq 0$,

$$|w(t-t_0) - u^{(\varepsilon)}(t)| \ll e^{-\frac{1}{2}\sqrt{\varepsilon}t}(\varepsilon\sqrt{\mu}\, t + \sqrt{\varepsilon\mu}) \ll \sqrt{\varepsilon\mu},$$

which completes the proof. □

We now construct one-sided minimizers and prove Proposition 5.4. We fix c, t_0, v_0 until the end of the section, and construct q^+, q^- is analogous. For a positive integer k, we consider the functional

$$\mathcal{L}_{\omega,k}(q) = \int_{t_0}^{t_0+k} L_\omega(u(t), v(t), t)dt. \tag{16.6}$$

We construct minimizers q_k of (16.6), and then obtain q^+ as their limit.

Lemma 16.2. *The functional $\mathcal{L}_{c,k}$ attains its minimum q_k over all $q = (u, v) \in H^1([t_0, t_0 + k])^2$ such that*

$$q(t_0) = (\pi, v_0), u(t_1) = 2\pi. \tag{16.7}$$

Then $q_k \in H^2([t_0, t_0 + k])^2$, and is a solution of Euler-Lagrange equations on $(t_0, t_0 + k)$.

Proof. The Tonelli theorem [33, Appendix 1] implies that for a fixed v_1, such a minimum is attained over all $q = (u, v)$ such that $v(t_1) = v_1$, as the conditions for the Tonelli theorem to hold in the non-autonomous case are satisfied. It is easy to show that it suffices to consider v_1 from a closed interval, which by compactness and continuity of the minimum of (16.6) in v_1 implies existence of such a minimizing q_k. Furthermore, by the Tonelli theorem, $q_k \in H^2([t_0, t_0+k])^2$, and q_k is a solution of the Euler-Lagrange equations on (t_0, t_0+k). □

We denote by q_k the (not necessarily unique) minimizer of (16.6) satisfying (16.7). We always set $u_k(t) = 2\pi$ for $t \geq t_0 + k$.

Lemma 16.3. *The minimizer $q_k = (u_k, v_k)$ satisfies for all $k \geq k_0$, k_0 sufficiently large:*
(i) For all $t \in [t_0, t_0 + k)$, $0 < u_k(t) < 2\pi$,
(ii) For all $t > t_0 + 4\mu/\sqrt{\varepsilon}$, $u_k(t) > \pi$,
(iii) For all $t \geq t_0$, $u_k(t) > \pi - 1/4$,
(iv) For all $t \geq t_0$, $u_k(t) \geq z^+(t - t_0)$. Furthermore, for all $0 \leq T \leq 3/(4\sqrt{\varepsilon})$,

$$u_k \geq z^+(t - t_0 - T).$$

Proof. (i) Assume that for some $t_0 < t^* < t_0 + k$, $u_k(t) = 2\pi$. We define

$$\tilde{q} = \begin{cases} \tilde{q}(t) = q_k(t) & \text{for } t \in [t_0, t^*], \\ \tilde{q}(t) = (2\pi, v^k(t^*)) & \text{for } t \in [t^*, t_0 + k]. \end{cases}$$

Now it is easy to check by direct calculation that $\mathcal{L}_{c,k}(\tilde{q}) \le \mathcal{L}_{c,k}(q_k)$, thus $\tilde{q}$ minimizes $\mathcal{L}_{c,k}$ and must be a solution of the Euler-Lagrange equations on $(t_0, t_0 + k)$. We deduce that $\tilde{u} \equiv 2\pi$ which is in contradiction to $\tilde{u}(t_0) = u_k(t_0) = \pi$. By continuity we get the right-hand side of (i). Similarly we show $u_k(t) > 0$, otherwise we replace the segment between two intersections of 0 with $u^k(t) = 0$ and get a contradiction.

(ii) By (5.2), we can find k_0 large enough so that for any $k \ge k_0$,

$$\mathcal{L}_{c,k}(q_k) \le 4\sqrt{\varepsilon(1 + 3\mu/2)}. \tag{16.8}$$

Assume $u_k(t_0 + d) = \pi$ for some $t_0 + d \in (t_0, t_0 + k)$. Again by (5.2), as q^k is a minimizer, we see that

$$\int_{t_0+d}^{t_0+k} L_\omega(q_k, (q_k)_t, t)dt \ge 4\sqrt{\varepsilon(1 - \mu)}. \tag{16.9}$$

From (16.8) and (16.9) and the standing assumption $\mu \le 1/16$ we obtain the upper bound

$$\int_{t_0}^{t_0+d} L_\omega(q_k, (q_k)_t, t)dt \le 4\sqrt{\varepsilon(1 + 3\mu/2)} - 4\sqrt{\varepsilon(1 - \mu)} \le 6\sqrt{\varepsilon}\mu. \tag{16.10}$$

Let $t^* = (d/2) \wedge 1$, and let $\pi - a$ be the minimal value of u_k on $[t_0, t_0 + t^*]$, $0 \le a < \pi$. It is easy to see that

$$\begin{aligned} \int_{t_0}^{t_0+t^*} L_\omega(q_k, (q_k)_t, t)dt &\ge \int_{t_0}^{t_0+t^*} \left(\frac{1}{2}((u_k)_t)^2 + \varepsilon(1 - \mu)(1 - \cos u_k) \right) dt \\ &\ge \frac{1}{2t^*}a^2 + \varepsilon(1 - \mu)\left(2 - \frac{a^2}{2}\right)t^* \\ &\ge 2\varepsilon(1 - \mu)t^* \ge \frac{3}{2}\varepsilon t^*. \end{aligned}$$

Repeating that over $[t_0 + d - t^*, t_0 + d]$ we get $\int_{t_0}^{t_0+d} L_\omega(q_k, (q_k)_t, t) \ge 3\varepsilon t^*$, thus by (16.10), $t^* \le 2\mu/\sqrt{\varepsilon} \le 1$. By definition, $d \le 4\mu/\sqrt{\varepsilon}$. The claim follows by continuity of u_k.

(iii) Using the same notation as in (ii) and by (16.10), we easily see that

$$6\sqrt{\varepsilon}\mu \ge \int_{t_0}^{t_0+d} L_\omega(q^k, q^k_t, t)dt \ge \frac{2}{d}a^2 \ge \frac{\sqrt{\varepsilon}}{2\mu}a^2,$$

thus $a^2 \le 12\mu^2$. By (A2), $a \le 1/4$.

(iv) We prove it by using Lemma 5.2 in two steps. First we show that

$$u_k(t) > z^+(t - t_0 - T) \text{ for all } t \ge t_0 + T - 3/(4\sqrt{\varepsilon}) \tag{16.11}$$

and all $T \ge 3/(4\sqrt{\varepsilon})$. By definition and Lemma 5.3, (i), (16.11) holds for $T = k + 3/(4\sqrt{\varepsilon})$. Assume the contrary and find the infimum T^* of $T \ge 3/(4\sqrt{\varepsilon})$ for which (16.11) holds. By compactness and continuity, we have that (16.11) holds for $T = T^*$ with $\ge$ instead of $>$. However, by construction, (ii),(iii) and Lemma 5.3, (v), the strict inequality in (16.11) holds for $t \in \{t_0 + T - 3/(4\sqrt{\varepsilon}), t_0 + k\}$, thus by Lemma 5.2 we must have strict inequality in (16.11) and $T^* = 3/(4\sqrt{\varepsilon})$.

Now we show that

$$u_k(t) > z^+(t - t_0 - T) \text{ for all } t \ge t_0 \tag{16.12}$$

for all $T \in (0, 3/(4\sqrt{\varepsilon})]$. Again we find the infimum T^* for which (16.12) holds. We obtain contradiction analogously to the previous step, using $u_k(t) > z^+(t - t_0 - T)$ for $t \in \{t_0, t_0 + k\}$; unless $T^* = 0$ and (16.12) holds with $\ge$ instead of $>$, which we need to show. □

Proof of Proposition 5.4. It is easy to check that for $k \geq k_0$, k_0 as in Lemma 16.3, q_k is uniformly bounded in $H^2_{\mathrm{loc}}([t_0,\infty))^2$, i.e. that for any $T > k_0$, we have that q_k, $k \geq T$, is uniformly bounded in $H^2([t_0, t_0+T])^2$. Indeed, uniform bounds in k on $||(q_k)_t||_{L^2([t_0,t_0+T])^2}$ follow from the fact that $\mathcal{L}_{\omega,k}(q_k)$ is by definition decreasing in k, and the uniform bound $|V(q,t)| \leq \varepsilon(1+\mu)$. The uniform bound in k on $||q_k||_{L^2([t_0,t_0+T])^2}$ follows from that and $q(t_0) = (\pi, v_0)$. The uniform bound in k on $||(q_k)_{tt}||_{L^2([t_0,t_0+T])^2}$ is deduced from the fact that q_k is a solution of Euler-Lagrange equations, and the uniform bounds $|V_u| \leq \varepsilon(1+\mu)$, $|V_v| \leq \varepsilon\mu$.

Now by diagonalization we find a convergent subsequence of q_k in $H^1_{\mathrm{loc}}([t_0,\infty))^2$ converging to some $q^0 \in H^1_{\mathrm{loc}}([t_0,\infty))^2$. It is straightforward to show that $\lim_{k\to\infty} \mathcal{L}_{\omega,k}(q_k) = S^+_\omega(t_0, v_0)$, as $\mathcal{L}_{\omega,k}(q_k)$ is decreasing, bounded by $S^+_\omega(t_0, v_0)$ from below, and we can arbitrarily well approximate $S^+_\omega(t_0, v_0)$ with $\mathcal{L}_{\omega,k}(q_k)$ for k large enough.

By construction and Lemma 16.3, (iv), $\lim_{t\to\infty} u^0(t) = 2\pi$. By the Fatou Lemma, $\int_{t_0}^{\infty} L_\omega(q^0, q^0_t, t) \leq S^+_\omega(t_0, v_0)$, thus by the definition of $S^+_\omega(t_0, v_0)$, the equality must hold, and the existence is proved.

Now, if q^+ is any one-sided minimizer at (c, t_0, v_0), it must be by the Tonelli theorem a solution of Euler-Lagrange equations on (t_0, ∞), and C^4 because of the regularity of solutions of ordinary differential equations.

The proof of (ii) is analogous to the proof of Lemma 16.3, (iv). □

16.2. **Proof from Section 14.** Let $\tilde{\varepsilon} = \sqrt{\varepsilon(1+2\mu^{1/2})}$, and let $\delta = 2\sqrt{\tilde{\varepsilon}\mu}$. Denote by $u^{(\tilde{\varepsilon})} = 4\,\mathrm{arctg}(e^{\sqrt{\tilde{\varepsilon}}\,t})$ the separatrix solution of the unperturbed pendulum $u_{tt} = \tilde{\varepsilon}\sin u$.

Lemma 16.4. *If $\tilde{u}$ is a solution of $u_{tt} = \tilde{\varepsilon}\sin u$, $u(0) = \pi$, $u_t(0) = 2\sqrt{\tilde{\varepsilon}}(1+\delta^2)$, then for some absolute $c_{30} > 0$, we can find $t_1 > 0$ such that*

$$t_1 \leq \frac{c_{30}}{\sqrt{\tilde{\varepsilon}}}|\log\delta|, \tag{16.13}$$

we have $2\pi + \delta/2 \leq \tilde{u}(t_1) \leq 2\pi + 32\delta$, and for all $t \in [0, t_1]$, $|\tilde{u}(t) - u^{(\tilde{\varepsilon})}(t)| \leq 32\delta$.

Proof. Let $w^{(1)} = \tilde{u} - u^{(\tilde{\varepsilon})}$, let $t \geq 0$, and let $t_1 > 0$ be the unique (by monotocity) t such that $\tilde{u}(t_1) = 2\pi + \delta$. Clearly $\tilde{u}(t) \geq u^{(\tilde{\varepsilon})}(t)$, thus $w^{(1)}(t) \geq 0$. By definition, $w^{(1)}$ is solves the differential inequality $w^{(1)}_{tt}(t) \leq \varepsilon|\tilde{u}(t) - u^{(\tilde{\varepsilon})}(t)| = \varepsilon w^{(1)}(t)$. Now if $w^{(2)} = w^{(1)}_t - \sqrt{\tilde{\varepsilon}}w^{(1)}$, $w^{(3)} = w^{(1)}_t + \sqrt{\tilde{\varepsilon}}w^{(1)}$, we have $w^{(2)}(0) = w^{(3)}(0) = 2\sqrt{\tilde{\varepsilon}}\cdot\delta^2$, and they are solutions of $w^{(2)}_t \leq -\sqrt{\tilde{\varepsilon}}w^{(2)}$, $w^{(3)}_t \leq \sqrt{\tilde{\varepsilon}}w^{(3)}$. Thus by the Gronwall Lemma,

$$w^{(2)}(t) \leq 2\sqrt{\tilde{\varepsilon}}\,\delta^2\; e^{-\sqrt{\tilde{\varepsilon}}\,t}, \qquad w^{(3)}(t) \leq 2\sqrt{\tilde{\varepsilon}}\,\delta^2\, e^{\sqrt{\tilde{\varepsilon}}\,t}. \tag{16.14}$$

By the conservation of energy $H(u) = u_t^2/2 - (1-\cos u)$, and by $\tilde{u}(t) \geq u^{(\tilde{\varepsilon})}(t)$ and $\tilde{u}_t, u^{(\tilde{\varepsilon})}_t \geq 0$ we easily obtain that $w^{(1)}_t \geq 0$, thus the right-hand side of (16.14) implies for all $t \in [0, t_1]$,

$$w^{(1)}(t) \leq 2\delta^2\, e^{\sqrt{\tilde{\varepsilon}}\,t_1}. \tag{16.15}$$

By the conservation of energy, it is easy to show that for all $t \geq 0$, $\tilde{u}_t(t) \geq 2\sqrt{\tilde{\varepsilon}}\,\delta$. Inserting that in the left-and side of (16.14) we obtain

$$w^{(1)}(t_1) \geq 2\delta - \frac{1}{\sqrt{\tilde{\varepsilon}}}u^{(\tilde{\varepsilon})}_t(t_1) - 2\delta^2\, e^{-\sqrt{\tilde{\varepsilon}}\,t_1}. \tag{16.16}$$

Now for all $t \geq 0$, as $2\pi - 4\,\mathrm{arctg}\,x < 4/x$, we have $2\pi - u^{(\tilde{\varepsilon})}(t) \leq 4e^{-\sqrt{\tilde{\varepsilon}}t}$. By the conservation of energy and $(1-\cos(2\pi - x)) \leq x^2/2$, we now get that $u^{(\tilde{\varepsilon})}_t(t) < 4e^{-\sqrt{\tilde{\varepsilon}}t}$. Thus choosing t_1 so that $4e^{-\sqrt{\tilde{\varepsilon}}t} = \delta/4$, we get $2\pi - \delta/4 < u^{(\tilde{\varepsilon})}(t_1) < 2\pi$, (16.15) becomes $w^{(1)}(t) \leq 32\delta$, and as $\delta^2 \leq 1$, (16.16) becomes $w^{(1)}(t) \geq \delta$, which implies the claim. □

Proof of Lemma 14.1. Take $\tilde{u}$ from Lemma 16.4, and set

$$w(t) = \begin{cases} \tilde{u}(t) & t \geq t_1 \\ \tilde{u}(t) - \varepsilon^{3/2}\mu^{1/2}(t_1 - t)^3 & t \leq t_1, \end{cases}$$

where t_1 is chosen so that $\tilde{w}(t) = \pi + 2\mu^{1/2}$. Then as in Lemma 16.1, we can find $0 < t_0 < \mu^{1/2}$ such that $w(t_0) = \pi$. We set $\tilde{z}^+(t) = w(t - t_0)$, $\tilde{T} = t_1 - t_0$, and $\tilde{z}^- = 2\pi - w(-t + t_0)$. The rest of the proof is analogous to the proof of Lemma 5.3, using also the relation $|u^{(\varepsilon)} - u^{(\tilde{\varepsilon})}| \ll \sqrt{\varepsilon\mu}$ to obtain (v). □

17. Appendix C: Proofs of bounds on derivatives

This Appendix is dedicated to the proof of Proposition 9.1. In the first subsection we obtain weighted a-priori bounds on q^0 and the potential V and its derivatives. In the second subsection we by careful differentiation we obtain a differential inequality which proves the invariance of sets.

17.1. **A-priori bounds on weighted integrals.** Throughout the subsection we assume $q \in \mathcal{B}_1$.

Lemma 17.1. *Let $j, m \in \{1, 2, 3, 4, 5, 6\}$. Then for some absolute constant $c_{31} > 0$,*

$$\varepsilon^{(j+1)/2} \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|} e^{-\frac{m}{2}\sqrt{\varepsilon}\|t\|} dt \leq c_{31}\lambda(\tau)^j, \tag{17.1}$$

$$\int_{\mathbb{R}} \frac{1}{L_{k(t)}^{j+1}} e^{-\lambda(\tau)|t-\tau|} dt \leq c_{31}\lambda(\tau)^j. \tag{17.2}$$

Proof. To show (17.1), without loss of generality let

$$\tau \in [\tilde{T}_k - 2L_{k-1}, \tilde{T}_{k+1} + 2L_k] \tag{17.3}$$

for a fixed $k \in \mathbb{Z}$, i.e. $||\tau|| = |\tau - \tilde{T}_k|$. By applying the definition of $\|t\|$ and substituting $t - T_j + \tilde{T}_k$ instead of t in the second line below, we get

$$\begin{aligned}\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|} e^{-\frac{m}{2}\sqrt{\varepsilon}\|t\|} dt &= \sum_{j=-\infty}^{\infty} \int_{T_j - 2L_{j-1}}^{T_j + 2L_j} e^{-\lambda(\tau)|t-\tau|} e^{-\frac{m}{2}\sqrt{\varepsilon}|t - T_j|} dt \\ &= \sum_{j=-\infty}^{\infty} \int_{\tilde{T}_k - 2L_{j-1}}^{\tilde{T}_k + 2L_j} e^{-\lambda(\tau)|t - \tilde{T}_k + T_j - \tau|} e^{-\frac{m}{2}\sqrt{\varepsilon}|t - \tilde{T}_k|} dt.\end{aligned} \tag{17.4}$$

Now by the assumption (17.3), for $j = k-1, k, k+1$ we have $|T_j - \tau| \geq |\tilde{T}_k - \tau|$. For $j \geq k+1$, $|T_j - \tau| \geq |T_j - \tilde{T}_{k+1}| + |\tilde{T}_k - \tau| \geq 4L(|j-k| - 1) + |\tilde{T}_k - \tau|$. By analogous consideration for $j \leq k-1$, we conclude that for all $j \in \mathbb{Z}$,

$$|T_j - \tau| \geq |\tilde{T}_k - \tau| + 4L(|j-k| - 1) = \|\tau\| + 4L((|j-k| - 1) \vee 0). \tag{17.5}$$

Now by definition of $\lambda(\tau)$, we have $\sqrt{\varepsilon}\, m/2 \geq \sqrt{\varepsilon}/2 \geq \lambda(\tau) + \sqrt{\varepsilon}/4$. Combining it with (17.5), we can estimate the exponent in (17.4):

$$\begin{aligned}\lambda(\tau)|t - \tilde{T}_k + T_j - \tau| + \frac{m}{2}\sqrt{\varepsilon}|t - \tilde{T}_k| &\geq \lambda(\tau)|T_j - \tau| - \lambda(\tau)|t - \tilde{T}_k| + \frac{m}{2}\sqrt{\varepsilon}|t - \tilde{T}_k| \\ &\geq \lambda(\tau)\|\tau\| + \lambda(\tau)4L((|j-k| - 1) \vee 0) + \frac{1}{4}\sqrt{\varepsilon}|t - \tilde{T}_k|.\end{aligned}$$

Thus (17.4) is less or equal than

$$e^{-\lambda(\tau)\|\tau\|} \left(2 + \sum_{j=-\infty}^{\infty} e^{-\lambda(\tau)4L\cdot j}\right) \int_{-\infty}^{\infty} e^{-\frac{1}{4}\sqrt{\varepsilon}|t - \tilde{T}_k|} = \frac{4}{\sqrt{\varepsilon}} e^{-\lambda(\tau)\|\tau\|} \frac{1}{1 - e^{-4\lambda(\tau)L}}. \tag{17.6}$$

If $\lambda(\tau) = \sqrt{\varepsilon}/4$, then $4L\lambda(\tau) \geq 1$, thus right-hand side of (17.6) is $\ll \varepsilon^{-1/2}$. This implies (17.1). Otherwise $\lambda(\tau) = 8\log\|\tau\|/\|\tau\|$, so the right-hand side of (17.6) becomes

$$\frac{4}{\sqrt{\varepsilon}} \frac{||\tau||^{-8}}{1 - \|\tau\|^{-4L/\|\tau\|}}. \tag{17.7}$$

Now by substitution $z = (1/||\tau||)\log(1/||\tau||)$, noting that by all the assumptions, $z \in (-1/2, 0)$, we easily show by Taylor expansion that the denominator in (17.7) is $\ll ||\tau|| \log(1/||\tau||)| \leq ||\tau||^2$. Thus (17.7) is $\ll \varepsilon^{-1/2}||\tau||^{-6} \leq \varepsilon^{-1/2}\lambda(\tau)^6$, which implies (17.1).

To show (17.2), let $\tau \in [\tilde{T}_k, \tilde{T}_{k+1}]$ for fixed $k \in \mathbb{Z}$. As $L_k \geq 4L \geq 1/\sqrt{\varepsilon}$,

$$\int_{\mathbb{R}} \frac{1}{L_{k(t)}^{j+1}} e^{-\lambda(\tau)|t-\tau|} dt \leq \frac{1}{L_k^{j+1}} \int_{\tilde{T}_k}^{\tilde{T}_{k+1}} e^{-\lambda(\tau)|t-\tau|} dt + \varepsilon^{\frac{j+1}{2}} e^{-\lambda(\tau)|\tau-\tilde{T}_k|} \int_{-\infty}^{\tilde{T}_k} e^{-\lambda\tau|t-\tilde{T}_k|} dt$$

$$+ \varepsilon^{\frac{j+1}{2}} e^{-\lambda(\tau)|\tau-\tilde{T}_{k+1}|} \int_{\tilde{T}_{k+1}}^{\infty} e^{-\lambda(\tau)|t-\tilde{T}_{k+1}|} dt$$

$$\leq \frac{2}{\lambda(\tau)\, L_k^{j+1}} + \frac{2\varepsilon^{\frac{j+1}{2}}}{\lambda(\tau)} e^{-\lambda(\tau)||\tau||}. \tag{17.8}$$

As $L_k \geq 1/\lambda(\tau)$, the first summand in (17.8) is always $\ll \lambda(\tau)^j$. If $\lambda(\tau) = \sqrt{\varepsilon}/8$, the second summand is $\ll \varepsilon^{j/2} \ll \lambda(\tau)^j$. Otherwise, as $\varepsilon \leq 1$, and as $||\tau|| \geq 1/(8\lambda(\tau))$,

$$\frac{\varepsilon^{2\frac{j+1}{2}}}{\lambda(\tau)} e^{-\lambda(\tau)||\tau||} \leq \frac{1}{\lambda(\tau)} e^{-8\log||\tau||} \leq \frac{1}{||\tau||^8 \lambda(\tau)} \ll \lambda(\tau)^6 \leq \lambda(\tau)^j,$$

which completes (17.2). □

Lemma 17.2. *There exists an absolute constant $c_{32} > 0$ so that*

$$||V_u||^2_{L^2_\tau(\mathbb{R})} \leq c_{32}\lambda(\tau)^3, \qquad ||V_v||^2_{L^2_\tau(\mathbb{R})} \leq c_{32}\lambda(\tau)^3. \tag{17.9}$$

Proof. As $q \in \mathcal{B}_1$, by (8.1), we get

$$|\sin u(t)| \leq |u - 2k(t)\pi| \ll e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|}, \tag{17.10}$$

$$|1 - \cos u(t)| \leq \frac{1}{2}|u - 2k(t)\pi|^2 \ll e^{-\sqrt{\varepsilon}\|t\|}. \tag{17.11}$$

By definition and uniform bounds on f and its derivatives, we have

$$|V(u, v, t) \leq 2\varepsilon(1 - \cos u(t)) \ll \varepsilon\, e^{-\sqrt{\varepsilon}\|t\|},$$

$$|V_u(u, v, t)| \leq 2\varepsilon(1 - \cos u(t) + |\sin u(t)|) \ll \varepsilon\, e^{-\frac{1}{2}\sqrt{\varepsilon}\|t\|}, \tag{17.12}$$

$$|V_v(u, v, t)| \leq \varepsilon\mu(1 - \cos u(t)) \ll \varepsilon\mu\, e^{-\sqrt{\varepsilon}\|t\|}. \tag{17.13}$$

It suffices now to apply (17.1). □

Lemma 17.3. *There exists an absolute constant $c_{33} > 0$ so that*

$$||u_t^0||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}\lambda(\tau), \qquad ||v_t^0 - \omega_{k(.)}||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}\lambda(\tau), \tag{17.14}$$

$$||u_{tt}^0||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}\lambda(\tau)^3, \qquad ||v_{tt}^0||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}\lambda(\tau)^3, \tag{17.15}$$

$$||u_{ttt}^0||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}(\varpi^2 + 1)\lambda(\tau)^3, \qquad ||v_{ttt}^0||^2_{L^2_\tau(\mathbb{R})} \leq c_{33}(\varpi^2 + 1)\lambda(\tau)^3. \tag{17.16}$$

Proof. This follows by a straightforward calculation from Lemmas 7.2 and 17.1. □

17.2. **Proof of invariance of $\mathcal{B}_2$, $\mathcal{B}_3$, $\mathcal{B}_4$.** By assumption that $q \in \mathcal{A}$, we have that $q \in H^3_{\text{loc}}(\mathbb{R})^2$. Whenever we require higher derivatives in the proofs to follow, we assume that we evaluate all on a dense, sufficiently smooth subset and then extend the claims of the Lemmas by continuity to the entire set as required.

Lemma 17.4. *If $q = (u, v) \in \mathcal{B}_2$, then*

$$||u_t - u_t^0||^4_{L^2_\tau(\mathbb{R})} \leq c_{34}\left(\lambda(\tau)^2 + \lambda(\tau)^{-1}||u_{tt} - u_{tt}^0||^2_{L^2_\tau(\mathbb{R})}\right), \tag{17.17}$$

$$||v_t - v_t^0||^4_{L^2_\tau(\mathbb{R})} \leq c_{35}\left(\lambda(\tau)^2 M^4 + \lambda(\tau)^{-1} M^2 ||v_{tt} - v_{tt}^0||^2_{L^2_\tau(\mathbb{R})}\right), \tag{17.18}$$

for some absolute constants $c_{34}, c_{35} \geq 1$.

Proof. Denote by $X = ||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})}$ and by $Y = ||u_{tt} - u_{tt}^0||^2_{L^2_\tau(\mathbb{R})}$. By partial integration and the Cauchy-Schwartz inequality, we get

$$X = \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_t(t) - u_t^0(t))^2 dt$$
$$\leq \lambda(t) \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}|u(t) - u^0(t)||u_t(t) - u^0(t)|dt + \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}|u(t) - u^0(t)||u_{tt}(t) - u^0_{tt}(t)|dt$$
$$\ll \lambda(\tau)^{1/2}||u - u^0||_{L^\infty(\mathbb{R})}\, X^{1/2} + \lambda(\tau)^{-1/2}||u - u^0||_{L^\infty(\mathbb{R})}\, Y^{1/2}. \tag{17.19}$$

Now either $X \ll \lambda(\tau)||u - u^0||^2_{L^\infty(\mathbb{R})}$ or $X \ll \lambda(\tau)^{-1/2}||u - u^0||_{L^\infty(\mathbb{R})}$, thus

$$X^2 \ll \lambda(\tau)^2||u - u^0||^4_{L^\infty(\mathbb{R})} + \lambda(\tau)^{-1}||u - u^0||^2_{L^\infty(\mathbb{R})}. \tag{17.20}$$

From (7.9) and (8.1) we have $||u - u^0||_{L^\infty(\mathbb{R})} \leq c_6 + c_7$, which gives (17.17).

To show (17.18), it suffices to note that by (8.2), $||v - v^0||_{L^\infty(\mathbb{R})} \leq c_8 M$. The rest of the proof is analogous. □

Lemma 17.5. *The set of all $q \in \mathcal{B}_1$ such that $||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} \leq c_9\lambda(\tau)$ is $\mathcal{A}$-relatively ξ-invariant for some absolute constant $c_9 > 0$.*

Proof. First note that

$$u_s - (u_{tt} - u^0_{tt}) = -V_u + u^0_{tt},$$

thus by squaring it we get

$$-2u_s(u_{tt} - u^0_{tt}) \leq -u_s^2 - (u_{tt} - u^0_{tt})^2 + 2V_u^2 + 2(u^0_{tt})^2, \tag{17.21}$$

where we write $V_u = \partial_u V(u, v, t)$. Differentiating with respect to s, by partial integration and taking into account that u^0 is constant , we obtain

$$\frac{d}{ds}||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} = 2\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_t - u_t^0)u_{ts}dt$$
$$\leq -2\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_{tt} - u^0_{tt})u_{ts}dt + 2\lambda(\tau)\left|\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_t - u_t^0)u_s dt\right| \tag{17.22}$$

The first summand is by inserting (17.21),

$$-2\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_{tt} - u^0_{tt})u_{ts}dt \leq -||u_s||^2_{L^2_\tau(\mathbb{R})} - ||u_{tt} - u^*_{tt}||^2_{L^2_\tau(\mathbb{R})} + 2||V_u||^2_{L^2_\tau(\mathbb{R})} + 2||u^0_{tt}||^2_{L^2_\tau(\mathbb{R})}. \tag{17.23}$$

By Young's inequality, we have that

$$2\lambda(\tau)\left|\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_t - u_t^0)u_s dt\right| \leq \lambda(\tau)^2||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} + ||u_s||^2_{L^2_\tau(\mathbb{R})}. \tag{17.24}$$

Inserting (17.23) and (17.24) into (17.22), we see that

$$\frac{d}{ds}||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} \leq -||u_{tt} - u^0_{tt}||^2_{L^2_\tau(\mathbb{R})} + \lambda^2(\tau)||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} + 2||V_u||^2_{L^2_\tau(\mathbb{R})} + 2||u^0_{tt}||^2_{L^2_\tau(\mathbb{R})}. \tag{17.25}$$

By substituting $X = ||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})}$, $Y = ||u_{tt} - u^0_{tt}||^2_{L^2_\tau(\mathbb{R})}$, and using (17.9) and (17.15), we obtain from (17.25)

$$\frac{d}{ds}X \leq -Y + \lambda^2(\tau)X + 2(c_{32} + c_{33})\lambda(\tau)^3. \tag{17.26}$$

(17.17) can be written as

$$-\frac{1}{2}Y \leq -\frac{\lambda(\tau)}{2c_{34}}X^2 + \frac{1}{2}\lambda(\tau)^3. \tag{17.27}$$

Clearly

$$0 \leq \frac{\lambda(\tau)}{2c_{34}}X^2 - 2\lambda(\tau)^2 X + 2c_{34}\lambda(\tau)^3, \tag{17.28}$$

thus by summing (17.27) and (17.28) we see that

$$-\frac{1}{2}Y + \lambda(\tau)^2 X \le -\lambda(\tau)^2 X + (2c_{34} + 1/2)\lambda(\tau)^3.$$

Summing it with (17.26) and substituting back X and Y, we obtain a differential inequality

$$\frac{d}{ds}||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} \le -\frac{1}{2}||u_{tt} - u^*_{tt}||^2_{L^2_\tau(\mathbb{R})} - \lambda(\tau)^2||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} + c_9\lambda(\tau)^3 \tag{17.29}$$

$$\le -\lambda(\tau)^2||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} + c_9\lambda(\tau)^3, \tag{17.30}$$

where $c_9 = 2(c_{32} + c_{33} + c_{34}) + 1/2$. The claim follows from the Gronwall's lemma and Lemma 8.1, i.e. $\mathcal{A}$-relative invariance of $\mathcal{B}_1$. □

Lemma 17.6. *The set of all $q \in \mathcal{B}_1$ such that $||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})} \le c_9(M^2+1)\,\lambda(\tau)$ is $\mathcal{A}$-relatively ξ-invariant for some absolute constant $c_9 > 0$.*

Proof. Denote by $\tilde{X} = ||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})}$, $\tilde{Y} = ||v_{tt} - v_{tt}^0||^2_{L^2_\tau(\mathbb{R})}$. All the steps in the proof are analogous to the proof of Lemma 17.5 up to the equations (17.27), (17.28), instead of which we have

$$-\frac{1}{2}\tilde{Y} \le -\frac{\lambda(\tau)}{2c_{35}M^2}\tilde{X}^2 + \frac{1}{2}\lambda(\tau)^3 M^2,$$

$$0 \le \frac{\lambda(\tau)}{2c_{35}M^2}\tilde{X}^2 - 2\lambda(\tau)^2\tilde{X} + 2c_{35}\lambda(\tau)^3 M^2.$$

We thus obtain the differential inequality

$$\frac{d}{ds}||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})} \le -\frac{1}{2}||v_{tt} - v_{tt}^0||^2_{L^2_\tau(\mathbb{R})} - \lambda(\tau)^2||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})} + c_9(M^2+1)\lambda(\tau)^3 \tag{17.31}$$

$$\le -\lambda(\tau)^2||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})} + c_9(M^2+1)\lambda(\tau)^3. \tag{17.32}$$

which completes the proof analogously as in Lemma 17.5. □

Lemmas 17.5 and 17.6 compete the proof of $\mathcal{A}$-relative ξ-invariance of $\mathcal{B}_2$.

Lemma 17.7. *There exists a constant $c_{10} > 0$ such that the set of all $q \in \mathcal{B}_2$ satisfying (9.3a), (9.3b) is an $\mathcal{A}$-relative ξ-invariant set.*

Proof. The constant c_{36} may change from line to line in the proof. As $u_{ts} - u_{ttt} = D_tV_u$, we get

$$-2u_{ttt}u_{ts} = -u_{ttt}^2 - u_{ts}^2 + (D_tV_u)^2,$$

thus

$$\frac{d}{ds}u_{tt}^2 = 2u_{tt}u_{tts} = 2(u_{tt}u_{ts})_t - 2u_{ttt}u_{ts} = 2(u_{tt}u_{ts})_t - u_{ttt}^2 - u_{ts}^2 + (D_tV_u)^2.$$

Calculating the weighted integral, we get by partial integration and the Young's inequality:

$$\begin{aligned}\frac{d}{ds}||u_{tt}||^2_{L^2_\tau(\mathbb{R})} &= -||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} - ||u_{ts}||^2_{L^2_\tau(\mathbb{R})} + 2\int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|}(u_{tt}u_{ts})_t dt + ||D_tV_u||^2_{L^2_\tau(\mathbb{R})}\\ &\le -||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} - ||u_{ts}||^2_{L^2_\tau(\mathbb{R})} + \lambda(\tau)^2||u_{tt}||^2_{L^2_\tau(\mathbb{R})} + ||u_{ts}||^2_{L^2_\tau(\mathbb{R})} + ||D_tV_u||^2_{L^2_\tau(\mathbb{R})}\end{aligned}$$

$$\le -||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} + \lambda(\tau)^2||u_{tt}||^2_{L^2_\tau(\mathbb{R})} + ||D_tV_u||^2_{L^2_\tau(\mathbb{R})}. \tag{17.33}$$

By careful differentiation, while using uniform bounds on f and its derivatives, and as $\mu \le 1$ and $\varpi \ge 1$ by definition, we get

$$\begin{aligned}|D_tV_u| &\ll \varepsilon\mu|u_t| + \varepsilon|u_t| + \varepsilon\mu(1 - \cos u + |\sin u|)|v_t| + \varepsilon\mu(1 - \cos u + |\sin u|))\\ &\ll \varepsilon|u_t - u_t^0| + \varepsilon|u_t^0| + \varepsilon|v_t - v_t^0| + \varepsilon|v_t^0 - \omega_{k(t)}| + \varepsilon(1 - \cos u + |\sin u|)\varpi.\end{aligned}$$

Now by weighted integration and applying (17.10), (17.11), (17.1), (9.2a) and (9.2b) we obtain

$$||D_tV_u||^2_{L^2_\tau(\mathbb{R})} \le c_{36}\,\varepsilon^2(M^2 + \varpi^2)\lambda(\tau). \tag{17.34}$$

From (17.15), (17.29) and $\lambda(\tau) \ll \sqrt{\varepsilon}$ we get

$$\frac{d}{ds}||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} \leq -\frac{1}{2}||u_{tt} - u_{tt}^0||^2_{L^2_\tau(\mathbb{R})} + c_{36}\varepsilon\lambda(\tau) \leq -\frac{1}{4}||u_{tt}||^2_{L^2_\tau(\mathbb{R})} + c_{36}\varepsilon\lambda(\tau). \tag{17.35}$$

Now, as $q \in \mathcal{B}_2$, we have that $0 \leq -||u_t - u_t^0||_{L^2_\tau(\mathbb{R})} + c_9\lambda(\tau)$. Multipying it with ε^2 and (17.35) with ε and summing it, we obtain

$$\frac{d}{ds}\varepsilon||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})} \leq -\frac{1}{4}\varepsilon||u_{tt}||^2_{L^2_\tau(\mathbb{R})} - \varepsilon^2||u_t - u_t^0||_{L^2_\tau(\mathbb{R})} + c_{36}\varepsilon^2\lambda(\tau). \tag{17.36}$$

Denote by $Z = ||u_{tt}||^2_{L^2_\tau(\mathbb{R})} + \varepsilon||u_t - u_t^0||^2_{L^2_\tau(\mathbb{R})}$. Summing (17.33), (17.34) and (17.36) and using $\lambda(\tau)^2 \leq \varepsilon/8$, we obtain

$$\begin{aligned}\frac{d}{ds}Z &\leq -||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} - \frac{1}{4}\varepsilon Z + c_{36}(M^2 + \varpi^2)\varepsilon^2\lambda(\tau) \\ &\leq -\frac{1}{4}\varepsilon Z + c_{36}\varepsilon^2(M^2 + \varpi^2)\lambda(\tau),\end{aligned} \tag{17.37}$$

which by the Gronwall's lemma and $\mathcal{A}$-relative ξ-invariance of $\mathcal{B}_2$ proves the first part of the claim.

Analogously we obtain

$$\begin{aligned}\frac{d}{ds}||v_{tt}||^2_{L^2_\tau(\mathbb{R})} &\ll -||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} + \lambda(\tau)^2||v_{tt}||^2_{L^2_\tau(\mathbb{R})} + ||D_tV_v||^2_{L^2_\tau(\mathbb{R})}, \\ |D_tV_v| &\ll \varepsilon\mu\left(|u_t - u_t^0| + |u_t^0| + |v_t - v_t^0| + |v_t^0 - \omega_{k(t)}| + (1-\cos u)\varpi\right),\end{aligned} \tag{17.38}$$

thus as $\mu \leq 1$,

$$||D_tV_v||^2_{L^2_\tau(\mathbb{R})} \leq c_{36}\,\varepsilon^2(M^2 + \varpi^2)\lambda(\tau). \tag{17.39}$$

If $\tilde{Z} = ||v_{tt}||^2_{L^2_\tau(\mathbb{R})} + \varepsilon||v_t - v_t^0||^2_{L^2_\tau(\mathbb{R})}$, we analogously to (17.37) obtain

$$\begin{aligned}\frac{d}{ds}\tilde{Z} &\leq -||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} - \frac{1}{4}\varepsilon\tilde{Z} + c_{36}\varepsilon^2(M^2 + \varpi^2)\lambda(\tau) \\ &\leq -\frac{1}{4}\varepsilon\tilde{Z} + c_{36}\varepsilon^2(M^2 + \varpi^2)\lambda(\tau),\end{aligned} \tag{17.40}$$

which completes the proof analogously as for Z. □

Lemma 17.8. *There exists a constant $c_{11} > 0$ such that the set of all $q \in \mathcal{B}_3$ satisfying (9.4a), (9.4b) is an $\mathcal{A}$-relative ξ-invariant set.*

Proof. Analogously as in (17.33), (17.38), and as $\lambda(\tau)$ we easily get

$$\frac{d}{ds}||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} \leq -||u_{tttt}||^2_{L^2_\tau(\mathbb{R})} + \frac{\varepsilon}{16}||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} + ||D_{tt}V_u||^2_{L^2_\tau(\mathbb{R})}, \tag{17.41}$$

$$\frac{d}{ds}||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} \leq -||v_{tttt}||^2_{L^2_\tau(\mathbb{R})} + \frac{\varepsilon}{16}||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} + ||D_{tt}V_v||^2_{L^2_\tau(\mathbb{R})}. \tag{17.42}$$

By the Sobolev inequalities and as $q \in \mathcal{B}_2$ and $\lambda(\tau) \leq \sqrt{\varepsilon}$, we easily deduce that for $q \in \mathcal{B}_3$,

$$||u||_{L^\infty(\mathbb{R})} \ll (M + \varpi)\sqrt{\varepsilon},$$
$$||v - \omega_{k(.)}||_{L^\infty(\mathbb{R})} \ll (M + \varpi)\sqrt{\varepsilon}.$$

Using that and analogously as in the Proof of Lemma 14.7 (the calculation is analogous and routine, thus omitted), we see that

$$||D_{tt}V_u||^2_{L^2_\tau(\mathbb{R})} \ll (M^4 + \varpi^4)\varepsilon^2\lambda(\tau), \tag{17.43}$$

$$||D_{tt}V_v||^2_{L^2_\tau(\mathbb{R})} \ll (M^4 + \varpi^4)\varepsilon^2\lambda(\tau). \tag{17.44}$$

Setting

$$W = ||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} + ||u_{tt}||^2_{L^2_\tau(\mathbb{R})} + \varepsilon||u_t - u^0_t||^2_{L^2_\tau(\mathbb{R})} = ||u_{ttt}||^2_{L^2_\tau(\mathbb{R})} + Z,$$
$$\tilde{W} = ||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} + ||v_{tt}||^2_{L^2_\tau(\mathbb{R})} + \varepsilon||v_t - v^0_t||^2_{L^2_\tau(\mathbb{R})} = ||v_{ttt}||^2_{L^2_\tau(\mathbb{R})} + \tilde{Z},$$

and summing (17.33),(17.41) and (17.43); respectively (17.38), (17.42) and (17.44), we obtain

$$\frac{d}{ds}W \leq -\frac{1}{4}\varepsilon W + c_{37}\varepsilon^2(M^4 + \varpi^4)\lambda(\tau),$$
$$\frac{d}{ds}\tilde{W} \leq -\frac{1}{4}\varepsilon \tilde{W} + c_{37}\varepsilon^2(M^4 + \varpi^4)\lambda(\tau).$$

which by the Gronwall's lemma and $\mathcal{A}$-relative ξ-invariance of $\mathcal{B}_3$ completes the proof. ☐

18. Appendix D: The parity lemma

This Appendix is dedicated to the proof of Lemma 12.4 in Section 12. For clarity of the argument, we give definitions and prove a generalized claim in an abstract setting.

Assume U is an open, bounded subset of $\mathbb{R}^2$. Let $\tilde{u}, \tilde{v} : [s_0, s_1] \times [t_0, t_1]$ be continuous functions satisfying the following properties:

(i) If $\tilde{u}(t, s) = 0$, then $(t, \tilde{v}(s, t)) \notin \partial U)$,

(ii) For all $s \in [s_0, s_1]$, $u(s_0, t_0) < 0$ and $u(s_0, t_1) > 0$.

We say that $\tilde{u}, \tilde{v}$ intersect U for some $s \in [s_0, s_1]$, if there exists $t \in [t_0, t_1]$ such that $\tilde{u}(s, t) = 0$ and $(t, \tilde{v}(s, t)) \in U$ (clearly by (i), $(t, \tilde{v}(s, t))$ is then in the interior of U). We can count the number of times $\tilde{u}, \tilde{v}$ intersect U in the following sense. For a fixed s, let $\Upsilon(s) = \{t_0, t_1\} \cup \{t \in [t_0, t_1],\ (t, v(s, t)) \in \partial U)\}$. By assumptions, $\tilde{u}(t, s) \neq 0$ for all $t \in \Upsilon(s)$. We define the relation of equivalence $\sim$ on $\Upsilon(s)$ with $t_1 \sim t_2$ whenever for all $t_3 \in \Upsilon(s)$ such that $t_1 \leq t_3 \leq t_2$, we have that $\tilde{u}(t_1, s)$, $\tilde{u}(t_2, s)$ and $\tilde{u}(t_3, s)$ have the same sign. Let $\tilde{\Upsilon}(s) = \Upsilon(s)/\sim$ with the induced topology. As by assumptions $\Upsilon(s)$ is a closed subset of a compact set, $\tilde{\Upsilon}(s)$ is compact. By (i) and the compactness of ∂U, we see that $\tilde{\Upsilon}(s)$ is totally disconnected, thus $\tilde{\Upsilon}(s)$ is finite. By (ii), $|\tilde{\Upsilon}(s)| \geq 2$.

Consider $|\tilde{\Upsilon}(s)| - 1$ segments $(t_k, t_{k+1}) \subset \Upsilon(s)^c$, where $t_k, t_{k+1} \in \Upsilon(s)$ and $t_k \not\sim t_{k+1}$. Then there exists at least one zero $t \in (t_k, t_{k+1})$ (i.e. $\tilde{u}(t, s) = 0$), and all the zeroes $t \in (t_k, t_{k+1})$ are either all in U or all in $\bar{U}^c$ (i.e. for all $t \in (t_k, t_{k+1})$ such that $\tilde{u}(t, s) = 0$, we have either that for all such t, $(t, \tilde{v}(s, t)) \in U$, or for all such t, $(t, \tilde{v}(s, t)) \in \bar{U}^c$).

Definition 18.1. We say that $\tilde{u}$, $\tilde{v}$ as above for a given $s \in [s_0, s_1]$ intersect U exactly $n(s)$ times, if $n(s)$ is the number of segments $(t_k, t_{k+1}) \subset \Upsilon(s)^c$ for which $t_k, t_{k+1} \in \Upsilon(s)$ and $t_k \not\sim t_{k+1}$, such that all the zeroes $t \in (t_k, t_{k+1})$ are in U.

Now we have:

Proposition 18.1. The Parity Lemma. *The parity of $n(s)$ is constant for all $s \in [s_0, s_1]$.*

In particular, we have that if $n(s_0)$ is odd, then $\tilde{u}, \tilde{v}$ intersect U for all $s \in [s_0, s_1]$.

Proof. Consider for some $\delta \geq 0$ the closed δ-neighborhood of ∂U, denoted by $U(\delta)$ (clearly $U(0) = \partial U$). By the assumptions, $U(\delta)$ is compact; and by compactness, there exists $\delta_0 > 0$ so that there are no zeroes in $U(\delta_0)$ (i.e. for all $(s, t) \in [s_0, s_1] \times [t_0, t_1]$, if $(t, \tilde{v}(s, t)) \in U(\delta_0)$, then $\tilde{u}(s, t) \neq 0$).

If $0 \leq \delta \leq \delta_0$, let $\Upsilon(s, \delta)$ be all t such that $(t, v(s, t)) \in U(\delta)$. Analogously as in the case $\delta = 0$, $\Upsilon(s, 0) = \Upsilon(s)$, we define $\tilde{\Upsilon}(s, \delta)$ which is again finite with cardinality ≥ 2, and $n(s, \delta) \leq |\tilde{\Upsilon}(s, \delta)| - 1$.

Fix $s \in [s_0, s_1]$ for now. Note that $\delta \mapsto |\tilde{\Upsilon}(s, \delta)|$, $\delta \mapsto n(s, \delta)$ are non-decreasing for $\delta \in [0, \delta_0]$. This follows from the construction and the natural continuous embedding $\Upsilon(s, \delta) \to \Upsilon(s, \delta')$ for $\delta < \delta'$. Thus $|\tilde{\Upsilon}(s, \delta)|$, $n(s, \delta)$ strictly increase for at most finitely many times $0 < \delta_1(s) < ... < \delta_{k(s)}(s) \in [0, \delta_0]$. Again by construction it is easy to see that these $\delta_j(s)$, $j = 1, ..., k(s)$ are characterized as these $\delta \in [0, \delta_0]$ for which at least one entire equivalence class in $\Upsilon(s, \delta)$ lies on the boundary of $U(\delta)$.

Furthermore, for some $\delta_j(s) > 0$, we have that $n(s, \delta_j(s)) - n(s, \delta_j(s)^-)$ is even. Indeed, by considering the equivalence classes which are in $\tilde{\Upsilon}(s, \delta_j(s))$, but are not in $\tilde{\Upsilon}(s, \delta_j(s)^-)$, we see that the number of equivalence classes must increase by an even number (as the signs of $u(s,t)$ alternate on equivalence classes), and also $n(s, \delta_j(s)) - n(s, \delta_j(s)^-)$ must be an even number (as the zeroes "appear" in pairs of zeroes in U, respectively $\bar{U}^c$).

Now, if $\delta \notin \{\delta_1(s), ...\delta_{k(s)}(s)\}$, it is easy to see that for some small neighborhood $(s - \nu, s + \nu)$, $\nu > 0$, we have that for all $s^* \in (s - \nu, s + \nu)$, $|\tilde{\Upsilon}(s^*, \delta)|$ and $n(s^*, \delta)$ are constant. This is clearly true, as then no equivalence classes in $\Upsilon(s, \delta)$ are entirely on the boundary of $U(\delta)$, thus persist for sufficiently small (uniformly by finiteness) perturbation. Also $|\tilde{\Upsilon}(s, \delta)| = |\tilde{\Upsilon}(s, \delta')|$ for sufficiently small $\delta' - \delta > 0$, so there can be no new equivalence classes in $\Upsilon(s^*, \delta)$ for sufficiently small ν.

We deduce from all that that the parity of $n(s)$ does not change on some small interval $(s - \nu, s + \nu)$. We see that by choosing some $\delta \neq \delta_j(s)$, $0 < \delta < \delta_0$. Then the parity of $n(s) = n(s, 0)$ and $n(s, \delta)$ is the same, and $n(s, \delta)$ does not change for sufficiently small $\nu > 0$. The claim follows easily by contradiction. □

Proof of Lemma 12.4. We set $\tilde{u}(s,t) = u(s,t) - (2k+1)\pi$, $\tilde{v}(s,t) = v(s,t)$, $t_0 = \tilde{T}_{k-1}$, $t_1 = \tilde{T}_{k+1}$, and $U = \tilde{\mathcal{N}}_k$. Now the property (i) follows from the definition of $\mathcal{A}$, and (ii) by Lemma 12.3. Proposition 18.1 thus implies Lemma 12.4. □

References

[1] S. Angenent, A variational interpretation of Mel'nikov's function and exponentially small separatrix splitting. Symplectic geometry, 5-35, London Math. Soc. Lecture Note Ser., **192**, Cambridge Univ. Press, 1993.

[2] V.I. Arnold, Instability of dynamical systems with several degrees of freedom, Soviet. Math. Dokl. **5** (1964), 581-585.

[3] A. Avila, S. Crovisier, A. Wilkinson, Diffeomorphisms with positive metric entropy, arXiv:1408.4252v3.

[4] P. Bernard, Perturbation d'un hamiltonien partiellement hyperbolique, C. R. Acad. Sc. Paris Ser. I Math. **323** (1996), 189-134.

[5] P. Bernard, The dynamics of pseudographs in convex Hamiltonian systems, J. Amer. Math. Soc **21** (2008), 615-669.

[6] P. Bernard, Arnold diffusion: from the a priori unstable to the a priori stable case. *International Congress of Mathematicians. Vol. III, 1680-1700, Hindustan Book Agency, 2010.*

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

[8] M. Berti, P. Bolle, A functional analysis approach to Arnold diffusion, Ann. Inst. H. Poncaré Anal. Non Linéaire **19** (2002), 395-450.

[9] M. Berti, P. Bolle, Fast Arnold diffusion in systems with three time scales, Discrete Contin. Dyn. Syst. **8** (2002), 795-811.

[10] M. Berti, L. Biasco and P. Bolle, Drift in phase space: A new variational mechanism with optimal diffusion time, J. Math. Pures Appl. **82** (2003), 613-664.

[11] U. Bessi, An approach to Arnold diffusion through the calculus of variations, Nonlinear Analysis, Theory, Models, Applications **26** (1996), 1115-1135.

[12] C. Cheng and J. Yan, Existence of diffusion orbits in a-priori unstable Hamiltonian systems, J. Differential Geom. **67** (2004), 457-517.

[13] C. Cheng and J. Yan, Arnold diffusion in Hamiltonian systems: a priori unstable case, J. Differential Geom. **82** (2009), 229-277.

[14] C.Q. Cheng, Variational construction of diffusion ofbits for positive definite Lagrangians. *International Congress of Mathematicians. Vol. III, 1714-1728, Hindustan Book Agency, 2010.*

[15] L Cherchia, G. Gallavotti, Drift and diffusion in phase space, Ann. Inst. H. Poincaré Phys. Théor. **60** (1994), 144 pp.

[16] L. C. Evans, Partial Differential Equations (2nd ed.), American Mathematical Society, 2010.

[17] A Fathi, Weak KAM theorem in Lagrangian dynamics, book preprint, *5th edn.*

[18] E. Fontich, P. Martin, Differentiable invariant manifolds for partially hyperbolic tori and a lambda lemma, Nonlinearity **13** (2000), 1561-1593.

[19] Th. Gallay and S. Slijepčević, Energy flow in formally gradient partial differential equations on unbounded domains. J. Dynam. Differential Equations 13 (2001), 757-789. Zbl 1003.35085, MR1860285

[20] Th. Gallay and S. Slijepčević, Distribution of Energy and Convergence to Equilibria in Extended Dissipative Systems, arXiv 1212.1573.

[21] Th. Gallay and S. Slijepčević, Energy bounds for the two-dimensional Navier-Stokes equations in an infinite cylinder, Comm. Partial Differential Equations **39** (2014), 1741-1769.

[22] Th. Gallay and S. Slijepčević, Distribution of energy and convergence to equilibria in extended dynamical systems, J. Dynam. Differential Equations **27** (2015), 653-682.

[23] V. Gelfreich, D. Turaev, Arnold diffusion in a priory chaotic Hamiltonian systems, preprint, arXiv1406.2945.

[24] V. Kaloshin and K. Zhang, Arnold diffusion for smooth convex systems of two and a half degrees of freedom, Nonlinearity **28** (2015), 2699-2720.
[25] O. Knill, The problem of positive Kologorov-Sinai entropy for the standard map, arXiv.math/9908014v2.
[26] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics **840**, Springe-Verlag, Berlin, 1981.
[27] V. Kaloshin and K. Zhang, A strong form of Arnold diffusion for two and a half degrees of freedom, preprint.
[28] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1995. Zbl 0878.58020, MR1326374
[29] R. de la Llave, Some recent progress in geometric methods in the instability problem in Hamiltonian mechanics. *International Congress of Mathematicians. Vol. II,1705-1729, Eur. Math. Soc., 2006.*
[30] P. Lochak, Arnold diffusion; a compendium of remarks and questions. *Hamiltonian systems with three or more degrees of freedom*, 168-183, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 533, Kluwer, 1999.
[31] J. Mather, More Denjoy minimal sets for area preserving diffeomorphisms, Comment. Math. Helv. **60** (1985), 508-557.
[32] R. S. MacKay, J. D. Meiss and I.C. Percival, Transport in Hamiltoonian systems, Phys. D **13** (1983), 55-81.
[33] J. Mather, Action minimizing measures for positive definite Lagrangian systems, Math. Zeitschrift **207** (1991), 169-207.
[34] J. Mather, Variational construction of connecting orbits, Ann. Inst. Fourier (Grenoble) **43** (1993), 1349-1386.
[35] J. Mather, Arnold diffusion by variational mathods. *Essays in mathematics and its applications*, 271-285, Springer, 2012.
[36] A. Miranville, S. Zelik, Attractors for dissipative partial differential equations in bounded and unbounded domains, Handbook of differential equations: evolutionary equations. Vol. IV, 103-200, Elsevier/North-Holland, Amsterdam, 2008.
[37] N. Nekhoroshev, An exponential estimate of the time of stability of Hamiltonian systems close to integrable, Usp. Mater. Nauk. **32** (1977), 5-66 (in Russian); Russian Math. Surveys **32** (1977), 1-65 (Engl. Transl.).
[38] S. I. Pohožaev, The set of critical values of functionals, Mat. Sb. (N.S.) **75** (1968), 106-111.
[39] C. Simo, C. Valls, A formal approximation of the splitting of separatrices in the classical Arnold's example of diffusion with two equal parameters, Nonlinearity **14** (2001), 1707-1760.
[40] S. Slijepčević, Monotone gradient dynamics and and Mather's shadowing, Nonlinearity **12** (1999), 969-986.
[41] S. Slijepčević, Extended gradient systems: dimension one, Discrete Contin. Dyn. Syst (2000), 503-518. Zbl 1009.37004, MR1757384
[42] S. Slijepčević, Construction of invariant measures of Lagrangian maps: minimisation and relaxation, Math. Z. **237** (2001), 469-504.
[43] S. Slijepčević, Variational construction of positive entropy invariant measures of Lagrangian systems and Arnold diffusion II, in preparation.
[44] A. Sorrentino, Lecture Notes on Mather's Theory for Lagrangian Systems, preprint.
[45] D. Treschev, Evolution of slow variables in a-priori unstable Hamiltonian systems, Nonlinearity qbf 17 (2004), 1803-1841.
[46] D. Treschev, Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems, Nonlinearity **25** (2012), 2717-2757.
[47] P. Walters, An Introduction to Ergodic Theory, Springer, 2000.
[48] K. Zhang, Speed of Arnold diffusion for analytic Hamiltonian systems, Invent. Mathematique **186** (2011), 255-290.

Authors' addresses: *Siniša Slijepčević*, Department of Mathematics, Bijenička 30, University of Zagreb, Croatia e-mail: `slijepce@math.hr`.