SLOW DYNAMICS VIA DEGENERATE VARIATIONAL ASYMPTOTICS

GEORG GOTTWALD AND MARCEL OLIVER

Abstract. We introduce the method of degenerate variational asymptotics for a class of singularly perturbed ordinary differential equations whose leading order behavior is dominated by gyroscopic forces. Such systems exhibit dynamics on two separate time scales which are dynamically linked with no explicit splitting into slow and fast subsystems. We derive approximate equations for the slow motion to arbitrary order by performing an asymptotic expansion of the Lagrangian rather than the Euler–Lagrange equations of motion themselves.

Rigorous justification of the method is provided in two different settings. For harmonic potentials, we show that the method can be understood explicitly in terms of perturbation theory for finite dimensional linear eigenvalue problems. In the general case, we resort to an indirect analysis involving a nonvariational auxiliary model. We illustrate our analytical results by numerical simulation.

1. Introduction

Many problems in science and engineering involve motion on different temporal and spatial scales. While often only features on the largest and slowest scales are of practical interest, fast modes may determine even leading order slow motion—as is the case in classical molecular dynamics—or pose prohibitive time step restrictions onto numerical simulations of the full dynamics even if the initial state is free of fast modes. The latter situation is typically encountered in free surface and shallow water fluid flow. Both from a practical point of view and as a theoretical tool for understanding fundamental processes, it is important to find effective reduced models that capture the large scale behavior without the need to explicitly compute on small scales.

The reduction to slow dynamics can often be formulated as the construction of a slow manifold for the parent system. The theory is well-developed in the normally hyperbolic case where a variety of methods, including asymptotic expansions, multiple scale expansions, and averaging over fast modes have been successful [7, 22, 24, 9, 26, 11, 27]. Hamiltonian system, on the other hand, are not generically normally hyperbolic and thus require different tools. One path toward the construction of slow manifolds for Hamiltonian systems is based on canonical transformations to normal form as, for example, in [25, 28]; see MacKay [13] for an excellent review on the construction of slow manifolds. When the equations of motion arise as Euler–Lagrange equations of a variational principle, it is also possible to find reduced models by approximating the underlying Lagrangian. This strategy is often referred to as *variational asymptotics*.

Date: January 22, 2007.

Hamiltonian and Lagrangian perturbation methods have the distinct advantage that the procedure leaves the conservation laws intact, or produces proper reduced analogs of the conservation laws so long as the approximation respects the symmetries of the variational formulation. This is often desirable for simulations of "typical" system behavior beyond the time horizon up to which model and data uncertainties are sufficiently small to allow prediction of individual trajectories.

Lagrangian perturbation theory is particularly natural in fluid mechanics, but has also been used elsewhere; Salmon [20] pioneered this point of view in the context of rotating shallow water equations where slow motion corresponds to large scale quasi two-dimensional transport while the fast modes correspond to so called gravity waves which describe oscillations of the stratification surfaces. Salmon noticed that some of the classical reduced models can be reinterpreted in terms of variational asymptotics. Moreover, he found new models with varying regularity of the solutions. More recently, Salmon's ideas were generalized in [8, 15]. The latter work, in particular, systematically leverages the observation that model reduction is achieved by letting the variational principle degenerate. The emerging degeneracy conditions are weaker than explicit constraints as have been used, for example, by Salmon. As a result, it is possible to generate continuous families of reduced models. From among these an "optimal" slow dynamics can be selected by taking the regularity properties of the reduced system into account. In particular, it was conjectured in [15] that there is a distinguished reduced model for rotating shallow water at first order whose solutions are most consistent with the *a priori* long-time large-scale modeling regime.

From the viewpoint of formal variational asymptotics, [15] gives a rather complete picture of the different pathways to model reduction for rotating shallow water, and ties together a number of well-explored approaches within a unified variational framework. In the infinite dimensional setting of rotating fluid flow, however, rigorous justification of the model reduction is harder and has only been done in special cases [1, 14]. Nevertheless, in the geophysical fluid dynamics community it is generally believed that balanced models as, for example, the quasi-geostrophic equations [16], approximate the full dynamics on long time-scales for well-balanced initial states.

In this paper, we study variational perturbation theory in the context of a structurally similar, simpler problem: systems of ordinary differential equations of the form

$$\varepsilon\, S\ddot{q}_\varepsilon - R(q_\varepsilon)\dot{q}_\varepsilon + \nabla V(q_\varepsilon) = 0\,, \tag{1}$$

where $q_\varepsilon \colon [0,T] \to \mathbb{R}^{2d}$, which arises as the Euler–Lagrange equation from the Lagrangian

$$L_\varepsilon = \frac{\varepsilon}{2}\,\dot{q}_\varepsilon^T M \dot{q}_\varepsilon - V(q_\varepsilon) - \frac{1}{2}\,\dot{q}_\varepsilon^T F(q_\varepsilon) \tag{2}$$

with

$$S = \frac{M + M^T}{2} \qquad \text{and} \qquad R(q_\varepsilon) = \frac{\nabla F(q_\varepsilon) - \nabla F(q_\varepsilon)^T}{2}\,. \tag{3}$$

When $d = 1$, the model describes the motion of a single charged particle in a planar potential V under the influence of a magnetic field normal to the plane of motion. The limit $\varepsilon \to 0$ corresponds to the mass of the particle going to zero while its charge is held constant. Moreover, the Lagrangian (2) is the natural finite dimensional analog of the rotating shallow water Lagrangian [21]—the magnetic term $R(q_\varepsilon)\dot{q}_\varepsilon$ corresponding to the Coriolis term and the potential term $\nabla V(q_\varepsilon)$

to the surface height field—so that our system can be considered a toy model for balance in rotating fluid flow. The derivation of (1) in this context is as follows. The rotating shallow water equations are

$$\frac{\mathrm{D}u}{\mathrm{D}t} + f_0\, Ju + g\,\nabla h = 0\,, \tag{4a}$$

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\,\nabla\cdot u = 0\,, \tag{4b}$$

where

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{5}$$

denotes the canonical symplectic matrix. 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 [16]. If we non-dimensionalize by introducing a typical velocity U, a typical horizontal length scale L and a typical layer depth H, we can rewrite the rotating shallow water equations as

$$\varepsilon\,\frac{\mathrm{D}u}{\mathrm{D}t} + Ju - \frac{B}{\varepsilon}\,\nabla h = 0\,, \tag{6a}$$

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\,\nabla\cdot u = 0\,, \tag{6b}$$

where $\varepsilon = U/f_0L$ is the Rossby number and $B = (L_R/L)^2$ the Burger number with Rossby radius of deformation $L_R = \sqrt{gH}/f_0$. Our model equation (1) corresponds to the so-called *semigeostrophic scaling* $B = O(\varepsilon)$ with $\varepsilon \ll 1$ [16, 21] for a prescribed layer depth potential and the setting $S = 1$, $R(q_\varepsilon) = J$. In this context, $\varepsilon \to 0$ is analogous to rapid rotation in the semigeostrophic scaling [16] where the fluid motion is dominated by large vortices as we know them from weather maps.

The defining feature of the system is that its formal leading order Lagrangian is affine, i.e. is linear in the velocity. The main idea behind our method is raising this property to any given order in ε through carefully chosen near-identity changes of coordinates. This pushes the non-affine contributions into higher order. By truncating the ε-expansion of the Lagrangian, dropping all non-affine terms, we derive Euler–Lagrange equations which approximate the slow dynamics to the chosen order and which, by construction, are conservative.

A second feature of (1) is that the separation of scales is not explicit. Heuristically, there is fast oscillatory dynamics through the balance between the inertia term $\varepsilon\,\ddot{q}_\varepsilon$ and the magnetic term $R(q_\varepsilon)\dot{q}_\varepsilon$ with corresponding fast time scale $\tau = t/\varepsilon$. There is also an underlying slow dynamics through the balance of the magnetic term and the potential force ∇V. The smaller ε, the larger is the separation between the two time-scales. (Numerical illustration is given toward the end of the paper in Figures 1 and 2, which show projected particle trajectories as well as the approximate slow manifolds.) One is typically interested in distilling the slow dynamics, and finding approximate equations for it. However, these two balances are intricately linked—there is no component of q_ε which is entirely fast or slow. Even in the case of a harmonic potential where the system splits into fast and slow linear eigenspaces, each of dimension $2d$, the eigenvectors do not align with coordinate axes and exhibit nontrivial dependence on ε. This feature prevents direct application of common techniques such as the method of averaging [26] or weak convergence techniques as in [3].

Our goals are, on the one hand, to take a first step toward a rigorous understanding of Hamiltonian balance models in fluid dynamics. On the other hand, to provide an asymptotic analysis of system (1) which is both explicit and rigorous. To the best of our knowledge, this has not been done elsewhere. Reich and Cotter [5] point out that abstract Hamiltonian normal form theory is applicable and prove that solutions near the slow manifold remain close to it for exponentially long times. In principle, normal form theory also provides tools to construct equations for the solution on the slow manifold, but these are hard to implement in practice beyond order one. Our method, in comparison, requires no integration and can thus be carried out by rote, for example on a computer algebra system. Moreover, it exposes the inherent freedom in the choice of coordinates and allows to easiy obtain models in non-canonical coordinates, thereby providing enough generality in an infinite dimensional setting to ensure well-posedness and regularity of the resulting slow dynamics.

We note that this problem is different from singular perturbation problems with a strong constraining force as have been studied by Takens [23], Bornemann [3], and others. In their case, leading order slow frequencies depend on averages over fast degrees of freedom while in our case the leading order slow motion completely decouples from the fast modes.

A crucial ingredient for our results is the existence of a conserved energy,

$$E_\varepsilon(t) = \frac{\varepsilon}{2}\,|\dot q_\varepsilon(t)|^2 + V(q_\varepsilon(t)) = E_\varepsilon(0)\,. \tag{7}$$

We assume throughout that $V(q)$ is strictly convex for large values of q, so that $q_\varepsilon(t)$ is bounded independent of ε for all times. We also assume that V is sufficiently smooth so that all necessary derivatives exist and are continuous. Finally, unless explicitly stated, we make the simplifying assumption that $M = I$ and $R(q) = Jq$, where J denotes the canonical symplectic $2d \times 2d$ matrix

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\,. \tag{8}$$

There is no inherent obstacle to studying more general cases, even with non-quadratic or degenerate energies. We can now state our main result.

Theorem 1. *For $n \geq 0$ and $q_0 \in \mathbb{R}^{2d}$ fixed, let $q(t)$ denote a solution to the "near-near-identity" slow limit system of order n*

$$\dot q = F^n_{\text{var}}(q) \tag{9}$$

with $q(0) = q_0$, where the algorithm for the construction of F^n_{var} is detailed in Section 2. Let $q_\varepsilon(t)$ solve the full parent dynamics (1) *consistently initialized via $q_\varepsilon(0) = q_0$ and $\dot q_\varepsilon(0) = F^n_{\text{var}}(q_0)$. Then for every fixed $T > 0$ there exists $\varepsilon_0 > 0$ and $c = c(q_0, T)$ such that*

$$\sup_{t\in[0,T]} \|q_\varepsilon(t) - q(t)\| \leq c\,\varepsilon^{n+1} \tag{10}$$

for all $0 < \varepsilon \leq \varepsilon_0$.

Remark 1. The general construction of Section 2 may entail changes of coordinates at an order less or equal to n. In this case, the transformation must be explicitly applied both to the consistent initialization and to the reduced solution before taking the difference in (10). See Theorem 6 for details.

The remainder of this paper is structured as follows. In Section 2, we introduce the formalism of degenerate variational asymptotics and apply it to our model problem. We analyze the linear case in Section 3, then turn to the general case of an arbitrary nonlinear potential in Section 4. The paper closes with a numerical demonstration and a short discussion of the results.

2. Degenerate variational asymptotics

Variational asymptotics refers to perturbation techniques that are performed at the level of the variational principle. The actual reduced equations of motion are obtained only at the very end as the Euler–Lagrange equations of the perturbed Lagrangian. The main idea in our situation is to construct a near-identity change of coordinates which renders the Lagrangian degenerate up to the desired order of accuracy. This imposed degeneracy of the Lagrangian removes the terms which contribute second or higher-order time derivatives in the associated Euler-Lagrange equations and which are responsible for the fast time scale. We will show later that the resulting slow equations approximate the dynamics on the slow manifold of the full system. The formal construction proceeds in four steps.

Step 1. Introduce the near-identity transformation

$$q_\varepsilon = q + \varepsilon\, q' + \tfrac{1}{2}\,\varepsilon^2\, q'' + \cdots + q^{[n]}\,, \tag{11}$$

where q_ε represents the old, physical coordinates and q the new, yet-to-be-determined coordinates. We will later choose $q^{[1]}, \ldots, q^{[n]}$ as functionals of q. At this point, it is important to note that (11) is not a Taylor expansion of a function but rather a transformation between arguments of the action functionals associated to the respective Lagrangians. However, it can be treated formally like any power series expansion.

Step 2. Plug the transformation into the parent Lagrangian L_ε, expand, and collect powers of ε. Arbitrary time derivatives, which are null-Lagrangians as they do not contribute to the equations of motion, may be added or subtracted as convenient.

Step 3. Choose, iteratively at order $\varepsilon^1, \ldots, \varepsilon^n$ the corresponding $q^{[1]}, \ldots, q^{[n]}$ as functions of q and its time derivatives such that each L_i for $i = 1, \ldots, n$ becomes affine. In other words, only terms which are at most linear in $\dot q$ must remain.

Step 4. Truncate terms of order larger than ε^n and compute the Euler–Lagrange equations of the resulting reduced Lagrangian.

We will now perform this procedure for our model system (1) with $M = I$ and $R(q) = Jq$. First, however, a remark on notation. 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.

$$(v^T \operatorname{Hess} q\, u)_i \equiv v^T \operatorname{Hess} q_i\, u\,. \tag{12}$$

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

The parent Lagrangian with $M = I$ and $R(q) = Jq$ reads

$$L_\varepsilon = \frac{\varepsilon}{2}\,|\dot q_\varepsilon|^2 - V(q_\varepsilon) - \frac{1}{2}\,\dot q_\varepsilon^T J q_\varepsilon \,. \tag{13}$$

Executing Steps 1 and 2, we obtain

$$\begin{aligned} L_\varepsilon &= L_0 + \varepsilon\, L_1 + \tfrac{1}{2}\,\varepsilon^2\, L_2 \\ &= -V(q) - \tfrac{1}{2}\,\dot q^T J q \\ &\quad + \varepsilon \left[\tfrac{1}{2}\,|\dot q|^2 - \mathrm{D}V(q)\, q' - \tfrac{1}{2}\,\dot q'^T J q - \tfrac{1}{2}\,\dot q^T J q'\right] \\ &\quad + \tfrac{1}{2}\,\varepsilon^2 \left[2\,\dot q^T \dot q' - q'^T \mathrm{D}\nabla V(q) q' - \mathrm{D}V(q)\, q'' - \tfrac{1}{2}\,\dot q''^T J q - \dot q'^T J q' - \tfrac{1}{2}\,\dot q^T J q''\right] \\ &= -V(q) - \tfrac{1}{2}\,\dot q^T J q \\ &\quad + \varepsilon \left[\tfrac{1}{2}\,|\dot q|^2 - \mathrm{D}V(q)\, q' - \dot q^T J q'\right] \\ &\quad + \tfrac{1}{2}\,\varepsilon^2 \left[2\,\dot q^T \dot q' - q'^T \mathrm{D}\nabla V(q) q' - \mathrm{D}V(q)\, q'' - \dot q'^T J q' - \dot q^T J q''\right] \end{aligned} \tag{14}$$

up to terms of order ε^3 and perfect time derivatives.

At the leading order, our model Lagrangian is already affine—this is clearly necessary for the method to work. The corresponding leading order Euler–Lagrange equations are

$$J\dot q = \nabla V(q)\,; \tag{15}$$

Details of this computation are provided in Appendix A.

At order ε, we need to execute Step 3: remove the quadratic term $\frac{1}{2}\,|\dot q|^2$ which contribute to fast motion on the time scale $1/\varepsilon$. This necessitates

$$q' = -\tfrac{1}{2}\, J\dot q + f(q)\,, \tag{16}$$

where $f(q)$ is so far arbitrary. We specialize in choices where all terms have the same homogeneity and which, in particular, may annihilate q' under the leading order dynamics (15). This singles out the family of transformations

$$q' = -\tfrac{1}{2}\, J\dot q + \mu\,\nabla V(q)\,. \tag{17}$$

We will show later that the order of the method is not affected by the choice of the parameter μ.

Inserting (17) back into (14) and computing the Euler–Lagrange equations up to $O(\varepsilon)$, we obtain the first-order slow reduced system

$$\left[1 + \varepsilon\,(\tfrac{1}{2} + \mu)\,\Delta V\right] J\dot q = \nabla V + 2\varepsilon\mu\,\mathrm{D}\nabla V\,\nabla V\,. \tag{18}$$

The choices $\mu = \pm\frac{1}{2}$ are of special interest. When $\mu = \frac{1}{2}$, the transformation (16) reduces to the identity up to terms of order ε^2 due to the leading order balance (15); we will refer to this case as the *near-near-identity* transformation. Theorem 1, as literally stated, only applies to this case. When $\mu = -\frac{1}{2}$, the reduced model is canonical. The two cases correspond, in the geophysical setting, to the two models which have been derived in Salmon [20]—cf. the discussion in [15].

At general order i, the computation proceeds as follows. First, note that $q^{[i]}$ always appears in the combination

$$L_i = -\mathrm{D}V(q)\, q^{[i]} - \dot q^T J q^{[i]} + R_i\,, \tag{19}$$

where R_i denotes remainder terms of lower order with respect to derivatives in ε. Second, note that the substitution of all previous $q^{[1]}, \dots, q^{[i-1]}$ into the expression for R_i yields terms containing at most $i{+}1$ time derivatives. We now let j denote the maximum number of time derivatives which appear in any term of the remainder, initially set $q^{[i]} = 0$, and iterate down from $j = i+1, \ldots, 1$ through the following procedure.

Step 3.1. Collect terms with j time derivatives and write them in the form $\dot{q}^T S_{j-1}$, where S_{j-1} contains terms with at most $j-1$ time derivatives. This can always be done by adding perfect time derivatives or, equivalently, by integration by parts under the action integral.

Step 3.2. Add the term $-JS_{j-1}$ to the expression for $q^{[i]}$ computed so far. It is easily checked that this will remove all terms with j time derivatives from L_i. At the same time, this will introduce, through $-\mathrm{D}V\, q^{[i]}$, further terms containing $j-1$ time derivatives, which have to be kept for the next step of the iteration.

Step 3.3. While $j > 1$, decrease j by one and go to Step 3.1.

Appendix B gives the details of the computation up to second order. Higher order computations are even more lengthy, though conceptionally straightforward. We expect that a relatively easy closed form recursive expression exists, but we have not yet been able to write it out.

Remark 2. When J is replaced by a more general skew term $R(q)$, the above recursion is guaranteed to be solvable so long as $R(q)$ is nondegenerate.

Remark 3. The general near-near-identity transformation at order i is obtained as follows. First, use the Euler–Lagrange equations at order $i-1$ to eliminate all time derivatives from the $q^{[i]}$ obtained as above. Since the resulting expression is only a function of q, subtracting it from the actual $q^{[i]}$ will yield a new $q^{[i]}$ that still satisfies the degeneracy condition and also vanishes formally to order $i-1$.

Remark 4. The reduced equations (18) can also be derived in the following *ad hoc* manner. Split the two time-derivatives in the kinetic energy term of the Lagrangian (13) according to

$$\dot{q}_\varepsilon = \mu_i\, \dot{q} - (1-\mu_i)\, J\nabla V + O(\varepsilon) \qquad \text{for } i = 1, 2\,. \tag{20}$$

The case $\mu_{1,2} = 1$ corresponds to no splitting and the case $\mu_{1,2} = 0$ corresponds to a complete substitution of the kinetic term with the leading order balance. The parent Lagrangian (13) then becomes

$$L_\varepsilon = \frac{\varepsilon}{2} \bigl[\mu_1\, \dot{q}_\varepsilon - (1-\mu_1)\, J\nabla V\bigr] \cdot \bigl[\mu_2\, \dot{q}_\varepsilon - (1-\mu_2)\, J\nabla V\bigr] - V(q_\varepsilon) - \frac{1}{2}\, \dot{q}_\varepsilon^T J q_\varepsilon\,. \tag{21}$$

The variational derivative of the Lagrangian reads, after elementary manipulations,

$$\begin{aligned} \frac{\delta L_\varepsilon}{\delta q} &= J\dot{q} - \nabla V \\ &\quad + \varepsilon\, \bigl((1-\mu_1)(1-\mu_2)\, \mathrm{D}\nabla V \nabla V + \tfrac{1}{2}\, (\mu_1 - 2\mu_1\mu_2 + \mu_2)\, \Delta V J\dot{q} - \mu_1\, \mu_2\, \ddot{q}\bigr)\,. \end{aligned} \tag{22}$$

The dynamics is slow when either $\mu_1 = 0$ or $\mu_2 = 0$. Without loss of generality we set $\mu_2 = 0$ and define $\mu_1 = \nu$, i.e. we express one time-derivative in the kinetic energy completely via the leading order balance and allow partly substitution for the remaining one. The definition of a parameter $\mu = -(1-\nu)/2$ recovers the variational first-order reduced model (18). Note that if the total kinetic energy is replaced entirely by the leading order balance we obtain the canonical case of (18) with $\mu = -\frac{1}{2}$. Higher order models may be deduced analogously. Note, however, that this derivation loses track of the underlying coordinate transformation. This transformation is essential for consistent initialization and postprocessing of the reduced model solution.

All of the above is formal; there is no *a priori* guarantee that solutions to the reduced equations remain close to solutions of the parent model. We will address this issue in the following two sections, first in the special case of a harmonic potential, then for general nonlinear potentials.

3. Harmonic potentials

In this section we consider the special case when the potential is quadratic, so that the motion is effectively a superposition of harmonic oscillators. Most of the actual results of this section are contained, at least qualitatively, in the general convergence theorem which will be proved in Section 4. However, a separate study of the linear case is useful for several reasons.

First, the linear case can be settled purely at the level of linear algebra, permitting a construction which is explicit enough that it illustrates the workings of the degenerate asymptotics at arbitrary order. Second, it constrains the type of result one may hope to prove in the general case. In particular, the linear case provides examples which show that Theorem 1 as well as the long-time result proved in [10] are essentially sharp. Third, it provides a starting point for addressing the question, first raised in [15], under which circumstances the near-near-identity case of Section 2 may provide more accurate solutions than models based on more general transformations.

We take the following point of view. We start with a parent Lagrangian of the form

$$L_\varepsilon = \frac{\varepsilon}{2} \, |\dot q_\varepsilon|^2 - \frac{1}{2} \, \dot q_\varepsilon^T F q_\varepsilon - \frac{1}{2} \, q_\varepsilon^T \Omega q_\varepsilon \tag{23}$$

where F is skew symmetric and Ω is symmetric, and both matrices are assumed nonsingular. We suppose now that we have already transformed the system such that contributions to the Lagrangian up to order ε^{n-1} are affine. Since any transformation which cancels the singular term also creates higher order time derivatives at higher orders in ε—see (115) and (116) in Appendix B—we shall consider, with slight abuse of notation, the transformed parent Lagrangian

$$L_\varepsilon = \varepsilon^n \, \dot q_\varepsilon^T \, M_\varepsilon[q_\varepsilon] - \frac{1}{2} \, \dot q_\varepsilon^T F_\varepsilon q_\varepsilon - \frac{1}{2} \, q_\varepsilon^T \Omega_\varepsilon q_\varepsilon \, , \tag{24}$$

where

$$M_\varepsilon[q] = \sum_{i=1}^m M_\varepsilon^i \, \frac{\mathrm{d}^i q}{\mathrm{d}t^i} \, , \tag{25}$$

and the matrices M_ε^i, F_ε, and Ω_ε are now finite expansions in powers of ε. F_ε and Ω_ε are not necessarily skew respectively symmetric, but have nonsingular skew respectively symmetric part for small enough ε. We write

$$M_\varepsilon^i = M_i + \varepsilon \, M_i' + \dots \tag{26}$$

with similar expressions for F_ε and Ω_ε. Note that for $n = m = 1$, we recover (23) by setting $M_1 = I$. This is so far exact. We will now show that one step in the transformation and approximation process yields a characteristic eigenvalue problem which is isospectral to the original one up to terms of order ε^{n+1}.

The transformation. Per our general procedure, we seek a transformation of the form

$$q_\varepsilon = q + \varepsilon^n \, \hat{q} \tag{27}$$

which renders the $O(\varepsilon^n)$ contribution of the Lagrangian affine. Inserting (27) into the full Lagrangian (24), we obtain

$$\begin{aligned} L_n &= -\tfrac{1}{2} \, \dot{q}^T F_\varepsilon q - \tfrac{1}{2} \, q^T \Omega_\varepsilon q \\ &\quad + \varepsilon^n \left(\dot{q}^T M[q] - \tfrac{1}{2} \, \dot{\hat{q}}^T F q - \tfrac{1}{2} \, \dot{q}^T F \hat{q} - \tfrac{1}{2} \, \hat{q}^T \Omega q - \tfrac{1}{2} \, q^T \Omega \hat{q} \right) + O(\varepsilon^{n+1}) \\ &= -\tfrac{1}{2} \, \dot{q}^T F_\varepsilon q - \tfrac{1}{2} \, q^T \Omega_\varepsilon q + \varepsilon^n \left(\dot{q}^T M[q] - \dot{q}^T R \hat{q} - q^T S \hat{q} \right) + O(\varepsilon^{n+1}) \,, \end{aligned} \tag{28}$$

where we dropped a perfect derivative in the second step, and where

$$R = \frac{F - F^T}{2} \qquad \text{and} \qquad S = \frac{\Omega + \Omega^T}{2} \tag{29}$$

with $F = F_\varepsilon|_{\varepsilon=0}$, $\Omega = \Omega_\varepsilon|_{\varepsilon=0}$, and $M = M_\varepsilon|_{\varepsilon=0}$.

Degeneracy condition. We now compute $\hat{q}$ such that all second and higher time derivatives from the $O(\varepsilon^n)$ term in (28) are removed. Writing

$$\hat{q} = \sum_{i=0}^{m} Z_i q^{(i)} \qquad \text{with} \qquad q^{(i)} = \frac{\mathrm{d}^i q}{\mathrm{d}t^i} \,, \tag{30}$$

we expand

$$\begin{aligned} \dot{q}^T M[q] - \dot{q}^T R \hat{q} - q^T S \hat{q} &= \sum_{i=1}^{m} \dot{q}^T M_i q^{(i)} - \sum_{i=0}^{m} \dot{q}^T R Z_i q^{(i)} - \sum_{i=0}^{m} q^T S Z_i q^{(i)} \\ &= \dot{q}^T (M_m - R Z_m) q^{(m)} + \sum_{i=1}^{m-1} \dot{q}^T (M_i - R Z_i + S Z_{i+1}) q^{(i)} \\ &\quad - \dot{q}^T (R Z_0 - S Z_1) q - q^T S Z_0 q \,, \end{aligned} \tag{31}$$

where we again consider Lagrangians to be equal if they differ by a perfect time derivative. The choice

$$Z_i = R^{-1} \sum_{j=i}^{m} (S R^{-1})^{j-i} \, M_j \tag{32}$$

for $i = 1, \ldots, m$ annihilates all but the last two terms in (31) and, hence, renders the Lagrangian affine at the specified order in ε; Z_0 is free and will be chosen later. The Lagrangian, transformed and truncated to $O(\varepsilon^n)$, then reads

$$L_{\text{slow}} = -\tfrac{1}{2} \, \dot{q}^T \hat{F}_\varepsilon q - \tfrac{1}{2} \, q^T \hat{\Omega}_\varepsilon q \,, \tag{33}$$

where

$$\hat{F}_\varepsilon = F_\varepsilon + 2 \, \varepsilon^n Z \,, \tag{34}$$

$$\hat{\Omega}_\varepsilon = \Omega_\varepsilon + 2 \, \varepsilon^n \, S Z_0 \,, \tag{35}$$

and

$$Z = R Z_0 - \sum_{i=1}^{m} (S R^{-1})^i \, M_i \,. \tag{36}$$

The near-near-identity case. We now wish to set Z_0 such that $\hat{q} = 0$ up to next order terms. First, we notice that the leading order dynamics

$$R\dot{q} = Sq\,, \tag{37}$$

which is of the same form for the full Lagrangian (24) and for the reduced affine Lagrangian (33), implies that

$$q^{(i)} = (R^{-1}S)^i q\,. \tag{38}$$

Thus, up to next order terms,

$$\begin{aligned} \hat{q} &= \sum_{i=0}^{m} Z_i (R^{-1}S)^i q \\ &= Z_0 q + \sum_{i=1}^{m} \sum_{j=i}^{m} R^{-1} (SR^{-1})^{j-i}\, M_j\, (R^{-1}S)^i q\,, \end{aligned} \tag{39}$$

and the condition for a near-near-identity transformation is

$$\begin{aligned} Z_0 &= -R^{-1} \sum_{i=1}^{m} \sum_{j=i}^{m} (SR^{-1})^{j-i}\, M_j\, (R^{-1}S)^i \\ &= -R^{-1} \sum_{i=1}^{m} \sum_{j=1}^{i} (SR^{-1})^{i-j}\, M_i\, (R^{-1}S)^j\,. \end{aligned} \tag{40}$$

Euler–Lagrange equations and spectral problem. The variational principle of the reduced Lagrangian (33), as detailed in Appendix A, yields the Euler–Lagrange equation

$$\hat{R}_\varepsilon \dot{q} = \hat{S}_\varepsilon q\,, \tag{41}$$

with

$$\hat{R}_\varepsilon = \tfrac{1}{2}\,(\hat{F}_\varepsilon - \hat{F}_\varepsilon^T) = R_\varepsilon + \varepsilon^n\,(Z - Z^T)\,, \tag{42}$$

$$\hat{S}_\varepsilon = \tfrac{1}{2}\,(\hat{\Omega}_\varepsilon + \hat{\Omega}_\varepsilon^T) = S_\varepsilon + \varepsilon^n\,(SZ_0 + Z_0^T S)\,. \tag{43}$$

The corresponding characteristic matrix equation is

$$(\hat{S}_\varepsilon - \lambda \hat{R}_\varepsilon) q = 0\,, \tag{44}$$

where the $O(\varepsilon^n)$ contributions read (again, we are freely changing next order terms)

$$\begin{aligned} &\bigl[SZ_0 + Z_0^T S - \lambda(Z - Z^T)\bigr]\, q \\ &\quad = \left[SZ_0 + Z_0^T S - \lambda R Z_0 + \lambda \sum_{i=1}^{m} (SR^{-1})^i\, M_i - \lambda Z_0^T R - \lambda \sum_{i=1}^{m} M_i^T\, (-R^{-1}S)^i\right] q \\ &\quad = \left[(S - \lambda R)\, Z_0 + \lambda \sum_{i=1}^{m} (SR^{-1})^i\, M_i + \sum_{i=1}^{m} M_i^T\, (-\lambda)^{i+1}\right] q \\ &\quad = \sum_{i=1}^{m} \left[(\lambda R - S)\, R^{-1} \sum_{j=1}^{i} (SR^{-1})^{i-j}\, M_i\, \lambda^j + \lambda (SR^{-1})^i\, M_i + M_i^T\, (-\lambda)^{i+1}\right] q \\ &\quad = \sum_{i=1}^{m} \left[\sum_{j=1}^{i-1} (SR^{-1})^{i-j}\, \lambda^{j+1} - \sum_{j=1}^{i} (SR^{-1})^{i-j+1}\, \lambda^j + \lambda (SR^{-1})^i\right] M_i\, q \end{aligned}$$

$$+ \sum_{i=1}^{m} \big[M_i \, \lambda^{i+1} + M_i^T \, (-\lambda)^{i+1} \big] \, q \, . \tag{45}$$

An index shift in one of the inner sums shows that all terms in the second last line cancel. Hence, the characteristic equation of the reduced slow system reads

$$\Big[S_\varepsilon - \lambda \, R_\varepsilon + \varepsilon^n \sum_{i=1}^{m} \big(M^i \, \lambda^{i+1} + (M^i)^T \, (-\lambda)^{i+1} \big) \Big] \, q = 0 \, . \tag{46}$$

This agrees up to $O(\varepsilon^n)$ with the characteristic matrix equation for the full problem,

$$\Big[S_\varepsilon - \lambda \, R_\varepsilon + \varepsilon^n \sum_{i=1}^{m} \big(M_\varepsilon^i \, \lambda^{i+1} + (M_\varepsilon^i)^T \, (-\lambda)^{i+1} \big) \Big] \, q = 0 \, , \tag{47}$$

which can be easily determined from the full Euler–Lagrange equation, computed in Appendix A.

The equations of motion of the full system (23) in Hamilton form read

$$\begin{pmatrix} F & -\varepsilon I \\ \varepsilon I & 0 \end{pmatrix} \begin{pmatrix} \dot q \\ \dot p \end{pmatrix} = \begin{pmatrix} \Omega & 0 \\ 0 & \varepsilon I \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} . \tag{48}$$

It is therefore an elliptic nondegenerate Hamilton system which, therefore, has semisimple, purely imaginary eigenvalues for any positive ε. (This is well known, but can be seen as follows. Any eigenvalue problem of the form $RSv = \lambda v$ with a nonsingular skew matrix R and a symmetric positive definite matrix S can be transformed into the problem $S^{1/2} R S^{1/2} w = \lambda w$; the claim follows from the normality of the left hand matrix.) Since (24) is obtained from (23) by an exact transformation, it too has the same set of semisimple eigenvalues (since the dimension has increased, there are other eigenvalues that are not of physical interest.) In particular, these eigenvalues are analytic in ε, so that $O(\varepsilon^{n+1})$ perturbations to the characteristic equation result in $O(\varepsilon^{n+1})$ changes in the eigenvalues and eigenvectors [12].

Remark 5. In [15] we observed, for a particular example, that the frequencies were actually correct to $O(\varepsilon^{n+1})$ in the near-near-identity case and $O(\varepsilon^n)$, as expected, in all others. When $n = 1$ so that $m = 1$ and $M_1 = I$, this observation is readily explained in the framework above. In this case, (44) reduces to

$$\big[S - 2\varepsilon \, S R^{-1} R^{-1} S - \lambda \, (R - 2\varepsilon \, (R^{-1} S + S R^{-1})) \big] \, q = 0 \, . \tag{49}$$

Instead of inserting the leading order balance $R^{-1} S q = \lambda q + O(\varepsilon)$ as before, we use the next order eigenvalue problem

$$R^{-1} S q = \lambda q - 2\varepsilon \, \lambda^2 \, R^{-1} q + O(\varepsilon^2) \, , \tag{50}$$

which we have already shown to be true via (46), to find that

$$R^{-1} S q - \lambda q + 2\varepsilon \, \lambda^2 \; R^{-1} q + 4 \, \varepsilon^2 \, \lambda^2 \, (R^{-1} S - \lambda) \, R^{-1} R^{-1} q + O(\varepsilon^3) \, . \tag{51}$$

The leading order balance (37) guarantees that the $O(\varepsilon^2)$ terms cancel in the special case that $R = J$, and we obtain indeed the observed jump in order. It is more difficult to repeat this argument at higher order as subsequent terms in the expansion of M_ε^i come into play, so that explicit tracking of all terms through the succession of transformations is required.

General transformations. When the transformation does not reduce to the identity at order ε^n, the eigenvectors will naturally be changed by the transformation at this order. However, we can easily compute the $O(\varepsilon^n)$ contribution to the eigenvalues using classical perturbation theory. Here we show that in the case when we do not employ a near-near identity transformation the eigenvalue problems of the full dynamics and the reduced slow dynamics are still isospectral up to the expected order.

Writing $q = q_0 + \varepsilon^n q' + \dots$ and $\lambda = \lambda_0 + \varepsilon^n \lambda' + \dots$, we obtain as the eigenvalue problem of the full parent equations at $O(\varepsilon^{n-1})$

$$[S_\varepsilon - \lambda_0 R_\varepsilon] q_0 = 0 \,. \tag{52}$$

As before, the eigenvalues are semisimple and come in complex conjugate pairs, hence, are analytic in ε. We denote a pair of eigenvalues by $\lambda_0^\pm$ and their corresponding eigenvectors by $q_0^\pm$. Note that λ_0 represents the leading order slow dynamics since we assume $\lambda_0 \sim O(1)$. At $O(\varepsilon^n)$, we obtain

$$[S_\varepsilon - \lambda_0 R_\varepsilon] q' = \left[\lambda' \, R_\varepsilon - \sum_{i=1}^m \Big(M_\varepsilon^i \, \lambda^{i+1} + (M_\varepsilon^i)^T \, (-\lambda)^{i+1} \Big) \right] q_0 \,. \tag{53}$$

The next order frequency correction λ' can be calculated upon left multiplication of the adjoint eigenvector. We obtain

$$\lambda' \, (q_0^-)^H \, R_\varepsilon \, q_0^+ = \sum_{i=1}^m (\lambda_0^+)^{i+1} \left((q_0^-)^H \, M_\varepsilon^i \, q_0^+ + (-1)^{i+1} \, (q_0^-)^T (M_\varepsilon^i)^H \, q_0^+ \right) , \tag{54}$$

where superscript H denotes Hermitian conjugation.

The corresponding spectral perturbation of the reduced problem (44), which we recall as

$$\left[R_\varepsilon + \varepsilon^n \, (Z - Z^T) \right] \dot q = \left[S_\varepsilon + \varepsilon^n \, (S Z_0 + Z_0^T S) \right] q \,, \tag{55}$$

where Z is given by (36), yields, analogously,

$$[S_\varepsilon - \lambda_0 R_\varepsilon] q' = \left[\lambda' \, R_\varepsilon - \lambda_0 \sum_{i=1}^m \Big((S R^{-1})^i \, M_i + (-\lambda_0)^i \, M_i^T \Big) + \lambda_0 \, (R - S) Z_0 \right] q_0 \,. \tag{56}$$

Upon left multiplication with an adjoint eigenvector of (52) we obtain again the same expression (54) as for the full problem.

The spectral analysis of this section already provides complete account of the linear case where errors can only be phase errors. The general case is explored next.

4. General potentials

In this section we prove the main convergence result, Theorem 1, which states that for finite times consistently initialized solutions of the full parent model are represented, as in the linear case, up to the expected formal order of accuracy by the reduced, slow dynamics. The proof is structured as follows.

In a first step, we construct a nonconservative limit system by classical means which is formally equivalent to any of the variational limit systems. In particular, we prove the analog of Theorem 1 for solutions to this auxiliary nonvariational limit system. As a consequence, time derivatives of the parent system are systematically bounded—the precise statement is given in Lemma 5 below. We will use these estimates to bound the consistency error of the variational limit system in transformed coordinates. Finally, it is easy to control the propagation of consistency errors for times of order one.

The nonvariational limit theorem, which does not yet involve a change of coordinates as compared to Theorem 6 further below, is the following.

Theorem 2. *For $n \geq 0$ fixed, set*

$$F_{\text{nv}}^n(q) = \sum_{i=0}^n f_i(q)\, \varepsilon^i \tag{57}$$

with coefficient functions f_i recursively defined via

$$f_0(q) = -J\nabla V(q)\,, \tag{58a}$$

$$f_k(q) = -J \sum_{i+j=k-1} \mathrm{D}f_i(q)\, f_j(q)\,. \tag{58b}$$

For fixed initial positions $q_0 \in \mathbb{R}^{2d}$, let $q(t)$ denote a solution to the nonvariational limit system

$$\dot q = F_{\text{nv}}^n(q) \tag{59}$$

with $q(0) = q_0$. Let $q_\varepsilon(t)$ solve the full parent dynamics (1) *consistently initialized via $q_\varepsilon(0) = q_0$ and $\dot q_\varepsilon(0) = F_{\text{nv}}^n(q_0)$. Then for every $T > 0$ there exists $\varepsilon_0 > 0$ and $c = c(q_0, T)$ such that*

$$\sup_{t\in[0,T]} \|q_\varepsilon(t) - q(t)\| \leq c\, \varepsilon^{n+1} \tag{60}$$

for all $0 < \varepsilon \leq \varepsilon_0$.

It is instructive to first consider the leading-order case $n = 0$ separately. The case $n \geq 1$ is then proved for an iteratively constructed vector field which involves boundary terms of successive partial integrations.

Remark 6. The recursion relations (58) for the non-variational vectorfield $F_{\text{nv}}^n(q)$ can be easily derived in the following direct way. As in [18], write (1) as

$$\dot q_\varepsilon = p_\varepsilon\,, \tag{61a}$$

$$\varepsilon\, \dot p_\varepsilon = Jp_\varepsilon - \nabla V(q_\varepsilon)\,, \tag{61b}$$

and introduce a new fast variable $w_n = p_\varepsilon - F_{\text{nv}}^n(q_\varepsilon)$ so that

$$\dot q_\varepsilon = w + F(q)\,, \tag{62a}$$

$$\dot w_n = \left(\frac{1}{\varepsilon}\, J - \mathrm{D}F_{\text{nv}}^n\right) w_n + \frac{1}{\varepsilon}\left(JF_{\text{nv}}^n(q_\varepsilon) - \nabla V(q_\varepsilon)\right) - \mathrm{D}F_{\text{nv}}^n\, F_{\text{nv}}^n\,. \tag{62b}$$

Substituting in the expansion (57), we can iteratively determine the f_i such that the inhomogeneity in (62b) is of order ε^n. Left-multiplying (62b) with w_n yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\|w_n\| \leq \|\mathrm{D}F\|\,\|w_n\| + O(\varepsilon^n) \tag{63}$$

so that, if $w_n = O(\varepsilon^n)$ initially, it will remain so for times of order one. The dynamics is then dominantly slow and can be approximated to $O(\varepsilon^n)$ by $\dot q = F_{\text{nv}}^n(q)$ over times of order one. Notice, however, that this simple argument does not yield the optimal $O(\varepsilon^{n+1})$ which is guaranteed by Theorem 2. To gain this extra order, we have to work in an integral formulation which allows for cancellation of fast oscillations.

Proof of Theorem 2 for $n=0$. Write the full parent dynamics in the form

$$\dot{q}_\varepsilon = F_\varepsilon(q_\varepsilon)\,, \tag{64}$$

where $F_\varepsilon(q_\varepsilon)$ is obtained from (1) by means of the method of integrating factors as

$$F_\varepsilon(q_\varepsilon) = \mathrm{e}^{\frac{Jt}{\varepsilon}}\,\dot{q}_\varepsilon(0) - \frac{1}{\varepsilon}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\nabla V(q_\varepsilon(\tau))\,\mathrm{d}\tau\,. \tag{65}$$

Integration by parts yields

$$w_0(t) = \mathrm{e}^{\frac{Jt}{\varepsilon}}\,w_0(0) + J\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\frac{\mathrm{d}}{\mathrm{d}\tau}\nabla V(q_\varepsilon(\tau))\,d\tau \tag{66}$$

where

$$w_0(t) = F_\varepsilon(q_\varepsilon(t)) + J\nabla V(q_\varepsilon(t)) \equiv F_\varepsilon(q_\varepsilon(t)) - F^0_{\mathrm{nv}}(q_\varepsilon(t))\,, \tag{67}$$

which we recognize as the leading order consistency error of our model reduction. Rearranging terms, we find that

$$\begin{aligned} w_0(t) &= \mathrm{e}^{\frac{Jt}{\varepsilon}}\,w_0(0) + J\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}\nabla V(q_\varepsilon(\tau))\,\dot{q}_\varepsilon(\tau)\,\mathrm{d}\tau \\ &= \mathrm{e}^{\frac{Jt}{\varepsilon}}\,w_0(0) + J\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}\nabla V(q_\varepsilon(\tau))\,w_0(\tau)\,\mathrm{d}\tau \\ &\quad + J\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}\nabla V(q_\varepsilon(\tau))\,F^0_{\mathrm{nv}}(\tau)\,\mathrm{d}\tau\,. \end{aligned} \tag{68}$$

Further integration by parts shows that the last integral on the right actually contributes only at $O(\varepsilon)$. Indeed, for any function $G(q_\varepsilon)$,

$$\begin{aligned} \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,G(q_\varepsilon(\tau))\,\mathrm{d}\tau &= \varepsilon\,JG(q_\varepsilon(t)) - \varepsilon\,\mathrm{e}^{\frac{Jt}{\varepsilon}}\,JG(q_\varepsilon(0)) \\ &\quad - \varepsilon\,J\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}G[q_\varepsilon(\tau)]\,\big(w_0(\tau) - J\nabla V(q_\varepsilon(\tau))\big)\,\mathrm{d}\tau\,. \end{aligned} \tag{69}$$

The integrand containing w_0 can be combined with the first integral in (68). Since $q_\varepsilon(t)$ is *a priori* bounded due to the conservation of energy, all other terms are $O(\varepsilon)$, so that plugging (69) into (68) and taking norms on both sides yields the estimate

$$\|w_0(t)\| \le c_0\,\|w_0(0)\| + c_1\int_0^t \|w_0(\tau)\|\,d\tau + \varepsilon\,(c_2 + c_3\,t) \tag{70}$$

with constants $c_0,\dots,c_3$ depending only on the initial data. Further, the assumption of consistent initialization implies that $w_0(0)=0$, so that, for a fixed interval of time,

$$\|w_0(t)\| \le \varepsilon\,c_4 + c_1\int_0^t \|w_0(\tau)\|\,d\tau \tag{71}$$

which, by the Gronwall lemma, yields

$$\|w_0(t)\| \le \varepsilon\,c_4\,\mathrm{e}^{c_1 t} \le \varepsilon\,c_5\,. \tag{72}$$

The growth rate of the total error splits in the usual way into the consistency error, already estimated above, and a stability part:

$$\frac{\mathrm{d}}{\mathrm{d}t}(q_\varepsilon - q) = F_\varepsilon(q_\varepsilon) - F^0_{\mathrm{nv}}(q) = w_0 + \big(F^0_{\mathrm{nv}}(q_\varepsilon) - F^0_{\mathrm{nv}}(q)\big)\,. \tag{73}$$

Using the smoothness of the potential and the *a priori* bounds on q_ε and q, we have

$$\|F^0_{\mathrm{nv}}(q_\varepsilon) - F^0_{\mathrm{nv}}(q)\| = \|J(\nabla V(q_\varepsilon) - \nabla V(q))\| \le c_6 \, \|q_\varepsilon - q\| \, . \tag{74}$$

We now integrate (73), take norms and use (72) and (74) to obtain

$$\begin{aligned} \|q_\varepsilon(t) - q(t)\| &= \left\| q_\varepsilon(0) - q(0) + \int_0^t w_0(\tau)\, d\tau + \int_0^t \big(F^0_{\mathrm{nv}}(q_\varepsilon) - F^0_{\mathrm{nv}}(q)\big)\, \mathrm{d}\tau \right\| \\ &\le \varepsilon\, c_7 + c_8 \int_0^t \|q_\varepsilon(\tau) - q(\tau)\| \, \mathrm{d}\tau \, . \end{aligned} \tag{75}$$

Estimate (60) for $n = 0$ is now a direct consequence of the Gronwall lemma. □

The proof for general $n \ge 1$ follows the same outline—successive integration by parts where the boundary terms at each order define the vector field F^n_{nv} of the nonvariational slow dynamics, and where time derivatives which appear in the integrand are re-expressed in terms of the consistency error

$$w_n(t) = F_\varepsilon(q_\varepsilon(t)) - F^n_{\mathrm{nv}}(q_\varepsilon(t)) \tag{76}$$

modulo remainders which are free of time derivatives. This, however, poses a chicken-and-egg problem: At the kth step, only $w_0, \dots, w_k$ have been constructed, but we would rather refer to the final w_n. In principle, we could update each w_k in terms of w_{k+1} and a remainder as soon as the next order boundary terms emerge. However, this procedure appears to cause a combinatorial explosion in the number of terms. We thus abandon a strict bottom-up construction, guessing the general form of the nonvariational limit system from the first few iterations of the scheme. We then prove the corresponding expressions for the consistency error via finite induction on k. This is the content of the following lemma.

Lemma 3. *Under the assumptions of Theorem 2, let $n \ge 0$ and $0 \le k \le n$. Then*

$$\begin{aligned} w_k(t) &= \mathrm{e}^{\frac{Jt}{\varepsilon}}\, w_k(0) - \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, \mathrm{D}F^k_{\mathrm{nv}}[q_\varepsilon]\, w_n \, \mathrm{d}\tau \\ &\quad - \sum_{i=k}^n \varepsilon^i \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \sum_{j=0}^k \mathrm{D}f_j\, f_{i-j}\, \mathrm{d}\tau + \varepsilon^{n+1} \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, R^k_{n+1}(q_\varepsilon)\, \mathrm{d}\tau \, , \end{aligned} \tag{77}$$

where $R^k_{n+1}(q_\varepsilon)$ is bounded uniformly in ε and t.

Proof. Let n be fixed. We proceed by (finite) induction on $k \le n$. When $k = 0$, we note that $F^0_{\mathrm{nv}} = f_0 = -J\nabla V(q_\varepsilon)$, so that equation (66) reads

$$\begin{aligned} w_0(t) &= \mathrm{e}^{\frac{Jt}{\varepsilon}}\, w_0(0) - \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, \frac{\mathrm{d}}{\mathrm{d}\tau} f_0 \, \mathrm{d}\tau \\ &= \mathrm{e}^{\frac{Jt}{\varepsilon}}\, w_0(0) - \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, \mathrm{D}f_0 \left(w_n + \sum_{i=0}^n f_i\, \varepsilon^i \right) \mathrm{d}\tau \, , \end{aligned} \tag{78}$$

where we used definition (57) for $F^n_{nv}(q)$. Separating out the different terms in the integrand, we indeed obtain the statement of the lemma for $k = 0$ and $R^0_{n+1} = 0$.

Let us now assume that the statement is already proved for some $k \ge 0$. We single out the term with $i = k$ within the outer sum of (77),

$$\varepsilon^k \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \sum_{j=0}^k \mathrm{D}f_j\, f_{k-j}\, \mathrm{d}\tau = \varepsilon^k \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, J f_{k+1}\, \mathrm{d}\tau \, . \tag{79}$$

Integration by parts yields

$$\begin{aligned}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} J f_{k+1}\,\mathrm{d}\tau + \varepsilon\, f_{k+1}(t) - \varepsilon\,\mathrm{e}^{\frac{Jt}{\varepsilon}}\, f_{k+1}(0) \\ &= \varepsilon \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \frac{\mathrm{d}}{\mathrm{d}\tau} f_{k+1}\,\mathrm{d}\tau \\ &= \varepsilon \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \mathrm{D}f_{k+1}\,\dot q_\varepsilon\,\mathrm{d}\tau \\ &= \varepsilon \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \left(\mathrm{D}f_{k+1}\,w_n + \sum_{i=0}^n \mathrm{D}f_{k+1}\,f_i\,\varepsilon^i\right)\mathrm{d}\tau\,. \end{aligned} \tag{80}$$

We choose to keep explicit only terms up to overall order ε^n so that, combining (79) and (80) and re-indexing the inner sum, we write

$$\begin{aligned}&\varepsilon^k \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \sum_{j=0}^k \mathrm{D}f_j\,f_{k-j}\,\mathrm{d}\tau \\ &\quad= -\varepsilon^{k+1}\, f_{k+1}(t) + \varepsilon^{k+1}\mathrm{e}^{\frac{Jt}{\varepsilon}}\, f_{k+1}(0) \\ &\qquad + \varepsilon^{k+1}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \left(\mathrm{D}f_{k+1}\,w_n + \sum_{i=0}^n \mathrm{D}f_{k+1}\,f_i\,\varepsilon^i\right)\mathrm{d}\tau \\ &\quad= -\varepsilon^{k+1}\, f_{k+1}(t) + \varepsilon^{k+1}\mathrm{e}^{\frac{Jt}{\varepsilon}}\, f_{k+1}(0) + \varepsilon^{k+1}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \mathrm{D}f_{k+1}\,w_n\,\mathrm{d}\tau \\ &\qquad + \sum_{i=k+1}^n \varepsilon^i \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \mathrm{D}f_{k+1}\,f_{i-k-1}\,\mathrm{d}\tau + \varepsilon^{n+1}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \hat R_{n+1}^{k+1}\,\mathrm{d}\tau\,, \end{aligned} \tag{81}$$

where terms appearing at order ε^{n+1} and higher are grouped into the remainder $\hat R_{n+1}^{k+1}$. Plugging this computation into the full expression for w_k, we obtain

$$\begin{aligned}&w_k(t) - \varepsilon^{k+1}\, f_{k+1}(t) = \mathrm{e}^{\frac{Jt}{\varepsilon}}\left(w_k(0) - \varepsilon^{k+1}\, f_{k+1}(0)\right) \\ &\quad - \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} (\mathrm{D}F_{\mathrm{nv}}^k + \varepsilon^{k+1}\,\mathrm{D}f_{k+1})\,w_n\,\mathrm{d}\tau - \sum_{i=k+1}^n \varepsilon^i \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \sum_{j=0}^k \mathrm{D}f_j\,f_{i-j}\,\mathrm{d}\tau \\ &\quad - \sum_{i=k+1}^n \varepsilon^i \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \mathrm{D}f_{k+1}\,f_{i-k-1}\,\mathrm{d}\tau + \varepsilon^{n+1}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} (R_{n+1}^k - \hat R_{n+1}^{k+1})\,\mathrm{d}\tau\,. \end{aligned} \tag{82}$$

Hence, using definitions (76) and (57), and setting $R_{n+1}^{k+1} = R_{n+1}^k - \hat R_{n+1}^{k+1}$, we finally obtain

$$\begin{aligned} w_{k+1}(t) = \mathrm{e}^{\frac{Jt}{\varepsilon}}\, w_{k+1}(0) &- \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \mathrm{D}F_{\mathrm{nv}}^{k+1}\,w_n\,\mathrm{d}\tau \\ &- \sum_{i=k+1}^n \varepsilon^i \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} \sum_{j=0}^{k+1} \mathrm{D}f_j\,f_{i-j}\,\mathrm{d}\tau + \varepsilon^{n+1}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}} R_{n+1}^{k+1}\,\mathrm{d}\tau\,. \end{aligned} \tag{83}$$

This completes the proof. □

Corollary 4. *Under the assumptions of Theorem 2,*

$$w_n(t) = \mathrm{e}^{\frac{Jt}{\varepsilon}}\, w_n(0) - \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, \mathrm{D}F_{\mathrm{nv}}^n\, w_n \, \mathrm{d}\tau + \varepsilon^{n+1} \int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\, R_{n+1}(q_\varepsilon)\, \mathrm{d}\tau \,, \tag{84}$$

where $R_{n+1}(q_\varepsilon)$ is bounded uniformly in ε and t.

Proof. Write out Lemma 3 with $n = k$. Then note that $w_n = w_{n-1} + \varepsilon^n f_n$ and move all terms of order ε^n and higher into the remainder. What is left is an expression of the form (84) for w_{n-1}. Since n is arbitrary, we can shift the index back to n. □

Proof of Theorem 2, general case. The assumption of consistent initialization implies that $w_n(0) = 0$. We now proceed as in the case when $n = 0$, taking the norm of (84) and applying the Gronwall lemma, so that

$$\|w_n(t)\| \leq \varepsilon^{n+1}\, c_1\, \mathrm{e}^{c_2 t} \leq c_3\, \varepsilon^{n+1} \,. \tag{85}$$

The evolution equation for the total error now reads

$$\frac{\mathrm{d}}{\mathrm{d}t}(q_\varepsilon - q) = F_\varepsilon(q_\varepsilon) - F_{\mathrm{nv}}^n(q) = w_n + \bigl(F_{\mathrm{nv}}^n(q_\varepsilon) - F_{\mathrm{nv}}^n(q)\bigr) \,, \tag{86}$$

and we proceed as before:

$$\begin{aligned} \|q_\varepsilon(t) - q(t)\| &= \left\| q_\varepsilon(0) - q(0) + \int_0^t w_n(\tau)\, d\tau + \int_0^t \bigl(F_{\mathrm{nv}}^n(q_\varepsilon) - F_{\mathrm{nv}}^n(q)\bigr)\, \mathrm{d}\tau \right\| \\ &\leq \varepsilon^{n+1}\, c_4 + c_5 \int_0^t \|q_\varepsilon(\tau) - q(\tau)\|\, \mathrm{d}\tau \,. \end{aligned} \tag{87}$$

An application of the Gronwall lemma concludes the proof of Theorem 2. □

Remark 7. The F_{nv}^n for $n \geq 1$ differ from any of the variational slow models derived in Section 2. It is easy to check, however, that the models agree up to higher order terms. When $n = 1$, for example, applying identity (114c) to the first order variational reduced dynamics (18) with $\mu = \frac{1}{2}$ we recover, up to terms of order $O(\varepsilon)$, the nonvariational slow equation

$$J\dot{q} = \nabla V + \varepsilon J \mathrm{D}\nabla V\, J \nabla V \,. \tag{88}$$

Substitutions of this kind into the equations of motion generally do not respect the variational structure; numerical evidence shows that solutions to (88) blow up in a finite time, larger than the time of validity of the estimate of Theorem 2.

Remark 8. It is easy to check that the limit system of Theorem 2 is nonvariational with respect to the canonical symplectic structure: the first order term on the right of (88) is not closed, hence cannot be written as the gradient of a potential. It is more difficult to exclude whether, at least for special V, (88) may be variational with respect to a noncanonical structure. Numerical evidence, however, indicates that this is not generally true.

Remark 9. The nonvariational equation (88) can also be deduced directly from the parent model (1) if we express the inertia term $\varepsilon\, \ddot{q}_\varepsilon$ through the leading order balance equation (15). This approximation is called the *geostrophic wind approximation* in geophysical fluid dynamics. There is no change of coordinates. This is consistent with the fact that (88) can be obtained by formally consistent substitutions from the near-near-identity case $\mu = \frac{1}{2}$.

What we have shown so far is that the parent dynamics with approximately balanced initial data evolves on the slow time scale up to contributions on the order of the initialization error. This observation is captured in following statement, which will be useful in the sequel.

Lemma 5. *Under the assumptions of Theorem 2, the following consistency estimates hold. For any* $i \geq 0$,

$$q_\varepsilon^{(i+1)}(t) = F_{\mathrm{nv}}^{n,i} + O(\varepsilon^{n-i+1}) \tag{89}$$

with $F_{\mathrm{nv}}^{n,i}$ *recursively defined via*

$$F_{\mathrm{nv}}^{n,0}(q) = F_{\mathrm{nv}}^{n}(q)\,, \tag{90a}$$

$$F_{\mathrm{nv}}^{n,i}(q) = \mathrm{D}F_{\mathrm{nv}}^{n-i,i+1}(q)\,F_{\mathrm{nv}}^{n-i,0}(q)\,. \tag{90b}$$

Proof. The proof is achieved by successive differentiation of the identity of Corollary 4. For $i = 0$, the result has already been proved. To proceed further, we recall that $w_n(0) = 0$ and compute

$$\dot{w}_n = -\mathrm{D}F_{\mathrm{nv}}^n\, w_n - \frac{J}{\varepsilon}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}F_{\mathrm{nv}}^n\, w_n\,\mathrm{d}\tau + \varepsilon^{n+1}\,\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,R_{n+1}(q_\varepsilon)\,\mathrm{d}\tau\,. \tag{91}$$

Integrating by parts in the second term on the right, we obtain

$$\dot{w}_n = -\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,\mathrm{D}^2F_{\mathrm{nv}}^n\,(w_n,\dot{w}_n)\,\mathrm{d}\tau + \varepsilon^{n+1}\,\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{e}^{\frac{J(t-\tau)}{\varepsilon}}\,R_{n+1}(q_\varepsilon)\,\mathrm{d}\tau\,. \tag{92}$$

The remainder term is clearly of overall order ε^n, so that, by the Gronwall lemma, $\dot{w}_n = O(\varepsilon^n)$. This proves the case $i = 1$.

This argument can be iterated, noting that each differentiation of the remainder integral lowers its order with respect to ε by one. □

For the variational limit systems constructed in Section 2, the situation is more complicated. If the change of variables is near-near identity up to the order of the model, the proof is still straightforward since it can be shown that, in this special case, the above estimates on the corresponding consistency error w_k differ only in the remainder terms which are functions of q_ε only. The boundedness of q_ε implies that the formal order of each term is maintained.

In general, however, the change of coordinates must be part of the consistency argument. First, it introduces additional time derivatives into the consistency argument which must be accounted for. Second, the invertibility of the transformation is not immediate, but holds true for solutions of the Euler–Lagrange equations. These arguments are detailed in the following.

As before, let $L_\varepsilon(q_\varepsilon, \dot{q}_\varepsilon)$ denote the parent Lagrangian, and define an abstract new coordinate system q implicitly via a family of transformations

$$q_\varepsilon = \Phi_\varepsilon(q, \varepsilon\,\dot{q}, \ldots, \varepsilon^n\, q^{(n)}) \tag{93}$$

which depend smoothly on ε and formally reduce to the identity as $\varepsilon \to 0$. Assume that Φ_ε is chosen such that L_ε has an ε-expansion of the form

$$L_\varepsilon(q_\varepsilon, \dot{q}_\varepsilon) = L(q, \dot{q}) + \varepsilon^{n+1}\, R\Big(q, \frac{\mathrm{d}}{\mathrm{d}t}, \varepsilon\,\frac{\mathrm{d}}{\mathrm{d}t}\Big) \tag{94}$$

where L carries all terms up to order ε^n (we follow the convention of Section 2 where, for notational convenience, this dependence on ε is not explicitly indicated) and R is a finite sum of remainder terms each of which is homogeneous of degree at most $n+2$ in its $\mathrm{d}/\mathrm{d}t$ argument (again, there may be higher order terms in ε which are not explicitly indicated). It is not difficult to check that the transformation constructed in Section 2 is of this kind.

Now let q_ε be given as a solution to the parent system in original coordinates—a solution to the Euler–Lagrange equations corresponding to the left hand side of (94)—with balanced initial data as in Theorem 2. (Note that the balance is, for now, defined via the nonvariational vector field.) We do not know whether the transformation (93) is invertible in general. However, since q_ε solves the left hand Euler–Lagrange equations, it suffices to find a solution to the right hand Euler–Lagrange equations with matching initial data, where the variations transform according to

$$\delta q_\varepsilon = \mathrm{D}_q \Phi_\varepsilon \, \delta q + \varepsilon \, \mathrm{D}_{\dot q} \Phi_\varepsilon \, \delta \dot q + \cdots + \varepsilon^n \, \mathrm{D}_{q^{(n)}} \Phi_\varepsilon \, \delta q^{(n)} \, . \tag{95}$$

Again, we do not know if this tangent map is onto, but it is sufficient to find a solution to the Euler–Lagrange equations computed with respect to arbitrary variations δq which vanish at the temporal end points. This computation yields

$$L_q(q, \dot q) - \frac{\mathrm{d}}{\mathrm{d}t} L_{\dot q}(q, \dot q) = \varepsilon^{n+1} \, S\Big(q, \frac{\mathrm{d}}{\mathrm{d}t}, \varepsilon \, \frac{\mathrm{d}}{\mathrm{d}t}\Big) \tag{96}$$

where S denotes a finite sum of remainder terms each of which is homogeneous of degree at most $n+2$ in $\mathrm{d}/\mathrm{d}t$. Moreover, it can be checked that for the expressions arising in Section 2, the resulting Euler–Lagrange equation (96) is explicit in the highest time derivative so that standard local existence and uniqueness theory applies.

Hence, we only need to match the initial data. This problem is entirely algebraic as the various time derivatives are independent quantities. Since Φ_ε reduces to the identity as $\varepsilon \to 0$, we can always choose ε small enough such that the contraction mapping theorem is applicable and we can find matching initial data for the transformed, nominally higher order problem.

We have thus established that, if q_ε is a solution of the untransformed Euler–Lagrange equation, there exists, locally in time, a q, solution to the full transformed Euler–Lagrange equations, which transforms back to the original problem via (93). It is immediate that q must satisfy the same set of derivative estimates (89) as q_ε does. As a consequence, any term contained in the remainder S of (96) is order one at worst. Since the left hand side of (96) is our slow variational limit problem, this proves that it is consistent to order ε^n as expected. Since the time of local existence depends only on norms of q_ε which are uniformly bounded, the argument automatically holds globally in time.

Finally, the nonvariational balanced initialization and the variational balanced initialization are also consistent to order ε^n, so that the stability estimate is a literal repetition of that in the proof of Theorem 2. We have therefore proved the following, general version of Theorem 1.

Theorem 6. *For $n \geq 0$ and $q_0 \in \mathbb{R}^{2d}$ fixed, let $q(t)$ denote a solution to any of the slow limit systems*

$$\dot q = F^n_{\mathrm{var}}(q) \tag{97}$$

with $q(0) = q_0$, derived via the construction as detailed in Section 2. Let $\Phi_\varepsilon(q) = \Phi_\varepsilon(q, \varepsilon \, F^n_{\mathrm{var}}(q), \dots)$ denote the corresponding transformation to physical coordinates, *and let* $q_\varepsilon(t)$ *solve the full parent dynamics* (1) *consistently initialized via*

$$q_\varepsilon(0) = \Phi_\varepsilon(q_0) \tag{98}$$

and

$$\dot q_\varepsilon(0) = \mathrm{D}\Phi_\varepsilon(q_0)\, F^n_{\mathrm{nv}}(q_0)\,. \tag{99}$$

Then for every $T > 0$ *there exists* $\varepsilon_0 > 0$ *and* $c = c(q_0, T)$ *such that*

$$\sup_{t\in[0,T]} \|q_\varepsilon(t) - \Phi_\varepsilon(q(t))\| \le c\,\varepsilon^{n+1} \tag{100}$$

for all $0 < \varepsilon \le \varepsilon_0$.

Remark 10. In this finite dimensional setting, the rate of convergence does not depend on the chosen transformation (although the constants certainly will); in the setting of Appendix B, in particular, it is independent of the parameters μ and $\nu_{1,2,3}$. In the corresponding setting for the shallow water equation, however, even well-posedness of the reduced system and the spaces in which the system is well posed depend on the choice of transformation [15].

5. Numerical Examples

We finally give a numerical example which illustrates our results. We take the planar quartic potential

$$V(q) = \frac{1}{4}\,\alpha\, q_1^4 + \frac{1}{4}\,\beta\, q_2^4\,, \tag{101}$$

where $q = (q_1, q_2)$. The slow reduced Euler–Lagrange equations and associated transformations are given in Appendix C.

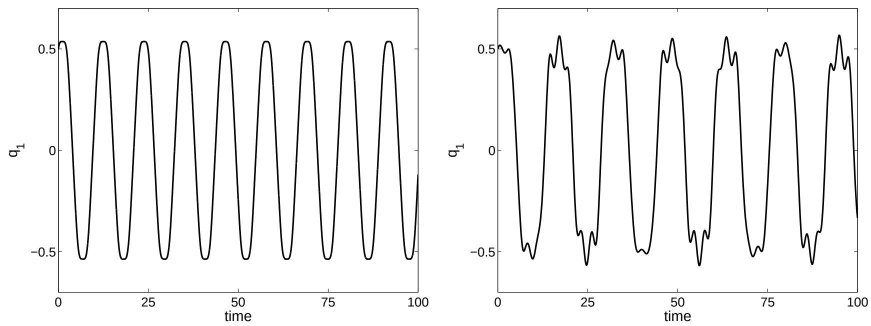

Figure 1. Dynamics of the parent model (1) with quartic nonlinear potential (101) where $\alpha = 3$ and $\beta = 1$. Left: $\varepsilon = 0.01$ and right: $\varepsilon = 0.75$.

Figures 1 and 2 give a visual illustration that consistent initialization indeed suppresses fast oscillations in the parent dynamics. The smaller ε the more accurately lie the initial data on the slow manifold and the less fast modes are initially excited.

In Figure 3, we show the error $\|q_\varepsilon - q\|$ as a function of ε for integration times $T = 1$ if q is the solution of the variational approximations. For the lowest order approximation (15) the slope is 0.98. For the first order slow dynamics (18) we measure a slope of 1.92, and for the second order Euler–Lagrange equations for the

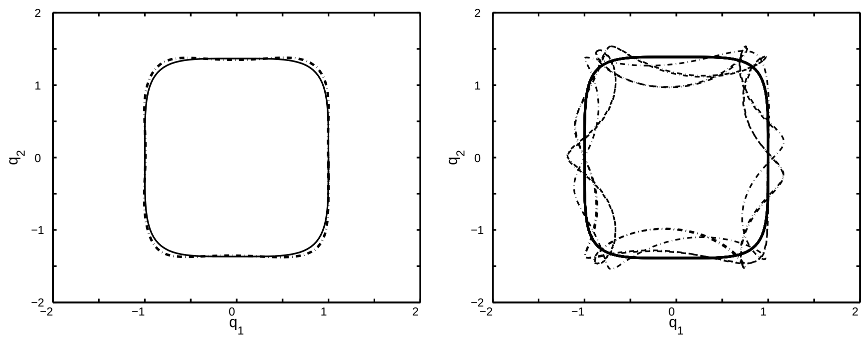

FIGURE 2. Illustration of the slow manifold of (1) as computed by the variational dynamics (18) (solid line), and the full dynamics of (1) (dashed line) with quartic nonlinear potential (101) where $\alpha = 3$ and $\beta = 1$. Left: $\varepsilon = 0.1$ and right: $\varepsilon = 0.2$.

slow dynamics (131) (see Appendix B) we measure a slope of 2.94. These results illustrate Theorem 1. For all cases we set $\mu = 0$ and $\nu_{1,2,3} = 0$. We checked that the rate of convergence does not change when varying the free parameters of the transformation. In Figure 4 we show the actual dynamics for $\varepsilon = 0.1$ for the leading order balance (15) and for the first order approximation (18). The better approximation of the higher order reduced models is evident. It is notable that the first order models give already excellent visual agreement at moderately small values of ε. This is consistent with the observation in geophysical fluid dynamics that the semi-geostrophic equations, which are a first order slow model, yield satisfactory results while subsequent higher order models are typically not considered worth the additional effort.

6. Discussion

We introduced a model problem exhibiting fast and slow dynamics interlinked in a non-obvious way. By means of variational asymptotics based on near-identity changes of variables we derived and analyzed approximate equations for the motion of the slow manifold. We proved that the reduced dynamics converges to the full dynamics on time scales of $O(1)$ with one order higher than its asymptotic order. Moreover for the case of a linear potential we could show that the eigenvalue problem is identical to the eigenvalue problem of the original parent equation at any order of ε.

We also derived a nonvariational limit system by means of a formal expansion of the variational system. We found that the nonvariational system has the same convergence properties as the variational system on times of order one.

While our approach lives on the Lagrangian side, one may also use Hamiltonian normal form theory [17, 2, 4, 5] to deduce the slow dynamics. By writing the parent equations (1) in terms of the fast time variable $\tau = t/\varepsilon$, the model can be viewed as a perturbed canonical Hamilton system

$$\ddot{q}_\varepsilon = J\dot{q}_\varepsilon - \varepsilon \, \nabla V(q_\varepsilon) \, . \tag{102}$$

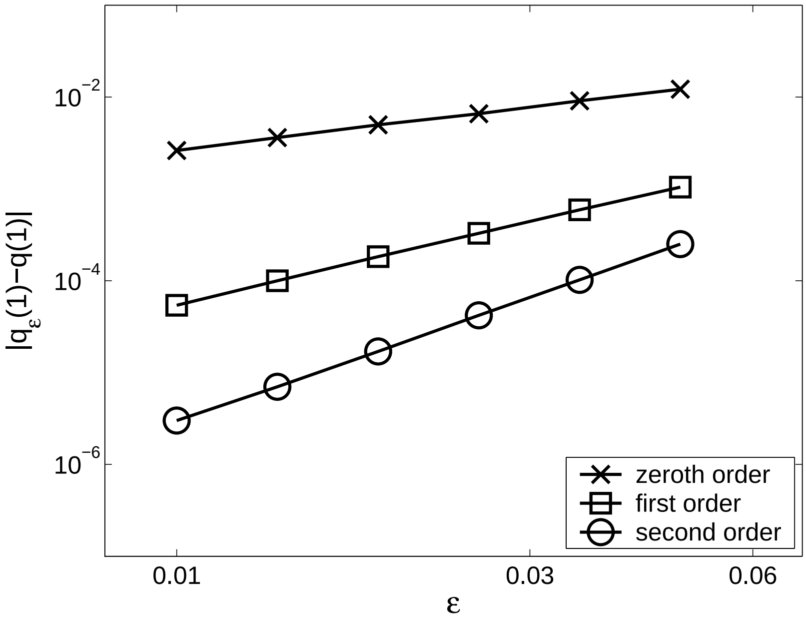

FIGURE 3. Log-log plot of the error $\|q_\varepsilon - q\|$ at $T = 1$ as a function of ε for $\varepsilon \in [0.001, 0.01]$ for the quartic potential with $\alpha = 3$ and $\beta = 1$.

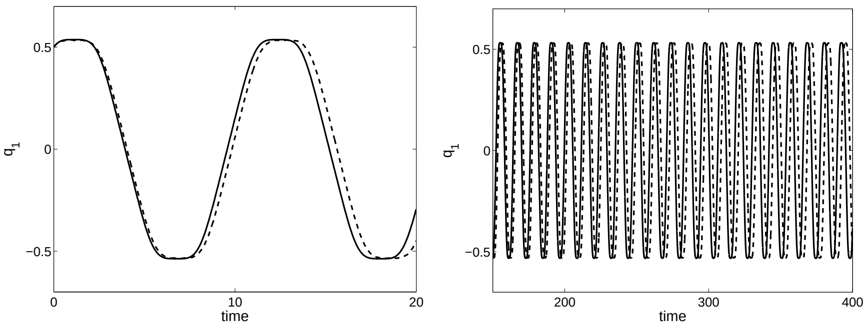

FIGURE 4. Parent dynamics (continuous lines) and variational near-identity reduced model (dashed lines) for the quartic potential with $\alpha = 3$ and $\beta = 1$ with $\varepsilon = 0.05$. Left: zeroth order variational reduced model (15). Right: first order variational reduced model (18).

In this framework, Nekhoroshev type estimates [19] can be used to prove exponential estimates for the tracking of the slow manifold as a phase space object. This has been implemented in a recent paper by Cotter and Reich [5], who 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 strong path-wise estimates for an explicitly computable slow model at fixed low order of an asymptotic series.

Wirosoetisno [29, 30] 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. In [6] a different normal form approach is taken employing center manifold theory.

Our Lagrangian framework has two advantages over the Hamiltonian framework. First, the derivation of higher-order models is more straightforward. Second, Hamiltonian normal form theory as it is commonly applied, for example in [5], will yield a reduced model in canonical variables. In the context of finite dimensional problems, this is just fine and covered as a special case of Theorem 6. For the full shallow water problem as discussed in [15], however, the model with canonical structure—first proposed by Salmon [20]—is ill-posed as a PDE. Thus, the extra degrees of freedom in choosing the transformation, which our approach naturally exposes, are crucial for deriving well-posed reduced equations.

The behavior in the linear case suggests that one may trade a power of ε in the trajectory error for an inverse power of ε in the time scale of validity of the approximation. For the nonlinear example of Section 5, the same type of behavior can be verified numerically. However, a detailed analysis of the first order case, which is given elsewhere [10], shows that this phenomenon holds true only under relatively restrictive assumptions—uniform convexity and symmetry of the potential and, more severely, strong assumptions on the stability of the reduced system ("no chaos") which are generic only in the planar case.

Appendix A. Euler–Lagrange Equations

We compute the Euler–Lagrange Equations for a Lagrangian of the form

$$L = \frac{1}{2}\,\dot q^T M \dot q - V(q) - \frac{1}{2}\,\dot q^T F(q) \tag{103}$$

as follows:

$$\begin{aligned} \delta S &= \delta \int_{t_1}^{t_2} L(q,\dot q)\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \left[\tfrac{1}{2}\,\delta \dot q^T M \dot q + \tfrac{1}{2}\,\dot q^T M \delta \dot q - \mathrm{D}V(q)\,\delta q - \tfrac{1}{2}\,\delta \dot q^T F(q) - \tfrac{1}{2}\,\dot q^T \mathrm{D}F(q)\,\delta q\right] \mathrm{d}t \\ &= \int_{t_1}^{t_2} \delta q^T \left[-\tfrac{1}{2}\,(M+M^T)\ddot q - \nabla V(q) + \tfrac{1}{2}\,(\mathrm{D}F(q) - \mathrm{D}F(q)^T)\,\dot q\right] \mathrm{d}t\,. \end{aligned} \tag{104}$$

Thus, the equations of motion read

$$S\ddot q - R(q)\dot q + \nabla V(q) = 0\,, \tag{105}$$

where

$$S = \frac{M+M^T}{2} \qquad \text{and} \qquad R(q) = \frac{\mathrm{D}F(q) - \mathrm{D}F(q)^T}{2}\,. \tag{106}$$

In Section 3, we also encounter higher order Lagrangians of the form

$$L_\varepsilon = \varepsilon^n\, \dot q_\varepsilon^T\, M_\varepsilon[q_\varepsilon] - \frac{1}{2}\,\dot q_\varepsilon^T F_\varepsilon q_\varepsilon - \frac{1}{2}\,q_\varepsilon^T \Omega_\varepsilon q_\varepsilon \tag{107}$$

with

$$M_\varepsilon[q] = \sum_{i=1}^m M_\varepsilon^i\, \frac{\mathrm{d}^i q}{\mathrm{d}t^i}\,. \tag{108}$$

The corresponding variation of the action integral reads

$$\begin{aligned}\delta S &= \delta \int_{t_1}^{t_2} L_\varepsilon \, \mathrm{d}t \\ &= \int_{t_1}^{t_2} \bigl[\varepsilon^n \, \delta\dot{q}_\varepsilon^T \, M_\varepsilon[q_\varepsilon] + \varepsilon^n \, \dot{q}_\varepsilon^T \, M_\varepsilon[\delta q_\varepsilon] \\ &\qquad - \tfrac{1}{2} \, \delta\dot{q}_\varepsilon^T F_\varepsilon q_\varepsilon - \tfrac{1}{2} \, \dot{q}_\varepsilon^T F_\varepsilon \delta q_\varepsilon - \tfrac{1}{2} \, \delta q_\varepsilon^T \Omega_\varepsilon q_\varepsilon - \tfrac{1}{2} \, q_\varepsilon^T \Omega_\varepsilon \delta q_\varepsilon \bigr] \, \mathrm{d}t \\ &= \int_{t_1}^{t_2} \delta q_\varepsilon^T \, \bigl[-\varepsilon^n \, M_\varepsilon[\dot{q}_\varepsilon] + \varepsilon^n \, M_\varepsilon^*[\dot{q}_\varepsilon] + \tfrac{1}{2} \, (F_\varepsilon - F_\varepsilon^T)\dot{q}_\varepsilon - \tfrac{1}{2} \, (\Omega_\varepsilon + \Omega_\varepsilon^T) q_\varepsilon \bigr] \, \mathrm{d}t \,, \end{aligned} \tag{109}$$

where

$$M_\varepsilon^*[q_\varepsilon] = \sum_{i=1}^m (-1)^i \, (M_\varepsilon^i)^T \, \frac{\mathrm{d}^i q_\varepsilon}{\mathrm{d}t^i} \,. \tag{110}$$

Thus, the Euler–Lagrange equations are

$$\varepsilon^n \sum_{i=1}^m \bigl(M_\varepsilon^i + (-1)^{i+1} \, (M_\varepsilon^i)^T \bigr) \, \frac{\mathrm{d}^{i+1}}{\mathrm{d}t^{i+1}} q_\varepsilon - R_\varepsilon \dot{q}_\varepsilon + S_\varepsilon q_\varepsilon = 0 \,, \tag{111}$$

where

$$R_\varepsilon = \frac{F_\varepsilon - F_\varepsilon^T}{2} \qquad \text{and} \qquad S_\varepsilon = \frac{\Omega_\varepsilon + \Omega_\varepsilon^T}{2} \,. \tag{112}$$

Appendix B. Derivation of the second order model

In the following, we derive explicit equations for general second order models. We assume that the configuration space is two-dimensional. This is not an essential restriction, but permits a number of simplifications due to the following set of vector identities:

$$J u v^T - v (J u)^T = u^T v \, J \,, \tag{113}$$

so that, in particular,

$$(J \nabla f) \otimes \nabla g - \nabla g \otimes (J \nabla f) = \nabla f \cdot \nabla g \, J \,, \tag{114a}$$

$$J \mathrm{D} \nabla h + \mathrm{D} \nabla h J = \Delta h \, J \,, \tag{114b}$$

$$\mathrm{D} \nabla h - J \mathrm{D} \nabla h J = \Delta h \, I \,, \tag{114c}$$

$$\mathrm{D} \nabla h \, J \, \mathrm{D} \nabla h = \det \operatorname{Hess} h \, J \,. \tag{114d}$$

We start out by inserting our choice of the first order transformation, equation (17), into the expanded Lagrangian (14),

$$\begin{aligned} L_\varepsilon &= -V(q) - \tfrac{1}{2} \, \dot{q}^T J q \\ &\quad + \varepsilon \, \bigl[-(\tfrac{1}{2} + \mu) \, \dot{q}^T J \nabla V - \mu \, |\nabla V|^2 \bigr] \\ &\quad + \tfrac{1}{2} \, \varepsilon^2 \, \bigl[-\tfrac{3}{4} \, \dot{q}^T J \ddot{q} + \tfrac{3}{2} \, \mu \, \dot{q}^T \mathrm{D} \nabla V \dot{q} + \tfrac{1}{4} \, \dot{q}^T J \mathrm{D} \nabla V \, J \dot{q} \\ &\qquad + \mu \, \nabla V^T \mathrm{D} \nabla V \, J \dot{q} - \mu^2 \, \nabla V^T \mathrm{D} \nabla V \, \nabla V - \mu^2 \, \dot{q}^T \mathrm{D} \nabla V \, J \nabla V \\ &\qquad + \tfrac{1}{2} \, \mu \, \ddot{q}^T \nabla V - \mathrm{D} V \, q'' - \dot{q}^T J q'' \bigr] \,. \end{aligned} \tag{115}$$

We once more impose a degeneracy conditions—we wish to remove all terms that contain more than one time derivative from the second order contribution of (115). By inspection,

$$q'' = -\tfrac{3}{4}\,\ddot{q} + \tfrac{1}{4}\,\mathrm{D}\nabla V\,J\dot{q} - (\mu - \tfrac{3}{4})\,J\mathrm{D}\nabla V\dot{q} + q_{\text{free}}\,, \tag{116}$$

where

$$q_{\text{free}} = \nu_1\,\nabla V\,\Delta V + \nu_2\,\mathrm{D}\nabla V\,\nabla V + \nu_3\,J\mathrm{D}\nabla V\,J\nabla V\,, \tag{117}$$

is again constructed to include terms of identical homogeneity which allow for the possibility of a near-near-identity transformation. Inserting this expression into (115), we obtain, at second order,

$$\begin{aligned} L_2 &= \dot{q}^T\big[(\tfrac{1}{4} - \mu - \nu_2)\,J\mathrm{D}\nabla V\,\nabla V + (\tfrac{3}{4} + \nu_3 - \mu - \mu^2)\,\mathrm{D}\nabla V\,J\nabla V - \nu_1\,J\nabla V\,\Delta V\big] \\ &\quad - \big[(\mu^2 + \nu_2)\,\nabla V^T\,\mathrm{D}\nabla V\,\nabla V + \nu_3\,\nabla V^T\,J\mathrm{D}\nabla V\,J\nabla V + \nu_1\,|\nabla V|^2\,\Delta V\big]\,. \end{aligned} \tag{118}$$

The overall transformation, including all terms up to second order, then reads

$$\begin{aligned} q_\varepsilon &= q + \varepsilon\,q' + \tfrac{1}{2}\,\varepsilon^2\,q'' \\ &= q + \varepsilon\,(\mu - \tfrac{1}{2})\,\nabla V + \tfrac{1}{2}\,\varepsilon^2\big[(\tfrac{1}{2} + \mu + \nu_1)\,\nabla V\,\Delta V \\ &\qquad + (\nu_2 - 2\mu + \tfrac{1}{4})\,\mathrm{D}\nabla V\,\nabla V + (\mu + \nu_3 - \tfrac{3}{2})\,J\mathrm{D}\nabla V\,J\nabla V\big] + O(\varepsilon^3)\,. \end{aligned} \tag{119}$$

Here we used the first-order slow dynamics (18) to expand $\dot{q}$ in powers of ε and write it as a function of q. The free parameters μ and $\nu_{1,2,3}$ can be used to impose the transformation to be canonical or near-near-identity. For near-near-identity we must require that

$$\mu = \tfrac{1}{2}\,, \quad \nu_1 = -1\,, \quad \nu_2 = \tfrac{3}{4} \quad \text{and} \quad \nu_3 = 1\,. \tag{120}$$

To calculate the second order Euler–Lagrange equations we write the second order Lagrangian as

$$L_2 = -\frac{1}{2}\,\dot{q}^T F_2 - V_2\,, \tag{121}$$

where $F_2 = (\mu + \nu_2 - \tfrac{1}{4})\,F_2^{(1)} + (\mu^2 + \mu - \nu_3 - \tfrac{3}{4})\,F_2^{(2)} + \nu_1\,F_2^{(3)}$ with

$$F_2^{(1)} = J\mathrm{D}\nabla V\,\nabla V\,, \tag{122a}$$

$$F_2^{(2)} = \mathrm{D}\nabla V\,J\nabla V\,, \tag{122b}$$

$$F_2^{(3)} = J\nabla V\,\Delta V\,. \tag{122c}$$

The second order contribution to the potential reads

$$\begin{aligned} V_2 &= \tfrac{1}{2}\,(\mu^2 + \nu_2)\,\nabla V^T\,\mathrm{D}\nabla V\,\nabla V + \tfrac{1}{2}\,\nu_3\,\nabla V^T\,J\mathrm{D}\nabla JV\,\nabla V + \tfrac{1}{2}\,\nu_1\,|\nabla V|^2\,\Delta V \\ &= \tfrac{1}{2}\,(\mu^2 + \nu_2 + \nu_3)\,\nabla V^T\,\mathrm{D}\nabla V\,\nabla V + \tfrac{1}{2}\,(\nu_1 - \nu_3)\,|\nabla V|^2\,\Delta V\,, \end{aligned} \tag{123}$$

so that in the near-near-identity case we have

$$V_2 = \nabla V^T\,\mathrm{D}\nabla V\,\nabla V - |\nabla V|^2\,\Delta V\,. \tag{124}$$

The second order contribution of the Euler–Lagrange equation is

$$R_2(q)\,\dot{q} = \nabla V_2\,, \tag{125}$$

see Appendix A, where

$$2\,R_2 = \mathrm{D}F_2 - \mathrm{D}F_2^T\,. \tag{126}$$

Setting $R_2 = R_2^{(1)} + R_2^{(2)} + R_2^{(3)}$ with

$$R_2^{(1)} \equiv \mathrm{D}F_2^{(1)} - (\mathrm{D}F_2^{(1)})^T$$

$$
\begin{aligned}
&= J\mathrm{D}\nabla\partial_i V\,\partial_i V + \mathrm{D}\nabla\partial_i V\,\partial_i V\,J + J\mathrm{D}\nabla V\,\mathrm{D}\nabla V + \mathrm{D}\nabla V\,\mathrm{D}\nabla V\,J \\
&= \mathrm{D}V\,\nabla\Delta V\,J + |\mathrm{Hess}\,V|^2\,J\,, \qquad (127a)\\
R_2^{(2)} &\equiv \mathrm{D}F_2^{(2)} - (\mathrm{D}F_2^{(2)})^T \\
&= 2\mathrm{D}\nabla V\,J\mathrm{D}\nabla V \\
&= 2\,\det\mathrm{Hess}\,V\,J\,, \qquad (127b)\\
R_2^{(3)} &\equiv \mathrm{D}F_2^{(3)} - (\mathrm{D}F_2^{(3)})^T \\
&= J\mathrm{D}\nabla V\,\Delta V + \mathrm{D}\nabla V\,J\,\Delta V + J\nabla V\otimes\nabla\Delta V - \nabla\Delta V\otimes J\nabla V \\
&= (\Delta V)^2\,J + \mathrm{D}V\,\nabla\Delta V\,J\,, \qquad (127c)
\end{aligned}
$$

where we used vector identities from (114), we obtain

$$
\begin{aligned}
2\,R_2 = \bigl[&(\mu+\nu_1+\nu_2-\tfrac14)\,\mathrm{D}V\,\nabla\Delta V + (\mu+\nu_2-\tfrac14)\,|\mathrm{Hess}\,V|^2 \\
&+ 2\,(\mu^2+\mu-\nu_3-\tfrac34)\,\det\mathrm{Hess}\,V + \nu_1\,(\Delta V)^2\bigr]\,J\,.
\end{aligned} \qquad (128)
$$

In particular, in the near-near-identity case, when

$$
\mu=\tfrac12\,,\quad \nu_1=-1\,,\quad \nu_2=\tfrac34 \quad\text{and}\quad \nu_3=1\,, \qquad (129)
$$

we obtain

$$
R_2 = \bigl[\tfrac12\,|\mathrm{Hess}\,V|^2 - \det\mathrm{Hess}\,V - \tfrac12\,(\Delta V)^2\bigr]\,J\,. \qquad (130)
$$

Combining terms at the various orders, we can write out the full second order Euler–Lagrange equations as

$$
\bigl[1+\varepsilon\,(\tfrac12+\mu)\,\Delta V - \varepsilon^2\,R_2 J\bigr]\,J\dot q = \nabla V + 2\varepsilon\mu\,\mathrm{D}\nabla V\,\nabla V + \varepsilon^2\,\nabla V_2\,. \qquad (131)
$$

Appendix C. Example: Quartic potential

Our numerical example uses the quartic potential

$$
V(q) = \frac{\alpha}{4}\,q_1^4 + \frac{\beta}{4}\,q_2^4\,. \qquad (132)
$$

We compute

$$
\nabla V = \begin{pmatrix}\alpha\,q_1^3\\ \beta\,q_2^3\end{pmatrix},\quad \mathrm{Hess}\,V = \begin{pmatrix}3\alpha\,q_1^2 & 0\\ 0 & 3\beta\,q_2^2\end{pmatrix},\quad \Delta V = 3\,(\alpha\,q_1^2+\beta\,q_2^2)\,, \qquad (133)
$$

so that the first order slow Euler–Lagrange equation reads

$$
\bigl[1+3\varepsilon\,(\tfrac12+\mu)\,(\alpha\,q_1^2+\beta\,q_2^2)\bigr]\,J\dot q = \begin{pmatrix}\alpha\,q_1^3\\ \beta\,q_2^3\end{pmatrix} + 6\varepsilon\mu\begin{pmatrix}\alpha^2 q_1^5\\ \beta^2 q_2^5\end{pmatrix}, \qquad (134)
$$

or

$$
\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}q_1\\ q_2\end{pmatrix} = \frac{1}{1+3\,\varepsilon\,(\frac12+\mu)\,(\alpha\,q_1^2+\beta\,q_2^2)}\begin{pmatrix}\beta\,q_2^3+6\varepsilon\mu\beta^2\,q_2^5\\ -\alpha\,q_1^3-6\varepsilon\mu\alpha^2\,q_1^5\end{pmatrix}. \qquad (135)
$$

To obtain the next order, we compute

$$
\begin{gathered}
\mathrm{D}V\,\nabla\Delta V = 6\,(\alpha^2 q_1^4+\beta^2 q_2^4)\,,\quad |\mathrm{Hess}\,V|^2 = 9\,(\alpha^2 q_1^4+\beta^2 q_2^4)\,,\\
\det\mathrm{Hess}\,V = 9\,\alpha\beta\,q_1^2\,q_2^2\,,\quad (\Delta V)^2 = 9\,(\alpha^2\,q_1^4+\beta^2\,q_2^4) + 18\,\alpha\beta\,q_1^2\,q_2^2\,,
\end{gathered} \qquad (136)
$$

so that

$$
R_2 = \bigl[\tfrac{15}{2}\,(\mu+\nu_1+\nu_2-\tfrac14)\,(\alpha^2 q_1^4+\beta^2 q_2^4) + 9\,(\mu^2+\mu+\nu_1-\nu_3-\tfrac34)\,\alpha\beta\,q_1^2\,q_2^2\bigr]\,J\,. \qquad (137)
$$

Further,

$$
\nabla V^T\,\mathrm{D}\nabla V\,\nabla V = 3\,(\alpha^3\,q_1^8+\beta^3\,q_2^8)\,, \qquad (138)
$$

$$|\nabla V|^2 \, \Delta V = 3 \, (\alpha^3 \, q_1^8 + \beta^3 \, q_2^8 + \alpha^2 \, \beta \, q_1^6 \, q_2^2 + \alpha \, \beta^2 \, q_1^2 \, q_2^6) \,, \tag{139}$$

so that

$$V_2 = \tfrac{3}{2} \, (\mu^2 + \nu_1 + \nu_2) \, (\alpha^3 \, q_1^8 + \beta^3 \, q_2^8) + \tfrac{3}{2} \, (\nu_1 - \nu_3) \alpha \beta q_1^2 q_2^2 \, (\alpha \, q_1^4 + \beta \, q_2^4) \,, \tag{140}$$

and

$$\nabla V_2 = 3 \begin{pmatrix} 4 \, (\mu^2 + \nu_1 + \nu_2) \, \alpha^3 \, q_1^7 + (\nu_1 - \nu_3) \alpha \beta \, (3 \, \alpha \, q_1^5 \, q_2^2 + \beta \, q_1 \, q_2^6) \\ 4 \, (\mu^2 + \nu_1 + \nu_2) \, \beta^3 \, q_2^7 + (\nu_1 - \nu_3) \alpha \beta \, (\alpha \, q_1^6 \, q_2 + 3 \, \beta \, q_1^2 \, q_2^5) \end{pmatrix} . \tag{141}$$

which can be inserted into (131). The associated second order transformation reads

$$\begin{aligned} q_\varepsilon &\equiv T(q) \\ &= q + \varepsilon \, (\mu - \tfrac{1}{2}) \begin{pmatrix} \alpha \, q_1^3 \\ \beta \, q_2^3 \end{pmatrix} \\ &\quad + \tfrac{3}{2} \, \varepsilon^2 \biggl[(\tfrac{3}{4} - \mu + \nu_1 + \nu_2) \begin{pmatrix} \alpha^2 \, q_1^5 \\ \beta^2 \, q_2^5 \end{pmatrix} + \alpha \, \beta \, (2 + \nu_1 - \nu_3) \begin{pmatrix} q_1^3 \, q_2^2 \\ q_1^2 \, q_2^3 \end{pmatrix} \biggr] . \end{aligned} \tag{142}$$

Moreover,

$$\begin{aligned} \dot q_\varepsilon &= \mathrm{D}T(q) \, \dot q \\ &= \biggl[I + 3 \, \varepsilon \, (\mu - \tfrac{1}{2}) \begin{pmatrix} \alpha \, q_1^2 & 0 \\ 0 & \beta \, q_2^2 \end{pmatrix} + \tfrac{15}{2} \, \varepsilon^2 \, (\tfrac{3}{4} - \mu + \nu_1 + \nu_2) \begin{pmatrix} \alpha^2 \, q_1^4 & 0 \\ 0 & \beta^2 \, q_2^4 \end{pmatrix} \\ &\quad + \tfrac{3}{2} \, \varepsilon^2 \, \alpha \beta \, (2 + \nu_1 - \nu_3) \begin{pmatrix} 3 \, q_1^2 \, q_2^2 & 2 \, q_1^3 \, q_2 \\ 2 \, q_1 \, q_2^3 & 3 \, q_1^2 \, q_2^2 \end{pmatrix} \biggr] \, \dot q \, . \end{aligned} \tag{143}$$

This transformation is explicitly needed to consistently initialize the numerical reference computation of the parent model and to diagnose the trajectory error.

Acknowledgments

We thank Sebastian Reich and Colin Cotter for stimulating discussions, and gratefully acknowledge hospitality and support of the *Mathematisches Forschungsinstitut Oberwolfach* where this work was performed under the Research-in-Pairs program. GAG acknowledges support by the Australian Research Council under grant number DP0452147, MO the kind hospitality of the Courant Institute of Mathematical Sciences during the finalization of this manuscript.

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] G. Benettin, L. Galgani and A. Giorgilli, Realisation of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory – part II, *Commun. Math. Phys.* **113** (1987), 87–103.

[3] F. Bornemann, Homogenization in time of singularly perturbed conservative mechanical systems. Lecture Notes in Mathematics **1687**, Springer-Verlag, 1998.

[4] C. Cotter, Model Reduction for Shallow Water Dynamics: Balance, Adiabatic Invariance and Subgrid Modelling. PhD Thesis, Imperial College, London, 2004

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

[6] S.M. Cox and A.J. Roberts, Initialisation and the quasigeostrophic slow manifold, Technical Report, http://arXiv.org/abs/nlin.CD/0303011.

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

[8] R. Ford, S.J.A. Malham, and M. Oliver, A new model for shallow water in the low Rossby-number limit, *J. Fluid Mech.* **450** (2002), 287–296.
[9] D. Givon, R. Kupferman, and A. Stuart, Extracting macroscopic dynamics: model problems and algorithms, *Nonlinearity* **17** (2004), R55–R127.
[10] G.A. Gottwald, M. Oliver, and N. Tecu, Long-time accuracy for approximate slow manifolds in a finite dimensional model of balance, *J. Nonlinear Sci.*, to appear.
[11] E.J. Hinch, Perturbation Methods, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1991.
[12] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, 1980.
[13] 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.
[14] A. Majda, Introduction to Pdes and Waves for the Atmosphere and Ocean, American Mathematical Society, Providence, RI, 2003.
[15] M. Oliver, Variational asymptotics for rotating shallow water near geostrophy: A transformational approach, *J. Fluid. Mech.* **551** (2006), 197–234.
[16] J. P. Pedlosky, Geophysical Fluid Dynamics, Springer-Verlag, New York, 1987.
[17] A. Neishtadt, Estimates in the Kolmogorov theorem on the conservation of conditionally periodic motions, *J. Appl. Math. Mech.* **45** (1981), 1016–1025.
[18] A.I. Neishtadt, *On delayed stability loss under dynamical bifurcations I*, Differ. Equ. **23** (1987), 1385–1390.
[19] N. Nekhoroshev, An exponential estimate of the time of stability of a nearly-integrable Hamiltonian system, *Russ. Math. Surv.* **32** (1977), 1–65.
[20] R. Salmon, New equations for nearly geostrophic flow, *J. Fluid. Mech.* **153** (1985), 461–477.
[21] R. Salmon, Lectures on geophysical fluid dynamics, Oxford University Press, New York, 1998.
[22] J.A. Sanders and F. Verhulst, "Averaging methods in nonlinear dynamical systems," Applied Mathematical Sciences **59**, Springer-Verlag, New York, 1985.
[23] 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-Verlag, 1980, pp. 425–445.
[24] A. N. Tikhonov, A. B. Vasil'eva and A. G. Sveshnikov, Differential equations, Springer-Verlag, New York, 1980.
[25] E. Tomboulis, *Canonical quantization of nonlinear waves*, Phys. Rev. D **12** (1975), 1678–1683.
[26] F. Verhulst, Nonlinear Differential Equations and Dynamical Systems, Springer-Verlag, Berlin, 2000.
[27] F. Verhulst, Methods and Applications of Singular Petrurbations, Texts in Applied Mathematics **50**, Springer-Verlag, Berlin, 2005.
[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.
[29] D. Wirosoetisno, Averaging, slaving and balance dynamics in a simple atmospheric model, *Physica D* **141** (2000), 37–53.
[30] D. Wirosoetisno, Exponentially accurate balance dynamics, *Adv. Diff. Eq.* **9** (2004), 177–196.

(G. Gottwald) School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia

(M. Oliver) School of Engineering and Science, International University Bremen, 28759 Bremen, Germany