EXPONENTIAL ESTIMATES OF SYMPLECTIC SLOW MANIFOLDS

K. U. KRISTIANSEN AND C. WULFF

Department of Mathematics,
Technical University of Denmark,
2800 Kgs. Lyngby,
DK
and
Department of Mathematics,
University of Surrey,
Guildford, GU2 7XH
UK

Abstract. In this paper we prove the existence of an almost invariant symplectic slow manifold for analytic Hamiltonian slow-fast systems with finitely many slow degrees of freedom for which the error field is exponentially small. We allow for infinitely many fast degrees of freedom. The method we use is motivated by a paper of MacKay from 2004. The method does not notice resonances, and therefore we do not pose any restrictions on the motion normal to the slow manifold other than it being fast and analytic. We also present a stability result and obtain a generalization of a result of Gelfreich and Lerman on an invariant slow manifold to (finitely) many fast degrees of freedom.

Contents

Date: October 1, 2015.

1. Introduction

Singularly perturbed systems involving different time and/or space scales arise in a wide variety of scientific problems. Important examples include: meteorology and short-term weather forecasting [20, 21, 39], molecular physics and the Born-Oppenheimer approximation [28], chemical enzyme kinetics and the Michaelis-Menten mechanism [5, 29], combustion [23], and the evolution and stability of the solar system [18, 19]. The main advantage of identifying slow and fast variables is dimension reduction by which all the fast variables are "slaved" to the slow ones through the *slow manifold*. In many dissipative systems, the rigorous foundation for a reduction to an attracting, invariant, lower dimensional manifold is provided by Fenichel's theory [6, 7]. In contrast, the conservative cases, exemplified by models of nonlinear waves [2, 21, 20, 25, 40] and of tethered satellites [14, 15, 16], with high frequency oscillations, do not support such a general theory. Nevertheless, such systems may possess *almost invariant slow manifolds*, see [10, 22, 24, 32, 40], on which the angle between the vector field and the tangent space is exponentially small (with respect to the time-scale separation). Orbits of the system may therefore spend a significant amount of time near such a slow manifold.

Dimension reduction is one of the main aims and tools for a dynamicist and the elimination of fast variables is very useful in, for example, numerical computations. In fact, since fast variables require more computational effort and evaluations, this reduction often bridges the gap between tractable and intractable computations. An example of this is the long time (*Gyear*s) integration of the solar system, see [18, 19]. See also [4, 9, 42] for a numerical treatment of slow-fast systems.

In this paper, we prove the existence of an exponentially accurate symplectic slow manifold for a general class of analytic slow-fast Hamiltonian systems. Here we allow the fast sub-system to be a semilinear evolution PDE. We also present a stability result, and obtain a generalization of a result in [11] on an invariant slow manifold to (finitely) many fast degrees of freedom.

1.1. Slow-fast systems.

Consider the system

$$\dot{w} = \epsilon W(w, z), \quad \dot{z} = Z(w, z), \tag{1.1}$$

where $\dot{()} = \frac{\mathrm{d}}{\mathrm{d}t}$ and with ϵ a small parameter. The analytic vector-fields W and Z will in general also depend upon ϵ, but we shall suppress this dependency throughout.

Taylor-expanding the second equation of (1.1) around $z = 0$ gives

$$\dot{w} = \epsilon W(w, z), \quad \dot{z} = r(w) + A(w)z + F(w, z), \tag{1.2}$$

where F is quadratic in z. Note that if $r \equiv 0$ then the manifold $\mathcal{M} = \{z = 0\}$ is invariant. Therefore $\mathcal{M} = \{z = 0\}$ is called a slow manifold in the sense of MacKay [24, Definition 1] if $\|r\|$ is small and $\|(\partial Z/\partial z)(w, 0)^{-1}\| = \|A(w)^{-1}\|$ exists. This ensures that $\mathcal{M}$ is close to being invariant, that the normal motion is truly fast, and that we can solve the equation $\dot{z} = 0$ for $z = \zeta(w)$. The latter property allows us to compute an approximately invariant slow manifold. In particular, if $A(w)$ is elliptic or hyperbolic then $\mathcal{M}$ is said to be normally elliptic respectively hyperbolic. One of the main tasks in singular perturbation theory is to determine the fate of $\mathcal{M}$

for $\epsilon > 0$ but small. When $\mathcal{M}$ is normally hyperbolic then there exists a perturbed, invariant slow manifold $\mathcal{M}(\epsilon)$ for $\epsilon \neq 0$ nearby [6, 7]. The invariance condition is, however, too much to aim for in the normally elliptic setting because normally elliptic manifolds, are unlikely to persist under typical perturbations [24, 26]. However the normally elliptic setting is an important case in Hamiltonian systems which are our main point of interest in this paper. In this case, the important question is how to improve the accuracy of approximately invariant slow manifolds and to determine how long orbits of the full system remain close to those of the approximate dynamics on the slow manifold.

1.2. **Improving a slow manifold.** MacKay, still in [24], presented a method for improving a given slow manifold and he sketched how, when applied successively to an analytic system of differential equations, this could lead to an exponentially accurate slow manifold $\mathcal{M}$ in the sense that the error field is $\|r\| = \mathcal{O}(e^{-c/\epsilon})$. The method of MacKay for improving a slow manifold does not separate normally hyperbolic from normally elliptic slow manifolds. Moreover MacKay allowed for an unbounded fast vector field and argued that one of the advantages of his method is that the improved slow manifold contained all nearby equilibria and conjectured that the error in fact behaved like $\mathcal{O}((\epsilon\|W(w,\cdot)\|)^n)$, for any n, for analytic systems eventually leading to $\mathcal{O}(e^{-c/(\epsilon\|W(w,\cdot)\|)})$, with W vanishing on equilibria. In the appendix we provide a simple counter example (Example A.4) to this last statement for $n > 1$.

The method of MacKay was previously presented by Fraser and Roussel in [8, 37] in a slightly different form. In the reduction method literature it is sometimes therefore also referred to as the iterative method of Fraser and Roussel, see e.g., [13].

Moreover, it seems that it was unknown to MacKay that Neishtadt in [32, Lemma 1] had already proven exponential estimates of the form $\mathcal{O}(e^{-c/\epsilon})$. Neishtadt did, however, not address equilibria and the method he used is (slightly) different from MacKay's. Notably it does not contain all equilibria near the slow manifold. Moreover, Neishtadt considered finite dimensional systems and did not consider the Hamiltonian case.

The method of MacKay can be sketched as follows: Setting $\dot{z} = 0$ in (1.1) gives, by applying the implicit function theorem (bearing in mind that F is quadratic in z), a solution $z = \zeta(w)$ close to $-A(w)^{-1}r(w)$ provided that r is sufficiently small. The graph $z = \zeta(w)$ will be the improved slow manifold (see Lemma A.2 below for details). To show that this is indeed an improved slow manifold, one straightens out the new slow manifold by introducing z_1 through $z = z_1 + \zeta$. Then the equations become

$$\dot{w} = \epsilon W_1(w, z_1), \quad \dot{z}_1 = Z_1(w, z_1) = r_1(w) + A_1(w)z_1 + F_1(w, z_1),$$

with

$$r_1(w) = -\epsilon\partial_w\zeta(w)W(w, \zeta(w)). \tag{1.3}$$

Here ∂_w is used to denote the (Frechet) partial derivatives $\frac{\partial}{\partial w}$, and we will continue to use this symbol regardless of what object is being differentiated. Since $\zeta = \mathcal{O}(\|r\|)$ the error vector field has been diminished by a factor $\mathcal{O}(\epsilon)$. Hence $\mathcal{M}_1 = \{z_1 = 0\}$ is an improved slow manifold. Note that $\mathcal{M}_1$ includes nearby equilibria, c.f. (1.3). Neishtadt in [32, Lemma 1] based his transformations on $z = z_1 + \tilde{\zeta}$,

$\tilde{\zeta}(w) = -A(w)^{-1}r(w)$, giving rise to the error

$$\tilde{r}_1 = -\epsilon\partial_w\tilde{\zeta}(w)W(w,\tilde{\zeta}(w)) + F(w,\tilde{\zeta}(w)).$$

Note that as opposed to (1.3), $\tilde{r}_1$ does not vanish at all nearby equilibria as the error now comes from two separate contributions.

MacKay conjectured that applying this procedure of constrained equilibria, setting $z_i = z_{i+1} + \zeta_i(w)$ with $z = \zeta_i(w)$ solving $Z_i(w,\zeta_i) = 0$, successively would lead to a slow manifold with an error-field of $\mathcal{O}(e^{-c/\epsilon})$. Note that w is never transformed and that only the inverse of the possibly unbounded operator A_i occurs. Therefore unbounded $A(w)$ can also be accounted for. We prove this result in Appendix A. Even though this result has to be attributed to Neishtadt, in contrast to Neishtadt's original result, we (i) follow MacKay and allow for an unbounded fast vector field where the $\dot{z}$ equation is a semilinear evolution equation, and (ii) using MacKay's method we obtain a slow manifold manifold that contains all nearby equilibria.

1.3. **Symplectic slow manifolds.** In this paper we focus on slow-fast Hamiltonian systems of the form (see also [10]):

$$H = H(w,z), \quad w = (u,v) \text{ slow}, \quad z = (x,y) \text{ fast}, \quad J = \operatorname{diag}(\epsilon J_{\mathcal{W}},\, J_{\mathcal{Z}}),$$

where, as opposed to regular perturbation theory, the main contribution to the perturbation comes from the symplectic structure operator J. This gives rise to the following equations of motions

$$\dot{w} = \epsilon J_{\mathcal{W}}\nabla_w H, \quad \dot{z} = J_{\mathcal{Z}}\nabla_z H.$$

We assume that the $\dot{z}$ equation is a semilinear Hamiltonian evolution equation, as detailed later. Normally elliptic slow manifolds are of particular interest in Hamiltonian systems as stability here is associated with oscillatory normal behavior. In Hamiltonian systems there are also generically invariant manifolds that are not normally hyperbolic, for example families of linearly stable periodic orbits parametrized by energy. Here we are interested in approximately invariant symplectic slow manifolds on which we can define a "slow" Hamiltonian system.

The Hamiltonian case has previously been considered by Gelfreich and Lerman in [10]. They restricted to one fast degree of freedom and also obtained an exponentially accurate slow manifold using an averaging method, similar to the one used by Neishtadt in [31].

For the Hamiltonian example

$$H = \frac{1}{2}x^2 + \frac{1}{2}y^2 + v + \epsilon y f(u), \tag{1.4}$$

with analytic $f(u) = \sum_{n=1}^{\infty} \mathrm{e}^{-n}\sin(nu)$ and $\omega = \mathrm{d}x \wedge \mathrm{d}y + \epsilon^{-1}\mathrm{d}u \wedge \mathrm{d}v$, Neishtadt showed that the slow manifold cannot be improved beyond such an estimate, see [10]. The exponential estimate is therefore the best one can aim for in a general non-hyperbolic setting. The method of averaging used in [10] does not extend to several fast variables primarily due to the general lack of control of resonances between the fast variables. On the other hand, it also aims at more than what MacKay and Neishtadt did in [24] and [31]: the results of [10] do not only provide exponential estimates of a slow manifold, they also provide an $\mathcal{O}(1)$-foliation, parametrized by the action variable, of almost invariant slow manifolds. The method therefore also addresses stability, not only existence of an accurate slow manifold. The reference [27] extends the results of [10] to infinite dimensional slow dynamics. The results of [27] hold true for *spatially* Gevrey smooth solutions, which allow for a Galerkin approximation that separates the vector-field into a bounded one and an exponential small remainder. The references [40, 41] also provide exponential estimates of particular slow manifolds in geophysical models by obtaining optimal truncations of the "super-balance equation" (invariance equation) of Lorenz [21].

1.4. **Improving symplectic slow manifolds.** MacKay, still in [24], suggested a separate method for improving slow manifolds in Hamiltonian systems. The proposed method was described as follows: Consider a Hamiltonian $H = H(p)$ with symplectic form ω and a slow manifold $\mathcal{M}_0$. Do the following:

- Compute an orthogonal symplectic foliation $\mathcal{F}_p$ so that for every $p \in \mathcal{M}_0$

$$\omega(p_1, p_2) = 0,\ p_1 \in \mathcal{T}_p\mathcal{F}_p,\ p_2 \in \mathcal{T}_p\mathcal{M}_0.$$

  Here $\mathcal{T}_p\mathcal{M}_0$ denotes the tangent space of $\mathcal{M}_0$ at p.
- Let $H_p = H|_{\mathcal{F}_p}$ and solve this for a nearby critical point $p_1 = p_1(p)$.
- Put $\mathcal{M}_1 = \{p_1(p)\}$ and $\omega_{\mathcal{M}_1} = \omega|_{\mathcal{M}_1}$.

Then $(H|_{\mathcal{M}_1}, \omega_{\mathcal{M}_1})$ is an improved slow system. For the further details see [24]. However, we believe that this method has some drawbacks. First of all, the method requires the computation of a new slow symplectic form at each step. In fact, we believe that the reason for suggesting an alternative to the general approach in the first place, is that one wishes to introduce transformations that preserve the symplectic structure. Moreover, MacKay's method also requires the computation of orthogonal symplectic foliations at each step.

We will therefore suggest an alternative method that circumvents these issues. Our method is then a symplectic extension of MacKay's general approach outlined above. We will at each step straighten out the improved manifold given as the solution $z = \zeta(w)$ of

$$J_{\mathcal{Z}}^{-1} Z(w, z) = \nabla_z H(w, z) = 0,$$

ensuring that the transformation involved in this procedure is symplectic. The slow symplectic form with symplectic structure matrix $\epsilon^{-1} J_{\mathcal{W}}^{-1}$ will therefore remain constant throughout the iteration. For analytic Hamiltonian systems where the fast system is a semilinear evolution equation we obtain a symplectic slow manifold with exponentially small error field containing an initially nearby equilibrium, as was also conjectured by MacKay. This is the main result of this paper. We will present this formally in Theorem 2.1 which we prove in Section 3. In Section 2.3 we also state and prove a stability result, see Corollary 2.2. As opposed to the general case in Appendix A, the symplectic nature of the problem requires us to transform the slow variables. Note that Lu [22] modified our approach, presented in a previous preprint version of this paper, and applied it to a more specific problem where the dynamics on the fastest scale is linear, see Remark 3.11 for more details. In Section 4 we present a dynamical consequence of this result on the persistence of a one degree of freedom slow manifold with exponentially small gaps.

2. Exponential estimates for symplectic slow manifolds

In this section we first introduce some notation (Section 2.1). Then, in Section 2.2, we introduce the setting we work in, in particular our assumptions on the fast Hamiltonian semilinear evolution equation. In Section 2.3 we present our main result and in Section 2.4 we consider two examples where the fast dynamics is a a nonlinear Schrödinger equation and a semilinear wave equation respectively.

2.1. **Some notation.** Let $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ and let $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ be real Banach spaces and $\mathcal{W}^{\mathbb{C}} = \mathcal{W} \oplus \mathrm{i}\mathcal{W}$ and let $\mathcal{Z}^{\mathbb{C}} = \mathcal{Z} \oplus \mathrm{i}\mathcal{Z}$, respectively, be their complexifications with norms $\|w_1 + \mathrm{i}w_2\|_{\mathcal{W}^{\mathbb{C}}} = (\|w_1\|_{\mathcal{W}}^2 + \|w_2\|_{\mathcal{W}}^2)^{1/2}$ and $\|z_1 + \mathrm{i}z_2\|_{\mathcal{Z}^{\mathbb{C}}} = (\|z_1\|_{\mathcal{Z}}^2 + \|z_2\|_{\mathcal{Z}}^2)^{1/2}$. Then $f : \mathcal{V}^{\mathbb{C}} \to \mathcal{Z}^{\mathbb{C}}$, with $\mathcal{V}^{\mathbb{C}}$ an open subset of $\mathcal{W}^{\mathbb{C}}$, is analytic if it is continuously differentiable, i.e., if there exists a continuous derivative $\partial_w f : \mathcal{V}^{\mathbb{C}} \to \mathcal{E}(\mathcal{W}^{\mathbb{C}}; \mathcal{Z}^{\mathbb{C}})$, where $\mathcal{E}(\mathcal{W}^{\mathbb{C}}; \mathcal{Z}^{\mathbb{C}})$ is the Banach space of complex linear operators from $\mathcal{W}^{\mathbb{C}}$ to $\mathcal{Z}^{\mathbb{C}}$ equipped with the operator norm, satisfying the following condition

$$\|f(w+h) - f(w) - \partial_w f(w)(h)\| = \mathcal{O}(\|h\|^2).$$

For later purpose we define $\mathcal{E}(\mathcal{Z}) := \mathcal{E}(\mathcal{Z}; \mathcal{Z})$. By a real analytic function we mean an analytic function which is real valued when its arguments are real. The higher order derivatives can be defined inductively and $\partial_w^n f$ becomes a map

$$\partial_w^n f : \mathcal{V}^{\mathbb{C}} \to \mathcal{E}^n(\mathcal{W}^{\mathbb{C}}; \mathcal{Z}^{\mathbb{C}}),$$

from $\mathcal{V}^{\mathbb{C}}$ into the Banach space $\mathcal{E}^n(\mathcal{W}^{\mathbb{C}}; \mathcal{Z}^{\mathbb{C}})$ of all bounded, n-linear maps from $\mathcal{W}^{\mathbb{C}} \times \cdots \times \mathcal{W}^{\mathbb{C}}$ (n times) into $\mathcal{Z}^{\mathbb{C}}$. See, for example, [33, Appendix A] for a reference on analytic function theory in Banach spaces. When $\mathcal{V}$ is an open subset of $\mathcal{W}$ and $\nu > 0$ then, as in [10], we define $\mathcal{V} + \mathrm{i}\nu$ to be the open complex ν-neighborhood of $\mathcal{V}$:

$$\mathcal{V} + \mathrm{i}\nu = \{w \in \mathcal{W}^{\mathbb{C}} \mid \mathrm{dist}_{\mathcal{W}^{\mathbb{C}}}(w, \mathcal{V}) < \nu\},$$

where $\mathrm{dist}_{\mathcal{W}^{\mathbb{C}}}$ is the metric induced from the Banach norm $\|\cdot\|_{\mathcal{W}^{\mathbb{C}}}$. In the following let $\mathcal{B}_r^{\mathcal{Z}^{\mathbb{C}}}(z) = \{u \in \mathcal{Z}^{\mathbb{C}}, \|u - z\| < r\}$ denote a $\mathcal{Z}^{\mathbb{C}}$-open ball of radius $r > 0$ around z in the Banach space $\mathcal{Z}^{\mathbb{C}}$.

2.2. **Setting and assumptions.** We consider a slow-fast Hamiltonian system with possibly infinitely many fast degrees of freedom of the form

$$\dot{w}_0 = \epsilon J_{\mathcal{W}} \nabla_{w_0} H_0(w_0, z_0), \quad \dot{z}_0 = J_{\mathcal{Z}} \nabla_{z_0} H_0(w_0, z_0) = J_{\mathcal{Z}} L z_0 + B_0(w_0, z_0). \quad (2.1)$$

Here $w_0 = (u_0, v_0) \in \mathcal{W}$, $\dim \mathcal{W} = 2d_{\mathcal{W}}$, are the slow variables and $z = (x, y) \in \mathcal{Z}$ are the fast variables, $\mathcal{Z}$ is a real Hilbert space, $\dim \mathcal{Z} = 2d_{\mathcal{Z}}$ and we allow for $d_{\mathcal{Z}} = \infty$. We use a subscript on H_0 and B_0 because we will transform this system into a more desirable form in the main theorem (Theorem 2.1). The symplectic structure operator is $J = \mathrm{diag}\,(\epsilon J_{\mathcal{W}}, J_{\mathcal{Z}})$ where $J_{\mathcal{W}}$ is the standard symplectic matrix on $\mathcal{W} = \mathbb{R}^{2d_{\mathcal{W}}}$, i.e.

$$J_{\mathcal{W}} = \begin{pmatrix} 0 & \mathrm{id} \\ -\mathrm{id} & 0 \end{pmatrix} \in \mathbb{R}^{2d_{\mathcal{W}} \times 2d_{\mathcal{W}}}$$

where id is the identity on $\mathbb{R}^{d_{\mathcal{W}}}$. We assume

(H0) $J_{\mathcal{Z}}$ and $J_{\mathcal{Z}} L$ are densely defined, closed, skew-symmetric invertible linear operators on $\mathcal{Z}$ and L is a bounded, self-adjoint operator.

Note that (H0) implies that $J_{\mathcal{Z}}$ and L commute.

In the following let $\mathcal{V} \subset \mathcal{W}$ be open and let $\mathcal{S} \subseteq \mathcal{Z}$ be an open neighbourhood of 0, and let $w \in \mathcal{V} + \mathrm{i}\nu_0$, $z \in \mathcal{S} + \mathrm{i}\sigma_0$ where $\nu_0, \sigma_0 > 0$. We assume that $\mathcal{Z}$ is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and that $B_0 : (\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \to \mathcal{Z}$ is analytic, as detailed in (H1) below. In this setting, by semigroup theory [35], the flow of (2.1) is well-defined. The Hamiltonian of (2.1) is

$$H_0(z_0, w_0) = \frac{1}{2}\langle Lz_0, z_0\rangle + V_0(z_0, w_0), \quad \text{where} \quad B_0 = J_{\mathcal{Z}}\nabla_{z_0}V_0. \tag{2.2}$$

Then $H_0 : (\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \to \mathbb{C}$ is analytic. We also allow H_0 to depend continuously on ϵ, but this dependence will be suppressed in the notation.

We define $\mathcal{Z}_1 := D(J_{\mathcal{Z}})$. For initial data in $\mathcal{Z}_1$ the solution of (2.1) is differentiable in time in $\mathcal{Z}$ and so a classical solution in $\mathcal{Z}$, see [35]. The space $\mathcal{Z}_1$ is a real Hilbert space with inner product

$$\langle\langle z_1, z_2\rangle\rangle = \langle J_{\mathcal{Z}} z_1, J_{\mathcal{Z}} z_2\rangle$$

and norm $|||\cdot|||$ which is stronger than $\|\cdot\|$, i.e.,

$$\|z\| \le |||z||| \quad \forall z \in \mathcal{Z}_1. \tag{2.3}$$

Here we assume without loss of generality

$$c_J := \|J_{\mathcal{Z}}^{-1}\|_{\mathcal{E}(\mathcal{Z})} \le 1.$$

(If $c_J > 1$ then we change the inner product on $\mathcal{Z}$ to $c_J\langle\cdot,\cdot\rangle$ and $J_{\mathcal{Z}}$ to $c_J^{-1}J_{\mathcal{Z}}$, ∇_z to $c_J\nabla_z$ to achieve $\|J_{\mathcal{Z}}^{-1}\|_{\mathcal{E}(\mathcal{Z})} \le 1$.) Moreover by definition $J_{\mathcal{Z}}^{-1}$ maps $\mathcal{Z}$ onto $\mathcal{Z}_1$. Note that $B_0 : \mathcal{W} \times \mathcal{Z} \to \mathcal{Z}$ and $\nabla_z V_0 = J_{\mathcal{Z}}^{-1}B_0$ imply that

$$\nabla_z V_0 : \mathcal{W} \times \mathcal{Z} \to \mathcal{Z}_1. \tag{2.4}$$

Taylor-expanding V_0 around $z = 0$ then gives

$$H_0(w_0, z_0) = h_0(w_0) + \langle r_0(w_0), z_0\rangle + \frac{1}{2}\langle (L + a_0(w_0))z_0, z_0\rangle + f_0(w_0, z_0), \tag{2.5}$$

where $f_0 = \mathcal{O}(\|z_0\|^3)$ and $(w_0, z_0) \in \mathcal{W} \times \mathcal{Z}$. Then (2.4) implies that

$$r_0 : \mathcal{W} \to \mathcal{Z}_1, \tag{2.6}$$

and

$$a_0 : \mathcal{W} \to \mathcal{E}(\mathcal{Z}; \mathcal{Z}_1), \quad F_0 = \nabla_z f_0 : \mathcal{W} \times \mathcal{Z} \to \mathcal{Z}_1.$$

We then consider Hamiltonians of the form (2.5) and assume the following:

(H1) The Hamiltonian H_0 is real analytic and uniformly bounded on $(w_0, z_0) \in (\mathcal{V}+\mathrm{i}\nu_0)\times(\mathcal{S}+\mathrm{i}\sigma_0)$, where $\mathcal{V} \subset \mathcal{W}$ is open, $\mathcal{S} \subset \mathcal{Z}$ is an open neighbourhood of 0 in $\mathcal{Z}$ and $\sigma_0, \nu_0 > 0$.

(H2) The map

$$a_0 : \mathcal{V} + \mathrm{i}\nu_0 \to \mathcal{E}(\mathcal{Z}; \mathcal{Z}_1)$$

is real analytic. Moreover,

$$\|a_0|||_{\nu_0} := \sup_{w_0 \in \mathcal{V}+\mathrm{i}\nu_0} \|a_0(w_0)\|_{\mathcal{E}(\mathcal{Z};\mathcal{Z}_1)} \le C_{a_0},$$

and

$$|||A_0(w_0)^{-1}|||_{\nu_0} := \sup_{w_0 \in \mathcal{V}+\mathrm{i}\nu_0} \|A_0(w_0)^{-1}\|_{\mathcal{E}(\mathcal{Z}_1)} \le K_0/2,$$

where

$$A_0(w) := L + a_0(w).$$

(H3) The functions

$$r_0 = \nabla_z H_0|_{z_0=0} : (\mathcal{V} + \mathrm{i}\nu_0) \to \mathcal{Z}_1^{\mathbb{C}},$$
$$h_0 = H_0|_{z_0=0} : (\mathcal{V} + \mathrm{i}\nu_0) \to \mathbb{C},$$

and f_0 with

$$\nabla f_0 = F_0 : (\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \to \mathcal{Z}_1^{\mathbb{C}},$$

are real analytic and uniformly bounded:

$$\begin{aligned} |||r_0|||_{\nu_0} &\leq \delta_0 < \infty \\ \|h_0\|_{\nu_0} &\leq C_{h_0} < \infty, \\ \|F_0|||_{\nu_0,\sigma_0} &\leq C_{F_0} < \infty, \\ \|f_0\|_{\nu_0,\sigma_0} &\leq C_{f_0} < \infty. \\ \|\partial_w h_0\|_{\nu_0} &\leq C'_{h_0} < \infty. \end{aligned}$$

Here we denote by $\|\cdot\|_{\nu_0,\sigma_0}$ the sup-norm taking over the domain $(\mathcal{V}+\mathrm{i}\nu_0)\times(\mathcal{S}+\mathrm{i}\sigma_0)$. Moreover we define $\|\cdot|||_{\nu_0,\sigma_0}$ to be the the sup-norm taking over the domain $(\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \subset \mathcal{W}^{\mathbb{C}} \times \mathcal{Z}^{\mathbb{C}}$ in the $\mathcal{Z}_1^{\mathbb{C}}$ norm.

Note that r_0 is measured in the $\mathcal{Z}_1$-norm $|||\cdot|||$ because of (2.6).

2.3. **Main result.** We are now ready to formulate our main result:

Theorem 2.1. *Assume (H0)-(H3) and let $\nu_0 > \nu > 0$, $\sigma_0 > \sigma > 0$. Then for $\delta_0 > 0$ and $\epsilon > 0$ sufficiently small the following holds true: There exists a symplectic transformation $(w, z) \mapsto (w_0, z_0)$, $(w, z) \in (\mathcal{V} + \mathrm{i}\nu) \times (\mathcal{S} + \mathrm{i}\sigma)$ with $|||z - z_0|||_{\nu,\sigma}$, $\|w - w_0\|_{\nu,\sigma} = \mathcal{O}(\delta_0)$ transforming* (2.2) *into*

$$H(w, z) = H_0(w_0, z_0) = h(w) + \langle r(w), z\rangle + \frac{1}{2}\langle (L + a(w))z, z\rangle + f(w, z), \tag{2.7}$$

with

$$\|h - h_0\|_{\nu}, \|a - a_0|||_{\nu}, \|f - f_0\|_{\nu,\sigma}, \|F - F_0|||_{\nu,\sigma} = \mathcal{O}(\delta_0)$$

and where

$$|||r|||_{\nu} \leq C_1 e^{-C_2/\epsilon}.$$

Here C_1 and C_2 are positive constants that depend solely on C_{a_0}, C'_{h_0}, K_0, C_{F_0}, $C_{\mathcal{S}}$, C_H, σ_0, ν_0, σ, ν, where $C_{\mathcal{S}} = \|z_0\|_{\sigma_0}$.

In other words: $\{z = 0\}$ is an almost invariant symplectic slow manifold. Note that the flow corresponding to the transformed Hamiltonian H is well-defined because the $\dot{z}$ equation is again a semlinear evolution equation of the form considered in [35].

Next we address the stability of the slow manifold:

Corollary 2.2. *Under the assumptions of Theorem 2.1 consider the transformed Hamiltonian* (2.7) *on the real domain $\mathcal{V} \times \mathcal{S}$. If $A_0(w) = L + a_0(w)$ is positive definite then*

$$\ell(w, z) = \frac{1}{2}\langle z, A(w)z\rangle + f(w, z),$$

is an approximate Lyapunov function for $\|z\|$ and ϵ sufficiently small, and there exist constants c_1 and c_2 so that

$$\|z(t)\| \leq \mathcal{O}(e^{-c_1/\epsilon}) \quad \text{for} \quad 0 \leq t \leq c_2\epsilon^{-2},$$

when $z(0) = 0$, provided that $w(t) \in \mathcal{V}$ for $0 \leq t \leq c_2\epsilon^{-2}$.

Proof. By assumption $A_0(w)$ is positive definite and hence, since by Theorem 2.1 we have $\|a - a_0\||_\nu = \mathcal{O}(\delta_0)$, so is $A(w)$ for δ_0 small. Therefore there exist constants $\lambda_1 > 0$ and $\lambda_2 > 0$ so that

$$\lambda_1\|z\|^2 \leq \ell(w, z) \leq \lambda_2\|z\|^2 \tag{2.8}$$

for $\|z\|$ small and all $w \in \mathcal{V} + \mathrm{i}\nu$. Differentiating $\ell(w, z)$ in t we then obtain

$$\begin{aligned}
\dot{\ell}(w(t), z(t)) &= \partial_z\ell J_{\mathcal{Z}}(r + Lz + az + \nabla_z f) + \epsilon\partial_w\ell J_{\mathcal{W}}\nabla_w H \\
&= \langle z + a(w)z + \nabla_z f, J_{\mathcal{Z}}(r + z + a(w)z + \nabla_z f)\rangle + \epsilon\partial_w\ell J_{\mathcal{W}}\nabla_w H \\
&= \langle z + a(w)z + \nabla_z f, J_{\mathcal{Z}} r\rangle + \epsilon\partial_w\ell J_{\mathcal{W}}\nabla_w H \\
&\leq C_3 e^{-C_2/\epsilon} + C_4\epsilon \sup_{w\in\mathcal{V}+\mathrm{i}\nu} \ell(w, z(t)),
\end{aligned}$$

for some constants C_3 and C_4 as long as $(w(t), z(t)) \in \mathcal{V} \times \mathcal{S}$. Here we have used that $\langle z_1, J_{\mathcal{Z}} z_1\rangle = 0$ for all z_1 and a Cauchy estimate on $\partial_w\ell$ and $\partial_w H$. Note that $\dot{\ell}(w(t), z(t))$ is defined at all $z(t) \in \mathcal{S}$. Integrating this inequality from $s = 0$ to t and using (2.8) we find that any initial data $z(0) = 0$

$$\begin{aligned}
\lambda_1\|z(t)\|^2 \leq \ell(w(t), z(t)) &\leq C_3 t\mathrm{e}^{-C_2/\epsilon} + C_4\epsilon \int_0^t \sup_{w\in\mathcal{V}+\mathrm{i}\nu} \ell(w, z(s))\mathrm{d}s \\
&\leq C_3 t e^{-C_2/\epsilon} + C_4\lambda_2\epsilon \int_0^t \|z(s)\|^2\mathrm{d}s.
\end{aligned}$$

We have here used that $\ell(w(t), z(t))|_{t=0} = 0$ since $z(0) = 0$ by assumption. Then by Gronwall's inequality in integral form [3] we obtain

$$\|z(t)\|^2 \leq C_3\lambda_1{}^{-1} t e^{-C_2/\epsilon} e^{C_4\lambda_2\lambda_1{}^{-1}\epsilon t},$$

and therefore while $0 \leq t \leq C_2\lambda_1/(2C_4\lambda_2\epsilon^2) = c_2/\epsilon^2$ we have

$$\|z(t)\| \leq \sqrt{\frac{C_3 C_2}{2C_4\lambda_2}} e^{-C_2/(4\epsilon)}/\epsilon = \mathcal{O}(\mathrm{e}^{-c_1/\epsilon}),$$

where $c_1 < C_2/4$, completing the proof. □

Note that this upper estimate $\mathcal{O}(\epsilon^{-2})$ on the time interval is large, even on the fast time scale $\tau = t/\epsilon$, where it is $\mathcal{O}(\epsilon^{-1})$.

2.4. **Examples.** In this section we present our two main examples. We consider PDEs with periodic boundary conditions, i.e., on the circle $\mathbb{S}^1 \simeq \mathbb{R}/(2\pi\mathbb{Z})$. We frequently use the Hilbert space $\mathcal{L}^2(\mathbb{S}^1; \mathbb{C}^d)$ of square integrable functions with inner product

$$\langle x_1, x_2\rangle_{\mathcal{L}^2(\mathbb{S}^1,\mathbb{C}^d)} = \int_0^{2\pi} x_1(s) \cdot \overline{x_2(s)}\mathrm{d}s$$

and the Sobolev spaces $\mathcal{H}_k(\mathbb{S}^1; \mathbb{C}^d)$ as the spaces containing the $\mathcal{L}^2(\mathbb{S}^1; \mathbb{C}^d)$ functions with k weak derivatives in the $\mathcal{L}^2(\mathbb{S}^1; \mathbb{C}^d)$ inner product

$$\langle x_1, x_2\rangle_{\mathcal{H}_k(\mathbb{S}^1;\mathbb{C}^d)} = \langle(1 - \partial_s^2)^k x_1, x_2\rangle_{\mathcal{L}^2(\mathbb{S}^1;\mathbb{C}^d)}.$$

The examples are also used later in Remark 3.6 to exemplify an abstract construction related to the introduction of an generating function.

2.4.1. *Nonlinear Schrödinger equation.* Consider the nonlinear Schrödinger equation

$$\mathrm{i}\,\partial_t z = \partial_s^2 z - z + \partial_{\overline{z}} U(w,z,\overline{z}), \tag{2.9}$$

on the circle $\mathbb{S}^1$ coupled to a slow system

$$\dot{w} = \epsilon J\left(\nabla_w h(w) + \int_0^{2\pi} \nabla_w U(w,z,\overline{z})\mathrm{d}s\right),$$

where $h:\mathcal{V}+\mathrm{i}\nu_0 \to \mathbb{C}$ is real analytic with $\nu_0>0$ and $\mathcal{V}\subset\mathcal{W}$ open. Furthermore, using the real coordinates $(x,y)\in\mathbb{C}$, where we identify $\mathbb{R}^2\simeq\mathbb{C}$ via $z=x+\mathrm{i}y=(x,y)$, we assume that $\mathcal{U}(w,x,y)=U(w,x+\mathrm{i}y,x-\mathrm{i}y)$ is analytic in $w\in\mathcal{V}+\mathrm{i}\nu_0$ and $x=\Re z$ and $y=\Im z$.

In the coordinates (x,y) the nonlinear Schrödinger equation (2.9) takes the form

$$\dot{x} = \partial_s^2 y - y + \tfrac{1}{2}\nabla_y \mathcal{U}(w,x,y), \quad \dot{y} = -\partial_s^2 x + x - \tfrac{1}{2}\nabla_x \mathcal{U}(w,x,y). \tag{2.10}$$

This follows from the fact that

$$\partial_{\bar{z}} U(w,z,\bar{z}) = \tfrac{1}{2}\partial_x \mathcal{U}(w,x,y) + \tfrac{\mathrm{i}}{2}\partial_y \mathcal{U}(w,x,y).$$

We can rewrite (2.10) as

$$\begin{pmatrix}\dot{x}\\ \dot{y}\end{pmatrix} = J\nabla H(w,x,y)$$

where $\nabla = \nabla^{\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)}$ is the gradient w.r.t. the $\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)$ pairing,

$$J = J_{\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)} = \begin{pmatrix} 0 & \mathrm{id}_{\mathcal{L}^2}\\ -\mathrm{id}_{\mathcal{L}^2} & 0\end{pmatrix},$$

is the standard symplectic structure matrix and

$$H(x,y,w) = \frac{1}{2}\int_{\mathbb{S}^1}\left(|\partial_s x|^2 + |\partial_s y|^2 + x^2 + y^2 + \mathcal{U}(w,x,y)\right)\mathrm{d}s + h(w). \tag{2.11}$$

Note that the Hamiltonian H is well defined on $\mathcal{Z} := \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}^2)$, but not on $\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)$. Defining

$$\langle J_{\mathcal{Z}}^{-1} z_1, z_2\rangle_{\mathcal{Z}} = \int_0^{2\pi}(x_1(s)y_2(s) - y_1(s)x_2(s))\mathrm{d}s = \langle J^{-1}z_1, z_2\rangle_{\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)} \tag{2.12}$$

we see that $J_{\mathcal{Z}} = (1-\partial_x^2)J$ so that $J_{\mathcal{Z}}^{-1}\colon \mathcal{H}_2(\mathbb{S}^1;\mathbb{R}^2)\to\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)$. The Laplacian is diagonal in the Fourier representation with eigenvalues $-k^2$. Hence, $\operatorname{spec} J_{\mathcal{Z}} = \{\pm\mathrm{i}(k^2+1)\colon k\in\mathbb{Z}\}$ so that $J_{\mathcal{Z}}$ generates a unitary group on $\mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)$ and on $\mathcal{Z} = \mathcal{H}_1(\mathbb{S}^1,\mathbb{R}^2)$. This shows that the Hamiltonian of the nonlinear Schrödinger equation (2.9) takes the form (2.2) with

$$L = \mathrm{id}, \quad V(w,x,y) = \frac{1}{2}\int_0^{2\pi}\mathcal{U}(w,x(s),y(s))\mathrm{d}s$$

where $z\in\mathcal{Z}$. The fact that $H(w,z)$ is analytic for $(w,z)\in(\mathcal{W}+\mathrm{i}\nu)\times(\mathcal{S}+\mathrm{i}\sigma)$ can be deduced from the fact that $B=J\nabla U$ is analytic as map from $(\mathcal{W}+\mathrm{i}\nu)\times(\mathcal{S}+\mathrm{i}\sigma)$ to $\mathcal{Z}$ from the fact that $\mathcal{H}_1(\mathbb{S}^1;\mathbb{R})$ is an algebra, see [1, Theorem 5.23]. Moreover

$$a_0(w) = (1-\partial_s^2)^{-1}\partial_z^2\mathcal{U}(w,0).$$

If $\partial_z^2\mathcal{U}(w,0)$ is small then $A(w) = \mathrm{id} + a_0(w)$ is invertible on $\mathcal{Z}_1 = \mathcal{H}_3(\mathbb{S}_1;\mathbb{R}^2)$ as required in (H2). Under this assumption conditions (H0-H3) are satisfied. By choosing $\partial_z \mathcal{U}(w,0)$ sufficiently small we can make δ_0 sufficiently small as required in Theorem 2.1.

2.4.2. *Semilinear wave equation.* Consider a semilinear wave equation of the form

$$\begin{aligned}\dot{x} &= y,\\ \dot{y} &= (\partial_s^2 - 1)x - \nabla_x U(u,x)\end{aligned}$$

with $z = (x,y)$ and where $x = x(t,s)$, $y = y(t,s)$ with $s \in \mathbb{S}^1$, for simplicity. This system is coupled to a set of slow ordinary differential equations for the evolution of $(u,v) = (u,v)(t) \in \mathbb{R}^{2\mathrm{d}_\mathcal{W}}$ of the form

$$\begin{aligned}\dot{u} &= \epsilon v,\\ \dot{v} &= -\epsilon\left(\nabla_u h(u) + \int_0^{2\pi} \nabla_u U(u,x)\mathrm{d}s\right).\end{aligned}$$

This system is Hamiltonian with Hamiltonian and symplectic structure operator given by

$$\begin{aligned}H(w,z) &= \frac{1}{2}|v|^2 + \frac{1}{2}\langle z,z\rangle + h(u) + \int_0^{2\pi} U(u,x)\mathrm{d}s,\\ J_\mathcal{Z} z &= \begin{pmatrix} y \\ \partial_s^2 x - x\end{pmatrix},\end{aligned}$$

Also $L = \mathrm{id}$ and $\mathcal{Z} = \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}) \times \mathcal{L}^2(\mathbb{S}^1;\mathbb{R})$, with

$$\langle z_1, z_2\rangle = \langle x_1, x_2\rangle_{\mathcal{H}_1(\mathbb{S}^1;\mathbb{R})} + \langle y_1, y_2\rangle_{\mathcal{L}^2(\mathbb{S}^1;\mathbb{R})}, \quad z_1 = \begin{pmatrix} x_1 \\ y_1\end{pmatrix},\, z_2 = \begin{pmatrix} x_2 \\ y_2\end{pmatrix}.$$

We assume that $U : \mathbb{R}^{\mathrm{d}_\mathcal{W}+1} \to \mathbb{R}$ is an analytic function. Then U is also analytic as a function from $\mathbb{R}^{d_\mathcal{W}} \times \mathcal{H}_1(\mathbb{S}^1;\mathbb{R})$ to $\mathbb{R}$, which follows from the theory of superposition operators, see [1, Theorem 5.23], and so the system is well-posed on $\mathcal{W} \times \mathcal{Z}$ [35]. Finally

$$D(J_\mathcal{Z}) = \mathcal{H}_2(\mathbb{S}^1;\mathbb{R}) \times \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}).$$

Then we note that due by the definition of $\nabla_z H$ we have for all $\tilde{z} \in \mathcal{Z}$ that

$$\partial_z H(w,z)\tilde{z} = \langle \nabla_z H(w,z), \tilde{z}\rangle$$

and hence, due to

$$\langle \nabla_x U(u,x), \tilde{x}\rangle_{\mathcal{L}^2(\mathbb{S}^1;\mathbb{R})} = \langle (1-\partial_x^2)^{-1}\, \nabla_x U(u,x), \tilde{x}\rangle_{\mathcal{H}_1(\mathbb{S}^1;\mathbb{R})}$$

we get

$$\nabla_z H = \begin{pmatrix} x + (1-\partial_x^2)^{-1}\nabla_x U \\ y\end{pmatrix}.$$

The assumptions on analyticity (H0-H3) are satisfied due to the analyticity assumption on $\mathcal{U}$ provided that $\partial_x^2 U(u,0)$ is sufficiently small.Note that the smoothing property in (H2), (H3) is a consequence of the appearance of the isomorphism $(1-\partial_s^2)^{-1} : \mathcal{L}^2(\mathbb{S}^1;\mathbb{R}) \to \mathcal{H}_2(\mathbb{S}^1;\mathbb{R})$ in the expression above. Moreover, $A_0(w_0)^{-1}$ exists and is bounded as an operator from $\mathcal{Z}_1$ into $\mathcal{Z}_1$ as required in (H2) if $\partial_x^2 U(u,0)$ is small so that $A_0(w_0) = \mathrm{id} + a_0(u_0)$ is a small $\mathcal{Z}_1$-perturbation of the identity. By choosing $\partial_x U(u,0)$ sufficiently small we can make δ_0 small as required in Theorem 2.1.

3. Proof of main result

We start with some preliminary lemmas which will be needed in the proof. Then, in Section 3.2 we introduce the generating functions for the symplectic transformations we consider. In Section 3.3 we prove that those symplectic transformations are well-posed. We then set up an iterative lemma (Section 3.4) which we use to prove the main theorem in Section 3.5.

3.1. **Preliminary lemmas.** We frequently need the following Cauchy estimate [33]:

Lemma 3.1. *Assume that* $\mathcal{W}^{\mathbb{C}}$ *and* $\mathcal{Z}^{\mathbb{C}}$ *are Banach spaces, let* $\mathcal{V}^{\mathbb{C}} \subseteq \mathcal{W}^{\mathbb{C}}$ *be open and assume that* $f : \mathcal{V}^{\mathbb{C}} \to \mathcal{Z}^{\mathbb{C}}$ *is analytic and that* f *is bounded on* $\mathcal{B}_\nu(w_0) \subset \mathcal{V}^{\mathbb{C}}$ *for some* $\nu > 0$. *Then for* $n \in \mathbb{N}$

$$\|\partial_w^n f(w_0)\| \leq n! \frac{\sup_{w\in\mathcal{B}_\nu(w_0)} \|f(w)\|}{\nu^n}. \tag{3.1}$$

Remark 3.2. *Let* $f : \mathcal{V} + \mathrm{i}\nu \to \mathcal{Z}^{\mathbb{C}}$ *be analytic and bounded and let* $\nu > \xi > 0$. *Then we can apply the estimate* (3.1) *to any* $w_0 \in \mathcal{V} + i(\nu - \xi)$ *to obtain:*

$$\sup_{w_0\in\mathcal{V}+\mathrm{i}(\nu-\xi)} \|\partial_w^n f(w_0)\| \leq n! \frac{\sup_{w\in\mathcal{V}+\mathrm{i}\nu} \|f(w)\|}{\xi^n},$$

which we will write compactly as

$$\|\partial_w^n f\|_{\nu-\xi} \leq n! \frac{\|f\|_\nu}{\xi^n}.$$

This is the form of Cauchy's estimate that we will be using.

We will also use the following generalized version of Taylor's theorem [33]:

Lemma 3.3. *If* $f : \mathcal{V}^{\mathbb{C}} \to \mathcal{Z}^{\mathbb{C}}$ *is* n *times continuously differentiable,* $n \geq 1$, *and if the segment* $w + sh$, $0 \leq s \leq 1$, *is contained in* $\mathcal{V}^{\mathbb{C}}$, *then*

$$\begin{aligned} f(w+h) = f(w) + \partial_w f(w)(h) + \cdots + \frac{1}{(n-1)!}\partial_w^{n-1} f(w) h^{n-1} \\ + \int_0^1 \frac{(1-s)^{n-1}}{(n-1)!} \partial_w^n f(w+sh)(h, \cdots, h) \mathrm{d}s. \end{aligned}$$

The integral remainder is bounded by $\frac{\|h\|^n}{n!} \sup_{0\leq s\leq 1} \|\partial_w^n f(w + sh)\|$.

Here, we write for $m \in \mathbb{N}$,

$$\partial_w^m f(w) h^m = \partial_w^m f(w)(h, \ldots, h).$$

To continue we introduce the following notation: Under assumption (H0-H3) for $R > 0$ such that $\mathcal{B}_R^{\mathcal{Z}^{\mathbb{C}}}(0) \subseteq \mathcal{S} + \mathrm{i}\sigma_0$, define

$$C''_{F_0}[\nu_0, R] := \sup_{\substack{w_0\in\mathcal{V}+\mathrm{i}\nu_0,\\ \|z_0\|\leq R}} \|\partial_z^2 F_0(w_0, z_0)\|_{\mathcal{E}(\mathcal{Z}\times\mathcal{Z};\mathcal{Z}_1)}. \tag{3.2}$$

We need the following lemma, which deals with solutions of the equation $\dot{z} = 0$, for the construction of improved slow manifolds:

Lemma 3.4. *Assume (H0-H3) with the subscripts dropped and assume that*
$$\delta < 2/(K^2 C_F''[\nu, K\delta])), \tag{3.3}$$
and that $K\delta < \sigma$. Then
$$0 = r(w) + A(w)z + F(w,z), \tag{3.4}$$
has a locally unique solution $z = \zeta(w) \in \mathcal{Z}_1^{\mathbb{C}}$ satisfying:
$$|||\zeta(w)||| \leq K|||r(w)|||,$$
for every $w \in \mathcal{V} + \mathrm{i}\nu$. Moreover $\zeta(w)$ is analytic in $w \in \mathcal{V} + \mathrm{i}\nu$:
$$\zeta \in C^\omega(\mathcal{V} + \mathrm{i}\nu; \mathcal{Z}_1^{\mathbb{C}}).$$

Proof. This is just a consequence of the uniform contraction theorem, see e.g., [12, Section 1.2.6], where (3.3) is the condition to ensure the contraction property. We include the proof to verify the estimates. Re-arranging (3.4) and applying the inverse $A(w)^{-1}$ gives
$$\zeta(w) = -A(w)^{-1}(r(w) + F(w,\zeta)). \tag{3.5}$$
Put $\tilde{\zeta}(w) = -A(w)^{-1}r(w)$ and write $\zeta = \tilde{\zeta}(w) + z$ so that
$$z = \Pi(w,z) := -A(w)^{-1}F(w,\tilde{\zeta}(w) + z) = -(L + a(w))^{-1}F(w,\tilde{\zeta}(w) + z)$$
for $w \in \mathcal{W} + \mathrm{i}\nu$. Note that $|||\tilde{\zeta}(w)||| \leq \frac{K}{2}|||r(w)|||$ by (H2). We will denote this upper bound by $\rho(w) = \frac{K}{2}|||r(w)|||$ and highlight that $\rho(w) \leq \frac{K}{2}\delta$. We will show that $\Pi(w,\cdot)$ is a contraction on $\mathcal{B}_{\rho(w)}^{\mathcal{Z}_1^{\mathbb{C}}}(0)$ for each $w \in \mathcal{W} + \mathrm{i}\nu$. Note that $\mathcal{B}_{\rho(w)}^{\mathcal{Z}_1^{\mathbb{C}}}(\tilde{\zeta}(w)) \subset \mathcal{S} + \mathrm{i}\sigma$ because $\sigma > K\delta > 2\rho(w)$ by assumption. By Taylor's formula we have for $\|z\| \leq K\delta$ that
$$|||F(w,z)||| = \| \int_0^1 (1-s)\partial_z^2 F(w,sz)z^2 \mathrm{d}s||| \leq \frac{1}{2}C_F''[\nu,K\delta]\|z\|^2.$$
Therefore for $z \in \mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0)$ and $w \in \mathcal{W} + \mathrm{i}\nu$
$$|||\Pi(w,z)||| \leq \frac{K}{4}C_F''[\nu,K\delta]\|\tilde{\zeta}(w) + z\|^2 \leq \frac{K}{4}C_F''[\nu,K\delta](2\rho)^2 \leq K^2 C_F''[\nu,K\delta]\delta\rho/2 < \rho$$
using that $2\rho \leq K\delta$ and assumption (3.3), and hence $\Pi(w,\cdot) : \mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0) \to \mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0)$. Next, by Taylor's formula
$$\partial_z F(w,z) = \int_0^1 \partial_z^2 F(w,tz)z\mathrm{d}t,$$
from which we obtain that for $z \in \mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0)$
$$|||\partial_z F(w,\tilde{\zeta}(w) + z)||| \leq C_F''[\nu,K\delta]\|\tilde{\zeta}(w) + z\| \leq C_F''[\nu,K\delta]K\delta.$$
Therefore for $z \in \mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0)$
$$|||\partial_z \Pi(w,z)||| \leq K^2 C_F''[\nu,K\delta]\delta/2 < 1,$$
using (3.3). This shows that $\Pi(w,\cdot)$ is a contraction on the ball $\mathcal{B}_\rho^{\mathcal{Z}_1^{\mathbb{C}}}(0)$ and therefore there exists a unique fixed point $z(w)$ of $\Pi(w,\cdot)$. In particular, $\zeta(w) = \tilde{\zeta}(w) + z(w)$ solves (3.5) and $|||\zeta(w)||| \leq 2\rho = K|||r(w)|||$. By [12, Section 1.2.6] the map $\zeta : \mathcal{V} + \mathrm{i}\nu \to \mathcal{Z}_1^{\mathbb{C}}$ is analytic. □

3.2. **Generating functions.** The proof of Theorem 2.1 is based on successive symplectic transformations. We will in the following consider the Hamiltonian $H(w,z)$ in place of $H_0(w_0,z_0)$ from (2.2), satisfying the assumptions (H0-H3), with subscripts removed, on $(w,z)\in(\mathcal{V}+\mathrm{i}\nu)\times(\mathcal{S}+\mathrm{i}\sigma)$. As for the general case considered in Appendix A and explained in the introduction we improve the manifold $\mathcal{M}_0=\{z=0\}$ by solving $\dot z=0$ for $z=\zeta(w)$ using Lemma 3.4. We generate a symplectic transformation Ψ from the non-canonical transformation $z=\zeta(w)+z_+$ through a generating function introduced in the following lemma.

In this section we formally define Ψ and show that it is symplectic. In the next section we will show that Ψ is well-defined.

Lemma 3.5. *Assume (H0-H3) with subscript dropped. Then there are projectors $\mathbb{P}_x$, $\mathbb{P}_y$ on $\mathcal{Z}$ such that*

$$\mathbb{P}_x+\mathbb{P}_y=\mathrm{id},\quad \mathbb{P}_x\mathbb{P}_y=0,\quad \mathbb{P}_x\mathcal{Z}\perp\mathbb{P}_y\mathcal{Z},$$

with the following property: Let $z=(x,y)$ with $x=\mathbb{P}_xz$, $y=\mathbb{P}_yz$ and let

$$g(u,v_+,x,y_+)=-\langle J_{\mathcal{Z}}^{-1}\zeta(u,v_+),(x,y_+)\rangle_{\mathcal{Z}}.$$

Then $(u_+,v_+,x_+,y_+)\mapsto\Psi(u_+,v_+,x_+,y_+)=(u,v,x,y)$ formally defines a a symplectic transformation given implicitly by the equations:

$$\begin{aligned} x_+&=x-\zeta^x(u,v_+),\quad y=y_++\zeta^y(u,v_+),\\ u_+&=u+\epsilon\nabla_{v_+}g(u,v_+,x,y_+),\quad v=v_++\epsilon\nabla_u g(u,v_+,x,y_+). \end{aligned}\tag{3.6}$$

Proof. To construct the generating function we first define a Hilbert space $\tilde{\mathcal{Z}}=\tilde{\mathcal{X}}\times\tilde{\mathcal{X}}$ such that (2.2) is Hamiltonian on $\tilde{\mathcal{Z}}$ with respect to the standard symplectic structure matrix

$$J_{\tilde{\mathcal{Z}}}=\begin{pmatrix}0 & \mathrm{id}_{\tilde{\mathcal{X}}}\\ -\mathrm{id}_{\tilde{\mathcal{X}}} & 0\end{pmatrix}.\tag{3.7}$$

We claim that $\tilde{\mathcal{Z}}=\mathcal{Z}_{-1/2}$ is the dual space of $\mathcal{Z}_{1/2}=D(|J_{\mathcal{Z}}|^{1/2})$ w.r.t. the $\mathcal{Z}$ pairing, so that the inner products on $\tilde{\mathcal{Z}}$ and $\mathcal{Z}$ are related as follows:

$$\langle z_1,z_2\rangle_{\mathcal{Z}}=\langle|J_{\mathcal{Z}}|z_1,z_2\rangle_{\tilde{\mathcal{Z}}}.$$

To define $\tilde{\mathcal{X}}$ we proceed as follows: We write

$$J_{\mathcal{Z}}=\int_{\lambda\in\mathrm{spec}(J_{\mathcal{Z}})}\lambda\mathrm{d}\mathbb{P}_\lambda$$

where $\mathrm{d}\mathbb{P}$ is the projection valued spectral measure of $J_{\mathcal{Z}}$, see e.g., [36, Theorem VIII.8], noting that $-\mathrm{i}J_{\mathcal{Z}}$ is self-adjoint on $\mathcal{Z}^{\mathbb{C}}$. Polar decomposition gives

$$J_{\mathcal{Z}}=J_{\tilde{\mathcal{Z}}}|J_{\mathcal{Z}}|.\tag{3.8}$$

Here

$$|J_{\mathcal{Z}}|=\int_{\mathrm{i}\omega\in\mathrm{spec}(J_{\mathcal{Z}})}|\omega|\mathrm{d}\mathbb{P}_{\mathrm{i}\omega}$$

is positive, self-adjoint and densely defined on $\mathcal{Z}$ and all three operators commute. Note that $J_{\tilde{\mathcal{Z}}}$ is by construction both skew-symmetric and unitary. Hence, $\mathrm{spec}(J_{\tilde{\mathcal{Z}}})=\{\mathrm{i},-\mathrm{i}\}$.

We will now first find spaces $\mathcal{X}$ and $\mathcal{Y}\simeq\mathcal{X}$ of $\mathcal{Z}=\mathcal{X}\oplus\mathcal{Y}$ such that $J_{\tilde{\mathcal{Z}}}$, when restricted to $\mathcal{Z}\subseteq\tilde{\mathcal{Z}}$ takes the form (3.7). Then we define $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}\simeq\tilde{\mathcal{X}}$ as dual spaces of $\mathcal{X}_{1/2}=D(|J_{\mathcal{Z}}|^{1/2})\cap\mathcal{X}$ and $\mathcal{Y}_{1/2}=D(|J_{\mathcal{Z}}|^{1/2})\cap\mathcal{Y}$ w.r.t. the $\mathcal{X}$ and $\mathcal{Y}$

inner product respectively and conclude that (3.7) also holds on $\tilde{\mathcal{Z}}$, the closure of $\mathcal{Z}$ in the $\tilde{\mathcal{Z}}$-norm.

To define $\mathcal{X}$, let $\mathbb{P}_\pm$ be the orthonormal spectral projectors onto the eigenspaces $\mathcal{Z}_\pm = \mathbb{P}_\pm \mathcal{Z}$ of $J_{\tilde{\mathcal{Z}}}$ to its eigenvalues $\pm\mathrm{i}$. Then $\mathcal{Z}_+ \simeq \mathcal{Z}_-$ and so there is an isomorphism $\iota : \mathcal{Z}_- \to \mathcal{Z}_+$ which we define as follows: let $\{e_j, j \in I\}$ be an orthonormal basis for $\mathcal{Z}_+$. Then $\{\bar{e}_j, j \in I\}$ is an orthonormal basis for $\mathcal{Z}_-$ and we define $\iota : \mathcal{Z} \to \mathcal{Z}$ by

$$\iota z = \sum_{j\in I} \langle z, e_j\rangle \bar{e}_j + \langle z, \bar{e}_j\rangle e_j. \tag{3.9}$$

Note that $\iota e_k = \bar{e}_k$, $\iota \bar{e}_k = e_k$ so that $\iota^2 = \mathrm{id}$ and $\iota^* = \iota$. (Technically, the isomorphism $\iota : \mathcal{Z}_- \to \mathcal{Z}_+$ above is the restriction of (3.9) to $\mathcal{Z}_-$.) We set $\mathbb{P}_x = \frac{1}{2}(\mathrm{id} + \iota)$ and $\mathbb{P}_y = \frac{1}{2}(\mathrm{id} - \iota)$. We then have $\mathbb{P}_x + \mathbb{P}_y = \mathrm{id}$ and $\mathbb{P}_x\mathbb{P}_y = 0$. Moreover we readily check that $\mathcal{X} \perp \mathcal{Y}$, where $\mathcal{X} = \mathbb{P}_x\mathcal{Z}$ and $\mathcal{Y} = \mathbb{P}_y\mathcal{Z}$. Finally $\mathcal{X} \simeq \mathcal{Y}$ because $\mathrm{i}(\mathbb{P}_+ - \mathbb{P}_-)$ is an isomorphism mapping $\mathcal{X}$ to $\mathcal{Y}$ and $\mathcal{Y}$ to $\mathcal{X}$. This follows from the fact that

$$\mathbb{P}_x \mathrm{i}(\mathbb{P}_+ - \mathbb{P}_-)\mathbb{P}_x = \mathbb{P}_y \mathrm{i}(\mathbb{P}_+ - \mathbb{P}_-)\mathbb{P}_y = 0$$

which is straightforward to check, using that $\iota \mathbb{P}_\pm \iota = \mathbb{P}_\mp$ and that $\iota\mathbb{P}_\pm = \mathbb{P}_\mp \iota$.

It remains to prove that the symplectic structure operator $J_{\mathcal{Z}}$ transforms to the operator $J_{\tilde{\mathcal{Z}}}$ on $\tilde{\mathcal{Z}}$, i.e., that (2.2) is Hamiltonian on $\tilde{\mathcal{Z}}$ with symplectic structure operator $J_{\tilde{\mathcal{Z}}}$. This holds true provided that

$$J_{\tilde{\mathcal{Z}}} \nabla_z^{\tilde{\mathcal{Z}}} = J_{\mathcal{Z}} \nabla_z$$

and this follows from

$$\nabla_z^{\tilde{\mathcal{Z}}} = |J_{\mathcal{Z}}| \nabla_z^{\mathcal{Z}}$$

which can be proved as follows: Let $f : \mathcal{Z} \to \mathbb{R}$ be differentiable so that

$$\mathrm{d}f(z)(\delta z) = \langle \nabla_z^{\mathcal{Z}} f(z), \delta z\rangle_{\mathcal{Z}}.$$

Since $\langle z_1, z_2\rangle_{\mathcal{Z}} = \langle |J_{\mathcal{Z}}| z_1, z_2\rangle_{\tilde{\mathcal{Z}}}$, we have

$$\langle \nabla_z^{\mathcal{Z}} f(z), \delta z\rangle_{\mathcal{Z}} = \langle |J_{\mathcal{Z}}| \nabla_z^{\mathcal{Z}} f(z), \delta z\rangle_{\tilde{\mathcal{Z}}},$$

and therefore

$$\nabla_z^{\tilde{\mathcal{Z}}} f(z) = |J_{\mathcal{Z}}| \nabla_z^{\mathcal{Z}} f(z).$$

The generating function we consider is then

$$G(u, v_+, x, y_+) = \langle x, y_+\rangle_{\tilde{\mathcal{X}}} + \epsilon^{-1}\langle u, v_+\rangle + g(u, v_+, x, y_+),$$

with

$$\begin{aligned} g(u, v_+, x, y_+) &= \langle J_{\tilde{\mathcal{Z}}} \zeta(u, v_+), (x, y_+)\rangle_{\tilde{\mathcal{Z}}} = \langle |J_{\mathcal{Z}}|^{-1} J_{\tilde{\mathcal{Z}}} \zeta(u, v_+), (x, y_+)\rangle_{\mathcal{Z}} \\ &= -\langle J_{\mathcal{Z}}^{-1} \zeta(u, v_+), (x, y_+)\rangle_{\mathcal{Z}}, \end{aligned}$$

using (3.8). The corresponding symplectic transformation defined as

$$x_+ = \nabla_{y_+}^{\tilde{\mathcal{X}}} G, \quad y = \nabla_x^{\tilde{\mathcal{X}}} G, \quad u_+ = \epsilon \nabla_{v_+} G, \quad v = \epsilon \nabla_u G$$

then gives the desired result. □

Remark 3.6. In the following we will illustrate the abstract construction used in Lemma 3.5 by re-considering our two examples from Section 2.4.

In the case of the nonlinear Schrödinger equation, see Section 2.4.1, we have $\mathcal{Z} = \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}^2)$ and $J_{\mathcal{Z}} = (1-\partial_s^2)J$ (with J the standard symplectic structure matrix) so that $\mathcal{Z}_{1/2} = D_{\mathcal{H}_1(\mathbb{S}^1;\mathbb{R}^2)}(|\partial_x|) = \mathcal{H}_2(\mathbb{S}^1;\mathbb{R}^2)$ and $\tilde{\mathcal{Z}} = \mathcal{Z}_{-1/2}$, the dual space of $\mathcal{Z}_{1/2}$ w.r.t. the $\mathcal{Z}$ inner product is $\tilde{\mathcal{Z}} = \mathcal{L}^2(\mathbb{S}^1;\mathbb{R}^2)$. In this case $\mathcal{X} = \mathcal{H}_1(\mathbb{S}^1;\mathbb{R})$ and $\tilde{\mathcal{X}} = \mathcal{L}^2(\mathbb{S}^1;\mathbb{R})$.

In the case of the semilinear wave equation, see Section 2.4.2, we have $\mathcal{Z} = \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}) \times \mathcal{L}^2(\mathbb{S}^1;\mathbb{R})$, $\mathcal{Z}_{1/2} = \mathcal{H}_{1.5}(\mathbb{S}^1;\mathbb{R}) \times \mathcal{H}_{0.5}(\mathbb{S}^1;\mathbb{R})$ and $\tilde{\mathcal{Z}} = \mathcal{H}_{0.5}(\mathbb{S}^1;\mathbb{R}) \times \mathcal{H}_{-0.5}(\mathbb{S}^1;\mathbb{R})$. In this case let

$$e_{\pm k}(s) = \frac{\mathrm{e}^{\mathrm{i}ks}}{2\sqrt{\pi}}\begin{pmatrix} \frac{\pm 1}{\sqrt{k^2+1}} \\ \mathrm{i} \end{pmatrix}, \quad k \in \mathbb{N}_0.$$

Then $\mathcal{Z}_+$ is spanned by e_k and $\mathcal{Z}_-$ by e_{-k}, $k \in \mathbb{N}_0$. Moreover

$$e_k^r(s) = \tfrac{1}{\sqrt{2}}\Re e_k(s) = \frac{1}{\sqrt{2\pi}}\begin{pmatrix} \frac{\cos ks}{\sqrt{k^2+1}} \\ -\sin ks \end{pmatrix}, \quad e_k^i(s) = \tfrac{1}{\sqrt{2}}\Im e_k(s) = \frac{1}{\sqrt{2\pi}}\begin{pmatrix} \frac{\sin ks}{\sqrt{k^2+1}} \\ \cos ks \end{pmatrix},$$

and so one copy of $\mathcal{X}$ has orthonormal basis $\{e_k^r, k \in \mathbb{N}_0\}$, the other, corresponding to $\mathcal{Y}$, has orthonormal basis $\{e_k^i, k \in \mathbb{N}_0\}$, both endowed with the $\mathcal{Z}$ inner product. Moreover the two isomorphic copies of $\tilde{\mathcal{X}}$ are again spanned by e_k^r and e_k^i, $k \in \mathbb{N}_0$, respectively, but now endowed with the $\tilde{\mathcal{Z}}$ inner product. Note that the decomposition $z = (x, y)$ for the semilinear wave equation used in the definition of the symplectic transformation, see (3.6), is such that $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and hence not the natural decomposition where $\mathcal{Z} = \mathcal{H}_1(\mathbb{S}^1;\mathbb{R}) \times \mathcal{L}^2(\mathbb{S}^1;\mathbb{R})$ and $x \in \mathcal{H}_1(\mathbb{S}^1;\mathbb{R})$, $y \in \mathcal{L}^2(\mathbb{S}^1;\mathbb{R})$. In this example we can identify $\mathcal{X}$ with the space of square summable sequences $\ell_2(\mathbb{N}_0;\mathbb{R})$ by identifying each $x \in \mathcal{X}$ with its component vector with respect to the orthonormal basis $\{e_k^r, k \in \mathbb{N}_0\}$, and similarly for $\mathcal{Y}$. We can then identify $\tilde{\mathcal{X}}$ with the space of sequences endowed with inner product with weight $1/(k^2+1)^{1/4}$ for the kth component.

It is important to highlight that the space $\tilde{\mathcal{Z}}$ is only used in the proof of the above lemma for the construction of the generating function. We will not refer to it further.

3.3. **Symplectic transformations.** In this section we prove that the symplectic transformation $(w_+, z_+) \mapsto \Psi(w_+, z_+) = (w, z)$ defined implicitly by the equations (3.6) is well-defined. We will write the transformation Ψ as

$$z = z_+ + \psi^z(w_+, z_+) = z_+ + \zeta(u, v_+), \quad w = w_+ + \epsilon\psi^w(w_+, z_+), \tag{3.10}$$

with $z_+ = (x_+, y_+)$ and $w_+ = (u_+, v_+)$. We also define $\psi = (\epsilon\psi^w, \psi^z)$. As before, let $\zeta(w) = (\zeta^x(w), \zeta^y(w))$ be the solution of (3.4). In the following we let $\mathcal{Z}_{-1}^{\mathbb{C}}$ be the dual to $\mathcal{Z}_1^{\mathbb{C}}$ with respect to the $\mathcal{Z}^{\mathbb{C}}$-inner product.

Lemma 3.7. *Assume (H0-H3) for $H(w, z)$. Let $\xi > 0$ be such that $\nu - \xi > 0$ and $\sigma - \xi > 0$ and assume*

$$\xi \geq \max(2K\delta, 8\epsilon C_{\mathcal{S}}), \quad 0 \leq \epsilon \leq 1/4 \tag{3.11}$$

where $\delta = |||r|||_\nu$ satisfies (3.3) and, as before, $C_{\mathcal{S}} = \|z\|_{\sigma_0}$. Then $(w_+, z_+) \mapsto \Psi(w_+, z_+) = (w, z)$ is an analytic symplectic transformation from $(\mathcal{V} + \mathrm{i}(\nu - \xi)) \times (\mathcal{S} + \mathrm{i}(\sigma - \xi))$ to $(\mathcal{V} + \mathrm{i}(\nu - \xi/2)) \times (\mathcal{S} + \mathrm{i}(\sigma - \xi/2))$. Moreover ψ^z is analytic from

$(\mathcal{V}+\mathrm{i}(\nu-\xi))\times(\mathcal{S}+\mathrm{i}(\sigma-\xi))$ *into* $\mathcal{Z}_1^{\mathbb{C}}$, ψ^w, ψ^z *are also analytic in* $z\in\mathcal{B}^{\mathcal{Z}_{-1}^{\mathbb{C}}}_{C_{\mathcal{S}}-\xi}(0)$ *and*

$$\|\psi^z|\|_{\nu-\xi,\sigma-\xi}\le\frac{\xi}{2},\quad \epsilon\|\psi^w\|_{\nu-\xi,\sigma-\xi}\le\frac{\xi}{2}, \tag{3.12}$$

where the norm $\|\cdot|\|$ *is defined in (H3).*

Proof. The functions $\psi^z=\psi^z(w_+,z_+)$, $\psi^u=\psi^u(w^+,z^+)$ and $\psi^v=\psi^v(w^+,z^+)$ are defined via the equations

$$\psi^z=\zeta(u_++\epsilon\psi^u,v_+), \tag{3.13a}$$

$$\psi^u=\langle J_{\mathcal{Z}}^{-1}\partial_v\zeta(u_++\epsilon\psi^u,v_+),\begin{pmatrix}x_++\zeta^x(u_++\epsilon\psi^u,v_+)\\ y_+\end{pmatrix}\rangle, \tag{3.13b}$$

$$\psi^v=-\langle J_{\mathcal{Z}}^{-1}\partial_u\zeta(u_++\epsilon\psi^u,v_+),\begin{pmatrix}x_++\zeta^x(u_++\epsilon\psi^u,v_+)\\ y_+\end{pmatrix}\rangle. \tag{3.13c}$$

When $\epsilon\ne0$ the function ψ^u is implicitly defined via (3.13b). We rewrite this equation as $\phi=\Pi(\phi,w_+,z_+)$ where $\phi=\psi^u$ and show that Π is a contraction on $\mathcal{B}^{\mathbb{R}^d}_\eta(0)$ where $\eta=\frac{\xi}{2\epsilon}$ and $d=d_{\mathcal{W}}$. Note that $\Pi(\phi,w_+,z_+)$ takes the form

$$\Pi(\phi,w_+,z_+)=\Pi_0(\phi,w_+)+\langle\Pi_1(\phi,w_+),z_+\rangle$$

and hence is affine in z_+. Moreover since $J_{\mathcal{Z}}^{-1}\in\mathcal{E}(\mathcal{Z};\mathcal{Z}_1)$ we see that $\Pi_1(\phi,w_+,z_+)\in\mathcal{Z}_1^{\mathbb{C}}$ and hence that Π is affine in $z_+\in\mathcal{Z}_{-1}^{\mathbb{C}}$.

We estimate using Lemma 3.4, Cauchy estimates (Lemma 3.1) and (3.11) that

$$\|\Pi(\phi,\cdot,\cdot)\|_{\nu-\xi,\sigma-\xi}\le\|\partial_v\zeta\|_{\nu-\xi/2}C_{\mathcal{S}}\le 2C_{\mathcal{S}}K\delta/\xi\le C_{\mathcal{S}}\le\frac{\xi}{8\epsilon}. \tag{3.14}$$

Hence $\Pi(\cdot,w_+,z_+)$ maps $\mathcal{B}^{\mathbb{R}^d}_\eta(0)$ to itself for all $w_+\in\mathcal{V}+\mathrm{i}(\nu-\xi)$, $z\in\mathcal{S}+\mathrm{i}(\sigma-\xi)$. Moreover,

$$\begin{aligned}\partial_\phi\Pi(\phi,w_+,z_+)&=\epsilon\langle J_{\mathcal{Z}}^{-1}\partial^2_{uv}\zeta(u_++\epsilon\phi,v_+),\begin{pmatrix}x_++\zeta^x(u_++\epsilon\psi^u,v_+)\\ y_+\end{pmatrix}\rangle\\&\quad+\epsilon\langle J_{\mathcal{Z}}^{-1}\partial_v\zeta(u_++\epsilon\psi^u,v_+),\mathbb{P}_x\partial_u\zeta(u_++\epsilon\phi,v_+)\rangle\end{aligned}$$

and so, again by Lemma 3.4, Cauchy estimates and (3.11)

$$\|\partial_\phi\Pi(\phi,\cdot,\cdot)\|_{\nu-\xi,\sigma-\xi}\le\epsilon\|\partial^2_{uv}\zeta\|_{\nu-\xi/2}C_{\mathcal{S}}+\epsilon\|\partial_u\zeta\|_{\nu-\xi/2}\|\partial_v\zeta\|_{\nu-\xi/2} \tag{3.15}$$

$$\le\epsilon\left(\frac{8K\delta C_{\mathcal{S}}}{\xi^2}+\left(\frac{2K\delta}{\xi}\right)^2\right)$$

$$\le\frac{4\epsilon C_{\mathcal{S}}}{\xi}+\epsilon\le\frac12+\epsilon\le\frac34. \tag{3.16}$$

Hence Π is a contraction on $\mathcal{B}^{\mathbb{R}^d}_\eta(0)$, and so by the contraction mapping theorem it has a unique fixed point $\psi^u(w_+,z_+):=\phi(w_+,z_+)$ which is analytic in $(w_+,z_+)\in(\mathcal{V}+\mathrm{i}(\nu-\xi))\times(\mathcal{S}+\mathrm{i}(\sigma-\xi))$. Note that estimates (3.14), (3.16) also hold for $z_+\in\mathcal{B}^{\mathcal{Z}_{-1}^{\mathbb{C}}}_{C_{\mathcal{S}}-\xi}(0)$ which proves that ψ^u is also analytic in $z_+\in\mathcal{B}^{\mathcal{Z}_{-1}^{\mathbb{C}}}_{C_{\mathcal{S}}-\xi}(0)$. By the chain rule, the same applies for ψ^v and ψ^z. □

Corollary 3.8. *The estimate for* $|\|\psi^z|\|_{\nu-\xi,\sigma-\xi}$ *in* (3.12) *can be improved to*

$$\|\psi^z|\|_{\nu-\xi,\sigma-\xi}\le|\|\zeta|\|_\nu\le K\delta, \tag{3.17a}$$

and we also have

$$\|\psi^w\|_{\nu-\xi,\sigma-\xi} \leq 4C_{\mathcal{S}}K\delta/\xi. \tag{3.17b}$$

Proof. From Lemma 3.7 we know that $(u|_{\mathcal{V}+\mathrm{i}(\nu-\xi),\mathcal{S}+\mathrm{i}(\sigma-\xi)}, v_+) \subset \mathcal{V}+\mathrm{i}(\nu-\xi/2)$. Hence,

$$\|\psi^z\|\|_{\nu-\xi,\sigma-\xi} = \|\zeta(u,v_+)\|\|_{\nu-\xi,\sigma-\xi} \leq \|\|\zeta\|\|_{\nu-\xi/2} \leq \|\|\zeta\|\|_{\nu}.$$

The other estimate follows as in the proof of Lemma 3.7 for both ψ^u and ψ^v upon replacing η by $2C_{\mathcal{S}}K\delta/\xi$. □

Let

$$\psi_0 = \psi|_{z_+=0}, \quad \psi_1 = \partial_z\psi|_{z_+=0}, \quad \psi_2 = \partial_z^2\psi|_{z_+=0},$$

and $\psi_i = (\psi_i^z, \psi_i^w)$, $i = 0,1,2$. Then

$$\epsilon\|\psi_0^w\|_{\nu-\xi} = \|w_+ - w(w_+,0)\|_{\nu-\xi} \leq \frac{\xi}{2}. \tag{3.18a}$$

In the next estimate we also interpret ψ_1^w and ψ_2^w through the $\mathcal{Z}^{\mathbb{C}}$-inner product so that

$$\langle\psi_1^w, z\rangle = \partial_{z_+}\psi^w|_{z_+=0}z, \quad \langle\psi_2^w z_1, z_2\rangle = \partial_{z_+}^2\psi^w|_{z_+=0}z_1z_2,$$

for all $z_1, z_2 \in \mathcal{Z}^{\mathbb{C}}$. Using Cauchy-estimates on $\psi_1^w = \partial_z\psi^w(w_+,0)$ and $\psi_2^w = \partial_z^2\psi^w(w_+,0)$, noting that ψ^w is analytic in $z \in \mathcal{Z}_{-1}^{\mathbb{C}}$ we obtain

$$\|\|\psi_1^w\|\|_{\nu-\xi} \leq \kappa^{-1}\|\psi^w\|_{\nu-\xi,\sigma-\xi}, \quad \|\psi_2^w\|\|_{\nu-\xi} \leq 2\kappa^{-2}\|\psi^w\|_{\nu-\xi,\sigma-\xi}, \tag{3.18b}$$

where $0 < \kappa \leq \sigma - \xi$. Moreover

$$\|\psi_1^z\|\|_{\nu-\xi} \leq \frac{K\delta}{\kappa} \leq \sigma/\kappa, \quad \|\psi_2^z\|\|_{\nu-\xi} \leq \frac{2K\delta}{\kappa^2}. \tag{3.18c}$$

The fact that $(w_+, z_+) \mapsto (w,z)$ is well-defined with domain $(\mathcal{V}+\mathrm{i}(\nu-\xi))\times(\mathcal{S}+\mathrm{i}(\sigma-\xi))$ and co-domain $(\mathcal{V}+\mathrm{i}(\nu-\xi/2))\times(\mathcal{S}+\mathrm{i}(\sigma-\xi/2))$ was crucial here and will be in the following. The $\nu-\xi/2$ and $\sigma-\xi/2$ terms in the co-domains allow for a step of $\xi/2$ to apply Lemma 3.1 to estimate derivatives on $(\mathcal{V}+\mathrm{i}(\nu-\xi/2))\times(\mathcal{S}+\mathrm{i}(\sigma-\xi/2))$ by function values on the larger domain $(\mathcal{V}+\mathrm{i}\nu)\times(\mathcal{S}+\mathrm{i}\sigma)$, c.f. the Cauchy estimate (3.1). This introduces a factor of $2\xi^{-1}$.

3.4. **Iterative lemma.** We are now ready to state and prove an Iterative Lemma which will is main ingredient of the proof of Theorem 2.1.

Lemma 3.9. ***(The Iterative Lemma for Hamiltonian systems)*** *Assume (H0-H3) for the Hamiltonian H, but relax the bounds on $\partial_w h$ and F slightly, so that h, a, r, f and F satisfy*

$$\|h\|_\nu \leq C_h,\ \|\partial_w h\|_{\nu-\xi/2} \leq C_h',$$
$$\|a\|\|_\nu \leq C_a,\ \|\|(L+a)^{-1}\|\|_\nu \leq K/2,$$
$$\|f\|_{\nu,\sigma} \leq C_f,\ \|F\|\|_{\nu,\sigma-\xi/2} \leq C_F,$$
$$\|\|r\|\|_\nu \leq \delta.$$

Here $\delta > 0$ satisfies (3.3) *and*

$$\min(\nu,\sigma) > \xi \geq \max\{8C_{\mathcal{S}}\epsilon, 2K\delta\}, \quad 0 \leq \epsilon \leq \frac{1}{4}, \tag{3.19a}$$

where $C_{\mathcal{S}} = \|z\|_\sigma$ as before. Then the symplectic transformation $\Psi : (w_+, z_+) \mapsto (w, z)$ from Lemma 3.7 mapping $(\mathcal{V} + \mathrm{i}(\nu - \xi)) \times (\mathcal{S} + \mathrm{i}(\sigma - \xi))$ into $(\mathcal{V} + \mathrm{i}(\nu - \xi/2)) \times (\mathcal{S} + \mathrm{i}(\sigma - \xi/2))$ transforms $H = H(w, z)$ into

$$H_+(w_+, z_+) = h_+(w_+) + \langle r_+(w_+), z_+\rangle + \frac{1}{2}\langle (L + a_+(w_+))z_+, z_+\rangle + f_+(w_+, z_+). \tag{3.19b}$$

Here

$$a_+ \in \mathcal{C}^\omega(\mathcal{V} + \mathrm{i}(\nu - \xi); \mathcal{E}(\mathcal{Z}^{\mathbb{C}}, \mathcal{Z}_1^{\mathbb{C}})),$$

and

$$F_+ = \nabla_z f_+ \in \mathcal{C}^\omega((\mathcal{V} + \mathrm{i}(\nu - \xi)) \times (\mathcal{S} + \mathrm{i}(\sigma - 3\xi/2); \mathcal{Z}_1^{\mathbb{C}}).$$

Moreover there is a constant c which is increasing in C_h', K, C_F, C_a, δ, C_H, $C_{\mathcal{S}}$ and $1/\kappa$, with κ satisfying $0 < \kappa \le \sigma - \xi$, and depends continuously on those constants only such that

$$\delta_+ = |||r_+|||_{\nu-\xi} \le c\frac{\epsilon\delta}{\xi}, \tag{3.19c}$$

and

$$\begin{aligned} \|h_+ - h\|_{\nu-\xi} \le c\delta, &\quad \|\partial_{w_+}(h_+ - h)\|_{\nu-3\xi/2} \le c\frac{\delta}{\xi}, \\ \|f_+ - f\|_{\nu-\xi,\sigma-\xi} \le c\frac{\delta}{\xi}, &\quad \|F_+ - F|||_{\nu-\xi,\sigma-3\xi/2} \le c\frac{\delta}{\xi^2}, \\ \|a_+ - a|||_{\nu-\xi} \le c\frac{\delta}{\xi}. & \end{aligned} \tag{3.19d}$$

Furthermore,

$$|||(L + a_+)^{-1}|||_{\nu-\xi} \le \frac{K_+}{2} := \frac{K}{2} + c\frac{\delta}{\xi}, \tag{3.19e}$$

provided

$$c\delta < \xi. \tag{3.19f}$$

Proof. We Taylor expand the new Hamiltonian $H_+(w_+, z_+) = H(w, z)$ around $z_+ = 0$ to put it into the form (3.19b) with

$$h_+(w_+) = H_+(w_+, 0), \quad r_+(w_+) = \nabla_{z_+} H_+(w_+, 0),$$

with

$$a_+(w_+) = \partial_{z_+}\nabla_{z_+} H_+(w_+, 0) - L,$$

and

$$f_+(w_+, z_+) = H_+(w_+, z_+) - h_+(w_+) - \langle r_+(w_+), z_+\rangle - \frac{1}{2}\langle (L + a_+)z_+, z_+\rangle.$$

We have for $w\in\mathcal{V}+\mathrm{i}(\nu-\xi)$

$$\begin{aligned}|h_+(w_+)-h(w_+)| &= |H_+(w_+,0)-H(w_+,0)| = |H(w|_{z_+=0},z|_{z_+=0})-H(w_+,0)|\\ &\le |H(w_+,0)-H(w|_{z_+=0},0)|\\ &+|H(w|_{z_+=0},0)-H(w|_{z_+=0},\psi^z|_{z_+=0})|,\\ &\le \epsilon\max_{s\in[0,1]}\|\partial_w h\circ(w_++s\epsilon\psi_0^w)\|_{\nu-\xi}\|\psi_w^0\|_{\nu-\xi}\\ &+\max_{s\in[0,1]}\|\partial_z H\circ(w_++\epsilon\psi_0^w,s\psi_0^z)\|_{\nu-\xi}\|\psi_0^z\|_{\nu-\xi},\\ &\le \epsilon\|\partial_w h\|_{\nu-\xi/2}\|\psi_0^w\|_{\nu-\xi}+\|\partial_z H\|_{\nu-\xi/2,\xi/2}\|\psi_0^z\|_{\nu-\xi},\\ &\le C_h'\epsilon\|\psi_0^w\|_{\nu-\xi}+\frac{C_H}{\kappa}\|\psi_0^z\|_{\nu-\xi}\le 4C_{\mathcal{S}}C_h'K\frac{\epsilon\delta}{\xi}+\frac{C_HK}{\kappa}\delta\\ &\le\left(C_h'+\frac{C_H}{\kappa}\right)K\delta. \end{aligned}\tag{3.20}$$

Here we used the mean value theorem, Lemma 3.7 and Corollary 3.8, and in the last inequality we used that $\xi\ge 8C_{\mathcal{S}}\epsilon$ by (3.19a).

Moreover, note that (3.20) and a Cauchy estimate give

$$\|\partial_w(h_+-h)\|_{\nu-3\xi/2}\le\frac{2}{\xi}\|h_+-h\|_{\nu-\xi}\le\left(C_h'+\frac{C_H}{\kappa}\right)\frac{2K\delta}{\xi}.$$

Let $\psi_0=(\epsilon\psi_0^w,\psi_0^z)$. Then, using (3.10), (3.18b) we obtain

$$\begin{aligned}|||r_+(w_+)||| &= |||\nabla_{z_+}H_+(w_+,0)||| = |||\nabla_{z_+}H(w|_{z_+=0},z|_{z_+=0})|||\\ &\le\|\partial_z H(w|_{z_+=0},z|_{z_+=0})\frac{\partial z}{\partial z_+}|_{z_+=0}\|_{\mathcal{E}(\mathcal{Z}_{-1},\mathbb{C})}\\ &+\|\partial_w H(w|_{z_+=0},z|_{z_+=0})\frac{\partial w}{\partial z_+}|_{z_+=0}\|_{\mathcal{E}(\mathcal{Z}_{-1},\mathbb{C})}\\ &\le |||\nabla_z H(w|_{z_+=0},\zeta(u|_{z_+=0},v_+))|||(1+|||\psi_1^z|||_{\nu-\xi})\\ &+\epsilon\|\partial_w H(w|_{z_+=0},\psi_0^z)\||||\psi_1^w|||_{\nu-\xi}.\end{aligned}\tag{3.21}$$

We estimate, with (3.12), that

$$\|(\partial_w H)\circ(w|_{z_+=0},\psi_0^z)\|_{\nu-\xi}\le\|\partial_w H\|_{\nu-\xi/2,\xi/2}.$$

Then, using the mean value theorem and Cauchy's estimate, this gives

$$\begin{aligned}\|\partial_w H\|_{\nu-\xi/2,\xi/2} &\le\|\partial_w h\|_{\nu-\xi/2}+\|\partial_{wz}^2 H\|_{\nu-\xi/2,\xi/2}\cdot\frac{\xi}{2}\\ &\le C_h'+\frac{\|\partial_w H\|_{\nu-\xi/2,\sigma}}{\kappa}\cdot\frac{\xi}{2}\le C_h'+\kappa^{-1}C_H.\end{aligned}$$

Moreover, since $\partial_z H(w,\zeta(w)) = 0$, using Cauchy's estimate and (3.17b) we obtain

$$\begin{aligned}
&|||\nabla_z H(w|_{z_+=0}, \zeta(u|_{z_+=0}, v^+))|||\\
&\quad = |||\nabla_z H(w|_{z_+=0}, \zeta(u|_{z_+=0}, v^+)) - \nabla_z H(w|_{z_+=0}, \zeta(w|_{z_+=0}))|||\\
&\quad \le \max_{\substack{s\in[0,1]\\ |||z|||\le\xi/2}} \|\partial_z \nabla_z H(w|_{z_+=0}, z)\|_{\mathcal{E}(\mathcal{Z}_1,\mathcal{Z}_1)} \cdot |||\zeta(u|_{z_+=0}, v^+) - \zeta(w|_{z_+=0})|||\\
&\quad \le |||L + \partial_z \nabla_z V|||_{\nu-\xi/2,\xi/2} |||\partial_v \zeta|||_{\nu-\xi/2} \epsilon \|\psi_0^v\|_{\nu-\xi}\\
&\quad \le (|||L||| + \|a|||_{\nu-\xi/2} + \|\partial_z F|||_{\nu-\xi/2,\xi/2}) \cdot \frac{2K\delta}{\xi} \cdot \frac{4C_{\mathcal{S}} K\delta\epsilon}{\xi}\\
&\quad \le (|||L||| + C_a + \kappa^{-1} C_F) \frac{4C_{\mathcal{S}} K\delta\epsilon}{\xi}.
\end{aligned}$$

Here we denote

$$\|\partial_z F|||_{\nu,\sigma} := \sup_{\substack{w\in\mathcal{V}+\mathrm{i}\nu,\\ z\in\mathcal{S}+\mathrm{i}\sigma}} \|\partial_z F(w,z)\|_{\mathcal{E}(\mathcal{Z};\mathcal{Z}_1)},$$

and, defining $\|\partial_z \nabla_z V|||_{\nu,\sigma}$, analogously we use that

$$|||\partial_z \nabla_z V|||_{\nu,\sigma} \le \|\partial_z \nabla_z V|||_{\nu,\sigma} \le \|a|||_\nu + \|\partial_z \nabla_z F|||_{\nu,\sigma}.$$

Plugging these estimates into (3.21), using (3.18b) and (3.18c), we obtain

$$\begin{aligned}
|||r_+|||_{\nu-\xi} &\le (|||L||| + C_a + \kappa^{-1} C_F) \frac{4C_{\mathcal{S}} K\delta\epsilon}{\xi} \cdot (1 + |||\psi_1^z|||_{\nu-\xi}) + \epsilon(C_h' + \kappa^{-1} C_H)|||\psi_1^w|||_{\nu-\xi}\\
&\le (|||L||| + C_a + \kappa^{-1} C_F) \frac{4C_{\mathcal{S}} K\delta\epsilon}{\xi} \cdot (1 + \sigma\kappa^{-1}) + (C_h' + \kappa^{-1} C_H) \frac{4\epsilon C_{\mathcal{S}} K\delta}{\xi\kappa}\\
&\le c\frac{\epsilon K\delta}{\xi}.
\end{aligned}$$

For $A_+ = L + a_+$ we have using (3.19b), that $\langle A_+ z_1, z_2\rangle = \partial_{z_+}^2 H_+(w_+, 0) z_1 z_2$ where

$$\begin{aligned}
\partial_{z_+}^2 H_+(w_+, 0) &= \partial_{z_+}(\partial_z H(w,z)\partial_{z_+} z + \partial_w H(w,z)\partial_{z_+} w)|_{z_+=0}\\
&= \partial_z^2 H(\partial_{z_+} z)^2 + \partial_z H \partial_{z_+}^2 z + \partial_w^2 H(\partial_{z_+} w)^2 + \partial_w H \partial_{z_+}^2 w\\
&+ 2(\partial_{zw}^2 H \partial_{z_+} z \partial_{z_+} w)_{\mathrm{sym}}\\
&= \partial_z^2 H(\mathrm{id} + \psi_1^z)^2 + \partial_z H \psi_2^z + \epsilon^2 \partial_w^2 H(\psi_1^w)^2 + \epsilon \partial_w H \psi_2^w\\
&+ 2\epsilon(\partial_{zw}^2 H(\mathrm{id} + \psi_1^z)\psi_1^w)_{\mathrm{sym}}\\
&= (L + \partial_z^2 V)(\mathrm{id} + \psi_1^z)^2 + \partial_z H \psi_2^z + \epsilon^2 \partial_w^2 H(\psi_1^w)^2 + \epsilon \partial_w H \psi_2^w\\
&+ 2\epsilon(\partial_{zw}^2 V(\mathrm{id} + \psi_1^z)\psi_1^w)_{\mathrm{sym}}.
\end{aligned}$$

Here we have used (3.10) and for any bilinear form $M : \mathcal{Z}^{\mathbb{C}} \times \mathcal{Z}^{\mathbb{C}} \to \mathbb{C}$ we define M_{sym} to be its symmetrization:

$$M_{\mathrm{sym}}(z_1, z_2) = \frac{1}{2}(M(z_1,z_2) + M(z_2,z_1));$$

all derivatives of H and V are evaluated at $(w|_{z_+=0}, z|_{z_+=0})$ and all derivatives w.r.t. z_+ of z and w are evaluated at $z_+ = 0$. Hence

$$\begin{aligned}
\langle\cdot, (a_+(w_+) - a(w_+))\cdot\rangle &= \partial_z^2 V - \partial_z^2 V(w_+, 0) + (\partial_z^2 V + L)(2 + \psi_1^z)\psi_1^z + \partial_z H \psi_2^z\\
&+ \epsilon^2 \partial_w^2 H(\psi_1^w)^2 + \epsilon \partial_w H \psi_2^w + 2\epsilon(\partial_{zw}^2 V(\mathrm{id} + \psi_1^z)\psi_1^w)_{\mathrm{sym}}.
\end{aligned}$$

Therefore, using Lemma 3.4, (3.17b), (3.18a), (3.18b), and (3.18c) we obtain

$$\begin{aligned}\|a_+(w_+) - a(w_+)\|\|_{\nu-\xi} &= \|\partial_z \nabla_z V(w|_{z_+=0}, z|_{z_+=0}) - \partial_z \nabla_z V(w_+, 0)\|\|_{\nu-\xi} \\ &+ (\|\|\partial_z \nabla_z V\|\|_{\nu-\xi/2,\xi/2} + \|\|L\|\|)(\|\|\psi_1^z\|\|_{\nu-\xi} + 2)\|\|\psi_1^z\|\|_{\nu-\xi} \\ &+ \|\partial_z H \psi_2^z\|\|_{\nu-\xi} + \epsilon^2 \|\partial_w^2 H\|_{\nu-\xi/2,\xi/2} \|\|\psi_1^w\|\|_{\nu-\xi}^2 \\ &+ \epsilon \|\partial_w H\|_{\nu-\xi/2,\xi/2} \|\|\psi_2^w\|\|_{\nu-\xi} \\ &+ 2\epsilon \|\partial_w \nabla_z V\|\|_{\nu-\xi/2,\xi/2} (1 + \|\psi_1^z\|\|_{\nu-\xi}) \|\|\psi_1^w\|\|_{\nu-\xi} \\ &\le \|\partial_z \nabla_z V(w|_{z_+=0}, z|_{z_+=0}) - \partial_z \nabla_z V(w_+, 0)\|\|_{\nu-\xi} \\ &+ (C_a + C_F \kappa^{-1} + \|\|L\|\|) \cdot (2 + \sigma\kappa^{-1}) \frac{K\delta}{\kappa} + \frac{C_H}{\kappa} \cdot \frac{2K\delta}{\kappa^2} \\ &+ \|\partial_w^2 H\|_{\nu-\xi/2,\xi/2} \left(\frac{4\epsilon K C_{\mathcal{S}} \delta}{\xi\kappa} \right)^2 + \|\partial_w H\|_{\nu-\xi/2,\xi/2} \frac{8\epsilon K C_{\mathcal{S}} \delta}{\xi\kappa^2} \\ &+ \|\partial_w \nabla_z V\|\|_{\nu-\xi/2,\xi/2} (1 + \sigma\kappa^{-1}) \frac{4\epsilon K C_{\mathcal{S}} \delta}{\xi\kappa}. \end{aligned} \tag{3.22}$$

Using Cauchy's estimate we get

$$\|\partial_w \nabla_z V\|\|_{\nu-\xi/2,\xi/2} \le \|\partial_w a z + \partial_w r + \partial_w F\|\|_{\nu-\xi/2,\xi/2} \le \frac{2(C_{\mathcal{S}} C_a + \delta + C_F)}{\xi}$$

and

$$\|\partial_w H\|_{\nu-\xi/2,\sigma} \le \frac{2C_H}{\xi}, \quad \|\partial_w^2 H\|_{\nu-\xi/2,\xi/2} \le \frac{4C_H}{\xi^2}.$$

Furthermore, using the mean value theorem, Cauchy's estimate and (3.17a), (3.17b) we obtain for $w \in \mathcal{V} + \mathrm{i}(\nu - \xi)$

$$\begin{aligned}&\|\partial_z \nabla_z V(w|_{z_+=0}, z|_{z_+=0}) - \partial_z \nabla_z V(w_+, 0)\|\| \\ &\quad \le \|\partial_z F(w|_{z_+=0}, z|_{z_+=0}) - \partial_z F(w|_{z_+=0}, 0)\|\| \\ &\qquad + \|a(w|_{z_+=0}) - a(w_+)\|\| \\ &\quad \le \|\partial_z^2 F\|\|_{\nu-\xi/2,\xi/2} \|\psi_0^z\|_{\nu-\xi} + \|\partial_w a\|\|_{\nu-\xi/2} \epsilon \|\psi_0^w\|_{\nu-\xi} \\ &\quad \le \frac{2C_F}{\kappa^2} \cdot K\delta + \frac{2C_a}{\xi} \cdot \frac{4\epsilon C_{\mathcal{S}} K\delta}{\xi}.\end{aligned}$$

Plugging these into (3.22) the estimate for $a_+ - a$ in (3.19d) then follows upon use of the conditions in (3.19a). Moreover for $A_+(w) = L + a_+(w)$ we have

$$\frac{2}{K}\|\|z\|\| \le \|\|A(w_+)z\|\| \le \|\|\|(a(w_+) - a_+(w_+))z\|\| + \|\|A_+(w_+)z\|\|$$

and so

$$\|\|A_+(w_+)z\|\| \ge \left(\frac{2}{K} - \|\|(a(w_+) - a_+(w_+))\|\| \right) \|\|z\|\|$$

and hence

$$\|\|A_+(w)^{-1}\|\| \le \frac{K}{2} \left(1 - \frac{Kc\delta}{2\xi} \right)^{-1} \le \frac{K}{2} \left(1 + \frac{cK\delta}{\xi} \right)$$

for $Kc\delta < \xi$. Here we have used that $(1-x)^{-1} = 1 + \frac{x}{1-x} \le 1 + 2x$ if $x \in [0, \frac{1}{2}]$, with $x = \frac{c\delta K}{2\xi}$. Redefining c to $\max(1, K^2)c$ then verifies (3.19e) if (3.19f) holds true.

Finally

$$\begin{aligned}\|F_+ - F\|_{\nu-\xi,\sigma-3\xi/2} \leq &\|\nabla_{z_+} H_+ - \nabla_{z_+} H\|_{\nu-\xi,\sigma-3\xi/2} + \|r_+ - r\|_{\nu-\xi} \\ &+ \|a_+ - a\|_{\nu-\xi} C_{\mathcal{S}}\end{aligned}$$

Using the estimates for $\|r_+ - r\|_{\nu-\xi}$ and $\|a_+ - a\|_{\nu-\xi}$ obtained above we only need to estimate the following

$$\begin{aligned}
\|\nabla_{z_+} H_+ - \nabla_{z_+} H\|_{\nu-\xi,\sigma-3\xi/2} &\leq \|(\mathrm{id} + \partial_{z_+}\psi^z)^* \nabla_z H \circ \Psi - \nabla_{z_+} H\|_{\nu-\xi,\sigma-3\xi/2} \\
&+ \epsilon \|(\partial_{z_+}\psi^w)^* (\partial_w H \circ \Psi)^*\|_{\nu-\xi,\sigma-3\xi/2} \\
&\leq \|(\nabla_z H)\circ\Psi - \nabla_{z_+} H\|_{\nu-\xi,\sigma-3\xi/2} + \|(\partial_{z_+}\psi^z)^*(\nabla_z H\circ\Psi)\|_{\nu-\xi,\sigma-3\xi/2} \\
&+ \epsilon\|\partial_{z_+}\psi^w\|_{\nu-\xi,\sigma-3\xi/2}\|\partial_w H\|_{\nu-\xi/2,\sigma-\xi} \\
&\leq \|\nabla_z\partial_w H\|_{\nu-\xi/2,\sigma-\xi}\epsilon\|\psi^w\|_{\nu-\xi,\sigma-3\xi/2} \\
&+ \|\nabla_z\partial_z H\|_{\nu-\xi/2,\sigma-\xi}\|\psi^z\|_{\nu-\xi,\sigma-3\xi/3} + \|\nabla_z H\|_{\nu-\xi/2,\sigma-\xi}\|\partial_{z_+}\psi^z\|_{\nu-\xi,\sigma-3\xi/2} \\
&+ \epsilon\|\partial_{z_+}\psi^w\|_{\nu-\xi,\sigma-3\xi/2}\|\partial_w H\|_{\nu-\xi/2,\sigma-\xi} \\
&\leq \|\partial_w r + \partial_w a z + \partial_w F\|_{\nu-\xi/2,\sigma-\xi}\frac{4\epsilon C_{\mathcal{S}} K\delta}{\xi} + \|a + L + \partial_z F\|_{\nu-\xi/2,\sigma-\xi} K\delta \\
&+ \|r + Lz + az + F\|_{\nu-\xi/2,\sigma-\xi}\|\partial_{z_+}\psi^z\|_{\nu-\xi,\sigma-3\xi/2} + \frac{2C_H}{\xi}\frac{4\epsilon C_{\mathcal{S}} K\delta}{\xi}\frac{2}{\xi} \\
&\leq (C_r + C_a C_{\mathcal{S}} + C_F)\frac{2}{\xi}\frac{4\epsilon C_{\mathcal{S}} K\delta}{\xi} + (C_a + \|L\| + 2C_F/\xi)K\delta \\
&+ (C_r + (C_a + \|L\|)C_{\mathcal{S}} + C_F)\frac{2K\delta}{\xi} + \frac{2C_H K\delta}{\xi^2} \\
&\leq c\delta/\xi^2,
\end{aligned}$$

where we used that $H\circ\Psi = H_+$ in the first inequality and the definition $\psi := (\psi^z, \epsilon\psi^w)$ in the third and fourth inequality. In the third inequality we use the mean value theorem and in the fourth inequality and final inequality Cauchy's estimate together with (3.17a) and (3.17b) and (3.19a). A similar estimate shows that $\|f_+ - f\|_{\nu-\xi,\sigma-\xi} \leq c\delta/\xi$. □

3.5. **Proof of Theorem 2.1.** To finish the proof of the theorem we successively apply the Iterative Lemma 3.9 as follows: We first apply the symplectic transformation from the iterative Lemma 3.9 three times to introduce $(w_3, z_3) \mapsto (w_2, z_2) \mapsto (w_1, z_1) \mapsto (w_0, z_0)$ taking

$$\xi = \xi_0 = \xi_1 = \xi_2 = \frac{1}{6}\min(\nu_0 - \nu, \sigma_0 - \sigma).$$

We choose $\delta_0 > 0$ and $\epsilon > 0$ sufficiently small to satisfy (3.3) and the conditions (3.19a) and (3.19f) of Lemma 3.9 for the above choice of ξ_0. Applying this lemma once we obtain $\delta_1 = \mathcal{O}(\epsilon)$. For all successive iterations we use the following bound for $C''_{F_n}[\nu_n, K_n\delta_n]$ from (3.2) where K_n, δ_n etc. denote the constants of Lemma 3.9 after n iterations: We set

$$C''_{F_n}[\nu_n, K_n\delta_n] = \frac{2C_{F_n}[\nu_n, K_n\delta_n + \kappa]}{\kappa^2}, \tag{3.23}$$

by applying the Cauchy estimate (3.1). Here $\kappa > 0$, $K_n\delta_n + \kappa \leq \sigma_n$ and $C_F[\nu_n, K_n\delta_n + \kappa]$ is such that $\|F_n\|_{\nu_n, K_n\delta_n+\kappa} \leq C_{F_n}[\nu_n, K_n\delta_n + \kappa]$. We let $\sigma_n - \xi_n \geq \sigma$ for all n

so that we can choose $\kappa = \sigma$ (noting that $K_n\delta_n \le \xi_n/2$ by (3.19a)). We then use the condition

$$\delta_n < \frac{\kappa^2}{K_n^2 C_{F_n}[\nu_n, K_n\delta_n + \kappa]} \tag{3.24}$$

instead of (3.3) in the following.

Since $\delta_1 = \mathcal{O}(\epsilon)$ we can satisfy (3.24) and the other conditions (3.19a) and (3.19f) of Lemma 3.9 for sufficiently small ϵ and therefore apply Lemma 3.9 twice to obtain $\delta_2 = \mathcal{O}(\epsilon^2)$ and $\delta_3 = \mathcal{O}(\epsilon^3)$. We can then ensure that δ_3 is small enough so that for $n \ge 3$ the conditions of Lemma 3.9, (3.24), (3.19a) and (3.19f), are satisfied for the choice $\xi_n = \mathcal{O}(\epsilon)$. We now apply Lemma 3.9 successively starting from $(\mathcal{V} + \mathrm{i}\nu_3) \times (\mathcal{S} + \mathrm{i}\sigma_3)$ with

$$\nu_3 - \nu \ge \frac{1}{2}(\nu_0 - \nu), \quad \sigma_3 - \sigma \ge \frac{1}{2}(\sigma_0 - \sigma). \tag{3.25}$$

Note that

$$\|h_3 - h_0\|_{\nu_3},\ \|a_3 - a_0\|\|_{\nu_3},\ \|f_3 - f_0\|_{\nu_3,\sigma_3},\ \|F_3 - F_0\|\|_{\nu_3,\sigma_3-\xi_0/2} = \mathcal{O}(\delta_0).$$

We then apply the transformations

$$\begin{aligned}\Psi_n :&(\mathcal{V} + \mathrm{i}\nu_{n+1}) \times (\mathcal{S} + \mathrm{i}\sigma_{n+1}) \to (\mathcal{V} + \mathrm{i}(\nu_n - \xi_n/2)) \times (\mathcal{S} + \mathrm{i}(\sigma_n - \xi_n/2)),\\ &(w_{n+1}, z_{n+1}) \quad \mapsto \quad (w_n, z_n),\end{aligned}$$

iteratively, with

$$\xi_n = 2c_*\epsilon \ge 2\epsilon \max(4C_{\mathcal{S}}, c_n), \tag{3.26}$$

with $\nu_n = \nu_3 - \sum_{i=3}^{n-1} \xi_i \ge \nu$, $\sigma_n = \sigma_3 - \sum_{i=3}^{n-1} \xi_i \ge \sigma$ and with c_* to be determined later. Here we choose $\epsilon > 0$ small enough such that $\xi_n \le 1$. This choice of ξ_n ensures that

$$\delta_{n+1} \le \frac{1}{2}\delta_n,$$

cf. (3.19c), and that

$$K_{n+1} - K_n, C_{f_{n+1}} - C_{f_n}, C_{a_{n+1}} - C_{a_n}, C_{h_{n+1}} - C_{h_n}, C'_{h_{n+1}} - C'_{h_n} \le \frac{c_n\delta_n}{\xi_n} \le \frac{\delta_n}{\epsilon},$$

and

$$C_{F_{n+1}} - C_{F_n} \le \frac{c_n\delta_n}{\xi_n^2} \le \frac{\delta_n}{\epsilon\xi_n} \le \frac{\delta_n}{8C_{\mathcal{S}}\epsilon^2},$$

c.f. (3.19d), (3.19e), (3.26), where we take ϵ small enough such that $\xi_n < \min(\nu_n - \nu, \sigma_n - \sigma)$ for $n \le N$, with N to be determined later. Then

$$\delta_{n+1} \le 2^{-n+2}\delta_3,$$

where $\delta_3 = \mathcal{O}(\epsilon^3)$. This proves that the constants from the Iterative Lemma 3.9 are bounded uniformly with respect to $3 \le n \le N$ with

$$K_n - K_3, C_{F_n} - C_{F_3}, C_{a_n} - C_{a_3}, C_{h_n} - C_{h_3}, C'_{h_n} - C'_{h_3} = \mathcal{O}(\epsilon). \tag{3.27}$$

Since the constant c from Lemma 3.9 is increasing and continuous in K, C_F, C_a, C_h and C'_h there is c_* such that $c_n \le c_*$ for all $n \le N$. We choose $c_* \ge 4C_{\mathcal{S}}$ and set $\xi_n = 2c_*\epsilon$, see (3.26). Because of the inequalities

$$\nu \le \nu_3 - 2c_*\epsilon(N-2) \le \nu_{N+1} = \nu_3 - \sum_{i=3}^{N} \xi_i,$$
$$\sigma \le \sigma_3 - 2c_*\epsilon(N-2) \le \sigma_{N+1} = \sigma_3 - \sum_{i=3}^{N} \xi_i, \tag{3.28}$$

noting that we want to define the transformed Hamiltonian on $\mathcal{V} + \mathrm{i}\nu \times \mathcal{S} + \mathrm{i}\sigma$ and using (3.25), we take N to be

$$N = \left\lceil \frac{M}{4c_*\epsilon} \right\rceil, \quad M = \min(\nu_0 - \nu, \sigma_0 - \sigma). \tag{3.29}$$

Here we denote by $\lceil x \rceil$ the smallest integer $\ge x$. This completes the proof of Theorem 2.1.

3.6. **Remarks on the proof of Theorem 2.1.** The following remark shows that we can construct the slow manifold such that it contains a given equilibrium of the Hamiltonian slow-fast system (2.1).

Remark 3.10. *Assume (H0-H3) and let $\delta_0 > 0$ and $\epsilon > 0$ be sufficiently small. In addition assume that there exists a locally unique equilibrium of* (2.1) *which in the (w_0, z_0)-coordinates takes the form $(w^e, 0) \in \mathcal{V} \times \mathcal{S}$. Then the equilibrium $(w^e, 0)$ is a fixed point of Ψ.*

Proof. Let (w_+^e, z_+^e) be such that $(w^e, 0) = \Psi(w_+^e, z_+^e)$ where Ψ is the symplectic transformation from (3.6), (3.10) used in the Iterative Lemma 3.9. First note that $z^e = 0$ implies that $\zeta(w^e) = 0$. Moreover from the definition of Ψ we have

$$z_+^e = z^e - \zeta(u^e, v_+^e) = -\zeta(u^e, v_+^e),$$
$$u_+^e = u^e - \epsilon\langle J_{\mathcal{Z}}^{-1} \partial_{v_+} \zeta(u, v_+), (x, y_+)\rangle,$$
$$v_+^e = v^e + \epsilon\langle J_{\mathcal{Z}}^{-1} \partial_{u} \zeta(u, v_+), (x, y_+)\rangle.$$

Insertion then proves the result. □

Also note that we can assume that the equilibrium (w^e, z^e) in the above remark is at $z^e = 0$ without loss of generality, by introducing the affine symplectic transformation $(w_0, z_0) \mapsto (w_0, z_0 - z^e)$. It is important to start our iteration from $z^e = 0$ - we can then ensure that slow manifolds, that we defined iteratively in the proof of Theorem 2.1, contain this equilibrium. Obviously we could also transform $w^e = 0$ but this is not necessary.

Remark 3.11. Lu [22] uses our method for obtaining a symplectic slow manifold as presented in an earlier preprint version of this paper to study breathers in a semilinear wave equation

$$u_{tt} = u_{xx} - u + f(u) \tag{3.30}$$

where $f(u)$ is odd, holomorphic and $f'(0) = 0$. He studies (3.30) on $2\pi/\omega$ odd periodic functions where the lowest Fourier mode $\sin x$ corresponds to the slow dynamics (after rescaling $x \to x/\omega$). He transforms $v = u_t$ such that the transformed system becomes well-posed on the subspace $\mathcal{Z}$ of odd functions in $\mathcal{H}_1 \times \mathcal{H}_1$. The resulting fast system (in the slow time) takes the form

$$\dot{z} = J_{\mathcal{Z}} L z + \epsilon^2 B(w, z). \tag{3.31}$$

So compared to (2.1) the nonlinearity is of order ϵ^2. Instead of solving $\dot{z} = 0$ in Lemma 3.4, Lu just solves $Lz = \epsilon^2 r(w)$, where $B(w,z) = J_{\mathcal{Z}} r(w) + \mathcal{O}(z)$, for $\hat{\zeta}(w) = \epsilon^2 L^{-1} r(w)$ and defines the symplectic transformation Ψ from Lemma 3.7 used in the iterative Lemma 3.9, with $\hat{\zeta}(w)$ instead of $\zeta(w)$. For the special case (3.31) the error δ of his construction of the slow manifold still shrinks by an order of ϵ in each step, and this simplifies the proof considerably.

4. An invariant two-dimensional slow manifold

In this section we prove the existence of a two dimensional normally elliptic slow manifold with exponentially small gaps under the following assumptions: Consider again a real analytic slow-fast Hamiltonian system with Hamiltonian $H_0(w_0, z_0)$, but now with a single slow degree of freedom, which in addition to (H0-H3) satisfies the following assumptions:

(I1) $d_{\mathcal{W}} = 1$, and $z_0 = 0$ is invariant for $\epsilon = 0$ for $w_0 \in \mathcal{V} + \mathrm{i}\nu$. Moreover $\{z_0 = 0, \epsilon = 0\}$ is filled with a family of non-degenerate periodic orbits parametrized by energy $E \in (E_1, E_2) + \mathrm{i}e_0$. Their frequency ω^E as a function of energy satisfies $\frac{\partial \omega^E}{\partial E} \neq 0$ for $E \in (E_1, E_2) + \mathrm{i}e_0$, $e_0 > 0$.

(I2) $\dim \mathcal{Z} = 2d_{\mathcal{Z}} < \infty$, $J_{\mathcal{Z}}$ is standard, and $A_0(w)$ is of the form

$$A_0(w) = L + a_0(w) = L + \epsilon M_0(w), \tag{4.1}$$

suppressing the ϵ-dependency in $M_0(w)$, with

$$L = \mathrm{diag}\,(\omega_1, \ldots, \omega_{d_{\mathcal{Z}}}, \omega_1, \ldots, \omega_{d_{\mathcal{Z}}}).$$

Moreover, $\delta_0 = \mathcal{O}(\epsilon)$.

(I3) We have $\omega_i \neq 0$ for all i and the following non-resonance condition holds:

$$\forall \ell \neq m \quad \omega_\ell \neq \omega_m. \tag{4.2}$$

Then cf. Theorem 2.1 there exists a symplectic map $(w, z) \mapsto (w_0, z_0)$ that transforms the Hamiltonian into

$$H = h(w) + r(w) \cdot z + \frac{1}{2} A(w) z \cdot z + f(w, z), \tag{4.3}$$

with $r = \mathcal{O}(\mathrm{e}^{-C/\epsilon})$ provided ϵ is sufficiently small. Here, as before, $f = \mathcal{O}(z^3)$, and from $\delta_0 = \mathcal{O}(\epsilon)$, we conclude that $A(w) = L + \epsilon M(w)$ has the same form as $A_0(w)$.

Note that if both (4.1) and (4.2) are not satisfied then "Takens chaos" can occur, see [38]. In this section we consider (2.1) on the fast time scale $\tau = t/\epsilon$.

In words, the result of this section is then the following: A periodic orbit for $h = h(w)$ can be continued into the full system given by (4.3) provided that there is no resonance with the fast system. If there is a resonance, then this only excludes exponentially small bands in the w-plane of periodic orbits. This result can be viewed as an extension of a result of Gelfreich and Lerman in [11] to several fast variables.

Remark 4.1. The setting considered in this section applies to the LK model in [40, Eqs. 2.7-2.10] and the generalized conservative versions [40, Eqs. 6.1-6.2] with $s \in \mathbb{R}^2$, and the main result (Theorem 4.6, below) therefore applies to these examples.

Before stating the result (Theorem 4.6 below), we perform a sequence of simplifications serving to bring the system into a form appropriate for application of the contraction mapping theorem. It is important to note that we are not connecting with $\epsilon = 0$. Instead we are introducing an artificial perturbation parameter μ.

First we transform $h = h(w)$ into action-angle variables (I, ϕ). We have

Lemma 4.2. *Assume (H0-3) and (I1-I3). Let $(I, \phi) \mapsto w$ be the symplectic change of coordinates which transforms $h(w)$ to $h(w) = \breve{h}(I)$. This transformation is analytic from $w \in \mathcal{V} + \mathrm{i}\nu$ to $(\phi, I) \in ([0, 2\pi] + \mathrm{i}\psi) \times ((I_1, I_2) + \mathrm{i}\iota)$ for some $\psi, \iota > 0$. Here $[0, 2\pi] + \mathrm{i}\psi$ is a complex neighbourhood of length ψ around $[0, 2\pi]$ and $(I_1, I_2) + \mathrm{i}\iota$ a complex neighbourhood of length ι around $(I_1, I_2) \subseteq \mathbb{R}^+$. This map transforms (4.3) into*

$$H(w, z) = \breve{H}(\phi, I, z) = \breve{h}(I) + \langle \breve{r}(\phi, I), z\rangle + \frac{1}{2}\langle \breve{A}(\phi, I)z, z\rangle + \breve{f}(\phi, I, z), \qquad (4.4)$$

where $\breve{h}(I) = h(w)$, $\breve{r}(\phi, I) = r(w)$, $\breve{A}(\phi, I) = L + \epsilon \breve{M}(\phi, I)$, $\breve{M}(\phi, I) = M(w)$, and $\breve{f}(\phi, I, z) = f(w, z)$.

Proof. The Hamiltonian system $h = h(w)$ is integrable since it is a one-degree of freedom system. This transformation does not depend upon the fast variables and can therefore directly be lifted to the full space. □

Next we reduce to an energy level: Since $\frac{\partial \omega^E}{\partial E} \neq 0$ by (I1) we have $\omega(I) := \partial_I \breve{h}(I) \neq 0$, and so we can solve the equation $H = E$ for $I = I^E(\phi, z)$ when $z \in \mathcal{B}_\sigma^{\mathcal{Z}^{\mathbb{C}}}(0)$ by potentially decreasing $\sigma > 0$, and we may introduce the angle ϕ as new time.

Lemma 4.3. *Under the above assumptions $z = z(\phi)$ solves the following non-autonomous Hamiltonian system of equations:*

$$\epsilon z'(\phi) = -J_{\mathcal{Z}} \nabla_z I^E(\phi, z). \qquad (4.5)$$

Proof. By definition

$$\epsilon \dot{z} = \epsilon \frac{\mathrm{d}z}{\mathrm{d}\phi} \partial_I \breve{H} = J_{\mathcal{Z}} \nabla_z \breve{H}.$$

Next, we differentiate

$$\breve{H}(\phi, I^E(\phi, z), z) = E, \qquad (4.6)$$

with respect to z to obtain

$$\partial_I \breve{H} \nabla_z I^E = -\nabla_z \breve{H}. \qquad (4.7)$$

This completes the result. □

Lemma 4.4. *Under the above assumption $I^E = I^E(\phi, z)$ takes the following form*

$$I^E(\phi, z) = \breve{h}^{-1}(E) - \frac{1}{2}\langle A^E(\phi)z, z\rangle - \langle r^E(\phi), z\rangle - f^E(\phi, z), \qquad (4.8)$$

with

$$r^E(\phi) = \breve{r}(\phi, \breve{h}^{-1}(E))/\omega^E, \quad f^E(\phi, z) = \mathcal{O}(\|z\|^3)$$

and

$$A^E(\phi) := \breve{A}(\phi, \breve{h}^{-1}(E))/\omega^E + \mathcal{O}(e^{-C/\epsilon}) = L^E + \epsilon M^E(\phi),$$

where

$$\omega^E = \omega(\breve{h}^{-1}(E)) \tag{4.9}$$

and

$$L^E = L/\omega^E, \quad M^E = \breve{M}(\phi, \breve{h}^{-1}(E))/\omega^E + \mathcal{O}(e^{-C/\epsilon}).$$

Here I^E, f^E, M^E and L^E are analytic in $\phi \in [0, 2\pi] + \mathrm{i}\psi$, $E \in (E_0, E_1) + \mathrm{i}e_0$ and $z \in \mathcal{B}_\sigma^{\mathcal{Z}^{\mathbb{C}}}(0)$. Finally

$$\|r^E\|_\psi := \max_{\phi\in[0,2\pi]+\mathrm{i}\psi} \|r^E(\phi)\| = \mathcal{O}(e^{-C/\epsilon}) \tag{4.10}$$

uniformly for $E \in (E_0, E_1) + \mathrm{i}e_0$.

Proof. Equation (4.6) with $z = 0$ gives

$$I^E(\phi, 0) = \breve{h}^{-1}(E),$$

cf. (4.4). Setting $z = 0$ in (4.7) then gives

$$\nabla_z I^E|_{z=0} = -\breve{r}(\phi, \breve{h}^{-1}(E))/\omega^E = -r^E(\phi) = \mathcal{O}(\mathrm{e}^{-C/\epsilon}). \tag{4.11}$$

If we differentiate (4.6) again with respect to z we get

$$\begin{aligned} 0 = \partial_z^2 \breve{H}|_{z=0} + \partial_I^2 \breve{H}|_{z=0}(\partial_z I^E|_{z=0})^2 + 2(\partial_I \partial_z \breve{H}|_{z=0}\partial_z I^E|_{z=0})_{\mathrm{sym}} \\ + \partial_I \breve{H}|_{z=0}\partial_z^2 I^E|_{z=0}, \end{aligned}$$

and so we find that

$$\begin{aligned} \partial_z \nabla_z I^E|_{z=0} = -(\breve{A}(\phi, \breve{h}^{-1}(E)) + \partial_I^2 \breve{h}(I)\nabla_z I^E|_{z=0}\partial_z I^E|_{z=0})/\omega^E, \\ - (\partial_I \breve{r}(\phi, I)\partial_z I^E|_{z=0} + (\partial_I \breve{r}(\phi, I)\partial_z I^E|_{z=0})^T)/\omega^E \end{aligned}$$

where $I = \breve{h}^{-1}(E)$. Using (4.11) this gives the result. □

Now we fix ϵ small and introduce $\|r^E\|_\psi \le \mu^2 = \mathcal{O}(\mathrm{e}^{-C/\epsilon})$ as a measure of the remainder in (4.8). Setting $z = \mu\hat{z}$ then transforms (4.5) into

$$\epsilon\hat{z}' = J_{\mathcal{Z}}(A^E(\phi)\hat{z} + \hat{r}^E(\phi) + \hat{F}^E(\phi, \hat{z})) \tag{4.12}$$

where

$$\hat{r}^E(\phi) := r^E(\phi)/\mu, \quad \hat{F}^E(\phi, \hat{z}) := F^E(\phi, \mu\hat{z})/\mu = \nabla_z f^E(\phi, \mu\hat{z})/\mu$$

Choosing $\epsilon > 0$ small enough such that $\mu < 1/2$ we see that $\hat{F}^E(\phi, \cdot)$ is analytic on $\mathcal{B}_{2\sigma}^{\mathcal{Z}}(0)$. Then due to (4.10) and due to the fact that $F^E(\phi, z) = \mathcal{O}(\|z\|^2)$ we obtain

$$\|\hat{r}^E\|_\psi := \sup_{\phi\in[0,2\pi]+\mathrm{i}\psi} \|\hat{r}^E(\phi)\| = \mathcal{O}(\mu), \quad \|\hat{F}^E\|_{\psi,2\sigma} := \sup_{\substack{\phi\in[0,2\pi]+\mathrm{i}\psi,\\ z\in\mathcal{B}_{2\sigma}^{\mathcal{Z}}(0)}} \|\hat{F}^E(\phi, \hat{z})\| = \mathcal{O}(\mu) \tag{4.13}$$

uniformly in $E \in (E_1, E_2)$. In the notation $\|\hat{F}^E\|_{\psi,2\sigma}$, and in what follows we adapt the definition from (H3), with $\mathcal{S} = \{0\}$.

Let

$$\Pi^{\mu,E} : \{\phi = 0\} \to \{\phi = 2\pi\}$$

be the stroboscopic mapping obtained from (4.12). It is symplectic since the system is Hamiltonian. Note that $\Pi^{0,E}(0) = 0$ due to (I1). The persistence of this fixed point for $\mu \neq 0$ provides the persistence of the periodic orbits, which we have parametrized by E. To study this mapping we consider the monodromy matrix

$\Psi^E(2\pi, 0)$ associated with the linear problem (obtained from (4.12) by setting $\mu = 0$)

$$\epsilon \hat{z}' = J_{\mathcal{Z}} A^E(\phi)\hat{z} = J_{\mathcal{Z}} \left(L^E + \epsilon M^E(\phi)\right)\hat{z}. \tag{4.14}$$

The eigenvalues of $\Psi^E(2\pi, 0)$, $\lambda_1^E, \ldots, \lambda_{2d_{\mathcal{Z}}}^E$, are the characteristic multipliers of $\Pi_{\mu,E}$ at $z = 0$. Exploiting the form of $A^E(\phi) = L^E + \epsilon M^E(\phi)$ and the non-resonance condition (4.2) we can approximate those very accurately using the following lemma:

Lemma 4.5. *Under the above assumptions for all sufficiently small $\epsilon > 0$ there exist $\tilde{C} > 0$ and a linear change of variables $\hat{z} \mapsto \tilde{z} = \hat{z} + \epsilon T^E(\phi)\hat{z}$, which is 2π-periodic in ϕ, and is analytic in $(\phi, E) \in [0, 2\pi] \times (E_1, E_2)$ which transforms (4.14) into*

$$\begin{aligned} \epsilon \tilde{z}_j' &= (\omega_j + \epsilon b_j^E(\phi))\tilde{z}_{j+d_{\mathcal{Z}}}/\omega^E + (R^E(\phi)z)_j, \\ \epsilon \tilde{z}_{j+d_{\mathcal{Z}}}' &= -(\omega_j + \epsilon b_j^E(\phi))\tilde{z}_j/\omega^E + (R^E(\phi)z)_{d_{\mathcal{Z}}+j}. \end{aligned} \tag{4.15}$$

Here $b_j^E(\phi)$, $j = 1, \ldots d_{\mathcal{Z}}$, are scalar analytic functions and $R^E(\phi)$ is a $(2d_{\mathcal{Z}}, 2d_{\mathcal{Z}})$ symmetric matrix which is analytic in $(\phi, E) \in [0, 2\pi] \times (E_1, E_2)$ and satisfies

$$\|R^E\| := \max_{\phi \in [0,2\pi]} \|R^E(\phi)\| = \mathcal{O}(\mathrm{e}^{-\hat{C}/\epsilon}) \tag{4.16}$$

uniformly in $E \in (E_1, E_2)$.

Proof. Consider the system

$$\dot{z} = J_{\mathcal{Z}}(L(\phi) + \epsilon M(\phi))z \tag{4.17}$$

where

$$L(\phi) = \mathrm{diag}(\lambda_1(\phi), \ldots, \lambda_{d_{\mathcal{Z}}}(\phi), \lambda_1(\phi), \ldots, \lambda_{d_{\mathcal{Z}}}(\phi))$$

is diagonal, $M(\phi)$ is symmetric and $L(\phi)$ and $M(\phi)$ are analytic in $\phi \in [0, 2\pi] + \mathrm{i}\psi$ and also depend on ϵ and E. Note that (4.14) is of the form (4.17), see Lemma 4.4, and we will use (4.17) to set up an iterative lemma.

For $\epsilon > 0$ small enough let $\mathrm{id} + \epsilon T(\phi)$ be the linear coordinate transformation such that

$$(\mathrm{id} + \epsilon T(\phi))J_{\mathcal{Z}}(L(\phi) + \epsilon M(\phi))(\mathrm{id} + \epsilon T(\phi))^{-1} = J_{\mathcal{Z}} L_+(\phi)$$

where $L_+(\phi)$ has the same form as $L(\phi)$. Then $z_+ := (\mathrm{id} + \epsilon T(\phi))z$ defines a symplectic change of coordinates. Moreover there are matrix valued functions G and F with $G(L, 0) = 0$, $F(L, 0) = 0$ such that

$$L_+ = L + G(L, \epsilon M), \quad T_+ = F(L, \epsilon M).$$

Both G and F are analytic as functions on $\mathcal{B}_1 \times \mathcal{B}_2$. Here $\mathcal{B}_1$ is ball of radius $r_1 > 0$ around $L_0 := L^E$ from (4.14) in the $d_{\mathcal{Z}}$ dimensional space of complex diagonal (n, n) matrices, where $n = 2d_{\mathcal{Z}}$, $\mathcal{B}_2$ a ball of radius $r_2 > 0$ around 0 in the space of hermitian (n, n) matrices, and r_1, r_2 are small enough such that $J_{\mathcal{Z}}(L + S)$ has disjoint eigenvalues for $L \in \mathcal{B}_1$, $S \in \mathcal{B}_2$. Therefore there are some smooth non-decreasing functions f and g mapping a neighbourhood of 0 in $\mathbb{R}^2$ into $\mathbb{R}$ such that

$$\begin{aligned} \|F(L, M)\|_\psi &\le \|M\|_\psi f(\|L - L_0\|_\psi, \|M\|_\psi), \\ \|G(L, M)\|_\psi &\le \|M\|_\psi g(\|L - L_0\|_\psi, \|M\|_\psi). \end{aligned}$$

Now $z_+ = (\mathrm{id} + \epsilon T(\phi))z$ satisfies

$$\begin{aligned}\epsilon z_+' &= (\mathrm{id} + \epsilon T(\phi))\epsilon z' + \epsilon^2 T'(\phi) z\\ &= (\mathrm{id} + \epsilon T(\phi)) J_{\mathcal{Z}} \left(L(\phi) + \epsilon M(\phi)\right) z + \epsilon^2 T'(\phi) z\\ &= J_{\mathcal{Z}} L_+(\phi) z_+ + \epsilon^2 T'(\phi)(\mathrm{id} + \epsilon T(\phi))^{-1} z_+\\ &= J_{\mathcal{Z}} L_+(\phi) z_+ + \epsilon M_+(\phi) z_+\end{aligned}$$

where

$$M_+(\phi) = \epsilon T'(\phi)(\mathrm{id} + \epsilon T(\phi))^{-1}.$$

Using Cauchy's estimate we get, with $B = L - L_0$,

$$\|M_+\|_{\psi-\xi} \le \epsilon \frac{\|T\|_\psi}{\xi(1-\epsilon\|T\|_\psi)} \le \frac{\epsilon\|M\|_\psi f(\|B\|_\psi, \epsilon\|M\|_\psi)}{\xi(1-\epsilon\|M\|_\psi f(\|B\|_\psi, \epsilon\|M\|_\psi))} \le \|M\|_\psi/2$$

if

$$\xi = 2\epsilon c, \quad \text{where} \quad c \ge \mathscr{C}(\|B\|_\psi, \|M\|_\psi) := \max\left(1, \frac{1-\epsilon\|M\|_\psi f(\|B\|_\psi, \epsilon\|M\|_\psi)}{f(\|B\|_\psi, \epsilon\|M\|_\psi)}\right).$$

When iterating this procedure we need to ensure that $c_n = \mathscr{C}(\|B_n\|_{\psi_n}, \epsilon\|M_n\|_{\psi_n})$ is bounded independent of n. For this we first show that $C_{B_n} := \|B_n\|_{\psi_n}$ where $B_n = L_n - L_0$, is sufficiently small and and $C_{M_n} := \|M_n\|_{\psi_n}$ bounded, so that $L_n \in \mathcal{B}_1$ and $\epsilon M_n \in \mathcal{B}_2$ for all $n \le N$ (with N to be determined later) provided $\epsilon > 0$ is sufficiently small. Here L_0 and M_0 are as in (4.14) and $\psi_n = \psi - \sum_{j=1}^{n-1} \xi_j$. This follows from the following estimates:

$$\begin{aligned}&C_{M_n} \le C_{M_{n-1}}/2 \le 2^{-n} C_{M_0},\\ &C_{B_n} \le \sum_{j=1}^{n} \|L_j - L_{j-1}\|_{\psi_j} \le \epsilon \sum_{j=0}^{n-1} C_{M_j} g_j \le 2\epsilon C_{M_0} \max_{j=0,\ldots n-1} g_j,\end{aligned}$$

where $g_j := g(C_{B_j}, \epsilon C_{M_j})$. This shows that c_n is bounded for all n if ϵ is small enough. Let $c_* = \sup_{n=0,\ldots N-1} c_n$, define $\xi_n = 2c_*\epsilon$ and let N be the largest number such that

$$\psi - N\xi = \psi - 2Nc_*\epsilon > 0,$$

so $N := \lfloor \frac{\psi}{2c_*\epsilon} \rfloor$, the largest integer $\le \frac{\psi}{2c_*\epsilon}$. Then the norm of $R^E(\phi) := \epsilon M_N(\phi)$ is bounded by $\epsilon C_{M_N} \le \epsilon 2^{-\lfloor \frac{\psi}{2c_*\epsilon} \rfloor} C_{M_0}$. We then set $T^E(\phi) = (T_N \circ \ldots \circ T_1)(\phi)$ and $b_j^E = \omega^E (L_N - L_0)_j/\epsilon$, $j = 1, \ldots d_{\mathcal{Z}}$. □

We can solve (4.15) to obtain

$$\tilde{z}(2\pi) = \tilde{\Psi}^E(2\pi, 0)\tilde{z}(0) = \tilde{\Phi}^E(2\pi, 0)\tilde{z}(0) + \mathcal{O}(\epsilon^{-1}\mathrm{e}^{-\tilde{C}/\epsilon})\tilde{z}(0), \tag{4.18}$$

where

$$\tilde{\Phi}^E(2\pi, 0) = \begin{pmatrix} \cos\left(\alpha^E/\epsilon\right) & \sin\left(\alpha^E/\epsilon\right)\\ -\sin\left(\alpha^E/\epsilon\right) & \cos\left(\alpha^E/\epsilon\right)\end{pmatrix},$$

where

$$\alpha^E = \mathrm{diag}\left(\alpha_1^E, \ldots, \alpha_\ell^E, \ldots, \alpha_{d_{\mathcal{Z}}}^E\right) := \int_0^{2\pi} \mathrm{diag}(\omega + \epsilon b^E(s))\mathrm{d}s/\omega^E. \tag{4.19}$$

The eigenvalues of $\tilde{\Psi}_E(2\pi, 0)$ are therefore

$$\lambda_\ell^E = \exp\left(\mathrm{i}\epsilon^{-1}\alpha_\ell^E\right) + \mathcal{O}(\epsilon^{-1}\mathrm{e}^{-\tilde{C}/\epsilon}), \quad \lambda_{\ell+d_{\mathcal{Z}}}^E = \bar{\lambda}_\ell^E, \tag{4.20}$$

for $\ell = 1, \ldots, d_{\mathcal{Z}}$. We write $\Pi^{\mu,E}(\hat{z}) = \tilde{\Pi}^{\mu,E}(\tilde{z})$ in the coordinates $\hat{z} \to \tilde{z} = \hat{z} + \epsilon T^E(\phi)\hat{z}$ from Lemma 4.5 using variations of constants in (4.12):

$$\tilde{\Pi}^{\mu,E}(\tilde{z}) = \tilde{\Phi}^E(2\pi, 0)\tilde{z} + \tilde{\rho}^E(2\pi, \tilde{z}),$$

where

$$\tilde{\rho}^E(\phi, \tilde{z}_0) = \int_0^\phi \tilde{\Phi}^E(\phi, s)(J_{\mathcal{Z}}\tilde{r}^E(s) + \epsilon^{-1}R^E(s)\tilde{z}(s) + J_{\mathcal{Z}}\tilde{F}^E(s, \tilde{z}(s))\mathrm{d}s$$

with $\tilde{z}(0) = \tilde{z}_0$ and

$$\tilde{r}^E(\phi) = (\mathrm{id}{+}\epsilon T^E(\phi))\hat{r}^E(\phi)/\epsilon, \quad \tilde{F}^E(\phi, \tilde{z}) = (\mathrm{id}{+}\epsilon T^E(\phi))\hat{F}^E(\phi, (\mathrm{id}{+}\epsilon T^E(\phi))^{-1}\hat{z})/\epsilon.$$

Choosing $\epsilon > 0$ small enough such that $\epsilon\|T^E\| \leq 1/2$ we obtain that $\tilde{F}^E(\phi, \cdot)$ is analytic on $\mathcal{B}_\sigma^{\mathcal{Z}}(0)$. Due to (4.13) we have

$$\|\tilde{F}^E\|_\sigma := \sup_{\substack{\|z\|\leq\sigma \\ \phi\in[0,2\pi]}} \|\tilde{F}^E(\phi, z)\| = \mathcal{O}(\epsilon^{-1}\mu), \quad \|\tilde{r}^E\| = \sup_{\phi\in[0,2\pi]} \|\tilde{r}^E(\phi)\| = \mathcal{O}(\epsilon^{-1}\mu),$$

uniformly for $E \in (E_1, E_2)$. and therefore also

$$\|\tilde{\rho}^E\|_{\tilde{\sigma}} = \mathcal{O}(\epsilon^{-1}\mu) \tag{4.21}$$

uniformly in $E \in (E_1, E_2)$, where we redefine $\mu^2 = \mathrm{e}^{-C/\epsilon}$, with $C \leq 2\tilde{C}$ and set $0 < \tilde{\sigma} < \sigma$ such that $\sigma - \tilde{\sigma} > \mathcal{O}(\epsilon^{-1}\mu)$. Whenever $(\mathrm{id} - \Phi^E(2\pi, 0))^{-1}$ exists the fixed points of $\tilde{\Pi}^{\mu,E}$ satisfy

$$\tilde{z} = \pi^{\mu,E}(\tilde{z}) := (\mathrm{id} - \tilde{\Phi}^E(2\pi, 0))^{-1}\tilde{\rho}^E(\phi, \tilde{z}). \tag{4.22}$$

Then we have:

Theorem 4.6. *Under assumptions (H0-H3), (I1-I3) for any $\epsilon > 0$ sufficiently small there is a two-dimensional manifold $\mathcal{M}_\epsilon$ of non-degenerate periodic orbits parametrized by energy $E \in (E_1, E_2) \setminus I$ where I is a union of $\mathcal{O}(\epsilon^{-1})$-many intervals, and the measure $|I|$ of I is exponentially small: $|I| = \mathcal{O}(\mathrm{e}^{-c/\epsilon})$ for some $c > 0$.*

Proof. Let $E = E_0$ be a bifurcation value, i.e., $\Phi^{E_0}(2\pi, 0)$ has at least two eigenvalues which are 1 (they come in pairs). Then, for $\epsilon > 0$ sufficiently small the following condition is satisfied:

$$\partial_E \alpha_\ell^E|_{E=E_0} \neq 0, \tag{4.23}$$

for all ℓ with $\lambda_\ell^{E_0} = 1$. Note that (4.23) follows from

$$\partial_E \omega^E|_{E=E_0} \neq 0,$$

cf. (4.9), for ϵ small which is guaranteed by the condition $\tau'(E) \neq 0$ on the family of periodic orbits from (I1). Let $\mathbb{P}_\ell$ be the projection to the space spanned by e_ℓ and $e_{\ell+d_{\mathcal{Z}}}$. Then $\mathbb{P}_\ell(\tilde{\Psi}^{E_0}(2\pi, 0) - \mathrm{id}) = 0$, by assumption, and with (4.18) we get

$$\begin{aligned}\mathbb{P}_\ell(\tilde{\Phi}^E(2\pi, 0) - \mathrm{id}) &= \begin{pmatrix} \cos\left(\alpha_\ell^E/\epsilon\right) & \sin\left(\alpha_\ell^E/\epsilon\right) \\ -\sin\left(\alpha_\ell^E/\epsilon\right) & \cos\left(\alpha_\ell^E/\epsilon\right) \end{pmatrix} - \mathrm{id}_{\mathbb{R}^2} \\ &= \epsilon^{-1}\partial_E \alpha_\ell^E|_{E=E_0}(E - E_0)J + \mathcal{O}((E - E_0)^2/\epsilon) + \mathcal{O}(\epsilon^{-1}\mu),\end{aligned}$$

where J is the standard $(2, 2)$ symplectic matrix. Due to (4.23) this implies that there is some $c_{\Phi,\ell} > 0$ such that for $\sqrt{\mu} \leq |E - E_0| \ll \epsilon$

$$\|\mathbb{P}_\ell(\tilde{\Phi}^E(2\pi, 0) - \mathrm{id}_{\mathbb{R}^2})\| \geq c_{\Phi,\ell}|E - E_0|/\epsilon.$$

Therefore for such E we get

$$\|(\tilde{\Phi}^E(2\pi,0)-\mathrm{id})^{-1}\| = \mathcal{O}(|E-E_0|^{-1}\epsilon).$$

Let us, for example, assume that

$$|E-E_0| \geq \sqrt{\mu}. \tag{4.24}$$

Then

$$\|(\tilde{\Phi}^E(2\pi,0)-\mathrm{id})^{-1}\| = \mathcal{O}(\epsilon/\sqrt{\mu}).$$

Due to (4.21) and (4.22) we see that $\tilde{\pi}_{\mu,E}$ from (4.22) maps $\mathcal{B}^{\mathcal{Z}}_{\tilde{\sigma}-\kappa}(0)$ to itself where $0<\kappa<\tilde{\sigma}$ provided $\epsilon>0$ is small enough such that $\mathcal{O}(\sqrt{\mu})<\tilde{\sigma}-\kappa$. We use a Cauchy estimate to obtain

$$\|\partial_{\tilde{z}}\tilde{\rho}^E\|_{\sigma-\kappa} \leq \|\partial_{\tilde{z}}\tilde{\rho}^E\|_{\sigma}/\kappa = \mathcal{O}(\epsilon^{-1}\mu)$$

uniformly in $E\in(E_1,E_2)$. We therefore conclude that the contraction mapping theorem applies to (4.22) for energy values E near E_0 satisfying (4.24). Next note that due to (4.23) for any fixed $\epsilon>0$ and every $\ell=1,\ldots,d_{\mathcal{Z}}$ there are $\mathcal{O}(\epsilon^{-1})$ many E values such that $\lambda_\ell(E)=1$. Therefore, since $d_{\mathcal{Z}}<\infty$, there are $\mathcal{O}(\epsilon^{-1})$ many critical E values; if we exclude an interval of length $\sqrt{\mu}$ around each of them then we can guarantee the persistence of periodic orbits for energy values on a complement of this set. □

Appendix A. Exponential accurate slow manifolds in general systems

In this appendix we consider the following system

$$\dot{w} = \epsilon W_0(w,z_0),\quad \dot{z}_0 = Z_0(w,z_0) = r_0(w) + Lz_0 + a_0(w)z_0 + F_0(w,z_0). \tag{A.1}$$

with $(w,z_0)\in(\mathcal{V}+\mathrm{i}\nu_0)\times(\mathcal{S}+\mathrm{i}\sigma_0)$ and where $\mathcal{V}$ and $\mathcal{S}$ are bounded and open sets in $\mathcal{W}=\mathbb{R}^{n_{\mathcal{W}}}$ and the Banach space $\mathcal{Z}$ respectively, and, as before, $\mathcal{S}$ is a neighbourhood of 0. Here the z equation is a semilinear evolution equation and the slow vector field W_0 is bounded as detailed below. We assume the following:

(G0) L is a densely defined closed operator which either generates a strongly continuous semigroup or an analytic semigroup.

In the following we set $\alpha=0$ in case L generates a strongly continuous semigroup and assume $\alpha\in[0,1)$ otherwise. If L generates an analytical semigroup let $\lambda_0\in\mathbb{R}$ be in the resolvent set of L and let λ_0 such that $\|(\lambda_0+L)^{-1}\|\leq 1$. Note that this is possible because $-L$ is sectorial and so there is some $M_L>0$, $\lambda_0\in\mathbb{R}$ and a sector $S=\{\lambda\in\mathbb{C};|\arg(\lambda-\lambda_0)|<\phi\}$, where $\phi<\pi/2$ such that any $\lambda\in\mathbb{C}\setminus S$ is in the resolvent set of $-L$ and satisfies $\|(\lambda+L)^{-1}\|\leq M_L/|\lambda+\lambda_0|$ [12]. Define the Banach space $\mathcal{Z}_\alpha := D((\lambda_0+L)^\alpha)$ with norm

$$|||z||| := \|z\|_{\mathcal{Z}_\alpha} := \|(\lambda_0+L)^\alpha z\|.$$

Then by construction $|||z|||\leq\|z\|$. We assume that $\mathcal{S}+\mathrm{i}\sigma_0\subseteq\mathcal{Z}_\alpha$. Similarly as before for a map $A\in\mathcal{E}(\mathcal{Z};\mathcal{Z}_\alpha)$ we define $\|A|||:=\|A\|_{\mathcal{E}(\mathcal{Z};\mathcal{Z}_\alpha)}$ and for a map $F:(\mathcal{W}+\mathrm{i}\nu_0)\times(\mathcal{S}+\mathrm{i}\sigma_0)\to\mathcal{Z}$ we define

$$|||F\|_{\nu_0,\sigma_0} = \sup_{\substack{w\in\mathcal{W}+\mathrm{i}\nu_0\\ w\in\mathcal{S}+\mathrm{i}\sigma_0}} \|F(w,z)\|.$$

Similarly we define $|||W_0\|_{\nu_0,\sigma_0}$ for a map $W:(\mathcal{W}+\mathrm{i}\nu_0)\times(\mathcal{S}+\mathrm{i}\sigma_0)\to\mathcal{W}$.

(G1) The functions $r_0 : (\mathcal{V} + \mathrm{i}\nu_0) \to \mathcal{Z}^{\mathbb{C}}$, $a_0 : (\mathcal{V} + \mathrm{i}\nu_0) \to \mathcal{E}(\mathcal{Z}_\alpha^{\mathbb{C}}; \mathcal{Z}^{\mathbb{C}})$, $F_0 : (\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \to \mathcal{Z}^{\mathbb{C}}$ and $W_0 : (\mathcal{V} + \mathrm{i}\nu_0) \times (\mathcal{S} + \mathrm{i}\sigma_0) \to \mathcal{W}^{\mathbb{C}}$ are real analytic and uniformly bounded by $\delta_0 = \|r_0\|_{\nu_0}$, $C_{a_0} = |||a_0\|_{\nu_0}$, $C_{F_0} = |||F_0\|_{\nu_0,\sigma_0}$, and $C_{W_0} = |||W_0\|_{\nu_0,\sigma_0}$. Here, as before, $F_0(z_0) = \mathcal{O}(|||z_0|||^2)$. Moreover, $C'_{W_0} = |||\partial_z W_0\|_{\nu_0,\sigma_0}$, $C''_{W_0} = |||\partial_z^2 W_0\|_{\nu_0,\sigma_0}$.

(G2) The operator $(L + a_0(\cdot))^{-1} : (\mathcal{V} + \mathrm{i}\nu_0) \to \mathcal{E}(\mathcal{Z}^{\mathbb{C}}, \mathcal{Z}_\alpha^{\mathbb{C}})$ is real analytic with

$$\|(L + a_0(\cdot))^{-1}|||_{\nu_0} \leq \frac{K_0}{2}.$$

Note that (G2) is true for some $K_0 > 0$ if if $|||a_0 - \lambda_0\|_{\nu_0}$ is sufficiently small. We then have the following result:

Theorem A.1. *Assume (G0-G2). Let $\sigma_0 > \sigma > 0$, $\nu_0 > \nu > 0$. Then for $\delta_0 \geq 0$ and $\epsilon > 0$ sufficiently small the following holds true: There exists a transformation of the fast variables $z_0 = \zeta(w) + z$, $(w, z) \in (\mathcal{V} + \mathrm{i}\nu) \times (\mathcal{S} + \mathrm{i}\sigma)$, $z_0 \in \mathcal{S} + \mathrm{i}\sigma_0$, with $|||\zeta|||_\nu = \mathcal{O}(\epsilon)$ so that*

$$\dot{z} = r(w) + \mathcal{O}(z),$$

where

$$\|r(w)\| \leq C_1 \|W_0(w, \zeta(w))\| e^{-C_2/\epsilon}.$$

Here C_1 and C_2 are positive constants which depend solely on C_{a_0}, K_0, C_{F_0}, $C_{\mathcal{S}}$, σ_0, σ, ν_0, ν, C_{W_0}, C'_{W_0} and C''_{W_0}.

In other words: $\{z = 0\}$ is an almost invariant slow manifold that contains all equilibria of (A.1) near $\{z_0 = 0\}$. This result was proved by Neishtadt in the case that $d_{\mathcal{Z}} < \infty$ and that L is bounded, using a slighly different iterative step, as outlined in the introduction, Section 1. The advantage of MacKay's method which we use in the proof is that the slow manifold we construct contains all nearby equilibria.

For the proof of the Theorem A.1 we need the following notion: For $R > 0$ such that $\mathcal{B}_R^{\mathcal{Z}_\alpha^{\mathbb{C}}}(0) \subseteq \mathcal{S} + \mathrm{i}\sigma_0$ we define $C''_F[\nu_0, R]$ as bound of

$$\sup_{\substack{w \in \mathcal{W} + \mathrm{i}\nu_0 \\ |||z||| \leq R}} \|\partial_z^2 F(w, z)\|_{\mathcal{E}(\mathcal{Z}_\alpha \times \mathcal{Z}_\alpha; \mathcal{Z})} \leq C''_F[\nu_0, R].$$

We need the following modification of Lemma 3.4 which is straightforward to prove:

Lemma A.2. *Assume (G0-G2), with the subscript dropped. Moreover assume that*

$$\delta < \min(K\sigma, 2/(K^2 C''_F[\nu, K\delta]). \tag{A.2}$$

Then

$$0 = r(w) + Lz + a(w)z + F(z, w) \tag{A.3}$$

has a locally unique solution $z = \zeta(w) \in \mathcal{Z}_\alpha^{\mathbb{C}}$ satisfying:

$$|||\zeta(w)||| \leq K\|r(w)\|, \tag{A.4}$$

for every $w \in \mathcal{V} + \mathrm{i}\nu$. Moreover $\zeta \in C^\omega(\mathcal{V} + \mathrm{i}\nu; \mathcal{Z}_\alpha^{\mathbb{C}})$.

Next we set up an iterative lemma.

Lemma A.3. *(**The Iterative Lemma**) Assume (G0-G2) with the subscript dropped and assume* (A.2)*. Let $\zeta=\zeta(w)$ be the solution from Lemma A.2. Let $\nu_+=\nu-\xi>0$ and $\sigma_+=\sigma-\xi\geq\kappa>0$. If*

$$2K\max(\epsilon C'_W,\delta)\leq\xi\leq\sigma \tag{A.5}$$

the map $z=\zeta(w)+z_+$, $(w,z_+)\in(\mathcal{V}+\mathrm{i}\nu_+)\times(\mathcal{S}+\mathrm{i}\sigma_+)$ from Lemma 3.4 transforms (A.1) *into*

$$\dot{w}=\epsilon W_+(w,z_+),\quad \dot{z}_+=r_+(w)+Lz+a_+(w)z_++F_+(w,z_+), \tag{A.6}$$

and there is a constant c which is continuous and increasing in $1/\kappa$, C_W, C'_W, C''_W, K, C_F and $C_{\mathcal{S}}=|||z|||_\sigma$ and depends on those constants only such that

$$\|r_+(w)\|\leq\frac{K\epsilon}{\xi}\|W_+(w,0)\|\delta\leq\frac{\epsilon c\delta}{\xi}\quad \textit{for}\quad w\in\mathcal{V}+\mathrm{i}\nu_+, \tag{A.7a}$$

$$|||a_+-a\|_{\nu_+}\leq c\delta, \tag{A.7b}$$

$$\|(L+a_+(w))^{-1}|||_{\nu_+}\leq\frac{K_+}{2}:=\frac{K}{2}+c\delta, \tag{A.7c}$$

$$|||F_+-F\|_{\nu_+,\sigma_+}\leq\frac{c\delta}{\xi} \tag{A.7d}$$

provided that $c\delta<1$.

Proof. The existence of $\zeta(w)$ follows from Lemma A.2 and yields

$$\begin{aligned}W_+(w,z_+)&=W(w,\zeta+z_+),\\ r_+(w)&=-\epsilon\partial_w\zeta(w)W(w,\zeta(w)),\\ a_+(w)&=a(w)-\epsilon\partial_w\zeta(w)\partial_zW(w,\zeta(w))+\partial_zF(w,\zeta(w)),\end{aligned} \tag{A.8}$$

and

$$F_+(w,z_+)=-\epsilon\partial_w\zeta(w)\int_0^1(1-s)\partial_z^2W(w,\zeta(w)+sz_+)\mathrm{d}s z_+^2+\tilde{F}(w,z_+), \tag{A.9}$$

where

$$\tilde{F}(w,z_+)=F(w,\zeta(w)+z_+)-F(w,\zeta(w))-\partial_zF(w,\zeta(w))z_+. \tag{A.10}$$

Using Lemma A.2 and a Cauchy estimate give

$$|||\partial_w\zeta|||_{\nu-\xi}\leq\frac{|||\zeta|||_\nu}{\xi}\leq\frac{K\delta}{\xi}, \tag{A.11}$$

for $\nu-\xi>0$ and $\xi>0$. Hence (A.8) directly gives (A.7a). To estimate a_+-a we first note that by (A.11)

$$\|\partial_w\zeta\partial_zW(w,\zeta)\|_{\mathcal{E}(\mathcal{Z}_\alpha)}\leq C'_W\frac{|||\zeta|||_\nu}{\xi}$$

for $w\in\mathcal{V}+\mathrm{i}(\nu-\xi)$. Moreover, since F is quadratic, by a Cauchy estimate,

$$|||\partial_zF(w,\zeta)\|=|||\int_0^1\partial_z^2F(w,s\zeta)\zeta\mathrm{d}s\|\leq 2C_F\kappa^{-2}|||\zeta|||_\nu,$$

for all $w\in\mathcal{V}+\mathrm{i}(\nu-\xi)$, and therefore for all such w, as $|||\zeta|||_\nu\leq K\delta$ by Lemma A.2,

$$|||a_+-a\|_{\nu-\xi}\leq C'_W\frac{\epsilon K}{\xi}\delta+2C_F\kappa^{-2}K\delta\leq\left(\frac{1}{2}+2C_F\kappa^{-2}K\right)\delta=c\delta,$$

where we have used (A.5). This proves (A.7b).

From

$$(L+a_+)^{-1} = (L+a)^{-1}(\mathrm{id} + (a_+ - a)(L+a)^{-1})^{-1}$$

we get

$$\begin{aligned}\|(L+a_+(w))^{-1}\|| &\le \|(L+a)^{-1}\||(1-\|(a_+-a)(L+a)^{-1}\|)^{-1}\\ &\le \frac{K}{2}\left(1-\frac{K}{2}c\delta\right)^{-1} \le \frac{K}{2}(1+cK\delta).\end{aligned}$$

for $Kc\delta < 1$. Here we have used that $(1-x)^{-1} \le 1+2x$ if $0 \le x \le \frac{1}{2}$ with $x = \frac{K}{2}c\delta$. Redefining c to $\max(1, K^2)c$ proves (A.7c) for $c\delta < 1$.

For $F_+ - F$, with F_+ from (A.9) we first estimate $\tilde{F} - F$ from (A.10). This gives:

$$\begin{aligned}|||\tilde{F} - F\|_{\nu-\xi,\sigma-\xi} &\le |||F(w,\zeta+z_+) - F(w,z_+)\|_{\nu-\xi,\sigma-\xi} + |||F(w,\zeta)\|_{\nu-\xi,\sigma-\xi}\\ &\quad + |||\partial_z F(w,\zeta)z_+\|_{\nu-\xi,\sigma-\xi}\\ &\le |||\partial_z F\|_{\nu,\sigma-\xi/2}|||\zeta|||_\nu + \left\| \int_0^1 (1-t)\partial_z^2 F(w,t\zeta(w))\zeta^2(w)\mathrm{d}t\right\|_{\nu-\xi}\\ &\quad + \left|\left|\left| \int_0^1 \partial_z^2 F(w,t\zeta(w))\zeta(w)z_+\mathrm{d}t\right\|\right.\right._{\nu-\xi,\sigma-\xi}\\ &\le |||\partial_z F\|_{\nu,\sigma-\xi/2}|||\zeta|||_\nu + \frac{1}{2}|||\partial_z^2 F\|_{\nu,\kappa}|||\zeta|||_\nu^2 + |||\partial_z^2 F\|_{\nu,\kappa}|||\zeta|||_\nu C_{\mathcal{S}}\\ &\le C_F\left(\frac{2K\delta}{\xi} + \kappa^{-2}K\delta\,(K\delta + 2C_{\mathcal{S}})\right).\end{aligned}$$

Here we have used that $|||\zeta|||_\nu \le K\delta$ by Lemma A.2 and that $\xi \ge 2K\delta$, see (A.5). Therefore

$$\begin{aligned}|||F_+ - F\|_{\nu-\xi,\sigma-\xi} &\le \frac{1}{2}\epsilon\|\partial_w\zeta\|_{\nu-\xi}\|\partial_z^2 W\|_{\nu,\sigma-\xi/2}C_{\mathcal{S}}^2 + |||\tilde{F} - F\|_{\nu-\xi,\sigma-\xi}\\ &\le \frac{\epsilon K C_{\mathcal{S}}^2\delta C_W''}{2\xi} + C_F\left(\frac{2K\delta}{\xi} + \kappa^{-2}K\delta\,(K\delta + 2C_{\mathcal{S}})\right) \le \frac{c\delta}{\xi}\end{aligned}$$

for a suitable choice of c with the given properties. Here we used (A.5). □

Proof of Theorem A.1. Let $\xi_0 = \xi_1 = \frac{1}{4}\min\{\nu_0 - \nu, \sigma_0 - \sigma\}$. For sufficiently small δ_0 and ϵ we can satisfy the conditions (A.5) and (A.2) of the iterative Lemma A.3 applied to (A.1) with ξ_0 as above to obtain that $\delta_1 = \mathcal{O}(\epsilon)$. For all successive iterations we use the bound $C''_{F_n}[\nu_n, K_n\delta_n] = 2C_{F_n}/\kappa^2$ for $\sigma_n - \kappa > K_n\delta_n$ with $\kappa \ge \sigma > 0$ which changes (A.2) to the condition

$$\delta_n < \min(K_n\sigma_n, \kappa^2 K_n^{-2} C_{F_n}^{-1}). \tag{A.12}$$

Here, as before K_n, δ_n etc. denote the constants of Lemma A.3 after n iterations, i.e., $K_n/2$ is the upper bound in (A.7c) of the operator $(L+a_n)^{-1}$ on $\mathcal{V}+\mathrm{i}\nu_n$ where $\nu_n = \nu_2 - \sum_{k=1}^{n-1}\xi_k > 0$. Furthermore C_{F_n} is the norm of F_n on $(\mathcal{V}+\mathrm{i}\nu_n)\times(\mathcal{S}+\mathrm{i}\sigma_n)$ and $\sigma_n = \sigma_2 - \sum_{k=1}^{n-1}\xi_k > 0$ and C_{a_n} is defined analogously.

Since $\delta_1 = \mathcal{O}(\epsilon)$ we can satisfy (A.5) and (A.2) by choosing for sufficiently small ϵ and therefore can apply Lemma A.3 again to obtain $\delta_2 = \mathcal{O}(\epsilon^2)$, cf. (A.7a).

Applying the Iterative Lemma A.3 successively we have

$$\max(C_{a_{n+1}} - C_{a_n}, \tfrac{1}{2}(K_{n+1} - K_n), C_{F_{n+1}} - C_{F_n}) \le c_n\delta_n/\xi_n, \tag{A.13}$$

and

$$\delta_{n+1} \leq c_n \delta_n \epsilon / \xi_n,$$

where we choose $\xi_n \leq 1$. Taking $\xi_n = 2c_n\epsilon > 0$ we get

$$\delta_{n+1} \leq 2^{-1}\delta_n \leq 2^{-n}\delta_2 \leq 2^{-n}C\epsilon^2.$$

Then from (A.13) we get

$$\max(\tfrac{1}{2}(K_n - K_2), C_{a_n} - C_{a_2}, C_{F_n} - C_{F_2}) \leq \sum_{n=2}^{N} \delta_n/\epsilon \leq 2C\epsilon.$$

by the geometric series formula. Since c_n is a continuous and increasing function of C_{F_n}, C_{a_n}, K_n and $1/\kappa$ and these are bounded in n, there is some c_* such that $c_n \leq c_*$ for all $n = 1, \ldots, N$. From the requirement

$$\sigma \leq \sigma_2 - 2c_*(N-1) = \sigma_{N+1}$$

we can take N to be as $N = \lceil \frac{\sigma_0 - \sigma}{4c_*\epsilon} \rceil$ which concludes the proof. □

The following example shows that the error $\|r_n(w)\|$ in the slow manifold after n steps does not behave like $\mathcal{O}((\epsilon\|W(w,\cdot)\|)^n)$ in general, as conjectured by MacKay [24], see the discussion in the introduction.

Example A.4. We consider the simple linear, two-dimensional example:

$$\dot{w} = \epsilon W(w,z) = \epsilon w, \quad \dot{z} = Z(w,z) = \epsilon w - z. \tag{A.14}$$

Here $z = 0$ is actually normally hyperbolic and there is an invariant slow manifold nearby:

$$z = \frac{\epsilon}{1+\epsilon} w. \tag{A.15}$$

Notice that $(w,z) = (0,0)$ is an equilibrium (saddle for $\epsilon > 0$). Applying MacKay's method n times to this example gives

$$z_n = 0 \quad \text{where} \quad z = z_n + \sum_{k=0}^{n} \zeta_k(w) = \sum_{k=1}^{n} (-1)^k \epsilon^{k+1} w, \tag{A.16}$$

as an approximately invariant slow manifold. In this case the approximation (A.16) also coincide with the nth degree Taylor polynomial of (A.15). The error field is $r_n(w) = \zeta_n(w) = (-1)^n\epsilon^{n+1}w$ which directly illustrates why MacKay's conjecture is incorrect. For a nonlinear example, one may replace $W(w,z) = w$ by $\mathscr{W}(w)$ satisfying $\mathscr{W}(0) = 0$, $\mathscr{W}'(0) \neq 0$. Then $(w,z) = (0,0)$ is still a hyperbolic equilibrium, and we have $r_1(w) = -\epsilon^2\mathscr{W}(w)$ and $r_2(w) = \epsilon^3\mathscr{W}'(w)\mathscr{W}(w)$ which cannot be bounded above from above by an expression with $|\mathscr{W}(w)|^2$ as a factor.

References

[1] Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. 2nd edn. Academic Press, 2003.

[2] C.J. Amick and K. Kirchgasser. A theory of solitary water-waves in the presence of surface tension. *Arch. Rational Mech. Anal.* 105:1–49, 1989.

[3] R. Bellmann. The stability of solutions of linear differential equations. 10:643–647, 1981.

[4] C. J. Cotter and S. Reich. Adiabatic invariance and applications to MD and NWP. *BIT Numerical Mathematics*, 3:439–455, 2003.

[5] A. Fehrst. Structure and Mechanisms. *Freeman, New York*, 2nd Edition, 1975.

[6] N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21:193–226, 1971.

[7] N. Fenichel. Asymptotic stability with rate conditions. *Indiana University Mathematics Journal*, 23:1109–1137, 1974.
[8] S.J. Fraser. The steady state and equilibrium approximations: a geometrical picture. *Journal of Chemical Physics*, 88:4732–4738, 1990.
[9] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A.M. Zagaris. Projecting to a slow manifold: Singularly perturbed systems and legacy codes. *SIAM Journal of Applied Dynamical Systems*, 4(3):711–732, 2005.
[10] V. Gelfreich and L. Lerman. Almost invariant elliptic manifold in a singularly perturbed Hamiltonian system. *Nonlinearity*, 15:447–557, 2002.
[11] V. Gelfreich and L. Lerman. Long-periodic orbits and invariant tori in a singularly perturbed Hamiltonian system. *Physica D*, 176:pp 125–146, 2003.
[12] D. Henry. *Geometric Theory of Semilinear Parabolic Equations*. Lecture Notes in Mathematics. Springer Verlag, Berlin, Heidelberg, New York, 1981.
[13] H. G. Kaper and T. J. Kaper. Asymptotic analysis of two reduction methods for systems of chemical reactions. *Physica D*, 165:66–93, 2002.
[14] K.U. Kristiansen, P. Palmer, and R. M. Roberts. A unification of models of tethered satellites. *SIAM Journal of Applied Dynamical Systems*, 10:1042–1069, 2011.
[15] K.U. Kristiansen, P. Palmer, and R. M. Roberts. The persistence of a slow manifold with bifurcation. *SIAM Journal of Applied Dynamical Systems*, 11:661–683, 2012.
[16] K.U. Kristiansen. Periodic orbits near a bifurcating slow manifold. *Journal of Differential Equations*, 259:4561–4614, 2015.
[17] M. Krupa, and P. Szmolyan. Extending geometric singular perturbation theory to nonhyperbolic points - Fold and canard points in two dimensions. *SIAM Journal on Mathematical Analysis*, 33: 286–314, 2001.
[18] J. Laskar. Large scale chaos in the solar system. *Astronomy and Astrophysics*, 287, 1994.
[19] J. Laskar and M. Gastineau. Existence of collisional trajectories of Mercury, Mars and Venus with the Earth. *Nature*, 459, 11 June 2009.
[20] E. N. Lorenz. The slow manifold - what is it? *American Meteorological Society*, 15 December, 1992.
[21] E. N. Lorenz. Existence of a slow manifold. *Journal of the Atmospheric Sciences*, 43(15):1547–1557, 1986.
[22] N. Lu. Small generalized breathers with exponentially small tails for KleinGordon equations. *J. Differential EquationsJ*, 256: 745–770, 2014.
[23] U. Maas, R.W. Dibble, J. Warnatz, and E. Zwicker. Combustion: Physical and Chemical Fundamentals, Modeling and Simulation, Experiments, Pollutant Formation, *Springer, Berlin*, 2nd Edition, 1999.
[24] R. S. MacKay. Slow manifolds. *In: "Energy Localisation and Transfer", eds T Dauxois, A Litvak-Hinenzon, RS MacKay, A Spanoudaki, World Scientific*, pp. 149–192, 2004.
[25] R. S. MacKay. Effective Hamiltonian for travelling discrete breathers. *Journal of Physics A: Mathematical and General*, 35:39–85, 2002.
[26] R. Mañé. Persistent manifolds are normally hyperbolic. *Bulletin of American Mathematical Society*, 80:90–91, 1980.
[27] K. Matthies and A. Scheel. Exponential averaging for Hamiltonian evolution equations. *Transactions of the American Mathematical Society*, 355(2):747–773, 2002.
[28] D. A. McQuarrie. *Physical Chemistry: A Molecular Approach*. Sausalito: University Science Books, 1997.
[29] L. Michaelis and M. Menten. Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift*, 49:333–369, 1913.
[30] J. D. Murray. *Mathematical Biology*, volume 19. Berlin: Springer-Verlag, 1993.
[31] A. Neishtadt. On the accuracy of conservation of adiabatic invariant. *Journal of Applied Mathematics and Mechanics*, 45(1):58–63, 1982.
[32] A. Neishtadt. Persistence of stability loss for dynamical bifurcation, i. *Journal of Differential Equations*, 23:1385–1390, 1987.
[33] J. Pöschel and E. Trubowitz. *Inverse Spectral Theory*, volume 130. Pure and Applied Mathematics, Academic Press, Inc., 1987.
[34] V. Rom-Kedar and D. Turaev. The symmetric parabolic resonance instability. *Nonlinearity*, 23:1325–1351, 2010.

[35] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
[36] M. Reed, B. Simon, *Methods of mathematical physics I: Functional Analysis*, Academic Press, London, 1980.
[37] M.R. Roussel, S.J. Fraser, Geometry of the steady-state approximation: perturbation and accelerated convergence methods, *J. Chem. Phys.* 93:1072-1081, 1990.
[38] F. Takens. Motion under the influence of a strong constraining force. In: Global Theory of Dynamical Systems, Evanston 1979. Z. Nitecki and C. Robinson, eds., Springer, New York, pp. 425–455, 1980.
[39] R. Temam. Inertial Manifolds. *The Mathematical Intelligencer*, 12(4):68–74, 1990.
[40] J. Vanneste. Asymptotics of a slow manifold. *SIAM Journal of Applied Dynamical Systems*, 7, 2008.
[41] D. Wirosoetisno. Exponentially accurate balance dynamics. *Advances in Differential Equations*, 9, 2004.
[42] A.M. Zagaris, C. W. Gear, T. J. Kaper, and I. G. Kevrekidis. Analysis of the accuracy and convergence of equation-free projection to a slow manifold. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43:757–784, 2009.