Reduced equations for models of laminated materials in thin domains. I.

R. E. Lee DeVille*
C. Eugene Wayne†

November 30, 2001

Abstract

In this paper, we consider a scalar wave equation on a thin, laminated, three-dimensional plate. We show that if the plate is sufficiently thin, then there is a hierarchy of two-dimensional equations whose dynamics model the dynamics of the full plate, each of which successively lengthens the time interval over which the approximation holds.

1 Introduction

Many problems involving the motion of elastic structures occur in domains where one or more of the dimensions of the body is significantly smaller than the others. One can think, for instance, of the vibrations of a long metal beam or rod, or the metal plates that cover the outside of an airplane or ship. In such situations, engineers have usually replaced the true equations of elasticity by simpler model equations – for instance, the Bernoulli or Timoshenko models of beams or the Kirchoff or Reisner-Mindlin models of plates. The reasons for such approximations were that the simplified models were more amenable to exact or asymptotic solution methods. Even with the advent of high speed computers numerical solution of the full three dimensional elasticity problem is very time-consuming and these model equations are still extensively used because of the savings of time that they represent. In the last decade, through the work of numerous authors (see [9],[10], [11] [2], [1], [3], [25], [18] for a small sampling of this literature), a number of results have been established that justify using these model equations as approximations in problems of *static* elastic behavior – that is to say, when the beam or plate is at rest. Far less is known about the validity of these model equations at describing the dynamic behavior of these

*Department of Mathematical Sciences, Rensellaer Polytechnic Institute, Troy NY 12180, USA. Email: devilr@rpi.edu

†Department of Mathematics and Statistics, Boston University, 111 Cummington St., Boston MA 02215, USA. Email: cew@math.bu.edu

elastic media. We are aware only of the work of Raoult [22], [23], Durdević [14], and Xiao [26].

In this paper, we begin a study of the dynamics of thin elastic media and the approximation of their motion by "reduced equations" – that is, by equations on lower dimensional spatial domains. Our approach differs from those mentioned in the previous paragraph in that we are interested in developing a hierarchical family of reduced equations which allow us to approximate the dynamics over longer and longer time intervals – in particular for a time interval longer than any fixed inverse power of the thickness of the domain. In this respect, our work is similar in spirit to that of Schwab and Babuška [3] on the analogous static problem (although our methods are completely different) who point out that in real applications one is always faced with a domain of some fixed thickness and one needs to be able to construct an approximation to the desired accuracy for that given thickness.

We take as our starting point a model for thin laminated materials studied in the static case by Babuška and Schwab, [3]. We prove that if one fixes a tolerance for the error, then for sufficiently thin beams or plates, one can construct a reduced equation which will approximate the dynamics of the three dimensional equation to within this error for an arbitrarily long time. As the length of time over which one wishes the approximation to be valid increases, the reduced equation becomes more complicated, but our method provides an algorithm for constructing refinements to the reduced equation inductively, and in a companion paper, [13], one of us shows that the coefficients in the reduced equations can be computed explicitly. We hope that using the experience gained in studying this problem we can use the same methods to derive similar reduced equations for the full equations of linear elasticity, just as in the static case the results of [3] were later extended to the static elasticity problem in [25].

The viewpoint we adopt in this paper is to think of the original three dimensional partial differential equation as an infinite dimensional Hamiltonian dynamical system. We identify in the (infinite dimensional) phase space of this dynamical system an (infinite dimensional) submanifold which is left invariant by the reduced PDE. We then show, by making appropriate canonical changes of variables in the Hamiltonian for our original PDE, that this invariant manifold is left approximately invariant by the flow of the full three-dimensional PDE. In fact, we make a sequence of canonical changes of variables which systematically: (1) change the submanifold slightly, leading to more and more refined approximating equations, and (2) show that, given initial conditions of our original PDE in some tubular neighborhood of this submanifold, the solution with these initial conditions is well approximated by the flow on the manifold (and hence by the reduced equation) for longer and longer time intervals.

The procedure is reminiscent of and inspired by the Nekhoroshev theory [19] of classical mechanics. That theory shows that for a nearly integrable Hamiltonian system, any solution behaves in a nearly quasi-periodic fashion for a very long time. More geometrically, one may interpret it as saying that initial conditions which lie close to an invariant torus in the phase space can remain close to that torus for a very long time, and their motion is well approximated by the motion on the torus. Our situation is similar in that the submanifold corresponding to the reduced equations corresponds to the invariant tori. Just as in the classical mechanics case, the flow near the invariant torus both stays near the torus and is well-approximated by the flow on the torus, we will show that the flow near the invariant submanifold both stays near the submanifold and is well-approximated by the flow on the invariant submanifold. Note that in contrast to the classical Nekhoroshev theory, and also in contrast to the extensions of the Nekhoroshev theory for nearly integrable PDEs ([4], [20], [21], [7]), the manifold to which we remain close is itself infinite dimensional. In order to apply the classical Nekhoroshev theory, it is necessary that there be some small parameter in the problem. In our case, we take advantage of the fact that vibrations of our laminar material in its "thin" direction have much higher frequency than those in its "long" directions. The inverse of the ratio of these two frequencies (which is, in fact, a positive power of the thickness of the three-dimensional domain) is effectively the small parameter in our expansion. Nekhoroshev-type theories in which one exploits a large separation in frequencies have also been developed for models of one-dimensional gases ([6]) and for models of diatomic molecules ([5]).

One complication of this approach is that the approximation equations which arise most naturally in this procedure are ill-posed. We circumvent that problem here by showing that we can still approximate solutions of our original partial differential equation if we choose initial conditions for the reduced equation in a subspace in which it is well-posed. In a companion paper [13], a method of deriving well-posed reduced equations is developed.

We conclude this introduction with a brief survey of the organization of the remainder of the paper. In the next subsection we define the problem we study and state our results. Then in Subsection 1.2 we outline our proof in some detail. Section 2 is concerned with the construction and estimation of the canonical transformations described in the description of our approach just above, while Section 3 studies the form of the Hamiltonian resulting from these canonical transformations. Finally, in Section 4, we prove that the transformed Hamiltonians give rise to the reduced equations we seek.

1.1 Notation and problem formulation

Let us consider $\omega \subset \mathbb{R}^2$ a bounded domain with $\mathcal{C}^1$ smooth boundary γ. We will use the coordinates x_1, x_2, y for $\mathbb{R}^3$. For simplicity, we will sometimes write $x = (x_1, x_2)$. Given a two dimensional domain ω and a positive thickness parameter d we define the three-dimensional domain

$$\Omega = \omega \times (0, \pi d)$$

with lateral boundary

$$\Gamma = \gamma \times (0, \pi d)$$

and the faces

$$R_- = \{(x_1, x_2, y) | (x_1, x_2) \in \omega, y = 0\},$$
$$R_+ = \{(x_1, x_2, y) | (x_1, x_2) \in \omega, y = \pi d\}.$$

In Ω we consider the hyperbolic problem with prescribed forcing terms on the faces, i.e.

$$\begin{aligned} u_{tt} &= Lu \text{ in } \Omega, \\ u &= 0 \text{ on } \Gamma, \\ \partial_n u &= f^{\pm} \text{ on } R_{\pm}, \end{aligned} \tag{1}$$

where the operator L is given by

$$Lu = \frac{\partial}{\partial y}\left(a(y/d)\frac{\partial u}{\partial y}\right) + b(y/d)\nabla_x \cdot (C(x)\nabla_x u),$$

with $\nabla_x = (\partial/(\partial x_1), \partial/(\partial x_2))^T$, and we choose an initial condition $(u^0, u_t^0) \in H^1 \times L^2$.

We further assume that $a(\cdot), b(\cdot) \in \mathcal{C}^1(0, \pi)$ are independent of d and that there are numbers $\mathbf{a}$, $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{B}$ so that

$$0 < \mathbf{a} \le a(z) \le \mathbf{A}, \quad 0 < \mathbf{b} \le b(z) \le \mathbf{B},$$

and that the matrix $C(x)$ is symmetric and uniformly positive-definite, i.e. that there are constants $0 < \mathbf{c} \le \mathbf{C} < \infty$ so that

$$\mathbf{c}\,|\xi|^2 \le \xi^T C(x)\xi \le \mathbf{C}\,|\xi|^2,$$

for all $\xi \in \mathbb{R}^2$, $x \in \omega$, and that C has $\mathcal{C}^\infty$ coefficients. The static version of (1) (i.e. the case in which u is independent of time) was introduced in [3] as a model for laminated materials.

The goal of this paper is to show that we can approximate the solutions to (1) by the solutions to a 2-dimensional PDE in the variables (x_1, x_2) (or by a system of such equations.) If we can find such a PDE which approximates the full 3-dimensional problem well, we will refer to it as a “reduced equation”.

We begin by studying an apparently simpler problem in which the inhomogeneous boundary conditions on the top and bottom face of the body are eliminated. After showing how we can approximate the motion of this problem by a reduced problem, we will show that the same is true of (1) by combining our results with those of Babuška and Schwab ([3]).

Thus, we first consider:

$$\begin{aligned} u_{tt} &= L \text{ in } \Omega, \\ u &= 0 \text{ on } \Gamma, \\ \partial_n u &= 0 \text{ on } R_{\pm}, \end{aligned} \tag{2}$$

with the initial condition $(u, u_t)|_{t=0} = (u^0, u_t^0)$. Denote

$$L_x u = \hat{b}(0)\nabla_x(C(x)\nabla_x u),$$
$$L_y u = (a(y/d)u_y)_y,$$

where $\hat{b}(0)$ is the average value of b, i.e.

$$\hat{b}(0) = \frac{1}{\pi d}\int_0^{\pi d} b(y/d)\, dy.$$

Note that L_x is a differential operator which depends only on x, and L_y is a differential operator which depends only on y. Our reduced equations will be equations defined on ω and thus we define the projection operator Π, via

$$(\Pi u)(x) = \frac{1}{\pi d}\int_0^{\pi d} u(x, y)dy\ .$$

For technical reasons, we will later need a new function space. Consider the eigenfunctions of L_x as defined in (19), denoted by ϕ_k. Fix an $\alpha \in (-1/2, 0)$. Then, given a Sobolev space $H^s(\omega)$, we know that the ϕ_k will form an orthogonal basis for H^s (orthonormal if we scale them properly). We consider the finite dimensional subspace of H^s spanned by $\{\phi_k\}_{k<d^\alpha}$, denoted H^s_α. It is proven in Lemma 6 that $H^1_\alpha \times L^2_\alpha$ is preserved under the flow of either (1) or (7), i.e. given initial conditions $(w^0, w_t^0) \in H^1_\alpha \times L^2_\alpha$, then $(u(t), u_t(t)) \in H^1_\alpha \times L^2_\alpha$ for all t.

If we choose $u^0 \in H^s$, where H^s refers to functions weakly differentiable in both the x and y directions, then we will abuse notation and say that $u^{\mathbf{r}0} = \Pi(u^0)$ as defined above will be in H^s, where H^s here refers to functions differentiable in the x directions. For clarity we could refer to these two different Hilbert spaces as $H^s(\Omega)$ and $H^s(\omega)$, respectively. If we choose $(u^0, u_t^0) \in H^2 \times H^1$, then we have $(u^{\mathbf{r}0}, u_t^{\mathbf{r}0}) \in H^2 \times H^1$. We will denote the projection of $u^{\mathbf{r}0}$ (resp. $u_t^{\mathbf{r}0}$) into the space $H^2_\alpha \times H^1_\alpha$ as $\overline{u^{\mathbf{r}0}}$ (resp. $\overline{u_t^{\mathbf{r}0}}$). Speaking colloquially: to get $\overline{u^{\mathbf{r}0}}$, we just expand $u^{\mathbf{r}0}$ in a Fourier series in ϕ_k and "chop off" all of the terms in the expansion with $k \geq d^\alpha$.

If we have a solution u smooth enough that all the derivatives in (2) exist and are continuous, and (2) is satisfied for all $0 \leq t \leq T$, then we say that u is a *strong solution* or *classical* solution of (2). We will typically study weak solutions of these equations, however. Define the time-dependent bilinear form

$$B[u, v; t] = \int_\Omega a(y/d)\frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + b(y/d)\left(\nabla_x u\right)^T C(x)\nabla_x v\, dV. \tag{3}$$

We define the Hilbert space

$$\dot{H}(\Omega) = \{u \in H^2(\Omega) \mid \text{Trace}(u)\text{=}0 \text{ on } \Gamma,\ \text{Trace}(\partial_n u)\text{=}0 \text{ on } y = 0,\ y = \pi d\}.$$

Then (following [15]) we say that u is a *weak solution* of (2) if $u \in L^2([0, T], \dot{H}(\Omega))$, $\dot{u} \in L^2([0, T], H^1(\Omega))$, and $\ddot{u} \in L^2([0, T], L^2(\Omega))$, and u satisfies

$$\langle \ddot{u}, v\rangle + B[u, v; t] = 0 \tag{4}$$

for all $v \in \dot{H}(\Omega)$ and for almost every time $0 \leq t \leq T$, and

$$u(0) = u^0, \qquad \dot{u}(0) = v^0.$$

Remark 1. *It is a standard exercise to show that with the assumptions that we have made on the coefficient functions in the equation, (2) always has a unique weak solution. Furthermore standard results (again see [15]) also show that if u is a strong solution of (2), then it is also a weak solution, and any weak solution smooth enough to be defined as a strong solution is also a strong solution.*

Remark 2. *It is more standard to impose Neumann boundary conditions by requiring that the solution take values in $H^1(\Omega)$, and in addition to satisfying the equation (4) one requires that the boundary terms*

$$\int_{R_\pm} a(y/d) \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} dx = 0 \ .$$

Since we require that the solution lie in H^2 for some later estimates we have chosen to enforce the boundary conditions in the definition of the function space $\dot{H}(\Omega)$.

We can also cast the reduced equations (which appear below in (5)) in a weak form. We will discuss this further after the introduction of the reduced equations.

As described earlier in the introduction, we actually derive a series of approximating equations for (2) which are accurate approximations over longer and longer time intervals. More precisely we have:

Theorem 1. *Fix $n > 0$, $C_0 > 0$, $\epsilon > 0$, and $\eta > 0$. Then there exist constants $C^{(2)}, C^{(3)}, \ldots,$ (which are computable in terms of the coefficient functions a, b, and C) and a function $N_\eta \colon \mathbf{N} \to \mathbf{N}$ with the following properties:*

- *$N_\eta(1) \geq 3 - \eta$, and $\lim_{n\to\infty} \frac{N_\eta(n)}{n} > 2 - \eta$.*

Furthermore, there exists $d_0 = d_0(C_0, \epsilon, \eta)$, and $C_T = C_T(C_0, \epsilon, \eta)$ such that for all $d < d_0$, one has:

- *If $u(x, y, t)$ is the solution of (2) with initial conditions $(u^0, u_t^0) \in H^2 \times H^1$ satisfying $\|u^0\|_{H^2} + \|u_t^0\|_{H^1} \leq C_0$, define $(\overline{u^{0}}^{\mathbf{r}}, \overline{u_t^{0}}^{\mathbf{r}}) \in H_\alpha^2 \times H_\alpha^1$ as the projections into H_α^2 and H_α^1 of $\Pi(u^0)$ and $\Pi(u_t^0)$. Let $\overline{u_n}^{\mathbf{r}}$ be the solution with these initial conditions of the equation*

$$\partial_t^2 \overline{u_n}^{\mathbf{r}} = D_n \overline{u_n}^{\mathbf{r}} \tag{5}$$

  where D_n is the differential operator $D_1 = L_x$, and

$$D_n = L_x + \sum_{q=2}^{n} C^{(q)} d^{2(q-1)} L_x^q,$$

  if $n > 1$, then

$$\|u(x, y, t) - \overline{u_n^{\mathbf{r}}}(x, t) \mathbf{1}(y)\|_{H^1 \times L^2} \leq \epsilon$$

  for all $t \leq C_T d^{-N_\eta(n)}$.

Remark 3. *Roughly speaking one can think of this theorem as follows. One fixes a tolerance for error (ϵ) and the order ($2n$) of the reduced equation. The theorem then guarantees that for a sufficiently thin plate, there exists a reduced equation which will accurately approximate the solution of the three dimensional equation (up to the chosen error) for a time scale of $\mathcal{O}(d^{-N_\eta(n)})$. Alternatively, if one wishes to approximate the solutions of the three dimensional problem for a predetermined length of time, one can choose n so that $C_T d^{-N_\eta(n)}$ is greater than this time, and the theorem guarantees that one can find a reduced equation that provides a good approximation for at least this long. Not surprisingly, the longer one wishes to have accurate approximation, the more complicated (i.e. the larger n) the reduced equation must be. However, in a companion paper [13], explicit expressions for the coefficients $C^{(j)}$ in the reduced equation will be derived for all j so that in principle, the reduced equations can be computed to any order.*

Remark 4. *The boundary conditions in equation (5) are those "inherited" from the definition of the operator L_x – i.e we require*

$$L_x^i \overline{u_n^{\mathbf{r}}}\big|_{\partial\omega} = 0, \text{ for } i < n.$$

Furthermore, as mentioned above, when we speak of a solution of (5) we mean a weak solution. More precisely, since $-L_x$ is a positive definite, densely defined, symmetric operator on $H_0^1(\omega)$, let $\mathcal{L}$ be its positive-definite square root. We define the bilinear form

$$B^{\mathbf{r}}[u,v;t] = \int_\Omega (\mathcal{L}u)^2 + \sum_{q=2}^{n} C_q d^{2(q-1)} (\mathcal{L}^q u)^2 \, dx. \tag{6}$$

A weak solution of (5) is then a function $u \in L^2(0,T,H_0^m(\omega))$, $\dot{u} \in L^2(0,T,L^2(\omega))$, and $\ddot{u} \in L^2(0,T,H^{-m}(\omega))$ that satisfies

$$\langle \ddot{u}, v\rangle + B^{\mathbf{r}}[u,v;t] = 0,$$

for all $v \in H_0^m(\omega)$ and for almost every time $0 \le t \le T$, where

$$u(0) = \overline{u^{\mathbf{r}}}, \qquad \dot{u}(0) = \overline{u_t^{\mathbf{r}}}.$$

(Here, $\langle \ddot{u}, v\rangle$ denotes the pairing of $\ddot{u} \in H^{-m}$, considered as a linear functional on H^m, acting on $v \in H^m$.) Note that if we consider the results of Theorem 1 for the case $n = 1$ we obtain a particularly simple reduced equation, at the expense of obtaining an approximation for only a relatively short length of time.

Corollary 1. *Fix $C_0 > 0$ and $\epsilon > 0$ and let $u(x,y,t)$ satisfy* (2) *with initial condition $(u^0, u_t^0) \in H^2 \times H^1$ satisfying $\|u^0\|_{H^2} + \|u_t^0\|_{H^1} \le C_I$, define $(\overline{u^{0\mathbf{r}}}, \overline{u_t^{0\mathbf{r}}})$ in $H_\alpha^2 \times H_\alpha^1$ as the projections into H_α^2 and H_α^1 of $\Pi(u^0)$ and $\Pi(u_t^0)$. There exists $d_0 = d_0(C_0,\epsilon,\eta)$, and $C_T = C_T(C_0,\epsilon,\eta)$ such that for all $d < d_0$, Let $\overline{u_n{}^{\mathbf{r}}}$ be the solution with these initial conditions of the equation*

$$\partial_t^2 u^{\mathbf{r}} = L_x u^{\mathbf{r}}, \tag{7}$$

subject to the boundary conditions

$$u^{\mathbf{r}}|_{\partial\omega} = 0. \tag{8}$$

Then one has

$$\|u(x,y,t) - \overline{u_n^{\mathbf{r}}}(x,t)\mathbf{1}(y)\|_{H^1\times L^2} \le \epsilon$$

for all $t \le C_T d^{-(3-\eta)}$.

We now turn to the (apparently) more difficult task of deriving reduced equations for our original equation (1). As we shall see, by combining Theorem 1 with results of [3], we immediately obtain an algorithm for deriving reduced equations for (1). In [3], Babuška and Schwab considered the static version of (1), namely the elliptic equation:

$$\begin{aligned} 0 &= Lv \text{ in } \Omega, \\ v &= 0 \text{ on } \Gamma, \\ \partial_n v &= f^{\pm} \text{ on } R_{\pm}. \end{aligned} \tag{9}$$

They derived an efficient dimensional reduction method which given an s and an ϵ, allows one to derive a system of (two dimensional) elliptic PDE's whose solution $\tilde{v} \in H^s$ and if v is the solution of (9) one has $\|v - \tilde{v}\|_{H^1} < \epsilon$.

We now proceed in the obvious fashion. Given a solution u of (1), and a solution v of (9), the function $w = u - v$ solves (2). We then use the results of [3] to derive a dimensionally reduced approximation to v and Theorem 1 to derive a dimensionally reduced approximation to w and then their sum gives a dimensionally reduced approximation to the solution of (1). There is one slight complication in this approach which is that since we do not know v exactly, we do not know the initial conditions of (2) exactly. To circumvent this difficulty we proceed as follows. Fix $\beta > 0$. By the results of [3] there exists a system of reduced equations for (9) whose solution $v^{\mathbf{r}} \in H^2$ and which satisfies

$$\|v - v^{\mathbf{r}}\|_{H^2} \le \beta \tag{10}$$

Now consider the equation

$$\begin{aligned} w_{tt} &= Lw \text{ in } \Omega, \\ w &= 0 \text{ on } \Gamma, \\ \partial_n w &= 0 \text{ on } R_{\pm}, \\ w|_{t=0} &= u^0 - v^{\mathbf{r}}\ , w_t|_{t=0} = u_t^0 \end{aligned} \tag{11}$$

We would like to apply Theorem 1 to derive reduced equations for (11), but since $v^{\mathbf{r}}$ does not exactly satisfy the boundary conditions $\partial_n v = f^{\pm}$ on $R_{\pm}$ on the top and bottom faces of the domain, $w|_{t=0} = u^0 - v^{\mathbf{r}}$ will not exactly satisfy $\partial_n w = 0$ on $\mathbb{R}_{\pm}$. To circumvent this difficulty we appeal to the following lemma.

Lemma 2. *For any* $\beta > 0$, *there exists a function* $w^0 \in H^2(\Omega)$ *such that*

- $\partial_n w^0 = 0$ *on* $R_\pm$ *and* $w^0|_{\omega\times(0,1)} = 0$
- $\|w^0 - (u^0 - v^{\mathbf{r}})\|_{H^1} < \beta$.

Proof: The proof is a straightforward approximation argument. Define $\mathcal{U} = (u^0 - v^{\mathbf{r}})$. Choose $\mathcal{U}^\infty \in C^\infty(\Omega) \cap H^2(\Omega)$ such that $\|\mathcal{U}^\infty - \mathcal{U}\|_{H^2} < \beta/4$. Furthermore since $\mathcal{U}|_{\omega\times(0,1)} = 0$, we can assume the same of $\mathcal{U}^\infty$. Since $\mathcal{U}^\infty$ is smooth, we have

$$\mathcal{U}^\infty(x,y) = \mathcal{U}^\infty(x^*,y^*) + \int_{(x^*,y^*)}^{(x,y)} D\mathcal{U}^\infty(\xi,\eta)\cdot d\ell\ ,$$

for any points (x,y) and (x^*,y^*) in Ω. Now replace $D\mathcal{U}^\infty$ by $\Delta \in C^\infty$ where $\Delta(x,y) = D\mathcal{U}^\infty(x,y)$ if $x \in \omega$, $y \in (\nu d, (1-\nu)\pi d)$, while $\Delta(x,y) = 0$ if $y \in (0, \frac{\nu}{2}d) \cup ((1-\frac{\nu}{2})d\pi, \pi d)$, and $\|D\mathcal{U}^\infty - \Delta\|_{L^2(\Omega)} < \beta/4$. Setting

$$w^0(x,y) = \mathcal{U}^\infty(x^*,y^*) + \int_{(x^*,y^*)}^{(x,y)} \Delta(\xi,\eta)\cdot d\ell\ ,$$

we find that $\partial_n w^0|_{R_\pm} = 0$ and $w^0|_{\omega\times(0,1)} = 0$ by construction while

$$\|w^0 - (u^0 - v^{\mathbf{r}})\|_{H^1} \le \|w^0 - (u^0 - v^{\mathbf{r}})\|_{L^2} + \|Dw^0 - D(u^0 - v^{\mathbf{r}})\|_{L^2}$$

The second term on the right hand side of this inequality is bounded $\|\Delta - D\mathcal{U}^\infty\|_{L^2} + \|D\mathcal{U}^\infty - D(u^0 - v^{\mathbf{r}})\|_{L^2} < \frac{\beta}{4} + \frac{\beta}{4}$. On the other hand, $w^0(x,y) - (u^0 - v^{\mathbf{r}})(x,y) = 0$ if $x \in \omega$ and $y \in (\nu\delta, (1-\nu)\pi\delta)$, so the first term can be made less than $\frac{\beta}{2}$ by making ν sufficiently small.

■

We now continue the discussion of deriving reduced equations for (1) by replacing (11) by

$$\begin{aligned} w_{tt} &= Lw \text{ in } \Omega, \\ w &= 0 \text{ on } \Gamma, \\ \partial_n w &= 0 \text{ on } R_\pm, \\ w|_{t=0} &= w^0\ , w_t|_{t=0} = u_t^0 \end{aligned} \tag{12}$$

Given any $n > 0$, Theorem 1, guarantees that there exists a reduced equation whose solution $\overline{w^{\mathbf{r}}}$ satisfies

$$\sup_{t\in[0,C_T d^{-N_\eta(n)}]} \|w(\cdot,t) - \overline{w^{\mathbf{r}}}(\cdot,t)\|_{H^1}\ < \beta. \tag{13}$$

Now consider the function $u^{\mathbf{r}} = v^{\mathbf{r}} + \overline{w^{\mathbf{r}}}$. Note that $u^{\mathbf{r}}$ is the sum of solutions of reduced equations, and we now show that it is also an accurate approximation to the solution of (1). Let $\tilde{w}$ be the solution of

$$\begin{aligned} \tilde{w}_{tt} &= L\tilde{w} \text{ in } \Omega, \\ \tilde{w} &= 0 \text{ on } \Gamma, \\ \partial_n \tilde{w} &= 0 \text{ on } R_\pm, \\ \tilde{w}|_{t=0} &= u^0 - v\ , \tilde{w}_t|_{t=0} = u_t^0 \end{aligned} \tag{14}$$

where v is the (exact) solution of (9). Note that by Lemma 2, plus the results of [3], we can assume that $\|w^0 - (u^0 - v)\|_{H^1} \leq \beta$. We will need another easy lemma, this one based on the fact that (12) and (14) are Hamiltonian equations and hence preserve energy.

Lemma 3. *Let w be the solution of the intial-boundary value problem (12), and let $\tilde{w}$ be the solution of (14). Then there exists a constant $C > 0$ such that*

$$\|w(\cdot,t) - \tilde{w}(\cdot,t)\|_{H^1} \leq C\beta \tag{15}$$

for all $t \geq 0$.

We postpone the proof of this lemma until the next section where the Hamiltonian nature of the problem is discussed. With this estimate in hand, we complete the proof that $u^{\mathbf{r}}$ provides a good approximation to the solution of (1)

$$\begin{aligned}
\|u - u^{\mathbf{r}}\|_{H^1} &= \|u - v - \overline{w^{\mathbf{r}}} + v - v^{\mathbf{r}}\|_{H^1} \leq \|\tilde{w} - \overline{w^{\mathbf{r}}}\|_{H^1} + \|v - v^{\mathbf{r}}\|_{H^1} \\
&= \|\tilde{w} - w + w - \overline{w^{\mathbf{r}}}\|_{H^1} + \|v - v^{\mathbf{r}}\|_{H^1} \\
&\leq \|\tilde{w} - w\|_{H^1} + \|w - \overline{w^{\mathbf{r}}}\|_{H^1} + \|v - v^{\mathbf{r}}\|_{H^1} \\
&\leq C\beta + \beta + \beta
\end{aligned}$$

Thus, if we pick $(C + 2)\beta < \epsilon$ we have proven that the solutions of (1) can be approximated by the solutions of reduced equations. More precisely, we have shown:

Theorem 2. *Fix any $n > 0$, $\epsilon > 0$, $\eta > 0$ and $C_0 > 0$. There exist constants $d_0 = d_0(C_0, \epsilon, \eta)$, and $C_T = C_T(C_0, \epsilon, \eta)$ and there exists a function $N_\eta(n)$ as in Theorem 1 such that for all $d \leq d_0$, if $u(x, y, t)$ is the solution of (1) with initial conditions $(u^0, u^0_t) \in H^2 \times H^1$ satisfying $\|u^0\|_{H^2} + \|u^0_t\|_{H^1} \leq C_0$, then there exist reduced equations whose solutions, which we denote $u^{\mathbf{r}}$, satisfy*

$$\|u(\cdot,t) - u^{\mathbf{r}}(\cdot,t)\|_{H^1} \leq \epsilon$$

for all $0 \leq t \leq C_T d^{-N_\eta(n)}$.

1.2 An outline of the proof

To prove Theorem 1, we will actually do most of our work in the frequency domain as opposed to the spatiotemporal. A common approach in applying KAM or Nekhoroshev theory to partial differential equations is to expand our solution in terms of a basis of the phase space (see [17]). We do that here and then rewrite the Hamiltonian in these new coordinates. We then describe a program which will allow us to change variables to get a Hamiltonian which will generate the type of flow we need to prove Theorem 1.

Our first step is to make the second-order system (1) into a first-order Hamiltonian system with the Hamiltonian function

$$\begin{aligned}
&\tilde{H} \colon H_1 \times L_2 \to \mathbb{R} \\
&\qquad (u, v) \mapsto \frac{1}{2} \int_\Omega \left\{ v^2 + a(y/d)(u_y)^2 + b(y/d) \left(\nabla_x u\right)^T C(x) \nabla_x u \right\} dV.
\end{aligned} \tag{16}$$

The canonical equations

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \delta_u \tilde{H}(u, u_t) \\ \delta_v \tilde{H}(u, u_t) \end{pmatrix} = \begin{pmatrix} v \\ (a(y/d)u_y)_y + b(y/d)\nabla_x(C(x)\nabla_x u) \end{pmatrix}$$

are easily seen to be equivalent to (1). Furthermore, it is also easily seen that our assumptions on a, b, and C make the function $\tilde{H}$ equivalent to the canonical norm on $H^1 \times L^2$.

Following the terminology of [17], we consider the Hilbert scale defined as $X_s = H^s \times H^{s-1}$. Then $\tilde{H}$ is a function defined on X_1, and $\nabla \tilde{H}$ is an isomorphism of the scale of order 1, i.e. $\nabla H\colon X_s \to X_{s-1}$. Defining $\tilde{J}$ to be the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ above, we also have that $\tilde{J}$ is an isomorphism of order 1. The canonical equations above, stated in this language, say that our PDE can be written as

$$\dot{u} = \tilde{J}\nabla\tilde{H}(u) =: V_{\tilde{H}}(u)$$

where $V_{\tilde{H}}$ is an isomorphism of order 2.

We can now easily prove Lemma 3.

Proof: By standard results from the theory of Hamiltonian PDE's, the value of $\tilde{H}$ is constant along solutions of (1). Furthermore, because of the assumptions on the coefficient functions, a, b, and c, there exist constants $\underline{c}$ and $\overline{c}$ such that

$$\underline{c}(\|u\|_{H^1} + \|v\|_{L^2}) \le \tilde{H}(u, v) \le \overline{c}(\|u\|_{H^1} + \|v\|_{L^2}) \ . \tag{17}$$

Thus, if $w(\cdot, t)$ is the solution of (12) and $\tilde{w}(\cdot, t)$ to solution of (14). Then the difference $w - \tilde{w}$ satisfies the same differential equation, with initial conditions which satisfy $\|(w - \tilde{w})|_{t=0}\|_{H^1} \le \beta$, and $(w_t - \tilde{w}_t)|_{t=0} = 0$. But then the conservation of the Hamiltonian along solutions and the estimate in (17) immediately imply Lemma 3.

■

We now chose a basis for this phase space, and expand the Hamiltonian with respect to that basis. We have defined the y-range of our domain to be $[0, \pi d]$, so we introduce the new variable $\eta = y/d$, which ranges over $[0, \pi]$. Consider the eigenproblems

$$\begin{aligned} L_\eta \psi = \frac{\partial}{\partial \eta}\left(a(\eta)\frac{\partial \psi}{\partial \eta}\right) &= -\lambda\psi, \qquad \frac{\partial \psi}{\partial \eta}(0) = \frac{\partial \psi}{\partial \eta}(\pi) = 0, \\ L_x \phi = \hat{b}(0)\nabla_x(C(x)\nabla_x \phi) &= -\mu\phi, \qquad \phi|_{\partial\omega} = 0. \end{aligned}$$

Classical results (see [16],[24], [8]) imply that the above non-singular Sturm-Liouville problems define complete L^2-orthonormal sequences of eigenfunctions and associated eigenvalues:

$$L_\eta \psi_l = -\lambda_l \psi_l, \tag{18}$$

$$L_x \phi_k = -\mu_k \phi_k, \tag{19}$$

and we can index the eigenvalues so that

$$0 = \lambda_0 \le \lambda_1 \le \lambda_2 \le \cdots$$
$$0 < \mu_1 \le \mu_2 \le \cdots$$

Furthermore, the eigenvalues satisfy the asymptotic estimates $\lambda_l = \mathcal{O}(l^2)$ and $\mu_k = \mathcal{O}(k)$.

To rewrite $\tilde{H}$ in frequency coordinates, we expand u in its Fourier series as

$$u(x,y,t) = \sum_{k,l} \hat{u}_{k,l}(t)\phi_k(x)\psi_l(y/d), \tag{20}$$

where ϕ_k and ψ_l are the above eigenfunctions. Note that this expansion (when coupled with the transform of $\dot{u}$) is a canonical change of variables.

Plugging (20) into (16) gives

$$\tilde{H} = \frac{d}{2}\sum_{k,l}\left(\dot{\hat{u}}_{k,l}(t)\right)^2 + \frac{1}{2d}\sum_{k,l}\lambda_l\left(\hat{u}_{k,l}(t)\right)^2 + \frac{d}{2}\sum_{k,l,l'}\mu_k \hat{u}_{k,l}(t)\hat{u}_{k,l'}(t)\frac{\hat{b}(l,l')}{\hat{b}(0)}, \tag{21}$$

where we define

$$\hat{b}(l,l') = \int_0^\pi b(\xi)\psi_l(\xi)\psi_{l'}(\xi)\,d\xi.$$

We now introduce complex coordinates for our Hamiltonian defining

$$z_{kl} = \frac{1}{\sqrt{2\omega_{kl}}}\dot{\hat{u}}_{k,l} + i\sqrt{\frac{\omega_{kl}}{2}}\hat{u}_{k,l}, \tag{22}$$

with

$$\omega_{kl}^2 = \frac{\lambda_l}{d^2} + \mu_k. \tag{23}$$

In terms of these new coordinates the Hamiltonian takes the form

$$\tilde{H} = d\sum_{k,l}\omega_{kl}\left|z_{kl}\right|^2 - \frac{d}{4}\sum_{k,l,l'}\frac{\mu_k}{\sqrt{\omega_{kl}\omega_{kl'}}}\beta_{l,l'}\left(z_{kl}-\overline{z}_{kl}\right)\left(z_{kl'}-\overline{z}_{kl'}\right),$$

where we define

$$\beta_{l,l'} = \frac{\hat{b}(l,l')}{\hat{b}(0)} - \delta_{l,l'}.$$

For ease of notation we define

$$\Gamma_{k,l,l'} = \frac{-\mu_k}{4\sqrt{\omega_{kl}\omega_{kl'}}}\beta_{l,l'}, \tag{24}$$

and thus our Hamiltonian finally takes the form

$$\tilde{H} = d\sum_{k,l}\omega_{kl}\left|z_{kl}\right|^2 + d\sum_{k,l,l'}\Gamma_{k,l,l'}\left(z_{kl}-\overline{z}_{kl}\right)\left(z_{kl'}-\overline{z}_{kl'}\right). \tag{25}$$

In these coordinates, Equation (1) is equivalent to the Hamiltonian system

$$\dot{z} = J\nabla \tilde{H}(z), \tag{26}$$

where we write $z = (z_{kl}, \overline{z}_{kl})$, and $J = i\begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$.

Note that, formally, $\tilde{H} = \mathcal{O}(d)$. We define a new Hamiltonian H so that $\tilde{H} = d \cdot H$. Note then that

$$\dot{z} = J\nabla \tilde{H}(z) = dJ\nabla H(z).$$

If we define $\tau = d \cdot t$, then we have $\frac{dz}{d\tau} = J\nabla H(z)$.

All we have done here is to rescale time from t to τ to remove a factor of d from the differential equation.

Recall that we are trying to find a dimensionally-reduced PDE, i.e. one which depends only on x_1 and x_2, and not on y. One way to do this would be to consider only the modes with $l = 0$, i.e. the z_{k0} modes. For example, if we consider the reduced Hamiltonian $H^{\mathbf{r}} = \sum_k \omega_{k0} |z_{k0}|^2 = \sum_k \frac{1}{2}\left(\dot{\hat{u}}_{k,0}^2 + \mu_k \hat{u}_{k,0}^2\right)$. then reverting to continuous variables, we have

$$H^{\mathbf{r}} = \frac{1}{2}\int_\Omega \left(\frac{\partial}{\partial t}(u^{\mathbf{r}})\psi_0(y/d)\right)^2 + \hat{b}(0)\left((\nabla_x u^{\mathbf{r}}\psi_0(y/d))\right)^T C(x)(\nabla_x u^{\mathbf{r}}\psi_0(y/d))\, dV,$$

whose canonical equations are easily seen to be equivalent to

$$u_{tt}^{\mathbf{r}} = \hat{b}(0)\nabla_x \cdot (C(x)\nabla_x u^{\mathbf{r}}) = L_x u^{\mathbf{r}},$$

an equation which depends only on the x_1, x_2 variables, and not y.

This is the theme. To dimensionally reduce our equation in the discrete variables, we will work with an equation in which terms with $l > 0$ are omitted. However, it is by no means clear that the resulting equations will provide an accurate approximation to the original dynamics. To make the replacement legitimate, we will have to show that the modes with $l = 0$ (the reduced solution) approximate all of the modes (the full solution) sufficiently well. In Section 2 we will make canonical changes of variables which ensure that the part of the Hamiltonian with $l = 0$ provides a better and better approximation to the dynamics of the full system. However, this introduces a new problem since after these changes of variables, it is no longer clear that the $l = 0$ terms in the (discrete) Hamiltonian represent a partial differential equation. Showing that this is the case will be done in Section 4.

We should note at this point that the second term in our Hamiltonian is the source of all of our difficulties. For example, if we knew that the second term in (25) did not couple any $l = 0$ terms with $l > 0$ terms, we would be done. Although the evolution of the $l > 0$ terms might be complicated under this Hamiltonian, it would not affect the $l = 0$ terms at all, and using the smoothness of the initial data, we can conclude that very little of the energy of the system is in the modes with $l > 0$. In this case, simply throwing away the $l > 0$ terms would give us a dimensionally-reduced PDE. Geometrically, what we would have shown is the submanifold generated by the $l = 0$ modes is invariant under the flow generated by this Hamiltonian.

This is, of course, too much to hope for. This could only happen if $\Gamma_{k,l,0} = 0$ for all l. This would be true only if $\hat{b}(l, 0) = 0$ for all l. But

$$\hat{b}(l, 0) = \int_0^\pi b(\eta)\psi_l(\eta)\psi_0(\eta)\, d\eta = \pi^{-1/2} \int_0^\pi b(\eta)\psi_l(\eta)\, d\eta$$

is (essentially) the lth Fourier coefficient of b, expanded with respect to the ψ_l. These will all vanish only in the case of constant b.

Thus, except in the most exceptional of cases, we are resigned to coupling between the $l = 0$ and the $l > 0$ terms. Our goal here is to make the coupling as small as possible. With this goal in mind, we write our Hamiltonian as

$$H = H^0_{\text{diag}} + H^0_1 + H^0_2,$$

where

$$\begin{aligned} H^0_{\text{diag}} &= \sum_{k,l} \omega_{kl} \left|z_{kl}\right|^2, \\ H_1 &= \sum_{\substack{k,l \\ l>0}} \left(\Gamma_{k,l,0} + \Gamma_{k,0,l}\right)\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{k0} - \overline{z}_{k0}\right), \\ H_2 &= \sum_{\substack{k,l,l' \\ l,l'>0}} \Gamma_{k,l,l'}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{kl'} - \overline{z}_{kl'}\right). \end{aligned}$$

Each of the above is a quadratic function defined on an appropriate Hilbert space $\mathcal{H}$ (to be specified below). We will refer to quadratic functions on this Hilbert space as one of three types: A term is of Type 0 if it couples only $l = 0$ and $l = 0$ terms, i.e. every quadratic summand is of the form $z_{k0}z_{k0}, z_{k0}\overline{z}_{k0}, etc.$. A term is of Type 1 if it couples $l = 0$ and $l > 0$ terms, and a term is of Type 2 if it couples $l > 0$ and $l > 0$ terms. Note that in our original Hamiltonian there are no Type 0 terms except for those already in the diagonal piece, but the canonical transformations we make in the course of the proof will generate such terms.

In light of the above comments, if we have a Hamiltonian with no Type 1 terms, the system would decouple, meaning that the points $\{z_{kl} \mid z_{kl} = 0 \text{ if } l > 0\}$ form an invariant submanifold. If we could find a change of coordinates in which our new Hamiltonian had no Type 1 terms, then, again, the system would effectively decouple, or, in the new coordinates, the $l = 0$ terms form an invariant manifold. We will find that we cannot do this, but we will be able to change coordinates so that the Hamiltonian in the new coordinates has a small Type 1 piece. Then, what we have done is find coordinates such that in these new coordinates, the submanifold generated by the $l = 0$ modes is nearly invariant, in that the submanifold interacts with the rest of the space only very weakly.

Briefly, we implement this program as follows. In the next section we first formally construct canonical transformations that reduce the size of the coupling between the modes with $l = 0$ and $l > 0$. In Subsection 2.2 we prove estimates showing that these formal canonical transformations are well defined. Then in Section 3 we prove estimates on the transformed Hamiltonian showing that these canonical transformations do actually reduce the size of the coupling between the $l = 0$ and $l > 0$ modes. Finally, in Section 4 we show that the $l = 0$ modes of the transformed Hamiltonian define dynamics which accurately approximate the dynamics of the original Hamiltonian for a very long time and then show that these $l = 0$ modes actually correspond to the Hamiltonian of the reduced equations in Theorem 1.

2 The transformed Hamiltonian

In this section we describe how the Hamiltonian (25) can be transformed so as to make the coupling between the modes with $l = 0$ and $l > 0$ smaller. In Subsection 2.1 we compute formally what the canonical transformation that reduces this coupling should be. Then in Subsection (2.2) we prove consider what happens if we iterate these canonical transformations.

2.1 The formal change of variables

In this section, we describe the type of canonical transformation we will use to reduce the size of the Type 1 terms, and prove a lemma which formally describes the effects of these canonical transformations.

If we have a Hamiltonian system with Hamiltonian χ, then the time-t map of the system, ϕ^t, is a symplectomorphism. Let us denote $\phi := \phi^1$, and we can calculate

$$H \circ \phi = H + \{\chi, H\} + \frac{1}{2}\{\chi, \{\chi, H\}\} + \cdots$$

where $\{A, B\} = \langle J\nabla A, \nabla B\rangle$. Recall that we write $H = H^0_{\text{diag}} + H_1 + H_2$, so that

$$\begin{aligned} H \circ \phi = &H^0_{\text{diag}} + H_1 + H_2 \\ &+ \left\{\chi, H^0_{\text{diag}}\right\} + \{\chi, H_1\} + \{\chi, H_2\} \\ &+ \frac{1}{2}\left\{\chi, \left\{\chi, H^0_{\text{diag}}\right\}\right\} + \frac{1}{2}\{\chi, \{\chi, H_1\}\} + \frac{1}{2}\{\chi, \{\chi, H_2\}\} + \cdots \end{aligned}$$

If we choose χ so that $\left\{\chi, H^0_{\text{diag}}\right\} = -H_1$, then we would also have

$$\left\{\chi, \left\{\chi, H^0_{\text{diag}}\right\}\right\} = -\{\chi, H_1\},$$

so that

$$\begin{aligned} H \circ \phi = &H^0_{\text{diag}} + H_2 + \{\chi, H_2\} + \{\chi, \{\chi, H_2\}\} + \cdots \\ &+ \frac{1}{2}\{\chi, H_1\} + \left(\frac{1}{2} - \frac{1}{6}\right)\{\chi, \{\chi, H_1\}\} + \cdots \end{aligned} \tag{27}$$

Although this is a complicated expression, we should note that the original Type 1 term has disappeared, and that everything in this expansion which may potentially be a Type 1 term is $\mathcal{O}(\chi)$. If χ is small, then this is an improved Hamiltonian from our standpoint, since the Type 1 terms are now smaller.

We can repeat this process. We identify the lowest-order Type 1 terms in $H \circ \phi$, then make a similar change of variables to cancel them. Hopefully, this improves our Hamiltonian further, in that the remaining Type 1 terms are yet smaller.

Lemma 4. *If the Type* 1 *term of* H *is of the form* $A = \sum_{\substack{k,l \\ l\neq 0}} \hat{A}_{k,l}\,(z_{kl} - \overline{z}_{kl})\,(z_{k0} - \overline{z}_{k0})$, *we can choose* χ *to solve*

$$\left\{\chi, H^0_{\text{diag}}\right\} = -A,$$

with the following properties:

1.
$$\chi = \sum_{\substack{k,l \\ l>0}} \chi_{kl}(z_{kl}z_{k0} - \overline{z}_{kl}\overline{z}_{k0}) + \chi_{k\bar{l}}(z_{kl}\overline{z}_{k0} - \overline{z}_{kl}z_{k0}),$$

   with
$$\chi_{kl} = \frac{-i\hat{A}_{k,l}}{\omega_{k0} + \omega_{kl}}\ , \quad \chi_{k\bar{l}} = \frac{-i\hat{A}_{k,l}}{\omega_{k0} - \omega_{kl}}. \tag{28}$$

2. *If* H *is a quadratic term of Type* 0, 1, *or* 2 *as defined above, then* $\{\chi, H\}$ *is a linear combination of terms of Type* 0, 1, 2, *i.e. the map* $\{\chi, \cdot\}$ *maps the set of these quadratic terms to itself. More specifically, we have the following table:*

Type of H	*Type of* $\{\chi, H\}$
0	1
1	$0+2$
2	1

Proof: To prove the first part, we simply make the Ansatz that

$$\chi = \sum_{k,l} \chi_{kl}z_{kl}z_{k0} + \chi_{\overline{k}l}\overline{z}_{kl}z_{k0} + \chi_{k\overline{l}}z_{kl}\overline{z}_{k0} + \chi_{\overline{kl}}\overline{z}_{kl}\overline{z}_{k0}, \tag{29}$$

and solve the equation

$$\left\{\chi, H^0_{\text{diag}}\right\} = -A.$$

An easy calculation then shows that we can choose $\chi_{\overline{k}l} = -\chi_{k\bar{l}}$ and $\chi_{\overline{kl}} = -\chi_{kl}$, and that

$$\chi_{kl} = \frac{-i\hat{A}_{k,l}}{\omega_{k0} + \omega_{kl}}\ , \quad \chi_{k\bar{l}} = \frac{-i\hat{A}_{k,l}}{\omega_{k0} - \omega_{kl}}. \tag{30}$$

If we write χ as

$$\chi = \sum_{\substack{k,l \\ l>0}} \chi_{kl}(z_{kl}z_{k0} - \overline{z}_{kl}\overline{z}_{k0}) + \chi_{k\bar{l}}(z_{kl}\overline{z}_{k0} - \overline{z}_{kl}z_{k0}).$$

and assume that we have a general quadratic function $H = H_{k,l,l'}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{kl'} - \overline{z}_{kl'}\right)$, then using the definition of the Poisson bracket, one has

$$\begin{aligned}\{\chi, H\} = & \sum_{\substack{k,l,l' \\ l>0}} (H_{k,0,l'} + H_{k,l',0}) S_{kl}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{kl'} - \overline{z}_{kl'}\right) \\ & + \sum_{k,l}\left(\sum_{l'>0}(H_{k,l,l'} + H_{k,l',l}) D_{kl'}\right)\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{k0} - \overline{z}_{k0}\right),\end{aligned} \tag{31}$$

where

$$S_{kl} = i\left(\chi_{kl} + \chi_{k\bar{l}}\right), \tag{32}$$

$$D_{kl} = i\left(\chi_{kl} - \chi_{k\bar{l}}\right). \tag{33}$$

If H is of Type 0, then the second sum disappears entirely, and the first sum only has terms with $l' = 0$. Thus $\{\chi, H\}$ is of Type 1. If H is of Type 2, then the first sum disappears, and the second sum has terms with $l > 0$. Thus $\{\chi, H\}$ is again of Type 1. If H is of Type 1, then the first sum has terms with $l' > 0$, making it Type 2, and the second sum has terms with $l = 0$, and is thus of Type 0.

Note that if H is of Type 1, then so is $\{\chi, \{\chi, H\}\}$. If H is of Type $0 + 2$, then so is $\{\chi, \{\chi, H\}\}$.

■

We make a few remarks about the previous lemma. First, we can calculate directly that

$$S_{kl} = i\left(\chi_{kl} + \chi_{k\bar{l}}\right) = \frac{\hat{A}_{kl}}{\omega_{k0} + \omega_{kl}} + \frac{\hat{A}_{kl}}{\omega_{k0} - \omega_{kl}} = -\frac{2d^2\omega_{k0}\hat{A}_{kl}}{\lambda_l}, \tag{34}$$

$$D_{kl} = i\left(\chi_{kl} - \chi_{k\bar{l}}\right) = \frac{\hat{A}_{kl}}{\omega_{k0} + \omega_{kl}} - \frac{\hat{A}_{kl}}{\omega_{k0} - \omega_{kl}} = \frac{2d^2\omega_{kl}\hat{A}_{kl}}{\lambda_l}. \tag{35}$$

Second, we point out a potentially confusing abuse of notation. Recall that we have defined H_1^0 as the Type 1 piece of the Hamiltonian H^0, and that we have

$$H_1^0 = \sum_{\substack{k \\ l>0}} (\Gamma_{k,l,0} + \Gamma_{k,0,l})\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{k0} - \overline{z}_{k0}\right).$$

Since we do have that $\Gamma_{k,l,0} = \Gamma_{k,0,l}$, we can write

$$H_1^0 = \sum_{\substack{k \\ l>0}} 2\Gamma_{k,l,0}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{k0} - \overline{z}_{k0}\right).$$

Throughout the paper, we will solve equations of the form

$$\left\{\chi, H^0_{\text{diag}}\right\} = G_1,$$

where G is some Hamiltonian, and G_1 its Type 1 piece. We will frequently choose coefficients for symmetry, i.e. so that $G_{1,kl0} = G_{1,k0l}$. In this case, if we apply Lemma 4, the term $\hat{A}_{kl}$ will be $2G_{1,kl0}$ and not $G_{1,kl0}$. Also, we calculate that

$$S_{kl} = -\frac{4d^2\omega_{k0}G_{kl0}}{\lambda_l}, \qquad D_{kl} = \frac{4d^2\omega_{kl}G_{kl0}}{\lambda_l}.$$

We now extend the previous lemma and compute the form of the terms in $\{\chi, H\}$:

Lemma 5. *If H is symmetric in the sense that*

$$H_{k,l,l'} = H_{k,l',l} \text{ for all } l, l',$$

then we can choose coefficients so that $\{\chi, H\}$ is also symmetric. Furthermore, the coefficients of $\{\chi, H\}$ can be found in the following chart:

H	*Type*	$\{\chi, H\}$
$\sum H_{k,0,0}\left(z_{k0} - \overline{z}_{k0}\right)\left(z_{k0} - \overline{z}_{k0}\right)$	0	$\{\chi, H\}_{kl0} = H_k S_{kl}$
$\sum H_{k,l,0}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{k0} - \overline{z}_{k0}\right)$	1	$\{\chi, H\}_{0,k00} = 2\sum_{l>0} H_{k,l,0} D_{kl}$ $\{\chi, H\}_{2,kll'} = H_{k,l',0}S_{kl} + H_{k,l,0}S_{kl'}$
$\sum H_{k,l,l'}\left(z_{kl} - \overline{z}_{kl}\right)\left(z_{kl'} - \overline{z}_{kl'}\right)$	2	$\{\chi, H\}_{kl0} = \sum_{l'>0} H_{k,l,l'} D_{kl'}$

where we define

$$S_{kl} = i\left(\chi_{kl} + \chi_{k\bar{l}}\right),$$
$$D_{kl} = i\left(\chi_{kl} - \chi_{k\bar{l}}\right).$$

Recall that if H is Type 1, then $\{\chi, H\}$ is Type $0 + 2$. We denote this in the above table by saying that $\{\chi, H\}$ has two pieces, one of Type 0, $\{\chi, H\}_0$, and one of Type 2, $\{\chi, H\}_2$.

The proof is a straightforward computation applying Lemma 4.

2.2 Iterating the canonical transformations

In this subsection we study (still largely on a formal level) what happens if we iterate the sort of canonical transformations constructed in the previous subsection. Let us define the Hilbert space $\mathcal{H}^s$ as

$$\mathcal{H}^s = \left\{ \{z_{kl}\}_{k,l=0}^{\infty} | \sum_{k,l} \omega_{kl}^{2s} |z_{kl}|^2 < \infty \right\},$$

with inner product

$$\langle z, w\rangle_{\mathcal{H}^s} = \sum_{k,l}^{\infty} \omega_{kl}^{2s} z_{kl} \overline{w}_{kl}.$$

Remark 5. *Note that $z \in \mathcal{H}^{1/2}$ if and only if $H^0_{\mathrm{diag}}(z) = \sum_{k,l} \omega_{kl} |z_{kl}|^2 < \infty$. Thus, $\mathcal{H}^{1/2}$ is a natural space in which to work.*

We say that H is a quadratic functional on $\mathcal{H}^s$ if $H \colon \mathcal{H}^s \to \mathbb{R}$, and $H(z) = \langle z, Az\rangle_{\mathcal{H}^0}$ where A is a linear map on $\mathcal{H}^s$. We denote the set of all $\mathcal{C}^1$ quadratic functionals on $\mathcal{H}^s$ as $\mathcal{QF}(\mathcal{H}^s)$. Thus $H \in \mathcal{QF}(\mathcal{H}^s)$ if H is quadratic, and H_*, called the derivative of H, exists. For each x, $H_*(x)$ is by definition a linear map. We should note that since H is quadratic, $H_*(x)$ is also linear in the argument x.

Since $H_*(x) \colon \mathcal{H}^s \to \mathbb{R}$ is a linear map, we can (by the Riesz Representation Theorem) identify it with a member of $(\mathcal{H}^s)^*$. We will define our dual spaces with respect to the $\mathcal{H}^0$ inner product, i.e.

$$\|z\|_{(\mathcal{H}^s)^*} := \sup_{\substack{w \in \mathcal{H}^s \cap \mathcal{H}^0 \\ \|w\|_{\mathcal{H}^s} = 1}} \langle z, w\rangle_{\mathcal{H}^0} . \tag{36}$$

There is a canonical identification of $(\mathcal{H}^s)^*$ and $\mathcal{H}^{-s}$ (see [[17], section 1.2]) so, if $H_* \colon \mathcal{H}^s \to \mathbb{R}$, we define ∇H to be the corresponding element of $\mathcal{H}^{-s}$, i.e.

$$\langle \nabla H(x), y\rangle_{\mathcal{H}^0} = H_*(x)y.$$

By the above remarks, if we have an $H \in \mathcal{QF}(\mathcal{H}^{1/2})$, then we know that $\nabla H(z)$ exists, and a priori, $\nabla H(z) \in \mathcal{H}^{-1/2}$, for $z \in \mathcal{H}^{1/2}$. Since ∇H happens to be linear in its argument (since H is quadratic), we can think of ∇H as a linear map from $\mathcal{H}^{1/2}$ to $\mathcal{H}^{-1/2}$. Thus if $H \in \mathcal{QF}(\mathcal{H}^{1/2})$, then ∇H is a priori an anti-smoothing linear map, i.e. the range of ∇H is a Sobolev space which contains functions less smooth than those in the domain. Let us consider $\mathcal{A} \subset \mathcal{QF}(\mathcal{H}^{1/2})$, where $H \in \mathcal{A}$ if and only if ∇H maps into $\mathcal{H}^{1/2}$, and $\|\nabla H\|_{\mathcal{H}^{1/2}, \mathcal{H}^{1/2}} < \infty$, and define $\|H\|_{\mathcal{A}} = \|\nabla H\|_{\mathcal{H}^{1/2}, \mathcal{H}^{1/2}}$. Our goal in the following will be more or less to show that the terms H^0_1 and H^0_2 are in $\mathcal{A}$, and that our changes of variables will leave us in $\mathcal{A}$. This will prevent us from loosing smoothness as the inductive argument proceeds.

If we have a Hamiltonian flow under χ with an associated flow map ϕ^t, then it is again a standard result that the canonical transformation $\phi = \phi^{t=1}$

$$H \circ \phi = \sum_{i=0}^{\infty} \frac{L^i(H)}{i!} \ , \tag{37}$$

provided the sum converges. (Here $L^i(H)$ is defined inductively by $L(H) = \{\chi, H\}$ and $L^{i+1}(H) = \{\chi, L^i(H)\}$.) Furthermore, one has the "partial expansion"

$$H \circ \phi = \sum_{i=0}^{K-1} \frac{L^i(H)}{i!} + \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{K-1}} L^K(H) \circ \phi^{t_K} dt_K \ldots dt_1 \tag{38}$$

Let's assume that we start with a Hamiltonian of the form $H^0_{\text{diag}} + H_{0+2} + H_1$ as above, where

$$H^0_{\text{diag}}(z) = \sum \omega_{kl} \left|z_{kl}\right|^2 ,$$

and H_{0+2} combines the remaining terms of Type $0+2$, and H_1 is the term of Type 1. The separation of Type 1 terms from Type 0 and Type 2 terms is natural in light of Lemma 4 and our goal of eliminating the Type 1 terms. Assume further that we've chosen χ so that $\left\{\chi, H^0_{\text{diag}}\right\} = -H_1$. Then, following (37),

$$\begin{aligned} H \circ \phi &= \sum_{i=0}^{\infty} \frac{L^i(H^0_{\text{diag}})}{i!} + \sum_{i=0}^{\infty} \frac{L^i(H_1)}{i!} + \sum_{i=0}^{\infty} \frac{L^i(H_{0+2})}{i!} \\ &= H^0_{\text{diag}} + \sum_{i=0}^{\infty} \frac{L^i(H_{0+2})}{i!} + \sum_{i=1}^{\infty} \left(\frac{1}{i!} - \frac{1}{(i+1)!}\right) L^i(H_1) \ , \end{aligned}$$

where the second line uses the fact that $L(H^0_{\text{diag}}) = -H_1$ implies $L^i(H^0_{\text{diag}}) = -L^{i-1}(H_1)$, and hence a cancelation occurs. Note further that since $0! = 1!$, there is no H_1 left in the second sum, only $L^i(H_1)$ for $i > 0$.

Choosing $\xi(i) = (i+1)!/i$, we can write the above as

$$H \circ \phi = H^0_{\text{diag}} + \sum_{i=0}^{\infty} \frac{L^i(H_{0+2})}{i!} + \sum_{i=1}^{\infty} \frac{L^i(H_1)}{\xi(i)}.$$

From Lemma 4, we know that if we start with a Type 1 term, and apply an even number of Poisson brackets to it, we are left with a Type 1 term. Also, if we start with a Type 0 or Type 2 term, and apply an odd number of Poisson brackets to it, we are left with a Type 1 term. Accordingly, if we want to calculate, for example, the Type $0+2$ term of $H \circ \phi$, it is

$$(H \circ \phi)_{0+2} = \sum_{i \text{ odd}} \frac{L^i(H_1)}{\xi(i)} + \sum_{i \text{ even}} \frac{L^i(H_{0+2})}{i!}. \tag{39}$$

We note for later that the second term includes a summand for $i = 0$, i.e. the term H_{0+2} itself. Similarly, the Type 1 term of $H \circ \phi$ is

$$(H \circ \phi)_1 = \sum_{i \text{ odd}} \frac{L^i(H_{0+2})}{i!} + \sum_{\substack{i \text{ even} \\ i>0}} \frac{L^i(H_1)}{\xi(i)}. \tag{40}$$

In our particular case, we start with our original Hamiltonian given by (25), which we denote as H^0. Separating by type as we did above, we write $H^0 = H^0_{\text{diag}} + H^0_1 + H^0_{0+2}$, where we note that H^0_{0+2} happens to be a term purely of Type 2 (see Section 1.2). Because of this, we will interchange the notation H^0_2 and H^0_{0+2} below.

We choose χ_0 so that $\left\{\chi_0, H^0_{\text{diag}}\right\} = -H^0_1$, and we define ϕ_0 to be the time-1 map of the flow under χ_0. Defining $L_0(H) = \{\chi_0, H\}$, we are left with

$$H^1 = H^0 \circ \phi_0 = \sum_{i=0}^{\infty} \frac{L_0^i(H^0)}{i!} = H^0_{\text{diag}} + \sum_{i=0}^{\infty} \frac{L_0^i(H^0_2)}{i!} + \sum_{i=1}^{\infty} \frac{L_0^i(H^0_1)}{\xi(i)}.$$

If we then choose to write

$$H^1 = H^0_{\text{diag}} + H^1_1 + H^1_{0+2},$$

then we know from (40) and (39) that

$$H^1_{0+2} = \sum_{i \text{ odd}} \frac{L_0^i(H_1^0)}{\xi(i)} + \sum_{i \text{ even}} \frac{L_0^i(H_2^0)}{i!}, \tag{41}$$

$$H^1_1 = \sum_{i \text{ odd}} \frac{L_0^i(H_2^0)}{i!} + \sum_{\substack{i \text{ even} \\ i>0}} \frac{L_0^i(H_1^0)}{\xi(i)}. \tag{42}$$

We stress again that in (41), the second sum includes a term with $i = 0$, i.e. that one of the contributions to H^1_{0+2} is H^0_2 itself. Let's assume that we have shown that L_0 is an operator of small norm, so that the above power series make sense. If we think of H^0_1 and H^0_2 as terms of about the same size, then (formally) the leading order term in H^1_1 is $L_0(H^0_2)$. So if we formally ignore all but the leading order term, then we expect that H^1_1 can be bounded as

$$\left\|H^1_1\right\|_{\mathcal{A}} \leq \|L_0\|_{\mathcal{A}} \left\|H^0_2\right\|_{\mathcal{A}}. \tag{43}$$

We will show below that all of these power series do in fact converge. We note here that we expect to formally improve our approximation by a factor depending on $\|L_0\|_{\mathcal{A},\mathcal{A}}$, in the sense that the Type 1 term of our new Hamiltonian is smaller by a factor of $\|L_0\|_{\mathcal{A},\mathcal{A}}$. Again, this assumes that L_0 is an operator of small norm, and that H^0_1 and H^0_2 are terms of about the same size at the start. We will justify these assumptions below.

In general, the induction will work as follows. Assume that we have done n steps of this process, so that we have the Hamiltonian

$$H^n = H^0 \circ \phi_0 \circ \cdots \circ \phi_{n-1},$$

which we write as

$$H^n = H^0_{\text{diag}} + H^n_1 + H^n_{0+2}. \tag{44}$$

We separate out the lowest-order (in d) terms of H^n_1 to form $H^n_{1,\text{low}}$, and then define

$$\widehat{H^n_1} = H^n_1 - H^n_{1,\text{low}}.$$

We choose χ_n to solve $\left\{\chi_n, H^0_{\text{diag}}\right\} = -H^n_{1,\text{low}}$. Note by Lemma 4 this means that

$$\chi_{n,kl} = \frac{-i2H^n_{1,\text{low},kl}}{\omega_{k0} + \omega_{kl}}, \quad \chi_{n,k\bar{l}} = \frac{-i2H^n_{1,\text{low},kl}}{\omega_{k0} - \omega_{kl}}. \tag{45}$$

Doing the coordinate change as above gives us

$$\begin{aligned} H^{n+1} = H^n \circ \phi_n &= \sum_{i=0}^{\infty} \frac{L_n^i(H^n)}{i!} \\ &= H^0_{\text{diag}} + \sum_{i=1}^{\infty} \frac{L_n^i(H^n_{1,\text{low}})}{\xi(i)} + \sum_{i=0}^{\infty} \frac{L_n^i(\widehat{H^n_1} + H^n_{0+2})}{i!}. \end{aligned}$$

Applying Lemma 4 once again, we have

$$H_{0+2}^{n+1} = \sum_{i \text{ odd}} \frac{L_n^i(H_{1,\text{low}}^n)}{\xi(i)} + \sum_{i \text{ odd}} \frac{L_n^i(\widehat{H_1^n})}{i!} + \sum_{i \text{ even}} \frac{L_n^i(H_{0+2}^n)}{i!}, \tag{46}$$

$$H_1^{n+1} = \sum_{i \text{ odd}} \frac{L_n^i(H_{0+2}^n)}{i!} + \sum_{\substack{i \text{ even} \\ i>0}} \frac{L_n^i(H_{1,\text{low}}^n)}{\xi(i)} + \sum_{i \text{ even}} \frac{L_n^i(\widehat{H_1^n})}{i!}. \tag{47}$$

By inspection, we see that the lowest-order terms in (47) will be $L_n(H_{0+2}^n)$ and $\widehat{H_1^n}$. The lowest-order term of H_{0+2}^n will be H_2^0, so we expect

$$\left\|L_n(H_{0+2}^n)\right\|_{\mathcal{A}} \leq \|L_n\|_{\mathcal{A}} \left\|H_2^0\right\|_{\mathcal{A}}.$$

Since $\widehat{H_1^n}$ was chosen to be the terms *not* of lowest order in H_1^n, we expect $\widehat{H_1^n}$ to be of higher order in d than H_1^n. So, if $\|L_n\|_{\mathcal{A},\mathcal{A}}$ is significantly smaller than $\|L_{n-1}\|_{\mathcal{A},\mathcal{A}}$, then we expect $\left\|H_1^{n+1}\right\|_{\mathcal{A}}$ to be significantly smaller than $\|H_1^n\|_{\mathcal{A}}$.

Unfortunately, there is one problem with the above argument, and that is that $\left\|H_2^0\right\|_{\mathcal{A}}$ is not finite. For example, we calculate:

$$\begin{aligned}
\left\|\nabla H_2^0\right\|_{\mathcal{H}^{1/2}}^2 &= \sum_{\substack{k \\ l>0}} \omega_{kl} \left|\nabla H_{2,kl}^0\right|^2 = \sum_{\substack{k \\ l>0}} \omega_{kl} \left|\sum_{l' \neq l} \Gamma_{k,l,l'} \left(z_{kl'} - \overline{z}_{kl'}\right)\right|^2 \\
&\leq \sum_{\substack{k \\ l>0}} \omega_{kl} \sum_{\substack{l' \neq l \\ l'>0}} \frac{\omega_{k0}^4}{\omega_{kl}\omega_{kl'}} \left|\beta_{l,l'}\right|^2 \left|(z_{kl'} - \overline{z}_{kl'})\right|^2 \\
&\leq \mathcal{B}_{2,\infty} \sum_{\substack{k \\ l'>0}} \omega_{k0}^4 \omega_{kl'}^{-1} \left|(z_{kl'} - \overline{z}_{kl'})\right|^2,
\end{aligned} \tag{48}$$

where we have defined $\mathcal{B}_{2,\infty} = \sup_{l'>0} \sum_{l>0} \left|\beta_{l,l'}\right|^2$, which is finite for $b(\cdot) \in C^1$, by the Plancherel theorem and the definition of $\beta_{l,l'}$.

Looking at the final line in (48), we see that since $\omega_{kl} > \omega_{k0}$ for $l > 0$,

$$\left\|\nabla H_2^0(z)\right\|_{\mathcal{H}^{1/2}} \leq \mathcal{B}_{2,\infty} \|z\|_{\mathcal{H}^{3/2}}.$$

Unfortunately, we're trying to calculate $\left\|H_2^0\right\|_{\mathcal{A}}$, which would require the right-hand side to involve $\|z\|_{\mathcal{H}^{1/2}}$ instead.

Looking more carefully, we see the obstruction to an estimate of this type is not the behavior of our sum for large l, but for large k. For example, if we choose an $\alpha < 0$, and only consider terms in the above series for $\omega_{k0} < d^{\alpha}$, then, using (48), we have that

$$\left\|H_2^0\right\|_{\mathcal{A}} \leq \mathcal{B}_{2,\infty} d^{2\alpha}. \tag{49}$$

To make this work, in all of our Hamiltonians above, we will chop off the terms for which $k > d^{\alpha}$. We will show that this will not affect our scheme much, since all of our Hamiltonians above have the nice property that they do not couple terms with different k values. Furthermore, it follows from Lemma 4 that none of the changes of variables give rise to a term which couples different k, i.e. there is no quadratic term in the sum of the form $z_{kl}z_{k'l}$ with $k \neq k'$.

Lemma 6. *Consider a Hamiltonian H of the form $H = \sum_K H_K$ with*

$$H_K = \sum_{l,l'>0} \hat{H}_{K,l,l'} \left(z_{Kl} - \overline{z}_{Kl}\right)\left(z_{Kl'} - \overline{z}_{Kl'}\right). \tag{50}$$

If $K \neq K'$, then

$$\{z_{Kl}, H_{K'}\} = 0 \text{ for all } l. \tag{51}$$

In short, for $K \neq K'$, the evolution of z_{Kl} does not depend at all on $H_{K'}$. So if we replace our Hamiltonian H with

$$H_\alpha = \sum_{K<d^\alpha} H_K,$$

then we will not affect the evolution of any z_{kl} for $k < d^\alpha$.

Proof: The proof of (51) is an easy calculation using $\frac{\partial H_K}{\partial z_{K'l}} = 0$ for $K \neq K'$. Since

$$\frac{\partial z_{kl}}{\partial z_{k'l}} = \delta_{k,k'},$$

(51) follows.

Furthermore, writing $H = H_\alpha + (H - H_\alpha)$, we see that for $k < d^\alpha$,

$$\{z_{kl}, H\} = \{z_{kl}, H_\alpha\}.$$

Thus it is clear that the evolution of z_{kl} is the same whether we use H or H_α as our Hamiltonian.

■

We define $\mathcal{H}^s_\alpha$ as

$$\mathcal{H}^s_\alpha = \left\{ \{z_{kl}\}_{k,l=0}^\infty | \sum_{\substack{k<d^\alpha \\ l>0}} \omega_{kl}^{2s} \left|z_{kl}\right|^2 < \infty \right\}.$$

It is a consequence of Lemma 6 that if we choose $z \in \mathcal{H}^s_\alpha$, i.e. a z with all modes zero for $k > d^\alpha$, and define a flow of the form

$$\dot{z} = J\nabla H(z),$$

where H satisfies the assumptions of Lemma 6, then the modes of z for $k > d^\alpha$ will stay zero for all time.

Thus if we choose our initial conditions in the space $\mathcal{H}_{\alpha}^{1/2}$, the flow keeps z in this space for all time. Thus it makes sense to measure not $\|H\|_{\mathcal{A}} = \|\nabla H\|_{\mathcal{H}^{1/2},\mathcal{H}^{1/2}}$, but $\|H\|_{\mathcal{A}_\alpha} = \|\nabla H\|_{\mathcal{H}_{\alpha}^{1/2},\mathcal{H}_{\alpha}^{1/2}}$.

If we can determine that we have a small Type 1 term, in the sense that $\|H_1^n\|_{\mathcal{A}_\alpha}$ is small, and we choose initial conditions in $\mathcal{H}_{\alpha}^{1/2}$, then the formal arguments earlier in this section show that the flow nearly decouples, and the $l = 0$ modes are a good approximation for the whole system. Also, if $H \in \mathcal{A}_\alpha \subset \mathcal{QF}(\mathcal{H}_{\alpha}^{1/2})$, we can again think of ∇H as a linear map from $\mathcal{H}_{\alpha}^{1/2}$ to itself.

Finally, the functions H_1^0 and H_2^0 are bounded functions on $\mathcal{H}_{\alpha}^{1/2}$

Lemma 7. *There exist constants C_1 and C_2, (independent of d) such that*

$$\|H_1^0\|_{\mathcal{A}_\alpha} \leq C_1 d^{1+\alpha} \ ; \quad \|H_2^0\|_{\mathcal{A}_\alpha} \leq C_2 d^{1+\alpha} \tag{52}$$

Proof: If we repeat the estimates of (48) we arrive this time at

$$\|\nabla H_2^0\|_{\mathcal{A}_\alpha}^2 \leq \mathcal{B}_{2,\infty} \sum_{\substack{k<d^\alpha \\ l'>0}} \omega_{k0}^4 \omega_{kl'}^{-2} \omega_{kl'} \left|(z_{kl'} - \overline{z}_{kl'})\right|^2 , \tag{53}$$

But $\omega_{k0}^4 = \mu_k^2 \leq Ck^2 < Cd^{2\alpha}$ using the asymptotics of the eigenvalues of μ_k and the fact that $k < d^\alpha$. Similarly, $\omega_{kl'}^{-2} \leq \frac{Cd^2}{(l')^2}$, and we see that

$$\|\nabla H_2^0\|_{\mathcal{A}_\alpha}^2 \leq \mathcal{B}_{2,\infty} d^{2+2\alpha} \sum_{\substack{k<d^\alpha \\ l'>0}} {l'}^{-2} \omega_{kl'} \left|(z_{kl'} - \overline{z}_{kl'})\right|^2 \leq 4\mathcal{B}_{2,\infty} d^{2+2\alpha} \|z\|_{\mathcal{H}_{\alpha}^{1/2}}^2 . \tag{54}$$

The calculation of $\|\nabla H_1^0\|_{\mathcal{A}_\alpha}^2$ is very similar and we omit the details.

■

Remark 6. *By very similar estimates one can show that*

$$|H_1^0(z)| \leq C_1 d^{1+\alpha} \|z\|_{\mathcal{H}_{\alpha}^{1/2}}^2 \ , \quad |H_2^0(z)| \leq C_2 d^{1+\alpha} \|z\|_{\mathcal{H}_{\alpha}^{1/2}}^2 . \tag{55}$$

Once again, the estimates are so similar to those above that we omit the details to save space.

3 Estimates on the transformed Hamiltonian

In this section we prove a number of estimates which justify the formal arguments of the previous section concerning the size of terms in the transformed Hamiltonian. In particular, in Subsection 3, we estimate the action of arbitrary compositions of the Poisson brackets of our Hamiltonian with the functions χ_n defined in (45). We then use these estimates to bound the size of the interaction terms in the Hamiltonian after making n canonical transformations. In Subsection 3.2, we collect a number of technical estimates which we will need in Section 4 to prove our approximation theorems.

3.1 Bounds on the interaction terms in the Hamiltonian

The main goal of this section is to calculate the size of $\|H_1^n\|_{\mathcal{A}_\alpha}$, specifically its order in powers of d.

We define

$$M_n = \{I \in \mathbb{Z}^n \mid\ I_1 \geq I_2 \geq \cdots \geq I_n \geq 0\}\,, \quad \text{with} \quad M_0 = \{\emptyset\},$$

and

$$M = \bigcup_{n\geq 0} M_n, \quad \text{and} \quad M^+ = \bigcup_{n>0} M_n.$$

We also define

$$\mathcal{M}_n = \{I \in M \mid I_1 \leq n\}\,.$$

Note that from the definition of M, if $I \in \mathcal{M}_n$, then $I_k \leq n$ for all k.

For any $I \in M_n$, we define

$$|I| = n, \quad and \quad p(I) = \sum_{k=1}^{n} 2(I_k + 1).$$

Note that $I \in \mathcal{M}_n$ can be written as a sequence of some number of n's, which we denote by $J(I)_n$, (where $J(I)_n$ could be zero), followed by $J(I)_{n-1}$ $n-1$'s, and so on and so forth down to $J(I)_0$ zeros. Thus, associated to each $I \in \mathcal{M}_n$ is another sequence $J(I)$ of integers where $J(I)_m$ gives the number of times m appears in I. Note that we can reconstruct I from $J(I)$.

We refer to an $I \in M$ as a *multiindex*. Given a function H and $I \in M_n$, we define

$$L_I(H) = L_{I_1} \circ \cdots \circ L_{I_n}(H).$$

We note here several instances of notation we use. First, as defined above, $\emptyset \in M_0$ is a multiindex. For convenience, we define $p(\emptyset) = 0$. We define $L_\emptyset$ to be the identity, so that $L_\emptyset(H) = H$. Also, we will sometimes write multiindices out in multiplicative notation, with multiplication meaning concatenation.

In particular, $n^\alpha m^\beta$ will refer to the multiindex which is α copies of the number n followed by β copies of the number m. Furthermore, this then implies

$$L_{n^\alpha m^\beta} = L_n^\alpha \circ L_m^\beta.$$

For example, this means that $L_{n^\alpha} = L_n^\alpha$, in either case being the map L_n iterated α times. We note that $I = n^\alpha m^\beta$ could also be specified as the multiindex with $J(I)_n = \alpha$, $J(I)_m = \beta$, and $J(I)_k = 0$ if $k \neq n, m$.

Now, let's say that we choose $G = H_1^0$. Then $L_I(H_1^0)$ is of Type 1 if $|I|$ is even, and of Type $0+2$ if $|I|$ is odd. On the other hand, if we choose $G = H_2^0$, then the reverse is true: $L_I(H_2^0)$ is of Type $0+2$ if $|I|$ is even, and of Type 1 if $|I|$ is odd.

From now on, G is always either H_1^0 or H_2^0. We define $L_I(G)_0$, $L_I(G)_1$, and $L_I(G)_2$ as follows. First, let $|I|$ be even. Then, if we want $L_I(G)$ to be a term of Type $0+2$, we need to choose $G = H_2^0$. We then define $L_I(G)_0$ and $L_I(G)_2$ to be the Type 0 and Type 2 pieces of $L_I(H_2^0)$. If we want $L_I(G)$ to be Type 1, we have to choose $G = H_1^0$, and so we define $L_I(G)_1$ to be $L_I(H_1^0)$. Now, if $|I|$ is odd, we have to make the opposite choices at each step.

In short, to define $L_I(G)_q$ for $q = 0, 1$, or 2, we choose G to be either H_1^0 or H_2^0 (whichever one will give us the correct type), and then compute $L_I(G)$. We stress that once we have chosen a multiindex I and a Type (0, 1, or 2), then G is chosen for us. We will assume this in all that follows. Whenever we write $L_I(G)_q$, we know that G is determined by I and q.

Lemma 8. *Every term in $H_{0+2}^n + H_1^n$ is of the form $L_I(H_1^0)$ or $L_I(H_2^0)$, for some $I \in \mathcal{M}_{n-1}$. Conversely, there exist coefficient functions $\Theta_1(I)$ and $\Theta_{0+2(I)}$ with the property that*

$$\max\left(|\Theta_1(I)|, |\Theta_{0+2}(I)|\right) \leq \frac{3^{p(I)}}{\prod_{j=0}^{n-1} J(I)_j!},$$

such that

$$H_1^n = \sum_{I \in \mathcal{M}_{n-1}} \Theta_1(I) L_I(G)_1,$$
$$H_{0+2}^n = \sum_{I \in \mathcal{M}_{n-1}} \Theta_{0+2}(I) L_I(G)_{0+2}.$$

We emphasize that G on the right hand side of these formulas is chosen to be either H_1^0 or H_2^0, whichever gives a term of the correct type.

Remark 7. *In [13] an explicit formula for the coefficients $\Theta_j(I)$ will be derived, but for the moment we need only to know that they exist.*

Proof: We prove the lemma inductively. Equations (42) and (41) show that it holds for $n = 1$. We assume that it holds up to some positive integer n and prove it holds for $n+1$. We provide the details for H_1^{n+1} – the proof is the same for H_2^{n+1}. Recalling (47) we have

$$H_1^{n+1} = \sum_{i \text{ odd}} \frac{L_n^i(H_{0+2}^n)}{i!} + \sum_{\substack{i \text{ even} \\ i>0}} \frac{L_n^i(H_{1,\text{low}}^n)}{\xi(i)} + \sum_{i \text{ even}} \frac{L_n^i(\widehat{H_1^n})}{i!}.$$

Recalling that $H_{1,\text{low}}^n$ is the sum of the lowest order terms in d of H_1^n, there is a subset $\mathcal{I} \subsetneq \mathcal{M}_{n-1}$ so that

$$H_{1,\text{low}}^n = \sum_{I \in \mathcal{I}} \Theta_1(I) L_I(G)_1, \quad \text{and} \quad \widehat{H_1^n} = \sum_{I \in \mathcal{M}_{n-1} \setminus \mathcal{I}} \Theta_1(I) L_I(G)_1.$$

Putting this together, we can write

$$H_1^{n+1} = \sum_{i\text{ odd}} \frac{L_n^i\left(\sum_{I\in\mathcal{M}_{n-1}} \Theta_{0+2}(I)L_I(G)_{0+2}\right)}{i!} + \sum_{\substack{i\text{ even}\\ i>0}} \frac{L_n^i\left(\sum_{I\in\mathcal{I}} \Theta_1(I)L_I(G)_1\right)}{\xi(i)} + \sum_{i\text{ even}} \frac{L_n^i\left(\sum_{I\in\mathcal{M}_{n-1}\setminus\mathcal{I}} \Theta_1(I)L_I(G)_1\right)}{i!}.$$

Let us consider the first sum.

$$\sum_{\substack{i\text{ odd}\\ I\in\mathcal{M}_{n-1}}} \frac{\Theta_{0+2}(I)}{i!} L_n^i(L_I(G)_{0+2}).$$

Since i is odd, $L_n^i(L_I(G)_{0+2})$ is of Type 1, so we can write it as $L_{\tilde{I}}(G)_1$, where $\tilde{I} = n^i I$. Clearly, $\tilde{I} \in \mathcal{M}_n$. Also, we can define $\Theta_1(\tilde{I}) = \Theta_{0+2}(I)/i!$. Clearly, if $|\Theta_{0+2}(I)| < 3^{p(I)}/(\prod_j J(I)_j!)$, then $\left|\Theta_1(\tilde{I})\right| < 3^{p(\tilde{I})}/(\prod_j J(I)_j!)$ also. In passing, we note that if we consider $\tilde{I} = n^0 I$, for some $I \in \mathcal{M}_{n-1}$, – i.e. if $\tilde{I} = I$, then $\Theta_1(\tilde{I}) = \Theta_{0+2}(I)$, so the functions Θ_j do not depend on n, the number of iterations we have performed.

We will get similar results if we consider the other two sums. Since we are applying some power of L_n, the multiindices that will arise are all in $\mathcal{M}_n$.

■

The next two lemmas give rigorous estimates for the size of the expressions that appear in H^n. Their proofs are long, and somewhat technical, and hence are deferred until Appendix A.

Lemma 9. *There exists a constant C_I and functions $Q_I(l), f_I(l), g_I(l)$ (depending on I) such that for all $I \in M^+$, we can write*

$$L_I(G)_{0,k00} = C_I d^{p(I)} \omega_{k0}^{p(I)+1}, \tag{56}$$

$$L_I(G)_{1,kl0} = Q_I(l) d^{p(I)} \omega_{k0}^{p(I)+1} \frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}}, \tag{57}$$

$$L_I(G)_{2,kl_1l_2} = d^{p(I)} \omega_{k0}^{p(I)+2} \left(\frac{f_I(l_1)g_I(l_2)}{\omega_{kl_1}^{1/2}\omega_{kl_2}^{1/2}} + \frac{f_I(l_2)g_I(l_1)}{\omega_{kl_1}^{1/2}\omega_{kl_2}^{1/2}} \right), \tag{58}$$

where Q_I, f_I, and g_I are all in ℓ^2 (with respect to l). Furthermore, for every n,

$$H_{1,\text{low}}^n = \sum_{\substack{I\in\mathcal{M}_{n-1}\\ p(I)=2n}} \Theta_1(I) L_I(G)_1.$$

Remark 8. *This lemma is a direct calculation using Lemma 4 and (31) – details of the proof are in Appendix A. Also we note that it is clear that H_1^n will have no terms which are of order d^m for $m < 2n$. This is because in the construction of the canonical transformation ϕ_{n-1}, we eliminated $H_{1,low}^{n-1}$, and by Lemma 9 this means that we eliminate all Type 1 terms of order $d^{2(n-1)}$ with the nth coordinate change.*

Now that we have the form for $L_I(G)_1$, we want to calculate $\|L_I(G)_1\|_{\mathcal{A}_\alpha}$. We use the following

Lemma 10. *We have the following estimates for $I \in M^+$ and d sufficiently small:*

$$\|L_I(G)_0\|_{\mathcal{A}_\alpha} \leq 8C_I d^{p(I)(1+\alpha)+\alpha},$$
$$\|L_I(G)_1\|_{\mathcal{A}_\alpha} \leq 4\,\|Q_I\|_{\ell^2}\, d^{p(I)(1+\alpha)+2\alpha},$$
$$\|L_I(G)_2\|_{\mathcal{A}_\alpha} \leq 4\,\|f_I\|_{\ell^2}\,\|g_I\|_{\ell^2}\, d^{p(I)(1+\alpha)+2\alpha}.$$

See Appendix A for a proof.

Note that this lemma justifies the formal arguments above. Recall that we chose $H^n_{1,\mathrm{low}}$, at each step, to be the terms of Type 1 of lowest order in d. Those arguments were formal, since we didn't really know the size of these operators exactly. This lemmas specifically tells us that, when we chose the term with the formally lowest power of d, we were choosing a term which actually was smallest in d in the sense of the operator norm on $\mathcal{A}_\alpha$.

Combining Lemmas 8, 9 and 10 gives us

Lemma 11. *For all n and for sufficiently small d, there exists $C_n > 0$ such that*

$$\left\|H^n_{1,\mathrm{low}}\right\|_{\mathcal{A}_\alpha} \leq C_n d^{2n(1+\alpha)+2\alpha}.$$

Proof: We know from Lemma 9 that $H^n_{1,\mathrm{low}} = \sum_{\substack{I\in\mathcal{M}_{n-1}\\ p(I)=2n}} \Theta_1(I) L_I(G)_1$. Using Lemma 10, we have that

$$\begin{aligned}\left\|H^n_{1,\mathrm{low}}\right\|_{\mathcal{A}_\alpha} &\leq 4d^{(2n)(1+\alpha)+2\alpha} \sum_{\substack{I\in\mathcal{M}_{n-1}\\ p(I)=2n}} |\Theta_1(I)|\,\|Q_I\|_{\ell^2} \\ &\leq 4C_{q,n} d^{(2n)(1+\alpha)+2\alpha} \sum_{\substack{I\in\mathcal{M}_{n-1}\\ p(I)=2n}} \frac{3^{p(I)}}{\prod J(I)_j!}, \end{aligned} \tag{59}$$

where the last inequality used Lemma 8 to bound the factor of Θ_1 and set $C_{q,n} = \max_{\substack{I\in\mathcal{M}_{n-1}\\ p(I)=2n}} \|Q_I\|_{\ell^2}$. The sum over I can be bounded by $3^{p(I)}(\sum_{\ell=0}^{\infty} \frac{1}{\ell!})^{(2n)} \leq (3e)^{(2n)}$ and the estimate of the Lemma follows.

■

3.2 Some technical details

In this subsection, we collect some estimates related to those of the previous subsection that we will need to prove Theorem 3 in the next section. In particular, we calculate the size of $\|\chi_n\|_{\mathcal{A}_\alpha}$ for any n and we show that our original Hamiltonian, H^0, and any subsequent Hamiltonian, H^n, define equivalent norms on $\mathcal{H}_\alpha^{1/2}$. Finally, we prove some lemmas about how a Hamiltonian flow affects the size of various norms.

Lemma 12. *If $k \leq d^\alpha$, then for d sufficiently small,*

$$|\omega_{kl} - \omega_{k0}| \geq \frac{|\omega_{kl}|}{2}, \text{ for } l > 0.$$

Proof: The proof is an immediate consequence of the asymptotics of the eigenvalues of (18) and (19).

■

Lemma 13. *For any n and for d sufficiently small,*

$$\|\chi_n\|_{\mathcal{A}_\alpha} \leq C d^{2n(1+\alpha)+1+2\alpha}.$$

Proof: By (45), we know that

$$\chi_{n,kl} = \frac{-i2H^n_{1,\text{low},kl}}{\omega_{k0} + \omega_{kl}}, \qquad \chi_{n,k\bar{l}} = \frac{i2H^n_{1,\text{low},kl}}{\omega_{k0} - \omega_{kl}}.$$

Using Lemma 12, we see that we have the two estimates

$$|\chi_{n,kl}| \leq CdH^n_{1,\text{low},kl}, \quad \text{and} \quad \left|\chi_{n,k\bar{l}}\right| \leq CdH^n_{1,\text{low},kl}.$$

From this it is clear that $\|\chi_n\|_{\mathcal{A}_\alpha} \leq Cd \left\|H^n_{1,\text{low}}\right\|_{\mathcal{A}_\alpha}$, which combined with Lemma 11 completes the proof.

■

Remark 9. *Recall that $L_n(G) = \{\chi_n, G\}$. Thus, the operator norm of L_n, considered as a linear operator on $\mathcal{A}_\alpha$, is bounded by $\|\chi_n\|_{\mathcal{A}_\alpha}$.*

Corollary 14. *Under the hypotheses of Lemma 13, there exists $C_n > 0$ such that the canonical transformation ϕ_n defined as the time one map of the Hamiltonian system*

$$\dot{z} = \{\chi_n, z\} \tag{60}$$

is a bounded linear map on $\mathcal{H}_\alpha^{1/2}$ satisfying

$$\begin{aligned} \|\phi_n\|_{\mathcal{H}_\alpha^{1/2},\mathcal{H}_\alpha^{1/2}} &\leq \exp(C_n d^{2n(1+\alpha)+1+2\alpha}) \\ \|\phi_n - \mathbf{1}\|_{\mathcal{H}_\alpha^{1/2},\mathcal{H}_\alpha^{1/2}} &\leq C_n d^{2n(1+\alpha)+1+2\alpha} \end{aligned}$$

Proof: These estimates follow immediately by applying Grönwall's inequality to the (linear) system of ordinary differential equations (60).

■

Remark 10. *We can also consider ϕ_n as a bounded, linear, canonical transformation on $\mathcal{H}^{1/2}$ by extending it as the identity on $\mathcal{H}^{1/2}\backslash\mathcal{H}^{1/2}_\alpha$.*

By combining Lemma 13 with (38) we obtain a representation of the Hamiltonian after n canonical transformations as a *finite* sum of terms of the form $L_I(H^0)$, plus a remainder that can be made arbitrarily small. More precisely, we have:

Proposition 15. *Fix $N > 0$ and $n \geq 1$. There exist positive constants d_0 and $C_{n,N}$, a finite subset $\mathcal{I}_{n,N} \subset \mathcal{M}_{n-1}$, and functions $\tilde{\Theta}(I)$ satisfying*

$$|\tilde{\Theta}(I)| \leq 1/(\prod_j J(I)_j!) \ , \tag{61}$$

such that if $0 < d < d_0$,

$$H^n(z) = H^0 \circ \phi^0 \circ \cdots \circ \phi^n(z) = \sum_{I \in \mathcal{I}_{n,N}} \tilde{\Theta}(I) L_I(H^0) + \mathcal{R}_{n,N}(z) \ , \tag{62}$$

where

$$\|\mathcal{R}_{n,N}\|_{\mathcal{A}_\alpha} \leq C_{n,N} d^{(1+2\alpha)N} \ . \tag{63}$$

Proof: The proof is inductive. For $n = 1$, (38) implies that

$$H^1(z) = H^0 \circ \phi^0(z) = \sum_{j=0}^{N-1} \frac{1}{j!} L_0^j(H^0) + \int_0^1 \cdots \int_0^{t_{N-1}} L_0^N(H^0) \circ (\phi_0)^{t_N} dt_N \dots dt_1 \ . \tag{64}$$

From Lemma 13 and Corollary 14, there exists $C > 0$ such that we have

$$\|L_0^N(H^0) \circ (\phi_0)^t\|_{\mathcal{A}_\alpha} \leq C d^{(1+2\alpha)(N-1)} \|L_0(H^0)\|_{\mathcal{A}_\alpha} \tag{65}$$

But $L_0(H^0) = \{\chi_0, H^0_{\text{diag}}\} + \{\chi_0, H^0_1\} + \{\chi_0, H^0_2\} = -H^0_1 + \{\chi_0, H^0_1\} + \{\chi_0, H^0_2\}$, and from Lemma 7 we know that $\|H^0_1\|_{\mathcal{A}_\alpha} \leq C d^{1+\alpha}$ while χ_0 is bounded with the aid of Lemma 13. Thus, the integral term in (64) can be bounded by $C d^{(1+2\alpha)N} \int_0^1 \cdots \int_0^{t_{N-1}} dt_N \dots dt_1 = \frac{C}{(N-1)!} d^{(1+2\alpha)N}$, and the lemma follows for $n = 1$ if we take $\mathcal{I}_{1,N} = \{\emptyset, 0^1, 0^2, \dots, 0^{N-1}\}$, and $\tilde{\Theta}(I = 0^j) = 1/j!$.

We now assume that the lemma holds for all $n = 1, \dots, M-1$ and prove that it holds for $n = M$.

$$\begin{aligned} H^M(z) &= H^{M-1} \circ \phi_{M-1}(z) = \sum_{j=0}^{N-1} \frac{1}{j!} L^j_{M-1}(H^{M-1}) \\ &\quad + \int_0^1 \cdots \int_0^{t_{N-1}} L^N_{M-1}(H^{M-1}) \circ (\phi_{M-1})^{t_N} dt_N \dots dt_1 \ . \end{aligned} \tag{66}$$

by (38). By the induction hypothesis we have

$$L_{M-1}^{j}(H^{M-1}) = \sum_{I\in\mathcal{I}_{M-1,N}} \frac{1}{j!}\tilde{\Theta}(I)L_{M-1}^{j}L_I(H^0) + L_{M-1}^{j}(\mathcal{R}_{M-1,N}) \tag{67}$$

From Lemma 13 and the inductive estimate on $\mathcal{R}_{M-1,N}$, we have

$$\|L_{M-1}^{j}(\mathcal{R}_{M-1,N})\|_{\mathcal{A}_\alpha} \leq Cd^{j[2(M-1)(1+\alpha)+1+2\alpha]}d^{(1+2\alpha)N} . \tag{68}$$

On the other hand $L_{M-1}^{j}L_I(H^0) = L_{\tilde{I}}(H^0)$, where $\tilde{I} = M-1^j I \in \mathcal{M}_{M-1}$, so we can define $\mathcal{I}_{M,N}$ to be those $\tilde{I} \in \mathcal{M}_{M-1}$ for which $\tilde{I} = (M-1)^j I$ for some $I \in \mathcal{I}_{M-1,N}$, $j = 0, 1, \ldots, N-1$ – in particular, this is a finite set. Defining $\tilde{\Theta}(\tilde{I}) = \frac{1}{j!}\tilde{\Theta}(I)$ we see that $\tilde{\Theta}(\tilde{I})$ satisfies (61).

By estimates very similar to those above, we can bound the integral term on the right hand side of (66) by $\int_0^1 \ldots \int_0^{t_{N-1}} Cd^{[2(M-1)(1+\alpha)+1+2\alpha](N-1)}dt_N \ldots dt_1 \leq \frac{C}{(N-1)!}d^{[2(M-1)(1+\alpha)+1+2\alpha](N-1)}$. Thus, if we define $\mathcal{R}_{M,N} = \sum_{j=0}^{N-1} L_{M-1}^{j}(\mathcal{R}_{M-1,N}) + \int_0^1 \ldots \int_0^{t_{N-1}} L_{M-1}^{N}(H^{M-1}) \circ (\phi_{M-1})^{t_N} dt_N \ldots dt_1$, we see that there exists a constant $C_{M,N}$ such that $\|\mathcal{R}_{M,N}\|_{\mathcal{A}_\alpha} \leq C_{M,N}d^{(1+2\alpha)N}$, and the lemma follows.

■

Remark 11. *If we compare the representation of $H^n(z)$ derived in the proof of Proposition 15 with that derived in Subsection 2.2 and Lemma 8, we see that in Proposition 15 we have just restricted the expansion to term $L_I(H^0)$ with $I \in \mathcal{M}_{n-1}$,* and $J(I)_l \leq N$ *for all l. Thus, there exists* a finite *subset $\tilde{\mathcal{I}}_{n,N} \subset \mathcal{M}_{n-1}$ such that*

$$\sum_{I\in\mathcal{I}_{n,N}} \tilde{\Theta}(I)L_I(H^0) = H^0_{\text{diag}} + \sum_{\tilde{I}\in\tilde{\mathcal{I}}_{n,N}} \Theta_1(I)L_I(G)_1 + \sum_{\tilde{I}\in\tilde{\mathcal{I}}_{n,N}} \Theta_{0+2}(I)L_I(G)_{0+2} \tag{69}$$

(Note: The H^0_{diag} term comes from the term on the left hand side with $I = \emptyset$.)

We now derive from this representation of the transformed Hamiltonian several corollaries concerning the size of the transformed Hamiltonian.

Lemma 16. *For all $n \geq 1$, there exists $d_0 > 0$ and $C_n > 0$ such that for $0 < d < d_0$,*

$$\|H_1^n\|_{\mathcal{A}_\alpha} = \|\nabla H_1^n\|_{\mathcal{H}_\alpha^{1/2},\mathcal{H}_\alpha^{1/2}} \leq C_n d^{2n(1+\alpha)+2\alpha} \tag{70}$$

$$\|H_0^n\|_{\mathcal{A}_\alpha} = \|\nabla H_0^n\|_{\mathcal{H}_\alpha^{1/2},\mathcal{H}_\alpha^{1/2}} \leq C_n d^{2+3\alpha} \tag{71}$$

$$\|H_2^n\|_{\mathcal{A}_\alpha} = \|\nabla H_0^n\|_{\mathcal{H}_\alpha^{1/2},\mathcal{H}_\alpha^{1/2}} \leq C_n d^{1+\alpha} \tag{72}$$

Proof: Choosing N such that $(1+2\alpha)N > 2n(1+\alpha)+2\alpha$, we can apply Proposition 15 to write

$$H^n(z) = \sum_{I\in\mathcal{I}_{n,N}} \tilde{\Theta}(I)L_I(H^0) + R_{n,N}(z) \tag{73}$$

where $\|R_{n,N}\|_{\mathcal{A}_\alpha} \leq C_n d^{2n(1+\alpha)+2\alpha}$. By Remark 11, there exists a finite subset $\tilde{\mathcal{I}}_{n,N} \subset \mathcal{M}_{n-1}$ such that

$$\sum_{I\in\mathcal{I}_{n,N}} \tilde{\Theta}(I)L_I(H^0) = H^0_{\text{diag}} + \sum_{\tilde{I}\in\tilde{\mathcal{I}}_{n,N}} \Theta_1(\tilde{I})L_I(G)_1 + \sum_{\tilde{I}\in\tilde{\mathcal{I}}_{n,N}} \Theta_{0+2}(\tilde{I})L_I(G)_{0+2} \ . \tag{74}$$

Thus,

$$H_1^n(z) = \sum_{\tilde{I}\in\tilde{\mathcal{I}}_{n,N}} \Theta_1(I)L_I(G)_1 + \mathcal{R}^{(1)}_{n,N}(z) \tag{75}$$

where $\mathcal{R}^{(1)}_{n,N}$ is the Type 1 part of $\mathcal{R}_{n,N}$ and hence satisfies the estimate (63). Since $\tilde{\mathcal{I}}_{n,N}$ is finite, there exists some integer $P_{\max}$ such that $p(\tilde{I}) \leq P_{max}$ for all $\tilde{I} \in \tilde{\mathcal{I}}_{n,N}$. On the other hand, the canonical transformations $\phi^0 \ldots, \phi^{n-1}$ were constructed so that all Type 1 terms in the Hamiltonian with $p(\tilde{I}) < 2n$ were canceled. Hence we can write

$$H_1^n(z) = \sum_{\substack{p=2n \\ p \ even}}^{P_{max}} \sum_{\substack{\tilde{I}\in\tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I})=p}} \Theta_1(I)L_I(G)_1 + \mathcal{R}^{(1)}_{n,N}(z) \tag{76}$$

Just as in the proof of Lemma 11, we can bound

$$\| \sum_{\substack{\tilde{I}\in\tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I})=p}} \Theta_1(I)L_I(G)_1\|_{\mathcal{A}_\alpha} \leq C_p d^{p(1+\alpha)+2\alpha} \ , \tag{77}$$

since there are only finitely many terms in the sum. The sum over p is also finite and hence easily bounded and $\|\mathcal{R}^{(1)}_{n,N}(z)\|_{\mathcal{A}_\alpha}$ is bounded by (63), yielding the estimate of (70).

To estimate $\|H_0^n(z)\|_{\mathcal{A}_\alpha}$ we proceed as above but since the construction of the canonical transformations does not eliminate the low order (in d) terms in H_0^n, we arrive at an expression for the Type 0 part of the Hamiltonian of the form

$$H_0^n(z) = \sum_{\substack{p=2 \\ p \ even}}^{P_{max}} \sum_{\substack{\tilde{I}\in\tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I})=p}} \Theta_{0+2}(I)L_I(G)_0 + \mathcal{R}^{(0)}_{n,N}(z) \tag{78}$$

where $\mathcal{R}^{(0)}_{n,N}$ is the Type 0 part of $\mathcal{R}_{n,N}$. With the aid of Lemma 10 we can again bound

$$\| \sum_{\substack{\tilde{I}\in\tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I})=p}} \Theta_1(I)L_I(G)_0\|_{\mathcal{A}_\alpha} \leq C_p d^{p(1+\alpha)+\alpha} \ . \tag{79}$$

and then the geometric series in p, combined with the estimate of (63) implies that (71) holds.

The estimate (72) follows in exactly the same fashion except that since the original Hamiltonian contains a term H_2^0, but no H_0^0 term, that additional term must be added to the right hand side of the analogue of (78), and it is estimated by Lemma 7.

$\blacksquare$

Remark 12. *If we proceed as in the previous lemma, using the expressions in Lemma 9 rather than the bounds in Lemma 10 we can bound the values of the Hamiltonian on $\mathcal{H}_\alpha^{1/2}$, as*

$$|H_1^n(z)| \leq Cd^{(2n)(1+\alpha)}\|z\|^2_{\mathcal{H}_\alpha^{1/2}} \tag{80}$$

$$|H_0^n(z)| \leq Cd^{2+\alpha}\|z\|^2_{\mathcal{H}_\alpha^{1/2}} \tag{81}$$

$$|H_2^n(z)| \leq Cd^{1+\alpha}\|z\|^2_{\mathcal{H}_\alpha^{1/2}} \tag{82}$$

just as we did in the proof of Lemma 16.

Note that if we define

$$\mathcal{O}_R(\mathcal{H}_\alpha^s) = \left\{ z \in \mathcal{H}_\alpha^s | \ \|z\|_{\mathcal{H}_\alpha^s} \leq R \right\}.$$

we have:

Lemma 17. *Let $B \subseteq \mathcal{H}_\alpha^{1/2}$ be bounded, i.e. there is an R with $B \subset \mathcal{O}_R(\mathcal{H}_\alpha^{1/2})$. Then there is a constant C such that for sufficiently small d,*

$$\left| H^n(z) - H^0(z) \right| \leq Cd^{1+\alpha}, \quad \textit{for all } z \in B.$$

Proof: This follows immediately from (80)–(82) and Remark 6.

$\blacksquare$

Lemma 18. *If we define H_0^n as the Type 0 piece of H^n, then $\langle z, z \rangle_{H^0_{\mathrm{diag}}+H_0^n} = H^0_{\mathrm{diag}}(z) + H_0^n(z)$, defines an inner product on $\mathcal{H}^{1/2}$ with the property that for any bounded set $B \in \mathcal{H}_\alpha^{1/2}$ there is a C with*

$$\frac{1}{C}\|z\|_{\mathcal{H}_\alpha^{1/2}} \leq \|z\|_{H^0_{\mathrm{diag}}+H_0^n} \leq C\|z\|_{\mathcal{H}_\alpha^{1/2}},$$

for all $z \in B$.

Proof: A positive-definite symmetric linear operator $A\colon \mathcal{H}^{1/2} \to \mathcal{H}^{1/2}$ defines an inner product via

$$\langle x, y \rangle_A = \langle Ax, y \rangle = \left\langle A^{1/2}x, A^{1/2}y \right\rangle, \tag{83}$$

where $A^{1/2}$ is the positive-definite square root of A.

Since the gradient of $H = H^0_{\text{diag}} + H^n_0$ is positive-definite and symmetric, we can define the inner product

$$H(z) = \langle \nabla H(z), z \rangle \,.$$

Thus, to complete the proof of the lemma, we need only show that

$$H^0_{\text{diag}}(z) + H^n_0(z) \sim H^0_{\text{diag}}(z),$$

for all $z \in B \subset \mathcal{H}^{1/2}_\alpha$, by which we mean that there exists a constant C with

$$\frac{1}{C}\left(H^0_{\text{diag}}(z) + H^n_0(z)\right) \leq H^0_{\text{diag}}(z) \leq C\left(H^0_{\text{diag}}(z) + H^n_0(z)\right),$$

for all $z \in B$.

Now, using Lemma 17, it is easy to see that for $z \in B$,

$$H^n(z) \sim H^0(z).$$

Inequalities (80) and (82) imply that there is a choice of d_0 such that for $d < d_0$, we have

$$|H^n_1(z) + H^n_2(z)| \leq \frac{1}{2}\|z\|^2_{\mathcal{H}^{1/2}_\alpha} = \frac{1}{2}H^0_{\text{diag}}(z) \text{ if } z \in \mathcal{H}^{1/2}_\alpha \,. \tag{84}$$

(Roughly, the larger n is, the easier it is to satisfy (84), and this is why the estimate can be made uniformly in n.) Using this for $n = 0$ gives us that since $H^0_{\text{diag}} = H^0 - H^0_1 - H^0_2$, then

$$\frac{1}{2}\left|H^0(z)\right| \leq \left|H^0_{\text{diag}}(z)\right| \leq \frac{3}{2}\left|H^0(z)\right|,$$

or that $H^0 \sim H^0_{\text{diag}}$. The same argument gives us that $H^n \sim H^0_{\text{diag}} + H^n_0$, and we are done.

■

The utility of the norms defined in the previous lemma comes from the fact that the Hamiltonian flow with Hamiltonian H preserves the value of H and hence the norm defined by H. More precisely we have:

Corollary 19. *Assume that we have a Hamiltonian H which generates an inner product, $\langle \cdot, \cdot \rangle_H$, as in the previous lemma. Then if we define a flow by*

$$\dot{z} = J\nabla H(z),$$

then this flow preserves the norm $\|\cdot\|_H$, i.e. $\frac{d}{dt}\|z\|_H = 0$, where we define $\|z\|^2_H = \langle z, z \rangle_H$.

As a consequence of this estimate, and the fact that the norm defined by H^0_{diag} is the $\mathcal{H}^{1/2}_\alpha$ norm, we have

Corollary 20. *We define, on* $\mathcal{H}_\alpha^{1/2}$*, the flow* $\dot{z} = J\nabla H(z)$*, with* $H = H^0_{\text{diag}} + H_>$*, where* $|H_>(z)| \leq \frac{1}{2}\left|H^0_{\text{diag}}(z)\right|$*, for all* $z \in \mathcal{H}_\alpha^{1/2}$*. Then*

$$\|z(t)\|^2_{\mathcal{H}^{1/2}} \leq 3\,\|z(0)\|^2_{\mathcal{H}^{1/2}}\,, \quad \text{for all } t.$$

Proof: Note that

$$\frac{1}{2}\left|H^0_{\text{diag}}(z(t))\right| \leq |H(z(t))| \leq \frac{3}{2}\left|H^0_{\text{diag}}(z(0))\right|.$$

Since $H(z(t)) = H(z(0))$ for all t, and $H^0_{\text{diag}}(z) = \|z\|^2_{\mathcal{H}_\alpha^{1/2}}$, the estimate follows.

■

Finally, we estimate the way in which modes with a fixed value of k evolve.

Lemma 21. *Let* $z \in \mathcal{H}^s$*, and define* $P^K(z)$ *as*

$$\left(P^K(z)\right)_{kl} = \delta_{k,K} z_{kl}.$$

We define $\|z\|_{\mathcal{H}^{1/2},K} := \left\|P^K(z)\right\|_{\mathcal{H}^{1/2}}$*. Then if we have the flow* $\dot{z} = J\nabla H^0(z)$*, there is a* C *(independent of* K*), such that*

$$\|z(t)\|_{\mathcal{H}^{1/2},K} \leq C\,\|z(0)\|_{\mathcal{H}^{1/2},K}\,.$$

Proof: The proof is an easy exercise using Lemma 6 which shows that the flow defined by H does not couple modes with different values of k and the sorts of energy estimates used in Corollary 20.

■

4 The Approximation Theorems

In this section we first prove (in Subsection 4.1) that the flow of the original Hamiltonian, H^0, can be approximated in $\mathcal{H}_\alpha^{1/2}$ for a very long time by the flow given by $H^0_{\text{diag}} + H^n_0$. In Subsection 4.2 we then revert to the continuous variables and show how this discrete approximation theorem implies the existence and approximation properties of the reduced equations defined in Theorem 1.

4.1 The approximation theorem in frequency space

In this section, we want to state and prove a theorem in the discrete variables which will imply Theorem 1. (We will establish this implication in Subsection 4.2.)

To state our theorem, we need to define one more Hilbert space. Recall that our standard Hilbert space $\mathcal{H}^{1/2}$ corresponds (roughly) to functions which are once differentiable in x and y. We want to consider functions which may have different degrees of differentiability in the x and y directions. So we define

$$\|z\|_{\mathcal{H}^{s,t}} = \sum_{k} \omega_{k0}^{2s} |z_{k0}|^2 + \sum_{\substack{k \\ l>0}} \omega_{kl}^{2t} |z_{kl}|^2 ,$$

which corresponds (if $s > t$) to functions with s derivatives in the long (x) direction, and t derivatives in the thin (y) direction.

Theorem 3. *Given H^0 as defined in (25), we make the coordinate changes defined in the iterative scheme of Section 2.1. Define $N(n) = 2n+1+(2n+2)\alpha$. We choose $z(0) \in \mathcal{H}^{s,r}$, with $s > 1, r > 1/2$, and define the flows*

$$\dot{z} = J\nabla H^0(z),$$
$$\dot{\gamma} = J\nabla(H^0_{\text{diag}} + H^n_0)(\gamma),$$

with $\gamma(0)$ chosen to be the projection of $z^{\mathbf{r}}(0)$ into $\mathcal{H}^{1/2}_\alpha$. Then, for any $\epsilon > 0$, there exists $C_n = C_n\left(a, b, C, \|z(0)\|_{\mathcal{H}^{s,t}}\right)$ and $d_0 = d_0(a, b, C, s)$ such that for $d < d_0$, we have the estimate

$$\|z(t) - \gamma(t)\|^2_{\mathcal{H}^{1/2}} \leq C_n d^{N(n)} t + 3\epsilon,$$

for $t \leq d^{-N(n)}$.

The main idea of this proof is as follows. First, we want to show that the $l = 0$ modes dominate the evolution of the equation. Most of the work we have done in previous sections has been to show that we can change variables so that the coupling between $l = 0$ and $l > 0$ modes is small. This means that energy that starts in a mode with $l > 0$ would take a long time to transfer to modes with $l = 0$. We need only require that the initial energy in modes with $l > 0$ is small, and we would be done.

The next step will be to show that the evolution of the modes with $k < d^\alpha, l = 0$ dominate all $l = 0$ modes. It is reasonable that this will work, since we are just throwing out a different set of high-frequency modes.

Proof: Choose $\epsilon > 0$. We define $z^{\mathbf{r}}$ and $z^{>}$ by

$$z^{\mathbf{r}}_{kl} = \delta_{l,0} z_{kl}, \ , \quad z^{>}_{kl} = z - z^{\mathbf{r}}.$$

We want to make both $\|z(t)\|_{\mathcal{H}^{1/2}, k \geq d^\alpha} < \epsilon$ and $\|z^{>}(t)\|_{\mathcal{H}^{1/2}} < \epsilon$. We know from Lemma 20 that if $\|z(0)\|_{\mathcal{H}^{s,t}} = R$, then $z(t) \in \mathcal{O}_{3R}(\mathcal{H}^{1/2}_\alpha)$ for all t. We assume implicitly in all that follows that z is in this bounded set.

According to Lemma 21, we know that there is a κ independent of K with

$$\|z(t)\|_{\mathcal{H}^{1/2},K} \leq \kappa \|z(0)\|_{\mathcal{H}^{1/2},K} . \tag{85}$$

Since this κ is independent of K, if we define

$$\|z\|_{\mathcal{H}^{1/2},K \geq d^\alpha} = \sum_{K \geq d^\alpha} \|z\|_{\mathcal{H}^{1/2},K} ,$$

then

$$\|z(t)\|_{\mathcal{H}^{1/2},K\geq d^\alpha} \leq \kappa \|z(0)\|_{\mathcal{H}^{1/2},K\geq d^\alpha}.$$

So we simply need to make sure that

$$\|z(0)\|_{\mathcal{H}^{1/2},K\geq d^\alpha} < \frac{\epsilon}{\kappa}.$$

Recall that we've chosen $z \in \mathcal{H}^{s,r}$, with $s > 1$, $r > 1/2$. Thus we know that

$$\sum_k \omega_{k0}^{2s} |z_{k0}|^2 \leq \|z\|_{\mathcal{H}^{s,r}}.$$

We note that the asymptotics of the eigenvalues in (19) imply that $\omega_{k0}^{2s} = \mu_k^s = M(b,C)k^s$ where $M(b,C)$ is a constant which depends only on $b(\eta), C(x)$. Thus

$$|z_{k0}| \leq \frac{\|z\|_{\mathcal{H}^{s,r}}}{M|k|^s},$$

and

$$\begin{aligned}\|z\|^2_{\mathcal{H}^{1/2},k>d^\alpha} &= \sum_{k>d^\alpha} \omega_{k0} |z_{k0}|^2 \leq \|z\|^2_{\mathcal{H}^{s,r}} \sup_{k>d^\alpha} \omega_{k0}^{1-2s} \\ &\leq M(b,C) \|z\|^2_{\mathcal{H}^{s,r}} d^{\frac{\alpha}{2}(1-2s)}.\end{aligned}$$

In summary,

$$\|z(0)\|^2_{\mathcal{H}^{1/2},k>d^\alpha} \leq \kappa d^{\frac{\alpha}{2}(1-2s)},$$

where κ is a constant which depends only upon $b(\eta)$, $C(x)$, and $\|z(0)\|_{s,r}$. This and (85) give

$$\|z(t)\|^2_{\mathcal{H}^{1/2},k>d^\alpha} = \tilde{\kappa} d^{\frac{\alpha}{2}(1-2s)}, \tag{86}$$

where κ is a constant which depends only upon $b(\eta)$, $C(x)$, and $\|z(0)\|_{s,r}$.

Thus, for $s > 1$, there is a d_0 such that for all $d < d_0$, we can make (86) smaller than ϵ. This d_0 depends on the coefficient functions $a(\eta)$, $b(\eta)$, and $C(x)$ defined in Equation (1), and also on s, since for larger s, the right hand side of (86) will shrink more quickly as d shrinks. Thus we note that if we choose our initial condition more smooth, the thickness of our domain can be larger.

Let us consider $z^>(t)$. We have chosen $z \in \mathcal{H}^{s,r}$ for some $s > 1, r > 1/2$. Now, according to (23), we have

$$\omega_{kl} > \frac{\kappa(a)}{d} \text{ for } l > 0,$$

where $\kappa(a)$ depends only on the function $a(\eta)$.

But

$$\begin{aligned}\left\|z^>(0)\right\|^2_{\mathcal{H}^{s,r}} &= \sum_{\substack{k\\ l>0}} \omega_{kl}^{2r} |z_{kl}|^2 = \sum_{\substack{k\\ l>0}} \omega_{kl}\omega_{kl}^{2(r-1/2)} |z_{kl}|^2 \\ &> \frac{\kappa}{d^{2(r-1/2)}} \left\|z^>(0)\right\|_{\mathcal{H}^{1/2}},\end{aligned}$$

so that $\|z^{>}(0)\|^2_{\mathcal{H}^{1/2}} \leq \kappa d^{2r-1} \|z^{>}(0)\|^2_{\mathcal{H}^{s,r}}$. Since $r > 1/2$, we can choose d_0 such that for $d < d_0$, we have

$$\left\|z^{>}(0)\right\|_{\mathcal{H}^{1/2}} < \epsilon, \tag{87}$$

for any $\epsilon > 0$. Again, note that if we choose r larger (and thus our initial condition more smooth), then the thickness of our domain can be larger.

We first note that we can approximate the solutions of $\dot{z} = J\nabla H^0(z)$ by those of $\dot{\gamma} = J\nabla H_0^n(\gamma)$.

Recall that we've defined $\phi_0, \dots, \phi_n$ so that $H^n = H^0 \circ \phi_0 \circ \dots \circ \phi_n$. Denote $\phi = \phi_0 \circ \dots \circ \phi_n$. If we define $\zeta = \phi^{-1}(z)$, then we have $\dot{\zeta} = J\nabla(H^0 \circ \phi)(\zeta) = J\nabla H^n(\zeta)$, with $\zeta(0) = \phi^{-1}(0)$. The $z(t)$ and $\zeta(t)$ equations express the same dynamics in different coordinates. However, now note that the transformation ϕ is a near identity transformation. More precisely, for any $R > 0$, we recall from Corollary 14 that there exists $C = C(R)$ such that on the ball of radius R in $\mathcal{H}^{1/2}_\alpha$,

$$\|\phi - \mathbf{1}\|_{\mathcal{H}^{1/2}_\alpha} \leq Cd^{1+2\alpha} \ . \tag{88}$$

Thus, we see that we so long as $z(t)$ remains within a ball of radius R, $\|z(t) - \zeta(t)\|_{\mathcal{H}^{1/2}_\alpha} \leq Cd^{1+2\alpha}R$, and hence if $Cd^{1+2\alpha}R << \epsilon$, we can approximate either the z-dynamics or the ζ-dynamics, whichever is more convenient, to within the error we allow. It turns out that it is easier to approximate the ζ dynamics because in the Hamiltonian H^n the coupling between the $l = 0$ and $l > 0$ modes is very weak.

For the following, we need some new notation. Recall that $H^0_{\text{diag}}(z) = \sum_{k,l} \omega_{kl} |z_{kl}|^2$, so we write

$$H^{0\mathbf{r}}_{\text{diag}}(z) = \sum_k \omega_{k0} |z_{k0}|^2 \ , \quad H^{0>}_{\text{diag}} = \sum_{\substack{k \\ l>0}} \omega_{kl} |z_{kl}|^2 .$$

We remind the reader that throughout the remainder of this subsection we will always be working in the Hilbert space $\mathcal{H}^{1/2}_\alpha$.

We need to verify that if we define

$$\dot{\zeta} = J\nabla H^n(\zeta), \tag{89}$$
$$\dot{\gamma} = J\nabla(H^0_{\text{diag}} + H^n_0)(\gamma), \tag{90}$$

with $\gamma(0)$ the projection of $z^{\mathbf{r}}(0)$ onto $\mathcal{H}^{1/2}_\alpha$, then γ tracks ζ sufficiently well. Since $\gamma(0)$ lies in the submanifold generated by the $\{z_{k0}\}$, and this submanifold is preserved by the Type 0 Hamiltonian, we can write the system of (89) and (90) as

$$\dot{\zeta} = J\nabla H^n(\zeta),$$
$$\dot{\gamma} = J\nabla(H^{0\mathbf{r}}_{\text{diag}} + H^n_0)(\gamma),$$

But if we write $\zeta = z^{\mathbf{r}} + z^{>}$, then the equation for ζ becomes

$$\dot{z^{\mathbf{r}}} = J\nabla H^{0\mathbf{r}}_{\text{diag}}(z^{\mathbf{r}}) + J\nabla H^{n}_{0}(z^{\mathbf{r}}) + J\nabla H^{n}_{1}(z^{>}), \tag{91}$$

$$\dot{z^{>}} = J\nabla H^{0>}_{\text{diag}}(z^{>}) + J\nabla H^{n}_{1}(z^{\mathbf{r}}) + J\nabla H^{n}_{2}(z^{>}). \tag{92}$$

First, we concentrate on (91). We define $\xi = z^{\mathbf{r}} - \gamma$, and this with (91) gives

$$\begin{aligned}\dot{\xi} &= J\nabla H^{0\mathbf{r}}_{\text{diag}}(\xi) + J\nabla H^{n}_{0}(\xi) + J\nabla H^{n}_{1}(z^{>}),\\ &= J\left(\nabla\left(H^{0\mathbf{r}}_{\text{diag}} + H^{n}_{0}\right)\right)(\xi) + J\nabla H^{n}_{1}(z^{>}).\end{aligned}$$

where $\|\xi(0)\|_{\mathcal{H}^{1/2}} = \mathcal{O}(d^{1+2\alpha})$ (from the fact that $\|\zeta(0) - z(0)\|_{\mathcal{H}^{1/2}} \le Cd^{1+2\alpha}$). We want to control $\|\xi(t)\|_{\mathcal{H}^{1/2}}$.

In Lemma 18, we proved that $\|\cdot\|_{\mathcal{H}^{1/2}_{\alpha}} \sim \|\cdot\|_{H^{0}_{\text{diag}}+H^{n}_{0}}$ for $z \in \mathcal{O}_{3R}\left(\mathcal{H}^{1/2}_{\alpha}\right)$. So it suffices to bound $\|\xi(t)\|_{H^{0}_{\text{diag}}+H^{n}_{0}}$ instead.

Then we calculate

$$\begin{aligned}\frac{d}{d\tau}\|\xi(\tau)\|^{2}_{H^{0}_{\text{diag}}+H^{n}_{0}} &= \left\langle \xi, J\nabla H^{n}_{1}(z^{>})\right\rangle_{H^{0}_{\text{diag}}+H^{n}_{0}} + \left\langle J\nabla H^{n}_{1}(z^{>}), \xi\right\rangle_{H^{0}_{\text{diag}}+H^{n}_{0}}\\ &\le 2\|\xi(\tau)\|_{H^{0}_{\text{diag}}+H^{n}_{0}}\left\|J\nabla H^{n}_{1}(z^{>}(\tau))\right\|_{H^{0}_{\text{diag}}+H^{n}_{0}}\\ &\le \kappa\|\xi(\tau)\|_{\mathcal{H}^{1/2}_{\alpha}}\|J\nabla H^{n}_{1}\|_{\mathcal{H}^{1/2}_{\alpha},\mathcal{H}^{1/2}_{\alpha}}\left\|z^{>}(\tau)\right\|_{\mathcal{H}^{1/2}_{\alpha}}\\ &\le \kappa_{n} d^{N(n)-1}(1 + \|\xi\|^{2}_{\mathcal{H}^{1/2}}).\end{aligned} \tag{93}$$

where the last inequality used the fact since $\|z(t)\|_{\mathcal{H}^{1/2}_{\alpha}} < 3R$ for all t, so is $\|z^{>}(\tau)\|_{\mathcal{H}^{1/2}_{\alpha}}$ and the fact that we showed in Lemma 16 that $\|J\nabla H^{n}_{1}\|_{\mathcal{H}^{1/2}_{\alpha},\mathcal{H}^{1/2}_{\alpha}} \le \kappa_{n} d^{N(n)-1}$.

Recall that we defined in Subsection 1.2 that $\tau = d \cdot t$, so that

$$\frac{d}{dt}\|\xi(\tau)\|^{2}_{H^{0}_{\text{diag}}+H^{n}_{0}} = d \cdot \frac{d}{d\tau}\|\xi(\tau)\|^{2}_{H^{0}_{\text{diag}}+H^{n}_{0}} \le \kappa_{n} d^{N(n)}(1 + \|\xi\|^{2}_{\mathcal{H}^{1/2}}). \tag{94}$$

Using Grönwall's Inequality and the fact that $\|\xi(0)\|^{2}_{\mathcal{H}^{1/2}} < C'd^{2+4\alpha}$ (from (88)), we have

$$\|\xi(t)\|^{2}_{\mathcal{H}^{1/2}_{\alpha}} \le e^{\kappa_{n} C}\left(C'd^{2+4\alpha} + C_{n} d^{N(n)} t\right), \tag{95}$$

so long as $t < Cd^{-N(n)}$.

For any $\epsilon > 0$, we can certainly choose d sufficiently small so that $C'd^{2+4\alpha} < \epsilon$, and thus

$$\|\xi(t)\|^{2}_{\mathcal{H}^{1/2}_{\alpha}} \le C'_{n} d^{N(n)} t + \epsilon. \tag{96}$$

Now we consider (92), the $z^{>}$ equation. An analysis similar to that in (93) gives us that

$$\frac{d}{dt}\left\|z^{>}\right\|^{2}_{\mathcal{H}^{1/2}_{\alpha}} \le \kappa\|J\nabla H^{n}_{1}\|_{\mathcal{H}^{1/2}_{\alpha},\mathcal{H}^{1/2}_{\alpha}}\left\|z^{>}\right\|_{\mathcal{H}^{1/2}_{\alpha}}.$$

A Grönwall argument combined with the estimate in (87) gives us that, for some constant C_n'',

$$\left\| z^{>} \right\|_{\mathcal{H}_\alpha^{1/2}} \leq C_n'' d^{N(n)} t + \epsilon. \tag{97}$$

Combining this estimate with Equation (86) to control the modes with large k gives us that

$$\|\zeta - z^{\mathbf{r}}\|_{\mathcal{H}^{1/2}} \leq \epsilon.$$

This combined with (96) and (97) gives us

$$\|\zeta(t) - \gamma(t)\|_{\mathcal{H}^{1/2}} \leq (C_n' + C_n'') d^{N(n)} t + 3\epsilon.$$

■

4.2 Proof of Theorem 1

We want to show here that Theorem 3 implies Theorem 1. We begin by noting the map embodied by transformations (20) and (22) (the map from continuous u variables to discrete z variables) is a continuous, linear map with continuous inverse. We know that z and γ stay close in Theorem 3. This together with Lemma 22 will show that the continuous analogs of z and γ also stay close.

Lemma 22. *We define the map Ξ, written as $\Xi_2 \circ \Xi_1$, where*

$$\Xi_1(u(x,y,t), u_t(x,y,t)) = \{\hat{u}_{k,l}, \dot{\hat{u}}_{k,l}\},$$
$$\Xi_2\left(\{\hat{u}_{k,l}, \dot{\hat{u}}_{k,l}\}\right) = \{z_{kl}, \overline{z}_{kl}\},$$

with

$$u(x,y,t) = \sum_{kl} \hat{u}_{k,l}(t)\phi_k(x)\psi_l(y/d),$$
$$z_{kl} = \frac{1}{\sqrt{2\omega_{kl}}}\dot{\hat{u}}_{k,l} + i\sqrt{\frac{\omega_{kl}}{2}}\hat{u}_{k,l}.$$

Then given s a half-integer, the map $\Xi \colon H^{s+1/2} \times H^{s-1/2} \to \mathcal{H}^s$ is a bounded invertible linear transformation with bounded inverse. Also, Ξ is a canonical transformation with multiplier i.

Proof: This is a straightforward calculation using the asymptotics of the eigenvalues μ_k and λ_l, analogous to Section 1.2 in Kuksin [17].

■

Remark 13. *If we restrict the domain of Ξ to $\mathcal{H}_\alpha^{s+1/2} \times \mathcal{H}_\alpha^{s-1/2}$, it remains a canonical transformation from this set onto its image.*

In Theorem 3, we assumed that the initial condition $z(0) \in \mathcal{H}^{s,r}$ for $s > 1$, $r > 1/2$. One possible choice would be to choose $s = r > 1$. We should note that in this case, the norm $\|z\|_{\mathcal{H}^{s,s}} = \|z\|_{\mathcal{H}^s}$. Then, the previous lemma tells us that choosing an initial condition $z(0) \in \mathcal{H}^s$ with $s > 1$ corresponds to choosing an initial pair $(u^0, u_t^0) \in H^{s+1/2} \times H^{s-1/2}$, with $s > 1$. For example, choosing $(u^0, u_t^0) \in H^2 \times H^1$ will do.

We will show just below that the equation

$$\dot{\gamma} = J\nabla H_0^n(\gamma)$$

corresponds to an equation of the form $u_{tt} = D_n u$, in the following sense. Let $(u(0), u_t(0)) \in H^2 \times H^1$, and let us assume that the equation $u_{tt} = D_n u$ is satisfied. Then for any time $\tau > 0$, the pair $(u(\tau), u_t(\tau))$ is uniquely determined. On the other hand, consider $z(0) = \Xi(u(0), u_t(0))$, and assume that it satisfies $\dot{z} = J\nabla H_0^n(z)$, thus specifying $z(\tau)$ for any $\tau > 0$. Then $z(\tau) = \Xi(u(\tau), u_t(\tau))$.

So, as in Theorem 1, let's assume that we have $(u^0(x,y), u_t^0(x,y)) \in H^2 \times H^1$, and compute $(\overline{u^{0\mathbf{r}}}, \overline{u_t^{0\mathbf{r}}}) \in H_\alpha^2 \times H_\alpha^1$ as the projections into H_α^2 and H_α^1 of $\Pi(u^0)$ and $\Pi(u_t^0)$. If we define $z = \Xi(u, u_t)$ and $\gamma = \Xi(\overline{u^{\mathbf{r}0}}, \overline{u^{\mathbf{r}0}_t})$, then $z, \gamma \in \mathcal{H}^s$ with $s > 1$, and Theorem 3 tells us that for any $\delta > 0$,

$$\|z(t) - \gamma(t)\|_{\mathcal{H}^{1/2}}^2 \leq C d^{N(n)} t + 3\delta.$$

From Lemma 22 and the linearity and continuity of Ξ^{-1} we have that

$$\left\|\Xi^{-1}\left(z(t)\right) - \Xi^{-1}\left(\gamma(t)\right)\right\|_{H^1 \times L^2} \leq C(d^{N(n)} t + 3\delta).$$

So, given an ϵ, we can find a C such that for all $t \leq C d^{-N(n)}$, we have

$$\|u(x,y,t) - {u_n}^{\mathbf{r}}(x,t)\mathbf{1}(y)\|_{H^1 \times L^2} = \left\|\Xi^{-1}\left(z(t)\right) - \Xi^{-1}\left(\gamma(t)\right)\right\|_{H^1 \times L^2} \leq \epsilon.$$

To finish our proof of the theorem, we need to calculate explicitly what the Type 0 term is after n changes of variables. From Theorem 3, we know that the dynamics of the Type 0 term mimic the dynamics of the whole system, and thus we would like to calculate more specifically what the Type 0 term actually is. We will show that this Type 0 term, when converted back to continuous coordinates, gives a PDE of the form (5).

From the proof of Theorem 3 we know that it suffices to consider the evolution in $\mathcal{H}_\alpha^{1/2}$, (since the modes with $k > d^\alpha$ are initially very small, and never grow beyond the size of our allowed error) and thus in the remaining discussion we restrict our attention to this space. Note that by combining (78), and Lemma 9 we see that

$$\begin{aligned} H_0^n(z) &= \sum_{\substack{p=2 \\ p \text{ even}}}^{P_{max}} \sum_{\substack{\tilde{I} \in \tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I}) = p}} \Theta_1(I) C_I d^p \sum_{k < d^\alpha} \omega_{k0}^{p+1} (z_{k0} - \overline{z}_{k0})^2 + \mathcal{R}_{n,N}^{(0)}(z) \\ &= -\frac{1}{2} \sum_{q=2}^{Q_{max}} \tilde{C}^{(q)} d^{2(q-1)} \sum_{k < d^\alpha} \omega_{k0}^{2q-1} (z_{k0} - \overline{z}_{k0})^2 + \mathcal{R}_{n,N}^{(0)}(z) \ . \quad (98) \end{aligned}$$

where in the last equality we set $p = 2(q-1)$ and defined $\tilde{C}^{(\ell)} = \sum_{\substack{\tilde{I} \in \tilde{\mathcal{I}}_{n,N} \\ p(\tilde{I}) = \ell}} (-2C_I)\Theta_{0+2}(I)$.

Remark 14. *Note that from the formula given above it appears that the coefficients $C^{(\ell)}$ could depend on n or N. In [13], a closed form expression for these coefficients is derived from which it follows that $C^{(\ell)}$ is independent of both n and N, but the conclusions of this paper would be unaffected even if they did depend on these parameters.*

We are not completely done. In order to derive our reduced PDE we will replace H_0^n by an approximation with finitely many terms and we must show that this can be done without affecting the dynamics too much. For this we use the following lemma:

Lemma 23. *Fix $R > 0$. There exists $\epsilon_0 > 0$ such that for $0 < \epsilon < \epsilon_0$ the following is true. Consider two quadratic Hamiltonians $H = H^{\mathbf{0r}}_{\rm diag} + h_1$ and $\tilde{H} = H^{\mathbf{0r}}_{\rm diag} + h_2$, both of which define norms equivalent to the norm defined by $H^{\mathbf{0r}}_{\rm diag}$, and assume that $\|h_1 - h_2\|_{\mathcal{A}_\alpha} < \epsilon$. If we consider the two flows*

$$\dot{u} = J\nabla(H^{\mathbf{0r}}_{\rm diag} + h_1) \quad \text{and} \quad \dot{v} = J\nabla(H^{\mathbf{0r}}_{\rm diag} + h_2)$$

with $u(0) = v(0)$ and $\|u(0)\|_{\mathcal{H}^{1/2}_\alpha} < R$, then the flow for v provides a good approximation to the flow for u on the space $\mathcal{H}^{1/2}_\alpha$ in the sense that there exists $C(R) > 0$ (independent of $u(0)$, so long as $\|u(0)\|_{\mathcal{H}^{1/2}_\alpha} < R$) such that

$$\|u(t) - v(t)\|_{\mathcal{H}^{1/2}_\alpha} \leq C\epsilon dt \ .$$

Proof: Noting that $u(t)$ and $v(t)$ are solutions of linear systems of Hamiltonian, ordinary differential equations, and that the difference of the corresponding vectorfields has Lipschitz constant bounded by ϵ, the estimate then follows from a straightforward application of Grönwall's inequality using the fact that the Hamiltonian's are conserved along their respective orbits. The extra factor of d on the right hand side of this inequality comes from the rescaling of time as in inequality (94).

■

We continue for the moment to assume that R is fixed and that we consider sequences $\{z_{k0}\}$ with $\|z\|_{\mathcal{H}^{1/2}_\alpha} < R$. In Section 4.1, we have already specified an order in d to which our reduced equation is accurate, so we will simply cut off the expression for H_0^n at exactly the same order in d. As stated in Theorem 3, after we have done n changes of variables, our approximation is accurate $\mathcal{O}(d^{N(n)})$ where $N(n) = 2n + 1 + (2n + 2)\alpha$. Thus, we first choose N such that $(1 + 2\alpha)N > N(n)$, which guarantees that $\|R^{(0)}_{n,N}\|_{\mathcal{A}_\alpha} \leq Cd^{(1+2\alpha)N} < Cd^{N(n)}$. We next note that if we define $h_q = \sum_{k<d^\alpha} \omega_{k0}^{2q-1}(z_{k0} - \overline{z}_{k0})^2$ an easy estimate like that leading to (52) implies $\|h_q\|_{\mathcal{A}_\alpha} \leq 4d^{(2q-1)\alpha}$. Thus if we define $\overline{C} = \max_{q=2,\ldots,Q_{max}} |\tilde{C}^{(q)}|$, we have

$$\| \sum_{q=M+1}^{Q_{max}} \tilde{C}^{(q)} d^{2(q-1)} \sum_{k<d^\alpha} \omega_{k0}^{2q-1}(z_{k0} - \overline{z}_{k0})^2 \|_{\mathcal{A}_\alpha} \leq 4Q_{max}\overline{C}d^{2M(1+\alpha)+\alpha} \ , \quad (99)$$

for d sufficiently small. Thus, if we choose $2M(1+\alpha)+\alpha > N(n)-1$, the expression on the left hand side of (99) is less than or equal to $Cd^{N(n)-1}$, and so if we define

$$\widetilde{H_0^n} = \sum_{q=2}^{M} \tilde{C}^q d^{2(q-1)} \sum_{k<d^\alpha} \omega_{k0}^{2q-1} (z_{k0} - \overline{z}_{k0})^2 \tag{100}$$

we have

$$\|H_0^n - \widetilde{H_0^n}\|_{\mathcal{A}_\alpha} < Cd^{N(n)-1} \ . \tag{101}$$

Remark 15. *Note that the inequality $2M(1+\alpha)+\alpha > N(n)-1$ can be satisfied by choosing $M = n$.*

As an immediate corollary of (101) and Lemma 23 we have

Lemma 24. *If we consider the two flows*

$$\dot{u} = J\nabla(H^{0\mathbf{r}}_{\text{diag}} + H_0^n)(u),$$
$$\dot{v} = J\nabla(H^{0\mathbf{r}}_{\text{diag}} + \widetilde{H_0^n})(v),$$

with $u(0) = v(0)$ on the space $\mathcal{H}^{1/2}_\alpha$, then the flow for v provides a good approximation for the flow for u in the sense that

$$\|u(t) - v(t)\|^2_{\mathcal{H}^{1/2}_\alpha} \leq Cd^{N(n)}t,$$

where M is chosen as in (100).

Finally, we examine the form of $\widetilde{H_0^n}$ in continuous coordinates. Since by 22, $(z_{k0} - \overline{z}_{k0}) = i\sqrt{2\omega_{k0}}\hat{u}_{k0}$, we have

$$\widetilde{H_0^n} = \sum_{q=2}^{M} \tilde{C}^{(q)} d^{2(q-1)} \sum_{k<d^\alpha} \omega_{k0}^{2q} \hat{u}_{k0}^2 = \sum_{q=2}^{M} \tilde{C}^{(q)} d^{2(q-1)} \sum_{k<d^\alpha} \mu_k^q \hat{u}_{k0}^2 \tag{102}$$

With this form of the Hamiltonian, we are finally in a position to prove:

Lemma 25. *Consider the form of $\widetilde{H_0^n}$ given in (102). The differential equation (on $\mathcal{H}^{1/2}_\alpha$) given by*

$$\dot{z} = J\nabla\left(H^{0\mathbf{r}}_{\text{diag}} + \widetilde{H_0^n}\right)(z),$$

if converted into continuous coordinates by Ξ^{-1} as defined in Lemma 22, gives a PDE of the form

$$u^{\mathbf{r}}_{tt} = L_x u^{\mathbf{r}} + \sum_{q=2}^{M} C^q d^{2(q-1)} L_x^q u^{\mathbf{r}} \ \text{on } \omega, \tag{103}$$

with the boundary conditions

$$L_x^i u\big|_{\partial\omega} = 0 \text{ for all } i < M, \tag{104}$$

where $C^q = (-1)^{q-1} 2\tilde{C}^q$.

Before proving this Lemma we note that from Section 4.2, we know that solutions to (103) will approximate well the solutions to (1), since the corresponding equations do so in the z coordinates, and the transformation Ξ from z variables to u variables is bounded with bounded inverse. Noting that (103) is of exactly the form of (5), and recalling our earlier remark that we can choose $M = n$, this completes the proof of Theorem 1.

Proof: Reexpressing $H^{0\mathbf{r}}_{\text{diag}}$ in continuous coordinates as we did $\widetilde{H^0_n}$ in (102) we see that $H^{0\mathbf{r}}_{\text{diag}} = \frac{1}{2}\sum_k (|\dot{\hat{u}}_{k0}|^2 + \mu_k |\hat{u}_{k0}|^2)$.

Now consider the subspace of $H^{2M}(\omega)$ defined by

$$W^M(\omega) = \{u \in H^{2M}(\omega) \mid L_x^j u|_{\partial\omega} = 0 \ , j = 0, 1, \ldots, M-1\} \ .$$

(Note that if $z \in \mathcal{H}_\alpha^{1/2}$, the corresponding u reconstructed by undoing the various changes of variables will be in $H^s(\omega)$ for any s.) Then an easy inductive argument using the definition of $\mathcal{L}$ shows that $\int_\omega (\mathcal{L}^j u)^2 dx = (-1)^M \int u L_x^j u dx$, for $j = 0, \ldots, M$.

From this observation it easily follows that if $u \in W^M$ and $q \leq M$, and we write $u = \sum \hat{u}_{k0}\phi_k(x)$ we have $\sum_k \mu^q \hat{u}_{k0}^2 = \int (u L_x^q u) dx = (-1)^q \int (\mathcal{L}^q u)^2 dx$.

$$H^{0\mathbf{r}}_{\text{diag}} + \widetilde{H^n_0} = \int \{\frac{1}{2}u_t^2 + \frac{1}{2}(\mathcal{L}u)^2 + \sum_{q=2}^{M} (-1)^q \tilde{C}^{(q)} d^{2(q-1)} (\mathcal{L}^q u)^2\} dx \ . \qquad (105)$$

Hamilton's equations for this Hamiltonian are readily seen to be the partial differential equation (103), with boundary conditions (104), and the Lemma follows.

■

A Proofs of Lemmas 9 and 10

Proof: (of Lemma 9). We prove this lemma by induction. We begin by considering the case where $I \in \mathcal{M}_0 \cap \mathcal{M}^+$, i.e. $I = 0^n$ for some $n > 0$. (Recall that if $I = 0^n$, and χ_0 is the Hamiltonian whose time 1 map gives the first canonical transformation, then $L_I(G) = \{\chi_0, \{\chi_0 \ldots \{\chi_0, G\} \ldots\}\}$, where there are a total of n Poisson brackets.)

We begin by explicitly calculating the Type 0 term when $n = 1$, i.e. $L_0(G)_{0,k00}$. By part 2 of Lemma 4, G must be H^0_1, and from the table in Lemma 5, we see that

$$\begin{aligned} L_0(G)_{0,k00} &= 2\sum_{l>0} H^0_{1,kl0} D_{0,kl} \\ &= 2\sum_{l>0} \frac{\omega_{k0}^{3/2}}{4\omega_{kl}^{1/2}} \beta_{l,0} \times d^2 \omega^1_{k0} (\omega_{k0}\omega_{kl})^{1/2} \frac{\beta_{l,0}}{\lambda_l} \\ &= \frac{1}{2} d^2 \omega^3_{k0} \sum_{l>0} \frac{(\beta_{l,0})^2}{\lambda_l} . \end{aligned}$$

This verifies the formula for $L_I(G)_{0,k00}$ in Lemma 9 in the case $I=0$ if we take $C_{0^1} = \frac{1}{2}\sum_{l>0}\frac{(\beta_{l,0})^2}{\lambda_l}$. (The convergence of the sum follows from the fact that the $\beta_{l,0}$ are the Fourier coefficients of an L^2 function.)

The derivation of the formulas for the Type 1 and 2 terms are very similar and are left as an exercise. (The details are also presented in [[12], Appendix H].)

We now assume that Lemma 9 holds for $I = 0, 0^2, \ldots, 0^n$ and prove that it holds for $I = 0^{n+1}$. Once again, we will perform the detailed calculations for the Type 0 terms and leave the analogous computations of the Type 1 and 2 terms as an exercise. We know that $L_{0^{n+1}}(G)_0 = L_0(L_{0^n}(G)_1)$. So we calculate

$$\begin{aligned} L_0(L_{0^n}(G)_1) &= 2\sum_{l>0} Q_{0^n}(l) d^{2n}\omega_{k0}^{2n+1}\frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}} \times -d^2\omega_{k0}^2\frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}}\frac{\beta_{l,0}}{\lambda_l} \\ &= -2d^{2n+2}\omega_{k0}^{2n+3}\sum_{l>0}\frac{Q_{0^n}(l)\beta_{l,0}}{\lambda_l}. \end{aligned}$$

Thus

$$C_{0^{n+1}} = -2\sum_{l>0}\frac{Q_{0^n}(l)\beta_{l,0}}{\lambda_l}.$$

Since $Q_{0^n}, \beta_{l,0} \in \ell^2$, this sum converges. We now know that the expressions for $L_I(G)_{j,kll'}$ in Lemma 9 holds for all $I \in \mathcal{M}_0 \cap \mathcal{M}^+$.

We next check that $H^1_{1,\text{low}}$ has the stated form. Recall that by definition, $H^1_{1,\text{low}}$ is the sum of the lowest order terms in d in H^1_1. But given the expression for $L_I(G)_{1,kl0}$ established above, these will be exactly those terms with $p(I)=2$ – i.e. those with $I_1 = 0$. Thus, if $n=1$, these are precisely the terms in $\mathcal{M}_0$ with $p(I)=2$ and the formula for $H^1_{1,\text{low}}$ follows.

We now complete the proof of Lemma 9 by showing that if it holds for $I \in \mathcal{M}_{n-1}\cap\mathcal{M}^+$, then it also holds for $I\in\mathcal{M}_n\cap\mathcal{M}^+$.

Note that by the induction hypothesis,

$$\begin{aligned} H^{n-1}_{1,\text{low},kl} &= \sum_{\substack{I\in\mathcal{M}_{n-1}\cap M^+ \\ p(I)=2n}} \Theta_1(I)Q_I(l)d^{p(I)}\omega_{k0}^{p(I)+1}\frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}} \\ &= d^{2n}\omega_{k0}^{2n+1}\frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}}\sum_{\substack{I\in\mathcal{M}_{n-1}\cap M^+ \\ p(I)=2n}}\Theta_1(I)Q_I(l) \end{aligned} \tag{106}$$

If we define

$$K^n(l) = 4\sum_{\substack{I\in\mathcal{M}_{n-1}\cap M^+ \\ p(I)=2n}}\Theta_1(I)Q_I(l), \tag{107}$$

Note that $\|K^n(\cdot)\|_{\ell^2} \le 4\sum_{\substack{I\in\mathcal{M}_{n-1}\cap M^+ \\ p(I)=2n}}|\Theta_1(I)|\|Q_I(l)\|_{\ell^2} < \infty$ since $\|Q_I(\cdot)\|_{\ell^2}$

is finite for all I and there are finitely many terms in the sum. Then

$$H^n_{1,\mathrm{low},kl} = \frac{1}{4} d^{2n} \omega_{k0}^{2n+1} \frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}} K^n(l).$$

Then we have (see Lemma 4)

$$D_{n,kl} = -d^{2n+2} \omega_{k0}^{2n+1} \left(\omega_{k0}\omega_{kl}\right)^{1/2} \frac{K^n(l)}{\lambda_l}, \tag{108}$$

$$S_{n,kl} = d^{2n+2} \omega_{k0}^{2n+2} \frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}} \frac{K^n(l)}{\lambda_l}. \tag{109}$$

To establish the formulas for $L_I(G)$ for $I \in \mathcal{M}_n$, note that if $I \in \mathcal{M}_n$, then either $I = n^q$ for some integer q or $I = n^q \tilde{I}$ for some $\tilde{I} \in \mathcal{M}_{n-1}$. We will first show that the formulas in Lemma 9 hold if $I = n$. This is again a straightforward calculation. For the Type 0 terms we have

$$\begin{aligned} L_n(G)_{0,k00} &= 2 \sum_{l>0} H^0_{1,kl0} D_{n,kl} \\ &= 2 \sum_{l>0} \frac{\omega_{k0}^{3/2}}{-4\omega_{kl}^{1/2}} \beta_{l,0} \times -d^{2n+2} \omega_{k0}^{2n+1} (\omega_{k0}\omega_{kl})^{1/2} \frac{K^n(l)}{\lambda_l} \\ &= \frac{1}{2} d^{2n+2} \omega_{k0}^{2n+3} \sum_{l>0} \frac{K^n(l)\beta_{l,0}}{\lambda_l}, \end{aligned}$$

and the (56) follows with $C_{I=n} = \sum_{l>0} \frac{K^n(l)\beta_{l,0}}{\lambda_l}$. Since both $K^n(l)$ and $\beta_{l,0}$ are in ℓ^2 the sum is bounded. Again, we leave the analogous computations of $L_n(G)_{1,kl0}$ and $L_n(G)_{2,kll'}$ as exercises.

Now suppose that we know that (56)-(58) hold for some I. We prove that they also hold for $I' = nI$. Since we know that they hold for $I = n$ or for $I \in \mathcal{M}_{n-1}$, the observation above about the form of $I \in \mathcal{M}_n$ and an inductive argument completes the proof. As above we compute in detail the formula for the Type 0 terms and leave the Type 1 and 2 terms as exercises. According to Lemma 5, we have that

$$\begin{aligned} L_{nI}(G)_{0,k00} &= 2 \sum_{l>0} d^{p(I)} \omega_{k0}^{p(I)+1} \frac{\omega_{k0}^{1/2}}{\omega_{kl}^{1/2}} Q_I(l) \times -d^{2n+2} \omega_{k0}^{2n+1} \omega_{k0}^{1/2} \omega_{kl}^{1/2} \frac{K^n(l)}{\lambda_l} \\ &= -2 d^{p(I)+2n+2} \omega_{k0}^{p(I)+2n+3} \sum_{l>0} \frac{Q_I(l) K^n(l)}{\lambda_l} \\ &= d^{p(nI)} \omega_{k0}^{p(nI)+1} C_{nI}, \end{aligned}$$

with

$$C_{nI} = -2 \sum_{l>0} \frac{Q_I(l) K^n(l)}{\lambda_l}. \tag{110}$$

Since $Q_I, K^n \in \ell^2$, this sum converges and, $L_{nI}(G)_{0,k00}$ satisfies the estimate in (56).

Note that now that we have proven (56)-(58), the formula for $H^n_{1,\mathrm{low}}$ follows immediately from the definition.

■

Proof: (of Lemma 10.) As in the proof of Lemma 9 we will provide all the details of the estimate of $L_I(G)_0$, and leave as exercises the very similar estimates of $L_I(G)_1$ and $L_I(G)_2$.

We use the fact that $L_I(G)_0 = C_I d^{p(I)} \omega_{k0}^{p(I)+1}$, and

$$\partial_{z_{k0}} L_I(G)_0(z) = 2C_I d^{p(I)} \omega_{k0}^{p(I)+1} \left(z_{k0} - \overline{z}_{k0}\right).$$

By definition,

$$\|\nabla L_I(G)_0(z)\|_{\mathcal{H}_\alpha^{1/2}} = \sum_{k<d^\alpha} \omega_{k0} \left|\partial_{z_{k0}} L_I(G)_0\right|^2.$$

Then we calculate

$$\begin{aligned}
\|\nabla L_I(G)_0(z)\|^2_{\mathcal{H}_\alpha^{1/2}} &= \sum_{k<d^\alpha} \omega_{k0} \left|\partial_{z_{k0}} L_I(G)_0\right|^2 \\
&\leq \sum_{k<d^\alpha} \omega_{k0} 8C_I d^{2p(I)} \omega_{k0}^{2p(I)+2} \left|z_{k0}\right|^2 \\
&\leq 8C_I d^{2p(I)+\alpha(2p(I)+2)} \sum_{k<d^\alpha} \omega_{k0} \left|z_{k0}\right|^2 \\
&= 8C_I d^{2p(I)+\alpha(2p(I)+2)} \|z\|^2_{\mathcal{H}_\alpha^{1/2}}.
\end{aligned}$$

Thus

$$\|\nabla L_I(G)_0\|_{\mathcal{H}_\alpha^{1/2}, \mathcal{H}_\alpha^{1/2}} \leq 8C_I d^{p(I)(1+\alpha)+\alpha}.$$

■

Acknowledgments: Many of the results in this paper appeared in the first author's Ph. D. thesis at Boston University. The second author thanks D. Arnold and A. Mielke for several helpful conversations related to this work. The research of the authors was supported in part by the National Science Foundation through grants DMS-9803164 and DMS-0103915. The first author was also supported in the VIGRE program under National Science Foundation grant DMS-9983646.

References

[1] Stephen M. Alessandrini, Douglas N. Arnold, Richard S. Falk, and Alexandre L. Madureira. Derivation and justification of plate models by variational methods. In *Plates and shells (Québec, QC, 1996)*, pages 1–20. American Mathematical Society, Providence, RI, 1999.

[2] Douglas N. Arnold and Richard S. Falk. Analysis of a linear-linear finite element for the Reissner-Mindlin plate model. *Mathematical Models & Methods in Applied Sciences*, 7(2):217–238, 1997.

[3] I. Babuška and C. Schwab. A posteriori error estimation for hierarchic models of elliptic boundary value problems on thin domains. *SIAM Journal on Numerical Analysis*, 33(1):221–246, 1996.

[4] Dario Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Mathematische Zeitschrift*, 230(2):345–387, 1999.

[5] G. Benettin, A. Carati, and G. Gallavotti. A rigorous implementation of the Jeans-Landau-Teller approximation for adiabatic invariants. *Nonlinearity*, 10(2):479–505, 1997.

[6] Giancarlo Benettin, Poul Hjorth, and Paolo Sempio. Exponentially long equilibrium times in a one-dimensional collisional model of classical gas. *Journal of Statistical Physics*, 94(5-6):871–891, 1999.

[7] Jean Bourgain. Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE. *International Mathematics Research Notices*, (11):475ff., approx. 21 pp. (electronic), 1994.

[8] Isaac Chavel. *Eigenvalues in Riemannian geometry.* Academic Press Inc., Orlando, Fla., 1984.

[9] P. G. Ciarlet and B. Miara. Justification of the two-dimensional equations of a linearly elastic shallow shell. *Comm. Pure Appl. Math.*, 45(3):327–360, 1992.

[10] Philippe G. Ciarlet and Hervé Le Dret. Justification of the boundary conditions of a clamped plate by an asymptotic analysis. *Asymptotic Analysis*, 2(4):257–277, 1989.

[11] Philippe G. Ciarlet and Véronique Lods. Asymptotic analysis of linearly elastic shells. I. Justification of membrane shell equations. *Archive for Rational Mechanics and Analysis*, 136(2):119–161, 1996.

[12] R. E. Lee DeVille. *Reduced equations for hyperbolic problems on thin domains.* PhD thesis, Boston University, 2001.

[13] R. E. Lee DeVille. Reduced equations for models of laminated materials in thin domains. II. preprint, 2001.

[14] Ivica Durdević. *Ein Beitrag zur Modellierung der Eigendynamik elastischer Schalen mit Hamiltonschen Strukturen.* PhD thesis, Universität Stuttgart, 2000.

[15] Lawrence C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, RI, 1998.

[16] E. L. Ince. *Ordinary Differential Equations.* Dover Publications, New York, 1944.

[17] Sergej B. Kuksin. *Analysis of Hamiltonian PDEs.* Oxford lecture series in mathematics and its applications; 19. Oxford University Press, Oxford, New York, 2000.

[18] Alexander Mielke. On the justification of plate theories in linear elasticity theory using exponential decay estimates. *J. Elasticity*, 38(2):165–208, 1995.

[19] N. N. Nekhorošev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Akademija Nauk SSSR i Moskovskoe Matematičeskoe Obščestvo. Uspehi Matematičeskih Nauk.*, 32(6(198)):5–66, 287, 1977.

[20] Jürgen Pöschel. On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi. *Nonlinearity*, 12(6):1587–1600, 1999.

[21] Jürgen Pöschel. On Nekhoroshev's estimate at an elliptic equilibrium. *International Mathematics Research Notices*, (4):203–215, 1999.

[22] A. Raoult. Construction d'un modèle d'évolution de plaques avec terme d'inertie de rotation. *Ann. Mat. Pura Appl. (4)*, 139:361–400, 1985.

[23] Annie Raoult. Asymptotic modeling of the elastodynamics of a multistructure. *Asymptotic Analysis*, 6(1):73–108, 1992.

[24] Jeffrey Rauch. *Partial Differential Equations.* Springer-Verlag, New York, 1991.

[25] C. Schwab. A posteriori modeling error estimation for hierarchic plate models. *Numer. Math.*, 74(2):221–259, 1996.

[26] Li-ming Xiao. Asymptotic analysis of dynamic problems for linearly elastic shells—justification of equations for dynamic membrane shells. *Asymptot. Anal.*, 17(2):121–134, 1998.