Variational construction of positive entropy invariant measures of Lagrangian systems and Arnold diffusion

SINIŠA SLIJEPČEVIĆ
Department of Mathematics, Bijenička 30, University of Zagreb, Croatia
(e-mail: slijepce@math.hr)

(Received 26 August 2016 and accepted in revised form 26 June 2018)

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 $2 \frac{1}{2}$ 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 Peierls barrier), then there exists a vast family of ‘horseshoes’, 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.

Contents

1 Introduction 800
   1.1 Statements of the main results 802
   1.2 The structure of the proof and notation 808
2 Variational construction of orbits and invariant measures 809
   2.1 Existence of solutions and the function spaces 809
   2.2 Existence of Euler–Lagrange orbits in invariant sets 811
   2.3 Construction of shadowing invariant measures 813
3. Introduction

Consider a $C^2$, Tonelli Lagrangian $L : \mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T} \to \mathbb{R}$ (see [39] 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, 12–14, 24, 31, 32, 40, 41], and on the fastest possible drift [8, 9, 43] (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, 26] 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. Furthermore, the geometric method typically relies on
transversal intersections of stable and unstable manifolds of the ‘scaffolding’ (e.g. the remaining Kolmogorov–Arnold–Moser (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 by leveraging the weak-KAM theory [17], or the description of the action-minimizing Mather sets and invariant measures and their extensions [31, 39], which can result with insightful descriptions of the regions of instability [5, 14].

Our aim is to propose an alternative technique for construction of Arnold diffusion, enabling, in addition to construction of individual orbits, a construction of ‘shadowing’ invariant measures in an \textit{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 ‘drift’ 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 [30] in the case of area-preserving twist diffeomorphisms on the cylinder, or equivalently of \(1\frac{1}{2}\) degrees of freedom Tonelli Lagrangians on the torus. In [36], we considered formally gradient dynamics of the action (see §2.1 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 the LaSalle principle the actual solution of the Euler–Lagrange equations. One of the novelties in [36] 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 2.2 for details). The construction in [36] 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 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 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 set. 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 [36]—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 [18], have recently matured enough [19–21], 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 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.

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 [19, 21], 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 without assuming decay to zero at infinity. 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.

After introducing the general method of constructing shadowing invariant measures, we apply it here to a family of $2\frac{1}{2}$ 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 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 Peierls barrier function, or of weak KAM solutions. For example, we applied an analogous approach to this paper in [38] to obtain lower bounds on the topological entropy in perhaps the simplest non-trivial example: a neighborhood of a hyperbolic fixed point of an area-preserving twist diffeomorphism.

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 $2\frac{1}{2}$ degrees of freedom, already considered in e.g. [2, 4, 8, 9, 40], given with

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

(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 $2\pi$-periodic in all the variables. Our standing assumptions are as
follows:

(A1) $f$ is $C^{4+\gamma}(\mathbb{R}^3)$, $\gamma > 0$ and $|f| \leq 1$, $|f_\nu| \leq 1$;
(A2) $0 \leq 16\mu \leq \epsilon \leq 1$.

Note that the bounds on $f$ in (A1) are not an essential restriction, as we can always scale $f$ and adjust $\mu$ 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 $\mu$, as well as other physically relevant cases such as those considered in [35].

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

$$
\begin{align*}
    u_s &= u_{tt} - \partial_u V(u,v,t), \\
    v_s &= v_{tt} - \partial_v V(u,v,t), \\
    u(0,t) &= u^0(t).
\end{align*}
$$

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)

$$
\begin{align*}
    u_{tt} &= V_u = \epsilon \sin u(1 - \mu f(u,v,t)) - \epsilon \mu (1 - \cos u)f_u(u,v,t), \\
    v_{tt} &= V_v = -\epsilon \mu (1 - \cos u)f_v(u,v,t).
\end{align*}
$$

Let $\phi$ 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 $\sigma(u,v,p_1,p_2) = \phi(u,v,p_1,p_2,2\pi)$ the time-2$\pi$ map, and then $\sigma$ is a diffeomorphism of $\mathbb{T}^2 \times \mathbb{R}^2$. As in the classical Arnold’s example [2], for each $\epsilon, \mu \in \mathbb{R}$, 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 $\sigma$-invariant (and the sets $(0,v,0,\omega,t)$, $(v,t) \in \mathbb{R} \times \mathbb{T}$ are $\phi$-invariant), i.e. these invariant tori persist for all perturbations.

We define regions of instability, generalizing an analogous notion for area-preserving twist maps [30] as follows. Given $\omega \in \mathbb{R}$, let $S_\omega$ be the Peierls barrier function (closely related to weak-KAM solutions of Hamilton–Jacobi equations [17, 39]), defined on $\mathbb{R}^2$ with

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

where $L_\omega(q,q_t,t) = u^2/2 + (v_t - \omega)^2/2 + V(u,v,t)$ is the adjusted Lagrangian. It is well known, and recalled in §3.1, that $S_\omega$ is continuous and 2$\pi$-periodic in each variable, that for each $(t_0,v_0)$ there exists a $q \in H^1_{loc}(\mathbb{R})^2$ for which the infimum is attained, and that, for $(t_0,v_0)$ which are critical points of $S_\omega$, 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 \( S_\omega \) is bounded. We interpret it as ‘no topological obstacle to diffusion’ (see Remark 1.1(1)). The set of non-degenerate \( \omega \) is open (see Remark 3.1), 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 \( S_\omega \) 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 \( S_\omega \), there exists a closed neighborhood \( N \) of \((t_0, v_0)\) such that, for each \((t_1, v_1) \in \partial N\),

\[
S_\omega(t_1, v_1) - S_\omega(t_0, v_0) \geq 3\Delta_0,
\]

where \( \Delta_0 > 0 \) is a uniform constant over \([\omega^-, \omega^+]\), and the diameter of \( N \) is not greater than \( R \), such that

\[
R \sqrt{\varepsilon} \leq 1/144.
\]

Apart from the technical restriction (1.5) (see Remark 1.1(2)), the condition (S1) is equivalent to the definition of the region of instability (see Remark 3.1). Our definition of the region of instability is closely related to the results of Cheng and Yan [12–14] and Bernard [5, 6] (see Remarks 1.1(3) and (4)). We note that the quantity \( 3\Delta_0 \) in (1.4) is analogous to the quantity \( 1W \) quantifying transport in the case of area-preserving twist maps [28] (see Remark 1.1(5)).

Importantly, we do not require any non-degeneracy of the minima of \( S_\omega \), for example we do not require that the second derivative of \( S_\omega \) 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 \( T_\omega \) intersect transversally, and would lead to the ‘geometric’ approach to diffusion phenomena ([26] 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 ‘horseshoes’, 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, \( \phi \)-invariant, positive entropy Borel probability measure \( \mu \) on \( \mathbb{T}^2 \times \mathbb{R} \times \mathbb{R} \), such that

\[
\bigcup_{\omega \in \mathcal{O}} T_\omega \subset \text{supp} \mu, \quad \bigcup_{\omega \in \mathbb{R} - \mathcal{O}} T_\omega \subset (\text{supp} \mu)^c.
\]

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 \( \bigcup_{\omega \in [\omega_1, \omega_2]} T_\omega \) (we choose \( \mathcal{O} = [\omega^-, \omega^+] \), and by ergodicity we find an orbit which is dense in the support of the measure \( \mu \)). Actually, by our method we also give a direct proof of existence of such and more elaborated shadowing orbits in Theorem 4.2. One can apply it to construct other complex structures in a vicinity of a chosen \( T_\omega \) for non-degenerate \( \omega \), such as an analogue of Mather’s construction in the case of twist maps [29] (see Remark 1.1(7)).
Importantly, we are also able to obtain estimates of the metric entropy of constructed measures, topological entropy of $\sigma$ and $\phi$ (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 $\Delta_1$, expressed as a certain extremal value of the norm of the Fréchet derivative of the action, and defined precisely in Proposition 3.21. The value $\Delta_1$ can be understood as a generalization of the lower bound on the angle of splitting of separatrices (see Remark 1.1(7)). Importantly, $\Delta_1 > 0$ whenever (S1) holds. Let $\sigma = \max(|\omega^-|, |\omega^+|, 1)$. Then we have the following.

To avoid repetition in the statements, we say that the topological entropy on $[\omega^-, \omega^+]$ is much greater than $h$, if there exists an absolute constant $1 \geq c_0 > 0$ such that the topological entropy of $\phi$ and $\sigma$ restricted to an invariant subset of

$$
|u_t| \leq c_0 \sigma \sqrt{\varepsilon},
$$

is at least $c_0 h$. We say that the drift acceleration on $[\omega^-, \omega^+]$ is much greater than $d$, if for each $\delta > 0$ there exists an absolute constant $1 \geq c(\delta) > 0$ (depending only on $\delta$), 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 3.21. Then we have the following.

(i) The topological entropy of $\phi$ and $\sigma$ on $[\omega^-, \omega^+]$ is much greater than

$$
\frac{\Delta_1}{\sigma^5 |\log \Delta_1|}.
$$

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

$$
\frac{\Delta_0 \Delta_1}{\sigma^6 (R \vee \mu) |\log \Delta_1|}.
$$

Our estimates are consistent with upper bounds on the drift acceleration in the cases considered by Nekhoroshev [33] (see Remark 1.1(8)), as well as on upper bounds in [8] (see Remark 1.1(9)).

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 $\mu_\omega$ in a neighborhood of $\mathbb{T}_\omega$ with the metric entropy as in Theorem 1.2(ii). By the Margulis–Ruelle inequality, there is a positive Lyapunov exponent with respect to $\mu_\omega$, 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 $\omega$ 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. [10, 22] and references therein).

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

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

(1.8)
where \( u^e = 4 \arctg e^{\sqrt{\varepsilon}} t \) 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 \( \mu \) and fixed \( \omega \), the oscillations of \( M_\omega \) approximate well the oscillations of \( S_\omega \), the minima of \( M_\omega \) approximate the minima of \( S_\omega \), and non-degeneracy of the minima of \( M_\omega \) implies non-degeneracy of minima of \( S_\omega \) (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 \( M_\omega \).

(S2) For each \( \omega \in [\omega^-, \omega^+] \), and for each global minimum \((t_0, v_0)\) of \( M_\omega \), there exists a closed neighborhood \( \mathcal{N} \) 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 diameter 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 each \( 0 < \mu \leq \mu_0 \), the conclusions of Theorem 1.1 and 1.2 hold, with \( \Delta_0 = \mu \tilde{\Delta}_0 \).

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 \( M_\omega \) 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 Arnold’s example for \( \varepsilon \leq 1 \) and sufficiently small \( \mu > 0 \) (see Remark 1.1(10)), without restrictions to \( \omega \), as long as the sign of \( \omega \) does not change. One can tailor the argument to also obtain accelerating orbits, i.e. orbits with unbounded \( \omega \) in that case (and any case when \( f \) does not depend on \( u \)).

**Remarks 1.1.**

1. Consider the stable and unstable manifolds of \( T_\omega \) of the time-2\( \pi \) map \( \sigma \). Then an unbounded family of global minima of \( S_\omega \) corresponds to an unbounded, connected family of homoclinic orbits of \( \sigma \), which cannot be ‘crossed’ by other orbits on two-dimensional stable and unstable manifolds of \( \sigma \). 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 3.31, at the end of §3.8, to assist in the cases of topologically complex \( \mathcal{N} \) in (S1). 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 \( \mu \) 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 \( \omega \).

3. The definition of the region of instability by Cheng and Yan [12–14] 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 \(\omega\) 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 \(\Delta W\) quantifying transport through gaps in Cantori \([28]\).

(6) Mather in \([29]\) constructed uncountably many minimal sets of twist maps with the same irrational angular rotation \(\omega\). One can adapt our construction in the proof of Theorem 1.2 to construct uncountably many minimal sets for irrational \(\omega\), 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 partially explain the relationship between \(\Delta_1\) and the angle of splitting of separatrices by analogy and will express the precise relationship in a future work. Consider a sufficiently smooth function \(S : \mathbb{R}^n \to \mathbb{R}\) with a local minimum \(x = 0\).

If the minimum is non-degenerate and \((D^2 S(0), x) = A ||x||^2\), for small enough \(\delta > 0\) on the level sets \(S(x) = A\delta^2\) in a neighborhood of \(0\) we trivially have that \(||D S(x)|| \geq 2A\delta + O(\delta^2)\). If, however, we merely assume that the set of local minima of \(S\) is bounded, we can, by the Morse–Sard theorem, find level sets of \(S\) arbitrarily close to the set of minima, such that \(||D S(x)|| \neq 0\) 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 \(||D S(x)|| \geq \Delta_1 > 0\) on any such level set. Now we take \(S_\omega\) instead of \(S\), acting on a suitable Banach space (see \([1]\) or §3.6).

If the stable and unstable manifolds of \(\mathbb{T}_\omega\) intersect transversally, \(D^2 S_\omega\) 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 \(S_\omega\) on a level set of \(S_\omega\) is proportional to \(A\) (we actually take the square of the norm and find a level set which maximizes \(\Delta_1\)). If, however, we only assume that the set of minima of \(S_\omega\) is bounded, we analogously to the finite-dimensional case obtain \(\Delta_1 > 0\) by an application of an infinite-dimensional analogue of the Morse–Sard theorem (see the proof of Proposition 3.21).

(8) In the cases considered by Nekhoroshev such as Arnold’s example \([2]\), both \(\Delta_0\) and \(\Delta_1\) are exponentially small in \(\epsilon\), which is consistent with \([33]\). They are, however, polynomial in \(\mu\)—see below.

(9) The ‘fast diffusion’ (with respect to the perturbation \(\mu\)) for sufficiently small \(\mu\) has the drift acceleration \(O(\mu/|\log \mu|)\), as conjectured by Lochak \([27]\), 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 \(\geq 1\)) with non-transversal intersection of whiskers by Berti, Biasco and Bolle, and by Treschev \([8, 40, 41]\). In \([8]\), 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 Arnold’s example to arbitrary $\mu$ (and other cases of (1.1)), by developing a computer-assisted proof verifying (S1) in the spirit of [35], by using all the a priori bounds we develop here.

1.2. The structure of the proof and notation. In §§2.1–2.3, 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 §2.1 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 §2.2 show that an invariant set with respect to the formally gradient dynamics bounded in norm contains in its closure a solution of the Euler–Lagrange equations. The key tool is then developed in §2.3, where we extend these ideas to construction of shadowing invariant measures.

In §§3.1–3.8, 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 $\xi$ is an abstract continuous semiflow on a separable metric space $X$, and let $A, \tilde{B}$ be subsets of $X$. Assume they satisfy the following.

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

(B2) There exists $\lambda > 0$ such that, if $q(s_0) \in A \cap \tilde{B}$, then, for all $s \in [s_0, s_0 + \lambda]$, $q(s) \in A$.

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

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

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

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

We construct the sets $A, \tilde{B}$ as follows. Let $\xi$ be the ‘formally gradient semiflow’ introduced in §2.1. We define the set $A$ in §3.3 by very roughly fixing the times of ‘jumps’ between invariant tori $T_\omega$. The set $\tilde{B}$ satisfying the conditions (B1), (B2) is then built using the ‘Russian doll’ approach. We construct a decreasing sequence of sets $B_1 \supset B_2 \supset \cdots \supset B_6$, showing inductively in each step that they are $A$-relatively $\xi$-invariant. Finally, in $B_6$ we establish sufficient control to also prove by an energy method the condition (B2), as required by Lemma 1.4. Specifically, in §§3.1 and 3.2 we recall the known results on homoclinic, heteroclinic orbits and the Peierls barrier, and prove a priori bounds on minimizing homoclinics and heteroclinics. In §3.3, we fix a sequence of tori $T_{\omega k}, k \in \mathbb{Z}$, 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 §§3.4 and 3.5, construct
sets $B_1, B_2, B_3, B_4$, by establishing $L^\infty$-bounds, as well as weighted $L^2$-bounds on the first, second and third derivatives. The core of the argument is then in §§3.6 and 3.7, where we establish local control of the dynamics $\xi$ 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 $\xi$, where ‘energy’ is in this case the Lagrangian action. The approach is inspired by the results from [18–21]. One of the novelties is the use of an infinite-dimensional version of the Morse–Sard theorem, enabling us to establish lower positive 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 cannot be crossed by $\xi$, 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 $\xi$, 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 $B$, as a function of a given sequence of tori to be shadowed, in §3.8. The last step of the construction is somewhat subtle (the set $B\theta$), and uses in a topological way the existence of the semiflow $\xi$.

In §§4.1 and 4.2, 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 $\xi$, introduced in §2.1. 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 $\xi$, which by correspondences $\pi_1$ and $\pi$ specified in §2.2 are the solutions of the Euler–Lagrange equations. The fixed velocities $v_\ell$ specifying the invariant tori $T_\omega$ are denoted by $\omega$. The sequence of ‘jump’ times is given by $\tilde{T}_k$, 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 §3.3. Various constants are fixed throughout the proof: the constants $L$ (minimal length of time $t$ between the ‘jumps’) and $M$ (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 $\sigma = \max\{|\omega^-|, |\omega^+|, 1\}$ depend on the choice of $f$ and the segment $[\omega^-, \omega^+]$ in a region of instability. We denote by $c_1, c_2, \ldots$ 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 $\varepsilon$ 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 $\sigma$-algebras, if $\mu$ 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 by $f^* \mu(C) = \mu(f^{-1}(C))$.

2. Variational construction of orbits and invariant measures

2.1. Existence of solutions and the function spaces. In this and the next two subsections, we consider the Lagrangian $L(q, q_1, t) = \frac{1}{2}q_1^2 + V(q, t)$, $L : \mathbb{T}^N \times \mathbb{R}^N \times \mathbb{T} \to \mathbb{R}$, where $\mathbb{T}$ is always parametrized with $[0, 2\pi)$, and $V$ is $C^2$, $2\pi$-periodic in all
coordinates. Here we prove existence of solutions of the formally gradient dynamics of the action, given by

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

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

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

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

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

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

**Theorem 2.1.** Assume \( q^0 \in \mathcal{X} \) at \( s_0 \) is the initial condition. Then we have the following.

(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 \( \xi \) on \( \mathcal{X}_{loc} \) and \( \mathcal{X}_{ul} \).

(iii) The semiflow \( \xi \) and the flow \( \varphi \) commute.

(iv) If \( V \in H^k(T^{N+1}) \), \( k \geq 1 \), 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 [23], only on a larger space than usual, and is given in Appendix A.

**Remark 2.1.** All the results of §2 hold if we consider a more general, \( C^2 \) Tonelli Lagrangian \( L(q, q_t, t) \) on \( T^N \times R^N \times T \) (see e.g. [31, 39] 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 [18, 20], and that the proofs of Theorems 2.2 and 2.9 can be generalized to hold. We develop the theory in a simpler case for clarity of the introduced ideas.
2.2. Existence of Euler–Lagrange orbits in invariant sets. Consider equilibria (or stationary solutions) of (2.1), i.e. the solutions of

\[ q_{tt} = \frac{\partial}{\partial q} V(q, t). \tag{2.3} \]

We denote by \( \mathcal{E} \) the set of all \( q \in \mathcal{X} \) satisfying (2.3). 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 (2.3), \( \pi_t \) is continuous in both of the topologies \( \mathcal{X}_{ul} \), \( \mathcal{X}_{loc} \) induced on \( \mathcal{E} \). Furthermore, we have that, for any \( t_1, t_2 \in \mathbb{R} \), \( \pi_t \circ \phi^{t_1} = \phi^{t_2} \circ \pi_{t_1} \), 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 (2.3). 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 \( S = \phi^{2\pi} \) 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 A.5 in Appendix A that if a solution of (2.3) satisfies \( \|q_t\|_{L^\infty(\mathbb{R})^N} < \infty \), then \( q \in \mathcal{X} \), and thus \( q \in \mathcal{E} \).

The result of this subsection is that, to construct elements of \( \mathcal{E} \), i.e. to find solutions of (2.3), it suffices to find an invariant set of (2.1).

**Theorem 2.2.** Assume \( \mathcal{B} \) is a non-empty, \( \xi \)-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 2.2.** We comment why Theorem 2.2 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 [18, 20]. These are dynamical systems which, when considered on bounded domains, are gradient-like; but on an unbounded domain they may behave differently. For example, consider the system of equations \( q_t = \Delta q - \nabla V(q) \), \( q : \mathbb{R}^M \to \mathbb{R}^N \), \( V : \mathbb{R}^N \to \mathbb{R} \) smooth and bounded from below. If we require that \( q \) at \( t = 0 \) decays fast enough at \( \infty \), the system is gradient-like, and \( \omega \)-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, then, for \( M = 1, 2 \), \( \omega \)-limit sets may contain non-equilibria, but always contain at least one equilibrium, and for \( M \geq 3 \), there are examples of \( \omega \)-limit sets without equilibria at all \([18, 20] \). As in (2.1), \( M = 1 \) (the dimension of the variable \( t \)), and Theorem 2.2 holds. We adapt the proof from [18, 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 2.3.** Assume \( q(s) \in \mathcal{B} \), \( s \geq s_0 \) is an orbit of \( \xi \). There exists an absolute constant \( c_1 > 0 \) and a sequence of relaxation times \( s_n, n \in \mathbb{Z} \), so that

\[
\int_{-n}^{n} q_s(t, s_n)^2 \, dt \leq \frac{c_1}{n} (\|q(s_0)\|_{\mathcal{X}_{ul}}^2 + 1).
\]
Proof. Let $\delta > 0$, and denote by $E_\delta$, $D_\delta$ the weighted action and action dissipation,

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

It is straightforward to verify that $E_\delta$, $D_\delta$ are well defined on $X$ (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 $X$ by continuity. We partially integrate and apply the Young’s inequality in the second line below:

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

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

$$e^{-1} E_\delta(q(s_0 + 1/\delta^2)) + \frac{1}{2} \int_{s_0}^{s_0 + 1/\delta^2} e^{-(s-s_0)\delta^2} D_\delta(q(s)) \, ds \leq E_\delta(q(s_0)). \quad (2.4)$$

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

$$E_\delta(q) \ll \frac{1}{\delta}(\|q\|_{X_{ul}}^2 + 1). \quad (2.4a)$$

Also by definition, $L(q, q_t, t) \geq 0$, and thus $E_\delta(q) \geq 0$. By combining this with $e^{-(s-s_0)\delta^2} \leq 1$ and (2.4a) and inserting it all in (2.4) we obtain

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

and thus, by definition of $D_\delta$,

$$\int_{s_0}^{s_0 + 1/\delta^2} \left( \int_{-1/\delta}^{1/\delta} q_s(t)^2 \, dt \right) \, ds \ll \frac{1}{\delta}(\|q(s_0)\|_{X_{ul}}^2 + 1). \quad (2.5)$$

Now set $\delta = 1/n$. From (2.5) 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 (2.5). □

Lemma 2.4. If $B$ is bounded in $X_{ul}$, then it is relatively compact in the closure $\bar{X}_{loc}$ of $X_{loc}$ in $H^1_{loc}(\mathbb{R})^N$.

Proof. It is easy to check from the definition of the $X_{ul}$-norm, that boundedness of $B$ in $X_{ul}$ implies boundedness of $q|_{[-n,n]}$ in $H^2([-n, n])^N$, uniformly for $q \in B$, for any $n > 0$. Thus for any sequence $q^{(j)}$ in $B$, 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 $H^1_{loc}(\mathbb{R})^N$ (which induces the $X_{loc}$ topology by definition). □
Lemma 2.5. 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 \( \bar{X}_{loc} \) is in \( \mathcal{E} \).

Proof. Fix \( m \in \mathbb{Z} \), and choose a test function \( g \in H^1_0([-m, m])^2 \) (that is, vanishing at \( t = -m, m \)). Now by partial integration and the Cauchy–Schwartz inequality,
\[
\int_{-m}^{m} \left( q_{t}(t, s_{n})g_{t}(t) + \frac{\partial V}{\partial q}(q(t, s_{n}), t)g(t) \right) \, dt = \int_{-m}^{m} (-q_{s}(t, s_{n})g(t)) \, dt \\
\leq \left( \int_{-m}^{m} q_{s}(t, s_{n})^2 \, dt \right)^{1/2} \| g \|_{L^2([-m, m])^2}.
\]
Now if \( q(s_n, \cdot) \) converges to some \( q^0 \) in \( \bar{X}_{loc} \), their restrictions to \([-m, m]\) converge in \( H^1([-m, m])^2 \). We deduce that
\[
\int_{-m}^{m} \left( q^0_{t}(t)g_{t}(t) + \frac{\partial V}{\partial q}(q^0(t), t)g(t) \right) \, dt = 0.
\]
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 (2.3). By construction, \( q_t \in L^2_{ul}(\mathbb{R}^N) \), and thus by Lemma A.5 we have that \( q \in \mathcal{E} \). \( \square \)

Theorem 2.2 follows by combining Lemmas 2.3, 2.4 and 2.5.

2.3. Construction of shadowing invariant measures. In §2.2, we showed how to construct a solution of (2.3), given an invariant set with respect to the semiflow \( \xi \). In this subsection we develop a measure-theoretical analogue to that. In the first subsubsection, we propose an abstract notion of a shadowing measure and derive its properties. In the second subsubsection, we prove existence of such measures, if a certain sub-algebra of Borel sets with certain invariance property with respect to the semiflow \( \xi \) is given.

2.3.1. Shadowing of invariant measures in an abstract setting. We propose an abstract definition of a shadowing invariant measure as follows. In this subsubsection we will always consider a measurable space \((\Omega, \mathcal{F})\), where \( \Omega \) is a compact metric space and \( \mathcal{F} \) the Borel \( \sigma \)-algebra. Let \( S \) be a homeomorphism on \( \Omega \) and \( \mu \) a \( S \)-invariant probability measure on \((\Omega, \mathcal{F})\). Recall that \( \mu \) is a factor of a \( S \)-invariant probability measure \( \nu \) 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 \( \theta \) pulls the measure \( \nu \) into \( \mu \), i.e. for any set \( D \in \mathcal{F} \), \( \nu(\theta^{-1}(D)) = \mu(D) \) (where we extended \( \theta \) to a measurable function on the entire \( \Omega \) in an arbitrary way).

Definition 2.1. Let \( \mathcal{G} \) be a \( \sigma \)-subalgebra of \( \mathcal{F} \). We say that a \( S \)-invariant Borel-probability measure \( \nu \) \( \mathcal{G} \)-shadows a \( S \)-invariant probability measure \( \mu \) on \((\Omega, \mathcal{F})\), if \( \mu \) is a factor of \( \nu \), and if, for each \( D \in \mathcal{G} \), we have \( \mu(D) = \nu(D) \).
We now show in several lemmas 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 \( \mu \) with respect to a \( \sigma \)-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 \( \bigcap_{j \in \mathbb{N}} \mathcal{D}_j = \{x\} \). 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

\[
\text{supp}(\mu|\mathcal{G}) \subseteq \text{supp}(\mu) \subseteq \text{supp}^c(\mu|\mathcal{G})^c. \tag{2.6}
\]

The following lemma follows directly from the definitions.

**Lemma 2.6.** Assume that \( \nu \) \( \mathcal{G} \)-shadows \( \mu \). 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 2.7.** Assume that \( \nu \) \( \mathcal{G} \)-shadows \( \mu \), that \( \mu \) 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 \( \nu \) \( \mathcal{G} \)-shadows \( \mu \).

**Proof.** Consider the ergodic decomposition of \( \nu \), i.e. a Borel-probability measure \( \chi \) on the compact, metrizable space of probability measures \( \mathcal{M}(\Omega) \) (equipped with the weak*-topology), such that \( \chi \)-a.e. measure (i.e. almost every measure with respect to \( \chi \)) is \( S \)-invariant and ergodic, and such that the usual representation formula for \( \nu \) in terms of \( \chi \) holds [42]. Then it is straightforward to check by verifying the definition of the ergodic decomposition [42] that \((\theta^*)^\ast \chi \) is the ergodic decomposition of \( \mu \), where \((\theta^*)^\ast \) is the double pull defined in a natural way. However, the ergodic decomposition is unique, and as \( \mu \) is ergodic, \((\theta^*)^\ast \chi \) must be concentrated on \( \mu \). That means that, for \( \chi \)-a.e. measure \( \tilde{\nu} \) (i.e. almost every measure in the ergodic decomposition of \( \nu \)), we have \( \theta^*(\tilde{\nu}) = \mu \). By construction, \( \mu \) is then a factor of \( \tilde{\nu} \).

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

\[
\tilde{\nu}(\mathcal{D}) \geq \tilde{\nu}(\theta^{-1}(\mathcal{M}_1 \cap \mathcal{D})) = \mu(\mathcal{M}_1 \cap \mathcal{D}) = \mu(\mathcal{D}), \tag{2.7}
\]

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 (2.7) must hold.

Finally, we establish relation of shadowing to the metric (or Kolmogorov–Sinai) entropy \( h_{\mu}(S) \) of a measure \( \mu \).

**Lemma 2.8.** If \( \nu \) \( \mathcal{G} \)-shadows \( \mu \), then \( h_\nu(S) \geq h_{\mu}(S) \).

**Proof.** This holds, as entropy is always non-increasing under factor maps and \( \mu \) is a factor of \( \nu \) [42].
2.3.2. Variational construction of shadowing measures. In this subsubsection, \( \mathcal{X} \) 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 by \( \pi(q) = (q \mod 2 \pi, q_t) \). 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 \( i : \mathcal{X} \to \hat{\mathcal{X}} \) be the canonical projection, and let \( \hat{\xi}, \hat{\varphi} \) be the induced semi-flow \( \xi \) and flow \( \varphi \) on \( \hat{\mathcal{X}} \). By (1.2) and by definition, \( \hat{\xi} \) and \( \hat{\varphi} \) are well defined. If \( \hat{\mathcal{E}} = i(\mathcal{E}) \), \( S = \varphi^{2\pi} \), \( \hat{S} = \hat{\varphi}^{2\pi} \) are the \( 2\pi \)-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:

\[
\begin{array}{cccc}
\mathcal{X} & \xrightarrow{i} & \hat{\mathcal{X}} \\
\downarrow \xi & & \downarrow \hat{\xi} \\
\mathcal{X} & \xrightarrow{i} & \hat{\mathcal{X}}
\end{array}
\quad
\begin{array}{cccc}
\mathcal{E} & \xrightarrow{i} & \hat{\mathcal{E}} & \xrightarrow{\hat{\pi}} & \mathbb{T}^N \times \mathbb{R}^N \\
\downarrow S & & \downarrow \hat{S} & & \downarrow \sigma \\
\mathcal{E} & \xrightarrow{i} & \hat{\mathcal{E}} & \xrightarrow{\hat{\pi}} & \mathbb{T}^N \times \mathbb{R}^N
\end{array}
\]

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 \( \cdot \). 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 \( \phi \)-, or equivalently \( \sigma \)-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 \( \sigma \)-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 \( \cdot^* \) 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:

\[
\begin{array}{cccc}
\mathcal{M}(\mathcal{X}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{X}}) \\
\downarrow \xi^* & & \downarrow \hat{\xi}^* \\
\mathcal{M}(\mathcal{X}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{X}})
\end{array}
\quad
\begin{array}{cccc}
\mathcal{M}(\mathcal{E}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{E}}) & \xrightarrow{\hat{\pi}^*} & \mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N) \\
\downarrow S^* & & \downarrow \hat{S}^* & & \downarrow \sigma^* \\
\mathcal{M}(\mathcal{E}) & \xrightarrow{i^*} & \mathcal{M}(\hat{\mathcal{E}}) & \xrightarrow{\hat{\pi}^*} & \mathcal{M}(\mathbb{T}^N \times \mathbb{R}^N)
\end{array}
\]

(by definition, \( S^*, \hat{S}^* \) and \( \sigma^* \) 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 \( \xi^* \) 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 \( \mu \), as an element of the \( \sigma \)-limit set of \( \mu \) 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 \( \sigma \)-subalgebra \( \mathcal{G} \) of the \( \sigma \)-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 \( \{ D \cap \mathcal{M}_1, \ D \in \mathcal{G} \} \) generates all Borel-measurable sets on \( \mathcal{M}_1 \). Specifically, for each \( q \in \mathcal{M}_1 \), there exists \( D_q \in \mathcal{G} \) such that if \( q, \tilde{q} \in \mathcal{M}_1, \ q \neq \tilde{q} \), then \( D_q \cap D_{\tilde{q}} = \emptyset \). Furthermore, for any \( q \in \mathcal{M}_1 \), \( D_{\hat{S}(q)} = \hat{S}(D_q) \).

(M2) The \( \xi \)-invariance. For each \( q \in \mathcal{M}_1 \) and each \( D \in \mathcal{G} \), if \( q \in D \), then for all \( s \geq 0 \), \( \xi^s(q) \in D \).
(M3) **Measurability.** If $\mathcal{M}_2 = \bigcup_{q \in \mathcal{M}_1} \mathcal{D}_q$, then the map $\hat{\theta} : \mathcal{M}_2 \to \mathcal{M}_1$ given by $\hat{\theta}(\mathcal{D}_q) = q$ 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 $\sigma$-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 $\mu(\bigcup_{n=1}^{\infty} \mathcal{D}_{i_n}) = 1$.

In applications, (M1), (M3) and (M4) will follow relatively easily from the construction of $\mu$, and the focus will be on ensuring the $\xi$-invariance of the constructed $\sigma$-algebra $\mathcal{G}$. An important tool in the construction of the shadowing measure, already suggested in [37], is that $\xi^*$ and $\hat{\xi}^*$ are gradient-like semiflows with the Lyapunov function (given below for $\hat{\xi}^*$)

$$\hat{\xi}^*(\mu) = \int \hat{\chi} \int_0^{2\pi} L(q(t), q_t(t), t) dt d\mu(q)$$

(2.8)

(see Lemma 2.10 below). We fix $\mu \in \mathcal{M}(\hat{\chi})$ and write $\mu(s) = \hat{\xi}^*(\mu, s)$ for $s \geq 0$, i.e. $\mu(0) = \mu$, and $\mu(s)$ is the pulled measure $\mu$ with respect to the map $\hat{\xi}^*$.

**Theorem 2.9.** (Variational construction of shadowing measures) Assume $\mu \in \mathcal{M}(\hat{\chi})$ such that $\|q_t\|_{\mathcal{H}_N^1(\mathbb{R})}$ is bounded on the support of $\mu(s)$, uniformly in $s \geq 0$. Assume $\mathcal{G}$ is a $\sigma$-subalgebra of Borel sets on $\hat{\chi}$ satisfying (M1)–(M4). Then there exists $\nu \in \mathcal{M}(\hat{\chi})$ which $\mathcal{G}$-shadows $\mu$.

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

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

**Lemma 2.10.** The function $s \to \hat{\xi}^*(\mu(s))$ is strictly decreasing, unless $\mu(0) \in \mathcal{M}(\hat{\chi})$, in which case it is constant. Furthermore,

$$\frac{d}{ds} \hat{\xi}^*(\mu(s)) = -\int \hat{\chi} \int_0^{2\pi} q_s(t)^2 dt d\mu(s)(q).$$

(2.9)

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

$$\frac{d}{ds} \int_0^{2\pi} L(q, q_t, t) dt = \int_0^{2\pi} \left( \frac{d}{ds} L(q, q_t, t) \right) dt = \int_0^{2\pi} \left( q_t q_{ts} + \frac{\partial}{\partial q} V(q, t) q_s \right) dt ds$$

$$= \int_0^{2\pi} \left( -q_t q_s + \frac{\partial}{\partial q} V(q, t) q_s \right) dt + q_t(2\pi) q_s(2\pi) - q_t(0) q_s(0)$$

$$= -\int_0^{2\pi} q_s^2 dt + q_t(2\pi) q_s(2\pi) - q_t(0) q_s(0).$$

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

$$\int_0^{s_1} L(q(s_0), q_t(s_0), t) dt - \int_0^{s_2} L(q(s_1), q_t(s_1), t) dt = -\int_{s_0}^{s_1} \int_0^{2\pi} q_s(s, t)^2 dt ds$$

$$+ \int_{s_0}^{s_1} (q_t(s, 2\pi) q_s(s, 2\pi) - q_t(s, 0) q_s(s, 0)) ds.$$
By the dominated convergence theorem, we can extend (2.10) also to $0 \leq s_0 < s_1$. By the assumptions we have for any $q \in \text{supp } \mu$,

\[
\int_{\hat{X}} \int_{s_0}^{s_1} \int_{0}^{2\pi} |q_t(s, t)q_s(s, t)| \, dt \, ds \, d\mu(q)
\leq \int_{\hat{X}} \int_{s_0}^{s_1} \left( \int_{0}^{2\pi} (q_t(s, t)^2 + q_s(s, t)^2) \, dt \right)^{1/2} \, ds \, d\mu(q)
\leq \int_{\hat{X}} \int_{s_0}^{s_1} \|q_t(s)\|_{H^1_{\text{ul}}(\mathbb{R})}^2 \, ds \, d\mu(q) \leq A \cdot s_1,
\]

where $A$ is the uniform bound $\|q_t(s)\|_{H^1_{\text{ul}}(\mathbb{R})}$. 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 $\mu$ (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 $\mu$, we now have for any $s_0, 0 \leq s_0 < s_1$,

\[
\int_{\hat{X}} \int_{s_0}^{s_1} q_t(s, 0)q_s(s, 0) \, ds \, d\mu(q) = \int_{\hat{X}} \int_{s_0}^{s_1} q_t(s, 2\pi)q_s(s, 2\pi) \, ds \, d\mu(q).
\]

Integrating (2.10) with respect to $\mu$, we now get for any $0 \leq s_0 < s_1$,

\[
\hat{L}^s(\mu(s_0)) - \hat{L}^s(\mu(s_1)) = -\int_{\hat{X}} \int_{s_0}^{s_1} \int_{0}^{2\pi} q_s(s, t)^2 \, dt \, ds \, d\mu(q).
\]

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

**Lemma 2.11.** There exists $\nu \in \mathcal{M}(\hat{X})$ which is a weak*-limit of a subsequence of $\mu(s)$, $s \geq 0$.

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

It suffices to show that $\nu \in \mathcal{M}(\hat{X})$. 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 $\mu$ in the third row below, we obtain

\[
\int_{\hat{X}} \left| \int_{-2n\pi}^{2n\pi} (q_t(s_n, t)h_t(t) + DV(q(s_n, t), t)h(t)) \, dt \right| \, d\mu(q)
\]
which by Lemma 2.10 converges to zero. We thus have that, for any $h \in H^1(\mathbb{R})^N$ with compact support, $\int_{\mathbb{R}^d} |\partial L(q, q_t, t) h| \, dv(q) = 0$. By choosing such a countable, dense set of $h$, and by continuity, we conclude that $v$ is supported on the solutions of (1.3). By construction, $v$ is supported on $q$ such that $q_t \in L^2_\text{ul}(\mathbb{R})^N$, and thus, by Lemma A.5, $v \in \mathcal{M}(\hat{S})$.

**Proof of Theorem 2.9.** Take $v$ constructed in Lemma 2.11. We apply results from §2.3.1 with $\Omega = \mathcal{Y}$, $\mathcal{Y}$ as in Lemma 2.11, thus compact and metrizable. As $v$ is a weak*-limit of $\hat{S}$-invariant measures and $\hat{S}$ is continuous, $v$ is $\hat{S}$-invariant. It suffices to show that $v$ shadows $\mu$. We take as the factor the function $\hat{Y} : M_2 \to M_1$ from the property (M3). Let $s_k > 0$ be the sequence from Lemma 2.11 such that $v$ is the weak* limit of $\mu(s_k)$.

First we show that, for each $D \in \mathcal{G}$, $v(D) = \mu(D)$. It suffices to show it for the generator $D_i$, $i \in \mathcal{I}$, 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 $\mu(D_{i_n}) = \mu(D_{i_n} \cap M_1) \leq \mu(s_k)(D_{i_n})$ for all $n \in \mathbb{N}$. However, by (M4) and $\sigma$-additivity of $\mu$ and $\mu(s_n)$, we get

$$1 = \sum_{n=1}^{\infty} \mu(D_{i_n}) \leq \sum_{n=1}^{\infty} \mu(s_k)(D_{i_n}) = \mu(s_k)\left( \bigcup_{n=1}^{\infty} D_{i_n} \right) \leq 1,$$

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

$$1 = \sum_{n=1}^{\infty} \mu(D_{i_n}) \leq \sum_{n=1}^{\infty} v(D_{i_n}) = v\left( \bigcup_{n=1}^{\infty} D_{i_n} \right) \leq 1,$$

and thus again equality must hold in all the terms.

By (M1) and the definition of $\hat{\sigma}$, $\hat{\theta}$ and $\hat{S}$ commute. As $v(D) = \mu(D) = \mu(D \cap M_1)$, $D \cap M_1$ generate all Borel-measurable sets on $M_1$, to show that $\hat{\theta}$ is measure-preserving it suffices to show that, for all $i \in \mathcal{I}$,

$$v(D_i) = v(\hat{\theta}^{-1}(D_i \cap M_1)).$$

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 $D_{i_n} \subset \hat{\theta}^{-1}(D_{i_n} \cap M_1)$, and thus $v(D_{i_n}) \leq v(\hat{\theta}^{-1}(D_{i_n} \cap M_1))$. By (M1) and (M3), the sets $\hat{\theta}^{-1}(D_{i_n} \cap M_1)$, $n \in \mathbb{N}$, are pairwise disjoint. Analogously as above, from all of this and $\sum_{n=1}^{\infty} v(D_{i_n}) = 1$ we conclude that (2.11) must hold.

If $\mu$ is $\hat{S}$-ergodic, we can find a $\hat{S}$-ergodic $v$ by Lemma 2.7.

3. Construction of invariant sets

3.1. The homoclinic orbits. As of this subsection, we focus on the *a priori* unstable case with the Lagrangian (1.1) and $N = 2$. In this subsection we recall the key properties
of the Peierls barrier function and stable and unstable manifolds of the invariant tori \( \mathbb{T}_\omega \). The results of this subsection are standard (see [17, 39] and references therein). As we were unable to find in the literature the \emph{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 Peierls barrier functions defined with

\[
S_\omega^-(t_0, v_0) = \inf \left\{ \int_{-\infty}^{t_0} L_\omega(q(t), q_1(t), t) \, dt, \; q = (u, v) \in H^1_{\text{loc}}((-\infty, t_0))^2, \quad q(t_0) = (\pi, v_0), \lim_{t \to -\infty} u(t) = 0 \right\},
\]

\[
S_\omega^+(t_0, v_0) = \inf \left\{ \int_{t_0}^{\infty} L_\omega(q(t), q_1(t), t) \, dt, \; q = (u, v) \in H^1_{\text{loc}}([t_0, \infty))^2, \quad q(t_0) = (\pi, v_0), \lim_{t \to \infty} u(t) = 2\pi \right\}.
\]

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 3.4(i) below). We call them one-sided (left-, respectively right-hand) sided minimizers at \( (\omega, t_0, v_0) \).

We first obtain \emph{a priori} bounds on \( S_\omega^-, S_\omega^+ \). We then introduce the notion and construct specific super- and sub-solutions of (1.2) required in this subsection and later, and finally construct one-sided minimizers and prove explicit \emph{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

\[
|u^-(t)| \leq C_2 e^{-\sqrt{2}|t-t_0|} \quad \text{for all } t \leq t_0, \quad |u^+(t) - 2\pi| \leq C_2 e^{-\sqrt{2}|t-t_0|} \quad \text{for all } t \geq t_0.
\]

Lastly, we prove continuity of \( S^- \), \( S^+ \) in \( \omega, t_0, v_0 \) and estimate the Lipschitz constant in \( \omega \).

\textbf{Lemma 3.1.} \textit{For all } \omega, t_0, v_0 \in \mathbb{R}, \textit{we have}

\[
4\sqrt{\varepsilon(1-\mu)} \leq S_\omega^-(t_0, v_0), \quad S_\omega^+(t_0, v_0) \leq 4\sqrt{\varepsilon(1+\mu)}.
\]

\textit{Proof.} By definition,

\[
\int_{t_0}^\infty L_\omega(q, q_1, t) \, dt \leq \int_{t_0}^\infty \left( \frac{1}{2} u_1^2 + \frac{1}{2} (v_1 - c)^2 + (\varepsilon(1+\mu))(1 - \cos u(t)) \right) dt.
\]

It is well known [2, 11] that the right-hand side of (3.3) attains a minimum for the separatrix solution of the pendulum equation \( u^0(t) = 4 \arctg e^{\sqrt{\varepsilon(1+\mu)} (t-t_0)}, u^0(t) = \omega(t - t_0) + v_0 \), and that the value of the integral on the right-hand side of (3.3) is then by direct calculation

\[
4\sqrt{\varepsilon(1+\mu)} \leq \int_{t_0}^\infty \left( \frac{1}{2} u_1^2 + \frac{1}{2} (v_1 - c)^2 + (\varepsilon(1-\mu))(1 - \cos u(t)) \right) dt
\]

\[
\leq \int_{t_0}^\infty L_\omega(q, q_1, t) \, dt,
\]

which completes the proof for \( S_\omega^+(t_0, v_0) \). The bounds for \( S_\omega^-(t_0, v_0) \) are analogous. \( \square \)
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 \(-\infty \leq t_0 < t_1 \leq \infty, -\infty \leq s_0 < s_1 \leq \infty\), if it is continuous on \( \bar{U} \) and for any \((s, t) \in U\), \( u_s - u_{tt} + \partial_u V(u, v, t) \leq 0\). Analogously we say that \( q = (u, v) \) is a sub-solution of (1.2a) on \( U \), if \( u_s - u_{tt} + \partial_u V(u, v, t) \geq 0\). 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 \( I = (t_0, t_1) \), \(-\infty \leq t_0 < t_1 \leq \infty\), 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 \( t \in (t_0, t_1)\),

\[
z_{tt} - \partial_u V(z, v, t) \leq 0.
\]

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 \( \varepsilon \). 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 3.2.** Assume \( z \) is a strict stationary super-solution of (1.2a) on \((t_0, t_1)\), \(-\infty \leq t_0 < t_1 \leq \infty\), assume \( q = (u, v) \in E\), and let \( u(t) \leq z(t) \) for all \( t \in (t_0, t_1)\). Then for all \( t \in (t_0, t_1)\), \( u(t) < z(t)\).

**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

\[
u_{tt}(t_2) - \partial_u V(u(t_2), v(t_2), t_2) \leq z_{tt}(t_2) - \partial_u V(z(t_2), v(t_2), t_2) < 0,
\]

which is in contradiction to \( q \in E\). \(\square\)

**Lemma 3.3.** There exist \( z^- : (-\infty, 3/(4\sqrt{\varepsilon})) \to \mathbb{R} \) and \( z^+ : [-3/(4\sqrt{\varepsilon}), \infty) \to \mathbb{R} \), depending only on \( \varepsilon, \mu \), satisfying the following 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 \( T \geq 0 \), \( z^-(t + T) \) and \( z^+(t - T) \) are strict stationary super-, respectively sub-solutions in the interior of their domain of definition.
(iv) \( z^-, z^+ \) are strictly increasing and we have

\[
\sqrt{\varepsilon}/2 < z^-_t(t), z^+_t(t) \text{ for all } t \in (-1/(4\sqrt{\varepsilon}), 1/(4\sqrt{\varepsilon})). \tag{3.4}
\]

(v) For all \( t \in [1/(4\sqrt{\varepsilon}), 3/(4\sqrt{\varepsilon})]\), \( z^-(t) \leq \pi + 1/4 \) and \( z^+(-t) \geq \pi - 1/4 \).
(vi) There exists an absolute constant \( c_2 > 0 \) such that, for all \( t \) in the domains of definition,

\[
|z^-_t(t)| \leq c_2 e^{-(1/2)\sqrt{\varepsilon}|t|}, \quad |z^+_t(t) - 2\pi| \leq c_2 e^{-(1/2)\sqrt{\varepsilon}|t|}. \tag{3.5}
\]
(vii) There exists an absolute constant $c_3 > 0$ such that
\[
|z^- (t) - u^{(e)} (t)| \leq c_3 \sqrt{\epsilon \mu} \quad t \leq 0, \quad |z^+ (t) - u^{(e)} (t)| \leq c_3 \sqrt{\epsilon \mu} \quad t \geq 0, \tag{3.6}
\]
where $u^{(e)} (t) = 4 \arctg e^{\sqrt{\epsilon} t}$ is the separatrix solution in the case $\mu = 0$.

An explicit construction of $z^-$, $z^+$ and the proof of Lemma 3.3 is given in Appendix B.

**Proposition 3.4.** 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 $(\omega, t_0, v_0)$ satisfy the following.

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

(ii) For all $0 \leq T \leq 3/(4 \sqrt{\epsilon})$,
\[
0 < u^- (t) \leq z^- (t - t_0 + T) \quad \text{for all } t \leq t_0, \tag{3.7}
\]
\[
z^+ (t - t_0 - T) \leq u^+ (t) < 2 \pi \quad \text{for all } t \geq t_0. \tag{3.8}
\]

The proof is in Appendix B. (Existence and (i) are a consequence of the Tonelli theorem [30, Appendix 1], and the a priori bounds follow from Lemma 3.2 applied to $z^-$, $z^+$ constructed in Lemma 3.3.)

Combining (3.5), (3.7) and (3.8) with $T = 0$ we get the following.

**Corollary 3.5.** Any one-sided minimizers $q^-$, $q^+$ at $(\omega, t_0, v_0)$ satisfy (3.1).

We finally deduce the Lipschitz constant for $S^-$ and $S^+$ in the variable $\omega$.

**Corollary 3.6.** The functions $S^-$, $S^+$ are continuous in $t_0$, $v_0$, $\omega$. Furthermore, there exists an absolute constant $c_4 \geq 1$ such that for any $(t_0, v_0) \in \mathbb{R}^2$,
\[
|S^-_{\omega} (t_0, v_0) - S^-_{\omega} (t_0, v_0)| \leq c_4 \mu |\bar{\omega} - \omega|, \quad |S^+_{\omega} (t_0, v_0) - S^+_{\omega} (t_0, v_0)| \leq c_4 \mu |\bar{\omega} - \omega|. \tag{3.9}
\]

**Proof.** We fix first $(t_0, v_0)$ and show that the Lipschitz constant of $S^+_{\omega} (t_0, v_0)$ in $\omega$. Choose $\omega, \bar{\omega} \in \mathbb{R}$, and let $q = (u, v)$ be a right-hand sided minimizer constructed in Proposition 3.4 at $(\omega, t_0, v_0)$ respectively. Define
\[
\tilde{q} (t) = (\tilde{u} (t), \tilde{v} (t)) := (u (t), v (t) + (\bar{\omega} - \omega) (t - t_0)),
\]
defined for $t \in [t_0, \infty)$. By definition of $q$, $\tilde{q}$, by applying $1 - \cos u \leq (u - 2 \pi)^2/2$, (3.1) and the standing assumption (A2), we get
\[
S^+_{\omega} (t_0, v_0) \leq \int_{t_0}^{\infty} L_{\tilde{\omega}} (\tilde{q} (t), \tilde{q} (t), t) \, dt = \int_{t_0}^{\infty} \left( \frac{1}{2} u^2 + \frac{1}{2} (v - \omega)^2 + V (\tilde{u}, \tilde{v}, t) \right) \, dt
\]
\[
= S^+_{\omega} (t_0, v_0) + \int_{t_0}^{\infty} (V (\tilde{u}, \tilde{v}, t) - V (u, v, t)) \, dt
\]
\[
\leq S^+_{\omega} (t_0, v_0) + \epsilon \mu \int_{t_0}^{\infty} \{ 1 - \cos (u (t)) \}. 
\]
Let \( L = \varpi \) the argument used in Lemma 3.12, and omitted as not needed in the following. We do not require it here, as we take (S1) as the standing assumption. Thus for every global minimum \((t_0, v_0)\) of \( S_\omega \) we can find a closed, bounded set \( \tilde{\mathcal{N}}(t_0, v_0) \subset \mathbb{R}^2 \), \((t_0, v_0) \in \tilde{\mathcal{N}}(t_0, v_0)\), such that, for each \((t_1, v_1) \in \partial \tilde{\mathcal{N}}(t_0, v_0)\), (1.4) holds, and such that there is a constant \( R \geq \sup\{|x - y|, x, y \in \tilde{\mathcal{N}}(t_0, v_0)\} \) satisfying (1.5).

Remark 3.1. By continuity and periodicity of \( S_\omega \), 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 \( S_\omega \). The proof is analogous to the argument used in Lemma 3.12, 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 \( \Delta_0, R \) 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

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

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

The main result of the subsection is the following.

**Proposition 3.7.** Assume that \( \omega, \tilde{\omega} \in [\omega^-,\omega^+] \) satisfy

\[
|\omega - \tilde{\omega}| \leq \frac{\Delta_0}{4c_4(R \lor \mu) \cdot \varpi}.
\]
(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 \( (\omega, t_0, v_0) \), respectively \( (\tilde{\omega}, t_0, v_0) \).

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

\[
\Sigma_{\omega, \tilde{\omega}}(t_1, v_1) - \Sigma_{\omega, \tilde{\omega}}(t_0, v_0) \geq 2\Delta_0 > 0. \tag{3.13}
\]

We denote by \( \mathcal{H} \) the set of all \( q \in \mathcal{E} \) satisfying parts (i), (ii) of Proposition 3.7 for some \( \omega, \tilde{\omega} \in [\omega^-, \omega^+] \) satisfying (3.12). We say that such \( q \in \mathcal{H} \) is a heteroclinic minimizer connecting \( \omega, \tilde{\omega} \). Within this subsection, denote by \( (t_0, v_0) \in [0, 2\pi)^2 \) a minimizer of \( S_\omega \), fixed if non-unique, and by \( N_\omega = N(t_0, v_0) \).

**Lemma 3.8.** If (3.12) holds, then \( \Sigma_{\omega, \tilde{\omega}} \) attains a local minimum \( (t_0, v_0) \) in the interior of \( N_\omega ^\circ \), such that, for any \( (t_1, v_1) \in \partial N_\omega ^\circ \), (3.13) holds.

**Proof.** Note first that \( \Sigma_{\omega, \tilde{\omega}} \) is continuous by the definition and Corollary 3.6. Thus by the definition and compactness of \( N_\omega ^\circ \), it suffices to show that for some \( (t_2, v_2) \) in the interior of \( N_\omega ^\circ \), and any \( (t_1, v_1) \in \partial N_\omega ^\circ \), \( \Sigma_{\omega, \tilde{\omega}}(t_1, v_1) - \Sigma_{\omega, \tilde{\omega}}(t_2, v_2) \geq 2\Delta_0 > 0 \). Let \( (t_2, v_2) = (t_0, v_0) \). Then by definition, because of \( |t_1 - t_\omega| \leq R \), \( |v_1 - v_\omega| \leq R \), \( c_4 \geq 1 \) and (3.12) we obtain

\[
\Sigma_{\omega, \tilde{\omega}}(t_1, v_1) - \Sigma_{\omega, \tilde{\omega}}(t_0, v_0) \geq S_\omega ^-(t_1, v_1) - S_\omega ^-(t_0, v_0) + S_\omega ^+(t_1, v_1) - S_\omega ^+(t_0, v_0) - |\tilde{\omega} - \omega| |v_1 - v_\omega| - \sigma |\tilde{\omega} - \omega| |t_1 - t_\omega| \\
\geq S_\omega ^-(t_1, v_1) - S_\omega ^-(t_0, v_0) + S_\omega ^+(t_1, v_1) - S_\omega ^+(t_0, v_0) - \Delta_0/2. \tag{3.14}
\]

From (3.9) and \( \sigma \geq 1 \) we deduce that

\[
S_\omega ^+(t_1, v_1) - S_\omega ^+(t_0, v_0) \geq S_\omega ^+(t_1, v_1) - S_\omega ^+(t_0, v_0) - \Delta_0/2,
\]

which combined with (3.14) and (1.4) gives

\[
\Sigma_{\omega, \tilde{\omega}}(t_1, v_1) - \Sigma_{\omega, \tilde{\omega}}(t_0, v_0) \geq S_\omega (t_1, v_1) - S_\omega (t_0, v_0) - \Delta_0 \geq 2\Delta_0. \tag*{□}
\]

**Lemma 3.9.** Assume \( (t_0, v_0) \) is a local minimum of \( \Sigma_{\omega, \tilde{\omega}} \), and let

\[
q(t) = \begin{cases} 
q^-(t), & t \leq t_0, \\
q^+(t), & t \geq t_0.
\end{cases}
\tag{3.15}
\]

where \( q^- \) is the left-hand sided minimizer at \( (\omega, t_0, v_0) \), 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 \( \tilde{q} = (\tilde{u}, \tilde{v}) \in H^1(\mathbb{R}) \) such that, for some \( (t_1, v_1) \in \mathbb{R}^2 \), \( \tilde{q}(t_1) = (\pi, v_1) \), then

\[
\mathcal{L}_{\omega, \tilde{\omega}}(\tilde{q}) - \mathcal{L}_{\omega, \tilde{\omega}}(q) \geq \Sigma_{\omega, \tilde{\omega}}(t_1, v_1) - \Sigma_{\omega, \tilde{\omega}}(t_0, v_0). \tag{3.16}
\]
Indeed, by definitions and the partial integration,
\[
\mathcal{L}_{\omega, \tilde{\omega}}(\tilde{q}) = \int_{t_1}^{t_0} L_\omega(\tilde{q}(t), \dot{\tilde{q}}(t), t) \, dt \\
+ \int_{t_1}^{\infty} L_\omega(\tilde{q}(t), \dot{\tilde{q}}(t), t) \, dt + (\tilde{\omega} - \omega)v_1 + \frac{1}{2}(\omega^2 - \tilde{\omega}^2)t_1 \\
\geq S_{\omega}^-(t_1, v_1) + S_{\tilde{\omega}}^+(t_1, v_1) + (\tilde{\omega} - \omega)v_1 + \frac{1}{2}(\omega^2 - \tilde{\omega}^2)t_1 = \Sigma_{\omega, \tilde{\omega}}(t_1, v_1).
\]
By the definition of \( q \), we obtain an equality in an analogous calculation for \( q \), thus \( \mathcal{L}_{\omega, \tilde{\omega}}(q) = \Sigma_{\omega, \tilde{\omega}}(t_0, v_0) \). This gives (3.16).

We now claim that, for any \( h \in H^1(\mathbb{R})^2 \), \( \partial \mathcal{L}(q)h = \lim_{\delta \to 0} (1/\delta)(\mathcal{L}(q + \delta h) - \mathcal{L}(q)) \) 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 \( \delta > 0 \), \( \mathcal{L}(q + \delta h) \geq \mathcal{L}(q) \), and that \( \partial \mathcal{L}(q)h \) exists. Consider an open neighborhood \( U \) of \( (t_0, v_0) \) in \( \mathbb{R}^2 \) such that \( \Sigma_{\omega, \tilde{\omega}}|U \geq \Sigma_{\omega, \tilde{\omega}}(t_0, v_0) \).

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 \( 0 \leq \delta \leq \delta_0 \), 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 \( \delta \) small enough, because (3.7), (3.8) with \( T = 0 \) and the fact that \( \tilde{z}^- \), \( \tilde{z}^+ \) are strictly increasing at \( t = 0 \). We find \( v_1 = \tilde{v}(t_1) \) sufficiently close to \( v_0 \) by finding \( \delta_0 \) small enough so that all the terms in
\[
|\tilde{v}(t_1) - v_0| \leq \int_{t_0}^{t_1} |\tilde{v}_t| \, dt + |\tilde{v}(t_0) - v_0| \\
\leq \sqrt{2}|t_1 - t_0|^{1/2}\left(\int_{t_0}^{t_1} (v_t^2(t) + \delta^2(v^h_t)^2) \, dt\right)^{1/2} + \delta|v^h(t_0)|
\]
are small enough. Now combining it with (3.16) and the fact that the right-hand side in (3.16) is \( \geq 0 \) on \( U \), we obtain \( \mathcal{L}(q + \delta h) \geq \mathcal{L}(q) \) for \( \delta \leq \delta_0 \). Now, a straightforward calculation and the fact from Proposition 3.4 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 \( t = t_1 \). \( \square \)

Proposition 3.7 now follows from Lemmas 3.8 and 3.9, with \( N_q := N_{\omega_0} \).

**Lemma 3.10.** 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 (3.8) and the properties of \( z^-, z^+ \) proved in Lemma 3.3. \( \square \)

**Lemma 3.11.** 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 \( t \in \mathbb{R} \),
\[
|u(t) - 2\pi 1_{[t_0, \infty)}(t)| \leq c_5 e^{-\left(1/2\sqrt{|v - t_0|}\right)},
\]
(3.17)
\[ |u_t(t)| \leq c_5 \sqrt{\varepsilon e^{-\left(1/2\right)\sqrt{\varepsilon}|t-t_0|}}, \tag{3.18} \]

\[ |v(t) - v_0 - \omega(t - t_0)| \leq c_5 \mu, \quad |v_t(t) - \omega| \leq c_5 \sqrt{\varepsilon \mu e^{-\sqrt{\varepsilon}|t-t_0|}} \quad \text{for } t \leq t_0, \tag{3.19} \]

\[ |v(t) - v_0 - \tilde{\omega}(t - t_0)| \leq c_5 \mu, \quad |v_t(t) - \tilde{\omega}| \leq c_5 \sqrt{\varepsilon \mu e^{-\sqrt{\varepsilon}|t-t_0|}} \quad \text{for } t \geq t_0, \tag{3.20} \]

\[ |u_{tt}(t)| \leq c_5 e^{-\left(1/2\right)\sqrt{\varepsilon}|t-t_0|}, \quad |v_t(t)| \leq c_5 \sqrt{\varepsilon \mu e^{-\sqrt{\varepsilon}|t-t_0|}}, \tag{3.21} \]

\[ |u_{ttt}(t)| \leq c_5 e^{\kappa} e^{-\left(1/2\right)\sqrt{\varepsilon}|t-t_0|}, \quad |v_{ttt}(t)| \leq c_5 \sqrt{\varepsilon \mu \sigma} e^{-\left(1/2\right)\sqrt{\varepsilon}|t-t_0|}. \tag{3.22} \]

**Proof.** The absolute constant \( c_5 \) may change from line to line in the proof. The relation (3.17) follows from the definition of \( q \) and (3.1). 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 (3.17), we see that

\[ |u_{tt}(t)| \leq \varepsilon (1 - \cos u(t) + |\sin u(t)|) \ll \varepsilon e^{-\left(1/2\right)\sqrt{\varepsilon}|t-t_0|}, \]

which is the left-hand side of (3.21). 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 (3.18). 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 (3.17), 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 (3.21). 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 (3.19) and (3.20). We use \( v(t_0) = v_0 \), integrate the right-hand side of (3.19) over \([t, t_0]\), respectively the right-hand side of (3.20) over \([t_0, t]\), and obtain the left-hand sides of (3.19) and (3.20). Finally, to bound the third derivatives, by careful differentiation, while using uniform bounds on \( f \) and its derivatives, and as \( \mu \leq 1 \) and \( \sigma \geq 1 \), we obtain

\[
|u_{ttt}| = |D_t V_u(u(t), v(t), t)| \ll \varepsilon \mu |u_t| + \varepsilon |u_t|
\]

\[
+ \varepsilon \mu (1 - \cos u + |\sin u|)(|v_t| + \varepsilon \mu (1 - \cos u + |\sin u|))
\ll \varepsilon |u_t| + \varepsilon |v_t - \omega 1_{(-\infty, t_0)}(t) - \tilde{\omega} 1_{[t_0, \infty)}(t)| + \varepsilon (1 - \cos u + |\sin u|)\sigma.
\]

By inserting the bounds (3.17), (3.18) and the right-hand sides of (3.19), (3.20), we obtain the left-hand side of (3.22). Similarly we get

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

which analogously as above implies the right-hand side of (3.22).

**Lemma 3.12.** The set \( \mathcal{H} \) is compact in \( H^2_{\text{loc}}(\mathbb{R})^2 \). Furthermore, for each \( q \in \mathcal{H} \) we have that \( q \in \mathcal{E} \) and \( q_t \in H^2_{\text{ul}}(\mathbb{R})^2 \).

**Proof.** It is straightforward to observe that the closure in \( H^2_{\text{loc}}(\mathbb{R})^2 \) of all \( q \) satisfying (3.17)–(3.22) and \((t_0, v_0) \in [0, 2\pi]^2\) is compact. Thus it suffices to show that \( \mathcal{H} \) is closed in \( H^2_{\text{loc}}(\mathbb{R})^2 \). Assume \( q_n \in \mathcal{H} \) connecting \( \omega_n \) and \( \tilde{\omega}_n \), \( q_n(t_n) = (\pi, v_n) \), is a sequence converging to \( q \in H^2_{\text{loc}}(\mathbb{R})^2 \). We first show that \( q \in \mathcal{E} \). By construction, \( q \) is a solution
of (1.3), and by Lemma 3.11 and the construction we easily show that \( q \in L^\infty(\mathbb{R})^2 \). Now \( q \in \mathcal{E} \) follows from Lemma A.5. It remains to show that \( q \in \mathcal{H} \). We see that \( q \) must satisfy condition (i) from the definition of \( \mathcal{H} \), as the sequences \( \int_{-\infty}^{\infty} L_{\omega_n}(q_n, (q_n)_{\tau}, t) \, dt \), \( \int_{-\infty}^{\infty} L_{\omega_n}(q_n, (q_n)_{\tau}, t) \, dt \) are convergent by (3.17)–(3.20) and the Lebesgue dominated convergence theorem, and as \( S^+, S^- \) are continuous.

To show that \( q \) satisfies the condition (ii) from the definition of \( \mathcal{H} \), note that \( N_{q_n} \) is a family of compact sets with a bounded union, thus we can find a convergent subsequence converging to a set \( N_{q} \) in the Hausdorff topology. By the construction, the sequence \((t_n, v_n)\) converges to some \((t_0, v_0) \in N_{q} \) such that \( q(t_0) = (\pi, v_0) \). If \((\tilde{t}_0, \tilde{v}_0) \in \partial N_{q_n} \), it is a limit of a subsequence of \((\tilde{t}_{n_1}, \tilde{v}_{n_1})\) lying on the boundaries of the convergent sub-sequence of \( \partial N_{q_n} \). The relation (3.13) now follows by the continuity of \((\omega, \tilde{\omega}, t, v) \mapsto \Sigma_{\omega, \tilde{\omega}}(t, v) \), established by the definition and Corollary 3.6. □

3.3. An approximate shadowing orbit. In this subsection 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 \( q^0 \in \mathcal{A}, \mathcal{B} \) 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 \( \Delta_0, R \) as in (S1) and \( \omega^* \) as in the introduction. Assume \( \omega_k, k \in \mathbb{Z} \) is a sequence in \([\omega^-, \omega^+]\) such that, for all \( k \in \mathbb{Z} \), (3.12) holds. The constant 4\( L \) 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 \( N_{q_k} \subset \mathbb{R}^2 \) be the sets associated to \( q_k \) as in the definition of \( \mathcal{H} \). Note that if \((t_1, v_1) \in \partial N_{q_k} \), and if \( \tilde{q} = (\tilde{u}, \tilde{v}) \in H_{1}^{\text{loc}}(\mathbb{R}^2) \) 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 (3.13) and (3.16) we have

\[
L_{\omega, \omega^*}(\tilde{q}) - L_{\omega, \omega^*}(q) \geq 2\Delta_0 > 0.
\] (3.23)

Also, by the definition of \( \Delta_0 \) and (3.2), we can easily deduce (using \( \mu \leq 1/16 \) by (A2)) the useful bound

\[
\Delta_0 \leq 9\sqrt{\varepsilon \mu}.
\] (3.24)

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

\[
\tilde{T}_k = T_k \mod 2\pi \text{ so that } -\pi < \tilde{T}_k - \tilde{L}_k \leq \pi,
\]

\[
\tilde{q}_k(t) = (u_k(t - \tilde{T}_k + T_k) + 2k\pi, v_k(t - \tilde{T}_k + T_k) + \tilde{V}_k - V_k),
\]

\[
\tilde{V}_k = V_k \mod 2\pi \text{ so that } -\pi < \tilde{v}_k - \tilde{v}_{k-1}(\tilde{T}_k) - \tilde{V}_k \leq \pi \text{ for } k \geq 1,
\]

\[
\tilde{V}_k = V_k \mod 2\pi \text{ so that } -\pi < \tilde{v}_k - \tilde{v}_{k+1}(\tilde{T}_k) - \tilde{V}_k \leq \pi \text{ for } k \leq -1,
\]

\[
\tilde{N}_k = N_{q_k} + (\tilde{T}_k - T_k, \tilde{V}_k - V_k),
\]

\[
L_k = \tilde{T}_{k+1} - \tilde{T}_k.
\]
where we always use the notation $\tilde{q}_k = (\tilde{u}_k, \tilde{v}_k)$. We now require ‘smoothening’ functions $\varphi^-, \varphi^+$, defined over an arbitrary interval $[a, b], a < b$:

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

$$
\varphi^{+}_{a, b}(t) = 1 - \varphi^{-}_{a, b}(t).
$$

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

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

where the implicit constant depends only on $\omega$. Let

$$
q^0(t) = \varphi^-_{\tilde{T}_{k-1}+L, \tilde{T}_{k}-L}(t)\tilde{q}_{k-1}(t) + \varphi^+_{\tilde{T}_{k-1}+L, \tilde{T}_{k}-L}(t)\tilde{q}_k(t)
$$

for all $t \in [\tilde{T}_{k-1}, \tilde{T}_k]$. (3.27)

**Remark 3.2.** Assume we fix a segment $[\omega^-, \omega^+]$ in a region of instability, and that for each $\omega, \tilde{\omega} \in [\omega^-, \omega^+]$ satisfying (3.12) 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 3.10). 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 §4.1.

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

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

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

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

$$
|v(\tilde{T}_k) - \tilde{V}_k| \leq M.
$$

**Lemma 3.13.** We have that $q^0 \in \mathcal{A}$.

**Proof.** The smoothness of $q^0$ and $q_t^0 \in H^2_{\text{ul}}(\mathbb{R})^2$ follow from the construction, Lemma 3.11 and Remark A.1. By definition, $u(\tilde{T}_k) = (2k + 1)\pi$ and $v(\tilde{T}_k) = \tilde{V}_k$, which trivially implies (3.29), (3.30). □

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

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

**Lemma 3.14.** There exist an absolute constant $c_6 \geq 1$ so that any $q^0 = (u^0, v^0)$ given by (3.27) satisfies

$$
|u^0(t) - 2k(t)\pi| \leq c_6e^{-(1/2)\sqrt{\|t\|}},
$$

(3.31)
Proof. We write (3.27) 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

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

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

To obtain (3.31), the left-hand side of (3.33), and (3.34), (3.35), it suffices to differentiate (3.36), (3.37) and insert (3.17)–(3.22) as required and (3.26).

From the left-hand sides of (3.19), (3.20) and the definition of $\tilde{v}_k$, $\tilde{T}_k$, $\tilde{V}_k$, we easily obtain that, for all $t \in [T_{k-1}, T_k]$,

$$|\tilde{v}_{k-1}(t) - \tilde{V}_{k-1} - \tilde{\omega}_k(t - \tilde{T}_{k-1})| \ll \mu,$$

$$|\tilde{v}_k(t) - \tilde{V}_k - \tilde{\omega}_k(t - \tilde{T}_k)| \ll \mu.$$  

By inserting $t = \tilde{T}_k$ in (3.38a), we get

$$|\tilde{\omega}_k - \tilde{\omega}_{k-1}(\tilde{T}_k)| \ll \frac{\mu}{\tilde{T}_k - \tilde{T}_{k-1}}.$$

By definition and (3.28) we know that $|\tilde{V}_k - \tilde{v}_{k-1}(\tilde{T}_k)| \ll 1 \land M$. As also $\mu \leq 1 \land M$, we have

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

Combining it with (3.38), 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 $\mu \leq 1 \land M$, we obtain

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

$$|\tilde{v}_k(t) - \tilde{V}_k - \tilde{\omega}_k(t - \tilde{T}_k)| \ll 1 \land M.$$  

Now (3.32) follows from (3.37) and (3.40). The right-hand side of (3.33) is obtained easily by differentiating (3.37), using the right-hand sides of (3.19), (3.20) and finally (3.39).  

□
3.4. Invariant sets with $L^\infty$ bounds. We now construct $A$-relatively $\xi$-invariant sets with respect to the dynamics (1.2) satisfying \textit{a priori} $L^\infty$ bounds. More specifically, we construct $B_1$ such that any $q = (u, v) \in B_1$ satisfies for all $t \in \mathbb{R}$

$$|u(t) - 2k(t)\pi| \leq c_7 e^{-(1/2)\sqrt{t}} \|t\|, \quad (3.41)$$

$$|v(t) - v^0(t)| \leq c_8 M, \quad (3.42)$$

where $c_7, c_8 > 0$ are absolute constants. We first define the set $B_1$, then show that it is $A$-relatively $\xi$-invariant, that $q^0 \in B_1$, and finally we deduce (3.41) and (3.42).

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

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

and

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

We define $B_1$ to be the set of all $q \in A$ satisfying

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

$$t \in [\tilde{T}_k, \tilde{T}_{k+1}], \quad (3.43)$$

$$v_k^+(t) \leq v(t) \leq v_k^-(t), \quad t \in [\tilde{T}_k, \tilde{T}_{k+1}]. \quad (3.44)$$

**Lemma 3.15.** The set $B_1$ is $A$-relatively $\xi$-invariant.

**Proof.** We apply twice the parabolic maximum principle [16, §7, Theorem 12]. Assume that $q(s_0) \in B_1$, and that for all $s \in [s_0, s_1]$, $q(s) \in A$. We have already shown in Lemma 3.3(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 3.2(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

$$(t, s) \in \{ [\tilde{T}_k, \tilde{T}_{k+1}] \cup [\tilde{T}_k, \tilde{T}_{k+1}] \cup [\tilde{T}_{k+1}, \tilde{T}_{k+1}], s \in [s_0, s_1], \quad (3.45)$$

and thus, by the parabolic maximum principle, (3.43) holds for all $s \in [s_0, s_1]$ and all $k \in \mathbb{Z}$.

Consider now the bounds on $z^-, z^+$. By Lemma 3.2(vi), we have that, for $t \in [\tilde{T}_k, \tilde{T}_{k+1}]$,

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

We obtain analogous bounds on $|z^- (t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon}))|$. We deduce that whenever $u(t)$ satisfies (3.43), then (3.41) holds for some absolute $c_7 = c_2 e$.

We now show that whenever $u(t, s)$ satisfies (3.41), $v^-$ 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 $(1 - \cos x) \leq (x - 2k(t)\pi)^2/2, (3.41)$ with $c_7 = c_2 e$ and the standing assumption (A1), we get that, for $t \in (\tilde{T}_k, \tilde{T}_{k+1})$,

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

$$\leq -c_2^2 e^{2} \varepsilon \mu e^{-\sqrt{\|t\|}} + c_2^2 e^{2} \varepsilon \mu e^{-\sqrt{\|t\|}} \leq 0.$$
Analogously we get that \((v^+_k)_{tt} - V_v(u, v^+_k, t) \geq 0\). Furthermore, by the definition of \(A\), we have that, for all \(k \in \mathbb{Z}\), \(v^+_k(\tilde{T}_k) \leq v_k^-(\tilde{T}_k)\) and \(v^+_k(\tilde{T}_{k+1}) \leq v_k^-(\tilde{T}_{k+1})\). It suffices now to apply the parabolic maximum principle to (1.2b) for all \(k \in \mathbb{Z}\), with the same parabolic boundary (3.45).

**Lemma 3.16.** 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 3.4, we see that, for \(t \in [\tilde{T}_k, \tilde{T}_{k+1}]\),

\[
\begin{align*}
z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2k\pi & \leq \tilde{u}_k(t) \\
& \leq 2(k + 1)\pi \leq z^+(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2(k + 1)\pi,
\end{align*}
\]

\[
\begin{align*}
z^+(t - \tilde{T}_k - 3/(4\sqrt{\varepsilon})) + 2k\pi & \leq 2(k + 1)\pi \\
& \leq \tilde{u}_{k+1}(t) \leq z^+(t - \tilde{T}_{k+1} + 3/(4\sqrt{\varepsilon})) + 2(k + 1)\pi.
\end{align*}
\]

As \(u^0\) is a convex combination of \(\tilde{u}_k(t), \tilde{u}_{k+1}(t)\) on \([\tilde{T}_k, \tilde{T}_{k+1}]\), (3.43) holds for \(u = u^0\).

The relation (3.44) for \(v = v^0\) follows from (3.32) and the definitions of \(v_k^-, v_k^+\) (as \(v_k^-\) is concave and \(v_k^+\) is convex).

**Lemma 3.17.** The relations (3.41), (3.42) hold for all \(q \in \mathcal{B}_1\).

**Proof.** The relation (3.41) has already been established in the proof of Lemma 3.15. As \(v(t), v^0(t)\) satisfy (3.44), we have that

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

To establish a bound on \(|v^+_k(t) - v^-_k(t)|\), we introduce \(w(t) = v^-_k(t) - \tilde{V}_k - (\tilde{V}_{k+1} - \tilde{V}_k)(t - \tilde{T}_k)\). 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 \(T \in [\tilde{T}, \tilde{T}_{k+1}]\). Then

\[
|w(T)| = c_2^2e^2 \varepsilon \mu \int_T^T \exp(-\sqrt{\varepsilon}(\tilde{T}_k - t)) dt \leq 8e^2 \sqrt{\varepsilon} \mu \exp(-\sqrt{\varepsilon}(\tilde{T}_k - T)),
\]

\[
|w(T) - w(\tilde{T})| \leq c_2^2e^2 \varepsilon \mu \int_T^T \exp(-\sqrt{\varepsilon}(\tilde{T}_k - t)) dt \leq c_2^2e^2 \mu.
\]

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

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

which is by (3.28) \(\leq c_8 M\), with \(c_8 = 2c_2^2e^2 + 2\). It suffices to insert this in (3.46). \(\square\)

**3.5. Bounds on the derivatives.** Here we show that there is an \(A\)-relatively \(\xi\)-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

\[
\lambda(\tau) = \frac{\sqrt{\varepsilon}}{4} \wedge \frac{8 \log ||\tau||}{||\tau||},
\]

(3.47)
where $\wedge$ is the minimum. In this subsection we use the weighted $L^2$ norm

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

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

$$\|u_t - u^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_9 \lambda(\tau), \quad (3.48a)$$

$$\|v_t - v^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_9 (M^2 + 1) \lambda(\tau), \quad (3.48b)$$

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

$$\varepsilon \|u_t - u^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|u_{ttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_{10} (M^2 + \sigma^2) \varepsilon \lambda(\tau), \quad (3.49a)$$

$$\varepsilon \|v_t - v^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|v_{ttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_{10} (M^2 + \sigma^2) \varepsilon \lambda(\tau). \quad (3.49b)$$

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

$$\varepsilon \|u_t - u^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|u_{ttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|u_{tttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_{11} (M^4 + \sigma^4) \varepsilon \lambda(\tau), \quad (3.50a)$$

$$\varepsilon \|v_t - v^0_t\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|v_{ttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 + \|v_{tttt}\|_{L^2_{\tau}^2(\mathbb{R})}^2 \leq c_{11} (M^4 + \sigma^4) \varepsilon \lambda(\tau). \quad (3.50b)$$

**Proposition 3.18.** 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 $A$-relatively $\xi$-invariant, and such that $q^0 \in \mathcal{B}_4$.

The proof of the proposition is routine but technical, and as such postponed to 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 C.4) which relies on the $L^\infty$ bounds obtained in the previous subsection. 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 3.19.** 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 $\tilde{h} = (\tilde{a}^k, \tilde{v}^k) \in H^3_{\text{loc}}(\mathbb{R})^2$ 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 $t \in \mathbb{R}$, $|\tilde{a}^h(t)| \leq c_{12} e^{-(1/2)\sqrt{\varepsilon} |t - \tilde{T}_k|}$.

(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 $T \geq 0$,

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

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

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

We will use the following simple lemma.
Lemma 3.20. 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 \( j = 0, 1, \ldots \) that \( y_j \geq e^{(1/2)\sqrt{x+j+1}} \), where \( x \) is chosen so that \( y_0 = e^{(1/2)\sqrt{x}} \), 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

\[
e^{(1/2)\sqrt{x+j+1}} = e^{(1/2)\sqrt{x+j}} + \frac{e^{1/2} e^{(1/2)\sqrt{x+j}}}{\sqrt{x+j}} \leq y_j + \frac{y_j}{\log y_j} = y_{j+1}.
\]

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

\[
\sum_{j=0}^{\infty} \frac{\log y_j}{y_j} \leq \frac{\log y_0}{y_0} + \int_{y_0}^{\infty} \frac{1}{2} e^{-2y \sqrt{x}} \, dz \ll \frac{\log y_0}{y_0} + xe^{-1/(2)\sqrt{x}} = \frac{\log y_0}{y_0}.
\]

Proof of Lemma 3.19. Let

\[
\tilde{h}(t) = \begin{cases} \varphi_{\tilde{T}_k - L(1 + 1/ \log L)}^{+} \cdot (q(t) - \tilde{q}_k(t)), & t \leq \tilde{T}_k, \\ \varphi_{\tilde{T}_k + L, \tilde{T}_k + L(1 + 1/ \log L)}^{-} \cdot (q(t) - \tilde{q}_k(t)), & t \geq \tilde{T}_k. \end{cases}
\]

Now, (i) and (ii) follow from the definition of the smoothening functions (3.25), and (iii) from (3.17), the definition of \( \tilde{q}_k \) and (3.41). We claim that, for each \( T \geq 4/\sqrt{\epsilon} \),

\[
\|\tilde{h}_t\|^2_{H^2([\tilde{T}_k - T, \tilde{T}_k - T + (1 + 1/ \log L)])} \ll (M^4 + \sigma^4) \frac{\log T}{T}. \tag{3.51}
\]

For \( T \leq L \), this follows directly from (3.50) 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 (3.19)–(3.22), and finally by using (3.26). To show (iv), we use (i), and the fact that, for all \( t \in [\tilde{T}_k - 2\pi, \tilde{T}_k + 2\pi] \), \( \tilde{q}_0(t) = \tilde{q}_k(t) \); thus by (3.42)

\[
|\tilde{v}(t)| = |v(t) - \tilde{v}(t)| = |v(t) - v^0(t)| \leq c_8 M.
\]

As, by (ii), \( h(t) \) vanishes for \( t \geq L(1 + 1/ \log L) \), the claim follows for \( T \geq L \). For \( T \in [L, L(1 + 1/ \log L)] \), (3.51) similarly follows from the case \( T = L \). Analogously we obtain for such \( T \),

\[
\|\tilde{h}_t\|^2_{H^2([\tilde{T}_k - T - T(1 + 1/ \log T), \tilde{T}_k - T])} \ll (M^4 + \sigma^4) \frac{\log T}{T}, \tag{3.52}
\]

and

\[
\|\tilde{h}_t\|^2_{H^2([-4/\sqrt{\epsilon}, 0])} \ll (M^4 + \sigma^4) \sqrt{\epsilon}, \quad \|\tilde{h}_t\|^2_{H^2([0, 4/\sqrt{\epsilon}])} \ll (M^4 + \sigma^4) \sqrt{\epsilon}. \tag{3.53}
\]
3.6. Lower bound on the action dissipation. In this subsection 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

$$M = 2\pi + 2(\sigma + 1)(R + \mu) + 6R^{1/2}e^{1/4}$$  \hspace{1cm} (3.54)

(clearly $M$ satisfies (3.28) as required). Let $C$ be the closure in $H^2_{\text{loc}}(\mathbb{R})^2$ of the set of all $h = (u^h, v^h) \in H^3_{\text{loc}}(\mathbb{R})^2$ satisfying, for all $t \in \mathbb{R}$ and all $T \geq 0$,

$$|u^h(t)| \leq c_{12} e^{-(1/2)\sqrt{T}|t|+2\pi},$$  \hspace{1cm} (3.55)

$$|v^h(0)| \leq c_8 M,$$  \hspace{1cm} (3.56)

$$\|h_t\|^2_{H^2((\infty, -T])^2} + \|h_t\|^2_{H^2((T, \infty))^2} \leq 2c_{12}(M^4 + \sigma^4)\left(\log T \land \frac{\sqrt{\varepsilon}|\log \varepsilon|}{8}\right).$$  \hspace{1cm} (3.57)

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

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

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

where we take $\omega^-(q) = \lim_{t \rightarrow -\infty} v_t$ and $\omega^+(q) = \lim_{t \rightarrow \infty} v_t$, 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 3.21.** There exists a constant $\Delta_1 > 0$, $0 \leq \Delta_1 \leq \Delta_0/2$, depending on the region of instability $[\omega^-, \omega^+]$, $R$, $\Delta_0$ and $f$, and constants $0 \leq \Delta_0(q) \leq \Delta_0/2$ defined for all $q \in \mathcal{H}$, so that, for all $(q, h) \in \mathcal{H} \times C$, if

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

then

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

To obtain $\Delta_1$ we introduce, for any $q \in \mathcal{H}$,

$$\Delta_1(q, e) = \inf \{D_q(h), \ h \in C, \ E_q(h) = E_q(0) + e\},$$  \hspace{1cm} (3.58)

$$\Delta_1(q) = \sup_{e \in [0, \Delta_0]} \Delta_1(q, e).$$  \hspace{1cm} (3.59)
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 \( E : \mathcal{Y} \to \mathbb{R} \) 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 \( \text{Ker} \ D^2 E(h), D^2 E(h) : \mathcal{Y} \to \mathcal{Y}^* \) is finite-dimensional for any \( h \in \mathcal{Y} \). (Equivalently, \( D^2 E \) is Fredholm, as in this case it suffices to check finite dimensionality of the kernel \( [34] \).) 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 3.22.** (Morse–Sard–Pohožaev Lemma \([34]\)) 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 \( \dim(\text{Ker} \ D^2 E(h)) \leq m < \infty \) 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 3.22 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

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

The space \( \mathcal{Y} \) is a Hilbert space, as the norm (3.60) 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 3.23.**

(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 3.12. 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 3.11, the definition of \( \mathcal{C} \) and the assumed localized topologies on the sets \( \mathcal{H} \) and \( \mathcal{C} \). \( \square \)

**Lemma 3.24.** For any \( q \in \mathcal{H} \), we have the following.

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

(ii) \( E_q \) is Fredholm and the dimension of \( \text{Ker} \ D^2 E_q \) 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 given by

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

(3.61)

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

(3.62)

$$
D^k E_q(h)(g^{(1)}, \ldots, g^{(k)}) = \int_{-\infty}^{\infty} \{D^k_{u,v} V(q(t) + h(t), t) (g^{(1)}(t), \ldots, g^{(k)}(t))\} dt,
$$

(3.63)

where $k = 3, 4$. We first show that the integrals on the right-hand sides are finite. Consider the terms containing $u_t, u_i^h, v_t, v_i^h$ in (3.61). They are absolutely integrable by the Cauchy–Schwarz inequality, as $\|u_i^h\|_{L^2(\mathbb{R})} < \infty, \|v_i^h\|_{L^2(\mathbb{R})} < \infty$ by the definition of $\mathcal{Y}$, and by Lemma 3.11 applied to $u_t, |v_t - \omega|, |v_t - \tilde{\omega}|$. Analogously the term $g_t^{(1)} g_t^{(2)}$ in (3.62) 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

$$
X := \int_{-\infty}^{\infty} \{\partial_i u \partial_i v V(q(t) + h(t), t) u^{(1)}(t) \ldots u^{(i)}(t) v^{(i+1)}(t) \ldots v^{(k)}(t)\} dt
$$

(3.64)

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 $V,$

$$
X \ll \int_{-\infty}^{\infty} |u^{(1)}(t)|(1 + |t|^{1/2})^{k-1} dt
$$

$$
\ll \|g^{(2)}\|_{\mathcal{Y}} \ldots \|g^{(k)}\|_{\mathcal{Y}} \left(\int_{-\infty}^{\infty} e^{-(1/4)\sqrt{|t|}}(1 + |t|^{k-1}) dt\right)^{1/2}
$$

$$
\times \left(\int_{-\infty}^{\infty} e^{(1/4)\sqrt{|t|}}|u^{(1)}(t)|^2 dt\right)^{1/2}
$$

$$
\ll \|g^{(1)}\|_{\mathcal{Y}} \|g^{(2)}\|_{\mathcal{Y}} \ldots \|g^{(k)}\|_{\mathcal{Y}}.
$$

(3.65)
Analogously, in the case $i = 0$,

\[
X \ll \|g^{(1)}\|_Y \cdot \|g^{(k)}\|_Y \int_{-\infty}^{\infty} |\partial_t^k V(q(t) + h(t), t)| (1 + |t|^{1/2})^k \, dt
\]

\[
\ll_f \|g^{(1)}\|_Y \cdot \|g^{(k)}\|_Y \int_{-\infty}^{\infty} (1 - \cos(u(t) + u^h(t))) (1 + |t|^{1/2})^k \, dt
\]

\[
\ll_f \|g^{(1)}\|_Y \cdot \|g^{(k)}\|_Y \int_{-\infty}^{\infty} e^{-\sqrt{t(t-t_0)}} (t^2) (1 + |t|^{1/2})^k \, dt,
\]

where in the last row we applied

\[
1 - \cos(u(t) + u^h(t)) \ll (u(t) - 2\pi 1_{(0, \infty)}(t))^2 + u^h(t)^2
\]

and then (3.17). Analogously to (3.65) we establish bound on the remaining terms containing $u^h$ and get

\[
X \ll_f, e \|g^{(1)}\|_Y \cdot \|g^{(k)}\|_Y (1 + \|h\|^2_Y).
\]  

(3.66)

We now show that (3.61) 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 $D^4 E_q$). By the mean value theorem and analogously as when evaluating (3.65), (3.66), we obtain

\[
|E_q(h + g^{(1)}) - E_q(h) - D E_q(h) g^{(1)}| \\
\leq \int_{-\infty}^{\infty} \left\{ \frac{1}{2} (g_t^{(1)})^2 + |D_{u,v}^2 V(q(t) + h(t), t)||g^{(1)}|^2 \right\} \, dt
\]

\[
\ll_f, e (1 + \|h\|^2_Y) \|g^{(1)}\|^2_Y,
\]

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

To show (ii), we observe that the kernel of $D^2 E_q(h)$ is by partial integration the set of all $g^{(1)} \in \mathcal{Y}$ such that $-g^{(1)}_t + D_{u,v}^2 V(q(t) + h(t), t)g^{(1)}(t) = 0$ 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 (3.61) as $DE_q(h)g^{(1)} = \int_{-\infty}^{\infty} -(q + h)_s g^{(1)} \, dt$, which is clearly always 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 3.25.** 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 3.24(i) and (ii), we can apply the Morse–Sard–Pohožaev lemma 3.22 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 3.24(iii), for any $h \in \mathcal{Y}$, $D_q(h) > 0$.

By Lemma 3.23(i) and (ii), the level set $\{E_q(h) = E_q(0) + e\} \cap C$ 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.  

\[\square\]
Lemma 3.26. 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 \(\{E_q(h) = E_q(0) + e\} \cap C\) are compact, we can find \(h_n \in C, n \in \mathbb{N}\) 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 \(C\). Let \(h = \lim_{n \to \infty} h_{k_n}\). Now because of continuity of \((q, h) \mapsto D_q(h)\) by Lemma 3.23(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 3.23(ii), \(E_q(h) = E_q(0) + e\), and thus by definition \(D_q(h) \geq \Delta_1(h, e)\), which completes the proof.

Proof of Proposition 3.21. We prove the proposition 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 3.25, \(\Delta_1(q) > 0\) 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 \(U_i \times B(e_i, r_i), i \in I\), 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 \(U_i, i \in I\) is an open cover of \(\mathcal{H}\). By compactness, we find its finite subcover. Denote it by \(U_1, \ldots, U_n\), and its associated open balls by \(B(e_1, r_1), \ldots, B(e_n, r_n)\).

We now set \(\Delta_1 = \min\{\tilde{\Delta}_1/2, r_1, \ldots, r_n\}\), 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.

3.7. Local upper bounds on the action. This subsection contains the core of the argument, as we show that the action within two ‘intersections’, or more precisely in the segment \([\tilde{T}_k - L, \tilde{T}_k + L]\), cannot 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

\[
\tilde{E}_k(q) = \int_{\tilde{T}_k - L}^{\tilde{T}_k} L_{o_k}(q, q_t, t) \, dt + \int_{\tilde{T}_k}^{\tilde{T}_k + L} L_{o_{k+1}}(q, q_t, t) \, dt + (\omega_{k+1} - \omega_k)v(\tilde{T}_k),
\]

\[
\tilde{D}_k(q) = \int_{\tilde{T}_k - L}^{\tilde{T}_k + L} \{u^2 + v^2\} \, dt,
\]

\[
\tilde{F}_k(q) = u_t(\tilde{T}_k + L)u_s(\tilde{T}_k + L) + (v_t(\tilde{T}_k + L) - \omega_{k+1})v_s(\tilde{T}_k + L) - u_t(\tilde{T}_k - L)u_s(\tilde{T}_k - L) - (v_t(\tilde{T}_k - L) - \omega_k)v_s(\tilde{T}_k - L).
\]
Let $B_5$ be the set of all $q \in B_4$ so that, for all $k \in \mathbb{Z}$,
\[
\tilde{E}_k(q) \leq \tilde{E}_k(q_k) + \Delta_0(q_k),
\] (3.67)
where $\Delta_0(q_k), \Delta_1$ are as constructed in Proposition 3.21.

**Proposition 3.27.** There exists an absolute constant $c_{13} > 0$ such that, if
\[
L \geq c_{13}(\sigma^5 + M^5) \frac{\|\log \Delta_1\|}{\Delta_1}
\] (3.68)
then $B_5$ is $A$-relatively $\xi$-invariant. Furthermore, $q^0 \in B_5$.

(Recall that $M$ in (3.68) is given by (3.54).) To prove it, we first establish the action-balance law (3.69), then find upper bounds on the action flux, lower bounds on the action dissipation on the energy level $\tilde{E}_k(q) = \tilde{E}_k(q_k) + \Delta_0(q_k)$, and then complete the proof. The constant $c_{13}$ may change throughout the subsection.

**Lemma 3.28.** For any $q \in \mathcal{X}$,
\[
\frac{d}{ds} \tilde{E}_k(q) = -\tilde{D}_k(q) + \tilde{F}_k(q).
\] (3.69)

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

**Lemma 3.29.** There exists an absolute constant $c_{13} > 0$ such that if (3.68) holds, then for any $q \in B_4$ and any $k \in \mathbb{Z}$, $|\tilde{F}_k(q)| \leq \Delta_1/4$.

*Proof.* By (3.33), (4.49a), the definition of $\lambda(\tau)$, using that $\|\tau\| \leq L_k(\tau)$ and $\varepsilon \leq 1$, we obtain
\[
|u_t(\tau)| \ll \|u_t\|_{H^1([\tau, \tau+1])} \ll (M + \sigma)\lambda(\tau)^{1/2} + \|u_t^0\|_{L^2([\tau, \tau+1])}
\]
\[
\ll (M + \sigma)\left(\frac{\log \|\tau\|}{\|\tau\|}\right)^{1/2},
\] (3.70)
and analogously
\[
|v_t(\tau) - \omega_{k(\tau)}| \ll (M + \sigma)\left(\frac{\log \|\tau\|}{\|\tau\|}\right)^{1/2}.
\] (3.71)
Because of (C.12) and (C.13) in Appendix C, we get, for all $q \in B_4 \subset B_1$, $|V_t| \ll e^{-\sqrt{\|\tau\|}/2} \ll (\log \|\tau\|/\|\tau\|)^{1/2}$, $|V_t| \ll e^{-\sqrt{\|\tau\|}/2} \ll (\log \|\tau\|/\|\tau\|)^{1/2}$. Using this and (3.50a), (3.50b) applied to $|u_t|$, $|v_t|$ analogously as above, we get
\[
|u_s| \leq |u_t| + |V_t| \ll (M^2 + \sigma^2)\left(\frac{\log \|\tau\|}{\|\tau\|}\right)^{1/2},
\] (3.72)
\[
|v_s| \leq |v_t| + |V_t| \ll (M^2 + \sigma^2)\left(\frac{\log \|\tau\|}{\|\tau\|}\right)^{1/2}.
\] (3.73)
As by definition, for $\tau = \tilde{T}_k \pm L$, $\|\tau\| = L$, we deduce that, for any $q \in B_4$,
\[
\tilde{F}_k(q) \leq c_{13}(M^3 + \sigma^3)\frac{\log L}{L}
\] (3.74)
for some absolute constant $c_{13} > 0$. Now it is straightforward to check that
\[
L \gg (\sigma^5 + M^5) \frac{|\log \Delta_1|}{\Delta_1} \gg (\sigma^3 + M^3) \frac{|\log \Delta_1|}{\Delta_1}
\]
suffices for the right-hand side of (3.74) to be $\ll \Delta_1$, which completes the proof for a large enough absolute constant $c_{13}$.

**Lemma 3.30.** There exists an absolute constant $c_{13} > 0$ such that if (3.68) 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
\[
|\hat{E}_k(q) - E_{q_k}(h) - T_k(\omega_k^2 - \omega_{k+1}^2)/2 - (\omega_{k+1} - \omega_k)(\hat{V}_k - V_k)| \leq \Delta_1/2, \quad (3.75)
\]
\[
|\hat{D}_k(q) - D_{q_k}(h)| \leq \Delta_1/2. \quad (3.76)
\]
Furthermore, for all $k \in \mathbb{Z}$,
\[
|\hat{E}_k(\tilde{q}_k) - E_{q_k}(0) - T_k(\omega_k^2 - \omega_{k+1}^2)/2 - (\omega_{k+1} - \omega_k)(\hat{V}_k - V_k)| \leq \Delta_1/2. \quad (3.77)
\]

**Proof.** Fix $k \in \mathbb{Z}$. We define $\tilde{h}$ as in Lemma 3.19, and let $h(t) = \tilde{h}(t + \tilde{T}_k - T_k)$. We first show that $h \in \mathcal{C}$. As by definition $|h| \leq 2\pi$, by inserting the definition of $h$ in the bounds in Lemma 3.19, we easily see that (3.55) and (3.56) follow from Lemma 3.19(iii) and (iv). Analogously we show that (3.57) follows from Lemma 3.19(vi).

Denote by $X_k = T_k(\omega_k^2 - \omega_{k+1}^2)/2 + (\omega_{k+1} - \omega_k)(\hat{V}_k - V_k)$ (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, we obtain
\[
E_{q_k}(h) = \int_0^\infty L_{ow_k}(q_k + h, (q_k)_t + h_t, t) \, dt
\]
\[
+ \int_0^\infty L_{ow_{k+1}}(q_k + h, (q_k)_t + h_t, t) \, dt + (\omega_{k+1} - \omega_k)(v_k(0) + v^h(0))
\]
\[
= \int_{-\tilde{T}_k}^{\tilde{T}_k} L_{ow_k}(q_k + h, (q_k)_t + h_t, t) \, dt + \int_{\tilde{T}_k}^{\tilde{T}_k} L_{ow_{k+1}}(q_k + h, (q_k)_t + h_t, t) \, dt
\]
\[
+ (\omega_{k+1} - \omega_k)(V_k + v^h(T_k)) + (\omega_k^2 - \omega_{k+1}^2)T_k/2
\]
\[
= \int_{-\tilde{T}_k}^{\tilde{T}_k} L_{ow}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t) \, dt
\]
\[
+ \int_{\tilde{T}_k}^{\tilde{T}_k-L} L_{ow_{k+1}}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t) \, dt + (\omega_{k+1} - \omega_k)v(\tilde{T}_k) + X_k
\]
\[
= \hat{E}_k(q) + \int_{-\tilde{T}_k-L}^{\tilde{T}_k-L} L_{ow}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t) \, dt
\]
\[
+ \int_{\tilde{T}_k+L}^{\tilde{T}_k+L} L_{ow_{k+1}}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t) \, dt + X_k, \quad (3.78)
\]

In the last line, we used Lemma 3.19(i), where we showed that for all $t \in [\tilde{T}_k - L, \tilde{T}_k + L]$, where we used that by construction, $\tilde{h}(t) = q(t) - \tilde{q}_k(t)$.  


By the definition of $V$, $\tilde{q}_k$, and then (3.17) and Lemma 3.19(iii), we obtain
\[
\int_{-\infty}^{\tilde{T}_k-L} V(\tilde{q}_k + \tilde{h}, t) \, dt \leq \varepsilon \int_{-\infty}^{\tilde{T}_k-L} (1 - \cos(\tilde{q}_k + \tilde{h})) \|t\| \, dt
\]
\[
\ll \varepsilon \int_{-\infty}^{\tilde{T}_k-L} q_k^2(t) \, dt + \varepsilon \int_{-\infty}^{\tilde{T}_k-L} h_k^2(t) \, dt
\]
\[
\ll \varepsilon \int_{-\infty}^{\tilde{T}_k-L} e^{-(1/2)\sqrt{\varepsilon}|t-T_k|} \, dt + \varepsilon \int_{-\infty}^{\tilde{T}_k-L} e^{-(1/2)\sqrt{\varepsilon}|t-\tilde{T}_k|} \, dt
\]
\[
\ll \sqrt{\varepsilon} e^{-(1/2)\sqrt{\varepsilon}L} \ll \frac{\log L}{L}.
\]

Inserting this in the definition of $L_{\omega k}$ and combining with (3.18), (3.19) and Lemma 3.19(vi), we get
\[
\int_{-\infty}^{\tilde{T}_k-L} L_{\omega k}(\tilde{q}_k + \tilde{h}, (\tilde{q}_k)_t + \tilde{h}_t, t) \, dt
\]
\[
\ll \|((\tilde{u}_k)_t)\|_{L^2((-\infty, \tilde{T}_k-L))}^2 + \|((\tilde{u}_k)_t - \omega_k)\|_{L^2((-\infty, \tilde{T}_k-L))}^2
\]
\[
+ \|\tilde{h}_t\|_{L^2([\tilde{T}_k-L(1+1/L), \tilde{T}_k-L])}^2 + \frac{\log L}{L}
\]
\[
\ll (M^4 + \sigma^4) \frac{\log L}{L}.
\]

By proving an analogous statement for the second integral in (3.78) and combining all the relations above, we see that
\[
|\tilde{E}_k(q) - E_{q_k}(h) - X_k| \ll (M^4 + \sigma^4) \frac{\log L}{L}.
\]

We complete the claim analogously as in Lemma 3.29. The proof for (3.76) is analogous, as we also, by Lemmas 3.11 and 3.19, control the second derivatives of $\tilde{q}_k$ and $\tilde{h}$. We get (3.77) by inserting $h = 0$ in (3.75).

\textbf{Proof of Proposition 3.27.} Let $c_{13}$ be the larger of the constants in Lemmas 3.28 and 3.30. Let for some $k \in \mathbb{Z}$, $s_1$ be the supremum of all the times $s$ such that (3.67) holds. Then, by continuity,
\[
\tilde{E}_k(q) = \tilde{E}_k(\tilde{q}_k) + \Delta_0(q_k).
\]
Combining (3.75), (3.77) and (3.79), we obtain $|E_{q_k}(h) - E_{q_k}(0) - \Delta_0(q_k)| \leq \Delta_1$. By Proposition 3.21, we thus obtain $D_{q_k}(h) \geq \Delta_1$, so by (3.76), $\tilde{D}_k(q) \geq \Delta_1/2$. Inserting this and the bound from Lemma 3.29 in the action balance law (3.69), we get
\[
\frac{d}{ds} \tilde{E}_k(q) \leq -\Delta_1/2 + \Delta_1/4 \leq -\Delta_1/4 < 0,
\]
which is in contradiction with the assumption. To show $q^0 \in \mathcal{B}_S$, 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 $q^0 \in \mathcal{B}_S$ by definition.

\qed
3.8. **Completion of the construction of an invariant set.** In this subsection we complete the construction of an invariant set \( B \) with respect to (1.2) by finally applying Lemma 1.4.

**Proposition 3.31.** Assume \( \omega_k \) is a sequence in a closed subset \([\omega^-, \omega^+]\) satisfying (3.12), and that \( L \) satisfies

\[
L \geq c_{14} \sigma^5 \frac{|\log \Delta_1|}{\Delta_1}
\]

for some large enough absolute constant \( c_{14} \). Then there exists a \( \xi \)-invariant set \( B \subset B_5 \subset \mathcal{X} \), such that \( q^0 \in B \).

**Remark 3.3.** Assume that as in the Remark 3.2 we fix a segment \([\omega^-, \omega^+]\) in a region of instability, and that for each \( \omega, \tilde{\omega} \in [\omega^-, \omega^+] \) satisfying (3.12) 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 \( B \) 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 \( B_1 - B_5 \). In the proofs of the main theorems, we thus use the notation \( B(L, (\tilde{L}_k)_{k \in \mathbb{Z}}, (\omega_k)_{k \in \mathbb{Z}}) \).

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 \( q \in B_5 \) cannot intersect \( \partial \mathcal{N}_k \) at \( k \), then define \( B_6 \) and prove its \( \xi \)-relative invariance, and finally complete the proof of Proposition 3.31.

**Lemma 3.32.** There exists an absolute constant \( c_{15} \) such that, if

\[
L \geq c_{15} \frac{|\log \mu|}{\sqrt{\varepsilon}},
\]

then the following holds: for each \( q \in B_5 \) intersecting \( \partial \mathcal{N}_k \) at \( k \), we have

\[
\tilde{E}_k(q) \geq \tilde{E}_k(\tilde{q}_k) + \Delta_0.
\]

**Proof.** To prove (3.82), 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 (3.23). Indeed, let \( \tilde{q}(t) = q(t - T_k + \tilde{T}_k) - (2k\pi, \tilde{V}_k - V_k) \) 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

\[
\bar{u}(t) = a \cdot e^{-(1/2)\sqrt{\varepsilon}|t-T_k|}, \quad t \leq T_k - L, \quad \bar{u}(t) = 2\pi - b \cdot e^{-(1/2)\sqrt{\varepsilon}|t-T_k|}, \quad t \geq T_k + L,
\]

where the constants \( a, b \) are uniquely chosen so that \( \bar{u} \) is continuous. By construction, (3.41) and the definition of \( V \), we have

\[
|\tilde{E}_k(q) - \mathcal{L}_{\omega_k, \omega_{k+1}}(\tilde{q})| = \int_{-\infty}^{T_k - L} L_{\omega_k}(\tilde{q}, \tilde{q}_t, t) \, dt + \int_{T_k + L}^{\infty} L_{\omega_{k+1}}(\tilde{q}, \tilde{q}_t, t) \, dt \lesssim \varepsilon \int_{-\infty}^{T_k - L} e^{-(1/2)\sqrt{\varepsilon}|t-T_k|} \, dt + \varepsilon \int_{T_k + L}^{\infty} e^{-(1/2)\sqrt{\varepsilon}|t-T_k|} \, dt \lesssim \sqrt{\varepsilon} e^{-(1/2)\sqrt{\varepsilon}L} \leq e^{-(1/2)\sqrt{\varepsilon}L}.
\]

Similarly, we obtain

\[
|\tilde{E}_k(\tilde{q}_k) - \mathcal{L}_{\omega_k, \omega_{k+1}}(q_k)| \lesssim e^{-(1/2)\sqrt{\varepsilon}L}.
\]
By (3.2) and the definition of $\Delta_0$, 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 of (3.83) and (3.84) are bounded by $\Delta_0/2$. Combining it with (3.23), we complete the proof.

\textbf{Lemma 3.33.} If $q \in B_5$, then, for all $k \in \mathbb{Z}$, $q$ cannot intersect $\partial N_k$ at $k$.

\textbf{Proof.} As $q \in B_5$, (3.67) holds. If $q$ would intersect $\partial N_k$ at $k$, this would contradict (3.82) and $\Delta_0(q_k) \leq \Delta_0/2$ established in Proposition 3.21.

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

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

By the definition of $A$, $u(\tilde{T}_{k-1}) < (2k + 1)\pi$ and $u(\tilde{T}_{k+1}) > (2k + 1)\pi$. Now by Lemma (3.33), for any $t \in \mathcal{Y}_k$, $u(t) \neq (2k + 1)\pi$. Let $\sim_k$ be a relation of equivalence on $\mathcal{Y}_k$ defined with $t_1 \sim_k t_2$ whenever, for all $t_3 \in \mathcal{Y}_k$ 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 $\hat{T}_k = \mathcal{Y}_k/\sim_k$ with the induced topology. As by assumptions, $\hat{T}_k$ is a closed subset of a compact set, $\hat{T}_k$ is compact. By definition, continuity of $q$ and compactness of $\partial N_k$, we see that $\hat{T}_k$ is totally disconnected and compact, thus finite, and $|\hat{T}_k| \geq 2$. Consider $|\hat{T}_k| - 1$ intervals $(t_j, t_{j+1}) \in \mathcal{Y}_k$, where $t_j, t_{j+1} \in \mathcal{Y}_k$ and $t_j \neq t_{j+1}$. We say that $q$ intersects $N_k$ at $k$ exactly $m$-number of times, if $m$ is the number of such intervals, for which $q$ intersects $N_k$ at $k$ for some $(t, v) \in N_k$ such that $t \in (t_j, t_{j+1})$.

Let $B_6$ be the set of all $q \in B_5$, such that for each $k \in \mathbb{Z}$, $q$ intersects $N_k$ at $k$ odd number of times (i.e. that $m$ in the definition of the number of intersections is odd).

\textbf{Lemma 3.34.} The set $B_6$ is $A$-relatively $\xi$-invariant, and $q^0 \in B_6$.

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

\textbf{Proof of Proposition 3.31.} Assume that $L$ satisfies (3.68) and (3.81), and that $M$ is given as in (3.54). Let $\tilde{B} = B_6$ and $B = \tilde{B} \cap A$. It suffices to show that $q^0 \in B$, and that the conditions (B1), (B2) hold, as the claim will then hold by Lemma 1.4. We have already shown (B1) and $q^0 \in B$ in Lemma 3.34. We now show (B2). The smoothness requirement in the definition of $A$ follows from Theorem 2.1(iv). We now show that (3.29) and (3.30) hold.

Let $q \in B_6$, fix $k \in B_6$, and find $(t, v(t)) \in \tilde{N}_k$ such that $u(t) = (2k + 1)\pi$ (this exists by the definition of $B_6$). By definition of $\tilde{N}_k$, $\tilde{T}_k$ and $\tilde{V}_k$ we have that $|t - \tilde{T}_k| \leq R$, $|v(t) - \tilde{V}_k| \leq R$. Without loss of generality, let $t \leq \tilde{T}_k$, $v(t) \leq \tilde{V}_k$ (the other cases are analogous). Then by the definition of $\Delta_0$ and (3.2),

$$\int_t^{\tilde{T}_k} \left(u^2 + (v_t - \omega_k)^2\right) dt \leq \max_{(t_0, v_0) \in \mathbb{R}^2} S_{\omega}(t_0, v_0) \leq 8\sqrt{\varepsilon}(1 + \mu) \leq 9\sqrt{\varepsilon}.$$
We thus have
\[
|u(\tilde{T}_k) - (2k + 1)\pi| \leq \int_t^{\tilde{T}_k} |u_t| \ dt \leq R^{1/2} \left( \int_t^{\tilde{T}_k} u_t^2 \ dt \right)^{1/2} \leq 3R^{1/2} \varepsilon^{1/4},
\]
\[
|v(\tilde{T}_k) - \tilde{V}_k| \leq R + \int_t^{\tilde{T}_k} |v_t| \ dt \leq R (1 + |\sigma|) + R^{1/2} \left( \int_t^{\tilde{T}_k} (v_t - \omega_k)^2 \ dt \right)^{1/2}
\]
\[
\leq R (1 + |\sigma|) + 3R^{1/2} \varepsilon^{1/4}.
\]

By the assumptions (1.5) and (3.54), we thus have \(|u(\tilde{T}_k) - (2k + 1)\pi| \leq 1/4\), \(|v(\tilde{T}_k) - \tilde{V}_k| \leq M/2\). Let \(q(s_0) = q\), and consider the solution \(q(s)\) of (1.2) with the initial condition \(q(s_0) = q\) at \(s = s_0\). By (3.73) and the definition of \(M\) in (3.54), we have \(|u_s| \ll \epsilon, \sigma\), \(|v_t| \ll \epsilon, \sigma\), thus the upper bound on \(|u_s(t)|, |v_t(t)|\) is independent of the choice of \(t = \tilde{T}_k\) and \(q = (u, v) \in B_0\). We conclude that, for any \(\tilde{\lambda} > 0\), there exists \(\lambda > 0\), independent of the choice of \(q(s_0) \in B_0\), such that, for all \(s \in [s_0, s_0 + \lambda]\), we have \(|u(\tilde{T}_k) - (2k + 1)\pi| \leq 1/4 + \tilde{\lambda}\), \(|v(\tilde{T}_k) - \tilde{V}_k| \leq M/2 + \tilde{\lambda}\). By the definition of \(\mathcal{A}\), this proves (B2). We conclude that \(\mathcal{B} := B_0 \cap \mathcal{A}\) is indeed \(\xi\)-invariant.

It remains to show that (3.80) suffices for (3.68) and (3.81) to hold. By definition of \(M\) in (3.54), \(c(\sigma) \ll \sigma^5\), thus (3.68) is satisfied. As by definition of \(\Delta_1\) in Proposition 3.21, (3.24) and (A2), we know that \(\Delta_1 \leq \Delta_0/2 \leq 4\sqrt{\varepsilon} \mu \leq \sqrt{\varepsilon}\), we deduce \(|\log \mu|/\sqrt{\varepsilon} \ll |\log \sqrt{\varepsilon}|/\sqrt{\varepsilon} \leq |\log \Delta_1|/\Delta_1\), which was required. \(\Box\)

4. Proofs of the main theorems

4.1. 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 \(T_{\tau}\), 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 subsection.

**Lemma 4.1.** Assume \(q \in \mathcal{B}\). Then, for all \(t \in \mathbb{R}\) and some absolute \(c_0 > 0\),
\[
|u_t(t)| \leq c_0 \sigma \left( \sqrt{\varepsilon} \wedge \left( \frac{|\log \|t\|}{\|t\|} \right) \right)^{1/2}, \quad |v_t(t) - \omega_k(t)| \leq c_0 \sigma \left( \sqrt{\varepsilon} \wedge \left( \frac{|\log \|t\|}{\|t\|} \right) \right)^{1/2}.
\]

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
\[
\|w\|^2_{L^2([\tau, \tau + 4/\sqrt{\varepsilon}])} \leq A \sqrt{\varepsilon}, \quad \|w_t\|^2_{L^2([\tau, \tau + 4/\sqrt{\varepsilon}])} \leq A \varepsilon^{3/2}.
\]

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 \(\tilde{w}(t) = w(t + 4t/\sqrt{\varepsilon})\), we obtain by direct calculation \(\|\tilde{w}\|_{H^1([0,1])} \ll A^{1/2} \sqrt{\varepsilon}\), which implies \(\|w\|_{L^\infty([0,1])} \ll A^{1/2} \sqrt{\varepsilon}\), thus the claim.

By (3.33), for any \(\tau \in \mathbb{R}\) we have that \(\|u^0_t\|^2_{L^2([\tau, \tau + 4/\sqrt{\varepsilon}])} \ll \sqrt{\varepsilon}\) and \(\|v^0_t - \omega_k(t)\|^2_{L^2([\tau, \tau + 4/\sqrt{\varepsilon}])} \ll \sqrt{\varepsilon}\). Inserting that in (3.49a) and (3.49b), and in the view of the definition (3.54) of \(M\) and that \(\lambda(\tau) \leq \sqrt{\varepsilon}/4\), we get that for any \(\tau \in \mathbb{R}\) the relations (4.2)
hold with \( w = u_t \), and also with \( w = v_t - \omega k(\cdot) \), with \( A = c_0 \sigma^2 \), for some absolute \( c_0 \).

This yields the \( c_0 \sigma \sqrt{\epsilon} \) upper bound in (4.1). The other bound has already been established (by an analogous argument) in (3.70) and (3.71).

**Theorem 4.2.** Let \([\omega^-, \omega^+]\) be such that (S1) holds. Let \( \omega_k, k \in \mathbb{Z} \) 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 \( k \in \mathbb{Z} \)

\[
|u(t_k) - 2k\pi| < \delta_k, \quad |u_t| < \delta_k, \quad |v_t(t_k) - \omega k| < \delta_k. \tag{4.3}
\]

**Proof.** Without loss of generality, assume that all the pairs \( \omega_k, \omega_{k+1} \) satisfy (3.12) (otherwise we complement the sequence). We set \( L \) large enough so that (3.80) holds. Choose \( \tilde{L}_k \) so that \( \tilde{L}_{k+1} - \tilde{L}_k \geq 4L \lor c_{16} \sigma^2 |\log \delta_k|^2/\delta_k^2 + 2\pi \) for some \( c_{16} > 0 \) to be determined later, and then \( L_k \geq 4L \lor c_{16} \sigma^2 |\log \delta_k|^2/\delta_k^2 \). Let \( q^0(L, (\tilde{L}_k)_{k \in \mathbb{Z}}, (\omega_k)_{k \in \mathbb{Z}}) \) as in Remark 3.2, and let \( \mathcal{B} = \mathcal{B}(L, (\tilde{L}_k)_{k \in \mathbb{Z}}, (\omega_k)_{k \in \mathbb{Z}}) \) as in Remark 3.3. Then by Proposition 3.31, \( \mathcal{B} \) is \( \xi \)-invariant, non-empty, and by Theorem 2.2, there is a \( \mathcal{B} \) in \( \mathcal{X} \) 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 (3.41) (the bound on \( u \)) and (4.1) (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 (4.3) holds. \( \square \)

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 \( c_{17}(|\log \mu|/\sqrt{\epsilon}) \leq L \leq c_{17}(|\log \mu|/\sqrt{\epsilon}) + 2\pi \) for some \( c_{17} \geq c_{15} \), i.e. such that (3.80) 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 Theorem 2.9 can be applied by using the construction of \( \xi \)-invariant sets in Proposition 3.31. Let \( (\Omega_0, \mathcal{F}_0, \mu_0, s) \) be the standard Bernoulli shift, where \( \Omega_0 \) is the set of all \( \chi = (\chi_j)_{j \in \mathbb{Z}}, \chi_j \in \{0, 1\}, \mathcal{F}_0 \) is the \( \sigma \)-algebra on \( \Omega_0 \) induced by finite cylinders, \( \mu_0 \) 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 \( t_0 : \Omega_0 \to \mathcal{X} \) and the induced map \( \hat{t}_0 : \Omega_0 \to \hat{\mathcal{X}} \) is defined as follows: we associate to each \( \chi \in \Omega_0 \) a sequence \( \omega_k(\chi) \) and \( \hat{L}_k(\chi), k \in \mathbb{Z} \). We then define \( q_x = q^0(L, (\hat{L}_k(\chi))_{k \in \mathbb{Z}}, (\omega_k(\chi))_{k \in \mathbb{Z}}) \) as in Remark 3.2, and the \( \xi \)-invariant sets \( \mathcal{B}_x \) containing \( q_x, \mathcal{B}_x = \mathcal{B}(L, (\hat{L}_k(\chi))_{k \in \mathbb{Z}}, (\omega_k(\chi))_{k \in \mathbb{Z}}) \) as in Remark 3.3. Let \( \hat{q}_x, \hat{\mathcal{B}}_x \) 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 \( \hat{L}_k \) so that \( q_x \) has a ‘jump’ at \( T = 4nL \) if and only if \( \chi_n = 0 \). More precisely, let \( \hat{\Omega}_0 \) be the subset of \( \Omega_0 \) of \( \chi \) with infinitely many 0 (clearly measurable and \( \mu_0(\hat{\Omega}_0) = 1 \)), and for \( \chi \in \hat{\Omega}_0 \), let \( k_j(\chi) \) be the increasing sequence of integers of ‘positions’ of zeros in \( \chi \), uniquely defined by the requirement \( k_0(\chi) \leq 0, k_1(\chi) \geq 1 \). We now set \( \hat{L}_j(\chi) = 4L \cdot k_j(\chi) \), and complete the definition of \( \hat{q}_x, \hat{\mathcal{B}}_x \). By the definition of the induced localized topology on \( \hat{\mathcal{X}} \), \( \hat{t}_0 : \hat{\Omega}_0 \to \hat{\mathcal{X}} \) is continuous (assuming the product topology on \( \Omega_0, \hat{\Omega}_0 \), thus measurable.

Recall that \( \hat{S} \) is the \( 2\pi \)-shift in the variable \( t \) introduced in §2.3.2. By construction, \( \hat{q}_{s(\chi)} = \hat{S}^{4L/2\pi}(\hat{q}_x), \hat{\mathcal{B}}_{s(\chi)} = \hat{S}^{4L/2\pi}(\hat{\mathcal{B}}_x) \). Let \( \hat{\mu} = (\hat{t}_0)^*\mu_0 \) be the pulled measure, and let
\(\hat{\mathcal{M}}_1 = \hat{\mathcal{I}}_0(\hat{\mathcal{O}}_0)\). As the entropy \(h_{\mu_0}(s) = \log 2\) \([42]\), by construction we have \(h_{\hat{\mu}}(\hat{S}^{4L/2\pi}) = \log 2\). Finally we define a measure \(\mu\) ‘to be shadowed’ by

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

By construction, \(\mu\) is \(\hat{S}\)-invariant. By \([42, \text{Theorem 4.13(i)}]\), \(h_{\hat{\mu}}(\hat{S}) = 2\pi \log 2/(4L)\).

To apply Theorem 2.9 and establish existence of a shadowing measure \(v\), we need to construct a \(\sigma\)-subalgebra \(\mathcal{G}\). First we define the sets \(\mathcal{D}_i \subset \hat{\mathcal{X}}\), \(i \in \mathcal{I}\), where \(\mathcal{I} = \{(j, k, n)\}, j \in \{0, 1\}, k \in \{0, 1, \ldots, 4L/2\pi - 1\}, n \in \mathbb{Z}\) as the set of all \(q = (u, v) \in \hat{\mathcal{X}}\) satisfying the following conditions

\[u(4nL + 2k\pi) \in \left[\pi - 1, \pi + 1\right] \mod 2\pi, \quad j = 0,\]
\[\left[-1, 1\right] \mod 2\pi, \quad j = 1.\]  

(4.4)

Now let \(\mathcal{G}\) be the \(\sigma\)-subalgebra generated by \(\mathcal{D}_i\) \(i \in \mathcal{I}\). As each \(q \in \mathcal{M}_1\) can be represented as \(\hat{S}^{k_0}q_X\) for some \(k_0 = 0, \ldots, 4L/2\pi - 1\), \(\hat{\mathcal{X}} \in \hat{\mathcal{O}}_0\), we define

\[
\mathcal{D}_q = \bigcap_{n=-\infty}^{\infty} \left\{ \mathcal{D}_{\chi_n, k_0, n} \bigcap_{k \neq k_0}^\infty \mathcal{D}_{0, k, n} \right\}. \quad \mathcal{B}_q = \hat{S}^k \mathcal{B}_X.
\]

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 (4.4), because of the relations (3.29), (3.43) and the properties of \(z^+, z^-\) (as there is some \(\hat{T}_m\) such that \(|\hat{T}_m - (4nL + 2k\pi)| \leq 2\pi\), and (ii) in the case \(j = 1\), because of (3.43) and the exponentially fast decay of \(z^+, z^-\) towards \(0 \mod 2\pi\). Thus by Proposition 3.31, 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 2.9 and obtain a \(v \in \mathcal{M}(\hat{\mathcal{E}})\) which shadows \(\mu\). By Lemma 2.8 we get \(h_v(\hat{S}) \geq h_\mu(\hat{S}) \sim 1/L\). The claim now follows from (3.80) and the variational principle for the topological and metric entropy \([42]\). \(\square\)

**Proof of Theorem 1.2(ii)**. Choose \(L \equiv 0 \mod 2\pi\) satisfying \(c_{15}(|\log \mu|) < L \leq c_{15}(|\log \mu|) + 2\pi\), i.e. such that (3.80) holds. Choose an increasing sequence \(\omega_1 = \omega^-; \omega_n = \omega^+\), such that \(\omega_k - \omega_{k-1} \leq \Delta_0/(4c_4(R \vee \mu) \sigma)\) for \(k = 2, \ldots, n\) (as required by (3.12)), with equality for all except possibly \(k = n\). Choose \(\omega_k\) for \(k \not\in \{1, \ldots, n\}\) in an arbitrary way as long as \(\omega_k \in [\omega^-, \omega^+]\) and as long as (3.12) holds. Let \(\tilde{L}_{k+1} - \tilde{L}_k = 4L + 2\pi\) for all except \(k = 1\) and \(k = n\) for which we set

\[
\tilde{L}_{k+1} - \tilde{L}_k = c_0 \sigma^{-2} \frac{|\log \delta|}{\delta^2} + 4L + 2\pi,
\]

(4.5)

for \(c_0\) large enough to be chosen later. By Proposition 3.31, \(\mathcal{B} = \mathcal{B}(L, (\tilde{L}_k(X)))_{k \in \mathbb{Z}},\)
\((\omega_k(X))_{k \in \mathbb{Z}}\) is \(\xi\)-invariant, and by Theorem 2.2, 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 (4.1), as \( \|t^-\| = L_1/2 \), \( \|t^+\| = L_n/2 \), as \( L_1, L_n \geq c_0\sigma_2^2 (|\log \delta|^2/\delta^2) \), and finally as \( \omega_k(t^-) = \omega_1 = \omega^- \) and \( \omega_k(t^+) = \omega_1 = \omega^+ \), we can choose \( c_0 \) in (4.5) to be a large enough absolute constant such that \( |v_t(t^-) - \omega^-| \leq \delta \), \( |v_t(t^+) - \omega^+| \leq \delta \). Now as \( n \sim \sigma(R \lor \mu)(\omega^+ - \omega^-)/\Delta_0 \), and by (3.80) and (4.5),

\[
|t^+ - t^-| \leq \sum_{k=1}^n L_k \ll \sigma^2 \frac{\log \delta}{\delta^2} + nL \\
\ll \sigma^2 \frac{\log \delta}{\delta^2} + \sigma^6 \frac{\log \Delta_1 (R \lor \mu)}{\Delta_0 \Delta_1} (\omega^+ - \omega^-) \\
\ll \sigma^6 \tilde{c}(\delta) \frac{\log \Delta_1 (R \lor \mu)}{\Delta_0 \Delta_1} (\omega^+ - \omega^-),
\]

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

**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 \( \tilde{\omega}^+ - \tilde{\omega}^- \leq \Delta_0/(4c_4(R \lor \mu) \cdot \sigma) \) (i.e. we can by Proposition 3.7 connect any two \( \omega, \tilde{\omega} \) in \( \mathcal{O} \) with a heteroclinic orbit with only one ‘jump’). The only modification as compared with the proof of Theorem 1.2 is now the choice of the sequence \( \omega_k \). Let \( \{\tilde{\omega}_j, j \in \mathcal{P}\} = \{\omega_j, j \in \mathcal{P}\} \), \( \mathcal{P} \) finite or countable, be a dense subset of \( \mathcal{O} \). Let \( p : \mathbb{N} \rightarrow \mathcal{P} \) 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 \( \chi \) 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 1.2, by which we obtain \( \chi \in \mathcal{M}(\mathcal{L}) \), which is supported on the closure of the union of \( \tilde{\xi} \)-invariant sets \( \mathcal{B}(L, (\tilde{L}_k)_{k \in \mathbb{Z}}, (\omega_k(\chi))_{k \in \mathbb{Z}}) \) (considered as subsets of \( \tilde{\mathcal{X}} \)), for \( \chi \in \Omega_0 \). It is easy to check by construction and (4.1) that supp \( \tilde{\pi}^* v \) 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 \( \tilde{\pi}^* v \cap \mathbb{T}_\omega = \emptyset \).

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

\[
n \geq \frac{4c_4(R \lor \mu) \cdot \sigma}{\Delta_0} (\omega^+ - \omega^-)
\]

(i.e. \( n \) is the minimal number of ‘jumps’ required by the Proposition 3.31 and (3.12) to cross the entire \( [\omega^-, \omega^+] \)). Consider the subshift of finite type \( (\Omega_1, \mathcal{F}_1, \mu_1, s) \) (with \( (\Omega_1, \mathcal{F}_1) \) a subspace of \( (\Omega_0, \mathcal{F}_0) \)), \( \mu_1 \) 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 \( \omega_1, \omega_2, \ldots, \omega_m \) in an arbitrary, \( s \)-invariant way, such that (3.12) 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. \( \square \)

**Remark 4.1.** 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 §2.3. More specifically, we could show that \( \tilde{\pi}(\text{supp}(\nu|\mathcal{G})) \) intersects all \( \mathbb{T}_\omega, \omega \in \mathcal{O} \), and that \( \bigcup_{\omega \in \mathcal{O}} \subset \tilde{\pi}(\text{supp}^\mathcal{G}(\nu|\mathcal{G})) \). This would, however, unnecessarily complicate the proof.
4.2. Proof of Theorem 1.3. In this subsection we first state an analogue of Lemma 3.3 which will enable us more precise control for small $\mu$, then evaluate the error term when approximating $S_\omega$ with $M_\omega$, and finally we prove Theorem 1.3.

**Lemma 4.3.** There exist a constant $\tilde{T} > 0$ and functions $\tilde{z}^-, \tilde{z}^+: [-\tilde{T}, 0] \rightarrow \mathbb{R}$ and $\tilde{z}^+: [0, \tilde{T}] \rightarrow \mathbb{R}$, depending only on $\varepsilon, \mu$, satisfying the following 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-, respectively 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

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

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

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

The proof is a relatively straightforward modification of the proof of Lemma 3.3, and thus also done in Appendix B.

**Lemma 4.4.**

(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 $t \in \mathbb{R}$,

$$|u(t) - u^{(\varepsilon)}(t - t_0)| \leq c_{19} \sqrt{\varepsilon \mu}. \quad \text{(4.7)}$$

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

$$|S_\omega(t_1, v_1) - S_\omega(t_0, v_0) - \mu(M_\omega(t_1, v_1) - M_\omega(t_0, v_0))| = O_f(\varepsilon \mu^{3/2}). \quad \text{(4.8)}$$

**Proof.** We first note that analogously to the proof of Proposition 3.4(ii) we can show that, for all $t \in [t_0 - \tilde{T}, t_0]$, we have $u(t) \geq \tilde{z}^-(t - t_0)$, and thus, by Proposition 3.4(ii), $\tilde{z}^-(t - t_0) \leq u(t) \leq \tilde{z}^-(t - t_0)$. We deduce that

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

which implies (4.7) by (3.6) and (4.6). The case $t < -\tilde{T} + t_0$ follows from $\tilde{z}^-(\tilde{T}) = -\sqrt{\varepsilon \mu}$ and $0 \leq u(t) \leq \tilde{z}^-(t - t_0)$, $\tilde{z}^-$ increasing. The proof for $t \geq t_0$ is analogous, which completes (i).

Let $q = (u, v) \in H^1_{\text{loc}}(\mathbb{R})^2$ be a two-sided minimizer at $(\omega, t_0, v_0)$, 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, \( \int_{-\infty}^{\infty} L_\omega(q, \tilde{q}, t) \, dt \geq S_\omega(t_1, v_1) \), we have, by straightforward calculation,

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

(4.9)

By the mean value theorem, (3.18), by the left-sides of (3.19), (3.20) (which also hold for two-sided minimizers which are not necessarily homoclinics) and (4.7), the absolute value of the difference of the integrands in (4.9) is \( \ll_f e^{-(1/2)\sqrt{\epsilon|t-t_0|}} \sqrt{\epsilon\mu} \), and thus (4.9) is \( \ll_f \varepsilon \mu^{3/2} \). The other inequality in (4.9) is obtained analogously, by starting with a two-sided minimizer at \((\omega, t_1, v_1)\).

\[\box empty\]

\textbf{Proof of Theorem 1.3.} This follows from (4.8), as we choose \( \mu_0 \) to be small enough, so that, for \( \mu \leq \mu_0 \), the \( O_f (\varepsilon \mu^{3/2}) \) term is \( \leq \mu \Delta_0 \). Then, for all \( 0 < \mu \leq \mu_0 \), (S1) holds with \( \Delta_0 = \mu \Delta_0 \).

\[\box empty\]

\textbf{Acknowledgement.} This study was partially funded by the Croatian Science Foundation, the grant No IP-2014-09-2285.

\[\box empty\]

\textbf{A. Appendix. 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 [18] 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 by

\[
\|q\|_{L^2_\text{ul}(\mathbb{R})^N} = \sup_{y \in \mathbb{R}} \left( \int_{\mathbb{R}} e^{-|t+y|} |q(t)|^2 \, dt \right)^{1/2},
\]

\[
L^2_\text{ul}(\mathbb{R})^N = \left\{ q \in L^2_{\text{loc}}(\mathbb{R})^N, \|q\|_{L^2_{\text{ul}}(\mathbb{R})^N} < \infty, \lim_{y \to 0} \|\varphi^y q - q\|_{L^2_{\text{ul}}(\mathbb{R})^N} = 0 \right\},
\]

\[
H^k(\mathbb{R})^N = \{ q \in L^2_{\text{ul}}(\mathbb{R})^N | \partial_j^j q \in L^2_{\text{ul}}(\mathbb{R})^N \text{ for all } j \leq k \},
\]

\[
\|q\|_{H^k_{\text{ul}}(\mathbb{R})^N} = \left( \sum_{j=0}^{k} \|\partial_j^j q\|_{L^2_{\text{ul}}(\mathbb{R})^N}^2 \right)^{1/2}.
\]

\textbf{Remark A.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 \( L^2_{\text{ul}}(\mathbb{R})^N \). Specifically, the Lipschitz continuity implies that \( \lim_{y \to 0} \|\varphi^y q - q\|_{L^2_{\text{ul}}(\mathbb{R})^N} = 0 \) holds, i.e. that \( \varphi^y(q) \) is continuous in \( y \in \mathbb{R} \).

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

\[
q_s = -AQ + F(q),
\]

(A.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

\[
\begin{align*}
\tilde{q}_s &= -\tilde{A}\tilde{q} + \tilde{F}(\tilde{q}), \\
\tilde{q}(0) &= 0,
\end{align*}
\]

(A.2)

where \( \tilde{F}(\tilde{q}) = F(\tilde{q} + q^0) - Aq^0 \), and \( \tilde{A} \) is the restriction of \( A \) to \( H^2_{ul}(\mathbb{R})_N \), \( \tilde{A} : H^2_{ul}(\mathbb{R})_N \rightarrow L^2_{ul}(\mathbb{R})_N \). It is straightforward to check that \( \tilde{F} \) is well defined as \( \tilde{F} : H^1_{ul}(\mathbb{R})_N \rightarrow L^2_{ul}(\mathbb{R})_N \), 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 A.1.** For some \( S > 0 \) small enough, there exists an unique solution \( \tilde{q} \) of (A.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 [23, Ch. 3]. First note that \( \tilde{A} \) generates an analytic semigroup \( S(s) = \exp(-s\tilde{A}) \) of bounded linear operators in \( L^2_{ul}(\mathbb{R})_N \), which can for example be verified by using the explicit expression of the heat kernel (see e.g. [18] for details). As \( \tilde{F} \) is uniformly, thus locally, Lipschitz in \( q \), and constant in \( s \), the claim follows from [23, Theorem 3.3.3] with (using the notation from [23]) \( X = L^2_{ul}(\mathbb{R})_N \), \( D(\tilde{A}) = H^2_{ul}(\mathbb{R})_N \), \( \alpha = 1/2 \), \( X^{1/2} = H^1_{ul}(\mathbb{R})_N \).

We now require the well-known fact [23] that \( \tilde{q} \) is a (‘classical’) solution of (A.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

\[
\tilde{q}(s_1) = \int_0^{s_1} e^{-\tilde{A}(s_1-s)} \tilde{F}(\tilde{q}(s)) \, ds.
\]

(A.3)

**Lemma A.2.** The solution \( \tilde{q} \) of (A.2) exists on \((0, \infty)\).

**Proof.** As \( \tilde{F}(\tilde{q}) \) is uniformly bounded in \( L^2_{ul}(\mathbb{R})_N \) by some constant \( A \), we have that if the solution of (A.2) exists on \((0, S)\), then, for all \( s \in (0, S) \), \( \|\tilde{q}(s)\|_{H^2_{ul}(\mathbb{R})_N} \leq e^S \cdot A \). The claim now follows from [23, Corollary 3.3.5], as ‘blow-up’ is not possible.

**Lemma A.3.** The solution of (A.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 (A.3), and obtain

\[
q(s_1) = q^0 + \int_0^{s_1} e^{-\tilde{A}(s_1-s)} (F(q(s)) - Aq^0) \, ds.
\]

(A.4)

Consider a sequence of initial conditions \( q^{0,(n)} \) converging in either \( \mathcal{X}_{ul} \) or \( \mathcal{X}_{loc} \) 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 (A.4), for \( q^{(n)}(s_1) \) and \( q(s_1) \) for \( n \) large enough and \( s_1 > 0 \) small enough, in either \( \mathcal{X}_{ul} \)- or \( \mathcal{X}_{loc} \)-norm (see also [23, Corollary 3.4.1] for details).
LEMMA A.4. If $V \in H^k(\mathbb{T}^{N+1})$, $k \geq 2$, then, for all $s > 0$, if $q$ is the solution of (A.1), then $q(s) \in H^k_{ul}(\mathbb{R})^N$.

Proof. We prove the lemma inductively. In the proofs of Lemmas A.1, A.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, $r(s) \in H^1_{ul}(\mathbb{R})^N$ for all $s > 0$. For arbitrarily small $\delta > 0$, consider the system of equations

$$r_s = -\tilde{A}r + F^{(k-2)}(s),$$

$$r(\delta) = \frac{d^{k-2}}{dt^{k-2}}q(\delta)$$

(A.5)

where $F^{(k-2)}(s) = (d^{k-2}/dt^{k-2})(\partial/\partial q)V(q(s, \cdot), \cdot)$ is a fixed function. One can verify, by using the inductive assumption $q \in H^k_{ul}(\mathbb{R})^N$, the assumed regularity of $V$ and the embedding properties of the uniformly local spaces \cite{18}, that $F^{(k-2)}(s) \in L^2_{ul}(\mathbb{R})^N$ for all $s \in (0, \infty)$, and that it is uniformly bounded in $L^2_{ul}(\mathbb{R})^N$ on $(\delta, S]$ for any $S > \delta$. Now by repeating the argument of existence and uniqueness of solutions as in Lemma A.1, we deduce that, for any $s \in (\delta, \infty)$, $r(s)$ is a solution of (A.5), thus $r(s) \in H^2_{ul}(\mathbb{R})^N$ and $q(s) \in H^k_{ul}(\mathbb{R})^N$. As $\delta > 0$ is arbitrarily small, the claim is proved.

Theorem 2.1 now follows from Lemmas A.1–A.4.

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

LEMMA A.5. If $q$ is a solution of (2.3) such that either $q_t \in L^\infty(\mathbb{R})^N$ or $q_{tt} \in L^2_{ul}(\mathbb{R})^N$, 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 $q_t \in L^2_{ul}(\mathbb{R})^N$, 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 A.1, we now have $q_t \in L^2_{ul}(\mathbb{R})^N$, $q_{tt} \in L^2_{ul}(\mathbb{R})^N$, thus $q \in \mathcal{X}$.

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

This appendix is dedicated to the proofs of Lemma 3.3 and Proposition 3.4 in §3.1 and Lemma 4.3 in §4.2, i.e. the construction of one-sided minimizers in §3.1 and calculation of a priori bounds.

B.1. Proofs from §3.1.

LEMMA B.1. Define $\tilde{w}(t) = 4 \arctg \exp(\sqrt{\varepsilon(1-2\mu^{1/2})}t)$, 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 we have the following.
(i) \(0 < t_1 \leq 2\sqrt{\mu/\varepsilon}\).
(ii) \(w\) is \(C^2\), for all \(t \in [-1/(2\sqrt{\varepsilon}), \infty)\), \(0 < w_t < 2\sqrt{\varepsilon}\), and for all \(t \in [-1/(2\sqrt{\varepsilon}), 1/(2\sqrt{\varepsilon})]\),
\[
\sqrt{\varepsilon}/2 \leq w_t.
\] (B.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 \(4\arctg e^t\) is greater than or equal to \(9/5\) 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 \(t^* = -\mu^{1/2}/2\), \(|t^* - t_1| \leq \sqrt{\mu}/\sqrt{\varepsilon}\), and thus the value of the polynomial in the definition of \(w\) is less than or equal to \(27\mu^2\). 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
\[
\tilde{F}(w) := w_{tt} - \varepsilon \sin w(t) - \varepsilon \mu |\sin w(t)| - \varepsilon \mu (1 - \cos w(t)) > 0
\]
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
\[
\tilde{F}(w) = (\varepsilon (1 - 2\mu^{1/2})) \sin \tilde{w}(t) + 6\varepsilon^{3/2}\mu^{1/2}((t_1 - t) \lor 0)
- \varepsilon \sin w(t) - \varepsilon \mu |\sin w(t)| - \varepsilon \mu (1 - \cos w(t))
\geq 6\varepsilon^{3/2}\mu^{1/2}((t_1 - t) \lor 0) - 2\varepsilon \mu^{1/2} \sin w(t) - \varepsilon \mu |\sin w(t)| - \varepsilon \mu (1 - \cos w(t)).
\] (B.2)

Denote the expression (B.2) by \(\tilde{F}(w(t))\). It suffices to show now that \(\tilde{F}(w(t)) > 0\).

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
\[
\tilde{F}(w) > \varepsilon \mu^{1/2} |\sin w| - \varepsilon \mu (1 - \cos w).
\] (B.3)

For any \(w \in (0, \pi]\), the inequality
\[
|\sin w| \geq \frac{1}{2} |w - \pi|(1 - \cos w).
\] (B.4)

holds. Inserting this and \(|w - \pi| \geq 2\mu^{1/2}\) in (B.3) we get \(\tilde{F}(w) > 0\).

Let \(w \in [\pi, \pi + 2\mu^{1/2}]\), or equivalently \(t \in [t_0, t_1]\). As \(\tilde{F}(w(t_1)) > 0\), it suffices to show that \(\tilde{F}(w(t)) - \tilde{F}(w(t_1)) \geq 0\). Calculating we get
\[
\tilde{F}(w(t)) - \tilde{F}(w(t_1)) \geq 6\varepsilon^{3/2}\mu^{1/2}[t - t_1| - (2\varepsilon \mu^{1/2} + \varepsilon \mu)|\sin w(t) - \sin w(t_1)|
- \varepsilon \mu |\cos w(t) - \cos w(t_1)|.
\] (B.5)
Using the mean-value theorem, (B.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 this with (B.5) to get $\mathcal{F}(w(t)) - \mathcal{F}(w(t_1)) \geq 0$.

We now also combine with (B.5) to get $\mathcal{F}(w(t)) - \mathcal{F}(w(t_1)) \geq 0$.

We consider the functional
\[ \mathcal{F}(w(t)) = \int_{t_0}^{t_1} \left( \frac{1}{2} \left| \frac{d^2 w}{dt^2} \right|^2 + \varepsilon \mu \left| \frac{d w}{dt} \right|^2 + \left| w(t) - \frac{1}{\sqrt{\varepsilon}} \right|^2 \right) dt. \]

Proof of Lemma 3.3. We take $w$ constructed in Lemma B.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 $2\pi - 4 \arctan x < 4/x$, $4 \arctan(1/x) = 2\pi - 4 \arctan x$, holding for all $x > 0$. By the mean value theorem, by noting that $|d(\arctan e^x)/dt| \leq 4e^{-|x|}$, the construction and bounds on $t_0$, $t_1$, we easily get that, for $t \geq 0$,
\[ |w(t - t_0) - u^{(\epsilon)}(t)| \ll e^{-(1/2)}\sqrt{\varepsilon} (\varepsilon \sqrt{\mu} t + \sqrt{\varepsilon \mu}) \ll \sqrt{\varepsilon \mu}, \]
which completes the proof. \qed

We now construct one-sided minimizers and prove Proposition 3.4. 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
\[ \mathcal{L}_{\omega,k} (q) = \int_{t_0}^{t_0+k} L_{\omega}(u(t), v(t), t) \, dt. \]
We construct minimizers $q_k$ of (B.6), and then obtain $q^+$ as their limit.

Lemma B.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
\[ q(t_0) = (\pi, v_0), \, u(t_1) = 2\pi. \] (B.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 [30, 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 (B.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 $(t_0, t_0 + k)$. \qed

We denote by $q_k$ the (not necessarily unique) minimizer of (B.6) satisfying (B.7). We always set $u_k(t) = 2\pi$ for $t \geq t_0 + k$.

Lemma B.3. The minimizer $q_k = (u_k, v_k)$ satisfies the following 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 \( t \geq t_0 \), \( u_k(t) \geq \varepsilon^+ (t - t_0) \). Furthermore, for all \( 0 \leq T \leq 3/(4\sqrt{\varepsilon}) \),

\[
u_k \geq \varepsilon^+ (t - t_0 - T).
\]

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 \( L_{c,k}(\tilde{q}) \leq L_{c,k}(q_k) \), thus \( \tilde{q} \) minimizes \( L_{c,k} \) and must be a solution of the Euler–Lagrange equations on \( (t_0, t_0 + k) \). We deduce that \( \tilde{a} = 2\pi \) which is in contradiction to \( \tilde{a}(t_0) = u_k(t_0) = \pi \). 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 (3.2), we can find \( k_0 \) large enough so that, for any \( k \geq k_0 \),

\[
L_{c,k}(q_k) \leq 4\sqrt{\varepsilon}(1 + 3\mu/2). \tag{B.8}
\]

Assume \( u_k(t_0 + d) = \pi \) for some \( t_0 + d \in (t_0, t_0 + k) \). Again by (3.2), as \( q^k \) is a minimizer, we see that

\[
\int_{t_0 + k}^{t_0 + d} L_\omega(q_k, (q_k)_t, t) \, dt \geq 4\sqrt{\varepsilon}(1 - \mu). \tag{B.9}
\]

From (B.8) and (B.9) and the standing assumption \( \mu \leq 1/16 \) we obtain the upper bound

\[
\int_{t_0}^{t_0 + d} L_\omega(q_k, (q_k)_t, t) \, dt \leq 4\sqrt{\varepsilon}(1 + 3\mu/2) - 4\sqrt{\varepsilon}(1 - \mu) \leq 6\sqrt{\varepsilon}\mu. \tag{B.10}
\]

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

\[
\int_{t_0}^{t_0 + t^*} L_\omega(q_k, (q_k)_t, t) \, dt \geq \int_{t_0}^{t_0 + t^*} \left( \frac{1}{2} (u_k)_t^2 + \varepsilon(1 - \mu)(1 - \cos u_k) \right) \, dt \\
\geq \frac{1}{2t^*} a^2 + \varepsilon(1 - \mu) \left( 2 - \frac{a^2}{2} \right) t^* \\
\geq 2\varepsilon(1 - \mu) t^* \geq \frac{3}{2} \varepsilon t^*.
\]

Repeating that over \( [t_0 + d - t^*, t_0 + d] \) we get \( \int_{t_0}^{t_0 + d} L_\omega(q_k, (q_k)_t, t) \, dt \geq 3\varepsilon t^* \), and thus, by (B.10), \( t^* \leq 2\mu/\sqrt{\varepsilon} \leq 1 \). By definition, \( d \leq 4\mu/\sqrt{\varepsilon} \). The claim follows by continuity of \( u_k \).

(iii) Using the same notation as in (ii) and by (B.10), we easily see that

\[
6\sqrt{\varepsilon}\mu \geq \int_{t_0}^{t_0 + d} L_\omega(q_k, q^k_t, t) \, dt \geq \frac{2}{d} a^2 \geq \frac{\sqrt{\varepsilon}}{2\mu} a^2,
\]

thus \( a^2 \leq 12\mu^2 \). By (A2), \( a \leq 1/4 \).
(iv) We prove this by using Lemma 3.2 in two steps. First we show that
\[ u_k(t) > z^+(t - t_0 - T) \quad \text{for all } t \geq t_0 + T - 3/(4\sqrt{\varepsilon}) \]  
(B.11)
and all \( T \geq 3/(4\sqrt{\varepsilon}). \) By definition and Lemma 3.3(i), (B.11) holds for \( T = k + 3/(4\sqrt{\varepsilon}). \)
Assume the contrary and find the infimum \( T^* \) of \( T \geq 3/(4\sqrt{\varepsilon}) \) for which (B.11) holds. By compactness and continuity, we have that (B.11) holds for \( T = T^* \) with \( \geq \) instead of \( >. \)
However, by construction, (ii), (iii) and Lemma 3.3(v), the strict inequality in (B.11) holds for \( t \in \{t_0 + T - 3/(4\sqrt{\varepsilon}), t_0 + k\} \), thus by Lemma 3.2 we must have strict inequality in (B.11) and \( T^* = 3/(4\sqrt{\varepsilon}). \)

Now we show that
\[ u_k(t) > z^+(t - t_0 - T) \quad \text{for all } t \geq t_0 \]  
(B.12)
for all \( T \in (0, 3/(4\sqrt{\varepsilon})]. \) Again we find the infimum \( T^* \) for which (B.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 (B.12) holds with \( \geq \) instead of \( >, \) which we need to show.

Proof of Proposition 3.4. It is easy to check that, for \( k \geq k_0, k_0 \) as in Lemma B.3, \( q_k \) is uniformly bounded in \( H^2_{\text{loc}}([t_0, \infty)) \), i.e. that, for any \( T > k_0, \) we have that \( q_k, k \geq T, \) is uniformly bounded in \( H^2([t_0, t_0 + T]) \). Indeed, uniform bounds in \( k \) in \( \|q(t)\|_{L^2([t_0, t_0 + T])} \) 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 \( \|q_k\|_{L^2([t_0, t_0 + T])} \) follows from that and \( q(t_0) = (\pi, v_0). \) The uniform bound in \( k \) on \( \|q_k(t)\|_{L^2([t_0, t_0 + T])} \) is deduced from the fact that \( q_k \) is a solution of Euler–Lagrange equations, and the uniform bounds \( |V_\mu| \leq \varepsilon(1 + \mu), |V_0| \leq \varepsilon\mu. \)

Now by diagonalization we find a convergent subsequence of \( q_k \) in \( H^1_{\text{loc}}([t_0, \infty)) \)
converging to some \( q^0 \in H^1_{\text{loc}}([t_0, \infty)). \) It is straightforward to show that \( \lim_{k \to \infty} \mathcal{L}_{\omega, k}(q_k) = S^+_0(t_0, v_0), \) as \( \mathcal{L}_{\omega, k}(q_k) \) is decreasing, bounded by \( S^+_0(t_0, v_0) \) from below, and we can arbitrarily well approximate \( S^+_0(t_0, v_0) \) with \( \mathcal{L}_{\omega, k}(q_k) \) for \( k \) large enough.

By construction and Lemma B.3(iv), \( \lim_{t \to \infty} u^0(t) = 2\pi. \) By the Fatou lemma, \( \int_{t_0}^{\infty} \mathcal{L}_{\omega}(q^0, q^0, t) \leq S^+_0(t_0, v_0), \) thus by the definition of \( S^+_0(t_0, v_0), \) 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 \( (t_0, \infty), \) and \( C^4 \) because of the regularity of solutions of ordinary differential equations.

The proof of (ii) is analogous to the proof of Lemma B.3(iv). \( \square \)

B.2. Proof from §4.2. Let \( \tilde{\varepsilon} = \sqrt{\varepsilon(1 + 2\mu)} \), and let \( \delta = 2\sqrt{\varepsilon}\mu. \) Denote by \( u^{(\tilde{\varepsilon})} = 4 \arctg(e^{\sqrt{\varepsilon} t}) \) the separatrix solution of the unperturbed pendulum \( u_{tt} = \tilde{\varepsilon} \sin u. \)

Lemma B.4. If \( \tilde{u} \) is a solution of \( u_{tt} = \tilde{\varepsilon} \sin u, \) \( u(0) = \pi, u_t(0) = 2\sqrt{\varepsilon}(1 + \delta^2), \) then for some absolute \( c_{30} > 0, \) we can find \( t_1 > 0 \) such that
\[ t_1 \leq \frac{c_{30}}{\sqrt{\varepsilon}} \log \delta, \]  
(B.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 \( w^{(1)} = \tilde{u} - u^{(\varepsilon)} \), let \( t \geq 0 \), and let \( t_1 > 0 \) be the unique (by monotocity) \( t \) such that \( \tilde{u}(t) = 2\pi + \delta \). Clearly \( \tilde{u}(t) \geq u^{(\varepsilon)}(t) \), thus \( w^{(1)}(t) \geq 0 \). By definition, \( w^{(1)} \) solves the differential inequality \( w^{(1)}(t) \leq \varepsilon \tilde{u}'(t) - u^{(\varepsilon)}(t) \). Now if \( w^{(2)} = w^{(1)}(t) - \sqrt{\varepsilon} w^{(1)}(1) \), \( w^{(3)} = w^{(1)}(t) + \sqrt{\varepsilon} w^{(1)}(1) \), we have \( w^{(2)}(0) = w^{(3)}(0) = 2\sqrt{\varepsilon} \cdot \delta^2 \), and they are solutions of \( w^{(2)}(t) \leq -\sqrt{\varepsilon} w^{(2)}(1) \), \( w^{(3)}(t) \leq \sqrt{\varepsilon} w^{(3)}(1) \). Thus by the Gronwall lemma,

\[
 w^{(2)}(t) \leq 2\sqrt{\varepsilon} \delta^2 e^{-\sqrt{\varepsilon} t}, \quad w^{(3)}(t) \leq 2\sqrt{\varepsilon} \delta^2 e^{\sqrt{\varepsilon} t}. \tag{B.14}
\]

By the conservation of energy, it is easy to show that, for all \( t \geq 0 \), \( \tilde{u}'(t) \geq 2\sqrt{\varepsilon} \delta \). Inserting that in the left-hand side of (B.14) implies, for all \( t \in [0, t_1] \),

\[
 w^{(1)}(t) \leq 2\delta^2 e^{\sqrt{\varepsilon} t_1}. \tag{B.15}
\]

Proof of Lemma 4.3. Take \( \tilde{u} \) from Lemma B.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 B.1, we can find \( 0 < t_0 < \mu^{1/2} \) such that \( w(t_0) = \pi \). We set \( \tilde{z}^+(t) = w(t - t_0) \), \( \tilde{z}^- = 2\pi - w(-t + t_0) \). The rest of the proof is analogous to the proof of Lemma 3.3, using also the relation \(|u^{(\varepsilon)} - u^{(\varepsilon)}| \ll \varepsilon^{1/2} \mu \) to obtain (v). \( \square \)

C. Appendix. Proofs of bounds on derivatives

This appendix is dedicated to the proof of Proposition 3.18. In the first subsection we obtain weighted \textit{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.

C.1. \textit{A priori} bounds on weighted integrals. Throughout the subsection we assume \( q \in B_1 \).

\begin{lemma}
 Let \( j, m \in \{1, 2, 3, 4, 5, 6\} \). Then for some absolute constant \( c_{31} > 0 \),

\[
 e^{(j+1)/2} \int_{\mathbb{R}} e^{-\lambda(\tau)|t-\tau|} e^{-\lambda(\mu/2)\sqrt{|t|}} dt \leq c_{31} \lambda(\tau)^j, \tag{C.1}
\]

\[
 \int_{\mathbb{R}} \frac{1}{L^{j+1}} e^{-\lambda(\tau)|t-\tau|} dt \leq c_{31} \lambda(\tau)^j. \tag{C.2}
\]

\end{lemma}
Proof. To show (C.1), without loss of generality let
\[ \tau \in [\tilde{T}_k - 2L_{k-1}, \tilde{T}_{k+1} + 2L_k] \]  \hspace{1cm} \text{(C.3)}
for a fixed \( k \in \mathbb{Z} \), i.e. \( \|\tau\| = |\tau - \tilde{T}_k| \). By applying the definition of \( \|\tau\| \) and substituting \( t - T_j + \tilde{T}_k \) instead of \( \tau \) in the second line below, we get
\[
\int_\mathbb{R} e^{-\lambda(\tau)|t-\tau|} e^{-(m/2)\sqrt{\|\tau\|}} \, dt = \sum_{j=-\infty}^{\infty} \int_{T_j-2L_{j-1}}^{T_j+2L_j} e^{-\lambda(\tau)|t-\tau|} e^{-(m/2)\sqrt{\|\tau-T_j\|}} \, dt \\
= \sum_{j=-\infty}^{\infty} \int_{\tilde{T}_k-2L_{k-1}}^{\tilde{T}_k+2L_k} e^{-\lambda(\tau)|t-\tilde{T}_k+T_j-\tau|} e^{-(m/2)\sqrt{\|\tau-\tilde{T}_k\|}} \, dt. \]  \hspace{1cm} \text{(C.4)}

Now by the assumption (C.3), for \( j = k-1, k, k+1 \) we have \( |T_j - \tau| \geq |\tilde{T}_k - \tau| \). For \( j \geq k+1 \), let \( |T_j - \tau| \geq |T_j - \tilde{T}_{k+1}| + |\tilde{T}_k - \tau| \geq 4L((j - k - 1) + |\tilde{T}_k - \tau| \). By analogous consideration for \( j \leq k-1 \), we conclude that, for all \( j \in \mathbb{Z} \),
\[
|T_j - \tau| \geq |\tilde{T}_k - \tau| + 4L(|j - k| - 1) = \|\tau\| + 4L(|j - k| - 1) \lor 0. \]  \hspace{1cm} \text{(C.5)}

Now by definition of \( \lambda(\tau) \), we have \( \sqrt{\varepsilon} \, m/2 \geq \sqrt{\varepsilon}/2 \geq \lambda(\tau) + \sqrt{\varepsilon}/4 \). Combining this with (C.5), we can estimate the exponent in (C.4):
\[
\lambda(\tau)|t - \tilde{T}_k + T_j - \tau| + \frac{m}{2} \sqrt{\varepsilon}|t - \tilde{T}_k| \geq \lambda(\tau)|T_j - \tau| - \lambda(\tau)|t - \tilde{T}_k| + \frac{m}{2} \sqrt{\varepsilon}|t - \tilde{T}_k| \\
\geq \lambda(\tau)\|\tau\| + \lambda(\tau)4L(|j - k| - 1) \lor 0 + \frac{1}{4} \sqrt{\varepsilon}|t - \tilde{T}_k|. \]

Thus (C.4) is less or equal than
\[
e^{-\lambda(\tau)\|\tau\|} \left( 2 + \sum_{j=-\infty}^{\infty} e^{-\lambda(\tau)4L-j} \right) \int_{-\infty}^{\infty} e^{-(1/4)\sqrt{\varepsilon}|t-\tilde{T}_k|} = \frac{4}{\sqrt{\varepsilon}} e^{-\lambda(\tau)\|\tau\|} \frac{1}{1 - e^{-4\lambda(\tau)\varepsilon}}. \]  \hspace{1cm} \text{(C.6)}

If \( \lambda(\tau) = \sqrt{\varepsilon}/4 \), then \( 4L\lambda(\tau) \geq 1 \), and thus the right-hand side of (C.6) is \( \ll \varepsilon^{-1/2} \). This implies (C.1). Otherwise \( \lambda(\tau) = 8 \log\|\tau\||/\|\tau\| \), so the right-hand side of (C.6) becomes
\[
\frac{4}{\sqrt{\varepsilon} (1 - \|\tau\||-4L/\|\tau\|}. \]  \hspace{1cm} \text{(C.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 reciprocal of the denominator in (C.7) is \( \ll \|\tau\| \log(1/\|\tau\|) \ll \|\tau\|^2 \). Thus (C.7) is \( \ll \varepsilon^{-1/2} \|\tau\|^{-6} \leq \varepsilon^{-1/2} \lambda(\tau)^6 \), which implies (C.1).

To show (C.2), let \( \tau \in [\tilde{T}_k, \tilde{T}_{k+1}] \) for fixed \( k \in \mathbb{Z} \). As \( L_k \geq 4L \geq 1/\sqrt{\varepsilon} \),
\[
\int_{\mathbb{R}} \frac{1}{L_k(t)} e^{-\lambda(\tau)|t-\tau|} \, dt \leq \frac{1}{L_k^{j+1}} \int_{\tilde{T}_k}^{\tilde{T}_{k+1}} e^{-\lambda(\tau)|t-\tau|} \, dt \\
+ \varepsilon^{(j+1)/2} e^{-\lambda(\tau)|\tau - \tilde{T}_k|} \int_{-\infty}^{\tilde{T}_k} e^{-\lambda(\tau)|t-\tilde{T}_k|} \, dt.
\]
Proof. As \( L_k \geq 1/\lambda (\tau) \), the first summand in (C.8) is always \( \ll \lambda (\tau)^j \). If \( \lambda (\tau) = \sqrt{\varepsilon}/8 \), the second summand \( \ll \varepsilon^{j/2} \ll \lambda (\tau)^j \). Otherwise, as \( \varepsilon \leq 1 \), and as \( \|\tau\| \geq 1/(8\lambda (\tau)) \),

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

which completes (C.2). \( \square \)

**Lemma C.2.** There exists an absolute constant \( c_{32} > 0 \) so that

\[
\| V_u \|_{L^2_t(\mathbb{R})}^2 \leq c_{32} \lambda (\tau)^3, \quad \| V_v \|_{L^2_t(\mathbb{R})}^2 \leq c_{32} \lambda (\tau)^3. \tag{C.9}
\]

**Proof.** As \( q \in B_1 \), by (3.41), we get

\[
|\sin u(t)| \leq |u - 2k(t)\pi| \ll e^{-(1/2)\sqrt{\|\tau\|}}, \tag{C.10}
\]

\[
|1 - \cos u(t)| \leq \frac{1}{2} |u - 2k(t)\pi|^2 \ll e^{-(1/2)\sqrt{\|\tau\|}}. \tag{C.11}
\]

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

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

\[
|V_u(u, v, t) | \leq 2\varepsilon (1 - \cos u(t) + |\sin u(t)|) \ll \varepsilon e^{-(1/2)\sqrt{\|\tau\|}}, \tag{C.12}
\]

\[
|V_v(u, v, t) | \leq \varepsilon \mu (1 - \cos u(t)) \ll \varepsilon \mu e^{-\sqrt{\|\tau\|}}. \tag{C.13}
\]

It suffices now to apply (C.1). \( \square \)

**Lemma C.3.** There exists an absolute constant \( c_{33} > 0 \) so that

\[
\| u_t^0 \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} \lambda (\tau), \quad \| v_t^0 - \omega_k(\tau) \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} \lambda (\tau), \tag{C.14}
\]

\[
\| u_{tt}^0 \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} \lambda (\tau)^3, \quad \| v_{tt}^0 \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} \lambda (\tau)^3. \tag{C.15}
\]

\[
\| u_{ttt}^0 \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} (\omega^2 + 1) \lambda (\tau)^3, \quad \| v_{ttt}^0 \|_{L^2_t(\mathbb{R})}^2 \leq c_{33} (\omega^2 + 1) \lambda (\tau)^3. \tag{C.16}
\]

**Proof.** This follows by a straightforward calculation from Lemmas 3.14 and C.1. \( \square \)

**C.2. Proof of invariance of \( B_2, B_3, B_4. \)** 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 C.4.** If \( q = (u, v) \in B_2 \), then

\[
\| u_t - u_t^0 \|_{L^2_t(\mathbb{R})}^4 \leq c_{34} (\lambda (\tau)^2 + \lambda (\tau)^{-1}) \| u_{tt} - u_{tt}^0 \|_{L^2_t(\mathbb{R})}^2, \tag{C.17}
\]

\[
\| v_t - v_t^0 \|_{L^2_t(\mathbb{R})}^4 \leq c_{35} (\lambda (\tau)^4 M^4 + \lambda (\tau)^{-1} M^2) \| v_{tt} - v_{tt}^0 \|_{L^2_t(\mathbb{R})}^2, \tag{C.18}
\]

for some absolute constants \( c_{34}, c_{35} \geq 1 \).
Proof. Denote by \( X = \|u_t - u_t^0\|_{L^2_t(L^2_x(\mathbb{R}))}^2 \) and by \( Y = \|u_{tt} - u_{tt}^0\|_{L^2_t(L^2_x(\mathbb{R}))}^2 \). By partial integration and the Cauchy–Schwarz inequality, we get

\[
X = \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_t(t) - u_t^0(t))^2 \, dt \\
\leq \lambda(\tau) \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}|u(t) - u^0(t)||u_t(t) - u_t^0(t)| \, dt \\
+ \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}|u(t) - u^0(t)||u_{tt}(t) - u_{tt}^0(t)| \, dt \\
\ll \lambda(\tau)^{1/2} \|u - u^0\|_{L^\infty_t(L^2_x)} X^{1/2} + \lambda(\tau)^{-1/2} \|u - u^0\|_{L^\infty_t(L^2_x)} Y^{1/2}.
\tag{C.19}
\]

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

\[
X^2 \ll \lambda(\tau)^2 \|u - u^0\|^4_{L^\infty_t(L^2_x)} + \lambda(\tau)^{-1} \|u - u^0\|^2_{L^\infty_t(L^2_x)}.
\tag{C.20}
\]

From (3.31) and (3.41) we have \( \|u - u^0\|_{L^\infty_t(L^2_x)} \leq c_6 + c_7 \), which gives (C.17).

To show (C.18), it suffices to note that by (3.42), \( \|v - v^0\|_{L^\infty_t(L^2_x)} \leq c_8 M \). The rest of the proof is analogous. \( \square \)

Lemma C.5. The set of all \( q \in B_1 \) such that \( \|u_t - u_t^0\|^2_{L^2_t(L^2_x)} \leq c_9 \lambda(\tau) \) is \( A \)-relatively \( \xi \)-invariant for some absolute constant \( c_9 > 0 \).

Proof. First note that

\[
u_s - (u_{tt} - u_{tt}^0) = -V_u + u_{tt}^0,
\]

thus by squaring it we get

\[
-2u_s (u_{tt} - u_{tt}^0) \leq -u_s^2 - (u_{tt} - u_{tt}^0)^2 + 2V_u^2 + 2(u_{tt}^0)^2,
\tag{C.21}
\]

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 as a function of \( s \), we obtain

\[
\frac{d}{ds}\|u_t - u_t^0\|^2_{L^2_t(L^2_x(\mathbb{R}))} = 2 \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_t - u_t^0) u_{ts} \, dt \\
\leq -2 \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_{tt} - u_{tt}^0) u_{ts} \, dt \\
+ 2\lambda(\tau) \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_t - u_t^0) u_s \, dt.
\tag{C.22}
\]

The first summand is estimated by inserting (C.21),

\[
-2 \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_{tt} - u_{tt}^0) u_{ts} \, dt \leq -\|u_s\|^2_{L^2_t(L^2_x(\mathbb{R}))} - \|u_{tt} - u_{tt}^0\|^2_{L^2_t(L^2_x(\mathbb{R}))} \\
+ 2\|V_u\|^2_{L^2_t(L^2_x(\mathbb{R}))} + 2\|u_{tt}^0\|^2_{L^2_t(L^2_x(\mathbb{R}))}.
\tag{C.23}
\]

By Young’s inequality, we have that

\[
2\lambda(\tau) \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|}(u_t - u_t^0) u_s \, dt \leq \lambda(\tau)^2 \|u_t - u_t^0\|^2_{L^2_t(L^2_x(\mathbb{R}))} + \|u_s\|^2_{L^2_t(L^2_x(\mathbb{R}))}.
\tag{C.24}
\]
Inserting (C.23) and (C.24) into (C.22), we see that
\[
\frac{d}{ds} \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 \leq -\| u_{tt} - u_{tt}^0 \|_{L^2_t(L^2_x)}^2 + \lambda^2(\tau) \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 + 2 \| V_u \|_{L^2_t(L^2_x)}^2 + 2 \| u_{tt}^0 \|_{L^2_t(L^2_x)}^2.
\] (C.25)
By substituting \( X = \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 \), \( Y = \| u_{tt} - u_{tt}^0 \|_{L^2_t(L^2_x)}^2 \), and using (C.9) and (C.15), we obtain from (C.25)
\[
\frac{d}{ds} X \leq -Y + \lambda^2(\tau) X + 2(c_{32} + c_{33}) \lambda(\tau)^3.
\] (C.26)
Equation (C.17) can be written as
\[
-\frac{1}{2} Y \leq -\frac{\lambda(\tau)}{2c_{34}} X^2 + \frac{1}{2} \lambda(\tau)^3.
\] (C.27)
Clearly
\[
0 \leq \frac{\lambda(\tau)}{2c_{34}} X^2 - 2\lambda(\tau)^2 X + 2c_{34} \lambda(\tau)^3,
\] (C.28)
thus by summing (C.27) and (C.28) we see that
\[
-\frac{1}{2} Y + \lambda(\tau)^2 X \leq -\lambda(\tau)^2 X + (2c_{34} + 1/2) \lambda(\tau)^3.
\]
Summing this with (C.26) and substituting back \( X \) and \( Y \), we obtain a differential inequality:
\[
\frac{d}{ds} \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 \leq -\frac{1}{2} \| u_{tt} - u_{tt}^0 \|_{L^2_t(L^2_x)}^2 + \lambda(\tau)^2 \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 + c_9 \lambda(\tau)^3
\] (C.29)
\[
\leq -\lambda(\tau)^2 \| u_t - u_t^0 \|_{L^2_t(L^2_x)}^2 + c_9 \lambda(\tau)^3,
\] (C.30)
where \( c_9 = 2(c_{32} + c_{33} + c_{34}) + 1/2 \). The claim follows from the Gronwall’s lemma and Lemma 3.15, i.e. \( A \)-relative invariance of \( B_1 \). \( \square \)

**Lemma C.6.** The set of all \( q \in B_1 \) such that \( \| v_t - v_t^0 \|_{L^2_t(L^2_x)}^2 \leq c_9(M^2 + 1) \lambda(\tau) \) is \( A \)-relatively \( \xi \)-invariant for some absolute constant \( c_9 > 0 \).

**Proof.** Denote by \( \tilde{X} = \| v_t - v_t^0 \|_{L^2_t(L^2_x)}^2 \), \( \tilde{Y} = \| v_{tt} - v_{tt}^0 \|_{L^2_t(L^2_x)}^2 \). All the steps in the proof are analogous to the proof of Lemma C.5 up to the equations (C.27), (C.28), instead of which we have
\[
-\frac{1}{2} \tilde{Y} \leq -\frac{\lambda(\tau)}{2c_{35}M^2} \tilde{X}^2 + \frac{1}{2} \lambda(\tau)^3 M^2,
\]
\[
0 \leq \frac{\lambda(\tau)}{2c_{35}} \tilde{X}^2 - 2\lambda(\tau)^2 \tilde{X} + 2c_{35} \lambda(\tau)^3 M^2.
\]
We thus obtain the differential inequality
\[
\frac{d}{ds} \| v_t - v_t^0 \|_{L^2_t(L^2_x)}^2 \leq -\frac{1}{2} \| v_{tt} - v_{tt}^0 \|_{L^2_t(L^2_x)}^2 + \lambda(\tau)^2 \| v_t - v_t^0 \|_{L^2_t(L^2_x)}^2 + c_9(M^2 + 1) \lambda(\tau)^3
\] (C.31)
\[
\leq -\lambda(\tau)^2 \| v_t - v_t^0 \|_{L^2_t(L^2_x)}^2 + c_9(M^2 + 1) \lambda(\tau)^3,
\] (C.32)
which completes the proof analogously as in Lemma C.5. \( \square \)
Lemmas C.5 and C.6 compete the proof of $A$-relative $\xi$-invariance of $B_2$.

**Lemma C.7.** There exists a constant $c_{10} > 0$ such that the set of all $q \in B_2$ satisfying (3.49a), (3.49b) is an $A$-relative $\xi$-invariant set.

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

$$-2u_{ttt}u_{ts} = -u_{tt}^2 - u_{ts}^2 + (D_t V_u)^2,$$

thus

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

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

$$\frac{d}{ds} \|u_{tt}\|_{L^2_t(\mathbb{R})}^2 = -\|u_{ttt}\|_{L^2_t(\mathbb{R})}^2 - \|u_{ts}\|_{L^2_t(\mathbb{R})}^2 + 2 \int_{\mathbb{R}} e^{-\lambda(\tau)|t-t'|} (u_{tt} u_{ts})_t dt + \|D_t V_u\|_{L^2_t(\mathbb{R})}^2 \leq -\|u_{ttt}\|_{L^2_t(\mathbb{R})}^2 - \|u_{ts}\|_{L^2_t(\mathbb{R})}^2 + \lambda(\tau)^2 \|u_{tt}\|_{L^2_t(\mathbb{R})}^2$$

$$+ \|u_{ts}\|_{L^2_t(\mathbb{R})}^2 + \|D_t V_u\|_{L^2_t(\mathbb{R})}^2 \|D_t V_u\|_{L^2_t(\mathbb{R})}^2 \leq -\|u_{ttt}\|_{L^2_t(\mathbb{R})}^2 + \lambda(\tau)^2 \|u_{tt}\|_{L^2_t(\mathbb{R})}^2 + \|D_t V_u\|_{L^2_t(\mathbb{R})}^2. \quad (C.33)$$

By careful differentiation, while using uniform bounds on $f$ and its derivatives, and as $\mu \leq 1$ and $\sigma \geq 1$ by definition, we get

$$|D_t V_u| \ll \epsilon \mu |u_t| + \epsilon |u_t| + \epsilon \mu (1 - \cos u + |\sin u|)|v_t| + \epsilon \mu (1 - \cos u + |\sin u|)$$

$$\ll \epsilon |u_t| - u_0^0 + \epsilon |u_t| + \epsilon |v_t| - v_0^0 + \epsilon |v_t| - \omega_0(t) + \epsilon (1 - \cos u + |\sin u|) \sigma.$$

Now by weighted integration and applying (10), (11), (1.1), (3.48a) and (3.48b) we obtain

$$\|D_t V_u\|_{L^2_t(\mathbb{R})}^2 \leq c_{36} \epsilon^2 (M^2 + \sigma^2) \lambda(\tau). \quad (C.34)$$

From (C.15), (C.29) and $\lambda(\tau) \ll \sqrt{\epsilon}$ we get

$$\frac{d}{ds} \|u_t - u_t^0\|_{L^2_t(\mathbb{R})}^2 \leq -\frac{1}{2} \|u_{tt} - u_{tt}^0\|_{L^2_t(\mathbb{R})}^2 + c_{36} \epsilon \lambda(\tau) \leq -\frac{1}{4} \|u_{tt} - u_{tt}^0\|_{L^2_t(\mathbb{R})}^2 + c_{36} \epsilon \lambda(\tau). \quad (C.35)$$

Now, as $q \in B_2$, we have that $0 \leq -\|u_t - u_t^0\|_{L^2_t(\mathbb{R})} + c_9 \lambda(\tau)$. Multiplying this with $\epsilon^2$ and (C.35) with $\epsilon$ and summing it, we obtain

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

Write $Z = \|u_{tt}\|_{L^2_t(\mathbb{R})}^2 + \epsilon \|u_t - u_t^0\|_{L^2_t(\mathbb{R})}^2$. Summing (C.33), (C.34) and (C.36) and using $\lambda(\tau)^2 \leq \epsilon/8$, we obtain

$$\frac{d}{ds} Z \leq -\|u_{tt}\|_{L^2_t(\mathbb{R})}^2 - \frac{1}{4} \epsilon Z + c_{36} (M^2 + \sigma^2) \epsilon^2 \lambda(\tau)$$

$$\leq -\frac{1}{4} \epsilon Z + c_{36} \epsilon^2 (M^2 + \sigma^2) \lambda(\tau), \quad (C.37)$$
which by the Gronwall’s lemma and $A$-relative $\xi$-invariance of $B_2$ proves the first part of the claim.

Analogously we obtain
\[
\frac{d}{ds} \| v_{tr} \|^2_{L^2_s(\mathbb{R})} \leq -\| v_{dr} \|^2_{L^2_s(\mathbb{R})} + \| D_s v \|^2_{L^2_s(\mathbb{R})} + 2\| D_s v \|_{L^2_s(\mathbb{R})},
\]
\[
\frac{d}{ds} \| v_{dr} \|^2_{L^2_s(\mathbb{R})} \leq -\| v_{dr} \|^2_{L^2_s(\mathbb{R})} + \frac{1}{16} \| v_{dr} \|^2_{L^2_s(\mathbb{R})} + 2\| D_s v \|_{L^2_s(\mathbb{R})},
\]
and thus, as $\mu \leq 1$,
\[
\| D_s v \|^2_{L^2_s(\mathbb{R})} \leq c_36 \epsilon^2 (M^2 + \sigma^2) \lambda(\tau).
\]
If $\tilde{Z} = \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \epsilon \| v_t - v_t^0 \|^2_{L^2_s(\mathbb{R})}$, we obtain, analogously to (C.37),
\[
\frac{d}{ds} \tilde{Z} \leq -\| v_{tr} \|^2_{L^2_s(\mathbb{R})} - \frac{1}{4} \epsilon \tilde{Z} + c_36 \epsilon^2 (M^2 + \sigma^2) \lambda(\tau)
\]
which completes the proof analogously as for $Z$.

**Lemma C.8.** There exists a constant $c_{11} > 0$ such that the set of all $q \in B_3$ satisfying (3.50a), (3.50b) is an $A$-relative $\xi$-invariant set.

**Proof.** Analogously as in (C.33), (C.38), and as $\lambda(\tau)$, we easily get
\[
\frac{d}{ds} \| v_{dr} \|^2_{L^2_s(\mathbb{R})} \leq -\| v_{dr} \|^2_{L^2_s(\mathbb{R})} + \frac{1}{16} \| v_{dr} \|^2_{L^2_s(\mathbb{R})} + 2\| D_s v \|_{L^2_s(\mathbb{R})},
\]
\[
\frac{d}{ds} \| v_{tr} \|^2_{L^2_s(\mathbb{R})} \leq -\| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \frac{1}{16} \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + 2\| D_s v \|_{L^2_s(\mathbb{R})}.
\]
By the Sobolev inequalities and as $q \in B_2$ and $\lambda(\tau) \leq \sqrt{\epsilon}$, we easily deduce that, for $q \in B_3$,
\[
\| u \|_{L^\infty(\mathbb{R})} \leq (M + \sigma)^\sqrt{\epsilon},
\]
\[
\| v - \omega_{k(\cdot)} \|_{L^\infty(\mathbb{R})} \leq (M + \sigma)^\sqrt{\epsilon}.
\]
Using that and analogously as in the Proof of Lemma 14.7 (the calculation is analogous and routine, thus omitted), we see that
\[
\| D_s v \|^2_{L^2_s(\mathbb{R})} \leq (M^4 + \sigma^4) \epsilon^2 \lambda(\tau),
\]
\[
\| D_s v \|^2_{L^2_s(\mathbb{R})} \leq (M^4 + \sigma^4) \epsilon^2 \lambda(\tau).
\]
Setting
\[
W = \| u_{tr} \|^2_{L^2_s(\mathbb{R})} + \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \epsilon \| u_t - u^0_t \|^2_{L^2_s(\mathbb{R})} = \| u_{tr} \|^2_{L^2_s(\mathbb{R})} + Z,
\]
\[
\tilde{W} = \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \epsilon \| v_t - v^0_t \|^2_{L^2_s(\mathbb{R})} = \| v_{tr} \|^2_{L^2_s(\mathbb{R})} + \tilde{Z},
\]
and summing (C.33), (C.41) and (C.43); respectively (C.38), (C.42) and (C.44), we obtain
\[
\frac{d}{ds} W \leq -\frac{1}{4} \epsilon W + c_{37} \epsilon^2 (M^4 + \sigma^4) \lambda(\tau),
\]
\[
\frac{d}{ds} \tilde{W} \leq -\frac{1}{4} \epsilon \tilde{W} + c_{37} \epsilon^2 (M^4 + \sigma^4) \lambda(\tau),
\]
which by the Gronwall’s lemma and $A$-relative $\xi$-invariance of $B_3$ completes the proof. \qed
D. Appendix. The parity lemma

This appendix is dedicated to the proof of Lemma 3.34 in §3.8. 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 $\tilde{u}(s, t) = 0$ and $(t, \tilde{v}(s, t)) \notin \partial U$.

(ii) For all $s \in [s_0, s_1]$, $u(s, 0) < 0$ and $u(s, t_1) > 0$.

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 $\mathcal{Y}(s) = \{t_0, t_1\} \cup \{t \in [t_0, t_1], (t, v(s, t)) \in \partial U\}$. By assumptions, $\tilde{u}(t, s) \neq 0$ for all $t \in \mathcal{Y}(s)$. We define the relation of equivalence $\sim$ on $\mathcal{Y}(s)$ with $t_1 \sim t_2$ whenever, for all $t_3 \in \mathcal{Y}(s)$ 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 $\tilde{\mathcal{Y}}(s) = \mathcal{Y}(s)/\sim$ with the induced topology. As by assumptions $\mathcal{Y}(s)$ is a closed subset of a compact set, $\tilde{\mathcal{Y}}(s)$ is compact. By (i) and the compactness of $\partial U$, we see that $\tilde{\mathcal{Y}}(s)$ is totally disconnected, thus $\tilde{\mathcal{Y}}(s)$ is finite. By (ii), $|\tilde{\mathcal{Y}}(s)| \geq 2$.

Consider $|\tilde{\mathcal{Y}}(s)| - 1$ segments $(t_k, t_{k+1}) \subset \tilde{\mathcal{Y}}(s)^c$, where $t_k, t_{k+1} \in \mathcal{Y}(s)$ and $t_k \neq t_{k+1}$. 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 $\tilde{U}^c$ (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 $t$, $(t, \tilde{v}(s, t)) \in \tilde{U}^c$).

Definition D.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 $(t_k, t_{k+1}) \subset \tilde{\mathcal{Y}}(s)^c$ for which $t_k, t_{k+1} \in \mathcal{Y}(s)$ and $t_k \neq t_{k+1}$, such that all the zeroes $t \in (t_k, t_{k+1})$ are in $U$.

Now we have the following.

PROPOSITION D.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 $\delta$-neighborhood of $\partial U$, 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 $(t, \tilde{v}(s, t)) \not\in U(\delta_0)$, then $\tilde{u}(s, t) \neq 0$).

If $0 \leq \delta \leq \delta_0$, let $\mathcal{Y}(s, \delta)$ be all $t$ such that $(t, v(s, t)) \in U(\delta)$. Analogously as in the case $\delta = 0$, $\mathcal{Y}(s, 0) = \mathcal{Y}(s)$, we define $\tilde{\mathcal{Y}}(s, \delta)$ which is again finite with cardinality greater than or equal to 2, and $n(s, \delta) \leq |\tilde{\mathcal{Y}}(s, \delta)| - 1$.

Fix $s \in [s_0, s_1]$ for now. Note that $\delta \mapsto |\tilde{\mathcal{Y}}(s, \delta)|, \delta \mapsto n(s, \delta)$ are non-decreasing for $\delta \in [0, \delta_0]$. This follows from the construction and the natural continuous embedding $\mathcal{Y}(s, \delta) \rightarrow \mathcal{Y}(s, \delta')$ for $\delta < \delta'$. Thus $|\tilde{\mathcal{Y}}(s, \delta)|, n(s, \delta)$ strictly increase for at most finitely many times $0 < \delta_1(s) < \cdots < \delta_{k(s)}(s) \in [0, \delta_0]$. Again by construction it is easy to see that these $\delta_j(s), j = 1, \ldots, k(s)$ are characterized as these $\delta \in [0, \delta_0]$ for which at least one entire equivalence class in $\mathcal{Y}(s, \delta)$ lies on the boundary of $U(\delta)$.

Furthermore, for some $\delta_j(s) > 0$, we have that $n(s, \delta_j(s)) - n(s, \delta_j(s)^-) = n(s, \delta_j(s)^+) - n(s, \delta_j(s)^-)$ is even. Indeed, by considering the equivalence classes which are in $\tilde{\mathcal{Y}}(s, \delta_j(s))$, but are not in
Proof of Lemma 3.34. We set 
\[ \bar{\nu}(s) = \nu(s) - (2k + 1)\pi, \quad \bar{\nu}(s, t) = v(s, t), \quad t_0 = \tilde{T}_{k-1}, \quad t_1 = \tilde{T}_{k+1}, \quad \text{and } U = \tilde{N}_k. \]
Now the property (i) follows from the definition of \( A \), and (ii) by Lemma 3.33. Proposition D.1 thus implies Lemma 3.34.

REFERENCES

[1] S. Angenent. A variational interpretation of Mel’nikov’s function and exponentially small separatrix splitting. Symplectic Geometry (London Mathematical Society Lecture Note Series, 192). Cambridge University Press, Cambridge, 1993, pp. 5–35.
[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 and A. Wilkinson. Diffeomorphisms with positive metric entropy. Publ. Math.Inst.
Hautes Études Sci. 124 (2016), 319–347.
[4] P. Bernard. Perturbation d’un hamiltonien partiellement hyperbolique. C. R. Acad. Sci. Paris Sér. I 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. Proceedings of the International Congress of Mathematicians (Hyderabad, India, 2010). Vol. 3. Hindustan Book Agency, New Delhi, 2010, pp. 1680–1700.
[7] P. Bernard, V. Kaloshin and K. Zhang. Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders. Acta Math. 217(1) (2016), 1–79.
[8] M. Berti, L. Biasco and P. Bolle. Drift in phase space: a new variational mechanism with optimal diffusion time. J. Math. Pures Appl. (9) 82 (2003), 613–664.
[9] M. Berti and P. Bolle. A functional analysis approach to Arnold diffusion. Ann. Inst. H. Poncaré Anal. Non Linéaire 19 (2002), 395–450.
[10] M. Berti and P. Bolle. Fast Arnold diffusion in systems with three time scales. Discrete Contin. Dyn. Syst. 8 (2002), 795–811.
[11] U. Bessi. An approach to Arnold diffusion through the calculus of variations. Nonlinear Anal. 26 (1996), 1115–1135.
[12] C. Cheng. Variational construction of diffusion orbits for positive definite Lagrangians. Proceedings of the International Congress of Mathematicians (Hyderabad, India, 2010). Vol. 3. Hindustan Book Agency, New Delhi, 2010, pp. 1714–1728.
[13] C. Cheng and J. Yan. Existence of diffusion orbits in a-priori unstable Hamiltonian systems. J. Differential Geom. 67 (2004), 457–517.
[14] C. Cheng and J. Yan. Arnold diffusion in Hamiltonian systems: a priori unstable case. *J. Differential Geom.* **82** (2009), 229–277.

[15] L. Chierchia and 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 edn. American Mathematical Society, Providence, RI, 2010.

[17] A. Fathi. *Weak KAM Theorem in Lagrangian Dynamics*, book preprint, 5th edn.

[18] T. 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, MR 1860285.

[19] T. 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.

[20] T. Gallay and S. Slijepčević. Distribution of energy and convergence to equilibria in extended dissipative systems. *J. Dynam. Differential Equations* **27** (2015), 653–682.

[21] T. Gallay and S. Slijepčević. Uniform boundedness and long-time asymptotics for the two-dimensional Navier–Stokes equations in an infinite cylinder. *J. Math. Fluid Mech.* **17** (2015), 23–46.

[22] V. Gelfreich and D. Turaev. Arnold diffusion in a priory chaotic Hamiltonian systems. *Preprint*, 2014, arXiv:1406.2945.

[23] D. Henry. *Geometric Theory of Semilinear Parabolic Equations* (Lecture Notes in Mathematics, 840). Springer, Berlin, 1981.

[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. *Preprint*, 1999, arXiv:math/9908014v2.

[26] R. de la Llave. Some recent progress in geometric methods in the instability problem in Hamiltonian mechanics. *Proceedings of the International Congress of Mathematicians* (Zürich, 2006), Vol. 2. European Mathematical Society, Zurich, 2006, pp. 1705–1729.

[27] P. Lochak. Arnold diffusion; a compendium of remarks and questions. *Hamiltonian Systems with Three or More Degrees of Freedom* (NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 533). Kluwer, Dordrecht, 1999, pp. 168–183.

[28] R. S. MacKay, J. D. Meiss and I. C. Percival. Transport in Hamiltonian systems. *Phys. D* **13** (1983), 55–81.

[29] J. Mather. More Denjoy minimal sets for area preserving diffeomorphisms. *Comment. Math. Helv.* **60** (1985), 508–557.

[30] J. Mather. Action minimizing measures for positive definite Lagrangian systems. *Math. Z.* **207** (1991), 169–207.

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

[32] J. Mather. Arnold diffusion by variational methods. *Essays in Mathematics and its Applications* *Ann. Inst. Fourier (Grenoble)*. J. Mather. Arnold diffusion by variational methods.

[33] N. Nekhoroshev. An exponential estimate of the time of stability of Hamiltonian systems close to integrable. *Uspekhi Mat. Nauk.* **32** (1977), 5–66 (in Russian); Engl. Trans. *Russian Math. Surveys* **32** (1977), 1–65.

[34] S. I. Pohožaev. The set of critical values of functionals. *Mat. Sb. (N.S.)* **75** (1968), 106–111.

[35] C. Simo and 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.

[36] S. Slijepčević. Monotone gradient dynamics and Mather’s shadowing. *Nonlinearity* **12** (1999), 969–986.

[37] S. Slijepčević. Extended gradient systems: dimension one. *Discrete Contin. Dyn. Syst.* (2000), 503–518. Zbl 1009.37004, MR 1757384.

[38] S. Slijepčević. A new measure of instability and topological entropy of area-preserving twist diffeomorphisms. *Rad Hrvat. Akad. Znan. Umjet. Mat. Znan.* **21** (2017), 117–141.

[39] A. Sorrentino. Lecture Notes on Mather’s Theory for Lagrangian Systems. *Publ. Mat. Urug.* **16** (2016), 169–192.

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

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

[42] P. Walters. *An Introduction to Ergodic Theory*. Springer, Heidelberg, 2000.

[43] K. Zhang. Speed of Arnold diffusion for analytic Hamiltonian systems. *Invent. Math.* **186** (2011), 255–290.