LONG-TIME ACCURACY FOR APPROXIMATE SLOW MANIFOLDS IN A FINITE DIMENSIONAL MODEL OF BALANCE

GEORG GOTTWALD, MARCEL OLIVER, AND NICOLAE TECU

Abstract. We study the slow singular limit for planar anharmonic oscillatory motion of a charged particle under the influence of a perpendicular magnetic field when the mass of the particle goes to zero. This model has been used by the authors as a toy model for exploring variational high order approximations to the slow dynamics in rotating fluids. In this paper, we address the long time validity of the slow limit equations in the simplest nontrivial case. We show that the first order reduced model remains $O(\varepsilon)$ accurate over a long $1/\varepsilon$ time scale. The proof is elementary, but involves subtle estimates on the nonautonomous linearized dynamics.

1. Introduction

We study the zero mass limit for the equations of a single charged particle in a planar anharmonic oscillator potential under the influence of an external magnetic field. The equations of motion are

$$\varepsilon\,\ddot{q} - J\dot{q} + \nabla V(q) = 0\,, \tag{1}$$

where $q \colon \mathbb{R} \to \mathbb{R}^2$ and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{2}$$

denotes the canonical symplectic matrix. This system is variational and can be obtained as the Euler–Lagrange equations associated with Lagrangian

$$L_\varepsilon = \frac{\varepsilon}{2}\,\dot{q}^T\dot{q} - V(q) - \frac{1}{2}\,\dot{q}^T Jq\,. \tag{3}$$

System (1) can also be viewed as a simple model of balance in rotating flow as exemplified by the rotating shallow water equations

$$\frac{Du}{Dt} + f_0\,Ju + g\,\nabla h = 0\,, \tag{4a}$$

$$\frac{Dh}{Dt} + h\,\nabla\cdot u = 0\,. \tag{4b}$$

Here $u = u(x,t)$ denotes the velocity field on the horizontal plane $x \in \mathbb{R}^2$, $h = h(x,t)$ the layer depth, ∇ the horizontal gradient, $\mathrm{D}/\mathrm{D}t = \partial_t + u\cdot\nabla$ the material derivative, g the constant of gravity, and $f_0/2$ the ambient angular velocity in the so-called f-plane approximation. To nondimensionalize the equations, let U be a

Date: November 29, 2006.

typical velocity, L be a typical horizontal length scale, and H a typical layer depth. Then, in new, nondimensional variables,

$$\varepsilon \frac{\mathrm{D}u}{\mathrm{D}t} + Ju - \frac{B}{\varepsilon} \nabla h = 0 \,, \tag{5a}$$

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h \nabla \cdot u = 0 \,, \tag{5b}$$

where $\varepsilon = U/f_0 L$ is the Rossby number and $B = (L_R/L)^2$ the Burger number with Rossby radius of deformation $L_R = \sqrt{gH}/f_0$. The *semigeostrophic scaling* $B = O(\varepsilon)$ with $\varepsilon \ll 1$ is one of the physically relevant scalings [15, 19]. In this situation, (1) can be regarded as the Lagrangian equation for one fluid parcel with externally prescribed layer depth potential, the magnetic force taking the role of the Coriolis force.

Observed large scale flows, for example the large vortices we know from weather maps, frequently dominate the behavior of the system, and there is a plethora of reduced so-called *balance models* which filter out the fast components of the motion. The present study is motivated by the desire to understand variational constructions of balance models which were initially proposed by Salmon [18] and generalized in [14]. The main advantage of going the variational route is the persistence of proper analogs of conserved quantities of the parent equations under model reduction. However, to our knowledge there is currently no proof that solutions to any of the semigeostrophic limit equations remain close to solutions of the full parent equations. Moreover, it has been numerically observed that balance models, for example the quasi-geostrophic equations [15], appear valid for much longer time scales than expected [5, 11]. For the full shallow water equations, rigorous estimates in the limit of rapid rotation have been obtained by Babin, Mahalov, and Nicolaenko [1] and Embid and Majda [3]. These estimates are challenging as resonant wave interactions have to be controlled; similar difficulties will arise when attempting to generalize our result to higher dimensions. On the other hand, their results are concerned with a time scale which is long with respect to the fast time scale rather than with respect to the slow time scale as in the present paper. Thus, our work here is a first step in understanding the conditions under which balance in the semigeostrophic limit is maintained over very long times.

Our model equation (1) is characterized by motion on two separate time scales. There is fast, nearly harmonic oscillatory motion through a balance between inertia and magnetic terms,

$$\varepsilon \, \ddot{q}_\varepsilon = J \dot{q}_\varepsilon \,, \tag{6}$$

and slow, generically anharmonic motion through a balance between magnetic and potential terms

$$J \dot{q}_\varepsilon = \nabla V(q_\varepsilon) \,. \tag{7}$$

The corresponding frequencies are of order ε^{-1} and 1, respectively. There is no explicit separation of fast and slow subsystems, so that the construction of a slow manifold becomes a nontrivial task.

The general theory and the construction of slow manifolds for normally hyperbolic systems is well developed [4, 20, 23, 24, 25]. However, most Hamiltonian systems, including our model, are not normally hyperbolic. Moreover, in the Hamiltonian case it is often desirable to retain a Hamiltonian vector field on the slow manifold and, consequently, the associated conservation laws. Structure preservation can be achieved by using normal form transformations on the Hamiltonian side or *variational asymptotics* on the Lagrangian side. The slow Hamilton or Euler–Lagrange equations of motion are then derived in a second, non-perturbative step. The construction of slow manifolds by canonical transformations has been addressed, for example, in [22, 28]. A construction by constraining the Lagrangian to a zero order approximation of the slow manifold has been introduced by Whitham [27] and applied to fluid mechanics by Salmon [17]. However, their approach does not generalize directly to higher order; see MacKay [10] for an excellent review. For unbalanced data where the full dynamics has nonnegligible fast contributions, averaging methods have to be used [20, 25].

The present paper is motivated by, though not dependent on, our previous work on generalizing Salmon's [18] constrained Lagrangian construction to higher order. The procedure has been developed both for the shallow water equations [14] and studied for the finite dimensional model (1) in [7]. This new construction provides a natural framework for working with noncanonical symplectic transformations, thereby providing enough generality in an infinite dimensional setting to ensure well-posedness and regularity of the resulting slow dynamics.

In this paper we address the question whether the dynamics on the slow manifold can shadow trajectories of the parent system over very long times. Specifically, we prove that the first order reduced models remain $O(\varepsilon)$ accurate representations of the slow parent dynamics over a long $1/\varepsilon$ time scale.

Related work has been done by a number of authors. Cotter and Reich [2] have used Hamiltonian normal form theory to show that solutions near the slow manifold of system (1) remain close to it for exponentially long times. Their proof is based on an optimal truncation of an asymptotic series, while the present paper aims at proving a long-time path-wise estimate for an explicitly computable slow model at fixed low order of an asymptotic series. Wirosoetisno [26] considers a different finite dimensional model of balance in which the splitting between fast and slow dynamics is explicitly built into the model. For equation (1), this is not the case, so that the problem of identifying the slow dynamics is nonobvious.

The paper is structured as follows. Section 2 introduces the first order limit systems and provides a brief sketch of their derivation. In Section 3, we state our main theorem and sketch the structure of the proof. Sections 4–7 provide the technical details of this proof. We conclude the paper with a short discussion of the result in the broader context of our original motivation.

2. First order slow equations

Throughout the paper, we write ∇ to denote the gradient, viewed as a column vector, and use D to denote the total derivative, viewed as a linear map acting on column vectors. The Hessian of a scalar function ϕ is written $\operatorname{Hess}\phi = \mathrm{D}\nabla\phi$; we write, in particular, $\operatorname{Hess} q$ when we need to distribute over components of a vector-valued function q, i.e.

$$(v^T \operatorname{Hess} q\, u)_i \equiv v^T \operatorname{Hess} q_i\, u \,. \tag{8}$$

Finally, Δ denotes the Laplacian which also distributes over components of vector-valued functions.

We are now ready to introduce the two first-order limit systems which are the main concern of this paper. First, we consider the variational slow equation

$$(1 + \varepsilon\, \Delta V(q))\, J\dot{q} = \nabla V(q) + \varepsilon\, \mathrm{D}\nabla V(q) \nabla V(q) \,. \tag{9}$$

It is derived by inserting a near-identity transformation of the form

$$q_\varepsilon = q + \varepsilon\, q' + \tfrac{1}{2}\,\varepsilon^2\, q'' + \dots \tag{10}$$

into the Lagrangian (3) and truncating to first order in ε. In this setting, q' is a free parameter which can be chosen to render the first order Lagrangian affine. The resulting dynamics is constrained to slow motion in a two-dimensional phase space. Equations (9) arise from the particular choice

$$q' = -\tfrac{1}{2}\, J\dot q + \tfrac{1}{2}\, \nabla V(q) \tag{11}$$

where, as can be seen by diagnosing (11) with the leading order balance $J\dot q = \nabla V(q)$, the q variables remain untransformed up to terms of second order in ε. For details and more general transformations, see [7].

Second, we consider the nonvariational slow equation

$$J\dot q = \nabla V(q) + \varepsilon\, J\mathrm{D}\nabla V(q) J \nabla V(q) \tag{12}$$

which can be derived in the following direct way. (An alternative derivation based on successive integration by parts is given as part of the set-up for our proofs in Section 4.) As in [12], write (1) as

$$\dot q = p\,, \tag{13a}$$

$$\varepsilon\, \dot p = Jp - \nabla V(q)\,, \tag{13b}$$

and introduce a new fast variable $w = p - F_{n+1}(q)$ so that

$$\dot q = w + F_{n+1}(q)\,, \tag{14a}$$

$$\dot w = \left(\frac{1}{\varepsilon}\, J - \mathrm{D}F_{n+1} \right) w + \frac{1}{\varepsilon} \left(JF_{n+1}(q) - \nabla V(q) \right) - \mathrm{D}F_{n+1}(q)\, F_{n+1}(q)\,. \tag{14b}$$

Expanding

$$F_n(q) = \sum_{i=0}^{n} f_i(q)\, \varepsilon^i\,, \tag{15}$$

we can iteratively determine the f_i such that the inhomogeneity in (14b) is of order ε^{n+1}. Left-multiplying (14b) with w yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \|w\| \leq \|\mathrm{D}F_{n+1}\|\, \|w\| + O(\varepsilon^{n+1}) \tag{16}$$

so that, if $w = O(\varepsilon^{n+1})$ initially, it will remain so for times of order one. The dynamics is then dominantly slow and can be approximated to $O(\varepsilon^{n+1})$ by

$$\dot q = F_n(q) \tag{17}$$

for times of order one. Figure 1 shows the projection onto the q_1-q_2 plane of a solution to the parent system initialized on the first-order approximate slow manifold $w = 0$ (dotted line) together with the corresponding first-order limit trajectory (solid line) for two different values of ε. When ε decreases, the approximation of the slow manifold becomes more accurate—evidenced by the decrease in amplitude of the fast components in the solution to the parent system. At the same time, the slow limit system becomes a more accurate description of the slow dynamics of the parent system. Note that the trajectory for the full system is not generally closed, even though it appears so in the left hand graph of Figure 1, while the slow limit system is topologically constrained to closed orbits.

The variational limit equations (9) differ from (12) by retaining different higher order terms. Thus, for the task of estimating the shadowing of trajectories of the slow limit systems, these differences are immaterial. Beyond the time of validity of the approximations, however, the nonvariational limit systems will generally blow up, while the variational limit system will continue its slow motion on the energy surface.

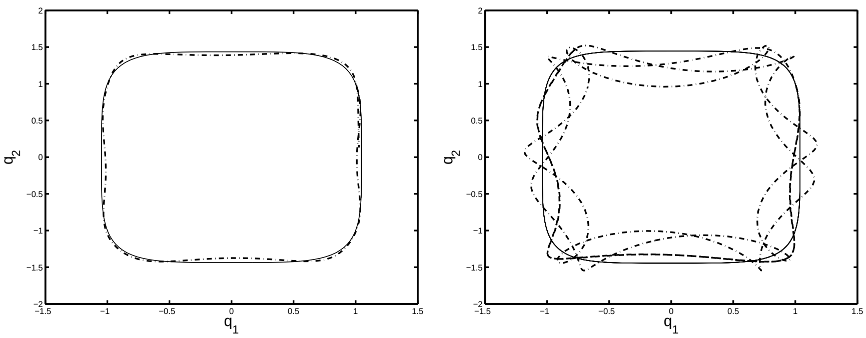

FIGURE 1. Full dynamics (1) (dotted line) and slow variational dynamics (9) (solid line) with quartic nonlinear potential $V(q) = \frac{3}{4}\,q_1^4 + \frac{1}{4}\,q_2^4$. Left: $\varepsilon = 0.1$ and right: $\varepsilon = 0.2$.

The present paper makes the point that under certain, more restrictive assumptions, the slow approximation is not only $O(\varepsilon^{n+1})$ on time scales of $O(1)$, but also $O(\varepsilon^n)$ on time scales of $O(\varepsilon^{-1})$. As the proofs for the long-time result require estimates on nonautonomous linear problems, they are considerably more involved. We therefore restrict ourselves to the case $n = 1$, stated in the following section. Numerical experiments, however, indicate that the result readily generalizes to higher order; see Figure 2.

3. Main Theorem

In the following, we assume that the potential V is even and convex. Physically, evenness of the potential means that the restoring force field is point symmetric about the equilibrium of the oscillator. Our main result then is the following.

Theorem 1. *Assume that V is even and convex. For $q(0) = q_0 \in \mathbb{R}^2$ fixed, let $q(t)$, implicitly depending on ε, be a sequence of solutions to either the variational slow equation* (9) *or to the nonvariational slow equation* (12)*, symbolically written as $\dot{q} = F(q)$. Further, let $q_\varepsilon(t)$ be a sequence of solutions to the full parent system* (1) *with balanced initial data $q_\varepsilon(0) = q_0$ and $\dot{q}_\varepsilon(0) = F(q_\varepsilon(0))$. Then there exist constants $c = c(q_0)$ and $\varepsilon_0 > 0$ such that*

$$\|q(t) - q_\varepsilon(t)\| \le c\,\varepsilon \tag{18}$$

for every $0 \le t \le \varepsilon^{-1}$ and $0 < \varepsilon \le \varepsilon_0$.

Remark 1. Within the setting of Theorem 1, there is no difference between the behavior of the variational and the nonvariational limit system. Beyond the $1/\varepsilon$ time scale, however, the nonvariational limit system generally blows up as the

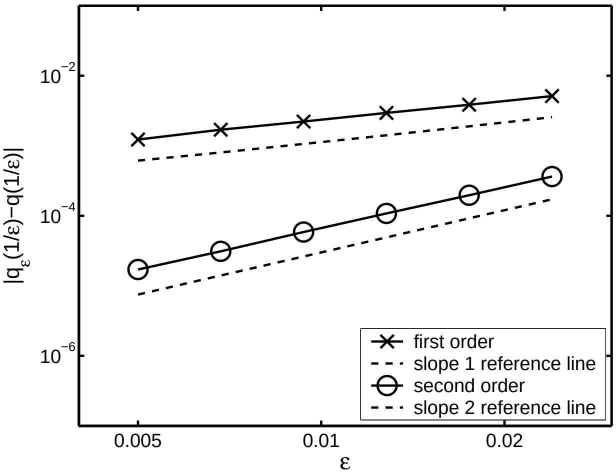

FIGURE 2. Error $\|q_\varepsilon - q\|$ at time $t = \varepsilon^{-1}$ vs. ε for the approximation of the slow manifold by first and second order limit systems with quartic potential $V(q) = \frac{3}{4}\,q_1^4 + \frac{1}{4}\,q_2^4$. The first order error scales very close to $O(\varepsilon)$ while the second order error scales very close to $O(\varepsilon^2)$.

dynamics is not constrained by a conserved energy. Theorem 1 thus provides a lower bound for the blow-up time of (12).

Remark 2. Numerical experiments show that the restriction to even potentials—which, as shown in Section 7, allows for a cancellation of error growth across consecutive half periods of the slow motion—is essential.

Remark 3. The result is optimal in the sense that the behavior is replicated in the linear case where the equations can be solved by explicit diagonalization [7, 14].

A crucial ingredient for the proof is the notion that the parent system (1) is Hamiltonian with conserved energy

$$E_\varepsilon(t) = \frac{\varepsilon}{2}\,\|\dot{q}_\varepsilon(t)\|^2 + V(q_\varepsilon(t)) = E_\varepsilon(0)\,. \tag{19}$$

Since the potential is convex, q_ε is bounded uniformly in t and ε. Moreover, the variational limit system is Hamiltonian with conserved energy

$$E(t) = V(q(t)) + \frac{\varepsilon}{2}\,\|\nabla V(q(t))\|^2 = E(0)\,. \tag{20}$$

We write our parent system symbolically as

$$\dot{q}_\varepsilon = F_\varepsilon(q_\varepsilon) \tag{21}$$

and, similarly, denote the vector fields corresponding to the non-variational limit system (12) and the variational limit system (9) by $F_{\rm nv}$ and $F_{\rm var}$, respectively. The evolution equation for the trajectory error $y = q_\varepsilon - q$ is split into three components,

$$\begin{aligned}\dot{y} &= F_\varepsilon(q_\varepsilon) - F_{\rm var}(q) \\ &= \big(F_\varepsilon(q_\varepsilon) - F_{\rm nv}(q_\varepsilon)\big) + \big(F_{\rm nv}(q_\varepsilon) - F_{\rm var}(q_\varepsilon)\big) + \big(F_{\rm var}(q_\varepsilon) - F_{\rm var}(q)\big)\,. \end{aligned} \tag{22}$$

The first two differences on the right are called the *consistency error* as is common in the analysis of numerical time integrators for differential equations (see, for example, [21]). The consistency errors are treated as driving terms to the linearization of the third term in (22). In other words, the evolution of the error is expressed as a *stability estimate* of the form

$$\dot{y} = G + \mathrm{D}F_{\mathrm{var}}(q)\, y\,. \tag{23}$$

Our situation is complicated by the fact that a straightforward application of the Gronwall inequality will miss by one order in ε. (This is in fact a restatement of the problem already encountered in the introduction.) To recover the missing order, we must carefully verify that the next order contribution of the consistency error actually gets averaged out over full periods of the slow motion. This averaging argument can be sketched as follows.

An oscillator $J\dot{q} = \nabla V(q)$ with even potential has trajectories which are anti-symmetric with respect to time shifts by half-periods. Thus, even functions h of the position q are $T/2$-periodic while odd functions $g(q)$ are $T/2$ anti-periodic. It turns out that the linearized growth of the difference in (18) can be cartooned by the scalar equation

$$\dot{y}(t) = g(q(t)) + h(q(t))\, y(t)\,. \tag{24}$$

Integrated over a full period T of $q(t)$, the contributions from h cancel identically. In the actual proof, this growth estimate is vector-valued and hence may have monodromy which needs to be carefully controlled.

The remainder of the paper provides the details of the proof. In Section 4, we re-derive the nonvariational limit system (12) by a formal argument which provides the raw structure for developing the rigorous proof later. The consistency error is analyzed in Section 5. The result, namely that the consistency error remains of the expected order on the long time scale, is stated as Theorem 6 at the end of Section 5.

The proof of Theorem 6 is performed by explicitly splitting off the fast dynamics which allows for averaging over one slow period by the application of the Gronwall lemma. As a by-product, we also get an estimate for the higher-order terms of the nonvariational system (12) which include fast components. In Section 6, we show that perturbations to the slow dynamics have at most secular growth and, moreover, do not resonate with the fast dynamics. This is expressed in Lemma 7 and in Corollary 10 respectively. In Section 7, these ingredients are combined to complete the proof of Theorem 1.

4. Derivation of the nonvariational limit system

In this section, we provide what effectively amounts to the variations-of-constants formulation of the derivation of the slow manifold $w = 0$ as sketched in the introduction. In this process, the nonvariational limit system arises from the boundary terms of successive integration by parts. Section 5 will take a different starting point which could be verified directly without going through this section. As such direct verification is non-constructive, however, it is worth taking the detour.

It is convenient to make the long time scale explicit. We thus introduce a slow time $\tau = \varepsilon\, t$, and write $u_\varepsilon(\tau) = q_\varepsilon(\tau/\varepsilon)$ and $u(\tau) = q(\tau/\varepsilon)$ to denote the solution as a function of slow time. This implies $\dot{q}(t) = \varepsilon\, \dot{u}(\tau)$. In these variables, $\dot{q}_\varepsilon(t) =$

$\varepsilon\,\dot u_\varepsilon(\tau)$ and $\ddot q_\varepsilon(t) = \varepsilon^2\,\dot u_\varepsilon(\tau)$, so that the parent dynamics (1) reads

$$\varepsilon^3\,\ddot u_\varepsilon(\tau) + \nabla V(u_\varepsilon(\tau)) - \varepsilon\,J\dot u_\varepsilon(\tau) = 0\,, \tag{25}$$

or

$$\ddot u_\varepsilon(\tau) - \frac{J}{\varepsilon^2}\,\dot u_\varepsilon(\tau) = -\frac{1}{\varepsilon^3}\,\nabla V(u_\varepsilon(\tau))\,. \tag{26}$$

Using an integrating factor and integrating once, we find

$$\dot u_\varepsilon(\tau) = \mathrm{e}^{\frac{J\tau}{\varepsilon^2}}\,\dot u_\varepsilon(0) - \frac{1}{\varepsilon^3}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,\nabla V(u_\varepsilon(\sigma))\,\mathrm{d}\sigma\,. \tag{27}$$

Writing

$$\varepsilon^2\,J\,\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}} = \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,, \tag{28}$$

we integrate by parts and obtain

$$\begin{aligned} w_0(\tau) &= \mathrm{e}^{\frac{J\tau}{\varepsilon^2}}\,w_0(0) + \frac{J}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,\mathrm{D}\nabla V(u_\varepsilon(\sigma))\,\dot u_\varepsilon(\sigma)\,\mathrm{d}\sigma \\ &= \mathrm{e}^{\frac{J\tau}{\varepsilon^2}}\,w_0(0) + \frac{J}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,\mathrm{D}\nabla V(u_\varepsilon(\sigma))\,w_0(\sigma)\,\mathrm{d}\sigma \\ &\quad - \frac{J}{\varepsilon^2}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,\mathrm{D}\nabla V(u_\varepsilon(\sigma))\,J\nabla V(u_\varepsilon(\sigma))\,\mathrm{d}\sigma\,, \end{aligned} \tag{29}$$

with

$$w_0(\tau) = \dot u_\varepsilon(\tau) + \frac{1}{\varepsilon}\,J\nabla V(u_\varepsilon(\tau))\,. \tag{30}$$

Leading order balance can now be written as $w_0(\tau) = 0$. To go to the next order, note that the last term in (29) does not contain any time derivatives on u_ε. Hence, we can continue integration by parts, thereby reducing the order of the prefactor to ε^{-1}. We regroup terms as before, obtaining

$$\begin{aligned} w_1(\tau) &= \mathrm{e}^{\frac{J\tau}{\varepsilon^2}}\,w_1(0) + \frac{J}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,\mathrm{D}\nabla V(u_\varepsilon(\sigma))\,w_1(\sigma)\,\mathrm{d}\sigma \\ &\quad - \int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,L_1(\sigma)[w_1(\sigma)]\,\mathrm{d}\sigma + \frac{1}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\,r_1(\sigma)\,\mathrm{d}\sigma\,, \end{aligned} \tag{31}$$

where

$$w_1(\tau) = \dot u_\varepsilon(\tau) + \frac{1}{\varepsilon}\,J\nabla V(u_\varepsilon(\tau)) - \mathrm{D}\nabla V(u_\varepsilon(\tau))J\nabla V(u_\varepsilon(\tau))\,. \tag{32}$$

Remainder terms which only depend on u_ε are lumped into the linear operator

$$L_1(\tau) = \mathrm{D}(\mathrm{D}\nabla V J\nabla V)(u_\varepsilon(\tau)) \tag{33}$$

and the vector

$$\begin{aligned} r_1(\tau) &= J\mathrm{D}\nabla V(u_\varepsilon(\tau))\,\mathrm{D}\nabla V(u_\varepsilon(\tau))\,J\nabla V(u_\varepsilon(\tau)) \\ &\quad + \mathrm{D}(\mathrm{D}\nabla V J\nabla V)(u_\varepsilon(\tau))\,[J\nabla V(u_\varepsilon(\tau)) - \varepsilon\,\mathrm{D}\nabla V(u_\varepsilon(\tau))\,J\nabla V(u_\varepsilon(\tau))]\,. \end{aligned} \tag{34}$$

The claim of Theorem 1 is that the dynamics on the slow manifold is given, to the order indicated, by the first-order nonvariational balance equation (12), which can now be written as $w_1(\tau) = 0$. Consistency of this approximation means that if w_1 is used as a diagnostic for the full dynamics and is small initially, it remains small for $\tau = O(1)$, or $t = O(1/\varepsilon)$ in physical time. It would thus be tempting to use the Gronwall inequality on (31) by estimating the norm of the first-order consistency error $\|w_1(\tau)\|$.

Indeed, the last two terms in (31) are benign: The term involving w_1 is already in good shape for a Gronwall inequality argument, while the final term again does not involve time derivatives, so that integration by parts can be used once more to improve the order of the prefactor from $1/\varepsilon$ to 1. The first integral on the right of (31), however, cannot be easily dealt with unless we restrict the argument to the short time scale ε. Our consistency argument, which is presented in Section 5, is based on the observation that this term can be understood as a perturbation of the unitary matrix group $\exp(J\tau/\varepsilon^2)$.

The stability argument in Section 6 is even more subtle—it depends on the odd symmetry of the next order correction to the nonvariational limit system. We must therefore expose these next order terms explicitly, which amounts to integration by parts, once again, on the last term of (31). We obtain

$$\begin{aligned} w_2(\tau) &= \mathrm{e}^{\frac{J\tau}{\varepsilon^2}}\, w_2(0) + \frac{J}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\, \mathrm{D}\nabla V(u_\varepsilon(\sigma))\, w_2(\sigma)\,\mathrm{d}\sigma \\ &\quad - \int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\, L_2(\sigma)[w_2(\sigma)]\,\mathrm{d}\sigma + \int_0^\tau \mathrm{e}^{\frac{J(\tau-\sigma)}{\varepsilon^2}}\, r_2(\sigma)\,\mathrm{d}\sigma\,, \end{aligned} \tag{35}$$

where

$$w_2(\tau) = w_1(\tau) + \varepsilon\, Z_1(u_\varepsilon(\tau)) = \dot u_\varepsilon(\tau) - F_{\mathrm{nv}}(u_\varepsilon(\tau)) + \varepsilon\, Z_1(u_\varepsilon(\tau))\,, \tag{36a}$$

$$F_{\mathrm{nv}}(u) = -\frac{1}{\varepsilon}\, J\nabla V(u) + \mathrm{D}\nabla V(u) J\nabla V(u)\,, \tag{36b}$$

$$Z_1(u) = -J\mathrm{D}(\mathrm{D}\nabla V J\nabla V)(u)[J\nabla V(u)] + \mathrm{D}\nabla V(u)\,\mathrm{D}\nabla V(u)\, J\nabla V(u)\,, \tag{36c}$$

and

$$L_2(\tau) = L_1(\tau) - \varepsilon\,\mathrm{D}Z_1(u_\varepsilon(\tau))\,. \tag{36d}$$

Remainder terms which only depend on u_ε, hence are uniformly bounded as $\varepsilon \to 0$, and those which are of $O(\varepsilon^2)$ are grouped into $r_2(\tau)$.

We proceed by left-multiplying (35) with $\exp(-J\tau/\varepsilon^2)$, and differentiate the resulting expression, so that

$$\begin{aligned} &-\frac{J}{\varepsilon^2}\,\mathrm{e}^{-\frac{J\tau}{\varepsilon^2}}\, w_2(\tau) + \mathrm{e}^{-\frac{J\tau}{\varepsilon^2}}\, \dot w_2(\tau) \\ &\quad = \frac{J}{\varepsilon}\,\mathrm{e}^{-\frac{J\tau}{\varepsilon^2}}\, \mathrm{D}\nabla V(u_\varepsilon(\tau))\, w_2(\tau) - \mathrm{e}^{-\frac{J\tau}{\varepsilon^2}}\, L_2(\tau)[w_2(\tau)] + \mathrm{e}^{-\frac{J\tau}{\varepsilon^2}}\, r_2(\tau)\,, \end{aligned} \tag{37}$$

which we write in its final form as

$$\dot w_2 = \frac{J}{\varepsilon^2}\big(I + \varepsilon\,\mathrm{D}\nabla V(u_\varepsilon)\big)\, w_2 - L_2[w_2] + r_2\,. \tag{38}$$

The analysis in Section 5 is based on formulation (38). Of course, we could have started with (38) right away and checked equivalence to (1) by direct calculation. This, however, would hide the structure of the argument, namely that an expansion for the slow equation emerges as an aggregation of boundary terms from successive integration by parts. The equation $w_2(\tau) = 0$ is the second-order nonvariational system for balanced motion. Although our main Theorem 1 is only concerned with the first-order balance model we need explicit expressions for the next order correction to prove that they do not impact the slow dynamics on the long time scale.

5. Consistency

In this section we prove an $O(\varepsilon)$ bound on the consistency error $w_1(\tau)$. Along the way, we obtain a representation of the explicit next order error $w_2(\tau)$ in terms of an oscillatory integral.

The proof is based on the fact that the fundamental matrix corresponding to the entire first term on the right of (38) is a small perturbation of a one-parameter unitary group. To make this statement rigorous, we must study the linear homogeneous problem

$$\dot{x}(\tau) = A(\tau)\, x(\tau) \tag{39}$$

where

$$A(\tau) = \frac{J}{\varepsilon^2}\big(I + \varepsilon\, \mathrm{D}\nabla V(u_\varepsilon(\tau))\big)\,. \tag{40}$$

The fundamental matrix $R = R(\tau;\sigma)$ is defined through $x(\tau) = R(\tau;\sigma)\, x(\sigma)$. We prove three auxiliary results.

Lemma 2. *$A^{-1}(\tau) = O(\varepsilon^2)$ as $\varepsilon \to 0$ uniformly in τ.*

Proof. A formal power series inversion yields

$$A^{-1} = -\varepsilon^2 \sum_{i=0}^{\infty} \big(-\varepsilon\, \mathrm{D}\nabla V(u_\varepsilon)\big)^i\, J\,. \tag{41}$$

Since u_ε remains bounded as $\varepsilon \to 0$, the series is absolutely convergent for ε small enough, hence uniformly bounded as $\varepsilon \to 0$. □

Lemma 3. *For any τ fixed, R solves the* adjoint equation

$$\partial_2 R(\tau;\sigma) = -R(\tau;\sigma)\, A(\sigma)\,. \tag{42}$$

Here, and in the following, ∂_j denotes the partial derivative with respect to the jth argument slot.

Proof. The fundamental matrix $R(\tau;\sigma)$ solves the nonautonomous linear matrix equation

$$\partial_1 R(\tau;\sigma) = A(\tau)\, R(\tau;\sigma)\,, \tag{43a}$$

$$R(\sigma;\sigma) = I\,. \tag{43b}$$

Moreover, the fundamental matrix satisfies

$$R(\tau;\sigma) = R(\tau;\rho)\, R(\rho;\sigma) \tag{44}$$

for arbitrary real numbers ρ, σ, and τ. In particular,

$$R(\tau;\sigma)\, R(\sigma;\tau) = I\,, \tag{45}$$

so that differentiation with respect to τ yields

$$\partial_1 R(\tau;\sigma)\, R(\sigma;\tau) + R(\tau;\sigma)\, \partial_2 R(\sigma;\tau) = 0\,. \tag{46}$$

Plugging in (43a), multiplying with $R(\sigma;\tau)$ from the left, and exchanging variable names, we obtain the statement of the lemma. □

Lemma 4. *There exists a constant c depending only on $u_\varepsilon(0)$ and an $\varepsilon_0 > 0$ such that*

$$\|R(\tau;\sigma)\| \leq 2\,\exp\biggl(c \int_\sigma^\tau \bigl(1 + \varepsilon\,\|w_1(\rho)\|\bigr)\,\mathrm{d}\rho\biggr) \tag{47}$$

for all $0 < \varepsilon \leq \varepsilon_0$. In particular, $\|R\| = O(\varepsilon^0)$ as long as w_1 remains a priori *bounded.*

Proof. We take a WKB-like approach and factor out the oscillations with frequency $O(\varepsilon^{-2})$ as follows. Without loss of generality, we take the initial time $\sigma = 0$ and drop the second argument of R henceforth. We separate trace and trace-free part of $\mathrm{D}\nabla V$ by setting

$$Q(\tau) = \mathrm{D}\nabla V(u_\varepsilon(\tau)) - \tfrac{1}{2}\,I\ \operatorname{Tr}\mathrm{D}\nabla V(u_\varepsilon(\tau))\,, \tag{48}$$

$$\phi(\tau) = 1 + \tfrac{1}{2}\,\varepsilon\ \operatorname{Tr}\mathrm{D}\nabla V(u_\varepsilon(\tau))\,, \tag{49}$$

and define

$$\Phi(\tau) = \int_0^\tau \phi(\sigma)\,\mathrm{d}\sigma\,. \tag{50}$$

We remark that as ϕ is a small perturbation of unity, Φ can be regarded as a new time-like variable. Then $R(\tau;\sigma)$ satisfies the equation

$$\dot{R}(\tau) = \frac{J}{\varepsilon^2}\,\bigl(\phi(\tau)\,I + \varepsilon\,Q(\tau)\bigr)\,R(\tau) \tag{51}$$

with $R(0) = I$. Further, define

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix} \qquad \text{and} \qquad \Omega = \frac{1}{\varepsilon^2}\begin{pmatrix} -\mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix}\,. \tag{52}$$

We introduce a "slow" fundamental matrix Y by explicitly factoring out fast oscillations,

$$R(\tau) = U\,\mathrm{e}^{\Omega\Phi(\tau)}\,Y(\tau)\,. \tag{53}$$

Since this transformation is unitary, bounding R is equivalent to bounding Y. Direct computation shows that Y must satisfy

$$\dot{Y} = \frac{1}{\varepsilon}\,MY\,, \tag{54}$$

with $Y(0) = U^{-1}$ and

$$M = \mathrm{e}^{-\Omega\Phi}\,U^{-1}\,JQ\,U\,\mathrm{e}^{\Omega\Phi}\,. \tag{55}$$

Noting that Q is a traceless symmetric matrix, we find that

$$M \equiv \mathrm{e}^{-\Omega\Phi}\,U^{-1}\,JQ\,U\,\mathrm{e}^{\Omega\Phi} = \begin{pmatrix} 0 & \mu\,\mathrm{e}^{\frac{2\mathrm{i}}{\varepsilon^2}\Phi} \\ \overline{\mu}\,\mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi} & 0 \end{pmatrix}\,, \tag{56}$$

where $\mu = -(Q_{12} + \mathrm{i}Q_{11})$ ultimately is a function of u_ε.

We complete the proof by a direct estimate on the growth of an arbitrary solution to the "slow" equation (54). Writing this equation out in components and integrating in τ, we have

$$y_1(\tau) = y_1(0) + \frac{1}{\varepsilon}\int_0^\tau \mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\sigma)}\,\mu(\sigma)\,y_2(\sigma)\,\mathrm{d}\sigma\,, \tag{57a}$$

$$y_2(\tau) = y_2(0) + \frac{1}{\varepsilon}\int_0^\tau \mathrm{e}^{\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\sigma)}\,\overline{\mu}(\sigma)\,y_1(\sigma)\,\mathrm{d}\sigma\,. \tag{57b}$$

Since

$$\mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\sigma)} = -\frac{\varepsilon^2}{2\mathrm{i}}\,\frac{1}{\phi(\sigma)}\,\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\sigma)}\,, \tag{58}$$

integration by parts in (57a) yields

$$\begin{aligned} y_1(\tau) &= y_1(0) - \frac{\varepsilon}{2\mathrm{i}}\,\mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\tau)}\,\frac{\mu(\tau)}{\phi(\tau)}\,y_2(\tau) + \frac{\varepsilon}{2\mathrm{i}}\,\frac{\mu(0)}{\phi(0)}\,y_2(0) \\ &\quad + \frac{1}{2\mathrm{i}}\int_0^\tau \biggl(\frac{|\mu(\sigma)|^2}{\phi(\sigma)}\,y_1(\sigma) + \varepsilon\,\mathrm{e}^{-\frac{2\mathrm{i}}{\varepsilon^2}\Phi(\sigma)}\,\frac{\mathrm{d}}{\mathrm{d}\sigma}\biggl(\frac{\mu(\sigma)}{\phi(\sigma)}\biggr)\,y_2(\sigma)\biggr)\,\mathrm{d}\sigma \end{aligned} \tag{59}$$

We now take the absolute value on both sides. Since u_ε is bounded uniformly in ε for all time, so are μ and the trace of $\mathrm{D}\nabla V(u_\varepsilon)$. As a consequence, $\phi(\tau)$ is bounded away from zero, e.g. $\phi(\tau) \geq \frac{1}{2}$, for ε small enough. Finally,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\sigma}\,\frac{\mu}{\phi} &= \mathrm{D}\biggl(\frac{\mu}{\phi}\biggr)(u_\varepsilon)\cdot\dot u_\varepsilon \\ &= \mathrm{D}\biggl(\frac{\mu}{\phi}\biggr)(u_\varepsilon)\cdot\biggl[w_1 - \frac{1}{\varepsilon}\,J\nabla V(u_\varepsilon) + \mathrm{D}\nabla V(u_\varepsilon)J\nabla V(u_\varepsilon)\biggr]\,. \end{aligned} \tag{60}$$

We thus find that

$$|y_1(\tau)| \leq |y_1(0)| + \varepsilon\,c\,|y_2(0)| + \varepsilon\,c\,|y_2(\tau)| + c\int_0^\tau \bigl(|y_1(\sigma)| + (1+\varepsilon\,\|w_1(\sigma)\|)\,|y_2(\sigma)|\bigr)\,\mathrm{d}\sigma\,, \tag{61}$$

where c denotes a constant that may only depend on the initial data $u_\varepsilon(0)$. Similarly,

$$|y_2(\tau)| \leq |y_2(0)| + \varepsilon\,c\,|y_1(0)| + \varepsilon\,c\,|y_1(\tau)| + c\int_0^\tau \bigl(|y_2(\sigma)| + (1+\varepsilon\,\|w_1(\sigma)\|)\,|y_1(\sigma)|\bigr)\,\mathrm{d}\sigma\,. \tag{62}$$

We set $\eta = |y_1| + |y_2|$ and add the previous two inequalities, so that

$$\eta(\tau) \leq (1+\varepsilon\,c)\,\eta(0) + \varepsilon\,c\,\eta(\tau) + c\int_0^\tau \bigl(2+\varepsilon\,\|w_1(\sigma)\|\bigr)\,\eta(\sigma)\,\mathrm{d}\sigma\,. \tag{63}$$

Choosing ε small enough, we finally obtain

$$\eta(\tau) \leq 2\,\eta(0) + 4\,c\int_0^\tau \bigl(1+\varepsilon\,\|w_1(\sigma)\|\bigr)\,\eta(\sigma)\,\mathrm{d}\sigma\,. \tag{64}$$

Note that the prefactor 2 in front of $\eta(0)$ is purely technical; bounds that are initially sharp are possible, but are not required for our purpose. Noting that Y and R are related via a unitary transformation we can prove the Lemma upon using the Gronwall inequality. □

The following lemma is not necessary for proving a bound of the consistency error. However, it will prove crucial for the stability error. This Lemma will allow to control the driving of the slow dynamics by the higher order fast terms.

Lemma 5. *Let $w_1(0) = 0$ and let R denote the fast fundamental matrix corresponding to* (39)*. Then there exists a bounded time dependent linear map L_3 such that*

$$w_2(\tau) = \varepsilon\,R(\tau;0)\,Z_1(u(0)) + \int_0^\tau R(\tau;\sigma)\,L_3(\sigma)[w_2(\sigma)]\,\mathrm{d}\sigma + O(\varepsilon^2)\,. \tag{65}$$

Proof. We rewrite (38) in its mild formulation

$$w_2(\tau) = R(\tau;0)\, w_2(0) + \int_0^\tau R(\tau;\sigma) \Big(-L_2(\sigma)[w_2(\sigma)] + r_2(\sigma) \Big)\, \mathrm{d}\sigma \,. \tag{66}$$

Taking norms of this expression would leave the last term under the integral at $O(1)$. We can improve its order by integration by parts, in effect retracing the steps of Section 4 with the fundamental matrix of the nonautonomously perturbed linear system in place of explicit matrix exponentials. Due to Lemma 2,

$$\varepsilon^2\, r_3(\tau) \equiv A^{-1}(\tau)\, r_2(\tau) = O(\varepsilon^2)\,. \tag{67}$$

Using the adjoint equation from Lemma 3, we compute

$$\begin{aligned} \int_0^\tau R(\tau;\sigma)\, r_2(\sigma)\, \mathrm{d}\sigma &= - \int_0^\tau \partial_\sigma R(\tau;\sigma)\, A^{-1}(\sigma)\, r_2(\sigma)\, \mathrm{d}\sigma \\ &= -\varepsilon^2\, R(\tau;\sigma)\, r_3(\sigma) \Big|_{\sigma=0}^{\sigma=\tau} + \varepsilon^2 \int_0^\tau R(\tau;\sigma)\, \mathrm{D}r_3(u_\varepsilon(\sigma))\, \dot u_\varepsilon(\sigma)\, \mathrm{d}\sigma \\ &= -\varepsilon^2\, R(\tau;\sigma)\, r_3(\sigma) \Big|_{\sigma=0}^{\sigma=\tau} + \varepsilon^2 \int_0^\tau R(\tau;\sigma)\, \mathrm{D}r_3(u_\varepsilon(\sigma))\, w_2(\sigma)\, \mathrm{d}\sigma \\ &\quad + \varepsilon \int_0^\tau R(\tau;\sigma)\, \mathrm{D}r_3(u_\varepsilon(\sigma))\, r_4(\sigma)\, \mathrm{d}\sigma \end{aligned} \tag{68}$$

with

$$r_4(\tau) \equiv \varepsilon\, (\dot u_\varepsilon(\tau) - w_2(\tau)) = \varepsilon\, F_{\mathrm{nv}}(u_\varepsilon(\tau)) - \varepsilon^2\, Z_1(u_\varepsilon(\tau)) = O(1)\,. \tag{69}$$

The order of the last term in (68) must still be improved. As this term has the same structure as the left hand side of (68), we iterate the argument and apply partial integration once more, concluding that this term contributes only at $O(\varepsilon^2)$. The second term on the right of (68) can be combined with the first part of the integrand in (66), resulting in an overall expression of the form (65) with $L_3(\tau) = -L_2(\tau) + O(\varepsilon^2)$. □

Theorem 6 (Consistency)**.** *Let $w_1(0) = 0$. Then there exist constants c and ε_0 such that*

$$\|w_1(\tau)\| \le c\,\varepsilon \qquad \textit{and} \qquad \|w_2(\tau)\| \le c\,\varepsilon \tag{70}$$

for $\varepsilon t = \tau \in [0,1]$ and for every $0 < \varepsilon \le \varepsilon_0$.

Proof. We make the dependence of w_2 on ε explicit by writing $w_2 = w_2(\tau;\varepsilon)$. Let τ_ε denote the maximum time such that $\|w_2(\tau;\varepsilon)\| < 1$ for every $\tau < \tau_\varepsilon$. If $\liminf \tau_\varepsilon > 0$, the claim is immediate. Otherwise, there exists a sequence $\varepsilon_k \to 0$ with $\tau_{\varepsilon_k} \to 0$ as $k \to \infty$; without loss of generality we may assume that $\tau_{\varepsilon_k} \le 1$. Then Lemma 4 implies that there exists a constant c which bounds $\|R(\tau,\sigma)\|$ for all $\sigma \in [0,\tau]$ with $\tau < \tau_{\varepsilon_k} < 1$ and all k. (Note that, since w_2 is bounded by 1, the difference between w_2 and w_1 is at most $O(\varepsilon_k)$. Hence, w_1 is also bounded *a priori*.) Taking norms on (65), we have

$$\|w_2(\tau;\varepsilon_k)\| \le c_1\, \varepsilon_k + c_2 \int_0^\tau \|w_2(\sigma;\varepsilon_k)\|\, \mathrm{d}\sigma \,. \tag{71}$$

The Gronwall inequality directly implies that there exists a constant c_3 such that $\|w_2(\tau;\varepsilon_k)\| \le c_3\, \varepsilon_k$ for all $\tau \in [0, \tau_{\varepsilon_k})$. For k large enough, $c_3\, \varepsilon_k \le \frac{1}{2}$. Continuity of w_2 then implies that there exists $\tau^*_{\varepsilon_k} > \tau_{\varepsilon_k}$ such that $\|w_2(\tau,\varepsilon_k)\| < 1$ for $\tau < \tau^*_{\varepsilon_k}$. This contradicts the maximality of τ_{ε_k}. □

Remark 4. To prove this theorem it would not have been necessary to carry along the next order contributions in w_2 explicitly—we could have derived an inequality of the form (71) for w_1 first, and deduced the w_2 bound from there. However, we chose to go this route as the intermediate equation (65) will be important later on.

A second contribution to the consistency error comes from the difference between the variational and nonvariational limit systems, $F_{\mathrm{nv}}(u_\varepsilon) - F_{\mathrm{var}}(u_\varepsilon)$. The computation is direct, with details as follows.

Recall that the variational slow dynamics in rescaled variables is given by

$$\varepsilon \left(1 + \varepsilon \, \Delta V(u)\right) J \dot{u} = \nabla V(u) + \varepsilon \, \mathrm{D}\nabla V(u) \, \nabla V(u) \, , \tag{72}$$

which we shall abbreviate

$$\dot{u} = F_{\mathrm{var}}(u) \equiv -\frac{1}{\varepsilon + \varepsilon^2 \, \Delta V(u)} \, J \left(\nabla V(u) + \varepsilon \, \mathrm{D}\nabla V(u) \nabla V(u)\right) . \tag{73}$$

For comparison and for future reference, we write out the nonvariational limit system,

$$\dot{u} = F_{\mathrm{nv}}(u) \equiv -\frac{1}{\varepsilon} \, J \nabla V(u) + \mathrm{D}\nabla V(u) J \nabla V(u) \, . \tag{74}$$

Expanding F_{var} in powers of ε and noting that $\Delta V J = J \, \mathrm{D}\nabla V + \mathrm{D}\nabla V \, J$, we find that

$$\begin{aligned} F_{\mathrm{var}} &= -\frac{1}{\varepsilon} \, J \nabla V + \mathrm{D}\nabla V J \nabla V - \varepsilon \, \Delta V \, \mathrm{D}\nabla V J \nabla V + O(\varepsilon^2) \\ &= F_{\mathrm{nv}} - \varepsilon \, \Delta V \, \mathrm{D}\nabla V J \nabla V + O(\varepsilon^2) \\ &\equiv F_{\mathrm{nv}} - \varepsilon \, Z_2(u) + O(\varepsilon^2) \, . \end{aligned} \tag{75}$$

6. Stability

In the following, we prove that perturbations to the slow variational limit system have secular rather than exponential growth—the limit system is a planar nonlinear oscillator. Moreover, when the potential is even, inhomogeneities which are odd functions of the slow dynamics will cancel on the time scale of the slow period and first appear on the long time scale. Odd potentials, on the other hand, would give rise to even forces and fast modes could nonlinearly interact to produce a slow mode of zero frequency.

The following lemma is our basic stability result.

Lemma 7. *Let $u(\tau)$ denote a solution to* (72)*, and let $S(\tau;\sigma)$ denote the fundamental matrix which governs the evolution of small perturbations about $u(\tau)$, i.e.*

$$\dot{S}(\tau;\sigma) = \mathrm{D}F_{\mathrm{var}}(u(\tau)) \, S(\tau;\sigma) \, , \tag{76a}$$

$$S(\sigma;\sigma) = I \, . \tag{76b}$$

Then

$$S(\tau;\sigma) = P(\tau;\sigma) \, \mathrm{e}^{\Lambda(\tau-\sigma)/T} \tag{77}$$

where $P(\,\cdot\,;\sigma)$ is periodic with period $T = O(\varepsilon)$, and $\Lambda = O(1)$ is nilpotent so that the growth of S can only be secular.

Proof. As V is convex, the planar phase space of the variational limit system (72) foliates into periodic orbits that are parameterized by their conserved energy. Thus, the existence of a factorization of the form (77) follows readily by the Floquet theorem. The scaling of the period, namely $T = O(\varepsilon)$, is evident from the leading order terms in (72), as the higher order terms are regular perturbations. It remains to be shown that the matrix of characteristic exponents Λ is nilpotent, i.e. that the growth of S is secular at most.

Guided by the Hamiltonian structure of the leading order slow dynamics, we recognize the energy as an action variable I and time as an angle variable σ. We then denote by $u(\tau; I, \sigma)$ the solution to (72) satisfying

$$u(\sigma; I, \sigma) = u_0(I)\,, \tag{78}$$

where $u_0(I)$ is a fixed curve of initial values normal to the periodic orbits. Differentiating (78) with respect to σ and I, respectively, we see that the phase error u_σ and amplitude error u_I evolve according to a differential equation of the form (76a), namely

$$\dot{u}_\sigma = \mathrm{D}F_{\mathrm{var}}(u(t))\, u_\sigma \qquad \text{and} \qquad \dot{u}_I = \mathrm{D}F_{\mathrm{var}}(u(t))\, u_I\,. \tag{79}$$

This is the same equation that a perturbation about the solution $u(\tau)$ of (73) would satisfy. In particular, u_σ and u_I are initially linearly independent, thus form a complete fundamental set of solutions for the evolution of small perturbations about $u(\tau)$.

The period of the oscillations for solutions of (72) generally depends on the orbit, so that $u(\tau; I, \sigma) = u(\tau + T(I); I, \sigma)$. Differentiation with respect to initial phase σ gives

$$u_\sigma(\tau; I, \sigma) = u_\sigma(\tau + T(I); I, \sigma)\,, \tag{80}$$

and, similarly,

$$u_I(\tau; I, \sigma) = u_I(\tau + T(I); I, \sigma) + \dot{u}(\tau + T(I); I, \sigma)\, T'(I)\,. \tag{81}$$

In particular, differentiating (78) with respect to σ and setting $\sigma = \tau = 0$, we find that

$$\begin{pmatrix} u_\sigma(T) \\ u_I(T) \end{pmatrix} = \begin{pmatrix} 1 & T'(I) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_\sigma(0) \\ u_I(0) \end{pmatrix}\,. \tag{82}$$

As the matrix on the right is clearly the exponential of a nilpotent matrix, the lemma is proved. □

Remark 5. Equation (82) states that there is no generation of amplitude errors u_I over one slow period. The exclusive generation of phase errors u_σ along the periodic orbit is due to the underlying Hamiltonian structure. Exponentially accurate normal form theory for perturbed Hamiltonian systems can yield even stronger bounds [13, 16]. This has been lately applied to the toy model (1) in [2]. However, our estimates are aimed at proving not just closeness of the slow manifold to the exact manifold of the parent system, but rather to prove convergence of actual trajectories on the manifolds and thereby we need to estimate the phase errors as well.

An immediate consequence of Lemma 7 is the following estimate.

Corollary 8. *Let S be defined as in Lemma 7. There exists a constant $c = c(u_0)$ such that*

$$\|S(\tau, \sigma)\| \leq \frac{c}{\varepsilon}\, |\tau - \sigma|\,. \tag{83}$$

Remark 6. In our setting, perturbations will enter as inhomogeneities of magnitude $O(\varepsilon)$. Thus, an estimate of the form (83) is not strong enough to conclude that errors remain small on the long time scale $\tau = O(1)$. For even potentials, however, there is a cancellation of errors between consecutive half-periods, and the order of the estimate improves by one. This will be detailed in Section 7.

We finally need a nonresonance result on the interaction between fast and slow modes.

Lemma 9. *Let S denote the slow fundamental matrix of Lemma 7 and R the fast fundamental matrix of Section 5. Then*

$$\int_\tau^\omega S(\omega;\rho)\, R(\rho;\sigma)\,\mathrm{d}\rho = O(\varepsilon^2) \tag{84}$$

for any $\tau, \sigma, \omega = O(1)$ with $|\omega - \tau| = O(\varepsilon)$.

Proof. The proof of the lemma is similar to that of Lemma 5. Recall that R satisfies the differential equation $\dot R(\rho;\sigma) = A(\rho)\, R(\rho,\sigma)$, where A is given by (40), so that

$$\begin{aligned}\int_\tau^\omega S(\omega;\rho)\, R(\rho;\sigma)\,\mathrm{d}\rho &= \int_\tau^\omega S(\omega;\rho)\, A^{-1}(\rho)\, \partial_\rho R(\rho;\sigma)\,\mathrm{d}\rho \\ &= S(\omega;\rho)\, A^{-1}(\rho)\, R(\rho;\sigma)\Big|_{\rho=\tau}^{\rho=\omega} - \int_\tau^\omega \partial_\rho S(\omega;\rho)\, A^{-1}(\rho)\, R(\rho;\sigma)\,\mathrm{d}\rho \\ &\quad - \int_\tau^\omega S(\omega;\rho)\, \partial_\rho A^{-1}(\rho)\, R(\rho;\sigma)\,\mathrm{d}\rho\,. \end{aligned} \tag{85}$$

Set $\delta = |\omega - \tau|$. Recall from Corollary 8 that $S(\omega;\rho) = O(\delta/\varepsilon)$, from Lemma 2 that $A^{-1} = O(\varepsilon^2)$, and from Lemma 4—with an $O(\varepsilon)$ bound on w_1 guaranteed by Theorem 6—that $R(\rho,\sigma) = O(1)$. Thus, the boundary term in (85) is $O(\delta\varepsilon)$ overall.

To estimate the second term on the right of (85), we formulate—as in Lemma 3—the adjoint equation corresponding to (76a),

$$\partial_\sigma S(\tau;\sigma) = -S(\tau;\sigma)\, \mathrm{D}F_{\mathrm{var}}(u(\sigma))\,. \tag{86}$$

We also note that $\mathrm{D}F_{\mathrm{var}} = O(1/\varepsilon)$. Then the application of the same set of estimates as used above for the boundary term yields an overall $O(\delta^2)$ bound.

To estimate the last term on the right of (85), we compute

$$\begin{aligned}\partial_\tau A^{-1} &= -A^{-1}\dot A A^{-1} \\ &= -A^{-1}\Big(\frac{J}{\varepsilon}\,\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{D}\nabla V(u_\varepsilon(\tau))\Big)A^{-1} \\ &= -A^{-1}\Big(\frac{J}{\varepsilon}\,\mathrm{D}\mathrm{D}\nabla V(u_\varepsilon(\tau))\big(w_1(\tau) + F_{\mathrm{nv}}(u_\varepsilon(\tau))\big)\Big)A^{-1} \\ &= O(\varepsilon^2)\,, \end{aligned} \tag{87}$$

since $w_1 = O(\varepsilon)$ and $F_{\mathrm{nv}}(u_\varepsilon) = O(1/\varepsilon)$. Thus, the last term on the right of (85) is $O(\varepsilon\delta^2)$ overall, and the proof of the lemma is complete. □

In the following Corollary we estimate the effect of the driving of the slow dynamics by higher order terms over one slow period T.

Corollary 10. *Let S denote the slow fundamental matrix of Lemma 7, $T = O(\varepsilon)$ the slow period, set $\tau_n = nT$, and let w_2 be defined through* (36a)*. Then*

$$\int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\sigma)\, w_2(\sigma)\,\mathrm{d}\sigma = O(\varepsilon^3) \tag{88}$$

on a time scale $\tau_n = O(1)$.

Proof. A crude estimate using $S(\omega;\rho) = O(1)$, $w_2 = O(\varepsilon)$, and $\tau_{n+1} - \tau_n = T = O(\varepsilon)$ yields an $O(\varepsilon^2)$-bound. This bound can be improved by one order as follows. According to Lemma 5,

$$\begin{aligned}
&\int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\sigma)\, w_2(\sigma)\,\mathrm{d}\sigma \\
&\quad = \varepsilon \int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\sigma)\, R(\sigma;0)\, Z_1(u(0))\,\mathrm{d}\sigma \\
&\qquad + \int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\sigma) \int_0^{\sigma} R(\sigma;\rho)\, L_3(\rho)[w_2(\rho)]\,\mathrm{d}\rho\,\mathrm{d}\sigma + O(\varepsilon^3) \\
&\quad = \varepsilon \int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\sigma)\, R(\sigma;0)\,\mathrm{d}\sigma\, Z_1(u(0)) \\
&\qquad + \biggl(\int_0^{\tau_n} \int_{\tau_n}^{\tau_{n+1}} + \int_{\tau_n}^{\tau_{n+1}} \int_{\rho}^{\tau_{n+1}} \biggr) S(\tau_{n+1};\sigma)\, R(\sigma;\rho)\,\mathrm{d}\sigma\, L_3(\rho)[w_2(\rho)]\,\mathrm{d}\rho \\
&\qquad + O(\varepsilon^3)\,.
\end{aligned} \tag{89}$$

Since $w_2 = O(\varepsilon)$ by Theorem 6, the result follows immediately from Lemma 9. □

7. Proof of Theorem 1

We can now proceed to prove our main result. We symbolically write the full system (1) as

$$\dot{u}_\varepsilon = F_\varepsilon(u_\varepsilon)\,, \tag{90}$$

with F_ε given by the integro-differential equation (27). Let $u(\tau)$ denote the solution to the variational slow equation

$$\dot{u} = F_{\mathrm{var}}(u)\,. \tag{91}$$

The evolution of the approximation error $y(\tau) = u_\varepsilon(\tau) - u(\tau)$ is split into three components,

$$\begin{aligned}
\dot{y} &= F_\varepsilon(u_\varepsilon) - F_{\mathrm{var}}(u) \\
&= \bigl(F_\varepsilon(u_\varepsilon) - F_{\mathrm{nv}}(u_\varepsilon)\bigr) + \bigl(F_{\mathrm{nv}}(u_\varepsilon) - F_{\mathrm{var}}(u_\varepsilon)\bigr) + \bigl(F_{\mathrm{var}}(u_\varepsilon) - F_{\mathrm{var}}(u)\bigr)\,.
\end{aligned} \tag{92}$$

We consider each of these components in turn. By definition, see equation (36), we write

$$F_\varepsilon(u_\varepsilon) - F_{\mathrm{nv}}(u_\varepsilon) = w_1 = w_2 - \varepsilon\, Z_1(u_\varepsilon) + O(\varepsilon^2)\,. \tag{93}$$

Next, quoting (75), we have

$$F_{\mathrm{nv}}(u_\varepsilon) - F_{\mathrm{var}}(u_\varepsilon) = \varepsilon\, Z_2(u_\varepsilon) + O(\varepsilon^2) \tag{94}$$

Last, we expand

$$F_{\mathrm{var}}(u_\varepsilon) - F_{\mathrm{var}}(u) = \mathrm{D}F_{\mathrm{var}}(u) y + \tfrac{1}{2}\, y^T\ \mathrm{Hess}\, F_{\mathrm{var}}(u)\, y + O\biggl(\frac{\|y\|^3}{\varepsilon}\biggr)\,. \tag{95}$$

(Recall that, in rescaled time, F_{var} and F_{nv} are $O(\varepsilon^{-1})$ vector fields.) We altogether obtain

$$\dot{y} = \mathrm{D}F_{\text{var}}(u)y + G\,, \tag{96}$$

where, by the use of the mean value theorem,

$$G = w_2 - \varepsilon(Z_1(u_\varepsilon) - Z_2(u_\varepsilon)) + \tfrac{1}{2}\, y^T \;\mathrm{Hess}\, F_{\text{var}}(u)\, y + O\biggl(\varepsilon^2, \frac{\|y\|^3}{\varepsilon}\biggr) \tag{97}$$

$$= w_2 + \varepsilon\, Z(u) + \varepsilon\, \mathrm{D}Z(u+\nu y)\, y + \tfrac{1}{2}\, y^T \;\mathrm{Hess}\, F_{\text{var}}(u)\, y + O\biggl(\varepsilon^2, \frac{\|y\|^3}{\varepsilon}\biggr) \tag{98}$$

for some $\nu = \nu(\tau) \in [0,1]$. Note that, at this stage, we only have $y = O(1)$. This is obvious since $y = u_\varepsilon - u$, both of which are $O(1)$ by conservation of energy. However, we also know that $Z \equiv Z_1 - Z_2$ is an odd function of u. This symmetry is crucial for proving that the contribution from Z averages out at $O(\varepsilon)$.

Let $T = O(\varepsilon)$ denote the slow period, set $\tau_n = nT$, $y_n = y(nT)$, and integrate in time,

$$y(\tau_n + \sigma) = S(\tau_n + \sigma; \tau_n)\, y_n + \int_{\tau_n}^{\tau_n+\sigma} S(\tau_n+\sigma;\rho)\, G(\rho)\, \mathrm{d}\rho\,. \tag{99}$$

We first perform a rough estimate on the integral in (99), then bootstrap to get the final, more refined estimate we need to prove the long-time result. Note that if $\sigma = O(\varepsilon)$, Corollary 8 provides an $O(1)$ bound on $S(\tau_n + \sigma; \rho)$. We now look at each of the terms contained in G under the integral. Due to Theorem 6,

$$\int_{\tau_n}^{\tau_n+\sigma} S(\tau_n+\sigma;\rho)\, w_2(\rho)\, \mathrm{d}\rho = O(\varepsilon^2)\,. \tag{100}$$

The next two terms in G also contribute at $O(\varepsilon^2)$ under the integral. Since $\mathrm{Hess}\, F_{\text{var}}(u) = O(1/\varepsilon)$,

$$\int_{\tau_n}^{\tau_n+\sigma} S(\tau_n+\sigma;\rho)\, y(\rho)^T \;\mathrm{Hess}\, F_{\text{var}}(u(\rho))\, y(\rho)\, \mathrm{d}\rho = O(\|y_n\|^2)\,. \tag{101}$$

Note that we have replaced $\sup_{\sigma\in[0,T]} \|y(\tau_n+\sigma)\|^2$ as would naturally appear on the right of (101) by $\|y_n\|^2$. This is justified so long as $\|y(\rho)\| \le 1$. Under this assumption, over a single slow period, $\|y(\tau_n+\sigma)\| \le c\, \|y_n\|$, as can easily be seen by taking the norm of (99) and using the Gronwall inequality.

All other terms contributing to (99) are of higher order than the terms discussed, so that, altogether, we obtain the first stage estimate

$$y(\tau_n+\sigma) = S(\tau_n+\sigma;\tau_n)\, y_n + O(\varepsilon^2, \|y_n\|^2)\,. \tag{102}$$

We can improve this estimate by one order as follows. Take $\sigma = T$, so that, by Lemma 7, $S(\tau_{n+1};\tau_n) = \mathrm{e}^\Lambda$. Again, we write

$$y_{n+1} = \mathrm{e}^\Lambda\, y_n + \int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\rho)\, G(\rho)\, \mathrm{d}\rho \tag{103}$$

and look at each of the contributions to the integral in turn. First,

$$\int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\rho)\, w_2(\rho)\, \mathrm{d}\rho = O(\varepsilon^3) \tag{104}$$

due to Corollary 10 which picks up cancellations by averaging fast motion over one slow period. No bootstrapping is required here; in fact, we could have used this estimate in the first step, too, but the comparison between the two stages clearly demonstrates that extra work is required on each of the terms to get to the longer time scale. Next,

$$\varepsilon \int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\rho)\, \mathrm{D}Z(u+\lambda y)(\rho)\, y(\rho)\, \mathrm{d}\rho = O(\varepsilon^2\, \|y_n\|) \tag{105}$$

by direct estimation; the higher order remainder terms also receive their obvious respective bounds by direct estimation.

Two terms remain which are not generally of the order required, and which we must carry through the computation explicitly. We split the error into a part which is propagated by the slow fundamental matrix S over one slow period T and a remainder which can be estimated. I.e., we write $y(\rho) = S(\rho;\tau_n)y_n + O(\varepsilon^2, \|y_n\|^2)$, see (102), so that

$$\begin{aligned} &\int_{\tau_n}^{\tau_{n+1}} S(\tau_{n+1};\rho)\, y(\rho)^T \ \operatorname{Hess} F_{\mathrm{var}}(u(\rho))\, y(\rho)\, \mathrm{d}\rho \\ &= \int_0^T S(T;\sigma)\, y_n^T\, S(\sigma;0)^T \ \operatorname{Hess} F_{\mathrm{var}}(u(\sigma))\, S(\sigma;0)\, y_n\, \mathrm{d}\sigma + O(\varepsilon^2\, \|y_n\|, \|y_n\|^3)\,. \end{aligned} \tag{106}$$

We have, so far, shown that

$$y_{n+1} = \mathrm{e}^{\Lambda}\, y_n + \int_0^T S(\tau_{n+1};\tau_n+\sigma)\, H(S(\sigma;0)\, y_n, \sigma)\, \mathrm{d}\sigma + O(\varepsilon^3, \|y_n\|^3) \tag{107}$$

where

$$H(w,\sigma) = \varepsilon\, Z(u(\sigma)) + \tfrac{1}{2}\, w^T \ \operatorname{Hess} F_{\mathrm{var}}(u(\sigma))\, w\,. \tag{108}$$

(Here and in the following we use Young's inequality to eliminate mixed terms from the higher order remainders.) Note that we have already used the T-periodicity of u to remove any n-dependence from (108). By recursive insertion, noting that $y_0 = 0$, we obtain the discrete mild formulation

$$y_n = \sum_{k=1}^{n} \mathrm{e}^{(n-k)\Lambda} \left(\int_0^T S(T;\sigma)\, H(S(\sigma;0)\, y_{k-1}, \sigma)\, \mathrm{d}\sigma + O(\varepsilon^3, \|y_{k-1}\|^3) \right). \tag{109}$$

The crucial observation is that, due to the even symmetry of the potential, there are cancellations between consecutive half-periods. This can be seen as follows.

When V is even, not only is u periodic with period $T = O(\varepsilon)$, but we also have $u(\tau) = -u(\tau + T/2)$. Since $\mathrm{D}F_{\mathrm{var}}$ is an even function of u, $\mathrm{D}F_{\mathrm{var}}(u(\sigma))$ is actually periodic with period $T/2$. We express S in its Floquet factorization afforded by Lemma 7 and note that P must then also be $T/2$-periodic. Moreover, by standard arguments, there also exists a "reverse" Floquet factorization

$$S(\tau;\sigma) = \mathrm{e}^{\Lambda(\tau-\sigma)/T}\, \tilde{P}(\tau;\sigma) \tag{110}$$

where the monodromy matrix e^{Λ} is as in (77), while $\tilde{P}$ is also $T/2$-periodic but may differ from P. Breaking the integral in (109) into consecutive half-periods and noting that $H(w,\sigma) = -H(w,\sigma+T/2)$ for fixed w, we obtain

$$\begin{aligned} \int_0^T & S(T;\sigma)\, H(S(\sigma;0)\, w, \sigma)\, \mathrm{d}\sigma \\ &= \int_0^{T/2} \mathrm{e}^{\Lambda(1-\sigma/T)}\, \tilde P(T/2;\sigma)\, H(P(\sigma;0)\, \mathrm{e}^{\Lambda\sigma/T}\, w, \sigma)\, \mathrm{d}\sigma \\ &\quad - \int_0^{T/2} \mathrm{e}^{\Lambda(1/2-\sigma/T)}\, \tilde P(T/2;\sigma)\, H(P(\sigma;0)\, \mathrm{e}^{\Lambda(1/2+\sigma/T)}\, w, \sigma)\, \mathrm{d}\sigma\,. \end{aligned} \tag{111}$$

We note, by changing into a basis in which Λ takes its Jordan normal form

$$\Lambda = \begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}\,, \tag{112}$$

that

$$\mathrm{e}^{\Lambda(1-\sigma/T)} - \mathrm{e}^{\Lambda(1/2-\sigma/T)} = \mathrm{e}^{\Lambda(1/2+\sigma/T)} - \mathrm{e}^{\Lambda\sigma/T} = \tfrac12\,\Lambda\,. \tag{113}$$

Thus, by adding and subtracting suitable intermediate terms in (111), we find that

$$\int_0^T S(T;\sigma)\, H(S(\sigma;0)\, w, \sigma)\, \mathrm{d}\sigma = \Lambda\, H_1(w,w) + H_2(w,\Lambda w)\,, \tag{114}$$

where, for $i=1,2$,

$$H_i(v,w) = O(\varepsilon^2, \|v\|\, \|w\|)\,. \tag{115}$$

Inserting (114) back into (109), we have

$$\begin{aligned} y_n &= \sum_{k=1}^n \mathrm{e}^{(n-k)\Lambda} \bigl[\Lambda\, H_1(y_{k-1},y_{k-1}) + H_2(y_{k-1},\Lambda y_{k-1}) + O(\varepsilon^3, \|y_{k-1}\|^3)\bigr] \\ &= \sum_{k=1}^n \bigl[\Lambda\, H_1(y_{k-1},y_{k-1}) + \mathrm{e}^{(n-k)\Lambda} \bigl(H_2(y_{k-1},\Lambda y_{k-1}) + O(\varepsilon^3, \|y_{k-1}\|^3)\bigr)\bigr] \end{aligned} \tag{116}$$

since $\Lambda^2 = 0$ and therefore $\mathrm{e}^{m\Lambda} = I + m\Lambda$.

Note that $\mathrm{e}^{(n-k)\Lambda}$ contributes secular growth, i.e. it is $O(1/\varepsilon)$. This contribution needs to the controlled by treating Λy_k special. We calculate

$$\varepsilon^{-1}\, \Lambda\, y_n = \sum_{k=1}^n \bigl[\varepsilon^{-1}\, H_2(y_{k-1},\Lambda y_{k-1}) + O(\varepsilon^2, \varepsilon^{-1}\, \|y_{k-1}\|^3)\bigr]\,. \tag{117}$$

Then, taking norms and setting

$$z_n = \|y_n\| + \varepsilon^{-1}\, \|\Lambda y_n\|\,, \tag{118}$$

we find that each of the terms within the summations in (116) and (117) is of the order

$$O\biggl(\varepsilon^2, \varepsilon\, z_k, z_k^2, \frac{z_k^3}{\varepsilon}\biggr) = O\biggl(\varepsilon^2, \frac{z_k^3}{\varepsilon}\biggr) \tag{119}$$

so long as $n = O(1/\varepsilon)$. In other words, there exists a constant c such that, using (116) and (117),

$$z_n \leq \frac{c}{\varepsilon} \sum_{k=0}^{n-1} z_k^3 + c\,\varepsilon\,. \tag{120}$$

If $z(\tau)$ is a monotonic interpolant of the z_k, we also have

$$z(\tau) \leq \frac{c}{\varepsilon^2} \int_0^\tau z^3(\sigma)\, \mathrm{d}\sigma + c\,\varepsilon \tag{121}$$

which, by the generalized Gronwall inequality [8], implies that

$$z(\tau) \leq \frac{\varepsilon}{\sqrt{c^{-2} - 2c\tau}} \,. \tag{122}$$

We conclude that there exists a τ^* independent of ε such that $z(\tau) \leq 2c\varepsilon$ for every $\tau \in [0, \tau^*]$. This bound on z directly implies a like bound on y. This completes the proof of Theorem 1.

8. Discussion

This paper demonstrates for a simple, low dimensional model of balance that trajectories near the slow manifold of the full equation are shadowed by trajectories of the balanced dynamics over very long times. While our main interest lies in variational slow systems, we used a nonvariational system as an intermediate step in the proof. The technical reason is that perturbative methods, the simple construction outlined in Section 2 or the version based on the variation of constants formula used in the actual proof, are effectively based on Taylor expansions in the equations of motion which generally destroy the variational structure. Even though this intermediate step appears unsatisfactory, we are not aware of a method to infer shadowing of trajectories directly from the properties of the two Lagrangians.

The generalization of the long-time shadowing result to larger systems is restricted by two main assumptions. First, we assume a "no chaos" condition on the slow dynamics, here guaranteed by the two-dimensionality of the reduced flow. Second, and perhaps more surprising, we need to assume a symmetry condition on the potential which suppresses slow-slow resonances.

In higher dimensions, we would expect non-resonance conditions to be much harder to prove. When generalizing to higher dimensional flows, the concept of trajectory stability of the slow dynamics will have to be replaced by stability of the slow manifold. In this sense, the statement that nearly balanced initial states remain nearly balanced over very long times may survive into the geophysical context, but further work is clearly needed.

Acknowledgments

This question addressed in this paper was raised while the first two authors were visiting the Mathematisches Forschungsinstitut Oberwolfach, whose hospitality we gratefully acknowledge, as part of the "Research in Pairs" program. MO also thanks the Courant Institute of Mathematical Sciences for its hospitality during the finalization of the manuscript. GAG is partly supported by ARC grant DP0452147.

References

[1] A. Babin, A. Mahalov, and B. Nicolaenko, *Global splitting and regularity of rotating shallow-water equations*, European J. Mech. B Fluids **16** (1997), 725–754.

[2] C. Cotter and S. Reich *Semi-geostrophic particle motion and exponentially accurate normal forms*, SIAM Multiscale Modelling **5** (2006), 476-496.

[3] P.F. Embid and A.J. Majda, *Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity*, Comm. Partial Differential Equations **21** (1996), 619–658.

[4] N. Fenichel, *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J. **21** (1971), 193–225.

[5] R. Ford, M. McIntyre and W. Norton, *Balance and slow quasimanifold: Some explicit results*, J. Atmos. Sci. **57** (2000), 1236–1254.

[6] D. Givon, R. Kupferman, and A. Stuart, *Extracting macroscopic dynamics: model problems and algorithms*, Nonlinearity **17** (2004), R55–R127.
[7] G. Gottwald and M. Oliver, *Slow dynamics via degenerate variational asymptotics*, submitted for publication.
[8] P. Hartmann, "Ordinary Differential Equations," Birkhäuser, Basel, 1982.
[9] E.J. Hinch, "Perturbation Methods," Cambridge Texts in Applied Mathematics, Cambridge University Press, 1991.
[10] R.S. MacKay, *Slow Manifolds*, in "Energy Localisation and Transfer", eds. T. Dauxois, A. Litvak-Hinenzon, R.S. MacKay, A. Spanoudaki, World Scientific, Singapore, 2004, pp. 149–192.
[11] M. McIntyre and W. Norton, *Potential vorticity inversion on a hemisphere*, J. Atmos. Sci. **57** (2000), 1214–1235; *Corrigendum* **58**, 949.
[12] A.I. Neishtadt, *On delayed stability loss under dynamical bifurcations I*, Differ. Equ. **23** (1987), 1385–1390.
[13] N. Nekhoroshev, *An exponential estimate of the time of stability of a nearly integrable Hamiltonian system*, Russ. Math. Surveys **32** (1977), 1–65.
[14] M. Oliver, *Variational asymptotics for shallow water near geostrophy: A transformational approach*, J. Fluid Mech. **551**, 197–234.
[15] J.P. Pedlosky, "Geophysical Fluid Dynamics," Springer-Verlag, 1987.
[16] J. Pöschel, *On Nekoroshev's estimates at an elliptic equilibrium*, Int. Math. Res. Not. **4** (1999), 203–215.
[17] R. Salmon, *Practical use of Hamilton's principle*, J. Fluid. Mech. **132** (1983), 431–444.
[18] R. Salmon, *New equations for nearly geostrophic flow*, J. Fluid Mech. **153** (1985), 461–477.
[19] R. Salmon, "Lectures on geophysical fluid dynamics," Oxford University Press, New York, 1998.
[20] J.A. Sanders and F. Verhulst, "Averaging methods in nonlinear dynamical systems," Applied Mathematical Sciences **59**, Springer-Verlag, New York, 1985.
[21] J. Stoer and R. Burlisch, "Introduction to numerical analysis," second edition, Springer Verlag, New York, 1993.
[22] E. Tomboulis, *Canonical quantization of nonlinear waves*, Phys. Rev. D **12** (1975), 1678–1683.
[23] A. N. Tikhonov, A. B. Vasil'eva, and A. G. Sveshnikov, "Differential equations," Springer-Verlag, New York, 1980.
[24] F. Verhulst, "Nonlinear Differential Equations and Dynamical Systems," Springer-Verlag, New York, 2000.
[25] F. Verhulst, "Methods and Applications of Singular Peturbations," Texts in Applied Mathematics **50**, Springer-Verlag, Berlin, 2005.
[26] D. Wirosoetisno, *Exponentially accurate balance dynamics*, Adv. Diff. Eq. **9** (2004), 177–196.
[27] G.B. Whitham, "Linear and Nonlinear Waves," Wiley, New York, 1974.
[28] C.R. Willis, M. El-Batanouny, S. Burdick, R. Boesch, and P. Sodano, *Hamiltonian dynamics of the double sine-Gordon kink*, Phys. Rev. B **35** (1987), 3496–3505.

(G. Gottwald) School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia
E-mail address, G. Gottwald: `gottwald@maths.usyd.edu.au`

(M. Oliver) School of Engineering and Science, International University Bremen, 28759 Bremen, Germany
E-mail address, M. Oliver: `oliver@member.ams.org`

(N. Tecu) Department of Mathematics, Yale University, New Haven, CT 06520, USA
E-mail address, N. Tecu: `nicolae.tecu@yale.edu`