Title
Finite Depth Gravity Water Waves in Holomorphic Coordinates

Journal
Annals of PDE, 3(1)

ISSN
2524-5317

Authors
Harrop-Griffiths, Benjamin
Ifrim, Mihaela
Tataru, Daniel

Publication Date
2017-06-01

DOI
10.1007/s40818-017-0022-z

Peer reviewed

FINITE DEPTH GRAVITY WATER WAVES IN HOLOMORPHIC COORDINATES

BENJAMIN HARROP-GRIFFITHS, MIHAELA IFRIM, AND DANIEL TATARU

Abstract. In this article we consider irrotational gravity water waves with finite bottom. Our goal is two-fold. First, we represent the equations in holomorphic coordinates and discuss the local well-posedness of the problem in this context. Second, we consider the small data problem and establish cubic lifespan bounds for the solutions. Our results are uniform in the infinite depth limit, and match our earlier infinite depth result in [8].

Contents

1. Introduction

This article is devoted to the study of the two dimensional finite bottom gravity water wave equations. Precisely, we consider an inviscid, incompressible, irrotational fluid evolving in the presence of gravity. The fluid occupies a time dependent domain $\Omega(t) \subset \mathbb{R}^2$ which has flat finite bottom $\{y = -h\}$ and a free upper boundary $\Gamma(t)$ which is asymptotically flat to $y \approx 0$. The two parameters in the problem, i.e., the gravity g and the depth h, are allowed to be arbitrary positive numbers. However, our results are only uniform in the range $g \lesssim h$, which includes the infinite depth limit but not the zero depth limit.

The first author was supported by a Junior Fellow award from the Simons Foundation.

The second author was supported by the Simons Foundation.

The third author was partially supported by the NSF grant DMS-1266182 as well as by the Simons Foundation.

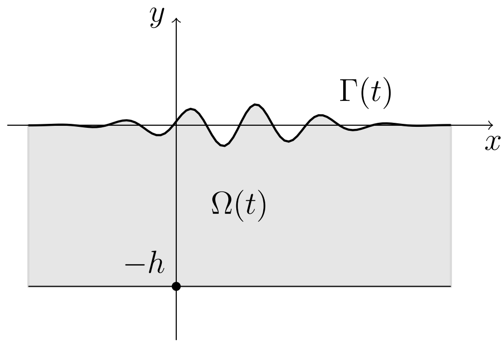

FIGURE 1. The fluid domain.

The fluid evolution is modeled by the incompressible Euler equations in $\Omega(t)$,

$$\begin{cases} u_t + u \cdot \nabla u = \nabla p - g\mathbf{j} \\ \text{div } u = 0 \\ u(0,x) = u_0(x). \end{cases} \tag{1.1}$$

On the bottom we have the boundary conditions for the velocity, namely

$$u \cdot \mathbf{j} = 0, \qquad y = -h. \tag{1.2}$$

On the free boundary $\Gamma(t)$, on the other hand, we have the dynamic boundary condition

$$p = 0 \quad \text{on } \Gamma(t), \tag{1.3}$$

and the kinematic boundary condition

$$\partial_t + u \cdot \nabla \text{ is tangent to } \bigcup_t \Gamma(t). \tag{1.4}$$

Under the additional assumption that the flow is irrotational, we can write u in terms of a velocity potential ϕ as $u = \nabla\phi$, where ϕ is a harmonic function whose normal derivative is zero on the bottom. Thus ϕ is determined by its trace $\psi = \phi|_{\Gamma(t)}$ on the free boundary $\Gamma(t)$. Denote by η the height of the water surface as a function of the horizontal coordinate. Then the fluid dynamics can be expressed in terms of a one-dimensional evolution of the free interface, Precisely, for the pairs of variables (η, ψ) we have

$$\begin{cases} \partial_t \eta - G(\eta)\psi = 0 \\ \partial_t \psi + g\eta + \dfrac{1}{2}|\nabla\psi|^2 - \dfrac{1}{2}\dfrac{(\nabla\eta \cdot \nabla\psi + G(\eta)\psi)^2}{1 + |\nabla\eta|^2} = 0, \end{cases} \tag{1.5}$$

where G represents the Dirichlet to Neumann map on the free boundary $\Gamma(t)$ associated to the Laplace equation inside the fluid domain with zero Neumann boundary condition on the bottom. This is the Eulerian formulation of the gravity water wave equations. The second equation above is known as Bernoulli's law.

While the above Eulerian formulation is easy to write, it is not so convenient to use due to the presence of the Dirichlet to Neumann map associated to the moving domain $\Omega(t)$. Instead, viewing the choice of the parametrization of the free boundary as a form of gauge freedom, we employ the holomorphic coordinates here. These are obtained using the so-called conformal method, where the domain $\Omega(t)$ is viewed as conformally equivalent to a strip. This will significantly simplify the analysis.

This system has received considerable attention over the years. The first steps toward understanding the local theory were due to Ovsjannikov, see [15], who used conformal coordinates in order to prove local well-posedness in spaces of analytic functions. Around the same time, in closely related work, Nalimov [14] proved the first small data result in Sobolev spaces in the infinite depth case. Somewhat later, his approach was extended to the finite bottom problem by Yosihara [20].

The Eulerian form of the equations, described above, emerged in [5, 21]; however, it was only much later that this led to a satisfactory local theory. For a good description of this we refer the reader to the more recent paper of Alazard-Burq-Zuily [1] as well as to Lannes's book [12].

Returning to the conformal method, the evolution equations restricted to the boundary were independently written by Wu [19] and Dyachenko-Kuznetsov-Spector-Zakharov [6] in the infinite bottom case in slightly different forms. Of these, it was Wu's paper [19] where this formulation was fully exploited to prove local well-posedness in the large data problem. Later, in [3] Choi and Camassa re-derive the equations for a perfect fluid in the finite depth case when taking both the gravity and capillary force into account. Their method is based on a direct manipulation of the Euler equations, whereas the method of Dyachenko et al. [6] is based on a variational approach. Holomorphic coordinates have been used subsequently in several other works, for example [7, 13].

One key feature of this evolution, which led to a very large body of work, is that it admits soliton solutions, which at low frequency/small amplitude are close to the KdV solitons. In the periodic regime these waves are called Stokes waves and there is a continuous family of such waves up to the maximum height wave, which has a profile with a 120 degree angle at the top. As this is only tangentially relevant to the present work, we simply refer the reader to the recent books of Lannes [12] for a good description of the KdV approximation, and of Constantin [4] for the study of solitary waves.

Our goal here is somewhat different, namely to initiate the study of the long time dynamics for the small data problem. As mentioned before, one difficulty in this regard is the presence of the Dirichlet to Neumann map in the Eulerian formulation of the equations. In order to bypass this difficulty we consider the equation in holomorphic coordinates, using a conformal map of the fluid domain into a flat strip. This strategy was previously implemented by the last two authors is several deep water scenarios, namely for gravity waves [8], capillary waves [10] and constant vorticity gravity waves [11].

As this is the first article fully developing the holomorphic coordinates in the finite bottom scenario, in the first part of the paper we carefully present the functional setting for our problem, and then derive the corresponding formulation for the water wave equations in this setting. In this article we only consider the case of the infinite strip. However, the periodic case is equally interesting, and has received perhaps more attention in the literature over the years.

In the holomorphic setting the coordinates are denoted by $\alpha + i\beta \in S := \mathbb{R} \times (-h, 0)$, and the fluid domain is parametrized using the conformal map

$$z : S \to \Omega(t),$$

which takes the bottom $\mathbb{R} - ih$ into the bottom, and the top $\mathbb{R}$ into the top $\Gamma(t)$. As such, the restriction to the real line $Z(\alpha) = z(\alpha - i0)$ can be viewed as a parametrization of the free boundary $\Gamma(t)$.

Our variables are the function $Z(\alpha) = \alpha + W(\alpha)$, which parameterizes the free surface, and the trace $Q(\alpha)$ of the holomorphic velocity potential on the free surface. Both (W, Q) are what we call here holomorphic functions, i.e., the trace on the upper boundary $\beta = 0$ of holomorphic functions in the strip S, which are purely real in the lower boundary $\beta = -1$. The space of holomorphic functions is a real algebra.

To algebraically describe the space of holomorphic functions we use the operator $\mathcal{T}_h$, which is the finite bottom analogue of the Hilbert transform arising in the description of the Dirichlet to Neumann map in the canonical domain. Precisely, $\mathcal{T}_h$ is the multiplier with symbol $-i\tanh(h\xi)$ and real kernel $-\frac{1}{2h}\operatorname{cosech}(\frac{\pi}{2h}\alpha)$, interpreted in the principal value sense. Then the holomorphic functions are described by the relation

$$\operatorname{Im} u = -\mathcal{T}_h \operatorname{Re} u.$$

The complex conjugates of holomorphic functions will be called antiholomorphic functions, and are described by the relation $\operatorname{Im} u = \mathcal{T}_h \operatorname{Re} u$. Arbitrary functions can be expressed as sums of holomorphic and antiholomorphic functions,

$$u = \mathbf{P}_h u + \bar{\mathbf{P}}_h u.$$

Here $\mathbf{P}_h$ projects onto the space of holomorphic functions and its complement $\bar{\mathbf{P}}_h = I - \mathbf{P}_h$ projects onto the space of antiholomorphic functions. Both can be viewed as orthogonal projections in the Hilbert space $\mathfrak{H}_h$ with inner product

$$\langle u, v\rangle_{\mathfrak{H}_h} = \int (\mathcal{T}_h \operatorname{Re} u \cdot \mathcal{T}_h \operatorname{Re} v + \operatorname{Im} u \cdot \operatorname{Im} v)\, d\alpha.$$

We note that $\mathfrak{H}_h$ is not a space of distributions as the $\mathfrak{H}_h$ norm does not see real constants. However, it can be viewed as a quotient space of distributions modulo real constants.

The water wave equations in holomorphic coordinates, derived in Section 3, have the form

$$\begin{cases} W_t + F(1 + W_\alpha) = 0 \\ Q_t + FQ_\alpha - g\mathcal{T}_h[W] + \mathbf{P}_h\left[\dfrac{|Q_\alpha|^2}{J}\right] = 0, \end{cases} \tag{1.6}$$

where

$$J = |1 + W_\alpha|^2, \qquad F = \mathbf{P}_h\left[\frac{Q_\alpha - \bar{Q}_\alpha}{J}\right].$$

We note here that one can freely add real constants to both W and Q; thus these equations are consistent with the low frequency structure of the space $\mathfrak{H}_h$.

The above system has a Hamiltonian structure. The Hamiltonian is the total energy of the system, which is closely related to the above inner product,

$$\mathcal{E} = \frac{g}{4}\langle W, W\rangle - \frac{1}{4}\langle Q, \mathcal{T}_h^{-1}[Q_\alpha]\rangle + \frac{g}{2}\langle WW_\alpha, W\rangle.$$

As written it is not immediately obvious that at low frequency the last term can be controlled by the $\mathfrak{H}$ norm of W. However, a direct computation shows that the Hamiltonian can be expressed in the form

$$\mathcal{E} = \frac{g}{4}\langle W, W\rangle - \frac{1}{4}\langle Q, \mathcal{T}_h^{-1}[Q_\alpha]\rangle + \frac{g}{2}\int |\operatorname{Im} W|^2 \operatorname{Re} W_\alpha \, d\alpha. \tag{1.7}$$

Here one can also see that $\mathcal{E}$ remains positive definite for as long as the curve $\Gamma(t)$ (i.e., the range of $W + \alpha$) remains non-intersecting.

For later use here it is will be useful to symmetrize the Q part of the above energy by introducing the positive self-adjoint operator

$$L_h = (-\mathcal{T}_h^{-1}\partial_\alpha)^{\frac{1}{2}},$$

so that the quadratic part of the energy is given by

$$E_0(w, r) := g\langle w, w\rangle + \langle L_h r, L_h r\rangle.$$

It is then natural to look for solutions (W, Q) in the Sobolev space $\mathcal{H}_h$ with norm

$$\|(W, Q)\|^2_{\mathcal{H}_h} := g\|W\|^2_{\mathfrak{H}_h} + \|L_h Q\|^2_{\mathfrak{H}_h},$$

which is similar to $\mathfrak{H}_h$ for both components at low frequency, and to $L^2 \times \dot{H}^{\frac{1}{2}}$ at high frequency.

For higher regularity we will use the spaces $\mathcal{H}_h^k = \langle D\rangle_h^{-k}\mathcal{H}_h$, where $\langle D\rangle_h = h^{-1}\langle hD\rangle$. However, these will not be applied directly to (W, Q). This is for the same reasons as in our previous work [8], namely that, after differentiation, the system for (W, Q) has a degenerate hyperbolic structure, so one needs to diagonalize it and work with diagonal variables instead. This is a well known feature of the water wave equation, and we refer the reader to [1] and [12] for the Eulerian version of this diagonalization, which is often carried out in a paradifferential fashion. In our case, as in [8], a convenient choice for the diagonal variables is given by

$$(\mathbf{W}, R) := \left(W_\alpha, \frac{Q_\alpha}{1 + W_\alpha}\right).$$

These are also physical variables that describe the slope of the free surface (given by $1+\mathbf{W}$), respectively the fluid velocity of the free surface.

Indeed, after differentiation one obtains a self-contained diagonal system in $(\mathbf{W}, R)$:

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \dfrac{1+\mathbf{W}}{1+\bar{\mathbf{W}}} R_\alpha = (1+\mathbf{W})M \\ R_t + bR_\alpha = i\dfrac{g\mathbf{W} - \mathfrak{a}}{1+\mathbf{W}}, \end{cases} \tag{1.8}$$

where the double speed (advection velocity) b is given by

$$b = 2\operatorname{Re}\left[R - \mathbf{P}_h[R\bar{Y}]\right], \qquad Y = \frac{\mathbf{W}}{1+\mathbf{W}}. \tag{1.9}$$

The other (real) parameters $\mathfrak{a}$ and M above are given by

$$\mathfrak{a} = 2\operatorname{Im}\mathbf{P}_h[R\bar{R}_\alpha] + g(1+\mathcal{T}_h^2)\operatorname{Re}\mathbf{W}, \tag{1.10}$$

$$M = 2\operatorname{Re}\mathbf{P}_h[R\bar{Y}_\alpha - \bar{R}_\alpha Y]. \tag{1.11}$$

The parameter $\mathfrak{a}$ also has a physical interpretation, in that $g + \mathfrak{a}$ is the normal derivative of the pressure on the free surface. It will be informative to write it in the form

$$\mathfrak{a} = a + a_1,$$

where the quadratic term

$$a := 2\,\mathrm{Im}\,\mathbf{P}_h[R\bar{R}_\alpha],$$

remains in the infinite depth limit (see [8]) whereas the linear term

$$a_1 := g(1+\mathcal{T}_h^2)\,\mathrm{Re}\,\mathbf{W},$$

is solely a feature of the finite depth case. The positivity of $g+\mathfrak{a}$ is also critical as a necessary well-posedness condition for the above system (the Taylor stability condition):

$$-\frac{\partial p}{\partial \nu}\bigg|_{\Gamma(t)} = \frac{g+\mathfrak{a}}{J} > 0. \tag{1.12}$$

The necessity of this condition is not immediately clear from the form of the system (1.8) above, as this is still a quasilinear system. However, it will become clear once we consider the linearized system in Section 5. In Section 3.4 we prove that this positivity condition remains satisfied as long as the free surface $\Gamma(t)$ remains a positive distance above the bottom; this provides an alternate, Fourier-based proof, of the similar result obtained in [12] in the Eulerian setting using a maximum principle based argument. Further, our proof does not depend on the fact that $\Gamma(t)$ is non-intersecting.

In the sequel we will consider solutions (W,Q) for the system (1.6) with the regularity properties

$$(W,Q) \in \mathcal{H}_h, \qquad (\mathbf{W},R) \in \mathcal{H}_h^k, \qquad k \geq 1.$$

To describe the lifespan of these solutions we introduce two control norms, namely

$$A := \|\mathbf{W}\|_{L^\infty} + \|Y\|_{L^\infty} + g^{-\frac{1}{2}}\|\langle D\rangle_h^{\frac{1}{2}} R\|_{L^\infty \cap B_2^{0,\infty}}, \tag{1.13}$$

respectively

$$B := g^{\frac{1}{2}}\|\langle D\rangle_h^{\frac{1}{2}}\mathbf{W}\|_{\mathrm{bmo}_h} + \|\langle D\rangle_h R\|_{\mathrm{bmo}_h}, \tag{1.14}$$

where, decomposing $f = f_{<h^{-1}} + f_{\geq h^{-1}}$ by frequency, the inhomogeneous space bmo_h is given by the norm

$$\|f\|_{\mathrm{bmo}_h} = \|f_{<h^{-1}}\|_{L^\infty} + \|f_{\geq h^{-1}}\|_{\mathrm{BMO}},$$

where BMO is the usual space of functions of bounded mean oscillation.

At high frequencies (i.e., larger than h^{-1}), these norms coincide with the norms in [8]. Here at least for small data A and B are controlled by the corresponding Sobolev norms of (W,Q) and $(\mathbf{W},R)$ as follows:

$$A \lesssim g^{-\frac{1}{2}}\left(\|(\mathbf{W},R)\|_{\mathcal{H}_h} + h^{-1}\|(W,Q)\|_{\mathcal{H}_h}\right)^{\frac{1}{2}}\left(\|(\mathbf{W}_\alpha,R_\alpha)\|_{\mathcal{H}_h} + h^{-1}\|(\mathbf{W},R)\|_{\mathcal{H}_h}\right)^{\frac{1}{2}}, \tag{1.15}$$

$$B \lesssim \|(\mathbf{W}_\alpha,R_\alpha)\|_{\mathcal{H}_h} + h^{-1}\|(\mathbf{W},R)\|_{\mathcal{H}_h} + h^{-2}\|(W,Q)\|_{\mathcal{H}_h}. \tag{1.16}$$

For large data some additional care is required due to the need to independently control Y uniformly in L^∞.

Before discussing well-posedness, we remark that as stated the problem (1.6) does not have unique solutions due to the gauge freedom

$$(W(t,\alpha),Q(t,\alpha)) \to (W(t,\alpha+\alpha_0(t))+\alpha_0(t), Q(t,\alpha+\alpha_0(t))+q_0(t)),$$

which corresponds to $F \to F + \alpha_0'(t)$ and a similar choice involving $q_0'(t)$ for the projector in the second equation.

At the initial time we cannot do more than make an arbitrary choice (unless we assume more decay at infinity for the initial data). However, we can fix the choice of α_0 and q_0 at later times by requiring that both F and the projector in the second equation have limit 0 at $-\infty$. This is allowed because the arguments of $\mathbf{P}_h$ are not only in L^2, but also in L^1.

Now we can state our local well-posedness result:

Theorem 1.
a) The system (1.6) *is locally well-posed for all initial data* (W_0, Q_0) *with regularity*

$$(W_0, Q_0) \in \mathcal{H}_h, \qquad (\mathbf{W}_0, R_0) \in \mathcal{H}_h^1, \qquad Y_0 \in L^\infty.$$

Further, the solutions can be continued as long as our control parameter $A(t)$ *remains finite, and* $\int B(t)\,dt$ *remains finite.*

b) This result is uniform with respect to our choice of parameters $g \lesssim h$ *as follows. If for a large parameter* C *the initial data satisfies*

$$g^{-1}h^{-1}\|(W_0, Q_0)\|_{\mathcal{H}_h} + g^{-1}\|(\mathbf{W}_0, R_0)\|_{\mathcal{H}_h} + \|(\mathbf{W}_{0,\alpha}, R_{0,\alpha})\|_{\mathcal{H}_h} + \|Y_0\|_{L^\infty} \leq C,$$

then there exists some $T = T(C)$*, independent of* g, h*, so that the solution exists on* $[-T, T]$ *with similar bounds.*

Here well-posedness should be interpreted in the sense of Hadamard as follows:

- Existence of solutions $(W, Q) \in C([-T, T]; \mathcal{H}_h)$, $(\mathbf{W}, R) \in C([-T, T]; \mathcal{H}_h^1)$.
- Uniqueness of solutions in the same class.
- Continuous dependence on the initial data in the same topology.
- Higher regularity: If the initial data has additional regularity (e.g. $\mathcal{H}_h^k$) then the solution has additional regularity as well.

Our second goal is to establish lifespan bounds for the small data problem. Given a generic quasilinear problem with data of size ϵ and quadratic interactions, the standard result is to obtain quadratic lifespan bounds, i.e., $T_{max} \gtrsim \epsilon^{-1}$. Here we show that for our problem, despite the presence of quadratic interactions, the lifespan is nevertheless cubic, i.e., $T_{max} \gtrsim \epsilon^{-2}$.

Theorem 2. *Consider the system* (1.6) *with small initial data* (W_0, Q_0),

$$g^{-1}h^{-1}\|(W, Q)(0)\|_{\mathcal{H}_h} + g^{-1}\|(\mathbf{W}, R)(0)\|_{\mathcal{H}_h} + \|(\mathbf{W}_\alpha, R_\alpha)(0)\|_{\mathcal{H}_h} \leq \epsilon.$$

Then the solution (W, Q) *exists and satisfies similar bounds on a time interval* $[-T_\epsilon, T_\epsilon]$ *with* $T_\epsilon \gtrsim \epsilon^{-2}$*. In addition, higher regularity also propagates uniformly on the same scale, i.e., for solutions as above we have*

$$\|(\mathbf{W}, R)\|_{C([-T_\epsilon, T_\epsilon]; \mathcal{H}_h^k)} \lesssim \|(\mathbf{W}, R)(0)\|_{\mathcal{H}_h^k} + \epsilon h^{1-k},$$

whenever the right hand side is finite.

We emphasize again that our results are uniform in the range of parameters $g \lesssim h$. In particular in the infinite depth limit it agrees with the result in [8]. Furthermore, our setting and our results are also invariant with respect to the scaling

$$(W(t,x), Q(t,x)) \to (\lambda^{-1}W(t, \lambda x), \lambda^{-1}Q(t, \lambda x)), \tag{1.17}$$

which corresponds to our parameters changing according to the law

$$(g, h) \to (\lambda g, \lambda h).$$

Because of this, in the proofs we can freely fix one of the parameters. Precisely, after deriving the equations we choose to fix $h = 1$ and work with g in the range $g \lesssim 1$. We will also write $\mathcal{T} = \mathcal{T}_1$ and similarly for other operators and function spaces.

We remark that one can also rescale time for a second degree of freedom in the choice of the parameters g and h. However, our results are not invariant with respect to this second scaling.

This result formally mirrors earlier results of the last two authors in [8] (together with John Hunter) in the infinite bottom case, as well as [10] for infinite bottom capillary waves, and [11] for constant vorticity gravity waves in deep water.

A common idea in all these papers is the use of the "quasilinear modified energy method," first introduced in [9], in order to establish long time bounds. This can be viewed as proxy for Shatah's normal form method [16], which cannot be directly implemented in quasilinear problems. Instead of correcting the quadratic terms in the equation via a normal form transformation, the basis of our "quasilinear modified energy method" is the idea that one can more readily modify the energy functional.

Despite the formal similarity to [8], the analysis here is considerably more difficult due to several crucial differences. In the infinite bottom case the null condition for resonant quadratic interactions is satisfied in a stronger form, i.e., the normal form transformation is nonsingular at zero frequency. Consequently, we are also able to obtain long time bounds for the linearized equation, and implicitly for the differences of solutions. By contrast, only short time bounds for the linearized equation are obtained in the present paper.

Another key difference between the two problems has to do with the existence of solitons, i.e., localized traveling waves. While the infinite bottom problem admits no solitons, in the finite bottom problem there are small solitons. This is most readily seen via the KdV approximation at low frequencies, which is widely discussed in the literature, see e.g. Lannes's book [12]. While these solitons do not play a significant role in the present paper, they are expected to be essential elements of any investigation of the nonlinear shallow water dynamics on any longer time scales.

We note that without surface tension these difficulties are essentially unique to the $2d$ problem and that the additional dispersion in $3d$ makes the analysis somewhat more straightforward. In the $3d$ case both enhanced lifespan bounds [2] and global well-posedness for small, smooth, localized initial data [17, 18] have been established previously.

We conclude the introduction with a brief overview of the paper. We begin in the next section with a detailed description of the conformal coordinates, as well the corresponding spaces of "holomorphic functions" where the evolution takes place. The fully nonlinear water wave system (1.6) is derived in Section 3, together with the differentiated quasilinear system (1.8). We also discuss the Hamiltonian formalism there, as well as the Taylor stability condition 1.12.

In Section 4 we study a model linear problem, which captures the quasilinear effects in our problem, but not the quadratic semilinear interactions. We will subsequently apply the estimates established here to both the linearized and differentiated equations.

The linearized problem is studied in Section 5. Unlike in our prior work on the infinite bottom problem, here we are only able to prove quadratic and not cubic bounds for the linearization. Thus, the estimates here are only useful for local well-posedness and not for the cubic lifespan result.

The study of the long time dynamics begins in the earnest in Section 6 with the normal form computation. As one can see there, the resonant interactions at zero frequency produce a zero frequency singularity in the the normal form transformation; thus one cannot use it directly even in the low frequency analysis. In Section 7 we compute the associated normal form energy, where repeated symmetrizations lead to cancellations of the singular part. This is the first step in the implementation of our modified energy method.

In Section 8 we show that the normal form energies admit good quasilinear modifications, which can be used to prove the long time bounds for the solutions. Finally, our main result is proved in the last section.

Many of the more technical estimates in the paper are relegated to the Appendix in order to keep the main arguments more clear and streamlined. This includes a number of Coifman-Meyer type commutator estimates, as well as their consequences for the various parameters in our water wave system.

2. Holomorphic coordinates

2.1. **Holomorphic functions in the canonical domain.** We start by considering solutions to the Laplace equation in the strip $S = \mathbb{R} \times (-h, 0)$ with mixed boundary conditions,

$$\begin{cases} -\Delta u = 0 \qquad \text{in} \;\; S \\ u(\alpha, 0) = f \\ \partial_\beta u(\alpha, -h) = 0. \end{cases} \tag{2.1}$$

The solution may be written in the form

$$u(\alpha, \beta) = \frac{1}{\sqrt{2\pi}} \int p(\xi, \beta) \hat{f}(\xi) e^{i\alpha\xi} \, d\xi,$$

where the Fourier multiplier p is given by

$$p(\xi, \beta) = \frac{\cosh((\beta + h)\xi)}{\cosh(h\xi)}.$$

We note that $p(D, \beta) f$ is well-defined for any $f \in \mathcal{S}'(\mathbb{R})$ and that

$$\partial_\beta^k p(\xi, \beta) = O(|\xi|^k e^{\frac{\beta}{h}\langle h\xi\rangle}).$$

Given a real-valued solution u to (2.1) we may find a harmonic conjugate v by solving the Cauchy-Riemann equations,

$$u_\alpha = v_\beta, \qquad u_\beta = -v_\alpha.$$

A solution is given by

$$v(\alpha, \beta) = \frac{1}{\sqrt{2\pi}} \int q(\xi, \beta) \hat{f}(\xi) e^{i\alpha\xi} \, d\xi,$$

where the Fourier multiplier $q(\xi, \beta)$ is given by

$$q(\xi, \beta) = \frac{i \sinh((\beta + h)\xi)}{\cosh(h\xi)}.$$

On the boundary $\{\beta = 0\}$ we have

$$v(\alpha, 0) = -\mathcal{T}_h f(\alpha),$$

where the Tilbert transform is

$$\mathcal{T}_h f(\alpha) = -\frac{1}{2h} \lim_{\epsilon \downarrow 0} \int_{|\alpha-\alpha'|>\epsilon} \operatorname{cosech}\left(\frac{\pi}{2h}(\alpha-\alpha')\right) f(\alpha')\, d\alpha',$$

is given by the Fourier multiplier $-i\tanh(h\xi)$. We remark that it takes real-valued functions to real-valued functions. We denote the inverse Tilbert transform by $\mathcal{T}_h^{-1}$. As discussed above there is some ambiguity in its definition. For concreteness we define it to be given by the Fourier multiplier $i\coth(h\xi + i0)$ such that $\mathcal{T}_h^{-1} f$ vanishes at $-\infty$ whenever $f \in L^1 \cap L^2$.

We will call functions on the line holomorphic if they are the restriction to the real line of holomorphic functions in the strip and satisfy the boundary condition on the bottom. This consists of functions u which satisfy

$$\operatorname{Im} u = -\mathcal{T}_h \operatorname{Re} u,$$

and forms a real algebra as can be seen from a simple application of the product formula

$$u\mathcal{T}_h[v] + \mathcal{T}_h[u]v = \mathcal{T}_h\big[uv - \mathcal{T}_h[u]\mathcal{T}_h[v]\big], \tag{2.2}$$

which follows from the corresponding identity for $\tanh\xi$. The complex conjugates of holomorphic functions are called antiholomorphic.

2.2. **Sobolev spaces.** On the space of all complex valued functions we define the real inner product

$$\langle u, v\rangle = \frac{1}{2}\operatorname{Re}\int (1-\mathcal{T}_h^2)u\cdot\bar{v} - (1+\mathcal{T}_h^2)u\cdot v\, d\alpha, \tag{2.3}$$

where we note that $-\mathcal{T}_h^2$ is a non-negative operator. The corresponding Hilbert space is denoted by $\mathfrak{H}_h$. Its norm can be rewritten in the form

$$\|u\|_{\mathfrak{H}_h}^2 = \int \left(\mathcal{T}_h \operatorname{Re} u \cdot \mathcal{T}_h \operatorname{Re} u + \operatorname{Im} u\cdot \operatorname{Im} u\right)\, d\alpha,$$

where one can easily see that this is non-negative, and thus a norm.

We denote by $\mathfrak{H}_h^{(h)}$, respectively $\mathfrak{H}_h^{(a)}$ the subspaces of $\mathfrak{H}_h$ consisting of holomorphic, respectively antiholomorphic functions. The interesting observation, which is in effect the motivation for our introducing the space $\mathfrak{H}_h$, is that its holomorphic and antiholomorphic subspaces are orthogonal complements of each other. We remark that, restricted to either $\mathfrak{H}_h^{(h)}$ or $\mathfrak{H}_h^{(a)}$, the $\mathfrak{H}_h$ norm can be rewritten as

$$\|u\|_{\mathfrak{H}_h}^2 = \int \left(|u|^2 - \frac{1}{2}u^2 - \frac{1}{2}\bar{u}^2\right)\, d\alpha.$$

We will also need the associated orthogonal projections, which are denoted by $\mathbf{P}_h$, respectively $\bar{\mathbf{P}}_h$. These are operators which are conjugated via the standard complex conjugation. We can define these two operators in two equivalent ways. In a real fashion, we can set

$$\mathbf{P}_h u = \frac{1}{2}\left[(1-i\mathcal{T}_h)\operatorname{Re} u + i(1+i\mathcal{T}_h^{-1})\operatorname{Im} u\right],$$

$$\bar{\mathbf{P}}_h u = \frac{1}{2}\left[(1+i\mathcal{T}_h)\operatorname{Re} u + i(1-i\mathcal{T}_h^{-1})\operatorname{Im} u\right].$$

In a complex fashion, we can write

$$\begin{aligned}\mathbf{P}_h u &= \frac{1}{4}\left[(2 - i\mathcal{T}_h + i\mathcal{T}_h^{-1})u - i(\mathcal{T}_h + \mathcal{T}_h^{-1})\bar{u}\right] \\ &= \frac{1}{4}\left[(1 - i\mathcal{T}_h)(1 + i\mathcal{T}_h^{-1})u + (1 - i\mathcal{T}_h)(1 - i\mathcal{T}_h^{-1})\bar{u}\right],\end{aligned}$$

respectively

$$\begin{aligned}\bar{\mathbf{P}}_h u &= \frac{1}{4}\left[(2 + i\mathcal{T}_h - i\mathcal{T}_h^{-1})u + i(\mathcal{T}_h + \mathcal{T}_h^{-1})\bar{u}\right] \\ &= \frac{1}{4}\left[(1 + i\mathcal{T}_h)(1 - i\mathcal{T}_h^{-1})u - (1 - i\mathcal{T}_h)(1 - i\mathcal{T}_h^{-1})\bar{u}\right].\end{aligned}$$

2.3. **Conformal mappings.** Given a fluid domain $\Omega = \Omega(t)$ with upper boundary $\Gamma = \Gamma(t)$ with a prescribed Sobolev regularity, and lower boundary $y = -h$, our goal here is to obtain a conformal map

$$z : S \to \Omega$$

with similar regularity. Here we do not assume that Γ is a graph, only that it is the upper boundary of a simply connected domain Ω which admits a parametrization with a suitable Sobolev regularity. Precisely, we represent the boundary $\Gamma(t)$ as a parametrized curve

$$s \to z(s)$$

with the following properties:

(i) Sobolev regularity: $z(s) - s \in H_h^k := \langle D \rangle_h^{-k} L^2$.
(ii) Nondegenerate and non-intersecting: The map $s \to z(s)$ is surjective and nondegenerate, $z'(s) \neq 0$.
(iii) Does not touch the bottom: $\operatorname{Im} z > -h$.

Then we have:

Theorem 3.
a) Let Ω be a simply connected domain whose lower boundary consists of the line $\operatorname{Im} z = -h$ and whose upper boundary is a curve Γ as above, with $k > \frac{3}{2}$. Then there exists a conformal map

$$z : S \to \Omega$$

taking the line $\beta = -h$ into itself and the line $\beta = 0$ into Γ. Further, the restriction of z to the upper boundary $\beta = 0$ has the regularity $z - \alpha \in \mathfrak{H}_h^k$ and is unique up to horizontal translations.

b) If in addition Γ admits a parametrization which satisfies the smallness condition

$$h^{-\frac{3}{2}}(\|z\|_{L^2} + h^k \|z\|_{H_h^k}) \ll 1,$$

then it is a graph $y = y(x)$ satisfying similar H_h^k bounds, and the following norms are comparable:

$$h^{-j}\|y\|_{L^2} + \|y\|_{H_h^j} \approx h^{-j}\|z - \alpha\|_{L^2} + \|z - \alpha\|_{\mathfrak{H}_h^j}, \qquad 0 \le j \le k.$$

Remark 2.1. We remark here on a minor downside to the use of holomorphic coordinates in the strip, namely that there is no canonical way to remove the horizontal translation symmetry (unless $z(s)-s$ has some L^1 integrability perhaps). We address this issue dynamically in our study of the water wave equations. Precisely, we make an arbitrary choice at the initial time, but we define a unique way to propagate this choice to later times.

Proof. By rescaling it suffices to assume $h = 1$. To clarify the geometric context, we note that the L^2 integrability condition on the parametrization guarantees that outside a compact set, the boundary Γ is the graph of a small H^s function.

It is easier to construct the inverse map

$$\Omega \ni z \to \zeta \in S.$$

For this we begin with the function $\beta(z)$, which is defined as the unique bounded solution to the elliptic boundary value problem

$$\begin{cases} \Delta_{x,y}\beta = 0 \qquad \text{in } \ \Omega \\ \beta(x,-1) = -1 \\ \beta(x,y) = 0 \qquad \text{on } \ \Gamma. \end{cases}$$

Maximum principle type arguments show that β is of class C^1 in Ω, and also that it has no critical points. Since Γ is asymptotically flat, it also easily follows that

$$\lim_{x\to\pm\infty} \nabla\beta(x,y) = (0,1).$$

Once we have the function β, its harmonic conjugate α is determined via the Cauchy-Riemann equations, and satisfies

$$\lim_{x\to\pm\infty} \frac{\alpha(x,y)}{x} = 0.$$

It is clear that α is uniquely determined up to constants.

The generated map $z \to \alpha + i\beta$ will then be a diffeomorphism from Ω to S. It remains to establish the regularity properties of this map restricted to Γ, and then of its inverse.

Our goal here is to show that the map

$$s \mapsto \frac{d\alpha}{ds}$$

(which so far is bounded, continuous and nonzero) has the regularity

$$\frac{d\alpha}{ds} - 1 \in H^{k-1}. \tag{2.4}$$

As $k-1 > \frac{1}{2}$, inverting we also have

$$\frac{ds}{d\alpha} - 1 \in H^{k-1}.$$

Hence by the chain rule we get

$$\frac{dz}{d\alpha} - 1 \in H^{k-1}, \qquad \operatorname{Im} z(\alpha) \in H^k,$$

as desired.

To prove (2.4) we use the Cauchy-Riemann equations to rewrite this in terms of the normal derivative of β, namely

$$\frac{d\alpha}{ds} = \frac{dz}{ds} \cdot \frac{d\beta}{d\nu}.$$

Hence we still need to show that

$$\frac{d\beta}{d\nu} - 1 \in H^{k-1}.$$

The function $\beta - y$ solves the Laplace equation in Ω with H^k Dirichlet data on the top Γ and zero Dirichlet data on the top. In addition, Γ also has H^k regularity (which implies also C^1 as $k \geq \frac{3}{2}$. Then we want its normal derivative on Γ to be in H^{k-1}. But this follows from standard elliptic theory; for an exposition of this which exactly fits the strip type of domains here we refer the reader to Chapter 3 of Lannes's book [12].

□

3. Derivation of the equations

3.1. **Derivation of the fully nonlinear system.** In this section we derive the fully nonlinear system (1.6) from the Euler equations (1.1), and the boundary conditions (1.2), (1.3) and (1.4).

We start by defining the holomorphic function w by

$$w(t, \alpha, \beta) = z(t, \alpha, \beta) - (\alpha + i\beta),$$

where $z = x + iy \colon S \to \Omega(t)$ is the conformal map constructed in Section 2.3 As z is holomorphic we have the Cauchy-Riemann equations

$$x_\alpha = y_\beta, \qquad x_\beta = -y_\alpha.$$

Let $\phi(t, x, y)$ be the velocity potential in Euclidean coordinates and take the potential in holomorphic coordinates to be

$$\psi(t, \alpha, \beta) = \phi(t, x(t, \alpha, \beta), y(t, \alpha, \beta)).$$

We take θ to be the harmonic conjugate of ψ and define $q = \psi + i\theta$. Applying the chain rule, the velocity $u = \nabla\phi$ is given by

$$u = \frac{1}{j}(x_\alpha \psi_\alpha + x_\beta \psi_\beta, y_\alpha \psi_\alpha + y_\beta \psi_\beta), \tag{3.1}$$

where the Jacobian j has the form

$$j = x_\alpha y_\beta - x_\beta y_\alpha = x_\alpha^2 + y_\alpha^2.$$

In this section we will use capital letters to denote the trace of functions on the boundary $\{\beta = 0\}$. In particular, by a slight abuse of notation, we will write $Y(t, \alpha) = y(t, \alpha, 0)$. We then have that $W(t, \alpha) = w(t, \alpha, 0)$ and $Q(t, \alpha) = q(t, \alpha, 0)$ are holomorphic and hence

$$Y = -\mathcal{T}_h[X - \alpha], \qquad Y_\alpha = -\mathcal{T}_h[X_\alpha], \qquad \Theta = -\mathcal{T}_h \Psi. \tag{3.2}$$

We observe that $1 - Z_\alpha^{-1} = \dfrac{W_\alpha}{1 + W_\alpha}$ is holomorphic, so by comparing real and imaginary parts we obtain

$$\frac{Y_\alpha}{J} = \mathcal{T}_h \left[\frac{X_\alpha}{J} - 1 \right] = \mathcal{T}_h \left[\frac{X_\alpha}{J} \right], \tag{3.3}$$

where $J(t,\alpha) = j(t,\alpha,0) = |1+W_\alpha|^2$.

Using the normal $(-Y_\alpha, X_\alpha)$ to the free boundary we write the kinematic boundary condition (1.4) in the form

$$(X_t, Y_t)\cdot(-Y_\alpha, X_\alpha) = U\cdot(-Y_\alpha, X_\alpha),$$

where $U(t,\alpha) = u(t,\alpha,0)$ is the restriction of the velocity to the free boundary. Using the expression (3.1) for the velocity in holomorphic coordinates and the Cauchy-Riemann equations we simplify the right hand side to obtain

$$X_\alpha Y_t - Y_\alpha X_t = -\Theta_\alpha. \tag{3.4}$$

Using (3.2) and (3.3) we write this in the form

$$\frac{X_\alpha}{J}\mathcal{T}_h[X_t] + \mathcal{T}_h\left[\frac{X_\alpha}{J}\right]X_t = \frac{\Theta_\alpha}{J}.$$

Applying the product formula (2.2) to the left-hand side we obtain

$$\mathcal{T}_h\left[\frac{X_\alpha X_t + Y_\alpha Y_t}{J}\right] = \frac{\Theta_\alpha}{J}. \tag{3.5}$$

Combining (3.4) and (3.5) we solve for X_t, Y_t to obtain

$$\begin{cases} X_t = \dfrac{\Theta_\alpha}{J}Y_\alpha + \mathcal{T}_h^{-1}\left[\dfrac{\Theta_\alpha}{J}\right]X_\alpha \\ Y_t = -\dfrac{\Theta_\alpha}{J}X_\alpha + \mathcal{T}_h^{-1}\left[\dfrac{\Theta_\alpha}{J}\right]Y_\alpha. \end{cases}$$

In terms of the holomorphic function $W = (X-\alpha) + iY$ we have

$$W_t = X_t + iY_t = -i(1 + i\mathcal{T}_h^{-1})\left[\frac{\Theta_\alpha}{J}\right](1+W_\alpha).$$

If we define

$$F = \mathbf{P}_h\left[\frac{Q_\alpha - \bar{Q}_\alpha}{J}\right],$$

then we may write this in the form

$$W_t + F(1+W_\alpha) = 0. \tag{3.6}$$

Next we use (1.1) to obtain the Bernoulli equation with dimensionless gravitational constant $g > 0$,

$$\phi_t + \frac{1}{2}|\nabla\phi|^2 + gy + p = 0. \tag{3.7}$$

From the dynamic boundary condition (1.3) we have $p = 0$.

Applying the chain rule and Cauchy-Riemann equations we obtain

$$\phi_t|_{\{\beta=0\}} = \Psi_t - \frac{1}{J}(X_\alpha X_t + Y_\alpha Y_t)\Psi_\alpha - \frac{1}{J}(Y_\alpha X_t - X_\alpha Y_t)\Theta_\alpha,$$
$$\frac{1}{2}|\nabla\phi|^2|_{\{\beta=0\}} = \frac{1}{2J}(\Psi_\alpha^2 + \Theta_\alpha^2).$$

Using the relations (3.4) and (3.5), we simplify the first of these to obtain

$$\phi_t|_{\{\beta=0\}} = \Psi_t - \mathcal{T}_h^{-1}\left[\frac{\Theta_\alpha}{J}\right]\Psi_\alpha - \frac{1}{J}\Theta_\alpha^2.$$

This leads to the equation

$$\Psi_t - \mathcal{T}_h^{-1}\left[\frac{\Theta_\alpha}{J}\right]\Psi_\alpha + \frac{1}{2J}(\Psi_\alpha^2 - \Theta_\alpha^2) + gY = 0.$$

We write this in terms of $Q = \Psi + i\Theta$ by applying $\mathbf{P}_h$ to obtain

$$Q_t - \mathbf{P}_h\left[\mathcal{T}_h^{-1}\left[\frac{\Theta_\alpha}{J}\right]\Psi_\alpha + \frac{\Theta_\alpha^2}{J}\right] + \mathbf{P}_h\left[\frac{1}{2J}(\Psi_\alpha^2 + \Theta_\alpha^2)\right] - g\mathcal{T}_h[W] = 0.$$

An application of the product formula (2.2) gives us

$$\mathcal{T}_h\left[\mathcal{T}_h^{-1}\left[\frac{\Theta_\alpha}{J}\right]\Psi_\alpha + \frac{\Theta_\alpha^2}{J}\right] = \frac{\Theta_\alpha}{J}\Psi_\alpha - \mathcal{T}_h^{-1}\left[\frac{\Theta_\alpha}{J}\right]\Theta_\alpha,$$

which leads to the equation

$$Q_t + FQ_\alpha - g\mathcal{T}_h[W] + \mathbf{P}_h\left[\frac{|Q_\alpha|^2}{J}\right] = 0. \tag{3.8}$$

Combining (3.6) and (3.8) we obtain at the fully nonlinear system (1.6).

3.2. **Symmetries.** Besides the gauge freedom, the system (1.6) has a number of symmetries:

(i) *Translation.* The equations are invariant under time and space translations, for $(t_0, \alpha_0) \in \mathbb{R}^2$

$$(W(t,\alpha), Q(t,\alpha)) \mapsto (W(t+t_0, \alpha+\alpha_0), Q(t+t_0, \alpha+\alpha_0)).$$

(ii) *Reflection.* We have a horizontal reflection symmetry given by

$$(W(t,\alpha), Q(t,\alpha)) \mapsto (-\bar{W}(t,-\alpha), \bar{Q}(t,-\alpha)).$$

(iii) *Time reversal.* We have a time reversal symmetry given by

$$(W(t,\alpha), Q(t,\alpha)) \mapsto (W(-t,\alpha), -Q(-t,\alpha)).$$

(iv) *Galilean invariance.* The system has a Galilean invariance, for $c \in \mathbb{R}$

$$(W(t,\alpha), Q(t,\alpha)) \mapsto (W(t,\alpha-ct), Q(t,\alpha-ct) - c\left((\alpha-ct) + W(t,\alpha-ct)\right) + \frac{1}{2}c^2t).$$

However, as our choice of spaces require R to vanish at $\pm\infty$ we break the Galilean symmetry as in terms of $(\mathbf{W}, R)$ the Galilean shift corresponds to the map

$$(\mathbf{W}(t,\alpha), R(t,\alpha)) \mapsto (\mathbf{W}(t,\alpha-ct), R(t,\alpha-ct) - c).$$

3.3. **Hamiltonian structure and conserved quantities.** If the free surface is given by $y = \eta(x)$, then the energy of the system in Euclidean coordinates is given by

$$\mathcal{E}(\eta, \phi) = \frac{g}{2} \int_{\mathbb{R}} |\eta|^2 \, dx + \frac{1}{2} \int_{\mathbb{R}} \int_{-h}^{\eta(x)} |\nabla \phi|^2 \, dy dx.$$

We may write this in terms of the holomorphic variables (W, Q) as

$$\mathcal{E}(W, Q) = \frac{g}{4} \langle W, W \rangle - \frac{1}{4} \langle Q, \mathcal{T}_h^{-1}[Q_\alpha] \rangle + \frac{g}{2} \langle W W_\alpha, W \rangle.$$

We note that the additional factor of $\frac{1}{2}$ appears here due to the use of the complex-valued functions.

It was first observed by Zakharov [21] that the water wave system is a Hamiltonian equation with Hamiltonian $\mathcal{E}$. To see this we consider the space of holomorphic functions $(W, Q) \in \mathcal{H}_h$ equipped with the inner product

$$\left\langle \begin{bmatrix} W_1 \\ Q_1 \end{bmatrix}, \begin{bmatrix} W_2 \\ Q_2 \end{bmatrix} \right\rangle := \frac{g}{2} \langle W_1, W_2 \rangle + \frac{1}{2} \langle L_h Q_1, L_h Q_2 \rangle. \tag{3.9}$$

With respect to this inner product we have

$$d\mathcal{E}(W, Q) = \begin{bmatrix} W + W W_\alpha - \mathcal{T}_h^{-1} \mathbf{P}_h \left[\bar{W} \mathcal{T}_h[W_\alpha] \right] \\ Q \end{bmatrix}.$$

We claim that the system (1.6) may then be written in the form

$$\begin{bmatrix} W_t \\ Q_t \end{bmatrix} = \begin{bmatrix} 0 & \mathfrak{A} \\ \mathfrak{C} & \mathfrak{B} \end{bmatrix} d\mathcal{E}(W, Q), \tag{3.10}$$

where the operators $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ are given by

$$\begin{aligned} \mathfrak{A}[w] &:= -(1 + W_\alpha) \mathbf{P}_h \left[\frac{w_\alpha - \bar{w}_\alpha}{J} \right], \\ \mathfrak{B}[q] &:= -Q_\alpha \mathbf{P}_h \left[\frac{q_\alpha - \bar{q}_\alpha}{J} \right] - \mathbf{P}_h \left[\frac{\mathbf{P}_h[\bar{Q}_\alpha q_\alpha] + \bar{\mathbf{P}}_h[Q_\alpha \bar{q}_\alpha]}{J} \right], \\ \mathfrak{C}[w] &:= g \mathbf{P}_h \left[\frac{\mathbf{P}_h \left[(1 + \bar{W}_\alpha) \mathcal{T}_h[w] \right] + \bar{\mathbf{P}}_h \left[(1 + W_\alpha) \mathcal{T}_h[\bar{w}] \right]}{J} \right]. \end{aligned}$$

Taking $\mathfrak{A}^*$ to be the adjoint of $\mathfrak{A}$ with respect to the inner product on the space of holomorphic functions in $\mathfrak{H}_h$, we apply Lemma A.7 to obtain

$$L_h^2 \mathfrak{C}[w] = -g \mathfrak{A}^*[w], \qquad L_h^2 \mathfrak{B}[q] = -\mathfrak{B}^*[L_h^2 q],$$

and hence the matrix operator

$$\begin{bmatrix} 0 & \mathfrak{A} \\ \mathfrak{C} & \mathfrak{B} \end{bmatrix},$$

is skew-adjoint with respect to the inner product (3.9). This skew-adjoint matrix is the representation in our setting of the symplectic form for the finite bottom system.

We now prove (3.10). We first note that

$$\mathfrak{A}[Q] = -F(1 + W_\alpha), \qquad \mathfrak{B}[Q] = -F Q_\alpha - \mathbf{P}_h \left[\frac{|Q_\alpha|^2}{J} \right].$$

It remains to consider the term involving $\mathfrak{C}$, which we may write in the form

$$g\mathbf{P}_h\left[\frac{\mathbf{P}_h\left[(1+\bar{W}_\alpha)\mathcal{T}_h[w]-W_\alpha\mathcal{T}_h[\bar{w}]\right]}{J}\right],\qquad w=W+WW_\alpha-\mathcal{T}_h^{-1}\mathbf{P}_h\left[\bar{W}\mathcal{T}_h[W_\alpha]\right].$$

Given holomorphic functions u,v we may write them in terms of their real parts and apply the product formula (2.2) to obtain the identity

$$\mathbf{P}_h\big[\mathcal{T}_h[uv]-\bar{u}\mathcal{T}_h[v]-\mathcal{T}_h[\bar{u}]v\big]=\mathcal{T}_h[u]v. \tag{3.11}$$

Taking $u=W$ and $v=W_\alpha$ we may apply this identity to the quadratic part of the numerator to obtain

$$\mathbf{P}_h\left[\mathcal{T}_h[WW_\alpha]-\bar{W}\mathcal{T}_h[W_\alpha]-W_\alpha\mathcal{T}_h[\bar{W}]+\bar{W}_\alpha\mathcal{T}_h[W]\right]=W_\alpha\mathbf{P}_h[\mathcal{T}_h[W]]+\mathbf{P}_h\left[\bar{W}_\alpha\mathcal{T}_h[W]\right].$$

Next we consider the cubic part of the numerator. Here we apply both the identity (3.11) and its complex conjugate with $u=W$ and $v=W_\alpha$ to obtain

$$\mathbf{P}_h\left[\bar{W}_\alpha\mathcal{T}_h[WW_\alpha]-\bar{W}_\alpha\bar{W}\mathcal{T}_h[W_\alpha]-W_\alpha\mathcal{T}_h[\bar{W}\bar{W}_\alpha]+W_\alpha\bar{\mathbf{P}}_h[W\mathcal{T}_h[\bar{W}_\alpha]]\right]=W_\alpha\mathbf{P}_h\left[\bar{W}_\alpha\mathcal{T}_h[W]\right].$$

Combining these with the linear part $\mathbf{P}_h[\mathcal{T}_h[W]]$ we obtain

$$\mathfrak{C}\left[W+WW_\alpha-\mathcal{T}_h^{-1}\mathbf{P}_h[\bar{W}\mathcal{T}_h[W_\alpha]]\right]=g\mathbf{P}_h\left[\frac{(1+W_\alpha)\mathbf{P}_h[(1+\bar{W}_\alpha)\mathcal{T}_h[W]]}{J}\right]=g\mathcal{T}_h[W],$$

where the second equality follows from the fact that $\dfrac{1+W_\alpha}{J}=\dfrac{1}{1+\bar{W}_\alpha}$ is antiholomorphic and hence we may discard the inner projection operator. This completes the proof of (3.10).

As the system is invariant under translation, $\alpha\mapsto\alpha+c$, via Noether's principle there will be a corresponding conserved quantity. This is the horizontal momentum,

$$\mathcal{I}(W,Q)=-\frac{1}{2}\langle L_hW,L_hQ\rangle=\frac{1}{2}\langle W,\mathcal{T}^{-1}Q_\alpha\rangle.$$

With respect to the above inner product on $\mathcal{H}_h$ we have

$$d\mathcal{I}(W,Q)=-\begin{bmatrix}g^{-1}L_h^2Q\\W\end{bmatrix}.$$

A further calculation gives us that

$$\begin{bmatrix}W_\alpha\\Q_\alpha\end{bmatrix}=\begin{bmatrix}0&\mathfrak{A}\\\mathfrak{C}&\mathfrak{B}\end{bmatrix}d\mathcal{I}(W,Q).$$

3.4. **Positivity of the normal derivative of the pressure.** As discussed above, a necessary condition for the well-posedness of (1.6) is the Taylor stability condition (1.12). In this section we first derive the expression for the normal derivative of the pressure in holomorphic coordinates, and then show that it remains positive for as long as the free surface remains a positive distance away from the bottom. We remark that an alternate proof of this property, using the maximum principle, can be found in Lannes [12]. Our proof here, based on a sum of squares representation, provides a different insight into this problem.

From the Bernoulli equation (3.7) we may write the normal derivative of the pressure as

$$-\left.\frac{\partial p}{\partial\nu}\right|_{\Gamma(t)}=\frac{1}{J}\partial_\beta\left(\phi_t+\frac{1}{2}|\nabla\phi|^2+gy\right)\Big|_{\{\beta=0\}}.$$

Using the Cauchy-Riemann equations we obtain

$$\partial_\beta \phi_t|_{\{\beta=0\}} = -\partial_\alpha \left(\mathcal{T}_h \left[\frac{\Psi_\alpha^2 + \Theta_\alpha^2}{2J} \right] + g\mathcal{T}_h[Y] \right).$$

A further application of the Cauchy-Riemann equations yields $g\partial_\beta y|_{\{\beta=0\}} = gX_\alpha$, so

$$-J \left. \frac{\partial p}{\partial \nu} \right|_{\Gamma(t)} = g + \frac{1}{2}(\partial_\beta - \mathcal{T}_h \partial_\alpha)(|\nabla \phi|^2)|_{\{\beta=0\}} + g((X_\alpha - 1) - \mathcal{T}_h[Y_\alpha]).$$

Next we define the holomorphic velocity,

$$r := \phi_x - i\phi_y = \frac{q_\alpha}{1+w_\alpha}. \tag{3.12}$$

From (3.1) we see that $|\nabla \phi|^2 = |r|^2$, and as r is holomorphic we obtain

$$\frac{1}{2} \left. (\partial_\beta - \mathcal{T}_h \partial_\alpha)(|r|^2) \right|_{\{\beta=0\}} = 2 \operatorname{Im} \mathbf{P}_h [R\bar{R}_\alpha] = a.$$

Further, $g\left((X_\alpha - 1) - \mathcal{T}_h[Y_\alpha]\right) = g(1+\mathcal{T}_h^2) \operatorname{Re} \mathbf{W} = a_1$. As a consequence,

$$-J \left. \frac{\partial p}{\partial \nu} \right|_{\Gamma(t)} = g + \mathfrak{a}. \tag{3.13}$$

We now show that the Taylor stability condition (1.12) is satisfied whenever the free surface $\Gamma(t)$ is a positive distance away from the bottom $\{y = -1\}$:

Lemma 3.1. *Assume that $(W, Q) \in \mathcal{H}_h$ are holomorphic, with $\operatorname{Im} W \geq c > -h$. Then we have the pointwise bound*

$$g + \mathfrak{a} \geq g(c+h). \tag{3.14}$$

Proof. Using the spatial scaling discussed in the introduction, it suffices to assume that $h = 1$. We recall the expression of $\mathfrak{a}$,

$$\mathfrak{a} = 2 \operatorname{Im} \mathbf{P}[R\bar{R}_\alpha] + g(1+\mathcal{T}^2) \operatorname{Re} \mathbf{W}.$$

We will consider the two terms separately, and prove that

$$(1+\mathcal{T}^2) \operatorname{Re} \mathbf{W} \geq c, \qquad \operatorname{Im} \mathbf{P}[R\bar{R}_\alpha] \geq 0. \tag{3.15}$$

For the first of these, we write it in terms of $\operatorname{Im} W$ as follows:

$$(1+\mathcal{T}^2) \operatorname{Re} \mathbf{W} = -(1+\mathcal{T}^2)\partial_\alpha \mathcal{T}^{-1}(\operatorname{Im} W).$$

The multiplier on the right has symbol

$$m(\xi) = 2\xi \operatorname{cosech} 2\xi.$$

As a consequence we may write

$$(1+\mathcal{T}^2) \operatorname{Re} \mathbf{W} = \int K(\alpha - \alpha') \operatorname{Im} W(\alpha')\, d\alpha',$$

where

$$K(\alpha) = \frac{1}{\sqrt{2\pi}} \check{m}(\alpha) = \frac{\pi}{8} \operatorname{sech}^2(\frac{\pi}{4}\alpha)$$

is non-negative, Schwartz and has integral 1. Then the first part of (3.15) follows.

For the second part we begin by writing

$$2\,\mathrm{Im}\,\mathbf{P}[R\bar{R}_\alpha] = -\frac{i}{2}\left[(1-i\mathcal{T})(R\bar{R}_\alpha) - (1+i\mathcal{T})(\bar{R}R_\alpha)\right].$$

Hence in Fourier space we have the representation

$$2\,\widehat{\mathrm{Im}\,\mathbf{P}[R\bar{R}_\alpha]}(\zeta) = \int_{\xi-\eta=\zeta} \hat{R}(\xi)\hat{\bar{R}}(\eta)K(\xi,\eta)\,d\eta,$$

where the kernel K is given by

$$K(\xi,\eta) = -\frac{1}{2}\left((\xi+\eta) + (\xi-\eta)\tanh(\xi-\eta)\right).$$

As in [8], a natural idea might be to obtain a sum of squares representation of the above integral. Naively, we could seek a decomposition of the kernel as below

$$K(\xi,\eta) = \int f_N(\xi)f_N(\eta)\,dN.$$

However, here we have the additional information that R is holomorphic, which naively allows us to estimate integrals mostly concentrated where $\xi,\eta > 0$ by symmetric integrals concentrated where $\xi,\eta < 0$. To eliminate this constraint we write everything in terms of the real part of R, which is an arbitrary function:

$$\hat{R}(\xi) = (1-\tanh\xi)\widehat{\mathrm{Re}\,R}(\xi).$$

Then the kernel K is replaced by

$$K_1(\xi,\eta) = (1-\tanh\xi)(1-\tanh\eta)K(\xi,\eta).$$

Further, for real functions we have the symmetry

$$\hat{f}(-\xi) = \hat{\bar{f}}(\xi),$$

so the above kernel can be further replaced by

$$K_2(\xi,\eta) = \frac{1}{2}(K_1(\xi.\eta) + K_1(-\xi,-\eta)).$$

We compute

$$\begin{aligned} K_2(\xi,\eta) &= \frac{1}{2}(\tanh\xi+\tanh\eta)(\xi+\eta) - \frac{1}{2}(1+\tanh\xi\tanh\eta)(\xi-\eta)\tanh(\xi-\eta) \\ &= \tanh\xi\tanh\eta\left(\frac{\xi}{\tanh\xi} + \frac{\eta}{\tanh\eta} - (\xi-\eta)\tanh(\xi-\eta)\right). \end{aligned}$$

On the other hand the following expression gives the symbol of a pointwise non-negative form

$$\begin{aligned} I(\xi,\eta) &:= \int (1+\tanh N)(1+\tanh(\xi-N))(1+\tanh(\eta-N))\,dN \\ &= \xi\left(1+\frac{1}{\tanh\xi}\right)\left(1+\frac{1}{\tanh(\eta-\xi)}\right) + \eta\left(1+\frac{1}{\tanh\eta}\right)\left(1+\frac{1}{\tanh(\xi-\eta)}\right), \end{aligned}$$

and after symmetrization

$$\frac{1}{2}\left(I(\xi,\eta) + I(-\xi,-\eta)\right) = \frac{\xi}{\tanh\xi} + \frac{\eta}{\tanh\eta} - \frac{\xi-\eta}{\tanh(\xi-\eta)}.$$

Then we can write

$$K_2(\xi,\eta) = \frac{1}{2}\tanh\xi\tanh\eta\,(I(\xi,\eta)+I(-\xi,-\eta)) + \tanh\xi\tanh\eta\,K_3(\xi,\eta),$$

where

$$K_3(\xi,\eta) = K_3(\xi-\eta) := 2(\xi-\eta)\operatorname{cosech}(2(\xi-\eta)).$$

The quadratic form determined by the first term in K_2 is non-negative. On the other hand for the second term we take advantage of its translation invariance to write

$$K_3(\xi-\eta) = \int g(\xi-N)g(\eta-N)\,dN,$$

with even, real-valued g. Indeed, taking the Fourier transform we get

$$\hat{g}^2 = \frac{1}{\sqrt{2\pi}}\hat{K}_3,$$

or equivalently

$$\hat{g}(\alpha)^2 = \frac{\pi}{8}\operatorname{sech}^2(\frac{\pi}{4}\alpha).$$

The right hand side is non-negative and its square root is a Schwartz function. This suffices for our purposes, and yields the desired representation for K_3.

□

3.5. **Derivation of the quasilinear system.** In this section we derive the quasilinear system (1.8) for the holomorphic variables

$$(\mathbf{W}, R) = \left(W_\alpha, \frac{Q_\alpha}{1+W_\alpha}\right),$$

where we recall that R is the restriction of the holomorphic velocity (3.12) to the free boundary.

As we expect mixed holomorphic-antiholomorphic terms to be lower order than purely holomorphic terms, we first introduce the (real-valued) advection velocity

$$b = 2\operatorname{Re}\mathbf{P}_h\left[\frac{Q_\alpha}{J}\right],$$

so that $F = b - \dfrac{\bar{Q}_\alpha}{J}$. We note that our earlier gauge fixing procedure corresponds to fixing the real constant in b so that

$$\lim_{\alpha\to-\infty} b(t,\alpha) = 0.$$

Evidently the similar condition at positive infinity does not need to hold.

Differentiating (1.6) we obtain a self-contained quasilinear system in (W_α, Q_α),

$$\begin{cases} W_{\alpha t} + bW_{\alpha\alpha} + b_\alpha(1+W_\alpha) = \left[\dfrac{\bar{Q}_\alpha}{1+\bar{W}_\alpha}\right]_\alpha \\ Q_{\alpha t} + bQ_{\alpha\alpha} + b_\alpha Q_\alpha - g\mathcal{T}_h[W_\alpha] = \bar{\mathbf{P}}_h\left[\dfrac{|Q_\alpha|^2}{J}\right]_\alpha. \end{cases}$$

As $b=\bar{b}$ we have

$$b_\alpha = \bar{F}_\alpha + \frac{1}{J}\left(Q_{\alpha\alpha} - \frac{Q_\alpha W_{\alpha\alpha}}{1+W_\alpha} - \frac{Q_\alpha\bar{W}_{\alpha\alpha}}{1+\bar{W}_\alpha}\right).$$

Grouping the highest order terms on the left hand side we obtain

$$\begin{cases} W_{\alpha t} + bW_{\alpha\alpha} + \dfrac{1}{1+\bar{W}_\alpha}\left(Q_{\alpha\alpha} - \dfrac{Q_\alpha W_{\alpha\alpha}}{1+W_\alpha}\right) = -\bar{F}_\alpha(1+W_\alpha) + \dfrac{Q_\alpha \bar{W}_{\alpha\alpha}}{(1+\bar{W}_\alpha)^2} + \left[\dfrac{\bar{Q}_\alpha}{1+\bar{W}_\alpha}\right]_\alpha \\ Q_{\alpha t} + bQ_{\alpha\alpha} + \dfrac{Q_\alpha}{J}\left(Q_{\alpha\alpha} - \dfrac{Q_\alpha W_{\alpha\alpha}}{1+W_\alpha}\right) - g\mathcal{T}_h[W_\alpha] = -\bar{F}_\alpha Q_\alpha + \dfrac{Q_\alpha^2 \bar{W}_{\alpha\alpha}}{J(1+\bar{W}_\alpha)} + \bar{\mathbf{P}}_h\left[\dfrac{|Q_\alpha|^2}{J}\right]_\alpha. \end{cases}$$

As in the infinite depth case, in order to obtain favorable estimates at high frequency we must diagonalize this system. To do this we define the operator

$$\mathbf{A}(w,q) = (w, q - Rw), \tag{3.16}$$

Taking $\mathbf{W} = W_\alpha$ we use the diagonal variables

$$(\mathbf{W}, R) = \mathbf{A}(W_\alpha, Q_\alpha).$$

We calculate

$$R_\alpha = \frac{1}{1+W_\alpha}\left(Q_{\alpha\alpha} - \frac{Q_\alpha W_{\alpha\alpha}}{1+W_\alpha}\right),$$

and obtain the equation

$$\mathbf{W}_t + b\mathbf{W}_\alpha + \frac{1+\mathbf{W}}{1+\bar{\mathbf{W}}}R_\alpha = \left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} - b_\alpha\right)(1+\mathbf{W}) + \bar{R}_\alpha. \tag{3.17}$$

Defining

$$M = \frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}} - b_\alpha,$$

we obtain the first part of (1.8).

For the second part of (1.8), we first write

$$Q_{\alpha t} + bQ_{\alpha\alpha} + \frac{1+\mathbf{W}}{1+\bar{\mathbf{W}}}RR_\alpha - g\mathcal{T}_h[\mathbf{W}] = \left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} - b_\alpha\right)(1+\mathbf{W})R + \bar{\mathbf{P}}_h\left[|R|^2\right]_\alpha, \tag{3.18}$$

and then calculate,

$$R_t = \frac{Q_{\alpha t}}{1+W_\alpha} - \frac{Q_\alpha W_{\alpha t}}{(1+W_\alpha)^2}.$$

Thus, using (3.17) and (3.18), we obtain the second part of (1.8).

4. Local well-posedness for a model equation

In this section we will study the local well-posedness for a model equation, which will play a key role, both in the study of the linearized problem in the next section, and in the study of the differentiated equations later on. Here, and for the rest of the paper, we will assume that $h = 1$, which we can do by scaling, and require uniformity with respect to g in the range $g \lesssim 1$.

Our model system has the form

$$\begin{cases} w_t + \mathfrak{M}_b w_\alpha + \mathbf{P}\left[\dfrac{r_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha \mathcal{T}^2 w}{1+\bar{\mathbf{W}}}\right] = G \\ r_t + \mathfrak{M}_b r_\alpha - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[w]}{1+\mathbf{W}}\right] = K, \end{cases} \tag{4.1}$$

where $\mathfrak{M}_b$ is the holomorphic multiplication operator

$$\mathfrak{M}_b f = \mathbf{P}(bf).$$

Here both the unknowns (w,r) and the inhomogeneous terms $(G,K) \in \mathcal{H}$ are holomorphic. The functions $(\mathbf{W},R)$ are solutions to the differentiated system (1.8) and b and $\mathfrak{a}$ are the associated advection velocity, respectively the frequency shift, which are given by the formulas (1.9), (1.10) in terms of $\mathbf{W}$ and R. For convenience we recall the expressions of b, $\mathfrak{a}$ and M below

$$b = 2\operatorname{Re}\left[R - \mathbf{P}[R\bar{Y}]\right],$$

and

$$\mathfrak{a} = g(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W} + 2\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha], \quad M = 2\operatorname{Re}\mathbf{P}[R\bar{Y}_\alpha - \bar{R}_\alpha Y].$$

Notably, in our analysis we will not use at all the Sobolev regularity of $(\mathbf{W},R)$. Instead we will only use the bounds for $(\mathbf{W},R)$ which are available in terms of the uniform control norms A and B. Similarly, for b and $\mathfrak{a}$ we use only the corresponding uniform bounds also in terms of A and B, see Lemmas A.13, A.12 in the Appendix.

A natural energy for this system is given by the quadratic part of the Hamiltonian,

$$E_0(w,r) = g\langle w,w\rangle - \langle r, \mathcal{T}^{-1}[r_\alpha]\rangle \approx_g \|(w,r)\|^2_{\mathcal{H}}. \tag{4.2}$$

However, as the equations above have variable coefficients, we instead work with an adapted energy functional

$$E^{(2)}_{lin}(w,r) = \langle w,w\rangle_{g+\mathfrak{a}} - \langle r, \mathcal{T}^{-1}[r_\alpha]\rangle = \langle w,w\rangle_{g+\mathfrak{a}} + \langle Lr, Lr\rangle, \tag{4.3}$$

where for a real valued weight ω we define the weighted inner product

$$\langle u,v\rangle_\omega = \int (\mathcal{T}\operatorname{Re} u\cdot \mathcal{T}\operatorname{Re} v + \operatorname{Im} u\cdot \operatorname{Im} v)\,\omega\, d\alpha.$$

We note that this inner product retains the orthogonality between holomorphic and antiholomorphic functions, and thus the projectors $\mathbf{P}$ and $\bar{\mathbf{P}}$ continue to play the same role. From the Taylor stability condition (1.12) in Lemma 3.1 , and the upper bound for $\mathfrak{a}$ in Lemma A.13, we have

$$E^{(2)}_{\mathrm{lin}}(w,r) \approx_A E_0(w,r)$$

for as long the fluid stays away from the bottom.

We also need a weighted form of the above energy functional. For a real valued weight ω we define

$$E^{(2)}_{\omega,lin}(w,r) := \langle w,w\rangle_{(g+\mathfrak{a})\omega} + \langle Lr, Lr\rangle_\omega. \tag{4.4}$$

Our main estimate for the model system is as follows:

Proposition 4.1. *Let I be a time interval where A is bounded and $B \in L^1$. Then in I the following properties hold:*

a) The system of equations (4.1) *is well posed in $\mathcal{H}$, and satisfies the estimate*

$$\frac{d}{dt}E^{(2)}_{lin}(w,r) = 2\langle G,w\rangle_{g+\mathfrak{a}} + 2\langle LK, Lr\rangle + O_A(B)E^{(2)}_{lin}(w,r). \tag{4.5}$$

b) Assume in addition that ω is a weight satisfying

$$\|\omega\|_{L^\infty} \le A, \qquad \|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}} \le B, \qquad \|(\partial_t + b\partial_\alpha)\omega\|_{L^\infty} \le B. \tag{4.6}$$

Then we also have

$$\frac{d}{dt}E^{(2)}_{\omega,lin}(w,r) = 2\left\langle G, w\right\rangle_{(g+\mathfrak{a})\omega} + 2\langle LK, Lr\rangle_\omega + O_A(B)E^{(2)}_{lin}(w,r). \tag{4.7}$$

Proof. We note that the bound (4.5) can be viewed as a special case of (4.7), so we will only prove the latter. We start by calculating

$$\begin{aligned}\frac{d}{dt}\left\langle w, w\right\rangle_{(g+\mathfrak{a})\omega} = &\left\langle w, w\right\rangle_{[(g+\mathfrak{a})\omega]_t} - 2\left\langle bw_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega} + 2\left\langle d\mathcal{T}^2 w, w\right\rangle_{(g+\mathfrak{a})\omega} + 2\left\langle G, w\right\rangle_{(g+\mathfrak{a})\omega}\\ &- 2\left\langle (1-\bar{Y})r_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega},\end{aligned}$$

where we define $d := \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}$. We complete the weight of the first term appearing in the expression above to

$$\left\langle w, w\right\rangle_{(\partial_t + b\partial_\alpha)[(g+\mathfrak{a})\omega]}.$$

Using also the relation $b_\alpha = M + 2\operatorname{Re} d$ we separate the above time derivative into

$$\frac{d}{dt}\left\langle w, w\right\rangle_{(g+\mathfrak{a})\omega} = 2\left\langle G, w\right\rangle_{(g+\mathfrak{a})\omega} + D^1_w + D^2_w + D^3_w + D^4_w,$$

where

$$\begin{aligned}D^1_w =\ & \left\langle w, w\right\rangle_{(\partial_t + b\partial_\alpha)[(g+\mathfrak{a})\omega]} - \langle M\mathcal{T}^2 w, w\rangle_{(g+\mathfrak{a})\omega},\\ D^2_w =\ & \left\langle 2i\operatorname{Im} d\mathcal{T}^2 w, w\right\rangle_{(g+\mathfrak{a})\omega},\\ D^3_w =\ & -2\left\langle bw_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega} + \left\langle b_\alpha \mathcal{T}^2 w, w\right\rangle_{(g+\mathfrak{a})\omega} - \left\langle w, w\right\rangle_{b\partial_\alpha[(g+\mathfrak{a})\omega]},\\ D^4_w =\ & -2\left\langle (1-\bar{Y})r_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega}.\end{aligned}$$

In a similar manner we expand the time derivative of the r term as

$$\frac{d}{dt}\left\langle Lr, Lr\right\rangle_\omega = 2\left\langle LK, Lr\right\rangle_\omega + D^1_r + D^2_r + D^3_r,$$

where

$$\begin{aligned}D^1_r =\ & \left\langle Lr, Lr\right\rangle_{(\partial_t + b\partial_\alpha)\omega},\\ D^2_r =\ & -2\left\langle L(br_\alpha), Lr\right\rangle_\omega + \left\langle Lr, Lr\right\rangle_{-b\partial_\alpha\omega},\\ D^3_r =\ & 2\left\langle L\left((1-Y)(g+\mathfrak{a})\mathcal{T}[w]\right), Lr\right\rangle_\omega.\end{aligned}$$

We now successively consider all the terms above:

1. The terms D^1_w and D^1_r are trivially estimated using the pointwise bounds for $\mathfrak{a}$ and its derivatives (see Lemma A.13) and ω, as well as the pointwise bound for M (see Lemma A.15).

2. The term D^2_w is expanded using the definition of our inner product as

$$D^2_w = 2\int -(g+\mathfrak{a})\omega\mathcal{T}(\operatorname{Im} d\,\mathcal{T}^2\operatorname{Im} w)\mathcal{T}\operatorname{Re} w + (g+\mathfrak{a})\omega\operatorname{Im} d\,\mathcal{T}^2\operatorname{Re} w\operatorname{Im} w\, d\alpha.$$

We use the relation $\operatorname{Im} w = -\mathcal{T}\operatorname{Re} w$ to eliminate $\operatorname{Re} w$ and obtain

$$\begin{aligned}D^2_w =\ & -2\int -(g+\mathfrak{a})\omega\mathcal{T}(\operatorname{Im} d\,\mathcal{T}^2\operatorname{Im} w)\operatorname{Im} w + (g+\mathfrak{a})\omega\operatorname{Im} d\,\mathcal{T}\operatorname{Im} w\operatorname{Im} w\, d\alpha\\ =\ & -2\int (g+\mathfrak{a})\omega\operatorname{Im} w\left(-[\mathcal{T},\operatorname{Im} d]\mathcal{T}^2\operatorname{Im} w + \operatorname{Im} d\,(1+\mathcal{T}^2)\mathcal{T}\operatorname{Im} w\right) d\alpha.\end{aligned}$$

Now we use a commutator bound

$$\|[\mathcal{T}, \operatorname{Im} d]\|_{L^2\to L^2} \lesssim \|d\|_{\mathrm{bmo}}$$

for the first term, see (A.9), and a Coifman-Meyer bound for the remaining product

$$\| \operatorname{Im} d\, (1+\mathcal{T}^2)\mathcal{T} \operatorname{Im} w\|_{L^2} \lesssim \| \operatorname{Im} d\|_{\mathrm{bmo}} \| \operatorname{Im} w\|_{L^2},$$

using the fact that the multiplier $1+\mathcal{T}^2$ has a rapidly decaying kernel, via (the dual of) (A.7).

3. The term D^3_w is similarly expanded as

$$\begin{aligned} D^3_w = & \int_{\mathbb{R}} (g+\mathfrak{a})\omega \operatorname{Im} w \left\{ -2\mathcal{T}(b\mathcal{T}^{-1} \operatorname{Im} w_\alpha) + \mathcal{T}(b_\alpha \mathcal{T} \operatorname{Im} w) + 2b \operatorname{Im} w_\alpha \right. \\ & \left. + b_\alpha \mathcal{T}^2 \operatorname{Im} w + 2b_\alpha \operatorname{Im} w \right\} \, d\alpha \\ = & \int_{\mathbb{R}} (g+\mathfrak{a})\omega \operatorname{Im} w \left\{ -2[\mathcal{T}, b]\mathcal{T}^{-1} \operatorname{Im} w_\alpha + [\mathcal{T}, b_\alpha]\mathcal{T} \operatorname{Im} w + 2b_\alpha(1+\mathcal{T}^2) \operatorname{Im} w \right\} \, d\alpha. \end{aligned}$$

To bound the integral above we use the L^∞ bounds for ω and a, together with Hölder's inequality. The desired bounds for this integral are a consequence of the commutator bounds

$$\|[\mathcal{T}, b]\|_{H^{-1}\to L^2} \lesssim \|b_\alpha\|_{\mathrm{bmo}}, \quad \|[\mathcal{T}, b_\alpha]\|_{L^2\to L^2} \lesssim \|b_\alpha\|_{\mathrm{bmo}},$$

which can be found in (A.9).

4. The term D^2_r*.* For simplicity we introduce the holomorphic variable $s := Lr$. Further, we expand

$$-\langle L(br_\alpha), Lr\rangle_\omega = \left\langle L(bL^2\mathcal{T}(r)), Lr\right\rangle_\omega = \langle L(bL\mathcal{T}(s)), s\rangle_\omega .$$

Then

$$\begin{aligned} D^2_r &= 2\int_{\mathbb{R}} \omega \operatorname{Im} s\, (\mathcal{T}LbL + LbL\mathcal{T}) \operatorname{Im} s - b\omega_\alpha (\operatorname{Im} s)^2 \, d\alpha \\ &= 2\int_{\mathbb{R}} \omega \operatorname{Im} s\, (\mathcal{T}LbL + LbL\mathcal{T} + \partial_\alpha b + b\partial_\alpha) \operatorname{Im} s \, d\alpha \\ &= 2\int_{\mathbb{R}} \omega \operatorname{Im} s\, ([\mathcal{T}L, b]L + L[b, L\mathcal{T}]) \operatorname{Im} s \, d\alpha. \end{aligned}$$

Thus we need an L^2 bound for the double commutator

$$\|[[\mathcal{T}L, b], L]\|_{L^2\to L^2} \lesssim \|b_\alpha\|_{\mathrm{bmo}},$$

which is established in the Appendix, see (A.6).

5. The term $D^4_w + D^3_r$*.* This has the form

$$\begin{aligned} D^4_w + D^3_r = & -2\left\langle (1-\bar{Y})r_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega} + 2\left\langle L\left((1-Y)(g+\mathfrak{a})\mathcal{T}[w]\right), Lr\right\rangle_\omega \\ = & \ 2\left\langle (1-\bar{Y})\mathcal{T}Ls, w\right\rangle_{(g+\mathfrak{a})\omega} + 2\left\langle L\left((1-Y)(g+\mathfrak{a})\mathcal{T}[w]\right), s\right\rangle_\omega . \end{aligned}$$

This has some commutator structure, so we expect to get the bound

$$|D^4_w + D^3_r| \lesssim (\|Y\|_{\mathrm{bmo}^{\frac12}} + \|\mathfrak{a}\|_{\mathrm{bmo}^{\frac12}} + \|\omega\|_{\mathrm{bmo}^{\frac12}})\|w\|_{\mathfrak{H}}\|s\|_{\mathfrak{H}},$$

with the implicit constant depending on the L^∞ norm of the same parameters $Y, \mathfrak{a}$ and ω. To see this we divide the analysis into several steps.

First we commute L across ω, and estimate the difference

$$|\langle Lz, s\rangle_\omega - \langle z, Ls\rangle_\omega| \lesssim \|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}}\|z\|_{\mathfrak{H}}\|s\|_{\mathfrak{H}}.$$

Expanding as above, this reduces to the commutator bound (see Lemma A.2)

$$\|[L, \omega]\|_{L^2\to L^2} \lesssim \|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}}.$$

We apply this to $z = (1-Y)(g+\mathfrak{a})\mathcal{T}[w]$ and $s = Lr$. This reduces our problem to estimating the difference

$$-2\left\langle(1-\bar{Y})r_\alpha, w\right\rangle_{(g+\mathfrak{a})\omega} - 2\left\langle(1-Y)(g+\mathfrak{a})\mathcal{T}[w], \mathcal{T}^{-1}r_\alpha\right\rangle_\omega.$$

Next we insert $g+\mathfrak{a}$ inside via the estimate

$$\langle(g+\mathfrak{a})z, w\rangle_\omega - \langle z, w\rangle_{(g+\mathfrak{a})\omega} \lesssim \|g+\mathfrak{a}\|_{\mathrm{bmo}^{\frac{1}{2}}}\|z\|_{H^{-\frac{1}{2}}}\|w\|_{\mathfrak{H}},$$

which reduces to the commutator bound (see (A.9))

$$\|[g+\mathfrak{a}, \mathcal{T}]\|_{H^{-\frac{1}{2}}\to L^2} \lesssim \|g+\mathfrak{a}\|_{\mathrm{bmo}^{\frac{1}{2}}}. \tag{4.8}$$

We apply this with $z = (1-\bar{Y})r_\alpha$ to reduce our problem to estimating the difference

$$-2\left\langle(1-\bar{Y})(g+\mathfrak{a})r_\alpha, w\right\rangle_\omega - 2\left\langle(1-Y)(g+\mathfrak{a})\mathcal{T}[w], \mathcal{T}^{-1}r_\alpha\right\rangle_\omega.$$

Finally, with $e = (1-\bar{Y})(g+\mathfrak{a}) \in \mathrm{bmo}^{\frac{1}{2}}$ and $z = \mathcal{T}^{-1}r_\alpha$, it remains to estimate the difference

$$|\langle e\mathcal{T}z, w\rangle_\omega + \langle z, \bar{e}\mathcal{T}w\rangle_\omega| \lesssim (\|e\|_{L^\infty}\|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}} + \|e\|_{\mathrm{bmo}^{\frac{1}{2}}}\|\omega\|_{L^\infty})\|w\|_{\mathfrak{H}}\|z\|_{\mathfrak{H}^{-\frac{1}{2}}}.$$

This vanishes if ω is constant. Else, writing $e = f + ig$, it reduces to the commutator bounds

$$\|[\omega, \mathcal{T}f + f\mathcal{T}]\|_{H^{-\frac{1}{2}}\to L^2} \lesssim (\|f\|_{L^\infty}\|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}} + \|f\|_{\mathrm{bmo}^{\frac{1}{2}}}\|\omega\|_{L^\infty}),$$

respectively

$$\|[\omega, \mathcal{T}g\mathcal{T}]\| \lesssim (\|f\|_{L^\infty}\|\omega\|_{\mathrm{bmo}^{\frac{1}{2}}} + \|f\|_{\mathrm{bmo}^{\frac{1}{2}}}\|\omega\|_{L^\infty}),$$

which follow by repeated application of bounds of the form (4.8).

□

5. The linearized equation

In this section we first calculate the linearization of (1.6) and then prove that the corresponding linearized system is well-posed in $\mathcal{H}$.

We take the linearized variables at (W, Q) to be $(w, q) = (\delta W, \delta Q)$ and compute

$$\delta R = \frac{q_\alpha - Rw_\alpha}{1+\mathbf{W}}, \qquad \delta F = \mathbf{P}\left[m - \bar{m}\right], \qquad \bar{R}\delta R = n,$$

where we define

$$m = \frac{q_\alpha - Rw_\alpha}{J} + \frac{\bar{R}w_\alpha}{(1+\mathbf{W})^2}, \qquad n = \frac{\bar{R}(q_\alpha - Rw_\alpha)}{1+\mathbf{W}}.$$

We then obtain the linearized equations

$$\begin{cases} w_t + Fw_\alpha + \mathbf{P}[m-\bar{m}](1+\mathbf{W}) = 0 \\ q_t + Fq_\alpha + \mathbf{P}[m-\bar{m}]Q_\alpha - g\mathcal{T}[w] + \mathbf{P}[n+\bar{n}] = 0. \end{cases} \tag{5.1}$$

As $F = b - \dfrac{\bar{R}}{1+\mathbf{W}}$ and $\mathbf{P} = 1 - \bar{\mathbf{P}}$ we may write this system in the form

$$\begin{cases} w_t + b w_\alpha + \dfrac{q_\alpha - R w_\alpha}{1+\bar{\mathbf{W}}} = 2(1+\mathbf{W}) \operatorname{Re} \bar{\mathbf{P}}[m] \\ q_t + b q_\alpha - g\mathcal{T}[w] + \dfrac{R(q_\alpha - R w_\alpha)}{1+\bar{\mathbf{W}}} = 2i \operatorname{Im} \bar{\mathbf{P}}[n] + 2Q_\alpha \operatorname{Re} \bar{\mathbf{P}}[m]. \end{cases} \tag{5.2}$$

This is a degenerate hyperbolic system with a double speed b, so in order to produce good energy estimates at high frequency we introduce diagonal variables. Following [8], a natural choice would be to take $(w, r) = \mathbf{A}(w, q) = (w, q - Rw)$. This would work at high frequencies, but not at low frequencies as we cannot make sense of the product Rw for $w \in \mathfrak{H}$. So instead we work with

$$(w, r) = (w, q + R\mathcal{T}^2 w).$$

We observe that $(w, r) \approx \mathbf{A}(w, q)$ when w is at frequencies $\gg 1$ whereas $(w, r) \approx (w, q)$ when w is at frequencies $\ll 1$.

In terms of the diagonalized variables (w, r) we have

$$m = -\frac{\bar{\mathbf{W}}(r_\alpha - R_\alpha \mathcal{T}^2[w] - R(1+\mathcal{T}^2)[w_\alpha])}{J} + \frac{\bar{R} w_\alpha}{(1+\mathbf{W})^2},$$

$$n = \frac{\bar{R}(r_\alpha - R_\alpha \mathcal{T}^2[w] - R(1+\mathcal{T}^2)[w_\alpha])}{1+\mathbf{W}},$$

where we have harmlessly removed the leading order holomorphic component of m that vanishes after projection to the space of antiholomorphic functions in (5.2). We then obtain the diagonalized system,

$$\begin{cases} w_t + b w_\alpha + \dfrac{r_\alpha}{1+\bar{\mathbf{W}}} - \dfrac{R_\alpha \mathcal{T}^2[w]}{1+\bar{\mathbf{W}}} = \mathcal{G} \\ r_t + b r_\alpha - \dfrac{(g+\mathfrak{a})\mathcal{T}[w]}{1+\mathbf{W}} = \mathcal{K}, \end{cases} \tag{5.3}$$

where

$$\mathcal{G} = 2(1+\mathbf{W}) \operatorname{Re} \bar{\mathbf{P}}[m] + \frac{R(1+\mathcal{T}^2)[w_\alpha]}{1+\bar{\mathbf{W}}},$$

$$\mathcal{K} = 2i \operatorname{Im} \bar{\mathbf{P}}[n] - R[1+\mathcal{T}^2, b] w_\alpha + R(1+\mathcal{T}^2)[w_t + b w_\alpha] + \frac{g\mathbf{W} - \mathfrak{a}}{1+\mathbf{W}}(1+i\mathcal{T})\mathcal{T}[w].$$

Here, for brevity in the notation, we have kept the $w_t + b w_\alpha$ term as a part of $\mathcal{K}$, rather then substituting it from the first equation. This is harmless since $1+\mathcal{T}^2$ has a Schwartz symbol so this term will only play a perturbative role.

While (w, r) are holomorphic, it is not immediately clear that (5.3) preserves the space of holomorphic functions so we apply the projection $\mathbf{P}$ to obtain

$$\begin{cases} w_t + \mathfrak{M}_b w_\alpha + \mathbf{P}\left[\dfrac{r_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha \mathcal{T}^2[w]}{1+\bar{\mathbf{W}}}\right] = \mathbf{P}\mathcal{G} \\ r_t + \mathfrak{M}_b r_\alpha - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[w]}{1+\mathbf{W}}\right] = \mathbf{P}\mathcal{K}, \end{cases} \tag{5.4}$$

which now has the form of the model equation (4.1). Our main result for the linearized system (5.4) is the following Theorem:

Theorem 4. *Suppose that there exists a solution (W, Q) to (1.6) on a time interval $[-T, T]$ such that $(W, Q) \in C([-T, T]; \mathcal{H})$ and $(\mathbf{W}, R) \in C([-T, T]; \mathcal{H}^1)$. Then the linearized equation (5.4) is locally well-posed in $\mathcal{H}$ on the interval $[-T, T]$, and the corresponding solution $(w, r) \in C([-T, T]; \mathcal{H})$ satisfies the estimate*

$$\|(w, r)(t)\|_{\mathcal{H}} \lesssim \exp\left(C \int_0^t \|(g^{\frac{1}{2}}\mathbf{W}, R)(s)\|_{H^1 \times H^{\frac{3}{2}}}\, ds\right) \|(w, r)(0)\|_{\mathcal{H}}, \tag{5.5}$$

where the implicit constant depends only on A and $\sup_{t\in[-T,T]} g^{-\frac{1}{2}} \|(g^{\frac{1}{2}}\mathbf{W}, R)(t)\|_{L^2 \times H^{\frac{1}{2}}}$.

We remark that $(w, q) = (W_\alpha, Q_\alpha)$ is a solution to (5.1), for which we will prove cubic lifespan bounds. Following [8], one might hope to also establish cubic lifespan bounds for small initial data for the linearized system (5.4). Unfortunately this is not the case and we expect that cubic lifespan bounds for the linearized system will fail on account of a breaking of symmetry when $(w, q) \neq (W_\alpha, Q_\alpha)$. One can view this as a reflection of the fact that the quadratic low frequency interactions are stronger here than in the infinite depth case.

In order to prove Theorem 4 it will suffice to obtain a priori estimates for $\|(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})\|_{\mathcal{H}}$ and apply Proposition 4.1. However, in stark contrast to the infinite depth case [8] we will be unable to control $\|(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})\|_{\mathcal{H}}$ only in terms of the pointwise norms A, B and the energy $E^{(2)}_{\text{lin}}(w, r)$. The difficulty arises due to the presence of nonlocal terms in the expression $\operatorname{Re} \bar{\mathbf{P}}[m]$ appearing in both $\mathcal{G}$ and $\mathcal{K}$. Here we will make use of the fact that $\mathbf{P}S_0 \colon L^1 \to L^\infty$, which leads to bounds in terms of the energy norms of $(\mathbf{W}, R)$.

As a consequence, we have the following Proposition:

Proposition 5.1. *We have the estimate*

$$\|(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})\|_{\mathcal{H}} \lesssim_{A, g^{-\frac{1}{2}}\|(g^{\frac{1}{2}}\mathbf{W}, R)\|_{L^2 \times H^{\frac{1}{2}}}} \left(B + \|(g^{\frac{1}{2}}\mathbf{W}, R)\|_{H^{\frac{1}{2}} \times H^1}\right) \|(w, r)\|_{\mathcal{H}}. \tag{5.6}$$

Proof. We decompose

$$\mathcal{G} = G_1 + G_2, \qquad \mathcal{K} = K_1 + K_2 + K_3 + K_4,$$

where,

$$\begin{aligned} G_1 &= 2(1 + \mathbf{W}) \operatorname{Re} \bar{\mathbf{P}}[m], & G_2 &= \frac{R(1 + \mathcal{T}^2)[w_\alpha]}{1 + \bar{\mathbf{W}}}, \\ K_1 &= 2i \operatorname{Im} \bar{P}[n], & K_2 &= -R[1 + \mathcal{T}^2, b]w_\alpha, \\ K_3 &= R(1 + \mathcal{T}^2)(w_t + bw_\alpha), & K_4 &= \frac{g\mathbf{W} - \mathfrak{a}}{1 + \mathbf{W}}(1 + i\mathcal{T})\mathcal{T}[w], \end{aligned}$$

and estimate each term separately.

1. Bounds for G_1. We may estimate

$$\|\mathbf{P}G_1\|_{\mathfrak{H}} \lesssim \|\bar{\mathbf{P}}[m]\|_{\mathfrak{H}} + \|\mathbf{W} \operatorname{Re} \bar{\mathbf{P}}[m]\|_{L^2}.$$

We first prove that

$$\|\bar{\mathbf{P}}[m]\|_{\mathfrak{H}} \lesssim_A g^{-\frac{1}{2}} B \|(w, r)\|_{\mathcal{H}}. \tag{5.7}$$

As $\bar{\mathbf{P}}$ vanishes when applied to holomorphic terms, we write

$$\bar{\mathbf{P}}[m] = -[\bar{\mathbf{P}}, \bar{Y}](1-Y)r_\alpha + [\bar{\mathbf{P}}, d\bar{\mathbf{W}}](1-Y)\mathcal{T}^2[w] \\ + \bar{\mathbf{P}}[\bar{Y}(1-Y)R(1+\mathcal{T}^2)w_\alpha] + [\bar{\mathbf{P}}, \bar{R}](1-Y)^2 w_\alpha.$$

For the first, second and fourth terms we apply the commutator estimate (A.14) and the product estimate (A.15) with the estimate (A.16) for Y and the estimate (A.23) for d. For the third term we simply use that $1+\mathcal{T}^2$ has Schwartz symbol and that $\|R\|_{L^\infty} \lesssim B$.

For the second term in G_1 we first decompose according to the frequency of $\operatorname{Re}\bar{\mathbf{P}}[m]$,

$$\mathbf{W}\operatorname{Re}\bar{\mathbf{P}}[m] = \mathbf{W}P_{\geq 1}\operatorname{Re}\bar{\mathbf{P}}[m] + \mathbf{W}S_0\operatorname{Re}\bar{\mathbf{P}}[m].$$

For the high frequency component we use the estimate (5.7) for $\bar{\mathbf{P}}[m]$ to obtain

$$\|\mathbf{W}P_{\geq 1}\operatorname{Re}\bar{\mathbf{P}}[m]\|_{L^2} \lesssim \|\mathbf{W}\|_{L^\infty}\|\bar{\mathbf{P}}[m]\|_{\mathfrak{H}} \lesssim_A g^{-\frac{1}{2}}AB\|(w,r)\|_{\mathcal{H}}.$$

For the low frequency component we are unable to estimate $S_0\operatorname{Re}\bar{\mathbf{P}}[m]$ in L^2, so instead we estimate

$$\|\mathbf{W}S_0\operatorname{Re}\bar{\mathbf{P}}[m]\|_{L^2} \lesssim \|\mathbf{W}\|_{L^2}\|S_0\operatorname{Re}\bar{\mathbf{P}}[m]\|_{L^\infty}.$$

It then remains to show that

$$\|S_0\operatorname{Re}\bar{\mathbf{P}}[m]\|_{L^\infty} \lesssim_A g^{-\frac{1}{2}}\|(g^{\frac{1}{2}}\mathbf{W}, R)\|_{H^{\frac{1}{2}}\times H^1}\|(w,r)\|_{\mathcal{H}}. \tag{5.8}$$

For the first term in m we use that we use that $S_0\mathbf{P}\colon L^1 \to L^\infty$ to obtain

$$\|S_0\mathbf{P}[\bar{Y}(1-Y)r_\alpha]\|_{L^\infty} \lesssim \|S_0(\bar{Y}(1-Y)r_\alpha)\|_{L^1}.$$

Considering this to be the product of $\bar{Y}(1-Y)$ and r_α we may only have high-high frequency interactions and hence

$$\begin{aligned} \|S_0(\bar{Y}(1-Y)r_\alpha)\|_{L^1} &\lesssim \sum_{k\approx k'} \|\mathbf{P}_k[\bar{Y}(1-Y)]\|_{L^2}\|\mathbf{P}_{k'}[r_\alpha]\|_{L^2} \\ &\lesssim \|\bar{Y}(1-Y)\|_{H^{\frac{1}{2}}}\|Lr\|_{\mathfrak{H}} \\ &\lesssim_A \|\mathbf{W}\|_{H^{\frac{1}{2}}}\|Lr\|_{\mathfrak{H}}, \end{aligned}$$

where the final line follows from the Moser estimate (A.10). For the second and third terms in m we may straightforwardly estimate

$$\left\|\frac{R_\alpha\bar{Y}}{1+\mathbf{W}}\mathcal{T}^2[w]\right\|_{L^1} + \left\|\frac{R\bar{Y}}{1+\mathbf{W}}(1+\mathcal{T}^2)w_\alpha\right\|_{L^1} \lesssim_A \|R\|_{H^1}\|w\|_{\mathfrak{H}}.$$

For the final term in m we consider it to be a product of $\bar{R}$ and $(1-Y)^2 w_\alpha$ to obtain

$$\|S_0(\bar{R}(1-Y)^2 w_\alpha)\|_{L^1} \lesssim \sum_{k\approx k'} \|R_k\|_{L^2}\|\mathbf{P}_k[(1-Y)^2 w_\alpha]\|_{L^2} \lesssim \|R\|_{H^1}\|(1-Y)^2 w_\alpha\|_{H^{-1}}.$$

The estimate (5.8) then follows from the product estimate (A.15).

2. Bounds for G_2. Here we simply use that $1+\mathcal{T}^2$ has Schwartz symbol to obtain

$$\left\|\mathbf{P}\left[\frac{R(1+\mathcal{T}^2)w_\alpha}{1+\bar{\mathbf{W}}}\right]\right\|_{\mathfrak{H}} \lesssim_A B\|w\|_{\mathfrak{H}}.$$

3. Bounds for K_1. As K_1 is purely imaginary we have

$$\|L\mathbf{P}[i\operatorname{Im}\bar{\mathbf{P}}[n]]\|_{\mathfrak{H}} \lesssim \|L\bar{\mathbf{P}}[n]\|_{\mathfrak{H}}.$$

We then write

$$\bar{\mathbf{P}}[n] = [\bar{\mathbf{P}}, \bar{R}](1-Y)r_\alpha - [\bar{\mathbf{P}}, \bar{R}](1-Y)R_\alpha \mathcal{T}^2[w] - \bar{\mathbf{P}}\left[\frac{|R|^2}{1+\mathbf{W}}(1+\mathcal{T}^2)w_\alpha\right],$$

and may estimate each term similarly to the proof of (5.7) to obtain

$$\|L\bar{\mathbf{P}}[n]\|_{\mathfrak{H}} \lesssim_A B\|(w,r)\|_{\mathcal{H}}.$$

4. Bounds for K_2. We start by dividing K_2 up according to frequency balance using the paraproduct operator T_R as

$$R[1+\mathcal{T}^2, b]w_\alpha = T_R[1+\mathcal{T}^2, b]w_\alpha + (R - T_R)[1+\mathcal{T}^2, b]w_\alpha.$$

When R is at low frequency we may estimate

$$\|T_R[1+\mathcal{T}^2, b]w_\alpha\|_{H^{\frac{1}{2}}} \lesssim \|R\|_{L^\infty}\|[1+\mathcal{T}^2, b]w_\alpha\|_{H^{\frac{1}{2}}}$$

and for the remaining terms we apply the paraproduct estimate (A.1) to obtain

$$\|(R-T_R)[1+\mathcal{T}^2, b]w_\alpha]\|_{H^{\frac{1}{2}}} \lesssim \|R\|_{\mathrm{bmo}^{\frac{1}{2}}}\|[1+\mathcal{T}^2, b]w_\alpha\|_{L^2}.$$

As a consequence,

$$\|L\mathbf{P}K_2\|_{\mathfrak{H}} \lesssim g^{\frac{1}{2}}A\|[1+\mathcal{T}^2, b]w_\alpha\|_{H^{\frac{1}{2}}}$$

We then decompose using paraproducts,

$$\begin{aligned}[1+\mathcal{T}^2, b]w_\alpha = [1+\mathcal{T}^2, T_b]w_\alpha + [1+\mathcal{T}^2, b_0]w_{\le 4} + (1+\mathcal{T}^2)T_{w_\alpha}b - T_{(1+\mathcal{T}^2)w_\alpha}b \\ + (1+\mathcal{T}^2)\Pi[b_{\ge 1}, w_\alpha] - \Pi[b_{\ge 1}, (1+\mathcal{T}^2)w_\alpha],\end{aligned}$$

and estimate each of these terms as follows: for the first two terms we apply the commutator estimate (A.8), for the third and fourth terms the estimate (A.7) and for the final two terms we apply the paraproduct estimate (A.1). The estimate for K_2 then follows from the estimate (A.18) for b.

5. Bounds for K_3. We may estimate similarly to K_2 to obtain

$$\begin{aligned}\|L\mathbf{P}K_3\|_{\mathfrak{H}} \lesssim \|R\|_{\mathrm{bmo}^{\frac{1}{2}}}\|(1+\mathcal{T}^2)[w_t + bw_\alpha - 2(1+\mathbf{W})S_0\operatorname{Re}\bar{P}[m]]\|_{H^{\frac{1}{2}}} \\ + \|R\|_{H^{\frac{1}{2}}}(1+\|\mathbf{W}\|_{L^\infty})\|S_0\operatorname{Re}\bar{P}[m]]\|_{L^\infty}.\end{aligned}$$

For the first term we estimate as for G_1 using that $1+\mathcal{T}^2$ has Schwartz symbol to obtain

$$\|(1+\mathcal{T}^2)[w_t + bw_\alpha - 2(1+\mathbf{W})S_0\operatorname{Re}\bar{P}[m]]\|_{H^{\frac{1}{2}}} \lesssim_A g^{-\frac{1}{2}}B\|(w,r)\|_{\mathcal{H}},$$

and for the second term we may simply apply the estimate (5.8).

6. Bounds for K_4. As $\mathcal{T}[w]$ is holomorphic we have

$$(1+i\mathcal{T})\mathcal{T}[w] = (1+\mathcal{T}^2)\operatorname{Re}\mathcal{T}[w].$$

We may then apply the paraproduct estimates (A.1) and (A.7) with the estimates (A.20) for $\mathfrak{a}$ and (A.16) for Y to obtain

$$\left\|\frac{g\mathbf{W}-\mathfrak{a}}{1+\mathbf{W}}(1+i\mathcal{T})\mathcal{T}[w]\right\|_{H^{\frac{1}{2}}} \lesssim \left(g\|Y\|_{\mathrm{bmo}^{\frac{1}{2}}} + \|\mathfrak{a}\|_{\mathrm{bmo}^{\frac{1}{2}}}\right)\|w\|_{L^2}.$$

This completes the proof of (5.6).

□

6. Normal forms

The goal of this section is to algebraically compute a normal form correction for the system (1.6) for (W, Q) as a translation invariant bilinear form. We recall that the aim of the normal form transformation is to eliminate the quadratic terms in the equation. Precisely, at least formally the normal form variables $(\tilde{W}, \tilde{Q})$ will solve a nonlinear equation where all the nonlinear terms are cubic and higher order. In this article we will not use such an equation directly for three reasons:

(i) The equation for the normal form variables $(\tilde{W}, \tilde{Q})$ is not self-contained, instead it still uses the original variables (W, Q) in the nonlinearity.
(ii) The system (1.6) is fully nonlinear and the normal form transformation does not mix well with the nonlinear structure.
(iii) The symbols for the normal form transformation are singular precisely when the output has frequency zero.

Instead, in the next section we use the normal form transformation in order to produce a cubic normal form energy that has the property that its time derivative along the flow is of quartic and higher order. Interestingly (and very usefully) the normal form symbol singularities do not carry over to the normal form energy; this is due to cancellations arising after repeated symmetrizations.

Incidentally, we remark that when considering the linearized equation some of these symmetrizations are lost, which is why we cannot prove cubic energy estimates for the linearized flow.

6.1. **The resonance analysis.** If we take $(W, Q) = 0$ in the linearized system (5.1) we obtain the system

$$\begin{cases} w_t + q_\alpha = 0 \\ q_t - g\mathcal{T}[w] = 0, \end{cases} \tag{6.1}$$

which has dispersion relation

$$\tau^2 = g\xi \tanh \xi.$$

As a consequence we see that solutions split into right-moving and left-moving components with dispersion relations $\tau = \pm g^{\frac{1}{2}}\omega(\xi)$, respectively, where

$$\omega(\xi) = -\operatorname{sgn} \xi \sqrt{\xi \tanh \xi}.$$

To understand bilinear resonant interactions we define the function

$$\Delta(\xi, \eta, \zeta) = \omega(\xi) + \omega(\eta) + \omega(\zeta).$$

Then resonant two wave interactions correspond to solutions to the system

$$\begin{cases} \Delta(\pm\xi, \pm\eta, \pm\zeta) = 0 \\ \xi + \eta + \zeta = 0. \end{cases}$$

As ω is sublinear, the only solutions occur when at least one of ξ, η, ζ vanishes.

Symmetrizing the Δ function, we define the resonance function Ω by

$$\begin{aligned} \Omega(\xi, \eta, \zeta) &= \Delta(\xi, \eta, \zeta)\Delta(\xi, -\eta, -\zeta)\Delta(\xi, -\eta, \zeta)\Delta(\xi, \eta, -\zeta) \\ &= J(\xi)^2 + J(\eta)^2 + J(\zeta)^2 - 2J(\xi)J(\eta) - 2J(\eta)J(\zeta) - 2J(\zeta)J(\xi), \end{aligned}$$

on the set $\mathcal{P} = \{\xi + \eta + \zeta = 0\}$, where $J(\xi) = \omega(\xi)^2 = \xi \tanh \xi$. This vanishes quadratically on each of the lines $\xi = 0$, $\eta = 0$, respectively $\zeta = 0$. The function Ω will play a key role in all computations which follow.

6.2. **Expansion to cubic order.** In order to construct normal forms for (W, Q) we first expand F to cubic order as

$$\Lambda^{\leq 3} F = Q_\alpha - Q_\alpha W_\alpha - \mathbf{P}\left[Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha\right] + Q_\alpha W_\alpha^2 + \mathbf{P}\left[(Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha)(W_\alpha + \bar{W}_\alpha)\right],$$

where, for a sufficiently smooth function $f\colon \mathbb{C}^2 \to \mathbb{C}$ we define $\Lambda^{\leq k} f$ to select the terms of polynomial order $\leq k$ in the Taylor expansion of f at zero.

We may then rewrite (1.6) as

$$\begin{cases} W_t + Q_\alpha = G^{[2]} + G^{[3]} + G^{[4+]} \\ Q_t - g\mathcal{T}[W] = K^{[2]} + K^{[3]} + K^{[4+]}, \end{cases} \tag{6.2}$$

where the quadratic terms are given by

$$G^{[2]} = \mathbf{P}[Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha], \quad K^{[2]} = -Q_\alpha^2 - \mathbf{P}\left[Q_\alpha \bar{Q}_\alpha\right],$$

the cubic terms are given by

$$\begin{aligned} G^{[3]} &= W_\alpha \mathbf{P}\left[Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha\right] - \mathbf{P}\left[(Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha)(W_\alpha + \bar{W}_\alpha)\right], \\ K^{[3]} &= Q_\alpha^2 W_\alpha + Q_\alpha \mathbf{P}\left[Q_\alpha \bar{W}_\alpha - \bar{Q}_\alpha W_\alpha\right] + \mathbf{P}\left[Q_\alpha \bar{Q}_\alpha (W_\alpha + \bar{W}_\alpha)\right], \end{aligned}$$

and $G^{[4+]}, K^{[4+]}$ contain only quartic and higher order terms.

6.3. **Normal forms.** By considering parity, we seek holomorphic normal form corrections of the form

$$\begin{cases} \tilde{W} = W + B^h[W, W] + \dfrac{1}{g} C^h[Q, Q] + B^a[W, \bar{W}] + \dfrac{1}{g} C^a[Q, \bar{Q}] \\ \tilde{Q} = Q + A^h[W, Q] + A^a[W, \bar{Q}] + D^a[Q, \bar{W}], \end{cases}$$

so that the normal form variables $(\tilde{W}, \tilde{Q})$ satisfy

$$\begin{cases} \Lambda^{\leq 2}[\tilde{W}_t + \tilde{Q}_\alpha] = 0 \\ \Lambda^{\leq 2}[\tilde{Q}_t - g\mathcal{T}\tilde{W}] = 0. \end{cases} \tag{6.3}$$

Here the operators B^h, C^h, B^a, C^a, A^h, A^a, D^a are translation invariant bilinear forms, which can be described via their symbols, as below:

$$\begin{aligned} B^h[W, W] &= \frac{1}{2\pi} \int B^h(\xi, \eta) \hat{W}(\xi) \hat{W}(\eta) e^{i(\xi + \eta)\alpha} \, d\xi d\eta \\ B^a[W, \bar{W}] &= \frac{1}{2\pi} \int B^a(\xi, \eta) \hat{W}(\xi) \hat{\bar{W}}(\eta) e^{i(\xi - \eta)\alpha} \, d\xi d\eta. \end{aligned}$$

To determine these symbols uniquely we assume that B^h, C^h are symmetric.

For the subsequent construction of the normal form energies we will interpret all symbols as functions on the plane $\mathcal{P} = \{\xi + \eta + \zeta = 0\}$. For notational convenience we will adopt this convention in the following computations. In the context of bilinear operators we may interpret $\zeta = \zeta(\xi, \eta) := -(\xi + \eta)$. For this reason we will compute holomorphic symbols at (ξ, η) and mixed holomorphic-antiholomorphic symbols at $(\xi, -\eta)$.

6.3.1. *Holomorphic products.* The holomorphic terms A^h, B^h and C^h are generated by the holomorphic part of the quadratic nonlinearity, i.e., the first term in $K^{[2]}$. Comparing holomorphic terms at the quadratic level we obtain a linear system for the symbols

$$\begin{bmatrix} \xi+\eta & -2\eta & -2\tanh\xi \\ -\xi & 0 & \tanh(\xi+\eta) \\ -\tanh\eta & \tanh(\xi+\eta) & 0 \end{bmatrix} \begin{bmatrix} A^h(\xi,\eta) \\ B^h(\xi,\eta) \\ C^h(\xi,\eta) \end{bmatrix} = \begin{bmatrix} 0 \\ i\xi\eta \\ 0 \end{bmatrix}.$$

From the first row we have

$$A^h = \frac{2\eta B^h}{\xi+\eta} + \frac{2\tanh\xi C^h}{\xi+\eta}.$$

We then calculate the symmetrizations

$$\begin{aligned} (\xi A^h)_{\mathrm{sym}} &= \frac{2\xi\eta B^h}{\xi+\eta} + \frac{(\xi\tanh\xi+\eta\tanh\eta)C^h}{\xi+\eta}, \\ (\tanh\eta A^h)_{\mathrm{sym}} &= \frac{(\xi\tanh\xi+\eta\tanh\eta)B^h}{\xi+\eta} + \frac{2\tanh\xi\tanh\eta C^h}{\xi+\eta}. \end{aligned}$$

Plugging this into the second row we obtain,

$$((\xi+\eta)\tanh(\xi+\eta) - \xi\tanh\xi - \eta\tanh\eta)C^h = i\xi\eta(\xi+\eta) + 2\xi\eta B^h,$$

and into the third row,

$$((\xi+\eta)\tanh(\xi+\eta) - \xi\tanh\xi - \eta\tanh\eta)B^h = 2\tanh\xi\tanh\eta C^h.$$

As a consequence we obtain the solutions

$$\begin{aligned} A^h(\xi,\eta) &= \frac{2i\eta J(\xi)\left(J(\zeta) - J(\xi) + J(\eta)\right)}{\Omega}, \\ B^h(\xi,\eta) &= -\frac{2i\zeta J(\xi)J(\eta)}{\Omega}, \\ C^h(\xi,\eta) &= -\frac{i\xi\eta\zeta\left(J(\zeta) - J(\xi) - J(\eta)\right)}{\Omega}. \end{aligned}$$

6.3.2. *Mixed terms.* The mixed terms A^a, B^a, C^a and D^a are generated by the mixed holomorphic-antiholomorphic part of the quadratic nonlinearity. As above, we write the mixed holomorphic-antiholomorphic terms as a linear system

$$\begin{bmatrix} \xi+\eta & -\eta & -\tanh\xi & 0 \\ 0 & -\xi & -\tanh\eta & \xi+\eta \\ -\xi & 0 & \tanh(\xi+\eta) & -\eta \\ -\tanh\eta & \tanh(\xi+\eta) & 0 & -\tanh\xi \end{bmatrix} \begin{bmatrix} A^a(\xi,-\eta) \\ B^a(\xi,-\eta) \\ C^a(\xi,-\eta) \\ D^a(\xi,-\eta) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}i(1-\coth(\xi+\eta))\xi\eta \\ -\frac{1}{2}i(1-\coth(\xi+\eta))\xi\eta \\ \frac{1}{2}i(1-\tanh(\xi+\eta))\xi\eta \\ 0 \end{bmatrix}$$

We solve this system to obtain

$$
\begin{aligned}
A^a(\xi,-\eta) &= -\frac{e^{2\zeta}}{e^{2\zeta}+1}\left\{(J(\eta)+\eta)\,\frac{B^h(\xi,\eta)}{\zeta\tanh\eta}+(J(\xi)-\xi)\,\frac{C^h(\xi,\eta)}{\xi\zeta}\right\},\\
B^a(\xi,-\eta) &= \frac{e^{2\zeta}}{e^{2\zeta}-1}\left\{(J(\zeta)-(\xi-\eta))\,\frac{B^h(\xi,\eta)}{\zeta}+(\eta J(\xi)-\xi J(\eta))\,\frac{C^h(\xi,\eta)}{\xi\eta\zeta}\right\},\\
C^a(\xi,-\eta) &= \frac{e^{2\zeta}}{e^{2\zeta}-1}\left\{(\eta J(\xi)-\xi J(\eta))\,\frac{B^h(\xi,\eta)}{\zeta\tanh\xi\tanh\eta}+(J(\zeta)-(\xi-\eta))\,\frac{C^h(\xi,\eta)}{\zeta}\right\},\\
D^a(\xi,-\eta) &= -\frac{e^{2\zeta}}{e^{2\zeta}+1}\left\{(J(\xi)-\xi)\,\frac{B^h(\xi,\eta)}{\zeta\tanh\xi}+(J(\eta)+\eta)\,\frac{C^h(\xi,\eta)}{\eta\zeta}\right\}.
\end{aligned}
$$

6.4. **Symbol classes and asymptotics for the normal form.** Here we consider the symbols arising in the normal form, and describe their size and regularity. These are needed in order to have good L^2 and L^p multilinear bounds.

From the perspective of high frequency bounds, we are interested in the interactions between one high negative frequency and one low frequency. Here we expect only the symbols B^h, B^a, A^h and D^a to play a role; the remaining symbols C^h, C^a and A^a (which do not appear at all in the infinite bottom case) will decay rapidly in the above regime. For the former symbols, on the other hand, we will need to compute second order expansions around $\xi=0$ (for B^h and A^h) respectively around $\eta=0$ (for B^a, A^h and D^a). However, due to the linear component of the normal derivative of the pressure, we will also require an expansion for C^h near $\eta=0$.

From the perspective of low frequency analysis, we do not have any low frequency pointwise control on $\operatorname{Re} W$ and $\operatorname{Re} Q$. Hence we will need to show that $\operatorname{Re} W$ and $\operatorname{Re} Q$ do not appear undifferentiated in our cubic energy functional. This requires certain cancellations to happen (akin to a null condition). For this we will need to exactly compute almost all of the above symbols at $\xi=0$ and at $\eta=0$.

We will interpret all symbols as functions on the plane $\mathcal{P}=\{\xi+\eta+\zeta=0\}$. In this plane we consider three distinguished lines $\xi=0$, $\eta=0$, $\zeta=0$. The symbol regularity will depend on the distance d to these lines and on the radius ρ,

$$
d = 1+\min\{|\xi|,|\eta|,|\zeta|\}, \qquad \rho = 1+\max\{|\xi|,|\eta|,|\zeta|\}.
$$

For a weight σ which is slowly varying with respect to these scales we denote by $S(\sigma)$ the class of symbols s on $\mathcal{P}$ which satisfy

$$
|(d\partial)^\alpha(\rho\partial_\rho)^\beta s| \lesssim c_{\alpha\beta}\sigma.
$$

We begin our discussion with the expression Ω, for which we have:

Lemma 6.1.

a) The symbol Ω restricted to $\mathcal{P}$ is non-positive and belongs to $S(d\rho)$.

b) The symbol Ω vanishes quadratically on the three lines and is elliptic elsewhere,

$$
\mathcal{T}^2(\xi)\mathcal{T}^2(\eta)\mathcal{T}^2(\zeta)\Omega^{-1}\in S(d^{-1}\rho^{-1}).
$$

c) We have the following expansion in the region $|\eta|\ll -\xi$:

$$
\Omega(\xi,\eta,\zeta) = 4J(\eta)\xi+(\eta+J(\eta))^2+S(e^{\xi}) = -4J(\eta)\zeta+(\eta-J(\eta))^2+S(e^{-\zeta}).
$$

d) On the line $\eta = 0$ *we have the limit*

$$\lim_{|\eta|\to 0} \eta^{-2}\Omega(\xi,\eta,\zeta) = J'(\xi)^2 - 4J(\xi) =: \Lambda(\xi) < 0.$$

The proof is a fairly straightforward algebraic computation and is omitted. We remark that part (b) is consistent with the fact that in our problem two wave resonances appear only when either an input frequency or the output frequency is zero. Part (c) is relevant in our high frequency analysis, while part (d) is needed for the low frequency cancellation.

Now we successively consider the symbols in our normal form analysis:

The symbol B^h. Here by inspection we see that all the zeros of Ω are canceled by the numerator, except for a simple zero at $\zeta = 0$. Then the natural regularity statement is obtained after multiplication with $\mathcal{T}(\zeta)$. Precisely, we have

$$\mathcal{T}(\zeta)B^h(\xi,\eta) \in S(\rho). \tag{6.4}$$

For the high frequency asymptotics in the region $|\eta| \ll -\xi$ we have the expansion

$$B^h(\xi,\eta) = -\frac{i}{2}\left(\xi - \frac{(\eta - J(\eta))^2}{4J(\eta)}\right) + S(d^2\rho^{-1}). \tag{6.5}$$

On the other hand, at $\eta = 0$ we have

$$B^h(\xi,0) = \frac{2i\xi J(\xi)}{\Lambda(\xi)}.$$

The symbol C^h. Again all the zeros of Ω are canceled by the numerator, except for a simple zero at $\zeta = 0$. Further, the difference $J(\xi) + J(\eta) - J(\zeta)$ decays exponentially if ξ and η have the same sign,

$$J(\xi) + J(\eta) - J(\zeta) = O(e^{-|\xi|} + e^{-|\eta|}), \quad \xi\eta > 0.$$

We then obtain the size of C^h as

$$\mathcal{T}(\zeta)C^h(\xi,\eta) \in \begin{cases} S(\rho \ \min\{|\xi|,|\eta|\}) & \xi\eta < 0 \\ S(d^{-N}\rho) & \xi\eta > 0. \end{cases} \tag{6.6}$$

The asymptotics in the region $|\eta| \ll -\xi$ are

$$C^h(\xi,\eta) = -\frac{i\eta(\eta + J(\eta))}{4J(\eta)}\left(\xi - \frac{(\eta - J(\eta))^2}{4J(\eta)}\right) + S(d^3\rho^{-1}).$$

Finally, we also need

$$C^h(\xi,0) = \frac{i\xi^2 J'(\xi)}{\Lambda(\xi)}.$$

The symbol A^h. As before, we remove the zero at $\zeta = 0$ to obtain the regularity

$$\mathcal{T}(\zeta)A^h(\xi,\eta) \in \begin{cases} S(|\eta|) & \eta\zeta < 0 \\ S(d^{-N}|\eta|) & \eta\zeta > 0. \end{cases} \tag{6.7}$$

Since A^h is not symmetric, we need asymptotics both near $\xi = 0$ and $\eta = 0$. First we consider the region $|\eta| \ll -\zeta$. Here we have

$$A^h(\xi,\eta) = -\frac{i\eta(\eta + J(\eta))}{2J(\eta)}\left(1 + \frac{(\eta - J(\eta))^2}{4\xi J(\eta)}\right) + S(d^3\rho^{-1}), \tag{6.8}$$

where the leading order term vanishes (which is consistent with the infinite bottom problem). Next, we consider the region $|\xi| \ll -\eta$:

$$A^h(\xi,\eta) = -i\left(\eta + \frac{J(\xi)^2 - \xi^2}{4J(\xi)}\right) + S(d^2\rho^{-1}). \tag{6.9}$$

Finally, we compute

$$A^h(\xi,0) = \frac{2iJ(\xi)J'(\xi)}{\Lambda(\xi)}, \qquad A^h(0,\eta) = \frac{4i\eta J(\eta)}{\Lambda(\eta)}.$$

Next we consider the symbols for the mixed terms, namely A^a, B^a, C^a and D^a. Here we will continue to consider $(\xi,\eta,\zeta) \in \mathcal{P}$ and compute the symbols at $(\xi,-\eta)$.

The symbol A^a. The symbol $A^a(\xi,-\eta)$ decays exponentially in all directions except near the half-lines $\{\xi = 0,\ \eta < 0\}$ and $\{\zeta = 0,\ \xi < 0\}$. Precisely, we have

$$\mathcal{T}(\zeta)A^a(\xi,-\eta) \in \begin{cases} S(d^{-N}|\xi|) & |\xi| \ll -\eta \ \text{ or } \ |\zeta| \ll -\xi \\ S(\rho^{-N}) & \text{elsewhere.} \end{cases}. \tag{6.10}$$

Finally, we have

$$A^a(\xi,0) = \frac{i}{\Lambda(\xi)(e^{2\xi}+1)}\left[2J(\xi) - \xi J'(\xi) + J(\xi)J'(\xi)\right].$$

The symbol B^a. This is similar to B^h, in that

$$\mathcal{T}(\zeta)B^a(\xi,-\eta) \in S(\rho). \tag{6.11}$$

In the region $|\eta| \ll -\xi$ we have the asymptotics

$$B^a(\xi,-\eta) = -i\xi + S(d^2\rho^{-1}). \tag{6.12}$$

Finally, we do not need the exact expressions for $B^a(\xi,0)$ and $B^a(0,-\eta)$, only the fact that they are purely imaginary.

The symbol C^a. The symbol $C^a(\xi,-\eta)$ decays exponentially away from the region $\{0 < \xi \ll -\eta\}$ and the half-line $\{\zeta = 0,\ \eta < 0\}$. Precisely,

$$\mathcal{T}(\zeta)C^a(\xi,-\eta) \in \begin{cases} S(|\xi|\rho) & 0 < \xi \ll -\eta \ \text{ or } \ |\zeta| \ll -\eta \\ S(d^{-N}|\xi|\rho) & 0 < -\xi \ll -\eta \ \text{ or } \ |\zeta| \ll -\xi \\ S(\rho^{-N}) & \text{elsewhere.} \end{cases}. \tag{6.13}$$

Finally, on the two lines we have

$$C^a(\xi,0) = \frac{i\xi}{\Lambda(\xi)(e^{2\xi}-1)}\left[2J(\xi) - \xi J'(\xi) + J(\xi)J'(\xi)\right],$$

$$C^a(0,-\eta) = -\frac{i\eta}{\Lambda(\eta)(e^{2\eta}-1)}\left[2J(\eta) - \eta J'(\eta) - J(\eta)J'(\eta)\right].$$

The symbol D^a. The symbol $D^a(\xi,-\eta)$ decays exponentially away from the region $\{0<-\xi\ll-\eta\}$ and the half-line $\{\eta=0,\ \xi<0\}$. Precisely,

$$\mathcal{T}(\zeta)D^a(\xi,-\eta)\in\begin{cases} S(|\xi|) & 0<-\xi\ll-\eta \ \text{ or } \ |\eta|\ll-\xi \\ S(d^{-N}|\xi|) & 0<\xi\ll-\eta \ \text{ or } \ |\zeta|\ll-\eta \\ S(\rho^{-N}) & \text{elsewhere.}\end{cases} \tag{6.14}$$

We will only require its high frequency asymptotics in the region $|\eta|\ll-\xi$:

$$D^a(\xi,-\eta)=-i\xi+S(d^2\rho^{-1}), \tag{6.15}$$

which are similar to those for B^a.

Finally, we also need

$$D^a(0,-\eta)=-\frac{i}{\Lambda(\eta)(e^{2\eta}+1)}\left[2J(\eta)-\eta J'(\eta)-J(\eta)J'(\eta)\right].$$

7. The normal form energy.

The aim of this section is to use the normal form computation in the previous section to produce a normal form energy, i.e., an energy functional which is accurate to quartic order. We summarize our result as follows:

Proposition 7.1. *For each $n\geq 1$ there exists a normal form energy $E^n_{NF}=E^n_{NF}(\mathbf{W},R)$ with the following properties:*

a) Algebraic properties. $E^n_{NF}(\mathbf{W},R)$ has only quadratic and cubic terms,

$$\Lambda^{\geq 4}E^n_{NF}(\mathbf{W},R)=0,$$

and its quadratic part is given by the linear energy

$$\Lambda^{\leq 2}E^n_{NF}(\mathbf{W},R)=E_0(\partial^{n-1}\mathbf{W},\partial^{n-1}R).$$

Further, $E^n_{NF}(\mathbf{W},R)$ is accurate to quartic order, i.e.,

$$\Lambda^{\leq 3}\frac{d}{dt}E^n_{NF}(\mathbf{W},R)=0 \tag{7.1}$$

along the flow of (1.6).

b) Qualitative description. E^n_{NF} has the form

$$E^n_{NF}(\mathbf{W},R)=E_0(\partial^{n-1}\mathbf{W},\partial^{n-1}R)+gB(\mathbf{W},\mathbf{W},\mathbf{W})+A(\mathbf{W},R,R),$$

where A and B are translation invariant trilinear forms. Further, there is a decomposition

$$E^n_{NF}=E^n_{NF,high}+E^n_{NF,low},$$

with

$$\begin{aligned} E^n_{NF,high}&=E_0(\partial^{n-1}\mathbf{W},\partial^{n-1}R)+gB_{high}(\mathbf{W},\mathbf{W},\mathbf{W})+A_{high}(\mathbf{W},R,R),\\ E^n_{NF,low}&=gB_{low}(\mathbf{W},\mathbf{W},\mathbf{W})+A_{low}(\mathbf{W},R,R),\end{aligned}$$

where the forms B_{high} and A_{high}, respectively B_{low} and A_{low} are characterized as follows:

(i) Case $n \geq 2$. Then the forms B_{high}, A_{high} are given by

$$
\begin{aligned}
&B_{high}(\mathbf{W},\mathbf{W},\mathbf{W}) := \langle \partial^{n-1}\mathbf{W}, \partial^{n-1}\mathbf{W}\rangle_{-4n\,\mathrm{Re}\,\mathbf{W}+\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}},\\
&A_{high}(\mathbf{W},R,R) := -\langle \partial^{n-1}R, \mathcal{T}^{-1}\partial^{n-1}R_\alpha\rangle_{-4n\,\mathrm{Re}\,\mathbf{W}-\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}}\\
&\qquad - 2\langle \mathbf{W}\partial^{(n-1)}R, \mathcal{T}^{-1}\partial^{(n-1)}R_\alpha\rangle + 2\langle \partial^{(n-2)}\mathbf{W}R_\alpha, \mathcal{T}^{-1}\partial^{(n-1)}R_\alpha\rangle,
\end{aligned}
\tag{7.2}
$$

whereas the forms B_{low} and A_{low} have symbols $B_{low}(\xi,\eta,\zeta)$, $A_{low}(\xi,\eta,\zeta)$ in the class

$$B_{low} \in S(d\rho^{2n-3}), \qquad A_{low} \in S(dd_1\rho^{2n-3}) + S(\rho^{2n-2}),$$

where d,ρ are defined as before and $d_1 = \min\{|\eta|,|\zeta|\}$ is the smaller of the two R frequencies.

(ii) Case $n = 1$. Then the forms B_{high}, A_{high} are given by

$$
\begin{aligned}
&B_{high}(\mathbf{W},\mathbf{W},\mathbf{W}) := \langle \mathbf{W},\mathbf{W}\rangle_{-4\,\mathrm{Re}\,\mathbf{W}+\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}},\\
&A_{high}(\mathbf{W},R,R) := -\langle R, \mathcal{T}^{-1}R_\alpha\rangle_{-4\,\mathrm{Re}\,\mathbf{W}-\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}} - 2\langle R\mathbf{W}, \mathcal{T}^{-1}R_\alpha\rangle,
\end{aligned}
\tag{7.3}
$$

and the forms B_{low} and A_{low} have symbols $B_{low}(\xi,\eta,\zeta)$, $A_{low}(\xi,\eta,\zeta)$ in the class

$$B_{low} \in S(\rho^{-1}), \qquad A_{low} \in S(1).$$

The remainder of this section is devoted to the proof of the above proposition. To start with we give a brief description of the types of trilinear forms B and A that we will work with. These trilinear forms are translation invariant so they can be described in terms of their symbols. Precisely, one can represent any such trilinear form $B(W,W,W)$ and $A(W,Q,Q)$ as

$$
\begin{aligned}
B(W,W,W) &= \frac{2}{\sqrt{2\pi}}\,\mathrm{Re}\int_{\xi+\eta+\zeta=0} B(\xi,\eta,\zeta)\hat{W}(\xi)\hat{W}(\eta)\hat{W}(\zeta)\,d\xi d\eta,\\
A(W,Q,Q) &= \frac{2}{\sqrt{2\pi}}\,\mathrm{Re}\int_{\xi+\eta+\zeta=0} A(\zeta,\xi,\eta)\hat{W}(\zeta)\hat{Q}(\xi)\hat{Q}(\eta)\,d\xi d\eta.
\end{aligned}
$$

At the same time, we also need trilinear forms which involve complex conjugates. However, the functions W and Q are holomorphic, and thus their Fourier transforms satisfy the relations

$$\bar{\hat{W}}(-\xi) = e^{2\xi}\hat{W}(\xi), \qquad \bar{\hat{Q}}(-\xi) = e^{2\xi}\hat{Q}(\xi). \tag{7.4}$$

These relations allow us to uniquely represent all the cubic terms in the normal form energy functional in the above form without any conjugates. The price to pay is that we need to allow such exponentials in our symbol classes. However, this happens in a very limited way. To account for this we introduce the following notation

Definition 7.2. Given any class of symbols $S(\sigma)$ on the plane $\mathcal{P} = \{\xi+\eta+\zeta = 0\}$, we denote by $ES(\sigma)$ the linear span of symbols in $\{S(\sigma), e^{\pm 2\xi}S(\sigma), e^{\pm 2\eta}S(\sigma), e^{\pm 2\zeta}S(\sigma)\}$.

We note than any trilinear form with symbols in the class ES, acting on holomorphic functions, can be written as a sum of trilinear forms with symbols in S, but where complex conjugation is also allowed.

Also we note that for any such trilinear form, its symbol is uniquely determined up to symmetries, i.e., for the symmetric part of the above symbols. Indeed, symmetrizations will play a crucial role in our computations because they will allow us to gain some critical cancellations.

The aim of this section is to determine the symbols A, B above, and to study their properties.

7.1. **From normal forms to normal form energies.** As a first step in the proof of the proposition, here we obtain a preliminary normal form energy $\tilde{E}^n_{NF}(W,Q)$ of the form

$$\tilde{E}^n_{NF}(W,Q) = E_0(\partial^n \tilde{W}, \partial^n \tilde{Q}), + g\tilde{B}(W,W,W) + \tilde{A}(W,Q,Q)$$

so that the key property (7.1) holds. The natural expression for the normal form energy is provided by the normal form transformation computed in the previous section. Precisely, we will take

$$\begin{aligned}\tilde{E}^n_{NF}(W,Q) &= \Lambda^{\leq 3} E_0(\partial^n \tilde{W}, \partial^n \tilde{Q}) \\ &= E_0(\partial^n W, \partial^n Q) + 2g\langle \partial^n W, \partial^n W^{[2]}\rangle - 2\langle \mathcal{T}^{-1}\partial^{n+1} Q, \partial^n Q^{[2]}\rangle.\end{aligned}$$

In view of the equations (6.3) the property (7.1) is automatically satisfied. It remains to express the trilinear forms above involving the normal form corrections $W^{[2]}$, $Q^{[2]}$ as trilinear forms $\tilde{B}(W,W,W)$ and $\tilde{A}(W,Q,Q)$.

Given the expressions for $W^{[2]}$ and $Q^{[2]}$, the trilinear form $\tilde{B}$ is as follows:

$$\begin{aligned}\tilde{B}(W,W,W) =& \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \zeta^{2n} (\bar{\hat{W}}(-\zeta)) - \hat{W}(\zeta)) B^h(\xi,\eta) \hat{W}(\xi)\hat{W}(\eta)\, d\xi d\eta \\ &+ \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \zeta^{2n} (\bar{\hat{W}}(-\zeta)) - \hat{W}(\zeta)) B^a(\xi,-\eta) \hat{W}(\xi)\bar{\hat{W}}(-\eta)\, d\xi d\zeta.\end{aligned}$$

We can put these two integrals together using the relation (7.4) to obtain

$$\tilde{B}(\xi,\eta,\zeta) = (e^{2\zeta}-1)\zeta^{2n}(B^h(\xi,\eta) + e^{2\eta}B^a(\xi,-\eta)).$$

Further, we can symmetrize $\tilde{B}$ with respect to the three variables, as well as with respect to the reflection symmetry[1]

$$\tilde{B}(\xi,\eta,\zeta) \to \bar{\tilde{B}}(-\xi,-\eta,-\zeta).$$

We denote the symmetrization of $\tilde{B}$ by $\tilde{B}^{sym}$, which can be used instead of $\tilde{B}$. As mentioned before, this symmetrization is very important, not only in order to uniquely describe the trilinear form, but also because it allows us to eliminate small denominators in the symbol for $\tilde{B}$ (even though such singularities do appear in the normal form).

[1] Here we use the fact that the trilinear forms $\tilde{A}$, $\tilde{B}$ are real valued.

We can perform a similar computation for $\tilde{A}$:

$$
\begin{aligned}
\tilde{A}(W,Q,Q) = & \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \zeta^{2n} (\bar{\hat{W}}(-\zeta)) - \hat{W}(\zeta)) C^h(\xi,\eta) \hat{Q}(\xi)\hat{Q}(\eta)\, d\xi d\eta \\
& + \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \zeta^{2n} (\bar{\hat{W}}(-\zeta)) - \hat{W}(\zeta)) C^a(\xi,-\eta) \hat{Q}(\xi)\bar{\hat{Q}}(-\eta)\, d\xi d\zeta \\
& + \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \coth\zeta\ \zeta^{2n+1} (\bar{\hat{Q}}(-\zeta)) - \hat{Q}(\zeta)) A^h(\xi,\eta) \hat{W}(\xi)\hat{Q}(\eta)\, d\xi d\eta \\
& + \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \coth\zeta\ \zeta^{2n+1} (\bar{\hat{Q}}(-\zeta)) - \hat{Q}(\zeta)) A^a(\xi,-\eta) \hat{W}(\xi)\bar{\hat{Q}}(-\eta)\, d\xi d\zeta \\
& + \frac{2}{\sqrt{2\pi}} \operatorname{Re} \int_{\xi+\eta+\zeta=0} \coth\zeta\ \zeta^{2n+1} (\bar{\hat{Q}}(-\zeta)) - \hat{Q}(\zeta)) D^a(\xi,-\eta) \hat{Q}(\xi)\bar{\hat{W}}(-\eta)\, d\xi d\zeta.
\end{aligned}
$$

This yields the symbol for $\tilde{A}$, namely

$$
\begin{aligned}
\tilde{A}(\zeta,\xi,\eta) = & \zeta^{2n}(e^{2\zeta}-1)\left(C^h(\xi,\eta) + e^{2\eta}C^a(\xi,-\eta)\right) \\
& + \xi^{2n+1}(e^{2\xi}+1)\left(A^h(\zeta,\eta) + e^{2\eta}A^a(\zeta,-\eta) + e^{2\zeta}D^a(\eta,-\zeta)\right).
\end{aligned}
$$

Again, this can be further symmetrized with respect to ξ and η, as well as with respect to the reflection symmetry to obtain the symbol $\tilde{A}^{sym}$.

7.2. **The properties of the symbols $\tilde{A}^{sym}$ and $\tilde{B}^{sym}$.** A crucial step in our analysis is to understand the properties of the symbols $\tilde{A}^{sym}$ and $\tilde{B}^{sym}$. In this we have two goals. In terms of low frequencies, we want to show that we can extract factors of $\xi\eta\zeta$, so that $\tilde{E}^n_{NF}$ depends only on the differentiated variables W_α and Q_α. In terms of high frequencies we seek to find the leading terms in the expansion of the symbols for $\tilde{A}$ and $\tilde{B}$ near the axis $\xi = 0$, $\eta = 0$ and $\zeta = 0$. These are as follows:

Lemma 7.3.

a) The symbols $\tilde{A}^{sym}$ and $\tilde{B}^{sym}$ can be expressed in the form

$$
\tilde{A}^{sym} \in \xi\eta\zeta ES(\rho^{2n-1}), \qquad \tilde{B}^{sym} \in \xi\eta\zeta ES(\rho^{2n-2}).
$$

b) The leading order terms in $\tilde{B}^{sym}$ in the region $|\eta| \ll \xi$ have the form

$$
\begin{aligned}
\tilde{B}^{sym} = & -\frac{i}{48}e^{2\xi}\xi^{2n}\left(8n\eta - \frac{(\eta+J(\eta))^2}{J(\eta)}\right) - \frac{i}{48}e^{-2\zeta}\zeta^{2n}\left(8n\eta + \frac{(\eta-J(\eta))^2}{J(\eta)}\right) \\
& + \eta ES(d\rho^{2n-1}).
\end{aligned} \tag{7.5}
$$

c) The leading order terms in $\tilde{A}^{sym}$ in the region $|\zeta| \ll \xi$ are as follows:

$$
\begin{aligned}
\tilde{A}^{sym} = & \frac{i}{16}e^{2\xi}\xi^{2n}\eta\left(8n\zeta + \frac{J(\zeta)^2-\zeta^2}{J(\zeta)}\right) - \frac{i}{16}e^{-2\eta}\eta^{2n}\xi\left(8n\zeta - \frac{J(\zeta)^2-\zeta^2}{J(\zeta)}\right) \\
& + \zeta ES(d\rho^{2n}).
\end{aligned} \tag{7.6}
$$

d) The leading order terms in $\tilde{A}^{sym}$ in the region $|\eta| \ll \zeta$ are as follows:

$$
\tilde{A}^{sym} = \frac{1}{4}ie^{2\eta+2\zeta}\xi^{2n+1}\eta + \eta ES(d^2\rho^{2n-1}) + \eta ES(\rho^{2n}). \tag{7.7}
$$

Proof. We successively establish the desired properties for $\tilde{A}^{sym}$ and $\tilde{B}^{sym}$. To simplify the bookkeeping we introduce the notation $\overset{\text{sym}}{=}$ to describe the relation between two symbols which have the same symmetrization.

1. The symbol $\tilde{B}^{sym}$. We recall that

$$\tilde{B}(\xi,\eta,\zeta) = (e^{2\zeta}-1)\zeta^{2n}\left(B^h(\xi,\eta) + e^{2\eta}B^a(\xi,-\eta)\right).$$

Using symmetries, the B^h contribution to $\tilde{B}^{sym}$ is given by obtained by symmetrizing the expression

$$\tilde{B}^{h,sym} \overset{\text{sym}}{=} \zeta^{2n}(e^{2\zeta}-1)B^h(\xi,\eta) \overset{\text{sym}}{=} 2\zeta^{2n}\sinh^2\zeta B^h(\xi,\eta) \overset{\text{sym}}{=} -i\zeta^{2n}(e^{2\zeta}-e^{-2\zeta})\frac{J(\xi)J(\eta)J(\zeta)}{\Omega(\xi,\eta,\zeta)}.$$

This is a smooth symbol. Further, since the exponential factor is odd and all other factors are even, its symmetrization vanishes on all three diagonals.

For the B^a part we simplify using the reflection symmetry,

$$\begin{aligned}\tilde{B}^{a,sym} &\overset{\text{sym}}{=} \zeta^{2n-1}e^{-2\xi}\{(J(\xi+\eta)-(\xi-\eta))B^h(\xi,\eta) + (\tanh\xi - \tanh\eta)C^h(\xi,\eta)\}\\ &\overset{\text{sym}}{=} i\zeta^{2n}(e^{2\xi}-e^{-2\xi})\frac{J(\xi)J(\eta)J(\zeta)}{\Omega(\xi,\eta,\zeta)} + i\frac{\zeta^{2n}(e^{2\xi}+e^{-2\xi})}{2\Omega(\xi,\eta,\zeta)}K(\xi,\eta),\end{aligned}$$

where

$$K(\xi,\eta) = 2(\xi-\eta)J(\xi)J(\eta) - (\eta J(\xi) - \xi J(\eta))(J(\zeta)-J(\xi)-J(\eta)).$$

The symmetrization of the first term vanishes on the diagonals as the first two factors are even, respectively odd, and the fraction is fully symmetric. The same applies for the last term, where all we need to use for K is that it is odd and antisymmetric.

Next we consider the high frequency asymptotics. Simply by considering separately the size of each component above, we obtain $\tilde{B}^{sym} \in ES(\rho^{2n+1})$, which suffices outside a small conical neighborhood of the diagonals. We need to improve this near the diagonals so we consider the case $|\eta| \ll |\xi|, |\zeta|$. Here we need to compute the principal part of $\tilde{B}^{sym}$ modulo lower order terms, i.e., symbols in $ES(d^2\rho^{2n-1})$.

The terms containing $e^{\pm 2\eta}$ are exponentially small compared to $e^{\pm 2\xi}$ and $e^{\pm 2\zeta}$ so we can neglect them. We can also neglect terms with the η^{2n} factor. Further, there can be no polynomial cancellation arising from the exponentials so we might as well consider them separately. Hence we consider the leading order coefficient L_ξ of $e^{2\xi}$ in the region where $\xi > 0$ (and thus $\zeta < 0$). Neglecting lower order terms we compute

$$\begin{aligned}-i\Omega L_\xi &= \frac{1}{3}(-\xi^{2n} + \frac{1}{2}\zeta^{2n})J(\xi)J(\eta)J(\zeta) + \frac{1}{12}\zeta^{2n}K(\xi,\eta)\\ &= -\frac{1}{3}(-\xi^{2n}+\frac{1}{2}\zeta^{2n})\xi\zeta J(\eta) + \frac{1}{12}\zeta^{2n}(2\xi(\xi-\eta)J(\eta) - \xi(\eta - J(\eta))^2)\\ &= \frac{1}{3}(\xi^{2n}-\zeta^{2n})\xi\zeta J(\eta) - \frac{1}{12}\zeta^{2n}\xi(\eta+J(\eta))^2\\ &= -\frac{1}{3}2n\xi^{2n}\zeta\eta J(\eta) - \frac{1}{12}\zeta^{2n}\xi(\eta+J(\eta))^2.\end{aligned}$$

Thus, dividing by Ω we obtain

$$-iL_\xi = -\frac{1}{12}2n\xi^{2n}\eta + \frac{1}{48}\zeta^{2n}\frac{(\eta+J(\eta))^2}{J(\eta)}.$$

There is a second relevant term in the same region, namely the one with the $e^{-2\zeta}$ factor, which is obtained by the reflection symmetry and yields the complex conjugate of the previous contribution. Thus we get the statement in the proposition.

2. The symbol $\tilde{A}^{sym}$. We recall the expression for $\tilde{A}$:

$$\begin{aligned}\tilde{A}(\zeta,\xi,\eta) = &(e^{2\zeta}-1)\zeta^{2n}C^h(\xi,\eta) + (e^{-2\xi}-e^{2\eta})\zeta^{2n}C^a(\xi,-\eta)\\ &+\xi^{2n+1}(e^{2\xi}+1)(A^h(\zeta,\eta)+e^{2\eta}A^a(\zeta,-\eta)+e^{2\zeta}D^a(\eta,-\zeta)).\end{aligned}$$

Using the reflection symmetry for the first term we have

$$\begin{aligned}\tilde{A}(\zeta,\xi,\eta) \overset{\text{sym}}{=} &\,2\zeta^{2n}\sinh^2\zeta\ C^h(\xi,\eta) + (e^{-2\xi}-e^{2\eta})\zeta^{2n}C^a(\xi,-\eta)\\ &+\xi^{2n+1}(e^{2\xi}+1)(A^h(\zeta,\eta)+e^{2\eta}A^a(\zeta,-\eta)+e^{2\zeta}D^a(\eta,-\zeta)).\end{aligned}$$

We first verify that the symbol $\tilde{A}^{sym}$ vanishes on the edges. The edge $\zeta = 0$ requires that $\tilde{A}^{sym}(0,\xi,-\xi) = 0$. This needs no computation, instead it is a consequence of the fact that $\tilde{A}$ above is smooth and purely imaginary. Indeed, the symmetry in (ξ,η) corresponds to $\tilde{A}(0,\xi,-\xi) \to \tilde{A}(0,-\xi,\xi)$, whereas the reflection symmetry corresponds to the transformation $\tilde{A}(0,\xi,-\xi) \to -\tilde{A}(0,-\xi,\xi)$.

It remains to compute the edge $\xi = 0$, i.e., $\tilde{A}^{sym}(-\eta,0,\eta)$. In view of the symmetries and the fact that $\tilde{A}$ is purely imaginary, we have

$$4\tilde{A}^{sym}(-\eta,0,\eta) = \tilde{A}(-\eta,0,\eta) + \tilde{A}(-\eta,\eta,0) - \tilde{A}(\eta,0,-\eta) - \tilde{A}(\eta,-\eta,0).$$

So we proceed to compute

$$\begin{aligned}\Lambda(\eta)\tilde{A}(-\eta,0,\eta) &= (e^{-2\eta}-1)\eta^{2n}\Lambda(\eta)C^h(0,\eta) + (1-e^{2\eta})\eta^{2n}\Lambda(\eta)C^a(0,-\eta)\\ &= i\eta^{2n+1}\left(-e^{-2\eta}J(\eta)J'(\eta)+2J(\eta)-\eta J'(\eta)\right) - 2i\eta^{2n+1}J(\eta)J'(\eta).\end{aligned}$$

A similar computation yields

$$\begin{aligned}\Lambda(\eta)\tilde{A}(-\eta,\eta,0) = &(e^{-2\eta}-1)\eta^{2n}\Lambda(\eta)\left(C^h(\eta,0)+C^a(\eta,0)\right)\\ &+(e^{2\eta}+1)\eta^{2n+1}\left(A^h(-\eta,0)+A^a(-\eta,0)+e^{-2\eta}D^a(0,\eta)\right)\\ =&-i\eta^{2n+1}\left(-e^{-2\eta}J(\eta)J'(\eta)+2J(\eta)-\eta J'(\eta)\right)\\ &-i(4+3e^{2\eta}+3e^{-2\eta})\eta^{2n+1}J(\eta)J'(\eta)i(e^{2\eta}-e^{-2\eta})\eta^{2n+1}(2J(\eta)-\eta J'(\eta)).\end{aligned}$$

Combining these two we get $\tilde{A}^{sym}(-\eta,0,\eta) = 0$.

Finally we compute the high frequency asymptotics for $\tilde{A}^{sym}$. To be precise, we have $\tilde{A}^{sym} \in ES(\rho^{2n+2})$ and we compute its symbol modulo lower order terms in $ES(d^2\rho^{2n})$ near the edge $\zeta = 0$, respectively $ES(d^2d_1\rho^{2n-1}) + ES(d_1\rho^{2n})$ near the edges $\xi = 0$ and $\eta = 0$. Here A^a and C^a do not contribute to the principal part so we drop them.

First we consider the case when ζ is small and ξ and η are large. Neglecting terms with a ζ^2 factor we are left with

$$\xi^{2n+1}(e^{2\xi}+1)(A^h(\zeta,\eta)+e^{2\zeta}D^a(\eta,-\zeta)).$$

We only need to retain the factors with $e^{\pm 2\xi}$ and $e^{\pm 2\eta}$, which leaves us with

$$\xi^{2n+1}(e^{2\xi}A^h(\zeta,\eta)+e^{-2\eta}D^a(\eta,-\zeta)).$$

In view of the symmetries it suffices to compute the coefficient L_ξ of $e^{2\xi}$ when $\xi > 0$. This is given by, after symmetrization,

$$\begin{aligned} L_\xi &= \frac{1}{4}(\xi^{2n+1}A^h(\zeta,\eta) + \eta^{2n+1}D^a(-\xi,\zeta)) \\ &= \frac{1}{4}\left[-i\xi^{2n+1}\left(\eta + \frac{J(\zeta)^2-\zeta^2}{4J(\zeta)}\right) + i\eta^{2n+1}\xi\right] \\ &= i\eta\xi^{2n}\left(\frac{n}{2}\zeta + \frac{J(\zeta)^2-\zeta^2}{16J(\zeta)}\right), \end{aligned}$$

as required in the proposition.

Lastly, we consider the case when η is small, neglecting A^a, C^a and all the η^3 terms. Here D^a also does not contribute. Thus, as both A^h and C^h are odd and purely imaginary, applying the symmetries we need to consider the expression

$$\frac{1}{2}\zeta^{2n}(e^{2\zeta} + e^{-2\zeta})C^h(\xi,\eta) + \frac{1}{4}\xi^{2n+1}(e^{2\xi} - e^{-2\xi})A^h(\zeta,\eta).$$

By symmetry it suffices to consider the case that $\zeta > 0$ and $\xi < 0$. Thus, the leading order terms in the region $|\eta| \ll \zeta$ are given by

$$e^{2\zeta}\left(\xi^{2n}e^{2\eta} - \zeta^{2n}\right)\frac{i\eta(J(\eta)+\eta)}{8J(\eta)}\left(\xi - \frac{(J(\eta)-\eta)^2}{4J(\eta)}\right).$$

Using that

$$(e^{2\eta}-1)\frac{i\eta(J(\eta)+\eta)}{8J(\eta)} = \frac{1}{4}i\eta e^{2\eta},$$

and ignoring lower order terms we are left with

$$\frac{1}{4}i\xi^{2n+1}\eta e^{2\eta+2\zeta} - \frac{n}{4}i\xi^{2n}\eta^2 e^{2\eta+2\zeta}\frac{J(\eta)-\eta}{J(\eta)} - \frac{1}{16}i\xi^{2n}\eta e^{2\eta+2\zeta}\frac{(J(\eta)-\eta)^2}{J(\eta)}.$$

The second and third symbols yield contributions in the class $\eta ES(\rho^{2n})$ and hence they can be neglected. Thus, we are left with only the leading term,

$$\frac{1}{4}i\xi^{2n+1}\eta e^{2\eta+2\zeta},$$

and the final claim of the lemma follows.

□

7.3. **The high-low decomposition of $\tilde{A}$ and $\tilde{B}$.** The normal form energy is conserved to quartic order but, as our problem is quasilinear, we expect that its time derivative will contain more derivatives of (W,Q) than we want. The idea is then to remedy this issue by adding quartic (and higher order) quasilinear corrections to the normal form energy. Fortunately, in this problem it suffices to correct only the leading order terms in the normal form energy. Because of this, it is convenient to split the normal form energy into a leading part plus a lower order part,

$$\tilde{E}^n_{NF} = \tilde{E}^n_{NF,high} + \tilde{E}^n_{NF,low},$$

which corresponds to the decomposition of the trilinear forms $\tilde{A}$ and $\tilde{B}$ as

$$\tilde{A} = \tilde{A}_{high} + \tilde{A}_{low}, \qquad \tilde{B} = \tilde{B}_{high} + \tilde{B}_{low}. \tag{7.8}$$

The decomposition of the symbols is already given in the previous lemma, here we just compute the terms in the leading order part. To understand this decomposition it is useful to separate the generic case $n \geq 2$ from $n = 1$. For larger n we have:

Lemma 7.4. *Let $n \geq 2$. Then the trilinear forms $\tilde{A}$, $\tilde{B}$ admit a decomposition as in* (7.8) *where the symbols of $\tilde{A}_{low}$, $\tilde{B}_{low}$ satisfy*

$$\tilde{B}_{low} \in \xi\eta\zeta ES(d\rho^{2n-3}), \qquad \tilde{A}_{low} \in \xi\eta\zeta ES(dd_1\rho^{2n-3}) + \xi\eta\zeta ES(\rho^{2n-2}), \tag{7.9}$$

and the forms $\tilde{A}_{high}$, $\tilde{B}_{high}$ are given by

$$\begin{aligned} \tilde{B}_{high}(W,W,W) = &\ \langle W^{(n)}, W^{(n)}\rangle_{-4n \operatorname{Re} W_\alpha + \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha}, \\ \tilde{A}_{high}(W,Q,Q) = &\ -\langle Q^{(n)}, \mathcal{T}^{-1}Q^{(n+1)}\rangle_{-4n \operatorname{Re} W_\alpha - \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha} \\ &+ 2\langle Q_\alpha W^{(n)}, \mathcal{T}^{-1}Q^{(n+1)}\rangle + 2n\langle Q_{\alpha\alpha} W^{(n-1)}, \mathcal{T}^{-1}Q^{(n+1)}\rangle. \end{aligned} \tag{7.10}$$

On the other hand for $n = 1$ we have the more accurate result

Lemma 7.5. *Let $n = 1$. Then the trilinear forms $\tilde{A}$, $\tilde{B}$ admit a decomposition as in* (7.8) *where $\tilde{A}_{low}$, $\tilde{B}_{low}$ satisfy*

$$\tilde{B}_{low} \in \xi\eta\zeta ES(\rho^{-1}), \qquad \tilde{A}_{low} \in \xi\eta\zeta ES(1), \tag{7.11}$$

and $\tilde{A}_{high}$, $\tilde{B}_{high}$ are given by

$$\begin{aligned} \tilde{B}_{high}(W,W,W) = &\ \langle W_\alpha, W_\alpha\rangle_{-4 \operatorname{Re} W_\alpha + \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha}, \\ \tilde{A}_{high}(W,Q,Q) = &\ -\langle Q_\alpha, \mathcal{T}^{-1}Q_{\alpha\alpha}\rangle_{-4 \operatorname{Re} W_\alpha - \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha}. \end{aligned} \tag{7.12}$$

We remark that the difference in sign in the coefficient of $\frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha$ above accounts exactly for the linear part of the normal derivative of the pressure, namely a_1. The second line in $\tilde{A}_{high}$ in (7.10) is also natural and is due to the fact that (W, Q) is not a good set of variables for the differentiated equations. Instead, in the next subsection we switch from Q_α to the diagonal variable R and the bulk of these terms will disappear.

Proof of Lemma 7.4. We successively consider all the contributions in the leading part of $\tilde{A}^{sym}$ and $\tilde{B}^{sym}$.

1. The contribution of $\tilde{B}_{high}$. This is given by the symbol

$$\tilde{B}_{high} = -\frac{i}{48} e^{2\xi}\xi^{2n}\left(8n\eta - \frac{(\eta + J(\eta))^2}{J(\eta)}\right) + \text{symmetries}.$$

There are twelve symmetries, and after applying them all we obtain

$$-\int |W^{(n)}|^2 \left(4n \operatorname{Re} W_\alpha - \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re} W_\alpha\right) d\alpha. \tag{7.13}$$

Modulo lower order terms which can be included in $\tilde{B}_{low}$ this agrees with the expression for $\tilde{B}_{high}$ in the lemma.

2. The contribution of $\tilde{A}_{high}$ with high frequencies on Q. This is given by the symbol

$$\tilde{A}_{high} = e^{2\xi} i\xi^{2n}\eta \left(\frac{n}{2}\zeta + \frac{J(\zeta)^2 - \zeta^2}{12J(\zeta)}\right) + \text{symmetries}.$$

There are four symmetries, and after applying them all we obtain

$$\int \operatorname{Re}(i\bar{Q}^{(n+1)}Q^{(n)}) \left(4n \operatorname{Re} W_\alpha + \frac{1}{2}(1+\mathcal{T}^2) \operatorname{Re} W_\alpha\right) d\alpha. \tag{7.14}$$

Modulo lower order terms which can be included in $\tilde{A}_{low}$ this agrees with the first term in the expression for $\tilde{A}_{high}$ in the lemma.

3. The contribution of $\tilde{A}_{high}$ with high frequency on W. This is given by the symbol

$$\frac{1}{4}e^{2\eta+2\zeta}(i\xi)^{n+1}(i\zeta)^n\eta + \frac{n}{4}ie^{2\eta+2\zeta}(i\xi)^{n+1}(i\zeta)^{n-1}\eta^2 + \text{symmetries}.$$

There are four symmetries so we get the expression

$$2\operatorname{Re}\int i\bar{Q}^{(n+1)}W^{(n)}Q_\alpha \, d\alpha + 2n \operatorname{Re}\int i\bar{Q}^{(n+1)}W^{(n-1)}Q_{\alpha\alpha}\, d\alpha, \tag{7.15}$$

which up to lower order terms is equivalent to the second line in $\tilde{A}_{high}$.

□

Proof of Lemma 7.5. This follows the same steps as in the previous proof, with the only difference that some terms which were previously distinct are now combining.

1. Contribution of $\tilde{B}_{high}$. We may write

$$B^{sym} = L_\xi e^{2\xi} + \text{symmetries},$$

where the full symbol is given by

$$\begin{aligned} -12i\Omega L_\xi = &-(6\zeta\eta + \xi^2)J(\xi)J(\eta)J(\zeta) - \xi\zeta J(\xi)^2 J(\eta) - \xi\eta J(\xi)^2 J(\zeta) \\ &- \zeta\eta J(\xi)^3 - \zeta^2 J(\xi)J(\eta)^2 - \eta^2 J(\xi)J(\zeta)^2. \end{aligned}$$

Again it suffices to consider the region $\xi \gg |\eta|$. A similar computation to before gives us that

$$L_\xi e^{2\xi} = \frac{1}{6}i\xi\eta\zeta e^{2\xi} - \frac{1}{24}i\xi\zeta(\eta + J(\eta))^2 J(\eta)^{-1}e^{2\xi} + \xi\eta\zeta ES(\rho^{-1}).$$

Applying the symmetries and observing that the leading order term is already symmetric in η, ζ we obtain $\tilde{B}_{high}$.

2. Contribution of $\tilde{A}_{high}$. Again we may write

$$\tilde{A}^{sym} = L_\xi e^{2\xi} + L_\zeta e^{2\zeta} + \text{symmetries},$$

where the full symbols

$$\begin{aligned} -4i\Omega L_\xi = &\ \xi\eta(\eta^2 + \zeta^2 - 2\xi^2)J(\zeta)(J(\zeta) - J(\xi) - J(\eta)) \\ &+ \xi^2\eta(\zeta - \eta)J(\xi)(J(\xi) - J(\eta) - J(\zeta)) \\ &+ \xi\eta^2(\zeta - \eta)J(\xi)(J(\xi) - J(\eta) + J(\zeta)), \end{aligned}$$

$$
\begin{aligned}
-4i\Omega L_\zeta = {} & \xi\eta(\xi^2+\eta^2-2\zeta^2)J(\zeta)(J(\zeta)-J(\xi)-J(\eta)) \\
& - \xi^2\eta(\xi-\eta)J(\zeta)(J(\xi)-J(\eta)-J(\zeta)) \\
& - \xi\eta^2(\xi-\eta)J(\zeta)(J(\xi)-J(\eta)+J(\zeta)).
\end{aligned}
$$

In the region where $\xi \gg |\zeta|$ a similar computation to before gives us that

$$
L_\xi e^{2\xi} = \frac{1}{4} i\xi\eta\zeta(\xi-\eta)e^{2\xi} + \frac{1}{16} i\xi^2\eta(J(\zeta)^2-\zeta^2)J(\zeta)^{-1}e^{2\xi} + \xi\eta\zeta ES(1).
$$

The first term gives us part of the term

$$
-\langle Q_\alpha, \mathcal{T}^{-1}Q_{\alpha\alpha}\rangle_{-4\,\mathrm{Re}\, W_\alpha},
$$

and the second term gives us

$$
-\langle Q_\alpha, \mathcal{T}^{-1}Q_{\alpha\alpha}\rangle_{-\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\, W_\alpha}.
$$

In the region where $\zeta \gg |\eta|$ the coefficients L_ζ and $L_{-\xi}$ of $e^{2\zeta}$ and $e^{-2\xi}$ respectively combine to give

$$
L_\zeta e^{2\zeta} + L_{-\xi}e^{-2\xi} = -\frac{1}{4} i\xi\eta\zeta(\xi-\eta)e^{2\zeta+2\eta} + \xi\eta\zeta ES(1),
$$

which combines with the the first part of $\tilde{A}_{high}$ to give the rest of the term

$$
-\langle Q_\alpha, \mathcal{T}^{-1}Q_{\alpha\alpha}\rangle_{-4\,\mathrm{Re}\, W_\alpha}.
$$

□

7.4. **The normal form energy in the diagonal variables** $(\mathbf{W}, R)$**.** The normal form energy $\tilde{E}^n_{NF}$ constructed so far is expressed in terms of the variables (W, Q). Since we can smoothly extract factors of $\xi\eta\zeta$ from the symbols $\tilde{A}$, $\tilde{B}$, it is clear that one can view both $\tilde{A}$ and $\tilde{B}$ as trilinear forms in (W_α, Q_α),

$$
\tilde{B}(W,W,W) = \tilde{B}_1(\mathbf{W},\mathbf{W},\mathbf{W}), \qquad \tilde{A}(W,Q,Q) = \tilde{A}_1(\mathbf{W},Q_\alpha,Q_\alpha),
$$

where their symbols satisfy

$$
\tilde{B}_1 \in ES(\rho^{2n-2}), \qquad \tilde{A}_1 \in ES(\rho^{2n-1}).
$$

The same procedure applied separately to the high frequency parts $\tilde{A}_{high}$, $\tilde{B}_{high}$ respectively the lower order terms $\tilde{A}_{low}$, $\tilde{B}_{low}$ yields the forms $\tilde{A}_{1,high}$, $\tilde{B}_{1,high}$, respectively $\tilde{A}_{1,low}$, $\tilde{B}_{1,low}$, where the former are given for $n \geq 2$ by (see Lemma 7.4)

$$
\begin{aligned}
\tilde{B}_{1,high}(\mathbf{W},\mathbf{W},\mathbf{W}) = {} & \langle \mathbf{W}^{(n-1)}, \mathbf{W}^{(n-1)}\rangle_{-4n\,\mathrm{Re}\,\mathbf{W}+\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}} \\
\tilde{A}_{high}(\mathbf{W},Q_\alpha,Q_\alpha) = {} & -\langle \mathcal{T}^{-1}Q_\alpha^{(n)}, Q_\alpha^{(n-1)}\rangle_{-4n\,\mathrm{Re}\,\mathbf{W}-\frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}} \\
& + 2\langle \mathcal{T}^{-1}Q_\alpha^{(n)}, Q_\alpha\mathbf{W}^{(n-1)}\rangle + 2n\langle \mathcal{T}^{-1}Q_\alpha^{(n-1)}, Q_{\alpha\alpha}\mathbf{W}^{(n-1)}\rangle,
\end{aligned}
\tag{7.16}
$$

and the symbols for the latter have regularity

$$
\tilde{B}_{1,low} \in ES(d\rho^{2n-3}), \qquad \tilde{A}_{1,low} \in ES(dd_1\rho^{2n-3}) + ES(\rho^{2n-2}).
$$

In order to conclude the proof of Proposition 7.1 we need one last step, namely to further switch from $(\mathbf{W}, Q_\alpha)$ to the diagonal variables $(\mathbf{W}, R)$. This is still a purely algebraic computation, where we only need to insure that the original normal form energy $\tilde{E}^n_{NF}(W,Q)$ and the new one $E^n_{NF}(\mathbf{W},R)$ agree to cubic order,

$$\Lambda^{\leq 3}\left(\tilde{E}^n_{NF}(W,Q) - E^n_{NF}(\mathbf{W},R)\right) = 0.$$

We caution the reader that at this point, by a slight abuse of notation, we switch the meaning of the Λ operators. Whereas previously these were taken with respect to the expansion in the (W,Q) variables, from here on we use instead the expansion in the $(\mathbf{W},R)$ variables. It is a simple observation that the above relation has identical meaning in both frames of reference, and this allows for a smooth transition between one setting and the other.

To fulfill the above requirement each of the terms in $\tilde{E}^n_{NF}$ is treated as follows, based on the relation $Q_\alpha = R(1+\mathbf{W})$:

- The term $\tilde{B}_1(\mathbf{W},\mathbf{W},\mathbf{W})$ is left unchanged.
- The term $\tilde{A}_1(\mathbf{W},Q_\alpha,Q_\alpha)$ is replaced by $\tilde{A}_1(\mathbf{W},R,R)$.
- The term $\langle W^{(n)}, W^{(n)}\rangle = \langle \mathbf{W}^{(n-1)}, \mathbf{W}^{(n-1)}\rangle$ is left unchanged.
- The term $\langle Q^{(n)}, \mathcal{T}^{-1}Q^{(n)}_\alpha\rangle$ is replaced by the expression

$$\Lambda^{\leq 3}\langle [R(1+\mathbf{W})]^{(n-1)}, \mathcal{T}^{-1}[R(1+\mathbf{W})]^{(n-1)}_\alpha\rangle.$$

Rewriting the last expression as

$$\Lambda^{\leq 3}\langle [R(1+\mathbf{W})]^{(n-1)}, \mathcal{T}^{-1}[R(1+\mathbf{W})]^{(n-1)}_\alpha\rangle = \langle R^{(n-1)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle + 2\langle [R\mathbf{W}]^{(n-1)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle,$$

we can write our final normal form energy as

$$E^n_{NF}(\mathbf{W},R) = E_0(\mathbf{W}^{(n-1)},R^{(n-1)}) - 2\langle [R\mathbf{W}]^{(n-1)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle + \tilde{B}_1(\mathbf{W},\mathbf{W},\mathbf{W}) + \tilde{A}_1(\mathbf{W},R,R),$$

which is as required in Proposition 7.1, with

$$B(\mathbf{W},\mathbf{W},\mathbf{W}) = \tilde{B}_1(\mathbf{W},\mathbf{W},\mathbf{W}), \quad A(\mathbf{W},R,R) = \tilde{A}_1(\mathbf{W},R,R) - 2\langle [R\mathbf{W}]^{(n-1)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle.$$

By construction this has all the properties in part (a) of Proposition 7.1, as well as the required symbol regularity properties for the trilinear part. It remains to compute the high frequency parts B_{high} and A_{high}. For B_{high} there is nothing to compute, as we can take

$$B_{high} = \tilde{B}_{1,high},$$

with $\tilde{B}_{1,high}$ as in 7.4.

For A_{high} on the other hand we expand

$$\begin{aligned}\langle [R\mathbf{W}]^{(n-1)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle = &\langle R^{(n-1)}\mathbf{W}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle + \langle \mathbf{W}^{(n-1)}R, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle \\ &+ (n-1)\langle \mathbf{W}^{(n-2)}R_\alpha, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle + \text{l.o.t.}\end{aligned}$$

The last two terms cancel with the last two terms in $\tilde{A}_{1,high}(\mathbf{W},R,R)$ so we obtain

$$\begin{aligned}A_{high} = &-\langle \mathcal{T}^{-1}R^{(n)}, R^{(n-1)}\rangle_{-4n\,\mathrm{Re}\,\mathbf{W} - \frac{1}{2}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}} \\ &- 2\langle (R^{(n-1)}\mathbf{W}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle + 2\langle R_\alpha\mathbf{W}^{(n-2)}, \mathcal{T}^{-1}R^{(n-1)}_\alpha\rangle\end{aligned}$$

as needed in part (b(i)) of Proposition 7.1. To complete the proof of Proposition 7.1 we apply an identical computation for the case $n=1$.

8. Higher order energy estimates

The main goal of this section is to establish two energy bounds for $(\mathbf{W}, R)$, and their higher derivatives. Precisely, we will seek to obtain first short and then long time bounds for the time dependent quantities

$$\mathbf{N}_n := \|(g^{\frac{1}{2}}\mathbf{W}, R)\|_{H^{n-1}\times H^{n-\frac{1}{2}}}, \qquad n \geq 1.$$

For $n = 0$ we will instead set

$$\mathbf{N}_0 := \|(W, Q)\|_{\mathcal{H}},$$

which is closely related to the conserved energy.

Our first result is a quadratic bound, which applies to all solutions independently of the size of the initial data. This is needed for our local well-posedness result in Theorem 1. Precisely, the large data result is as follows:

Proposition 8.1. *For any $n \geq 1$ there exists an energy functional $E^{n,(2)}(\mathbf{W}, R)$ with the following properties:*

(i) Norm equivalence:

$$E^{n,(2)}(\mathbf{W}, R) \approx_A E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R) + O_A(\mathbf{N}_{n-1}^2).$$

(ii) Quadratic energy estimates for solutions to (8.3)*:*

$$\frac{d}{dt}E^{n,(2)}(\mathbf{W}, R) \lesssim_A B\mathbf{N}_n^2.$$

Here we allow lower order errors in the energy equivalence, and thus, the bound for $\mathbf{N}_k$ for instance is obtained by reiterating the above estimates for $1 \leq n \leq k$, and using the energy conservation as a starting point which corresponds to $n = 0$.

Our second estimate is a cubic bound which only applies for small solutions, and is used to prove our cubic lifespan result in Theorem 2. The small data result is as follows:

Proposition 8.2. *For any $n \geq 1$ there exists an energy functional $E^{n,(3)}$ which has the following properties as long as $A \ll 1$:*

(i) Norm equivalence:

$$E^{n,(3)}(\mathbf{W}, R) = E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R) + O(A)\mathbf{N}_n^2. \tag{8.1}$$

(ii) Cubic energy estimates:

$$\frac{d}{dt}E^{n,(3)}(\mathbf{W}, R) \lesssim_A AB\mathbf{N}_n^2. \tag{8.2}$$

The first step in the analysis will be to isolate the main part of the systems for $(\mathbf{W}, R)$ and for their derivatives, and derive quadratic energy estimates for it. A key part in this will be played by the model system studied in Section (4). This model system plays the same role in this paper as the linearized system played in the analysis of the infinite depth water waves, (see [8]). However, this correspondence is incomplete, in that here we do not have cubic estimates for the linearized system, but we do have them for the system in $(\mathbf{W}, R)$ and also for its higher derivatives.

We will first differentiate the equations, and prove the large data result using the bounds for the model problem in Proposition 4.1. Then we consider the small data problem, and combine the prior high frequency analysis with the normal form energy derived in the previous section.

8.1. **The case** $n = 1$**.** We begin by looking at the $(\mathbf{W}, R)$ system (1.8). This is a self-contained diagonal system in these variables, which we rewrite in a form which is similar to the model problem in Proposition 4.1:

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \dfrac{1}{1+\bar{\mathbf{W}}}R_\alpha - \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathcal{T}^2\mathbf{W} = \mathcal{G} \\ R_t + bR_\alpha - \dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}} = \mathcal{K}, \end{cases} \tag{8.3}$$

where

$$\mathcal{G} := (1+\mathbf{W})M - \frac{R_\alpha}{1+\bar{\mathbf{W}}}(1+\mathcal{T}^2)[\mathbf{W}], \quad \mathcal{K} := -2i\frac{\operatorname{Im}\mathbf{P}\left[R\bar{R}_\alpha\right]}{1+\mathbf{W}} - \frac{\mathfrak{a}\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}}.$$

In order to view this system as an evolution in the space of holomorphic functions, we project (8.3) onto the space of holomorphic functions via the projection operator $\mathbf{P}$:

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \mathbf{P}\left[\dfrac{1}{1+\bar{\mathbf{W}}}R_\alpha\right] - \mathbf{P}\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathcal{T}^2\mathbf{W}\right] = \mathbf{P}\mathcal{G} \\ R_t + bR_\alpha - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}}\right] = \mathbf{P}\mathcal{K}. \end{cases} \tag{8.4}$$

We also recall here the expressions of b, a and M

$$b = 2\operatorname{Re}\left[R - \mathbf{P}[R\bar{Y}]\right],$$

and

$$\mathfrak{a} = g\mathbf{W} + ig\mathcal{T}[\mathbf{W}] + 2\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha], \quad M = 2\operatorname{Re}\mathbf{P}[R\bar{Y}_\alpha - \bar{R}_\alpha Y].$$

To this system we associate the positive definite linear functional energy $E^{(2)}_{lin}(\mathbf{W}, R)$ given by (4.3). The main result of this subsection establishes energy bounds for the system (8.4), thus proving the $n = 1$ part of Proposition 8.1

Proposition 8.3. *The above energy applied to solutions of projected system* (8.4) *satisfies the following estimates:*

i) *Norm equivalence:*

$$E^{(2)}_{lin}(\mathbf{W}, R) \approx_A \|(\mathbf{W}, R)\|^2_{\mathcal{H}}.$$

ii) *Cubic energy estimates:*

$$\frac{d}{dt}E^{(2)}_{lin}(\mathbf{W}, R) \lesssim_A B\mathbf{N}_1^2.$$

Here the energy equivalence follows directly from the positivity and boundedness of $\mathfrak{a}$, see Proposition 1.12 and Lemmma A.13. The second estimate in the proposition relies on the estimates obtained in (4.5). Precisely, in order to obtain the quadratic energy estimates for the large data, it suffices to prove a priori bounds for $\|(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})\|_{\mathcal{H}}$. These a priori bounds will be in terms of the pointwise control norms A, B and the energy $E^{(2)}_{lin}(\mathbf{W}, R)$:

Lemma 8.4. *The following estimates for the lower order terms* $(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})$*:*

$$\|(\mathbf{P}\mathcal{G}, \mathbf{P}\mathcal{K})\|_{\mathcal{H}} \lesssim_A B\mathbf{N}_1. \tag{8.5}$$

Proof. We begin with the estimate for $\mathcal{G}$:

$$\|\mathbf{P}\mathcal{G}\|_{\mathcal{H}} \lesssim \|\mathbf{P}M\|_{\mathcal{H}} + \|\mathbf{P}[\mathbf{W}M]\|_{L^2} \lesssim_A (B + AB)\|R\|_{H^{\frac{1}{2}}}.$$

To bound $\mathcal{K}$ we estimate each of the terms separately. Thus, for the first term we estimate

$$\begin{aligned}\| \langle D\rangle^{\frac{1}{2}} \left[\mathbf{P}\left[\operatorname{Im} \bar{\mathbf{P}}[\bar{R}R_\alpha]\right](1-Y)\right] \|_{\mathcal{H}} &\lesssim \| \langle D\rangle^{\frac{1}{2}} \bar{\mathbf{P}}[\bar{R}R_\alpha]\|_{\mathcal{H}} + \| \langle D\rangle^{\frac{1}{2}} \bar{\mathbf{P}}[\bar{R}R_\alpha]Y\|_{\mathcal{H}} \\ &\lesssim \|R\|_{\text{bmo}^1}\|R\|_{H^{\frac{1}{2}}} + \|Y\|_{\text{bmo}^{\frac{1}{2}}}\| \langle D\rangle^{\frac{1}{2}} R\|_{L^\infty}\|R\|_{H^{\frac{1}{2}}} \\ &\lesssim_A (B + AB)\|R\|_{H^{\frac{1}{2}}}.\end{aligned}$$

This is a direct consequence of the commutator estimate (A.14) together with the Y estimate derived in (A.16). As for the second term, we use the estimates derived for $\mathfrak{a}$ and Y in (A.20) respectively (A.16), to arrive at

$$\|\mathfrak{a}\mathcal{T}[\mathbf{W}](1-Y)\|_{H^{\frac{1}{2}}} \lesssim (\|\mathfrak{a}\|_{\text{bmo}^{\frac{1}{2}}} + \|\mathfrak{a}\|_{L^\infty}\|Y\|_{\text{bmo}^{\frac{1}{2}}})\|\mathbf{W}\|_{L^2}.$$

□

For the small data problem it is of further interest to track the solutions on larger time scales in order to prove Proposition 8.2. This is done at the end of this section.

8.2. **The case $n = 2$.** We recall that the system (1.8) for $(\mathbf{W}, R)$ is given by

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \dfrac{1}{1+\bar{\mathbf{W}}}R_\alpha + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W} = (1+\mathbf{W})M \\ R_t + bR_\alpha - \dfrac{g\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}} + \dfrac{ia}{1+\mathbf{W}} = 0, \end{cases}$$

where a is the same as in the infinite depth gravity water waves

$$a := 2\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha].$$

We differentiate with respect to α in order to obtain a system for $(\mathbf{W}_\alpha, R_\alpha)$,

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{[(1+\mathbf{W})R_\alpha]_\alpha}{1+\bar{\mathbf{W}}} = -b_\alpha\mathbf{W}_\alpha - (1+\mathbf{W})R_\alpha\bar{Y}_\alpha + \mathbf{W}_\alpha M + (1+\mathbf{W})M_\alpha \\ R_{\alpha t} + bR_{\alpha\alpha} - \dfrac{g\mathcal{T}[\mathbf{W}_\alpha]}{1+\mathbf{W}} + \dfrac{g\mathcal{T}[\mathbf{W}]}{(1+\mathbf{W})^2}\mathbf{W}_\alpha + \dfrac{ia_\alpha}{1+\mathbf{W}} - \dfrac{ia}{(1+\mathbf{W})^2}\mathbf{W}_\alpha = -b_\alpha R_\alpha, \end{cases}$$

and rewrite it as follows

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{[(1+\mathbf{W})R_\alpha]_\alpha}{1+\bar{\mathbf{W}}} = -b_\alpha\mathbf{W}_\alpha - (1+\mathbf{W})R_\alpha\bar{Y}_\alpha + \mathbf{W}_\alpha M + (1+\mathbf{W})M_\alpha \\ R_{\alpha t} + bR_{\alpha\alpha} - \left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}_\alpha]}{(1+\mathbf{W})^2}\right] = \dfrac{ia - g\mathcal{T}[\mathbf{W}]}{(1+\mathbf{W})^2}(1+i\mathcal{T})\mathbf{W}_\alpha - \dfrac{ia_\alpha}{1+\mathbf{W}} - b_\alpha R_\alpha. \end{cases}$$

We recall that

$$M = \frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}} - b_\alpha, \tag{8.6}$$

and use this definition to simplify the system above

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{[(1+\mathbf{W})R_\alpha]_\alpha}{1+\bar{\mathbf{W}}} = \left(M - \dfrac{R_\alpha}{1+\bar{\mathbf{W}}} - \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right)\mathbf{W}_\alpha \\ \qquad\qquad\qquad\qquad\qquad - (1+\mathbf{W})R_\alpha \bar{Y}_\alpha + \mathbf{W}_\alpha M + (1+\mathbf{W})M_\alpha \\ R_{\alpha t} + bR_{\alpha\alpha} - \left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}_\alpha]}{(1+\mathbf{W})^2}\right] = \dfrac{ia - g\mathcal{T}[\mathbf{W}]}{(1+\mathbf{W})^2}(1+i\mathcal{T})\mathbf{W}_\alpha - \dfrac{ia_\alpha}{1+\mathbf{W}} \\ \qquad\qquad\qquad\qquad\qquad + \left(M - \dfrac{R_\alpha}{1+\bar{\mathbf{W}}} - \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right)R_\alpha. \end{cases}$$

To align this more closely with the linearized equation and express the system in a manner similar to [8], we introduce the auxiliary holomorphic variable

$$\mathbf{R} = (1+\mathbf{W})R_\alpha.$$

Then it becomes

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}_\alpha = -\dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}_\alpha + \mathbf{R}\bar{Y}_\alpha + 2\mathbf{W}_\alpha M + (1+\mathbf{W})M_\alpha \\ \\ \mathbf{R}_t + b\mathbf{R}_\alpha - \left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}_\alpha]}{(1+\mathbf{W})}\right] = \dfrac{ia - g\mathcal{T}[\mathbf{W}]}{(1+\mathbf{W})}(1+i\mathcal{T})\mathbf{W}_\alpha + (\bar{R}_\alpha R_\alpha - ia_\alpha) + 2\mathbf{R}M \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad - 2\left(\dfrac{R_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right)\mathbf{R}. \end{cases}$$

Here, we have isolated on the left the leading part of our equations. The goal is to interpret the terms on the right hand side as perturbative (with one exception, which is only due to the low regularity setting, see below). In addition, for the cubic bound we will also need to pay attention to the quadratic part of the terms in the equations.

In order to simplify our bookkeeping we define two types of error terms for the above system. These are denoted by $\mathbf{err}(L^2)$ and $\mathbf{err}(H^{\frac{1}{2}})$, which correspond to the two equations. A similar strategy was employed in [8]. However, unlike in [8], here we also include bounded quadratic terms into the error, rather than explicitly keeping track of them. This simplifies the argument somewhat, at the expense of getting a less precise expression for the normal form energy.

The bounds for these errors are in terms of the control variables A, B, as well as the L^2 type norm $\mathbf{N}_2$, where

$$\mathbf{N}_2 := \|(g^{\frac{1}{2}}\mathbf{W}, R)\|_{H^1\times H^{\frac{3}{2}}}.$$

The acceptable errors in the $\mathbf{W}_\alpha$ equation are denoted, by $\mathbf{err}(L^2)$ and are of two types, $\mathbf{err}(L^2)^{[2]}$ and $\mathbf{err}(L^2)^{[3]}$. The first one, $\mathbf{err}(L^2)^{[2]}$, consists of quadratic terms which satisfy the bounds

$$\|\mathcal{T}\mathbf{P}\mathcal{G}\|_{L^2} \lesssim B\mathbf{N}_2, \quad \|\mathcal{G}\|_{H^{-\frac{1}{2}}} \lesssim A\mathbf{N}_2.$$

By $\mathbf{err}(L^2)^{[3]}$ we denote the cubic and higher counterpart of $\mathbf{err}(L^2)^{[2]}$, which contains terms $\mathcal{G}$ which satisfy the estimate

$$\|\mathcal{T}\mathbf{P}\mathcal{G}\|_{L^2} \lesssim_A AB\mathbf{N}_2, \quad \|\mathcal{G}\|_{H^{-\frac{1}{2}}} \lesssim_A A^2\mathbf{N}_2.$$

The acceptable errors in the $\mathbf{R}$ equation are denoted by $\mathbf{err}(H^{\frac{1}{2}})$ and are of two types, $\mathbf{err}(H^{\frac{1}{2}})^{[2]}$ and $\mathbf{err}(H^{\frac{1}{2}})^{[3]}$. The first one, $\mathbf{err}(H^{\frac{1}{2}})^{[2]}$, consists of quadratic terms $\mathcal{K}$ that satisfy the bounds

$$\|\mathcal{T}\mathbf{P}\mathcal{K}\|_{H^{\frac{1}{2}}} \lesssim B\mathbf{N}_2, \quad \|\mathcal{K}\|_{L^2} \lesssim A\mathbf{N}_2.$$

By $\mathbf{err}(H^{\frac{1}{2}})^{[3]}$ we denote terms in K which satisfy the estimates

$$\|\mathcal{T}\mathbf{P}\mathcal{K}\|_{H^{\frac{1}{2}}} \lesssim_A AB\mathbf{N}_2, \quad \|\mathcal{K}\|_{L^2} \lesssim_A A^2\mathbf{N}_2.$$

Remark 8.5. Compared to [8], above we define $\mathbf{N}_2$ in a more relaxed, inhomogeneous fashion. This is in part caused by the lack of scaling. It is reasonable because here we work with the system for the differentiated variables $(\mathbf{W}, R)$ or their higher counterparts, which is used to bound the high frequencies of the solutions.

A key property of the space of errors is contained in the following lemma:

Lemma 8.6. *Let Φ be a function which satisfies*

$$\|\Phi\|_{L^\infty} \lesssim A, \quad \|\Phi\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim B. \tag{8.7}$$

Then, we have the multiplicative bounds

$$\Phi \cdot \mathbf{err}(L^2) = \mathbf{err}(L^2), \quad \Phi \cdot \mathbf{err}(H^{\frac{1}{2}}) = \mathbf{err}(H^{\frac{1}{2}}). \tag{8.8}$$

The proof of the lemma is relatively straightforward and is left for the reader.

We now return to system above and expand some of the terms. We begin with the terms containing M. For this we will make use of the bounds we have established for M in the Appendix:

$$\|M\|_{L^\infty} \lesssim AB, \quad \|M\|_{H^{\frac{1}{2}}} \lesssim A\mathbf{N}_2, \quad \|M_\alpha\|_{L^2} \lesssim A\mathbf{N}_2. \tag{8.9}$$

Precisely, the M terms in the equations satisfy

$$M(1 + \mathbf{W}_\alpha) + M_\alpha \mathbf{W} = \mathbf{err}(L^2), \quad M\mathbf{R} = \mathbf{err}(H^{\frac{1}{2}}).$$

The first claim is a straightforward consequence of the pointwise bound for M and the L^2 bound for M_α. For the second, we recall that $\mathbf{R} = R_\alpha(1+\mathbf{W})$, which together with Lemma 8.6 allows us to only estimate MR_α. The $H^{\frac{1}{2}}$ bound for MR_α follows after a Littlewood-Paley decomposition of the product: the bounds for the low-high and balanced interactions are a direct consequence of (8.9), and the bounds for high-low interactions are obtained by combining (8.9) and Lemma A.9.

Next we consider the expression

$$\frac{ia - g\mathcal{T}[\mathbf{W}]}{(1 + \mathbf{W})}(1 + i\mathcal{T})\mathbf{W}_\alpha = \frac{ia - g\mathcal{T}[\mathbf{W}]}{(1 + \mathbf{W})}(1 + \mathcal{T}^2)\operatorname{Re}\mathbf{W}_\alpha,$$

which we claim belongs to $\mathbf{err}(H^{\frac{1}{2}})$. To prove our claim, we split the above expression into a quadratic part and a cubic and higher term,

$$-g\mathcal{T}[\mathbf{W}](1 + \mathcal{T}^2)\operatorname{Re}\mathbf{W}_\alpha + (ia + g\mathcal{T}[\mathbf{W}]\mathbf{W})(1 - Y)\left[(1 + \mathcal{T}^2)\operatorname{Re}\mathbf{W}_\alpha\right].$$

We may then apply the paraproduct estimates (A.1) and (A.7) with the estimates (A.20) for $\mathfrak{a}$ (which also applies to the component a of $\mathfrak{a}$) and (A.16), (A.17) for Y to obtain:

$$\begin{aligned}&\|(ia+g\mathcal{T}[\mathbf{W}]\mathbf{W})(1-Y)\left[(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}_\alpha\right]\|_{H^{\frac{1}{2}}}\\ \lesssim&\left\{A\left(\|a\|_{\mathrm{bmo}^{\frac{1}{2}}}+\|Y\|_{\mathrm{bmo}^{\frac{1}{2}}}\right)+\left(g\|\langle D\rangle^{\frac{1}{2}}\mathbf{W}\|_{\mathrm{bmo}}\|\mathbf{W}\|_{L^\infty}+\|\mathbf{W}\|_{L^\infty}\|Y\|_{\mathrm{bmo}^{\frac{1}{2}}}\right)\right\}\|\mathbf{W}_\alpha\|_{L^2}.\end{aligned}$$

Similarly, for the quadratic part we obtain

$$\|g\mathcal{T}[\mathbf{W}](1+\mathcal{T}^2)(\operatorname{Re}\mathbf{W}_\alpha)\|_{H^{\frac{1}{2}}}\lesssim_A A\mathbf{N}_2.$$

Next we consider the difference

$$\bar{R}_\alpha R_\alpha - ia_\alpha = 2\bar{\mathbf{P}}[R_\alpha\bar{R}_\alpha]+i\operatorname{Im}\mathbf{P}[R\bar{R}_{\alpha\alpha}].$$

For this we bound

$$\|\bar{R}_\alpha R_\alpha - ia_\alpha\|_{L^2}\lesssim A\mathbf{N}_2,\qquad \|\mathbf{P}\mathcal{T}(\bar{R}_\alpha R_\alpha - ia_\alpha)\|_{H^{\frac{1}{2}}}\lesssim B\mathbf{N}_2.$$

Taking into account all the above bounds, it follows that our system can be rewritten in the form

$$\begin{cases}\mathbf{W}_{\alpha t}+b\mathbf{W}_{\alpha\alpha}+\dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}}+\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}_\alpha = 2\mathbf{R}\bar{Y}_\alpha - 2\dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}_\alpha+\mathbf{err}(L^2)\\[2ex] \mathbf{R}_t+b\mathbf{R}_\alpha-\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}_\alpha]}{(1+\mathbf{W})}\right]=-4\operatorname{Re}\left(\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R}+\mathbf{err}(H^{\frac{1}{2}}).\end{cases}$$

One might wish to compare this system to the model system (4.1), for which we obtained the nice energy estimates in (4.5), and use these estimates to prove quadratic energy bounds provided that the right hand side terms are bounded in L^2, and $H^{\frac{1}{2}}$ respectively.

Unfortunately we still have terms on the right which cannot be bounded as error terms, i.e., in $L^2\times H^{\frac{1}{2}}$. This matches similar issues appearing in the infinite depth case in [8]. To deal with these terms we use the same conjugation with respect to a real exponential weight $e^{2\phi}$, where $\phi=-2\operatorname{Re}\log(1+\mathbf{W})$, which was previously used in [8]. When implementing such a transformation, we are not only able to eliminate the unbounded terms pointed out above, but we also manage to cast our system in a similar form as the model system in (4.1).

To see this, we compute

$$\phi_\alpha = -2\operatorname{Re}\frac{\mathbf{W}_\alpha}{1+\mathbf{W}},\qquad (\partial_t+b\partial_\alpha)\phi = 2\operatorname{Re}\frac{R_\alpha}{1+\bar{\mathbf{W}}}-2M.$$

We denote the weighted variables by

$$w:=e^{2\phi}\mathbf{W}_\alpha,\quad r:=e^{2\phi}\mathbf{R}.$$

Before explicitly writing down the resulting equations, we remark that by Lemma 8.6 we have

$$e^{2\phi}\mathbf{err}(L^2)=\mathbf{err}(L^2),\quad e^{2\phi}\mathbf{err}(H^{\frac{1}{2}})=\mathbf{err}(H^{\frac{1}{2}}),$$

which simplifies the transformed system to

$$\begin{cases} w_t + bw_\alpha + \dfrac{r_\alpha}{1+\bar{\mathbf{W}}} - \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathcal{T}^2 w = \mathbf{err}(L^2) \\ \\ r_t + br_\alpha - \left[\dfrac{(g+\mathfrak{a})\mathcal{T}[w]}{(1+\mathbf{W})}\right] = \mathbf{err}(H^{\frac{1}{2}}). \end{cases}$$

Here we have also harmlessly replaced w by $-\mathcal{T}^2 w$ in the last term on the left in the first equation. The difference is easily included in the error as $1+\mathcal{T}^2$ has a Schwartz symbol. This is done in order to bring the above equations more in line with the model linear problem.

Unfortunately our new variables (w, r), are not exactly holomorphic; the last system contains both holomorphic and also antiholomorphic components. To remedy this issue we need to project the system via the projection $\mathbf{P}$, and also work with the projected variables $(\mathbf{P}w, \mathbf{P}r)$. At this point one might legitimately be concerned that restricting to the holomorphic part would remove a good portion of our variables. However this is not the case, as one can verify that the a similar argument as the one in Lemma 3.4 from [8] applies to the finite depth case:

Proposition 8.7. *The energy of* $(\mathbf{P}w, \mathbf{P}r)$ *above is equivalent to the energy of* $(\mathbf{W}_\alpha, R_\alpha)$*:*

$$\|(\mathbf{P}w, \mathbf{P}r)\|_{\mathcal{H}} \sim_A \|(w,r)\|_{\mathcal{H}} \sim_A \|(\mathbf{W}_\alpha, R_\alpha)\|_{\mathcal{H}} \text{ modulo } A\mathbf{N}_2. \tag{8.10}$$

Unlike in [8], here we allow for lower order errors in order to account for the L^2 unboundedness of $\mathbf{P}$ at low frequencies. Once we do that, it remains to prove only a high frequency bound, for which the same argument as in [8] applies.

We are now ready to write the system for (Pw, Pr), namely

$$\begin{cases} \mathbf{P}w_t + \mathbf{P}\left[b\mathbf{P}w_\alpha\right] + \mathbf{P}\left[\dfrac{\mathbf{P}r_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathcal{T}^2\mathbf{P}w\right] = G_2 + \mathbf{P}\mathbf{err}(L^2) \\ \\ \mathbf{P}r_t + \mathbf{P}\left[b\mathbf{P}r_\alpha\right] - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{P}w]}{(1+\mathbf{W})}\right] = K_2 + \mathbf{P}\mathbf{err}(H^{\frac{1}{2}}), \end{cases} \tag{8.11}$$

where (G_2, K_2) contain all the additional terms,

$$\begin{cases} G_2 := -\mathbf{P}\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}(1+\mathcal{T}^2)\mathbf{P}w\right] - \mathbf{P}\left[b\bar{\mathbf{P}}w_\alpha\right] - \mathbf{P}\left[\dfrac{\bar{\mathbf{P}}r_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\bar{\mathbf{P}}w\right] \\ K_2 := -\mathbf{P}\left[b\bar{\mathbf{P}}r_\alpha\right] + \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\bar{\mathbf{P}}w]}{(1+\mathbf{W})}\right]. \end{cases}$$

The goal here is to prove that $G_2 = \mathbf{err}(L^2)$ and $K_2 = \mathbf{err}(H^{\frac{1}{2}})$, but this is straightforward as they all have a nice commutator structure; the proof is left for the reader. We denote the last set of variables by $(Pw, Pr) := (\mathfrak{w}, \mathfrak{r})$; these solve the system

$$(8.12)\qquad \begin{cases} \mathfrak{w}_t + \mathbf{P}\left[b\mathfrak{w}_\alpha\right] + \mathbf{P}\left[\dfrac{\mathfrak{r}_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathcal{T}^2\mathfrak{w}\right] = \mathbf{P}\,\mathbf{err}(L^2) \\ \\ \mathfrak{r}_t + \mathbf{P}\left[br_\alpha\right] - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathfrak{w}]}{(1+\mathbf{W})}\right] = \mathbf{P}\,\mathbf{err}(H^{\frac{1}{2}}). \end{cases}$$

Therefore, we can now apply the energy bounds obtained for the toy model (4.1) to the system (8.12). Now the result of Proposition (8.1) follows from the energy estimates for the model system (4.1), namely (4.5); further, if $n = 2$ then we can take

$$E^{n,(2)}(\mathbf{W}, R) = E^{(2)}_{lin}(\mathfrak{w}, \mathfrak{r}).$$

The last goal is to obtain cubic lifespan bounds for the small data problem, which would correspond to proving Proposition (8.2). We address this question later in this section.

8.3. **The case $n \geq 3$.** We follow the same strategy as in the case $n = 2$ and derive the equations for $(\mathbf{W}^{(n-1)}, R^{(n-1)})$. For this, we start with the system (1.8) and differentiate $(n-1)$ times. For this we will estimate the errors in terms of $\mathbf{N}_n$ which measures $(n-1)$ derivatives of $\mathbf{W}$ and R, with constants that depend on the control norms A and B.

The acceptable errors in the $\mathbf{W}^{(n-1)}$ equation are denoted, as before, by $\mathbf{err}(L^2)$ and are of two types, $\mathbf{err}(L^2)^{[2]}$ and $\mathbf{err}(L^2)^{[3]}$. The first one, $\mathbf{err}(L^2)^{[2]}$, consists of holomorphic quadratic terms in G of the form that satisfy the bound

$$\|\mathcal{T}\mathbf{P}G\|_{L^2} \lesssim B\mathbf{N}_n \quad \text{and } \|G\|_{H^{-\frac{1}{2}}} \lesssim A\mathbf{N}_n.$$

By $\mathbf{err}(L^2)^{[3]}$ we denote the cubic counterpart of $\mathbf{err}(L^2)$ of G, which satisfies the estimate

$$\|\mathcal{T}\mathbf{P}G\|_{L^2} \lesssim_A AB\mathbf{N}_n, \quad \|G\|_{H^{-\frac{1}{2}}} \lesssim_A A^2\mathbf{N}_n.$$

The acceptable errors in the $R^{(n-1)}$ equation are denoted, as before, by $\mathbf{err}(H^{\frac{1}{2}})$ and are of two types, $\mathbf{err}(H^{\frac{1}{2}})^{[2]}$ and $\mathbf{err}(H^{\frac{1}{2}})^{[3]}$. The first one, $\mathbf{err}(H^{\frac{1}{2}})$, consists of holomorphic quadratic terms in K that satisfy the bound

$$\|\mathcal{T}\mathbf{P}K\|_{H^{\frac{1}{2}}} \lesssim B\mathbf{N}_n, \quad \|K\|_{L^2} \lesssim A\mathbf{N}_n.$$

By $\mathbf{err}(H^{\frac{1}{2}})^{[3]}$ we denote terms in K which satisfy the estimates

$$\|\mathcal{T}\mathbf{P}K\|_{H^{\frac{1}{2}}} \lesssim_A AB\mathbf{N}_n, \quad \|K\|_{L^2} \lesssim_A A^2\mathbf{N}_n.$$

We begin by differentiating the terms in the $\mathbf{W}$ equation. For the b term, after standard estimates, we have

$$\begin{aligned} \partial_\alpha^{(n-1)}(b\mathbf{W}_\alpha) &= b\mathbf{W}_\alpha^{(n-1)} + (n-1)b_\alpha \mathbf{W}^{(n-1)} + \mathbf{err}(L^2) \\ &= b\mathbf{W}_\alpha^{(n-1)} + (n-1)\left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}}{1+\mathbf{W}}\right)\mathbf{W}^{(n-1)} + \mathbf{err}(L^2). \end{aligned}$$

Here we have used the relation (8.6), and also the L^∞ bound for M.

Continuing, we apply the same analysis for the $(n-1)$ derivative of the next term appearing in the $\mathbf{W}$ equation

$$\partial_\alpha^{(n-1)}\frac{(1+\mathbf{W})R_\alpha}{1+\bar{\mathbf{W}}} =\frac{\left[(1+\mathbf{W})R^{(n-1)}\right]_\alpha}{1+\bar{\mathbf{W}}}+\frac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}^{(n-1)}+\mathbf{err}(L^2). \tag{8.13}$$

Here we have again isolated the terms which cannot be placed into the error.

Similarly, using the bounds for M in Lemma A.15, the last component of the $\mathbf{W}$ equation is

$$\partial_\alpha^{(n-1)}\left[(1+\mathbf{W})M\right]=(1+\mathbf{W})\partial_\alpha^{(n-1)}M+\mathbf{err}(L^2)=(1+\mathbf{W})\partial_\alpha^{(n-1)}2\operatorname{Re}\mathbf{P}\left[R\bar{Y}_\alpha-\bar{R}_\alpha Y\right].$$

Because of the differentiation, there are no low frequency issues here. Distributing derivatives inside, the terms with derivatives on the antiholomorphic factors are all errors, so we are left only with the terms where all derivatives apply to the holomorphic factors. Harmlessly discarding the projection we arrive at

$$\begin{aligned}\partial_\alpha^{(n-1)}\left[(1+\mathbf{W})M\right]&=2(1+\mathbf{W})\operatorname{Re}\left[R^{(n-1)}\bar{Y}_\alpha-\bar{R}_\alpha Y^{(n-1)}\right]+\mathbf{err}(L^2)\\&=-\frac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)}+\mathbf{err}(L^2).\end{aligned}$$

Next, we differentiate the R equation. We begin with

$$\partial_\alpha^{(n-1)}(bR_\alpha) =bR_\alpha^{(n-1)}+(n-1)b_\alpha R^{(n-1)}+b^{(n-1)}R_\alpha+\mathbf{err}(H^{\frac{1}{2}}).$$

In the second term we use again the relation (8.6) and the boundedness of M. In the third term, only the holomorphic part of b yields a nontrivial contribution, and that only when all derivatives apply to Y. Discarding again the projection, we obtain

$$\partial_\alpha^{(n-1)}(bR_\alpha)=bR_\alpha^{(n-1)}+(n-1)\left(\frac{R_\alpha}{1+\bar{\mathbf{W}}}+\frac{\bar{R}_\alpha}{1+\mathbf{W}}\right)R^{(n-1)}+\frac{R_\alpha}{1+\bar{\mathbf{W}}}R^{(n-1)}+\mathbf{err}(H^{\frac{1}{2}}).$$

For the remaining terms in the R equation we write

$$\partial_\alpha^{(n-1)}\left(\frac{g\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}}\right)=\frac{g\mathcal{T}[\mathbf{W}^{(n-1)}]}{1+\mathbf{W}}-\frac{g\mathcal{T}[\mathbf{W}]}{(1+\mathbf{W})^2}\mathbf{W}^{(n-1)}+\mathbf{err}(H^{\frac{1}{2}}). \tag{8.14}$$

Lastly, using the bound for $\mathfrak{a}$ in Lemma A.13, we have

$$\partial_\alpha^{(n-1)}\left(\frac{i\mathfrak{a}}{1+\mathbf{W}}\right)=\frac{i\mathfrak{a}^{(n-1)}}{1+\mathbf{W}}-\frac{i\mathfrak{a}}{(1+\mathbf{W})^2}\mathbf{W}^{(n-1)}+\mathbf{err}(H^{\frac{1}{2}}). \tag{8.15}$$

In the first term we can discard the a_1 component of $\mathfrak{a}$ into the error. In the contribution of $a=2\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha]$, only the holomorphic part has an interesting component, precisely when all the derivatives fall on R. Hence we obtain

$$\frac{ia^{(n-1)}}{1+\mathbf{W}} =\frac{\bar{R}_\alpha}{1+\mathbf{W}}R^{(n-1)}+\mathbf{err}(H^{\frac{1}{2}}). \tag{8.16}$$

In the second term in (8.15), we substitute $i\mathbf{W}^{(n-1)}$ with $-\mathcal{T}[\mathbf{W}^{(n-1)}]$ modulo a negligible error. Thus, together with(8.14), (8.15) and (8.16) we arrive at

$$-\partial_\alpha^{(n-1)}\left(\frac{g\mathcal{T}[\mathbf{W}]}{1+\mathbf{W}}-\frac{i\mathfrak{a}}{1+\mathbf{W}}\right)=-\frac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}^{(n-1)}]}{(1+\mathbf{W})^2}-\frac{\bar{R}_\alpha}{1+\mathbf{W}}R^{(n-1)}+\mathbf{err}(H^{\frac{1}{2}}).$$

Combining the above equations we obtain the differentiated system

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{\left((1+\mathbf{W})R^{(n-1)}\right)_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}^{(n-1)} = G \\ R_t^{(n-1)} + bR_\alpha^{(n-1)} - \dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}^{(n-1)}]}{(1+\mathbf{W})^2} = K, \end{cases}$$

where

$$\begin{cases} G = -n\dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} - (n-1)\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}^{(n-1)} + \mathbf{err}(L^2) \\ K = -n\left(\dfrac{R_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right) R^{(n-1)} + \mathbf{err}(H^{\frac{1}{2}}). \end{cases}$$

The following step is to better diagonalize the system, and for this we only need to modify the $R^{(n-1)}$ equation by using the known substitution $\mathbf{R} := (1+\mathbf{W})R^{(n-1)}$ (see [8]). We obtain

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} = G \\ \mathbf{R}_t + b\mathbf{R}_\alpha - \dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}^{(n-1)}]}{1+\mathbf{W}} = K_1, \end{cases}$$

where

$$K_1 = -(n+1)\frac{R_\alpha \mathbf{R}}{1+\bar{\mathbf{W}}} - n\frac{\bar{R}_\alpha \mathbf{R}}{1+\mathbf{W}} + \mathbf{err}(H^{\frac{1}{2}}).$$

To deal with the mildly unbounded terms on the right we proceed in two steps using the same idea as in [8]. First we implement a new holomorphic substitution

$$\tilde{\mathbf{R}} := \mathbf{R} - R_\alpha \mathbf{W}^{(n-2)} + (2n-1)\mathbf{W}_\alpha R^{(n-2)}.$$

With the exception of a couple of terms (see also [8]), the contribution of the added quadratic correction is cubic and lower order, so we obtain

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{\tilde{\mathbf{R}}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} = -n\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{W}^{(n-1)} + \mathbf{err}(L^2) \\ \tilde{\mathbf{R}}_t + b\tilde{\mathbf{R}}_\alpha - \dfrac{(g+\mathfrak{a})\mathcal{T}[\mathbf{W}^{(n-1)}]}{1+\mathbf{W}} = -n\left(\dfrac{R_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right)\tilde{\mathbf{R}} + \mathbf{err}(H^{\frac{1}{2}}). \end{cases}$$

At this point we are in a similar situation as we were in the case $n = 2$. Precisely, we still have unbounded terms on the right, and the goal is to eliminate them. The second step is to use the same procedure as in the case $n = 2$, which is to multiply the equations by $e^{n\phi}$, where $\phi = -2\,\mathrm{Re}\log(1+\mathbf{W})$. After standard estimates, and using also Lemma 8.6, we can write a system for $(w := e^{n\phi}\mathbf{W}^{(n-1)}, r := e^{n\phi}\tilde{\mathbf{R}})$:

$$\begin{cases} w_t + bw_\alpha + \dfrac{r_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}w = \mathbf{err}(L^2) \\ r_t + br_\alpha - \dfrac{(g+\mathfrak{a})\mathcal{T}[w]}{1+\mathbf{W}} = \mathbf{err}(H^{\frac{1}{2}}). \end{cases}$$

As (w, r) are no longer holomorphic, we will need to project them via the projection $\mathbf{P}$. We denote the projected variables $(\mathbf{P}w, \mathbf{P}r)$ by $(\mathfrak{w}, \mathfrak{r})$, and write the equations for them. As we have seen in the case $n=2$, we obtain some additional terms which we can express as commutators. Moreover, these additional terms can be easily bounded using the commutators estimates obtained in the Appendix to obtain the system

$$\begin{cases} \mathfrak{w}_t + \mathbf{P}[b\mathfrak{w}]_\alpha + \mathbf{P}\left[\dfrac{\mathfrak{r}_\alpha}{1+\bar{\mathbf{W}}}\right] + \mathbf{P}\left[\dfrac{R_\alpha}{1+\mathbf{W}}\mathfrak{w}\right] = \mathbf{P}[\mathbf{err}(L^2)] \\ \mathfrak{r}_t + P\left[b\mathfrak{r}_\alpha\right] - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathfrak{w}]}{1+\mathbf{W}}\right] = \mathbf{P}[\mathbf{err}(H^{\frac{1}{2}})]. \end{cases} \tag{8.17}$$

Now, we modify this system one last time in order to be able to compare it with the model system (4.1), and after one rather straightforward estimates we can rewrite it as

$$\begin{cases} \mathfrak{w}_t + \mathbf{P}[b\mathfrak{w}]_\alpha + \mathbf{P}\left[\dfrac{\mathfrak{r}_\alpha}{1+\bar{\mathbf{W}}}\right] - \mathbf{P}\left[\dfrac{R_\alpha}{1+\mathbf{W}}\mathcal{T}^2[\mathfrak{w}]\right] = \mathbf{P}[\mathbf{err}(L^2)] \\ \mathfrak{r}_t + P\left[b\mathfrak{r}_\alpha\right] - \mathbf{P}\left[\dfrac{(g+\mathfrak{a})\mathcal{T}[\mathfrak{w}]}{1+\mathbf{W}}\right] = \mathbf{P}[\mathbf{err}(H^{\frac{1}{2}})]. \end{cases} \tag{8.18}$$

In order to be able to apply the estimates obtained for the model system in (4.1), we need to ensure that the energy of $(\mathfrak{w},\mathfrak{r})$ is equivalent to the one of $(\mathbf{W}^{(n-1)}, R^{(n-1)})$. This is summarized in the following proposition:

Proposition 8.8. *The energy of of* $(\mathfrak{w},\mathfrak{r})$ *above is equivalent to the one of* $(\mathbf{W}^{(n-1)}, R^{(n-1)})$,

$$\|(\mathfrak{w},\mathfrak{r})\|_{\mathcal{H}} \approx_A \|(w,r)\|_{\mathcal{H}} \approx_A \|(\mathbf{W}^{(n-1)}, R^{(n-1)})\|_{\mathcal{H}} \approx_A \ \textit{modulo } A\mathbf{N}_n. \tag{8.19}$$

A similar result can be found in [8] (see Lemma 3.5). The proof of the above proposition is quite similar, and we leave it as an exercise for the reader.

Now the result of Proposition (8.1) follows from the energy estimates for the model system (4.1), namely (4.5) applied to (8.18); to obtain the result we use the energy functional

$$E^{n,(2)}_{high}(\mathbf{W},R) = E^{(2)}_{lin}(\mathfrak{w},\mathfrak{r}).$$

The further goal is to obtain cubic lifespan bounds, which would correspond to proving Proposition (8.2). The key to that is to produce a suitable modified cubic energy. This is done in the next subsection. However, here we will discuss the leading part of the modified cubic energy, which is given by

$$E^{n,(3)}_{high}(\mathbf{W},R) = E^{(3)}_{high}(\mathfrak{w},\mathfrak{r}) := E^{(2)}_{lin}(w,r) - E^{(2)}_{\omega,lin}(w,r), \quad \text{where} \quad \omega = \frac{1}{4}(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}.$$

Remark 8.9. Comparing $E^{n,(3)}_{high}(w,r)$ with the corresponding version appearing in the infinite depth case, one will notice that the are some differences. On one hand, the second component of the above energy is specific to the finite bottom case, and does not appear at all in the infinite bottom problem. On the other hand, the last three terms in the quasilinear cubic energy from [8] are no longer showing up in the above leading energy. Mainly, this is because here we use better bookkeeping of the errors, and those terms are now reclassified as admissible error terms. In other words, here they are incorporated into the lower order component of the quasilinear modified energy we seek to construct.

We claim that we have favourable bounds for the time evolution of this energy. Precisely, we have

Proposition 8.10. *Let* $(\mathfrak{w}, \mathfrak{r})$ *be defined as above. Then*

a) *Assuming that* $A \ll 1$, *we have*

$$E^{(3)}_{high}(\mathfrak{w}, \mathfrak{r}) = E_0(\mathbf{W}^{(n-1)}, R^{(n-1)}) + O(A)\mathbf{N}_n^2. \tag{8.20}$$

b) *The solution of* $(\mathfrak{w}, \mathfrak{r})$ *of* (8.17) *satisfies the following energy estimate*

$$\frac{d}{dt} E^{(3)}_{high}(\mathfrak{w}, \mathfrak{r}) \lesssim_A B\mathbf{N}_n^2, \quad \Lambda^{\geq 4} \frac{d}{dt} E^{(3)}_{high}(\mathfrak{w}, \mathfrak{r}) \lesssim_A AB\mathbf{N}_n^2. \tag{8.21}$$

The proof is a straightforward application of Proposition (4.1). To see that, one needs to verify that the real weight $\omega = \frac{1}{4}(1+\mathcal{T}^2)\,\mathrm{Re}\,\mathbf{W}$ satisfies the required bounds (4.6). But this is true in view of Lemma A.13, as ω is a multiple of the a_1 component of $\mathfrak{a}$.

8.4. **The quasilinear modified energy for** $n \geq 2$**, small data.** In this section we construct an n-th order energy with cubic estimates, $E^{n,(3)}$, which satisfies the bounds in Proposition 8.2. This energy is obtained following the method introduced in [8](*the quailinear modified energy method*), which we now describe by splitting it into several steps:

1. Construct the normal form energy. This has been accomplished in the previous Section 7, but for convenience we outline the process here. Formally, it begins with the construction of a normal form transformation whose aim is to eliminate the quadratic terms in the equation (1.6) for (W, Q). The normal form variables $(\tilde{W}, \tilde{Q})$ are given by

$$\begin{cases} \tilde{W} = W + W^{[2]} = W + B^h[W, W] + \frac{1}{g}C^h[Q, Q] + B^a[W, \bar{W}] + \frac{1}{g}C^a[Q, \bar{Q}] \\ \tilde{Q} = Q + Q^{[2]} = Q + A^h[W, Q] + A^a[W, \bar{Q}] + D^a[Q, \bar{W}], \end{cases}$$

where the bilinear multipliers arising here are defined in Section 6. A full description of these symbols is given later in the same section. What matters is that the normal form variables $(\tilde{W}, \tilde{Q})$ solve an equation of the form

$$\begin{cases} \Lambda^{\leq 2}(\tilde{W}_t + \tilde{Q}_\alpha) = 0 \\ \Lambda^{\leq 2}(\tilde{Q}_t - g\mathcal{T}\tilde{W}) = 0. \end{cases}$$

Following Section 7, its associated cubic normal form energy functional is

$$\begin{aligned} \tilde{E}^n_{NF}(W, Q) &= \Lambda^{\leq 3} E_0(\partial^n \tilde{W}, \partial^n \tilde{Q}) \\ &= E_0(\partial^n W, \partial^n Q) + 2g\langle \partial^n W, \partial^n W^{[2]}\rangle - 2\langle \mathcal{T}^{-1}\partial^{n+1} Q, \partial^n Q^{[2]}\rangle. \end{aligned}$$

This is chosen so that the following relation holds

$$\Lambda^{\leq 3} \frac{d}{dt} \tilde{E}^n_{NF}(W, Q) = 0. \tag{8.22}$$

Here we discard the quartic terms in $E_0(\partial^n \tilde{W}, \partial^n \tilde{Q})$ as on one hand they are both highly unbounded, and on the other hand they do not affect the last relation above. Further, unlike the cubic terms, the quartic terms carry no intrinsic meaning as the normal form transformation is only uniquely determined up to cubic terms.

As we show in the proof Proposition 7.1, the normal form energy $\tilde{E}^n_{NF}(W, Q)$ can be expressed up to quartic terms as a function of diagonal variables $(\mathbf{W}, R)$ in the form

$$E^n_{NF}(\mathbf{W}, R) = E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R) + gB(\mathbf{W}, \mathbf{W}, \mathbf{W}) + A(\mathbf{W}, R, R).$$

with trilinear forms A and B whose symbols we have computed.

We further remark that while the normal form expression has singularities at frequency zero, no such singularities are present in the normal form energy. This is due to symmetrization cancellations akin to some form of null condition. Even better, neither W nor Q can appear undifferentiated in the above cubic terms.

The chief disadvantage of the normal form energy, which due to the fact that our problem is quasilinear, is that the quartic and higher terms in its time derivative d/dtE^n_{NF} are highly unbounded. Thus there is no hope to prove the bound (8.2) for it, neither does (8.1) hold, for that matter.

To better isolate the above difficulty, we have decomposed the normal form energy into two parts,

$$E^n_{NF} = E^n_{NF,high} + E^n_{NF,low},$$

where

$$\begin{aligned} E^n_{NF,high}(\mathbf{W}, R) = &\ E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R) + gB_{high}(\mathbf{W}, \mathbf{W}, \mathbf{W}) + \tilde{A}_{high}(\mathbf{W}, R, R), \\ E^n_{NF,low}(W, Q) = &\ gB_{low}(\mathbf{W}, \mathbf{W}, \mathbf{W}) + A_{low}(\mathbf{W}, R, R). \end{aligned}$$

Here the lower order part is quite complicated algebraically, but has the virtue that it does not cause difficulties neither in (8.1) nor in (8.2). The high frequency part, on the other hand, has the advantage that we can compute it explicitly. Precisely, by Proposition 7.1 we have

$$\begin{aligned} &B_{high}(\mathbf{W}, \mathbf{W}, \mathbf{W}) := \langle \partial^{n-1}\mathbf{W}, \partial^{n-1}\mathbf{W}\rangle_{-4n\operatorname{Re}\mathbf{W} + \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}}, \\ &A_{high}(\mathbf{W}, R, R) := -\langle \partial^{n-1}R, \mathcal{T}^{-1}\partial^{n-1}R_\alpha\rangle_{-4n\operatorname{Re}\mathbf{W} - \frac{1}{2}(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}} \\ &\quad - 2\langle \mathbf{W}\partial^{(n-1)}R, \mathcal{T}^{-1}\partial^{(n-1)}R_\alpha\rangle + 2\langle \partial^{(n-2)}\mathbf{W}R_\alpha, \mathcal{T}^{-1}\partial^{(n-1)}R_\alpha\rangle. \end{aligned} \tag{8.23}$$

This is the part we need to further modify and adapt to the quasilinear structure of our problem.

2. Construct the quasilinear modified energy. Here we construct *the quasilinear modified energy* $E^{n,(3)}$, starting from the normal form energy $E^n_{NF}(\mathbf{W}, R)$. Inspired by the expression for the high frequency part $E^n_{NF,high}(\mathbf{W}, R)$ of the normal form energy, one is naturally led to consider the high frequency quasilinear modified energy $E^{(3)}_{high}(\mathfrak{w}, \mathfrak{r})$ where

$$E^{(3)}_{high}(w, r) := E^{(2)}_{lin}(w, r) - \frac{1}{4}E^{(2)}_{\omega,lin}(w, r). \tag{8.24}$$

Comparing the two, we would like them to agree to cubic order. This is not exactly the case, however the next best thing happens, namely that the cubic part of the difference is lower order:

Lemma 8.11. *The trilinear form* $\Lambda^{\leq 3}(E^n_{NF}(\mathbf{W}, R) - E^{(3)}_{high}(\mathfrak{w}, \mathfrak{r}))$ *is a lower order form in* $(\mathbf{W}, R)$*, where* $n \geq 1$.

The lemma is proved later in this section.

Based on this, we define

$$E^{n,(3)} = E^{n,(3)}_{high} + E^{n,(3)}_{low},$$

where

$$E^{n,(3)}_{low} = E^n_{NF,low} + \Lambda^{\leq 3}(E^n_{NF} - E^{n,(3)}_{high}(w,r)).$$

This guarantees that we have the relation

$$\Lambda^{\leq 3} E^{n,(3)} = \Lambda^{\leq 3} E^n_{NF}. \tag{8.25}$$

3. $E^{n,3}$ **is a good quasilinear cubic energy.** In other words we want to prove that the estimate in Proposition (8.10) holds. In view of (8.22) and (8.25) it follows that for solutions to (1.6) we have

$$\Lambda^{\leq 3}\frac{d}{dt}E^{n,(3)} = 0. \tag{8.26}$$

Thus, we obtain

$$\frac{d}{dt}E^{n,(3)} = \Lambda^{\geq 4}\frac{d}{dt}E^{n,(3)}_{low} + \Lambda^{\geq 4}\frac{d}{dt}E^{n,(3)}_{high}.$$

This relation allows us to split the task of proving bounds for $E^{n,(3)}$ into separate bounds for the high, respectively the low frequency part. Precisely, it remains to establish the following:

Lemma 8.12. *The high frequency part* $E^{n,(3)}_{high}$ *satisfies the bounds*

$$E^{n,(3)}_{low} = \|(\partial^{n-1}\mathbf{W}, \partial^{n-1}R)\|^2_{\mathcal{H}} + O(A)\mathbf{N}_n^2, \tag{8.27}$$

respectively

$$\left|\Lambda^{\geq 4}\left(\frac{d}{dt}E^{n,(3)}_{high}\right)\right| \lesssim_A AB\mathbf{N}_n^2. \tag{8.28}$$

Lemma 8.13. *The (cubic) low frequency part* $E^{n,(3)}_{low}$ *satisfies the bounds*

$$E^{n,(3)}_{low} = O(A)\mathbf{N}_n^2, \tag{8.29}$$

respectively

$$\left|\Lambda^{\geq 4}\left(\frac{d}{dt}E^{n,(3)}_{low}\right)\right| \lesssim_A AB\mathbf{N}_n^2. \tag{8.30}$$

To conclude the proof of Proposition 8.2 it remains to prove the three lemmas above. This is the same argument as in [8], but here it is slightly more complicated, at least at the computational level.

Proof of Lemma 8.11. We first expand the expression $\Lambda^{\leq 3}(E^n_{NF,high}(\mathbf{W},R) - \frac{1}{2}E^{(3)}_{high}(\mathfrak{w},\mathfrak{r}))$ for the case $n \geq 3$ and express the result in terms of $(\mathbf{W},R)$. Up to cubic terms the expansion of $(\mathfrak{w},\mathfrak{r})$ is

$$\begin{cases} \Lambda^{\leq 3}\,\mathfrak{w} = \mathbf{W}^{(n-1)} - 2n\mathbf{P}[\operatorname{Re}\mathbf{W}\cdot\mathbf{W}^{(n-1)}] \\ \Lambda^{\leq 3}\,\mathfrak{r} = R^{(n-1)} - 2n\mathbf{P}[\operatorname{Re}\mathbf{W}\cdot R^{(n-1)}] + \mathbf{W}R^{(n-1)} - R_\alpha\mathbf{W}^{(n-2)} + (2n-1)\mathbf{W}_\alpha R^{(n-2)}. \end{cases}$$

Before substituting the expansion of $(\mathfrak{w},\mathfrak{r})$ into the energy formulas, we observe that the projection $\mathbf{P}$ can be dropped off; moreover the last term in the quadratic expansion of $\mathfrak{r}$ only contributes to lower order terms based on the definition provided in the earlier section.

Thus, we can also omit this term. The explicit quadratic and cubic terms showing up in the expression of $E^{(3)}_{high}(\mathfrak{w},\mathfrak{r}))$ are

$$
\begin{aligned}
\Lambda^{\leq 3}E^{(3)}_{high}(\mathfrak{w},\mathfrak{r})) &= \Lambda^{\leq 3}\left(E^{(2)}_{lin}(\mathfrak{w},\mathfrak{r}) - \frac{1}{2}E^{(2)}_{\omega,lin}(\mathfrak{w},\mathfrak{r})\right)\\
&= \left\langle \mathbf{W}^{(n-1)},\mathbf{W}^{(n-1)}\right\rangle_g + \left\langle LR^{(n-1)},LR^{(n-1)}\right\rangle\\
&\quad + \left\langle \mathbf{W}^{(n-1)},\operatorname{Re}\mathbf{W}\cdot\mathbf{W}^{(n-1)}\right\rangle_{-4ng} + \left\langle \mathbf{W}^{(n-1)},\mathbf{W}^{(n-1)}\right\rangle_{g\omega}\\
&\quad - \left\langle \mathcal{T}^{-1}R^{(n)}, -4n\operatorname{Re}\mathbf{W}\cdot R^{(n-1)} + 2R^{(n-1)}\mathbf{W} - 2R_\alpha\mathbf{W}^{(n-2)}\right\rangle\\
&\quad - \frac{1}{2}\left\langle \mathbf{W}^{(n-1)},\mathbf{W}^{(n-1)}\right\rangle_{g\omega} - \frac{1}{2}\left\langle LR^{(n-1)},LR^{(n-1)}\right\rangle_{\omega},
\end{aligned}
\tag{8.31}
$$

where $\omega = (1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}$.

It remains to compare the result with the expression of $\Lambda^{\leq 3}E^n_{high,NF}(\mathbf{W},R)$, which we recall below:

$$
\Lambda^{\leq 3}E^n_{NF,high}(\mathbf{W},R) = \; E_0(\partial^{n-1}\mathbf{W},\partial^{n-1}R) + gB_{high}(\mathbf{W},\mathbf{W},\mathbf{W}) + A_{high}(\mathbf{W},R,R),
$$

where $B_{high}(\mathbf{W},\mathbf{W},\mathbf{W})$, $A_{high}(\mathbf{W},R,R)$ are given in (7.2).

First we observe that the first line of the expansion in (8.31) is in fact $E_0(\partial^{n-1}\mathbf{W},\partial^{n-1}R)$. The terms on the second line in (8.31) together with the first term on la last line are the terms appearing in gB_{high} modulo a commutator, which yields a lower order term; the commutator is

$$
[\mathcal{T},\operatorname{Re}\mathbf{W}]\operatorname{Re}\mathbf{W}^{(n-1)}.
$$

We return to the remaining terms in (8.31) and observe that the first term in the expansion of the inner product on the third line together with the last term on la last line match (after integrating by parts) the first term in the expansion of A_{high}, (7.2), up to the commutators

$$
[L,\omega]\operatorname{Im}(LR^{(n-1)}), \quad [L,\omega]\,\mathcal{T}\operatorname{Re}(LR^{(n-1)}),
$$

which are again lower order terms.

Lastly, the last two terms, $-\left\langle \mathcal{T}^{-1}R^{(n)}, 2R^{(n-1)}\mathbf{W}\right\rangle$ and $\left\langle \mathcal{T}^{-1}R^{(n)}, 2R_\alpha\mathbf{W}^{(n-2)}\right\rangle$, are a perfect match to the remaining terms in A_{high}.

For the case $n=2$ the computation is similar but simpler. The last three terms in $\Lambda^{\leq 3}$ no longer appear, whereas in the expression for $A_{high}(\mathbf{W},R,R)$ in (7.2) the last two terms also cancel.

□

Proof of Lemma 8.12. This is a direct consequence of Lemma (8.10).

□

Proof of Lemma 8.13. We recall that $E^{n,3}_{low}$ is a trilinear expression of the form

$$
E^{n,3}_{low} = gB_{low}(\mathbf{W},\mathbf{W},\mathbf{W}) + A_{low}(\mathbf{W},R,R),
$$

where B_{low} and A_{low} are translation invariant trilinear forms. To begin with, we note that the exact form of the terms in $E^{n,3}_{low}$ is irrelevant here. All that matters is their symbol class, which we now recall. In the case of B_{low}, the symmetric symbol $B_{low}(\xi,\eta,\zeta)$ satisfies

$$
B_{low} \in ES(d\rho^{2n-3}),
$$

while in the case of A, the symbol $A_{low}(\xi,\eta,\zeta)$ is only symmetric in the last two variables and satisfies

$$A_{low} \in ES(\rho^{2n-2}) + ES(dd_1\rho^{2n-3}).$$

Here d, d_1 measure the distance to the axes as follows:

$$d = 1 + \min\{|\xi|,|\eta|,|\zeta|\}, \qquad d_1 = 1 + \min\{|\eta|,|\zeta|\}.$$

We recall that we can eliminate the exponentials in the symbols at the expense of replacing some of the arguments $(\mathbf{W}, R)$ by their complex conjugates.

We begin with the estimate (8.29). By applying a standard trilinear Littlewood-Paley decomposition combined with a standard separation of variables argument we can thus write B_{low} as a sum of a rapidly convergent series

$$B_{low}(\mathbf{W},\mathbf{W},\mathbf{W}) = \sum_{1\le j\le k} 2^j 2^{(2n-3)k} \sum_m \int \chi^{m,1}_{j,k}(D)\mathbf{W}_j \cdot \chi^{m,2}_{j,k}(D)\mathbf{W}_k \cdot \chi^{m,3}_{j,k}(D)\mathbf{W}_k \, d\alpha.$$

Here complex conjugates are also allowed, and the symbols $\chi^{m,i}_{j,k}(\xi)$ have the following properties:

(i) They are smooth on the respective dyadic scales 2^j, respectively 2^k uniformly with respect to j,k.

(ii) They are rapidly decaying in m, also uniformly with respect to j,k.

In particular the multipliers $\chi^{m,i}_{j,k}(D)$ are uniformly bounded in all L^p spaces and rapidly decaying with respect to m. Hence, we immediately obtain the following bound for $\tilde{A}_{low}$:

$$\begin{aligned}
|B_{low}(\mathbf{W},\mathbf{W},\mathbf{W})| &\lesssim \sum_{1\le j\le k} 2^j 2^{(2n-3)k} \|\mathbf{W}_j\|_{L^\infty} \|\mathbf{W}_k\|^2_{L^2} \\
&\lesssim \sup_j \|\mathbf{W}_j\|_{L^\infty} \sum_k 2^{(2n-2)k} \|\mathbf{W}_k\|^2_{L^2} \\
&\le g^{-1} A\mathbf{N}_n^2.
\end{aligned}$$

The computation is only slightly more involved for A_{low}. We only discuss the $ES(dd_1\rho^{2n-3})$ part, as the analysis for the lower homogeneity part $ES(\rho^{2n-2})$ is similar but simpler. We need to consider two cases depending on whether the $\mathbf{W}$ factor or an R factor is low frequency. We obtain

$$\begin{aligned}
|A_{low}(\mathbf{W},R,R)| &\lesssim \sum_{1\le j\le k} 2^j 2^{(2n-2)k} \|\mathbf{W}_j\|_{L^\infty} \|R_k\|^2_{L^2} + 2^{2j} 2^{(2n-3)k} \|R_j\|_{L^\infty} \|\mathbf{W}_k\|_{L^2} \|R_k\|_{L^2} \\
&\lesssim \sup_j \|\mathbf{W}_j\|_{L^\infty} \sum_k 2^{(2n-1)k} \|R_k\|^2_{L^2} + \sup_j 2^{\frac{j}{2}} \|R_j\|_{L^\infty} \sum_k 2^{(2n-\frac{3}{2})k} \|\mathbf{W}_k\|_{L^2} \|R_k\|_{L^2} \\
&\le A\mathbf{N}_n^2.
\end{aligned}$$

Now consider the bound (8.30), where we write

$$\Lambda^{\ge 4} \frac{d}{dt} B_{low}(\mathbf{W},\mathbf{W},\mathbf{W}) = 3B_{low}(\Lambda^{\ge 2}\partial_t \mathbf{W},\mathbf{W},\mathbf{W}),$$

respectively

$$\Lambda^{\ge 4} \frac{d}{dt} A_{low}(\mathbf{W},R,R) = A_{low}(\Lambda^{\ge 2}\partial_t \mathbf{W},R,R) + 2A_{low}(\partial_t \mathbf{W},\Lambda^{\ge 2}R,R).$$

For the time derivatives of $\mathbf{W}$ and R we separate the leading order transport term, precisely its paraproduct part, writing

$$\Lambda^{\geq 2}\partial_t \mathbf{W} = (\partial_t + T_b\partial_\alpha)\mathbf{W} - T_b\partial_\alpha \mathbf{W},$$

and similarly for R. Here by a slight abuse of notation we include the contribution of the low frequencies in b in T_b. This is because we do not have good control over the low frequencies of b, so these cannot be bounded perturbatively, and instead must be treated only in a commutator type fashion.

The first term has better regularity, and its contribution is treated perturbatively. Precisely, a computation similar to the one above applies provided we can establish the pointwise bounds

$$\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)\mathbf{W}\|_{B^{0,\infty}_\infty} + g^{-\frac{1}{2}}\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)R\|_{B^{\frac{1}{2},\infty}_\infty} \lesssim_A AB, \tag{8.32}$$

respectively the L^2 bounds

$$\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)\mathbf{W}\|_{H^{n-\frac{3}{2}}} + g^{-\frac{1}{2}}\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)R\|_{H^{n-1}} \lesssim_A A\mathbf{N}_n. \tag{8.33}$$

Both of these are proved in Lemma A.16 in the Appendix.

For the contribution of the transport term, on the other hand, we need to capture some cancellation. We discuss the case of the form B_{low}, as A_{low} is similar. In the product case, this cancellation is a simple integration by parts, based on the formula

$$\int b\partial_\alpha \mathbf{W}_1\mathbf{W}_2\mathbf{W}_3 + \mathbf{W}_1 b\partial_\alpha\mathbf{W}_2\mathbf{W}_3 + \mathbf{W}_1\mathbf{W}_2 b\partial_\alpha \mathbf{W}_3 \, d\alpha = -\int b_\alpha \mathbf{W}_1\mathbf{W}_2\mathbf{W}_3 \, d\alpha,$$

where the derivative is moved onto b. In our case, however, we need to contend instead with factors which at frequency 2^j have the form $\chi_j(D)(b_{<j}\partial_\alpha \mathbf{W}_j)$.

As a preliminary observation, we remark that we can commute out the coefficient $b_{<j}$, by writing

$$\chi_j(D)(b_{<j}\partial_\alpha \mathbf{W}_j) = b_{<j}\chi_j(D)\partial_\alpha \mathbf{W}_j + [\chi_j(D), b_{<j}]\partial_\alpha \mathbf{W}_j.$$

Here the commutator term can be expressed in the form

$$[\chi_j(D), b_{<j}]\partial_\alpha \mathbf{W}_j = L(\nabla b_{<j}, \mathbf{W}_j),$$

where L stands for a translation invariant bilinear form with integrable kernel. Then one can directly use the bounds in Lemma A.12 for b to show that this term satisfy the same bounds as in (8.32), (8.33), and thus can be treated perturbatively.

Once we have discarded the commutator term, we can include $\chi_j(D)$ into $\mathbf{W}_j$ for brevity, and then we are left with having to estimate an expression of the form

$$I = \int b_{<j}\partial_\alpha \mathbf{W}_j \cdot \mathbf{W}_k \cdot \mathbf{W}_k + \mathbf{W}_j \cdot b_{<k}\partial_\alpha \mathbf{W}_k \cdot \mathbf{W}_k + \mathbf{W}_j \cdot \mathbf{W}_k b_{<k}\partial_\alpha \mathbf{W}_k \, d\alpha.$$

Separating the expression $b_{<k}$ in all factors we can integrate by parts and obtain

$$I = -\int \partial_\alpha b_{<k}\mathbf{W}_j \cdot \mathbf{W}_k \cdot \mathbf{W}_k \, d\alpha - \int b_{[j,k]}\partial_\alpha \mathbf{W}_j \cdot \mathbf{W}_k \cdot \mathbf{W}_k \, d\alpha.$$

Now in the first integral we group the product $\partial_\alpha b_{<k}\mathbf{W}_k$, which again satisfies the same bounds as in (8.32), (8.33). In the second integral the derivative yields a 2^j factor, and now the expression $2^j b_{[j,k]}$ is even better than $\partial_\alpha b_{<k}$. □

8.5. **The quasilinear modified energy for $n = 1$, small data.** In this section we construct a first order energy with cubic estimates, $E^{1,(3)}$, which satisfies the bounds in Proposition 8.2. This energy is obtained following the same procedure as in the case $n \geq 2$ presented before, but with some minor computational differences, which we now describe.

One main source of differences is the expression for A_{high} which is slightly different here. Also in this case it is no longer meaningful to do the exponential conjugation. Because of this, it is now convenient to set up the quasilinear correction to the normal form energy in a more direct fashion,

$$E^{1,(3)}_{high} = E^{(3)}_{high}(\mathbf{W}, R) + E^{(3),a}(\mathbf{W}, R),$$

where the extra component

$$E^{(3),a}(\mathbf{W}, R) = -2\langle \mathbf{W}, \mathbf{W}^2\rangle + 2\langle R, \mathbf{W}\mathcal{T}^{-1}R_\alpha\rangle$$

mirrors the similar correction in the infinite bottom case [8].

An advantage of doing this is that the remaining lower order cubic part

$$E^{1,(3)}_{low} = B_{low}(\mathbf{W}, \mathbf{W}, \mathbf{W}) + A_{low}(R, R, \mathbf{W})$$

contains only terms whose symbol is not only lower order on the diagonals but also away from them, namely their symbols satisfy

$$B_{low} \in S(\rho^{-1}), \qquad A_{low} \in S(1). \tag{8.34}$$

This is due to the similar gain in Proposition 7.1.

With these definitions we remark that Lemma 8.11 is still valid. For that we need to match the terms in $\Lambda^{\leq 3}(E^n_{NF,high}(\mathbf{W}, R))$ to the terms in $\Lambda^{\leq 3}(\frac{1}{2}E^{1,(3)}_{high}(\mathbf{W}, R))$. The computations are similar to the ones we did for the case $n \geq 2$ but simpler.

Further, the statements of Lemmas 8.12, 8.13 remain unchanged. It remains to prove Lemmas 8.12,8.13 in this context.

Proof of Lemma 8.12, $n = 1$. The bound (8.27) is straightforward. The $E^{(3)}_{high}$ part of (8.28) is also exactly as before in view of Lemma 8.4. It remains to prove the extra correction $E^{(3),a}(\mathbf{W}, R)$ also satisfies (8.28). For convenience we state this in a separate lemma:

Lemma 8.14. *The cubic correction $E^{(3),a}$ satisfies the bounds*

$$E^{(3),a} \lesssim_A A\mathbf{N}_1^2, \tag{8.35}$$

respectively

$$\left|\Lambda^{\geq 4}\left(\frac{d}{dt}E^{(3),a}\right)\right| \lesssim_A AB\mathbf{N}_1^2. \tag{8.36}$$

Proof. The first bound is straightforward, but the second does require some computations. We consider both correction terms

$$I_1 = \left\langle \mathbf{W}, \mathbf{W}^2\right\rangle, \quad I_2 = \left\langle R, \mathbf{W}\mathcal{T}^{-1}R_\alpha\right\rangle,$$

and discuss each of them separately.

To estimate their derivatives it is easiest to use the unprojected form (1.8) of the equations for $\mathbf{W}$ and R, which for our purposes here we write in the form

$$
\begin{cases} (\partial_t + b\partial_\alpha)\mathbf{W} = -b_\alpha(1+\mathbf{W}) + \bar{R}_\alpha := G \\ (\partial_t + b\partial_\alpha)R = i\dfrac{g\mathbf{W} - \mathfrak{a}}{1+\mathbf{W}} := K. \end{cases} \tag{8.37}
$$

For G and K we only need their quadratic parts and higher,

$$
G^{2+} = -b_\alpha \mathbf{W} + \mathbf{P}[R\bar{Y}]_\alpha, \quad K^{2+} = -\frac{(ig\mathbf{W} - a_1)\mathbf{W} + a}{1+\mathbf{W}}.
$$

Then we have

$$
\Lambda^{\geq 4}\left(\frac{d}{dt}I_1\right) = -\left\langle b\mathbf{W}_\alpha, \mathbf{W}^2\right\rangle - \langle \mathbf{W}_\alpha, 2b\mathbf{W}\mathbf{W}_\alpha\rangle + \left\langle G^{2+}, \mathbf{W}^2\right\rangle + 2\left\langle \mathbf{W}, \mathbf{W}G^{2+}\right\rangle.
$$

Distributing derivatives and using Corollary A.8, we separate the terms with undifferentiated b as

$$
\begin{aligned}
-\left\langle b\mathbf{W}_\alpha, \mathbf{W}^2\right\rangle - \left\langle \mathbf{W}_\alpha, b\partial_\alpha(\mathbf{W}^2)\right\rangle &= -\left\langle b\mathbf{W}_\alpha, \mathbf{W}^2\right\rangle + + \left\langle \mathcal{T}^{-1}\partial_\alpha[b\mathcal{T}\mathbf{W}], \mathbf{W}^2\right\rangle \\
&= \left\langle (-b\partial_\alpha + \mathcal{T}^{-1}\partial_\alpha b\mathcal{T})\mathbf{W}, \mathbf{W}^2\right\rangle.
\end{aligned}
$$

Note that we can express this as the sum of two terms, as shown below

$$
\left\langle \mathcal{T}^{-1}\partial_\alpha[b\,,\,\mathcal{T}]\mathbf{W}, \mathbf{W}^2\right\rangle + \left\langle b_\alpha\mathbf{W}, \mathbf{W}^2\right\rangle,
$$

where both can be easily controlled by $AB\mathbf{N}_1$ using Lemma A.9 followed by Lemma A.18. The contribution of G^{2+} is harmless since all the terms in G^{2+} are bounded in L^2,

$$
\|G^{2+}\|_{L^2} \lesssim_A B\mathbf{N}_1.
$$

We now return to the last correction term, I_2:

$$
\Lambda^{\geq 4}\left(\frac{d}{dt}I_2\right) = \langle R_t, \mathbf{W}\mathcal{T}^{-1}R_\alpha\rangle + \langle R_t, \mathbf{W}\mathcal{T}^{-1}R_\alpha\rangle + \langle R_t, \mathbf{W}\mathcal{T}^{-1}R_\alpha\rangle.
$$

The argument for this expression is slightly more involved. We proceed as in the proof of Lemma 8.13, but with some extra care. We begin with a Littlewood-Paley decomposition

$$
\langle R, \mathbf{W}\mathcal{T}R_\alpha\rangle = \sum_{k,k_1,k_2\geq 0} \langle P_{k_1}R, P_{k_2}\mathbf{W}\mathcal{T}^{-1}P_{k_3}R_\alpha\rangle,
$$

and similarly for the time derivative. For the above summand to be nonzero, we need the two highest frequencies to be comparable. We first distinguish two easier cases:

(i) If $\min\{k, k_1, k_2\} \lesssim 1$, and the time derivative applies to the low frequency. Then the time differentiated factor is bounded in $L^2 \cap L^\infty$, and the two remaining factors are estimated in L^2 or L^∞ as needed.

(ii) If $k < k_1 = k_2$, then we take advantage of the fact that our factors are holomorphic, and thus have exponential decay at positive frequencies. Thus we obtain an e^{-Nk_1} gain which is more than enough for all our estimates.

This leaves us with two principal cases, namely the sums:

$$
J_1 = \partial_t \sum_{k>4} \int \bar{R}_k \mathbf{W}_k R_{\leq k,\alpha}\, d\alpha, \qquad J_2 = \partial_t \sum_{k>4} \int \bar{R}_k \mathbf{W}_{\leq k} R_{k,\alpha}\, d\alpha.
$$

To estimate their time derivatives we use again the decomposition

$$\partial_t \mathbf{W} = (\partial_t + T_b\partial_\alpha)\mathbf{W} - T_b\partial_\alpha \mathbf{W}, \qquad \partial_t R = (\partial_t + T_b\partial_\alpha)R - T_b\partial_\alpha R.$$

For the first term in each decomposition we have the estimates in Lemma A.16. Using them, the bounds for the corresponding contributions to J_1 and J_2 are somewhat tedious but routine. It remains to consider the T_b contributions, which are

$$J_1^b = \sum_{k>4} \int b_{<k}\bar{R}_{k,\alpha}\mathbf{W}_k R_{\leq k,\alpha} + \bar{R}_k b_{<k}\mathbf{W}_{k,\alpha}R_{\leq k,\alpha} + \bar{R}_k\mathbf{W}_k\partial_\alpha(T_b\bar{R}_{\leq k,\alpha})\, d\alpha,$$

$$J_2^b = \sum_{k>4} \int b_{<k}\bar{R}_{k,\alpha}\mathbf{W}_{\leq k}R_{k,\alpha} + \bar{R}_k T_b W_{\leq k,\alpha}R_{k,\alpha} + \bar{R}_k\mathbf{W}_{\leq k}\partial_\alpha(b_{<k}R_{k,\alpha})\, d\alpha.$$

Integrating by parts we rewrite these integrals as

$$J_1^b = \sum_{k>4} \int \bar{R}_k\mathbf{W}_k\partial_\alpha((T_b - b_{<k})\bar{R}_{\leq k,\alpha})\, d\alpha,$$

$$J_2^b = \sum_{k>4} \int \bar{R}_k(T_b - b_{<k})\mathbf{W}_{\leq k,\alpha}R_{k,\alpha}\, d\alpha.$$

Here the expressions $(T_b - b_{<k})\bar{R}_{\leq k,\alpha}$, respectively $(T_b - b_{<k})W_{\leq k,\alpha}$ are of the same type as the expressions considered in Lemma A.16 as part of $(\partial_t + T_b\partial_\alpha)R$, respectively $(\partial_t + T_b\partial_\alpha)\mathbf{W}$. Thus they also satisfy the bounds in Lemma A.16, and the desired conclusion follows.

□

Proof of Lemma 8.13, $n = 1$. Because of the better bounds for the lower order terms in (8.34), this proof is straightforward and is omitted. □

□

9. Proof of the main results

Given the estimates obtained in the previous sections both for the main evolution 1.6 and for the linearized equation, the proof of the main results in Theorem 1 and Theorem 2 are fairly routine. Thus, in this section we provide an outline of the proofs only. For a more in-depth exposition of arguments of this type we refer the reader to the earlier article [8] devoted to the infinite depth problem. We will however emphasize the differences between the finite and infinite depth case.

Proof of Theorem 1, outline. Due to scaling considerations we can work with $h = 1$ and $g \lesssim 1$. The main steps in the proof are as follows:

1. Existence of regular solutions. Here we start with initial data $(W, Q)(0) \in L^2 \times H^{\frac{1}{2}}$ and $(\mathbf{W}, R)(0) \in H^n \times H^{n+\frac{1}{2}}$ with $n \geq 2$, which has extra regularity both at low frequency and at high frequency. For such data, local in time solutions are constructed as weak limits of solutions for a frequency localized system. In doing this it is convenient to work with the differentiated equation (1.8), in order to have the equations in diagonalized form. For this the argument in [8] applies almost identically.

We note one advantage of working with the holomorphic coordinates, namely that the free water surface is not required to be a graph. If it were not for this, we could simply use the local well-posedness result in [1] or [12].

2. Uniqueness of regular solutions. Here we consider two solutions (W_1, Q_2) and (W_2, Q_2) with regularity $(W_j, Q_j) \in C([0,T]; \mathcal{H})$ and $(\mathbf{W}_j, R_j) \in C([0,T]; \mathcal{H}^n)$ with $n \geq 2$, and show that if their initial data agree then the two solutions must be equal. Note that while more regularity is assumed at high frequency, that is no longer the case at low frequency.

For the proof one subtracts the two sets of equations, estimating the difference of the two solutions for the differentiated equation (1.8). The key point is that up to perturbative terms, the difference $(w, r) = (\mathbf{W}_1 - \mathbf{W}_2, R_1 - R_2)$ solves a linear system similar to our model evolution for the linearized equation (5.4). Then one can conclude the proof of uniqueness in a standard manner using Gronwall's inequality.

3. Lifespan bounds in terms of the $\mathcal{H}^1$ size of the data. The lifespan of solutions constructed above depends both on the $\mathcal{H}^n$ size of the data $(\mathbf{W}, R)(0)$ and on g. Here we show that we can in effect obtain lifespan bounds which depend only on the $\mathcal{H}^1$ size of the data and which are independent of g. To be precise, we take initial data which satisfy the bounds

$$\|(W,Q)(0)\|_{\mathcal{H}} \leq g\mathcal{M}_0, \qquad \|(\mathbf{W}, R)(0)\|_{\mathcal{H}} \leq g\mathcal{M}_0, \qquad \|(\mathbf{W}_\alpha, R_\alpha)(0)\|_{\mathcal{H}} \leq \mathcal{M}_0, \tag{9.1}$$

as well as the pointwise bounds

$$\|Y(0)\|_{L^\infty} \leq \mathcal{K}_0, \qquad \operatorname{Im} W + 1 \geq c_0 > 0. \tag{9.2}$$

Then we will show that there exists $T = T(\mathcal{M}_0, Y_0, c_0)$ so that the solutions exist on $[-T, T]$ with similar bounds.

For the proof we use a bootstrap argument, assuming that the following bounds hold in $[0, T]$:

$$\|(W,Q)\|_{\mathcal{H}} \leq g\mathcal{M}, \qquad \|(\mathbf{W}, R)\|_{\mathcal{H}} \leq g\mathcal{M}, \qquad \|(\mathbf{W}_\alpha, R_\alpha)\|_{\mathcal{H}} \leq \mathcal{M}, \tag{9.3}$$

as well as the pointwise bounds

$$\|Y\|_{L^\infty} \leq \mathcal{K}, \qquad W + 1 \geq c > 0. \tag{9.4}$$

Then we need to show that for a suitable choice of $\mathcal{M}$, $\mathcal{K}$, c depending on $\mathcal{M}_0$, $\mathcal{K}_0$ and c_0 but not on g we can improve all these bounds. Through the following computations we denote by C_0 various constants which only depend on $\mathcal{M}_0$ and K_0.

We begin by observing that by Sobolev embedding our control parameters satisfy

$$A, B \leq C(\mathcal{M}, \mathcal{K}), \qquad a \geq cg.$$

Hence by the energy estimates for the differentiated equation in Proposition 8.1 we obtain

$$\|(\mathbf{W}, R)(0)\|_{\mathcal{H}} \leq gc^{-1}C_0(1 + tC(\mathcal{M}, \mathcal{K})), \qquad \|(\mathbf{W}_\alpha, R_\alpha)\|_{\mathcal{H}} \leq c^{-1}C_0 C(\mathcal{K})(1 + tC(\mathcal{M}, \mathcal{K})),$$

where the $\mathcal{K}$ dependence in the second bound is caused by the need to invert a $1 + \mathbf{W}$ factor, see [8] for a full argument.

To bound (W, Q) in time we use the equations directly to obtain

$$\|(W,Q)\|_{\mathcal{H}} \leq g(C_0 + tC(\mathcal{M}, \mathcal{K})).$$

To bound Y in L^∞ we reuse the argument in [8], which yields

$$\|Y(t)\|_{L^\infty}^2 \leq C_0(1 + tC(\mathcal{M}, \mathcal{K})).$$

Finally, to bound $\operatorname{Im} W$ from below we use directly the W equation to obtain

$$(\partial_t + \operatorname{Re} F \partial_\alpha) \operatorname{Im} W = (1 + \operatorname{Re} W_\alpha) \operatorname{Im} \left(\frac{R}{1 + \bar{\mathbf{W}}} \right),$$

which yields

$$\inf_{\alpha \in \mathbf{R}} 1 + \operatorname{Im} W(t, \alpha) \geq c_0 - tC(\mathcal{M}, \mathcal{K}).$$

Summarizing, in order to close the bootstrap we need to have the bounds

$$\mathcal{M} > C_0 c^{-1} C(\mathcal{K})(1 + tC(\mathcal{M}, \mathcal{K})), \qquad \mathcal{K}^2 > C_0 c^{-1}(1 + tC(\mathcal{M}, \mathcal{K})), \qquad c < c_0 - tC(\mathcal{M}, \mathcal{K}).$$

This is achieved by first choosing $c = c_0/2$, then $\mathcal{K}$ large enough $\mathcal{K}^2 = 2C_0 c^{-1}$ next $\mathcal{M}$ large enough $\mathcal{M} = 2C_0 C(\mathcal{K}) c^{-1}$, and finally a small enough $T < T(\mathcal{M}, \mathcal{K}, c)$.

4. $\mathcal{H}^n$ solutions for $n \geq 2$. Here we relax our low frequency regularity assumption for the data to $(W, Q)(0) \in \mathcal{H}$, while keeping the high frequency regularity $(\mathbf{W}, R)(0) \in \mathcal{H}^n$, $n \geq 2$, and prove that solutions still exist. By Step 2, such solutions are also unique. To obtain such solutions we consider a sequence of data $(W_n, Q_n)(0)$ with regularity $(W_n, Q_n)(0) \in L^2 \times H^{\frac{1}{2}}$ so that

$$(W_n, Q_n)(0) \to (W, Q)(0) \text{ in } \mathcal{H}, \qquad (\mathbf{W}_n, R_n)(0) \to (\mathbf{W}, R)(0) \text{ in } \mathcal{H}^2.$$

This is easily achieved by cutting off the low frequencies

$$(W_n, Q_n)(0) = P_{>-n}(W, Q)(0).$$

For n large enough this family of data is uniformly bounded in the sense of (9.1), so by the previous step they generate solutions (W_n, Q_n) with uniform bounds life-span. But then the estimates on the linearized equation in Section 5 show that the sequence (W_n, Q_n) converges to some (W, Q) uniformly in the $\mathfrak{H}$ topology. Due to the uniform bounds on (W_n, Q_n) this linearly yields $(W_n, Q_n) \to (W, Q)$ in $\mathfrak{H}^{2-}$. Thus R is well defined and we also have $(\mathbf{W}_n, R_n) \to (\mathbf{W}, R)$ in $\mathcal{H}^1$. Using now the uniform bounds on $(\mathbf{W}_n, R_n)$ we obtain weak convergence $(\mathbf{W}_n, R_n) \to (\mathbf{W}, R)$ in $\mathcal{H}^2$, and strong convergence in all weaker topologies. Thus we have obtained the desired solutions (W, Q).

5. Rough solutions. Here we show that the solution operator constructed above for data $(W, Q)(0) \in \mathcal{H}$ with $(\mathbf{W}, R)(0) \in \mathcal{H}^2$ extends continuously to data with only $(W, Q)(0) \in \mathcal{H}$ and $(\mathbf{W}, R)(0) \in \mathcal{H}^1$.

Indeed, consider some data which only satisfies the latter requirement. Then we regularize the data $(W, Q)(0)$ to $(W_n, Q_n)(0) = P_{<n}(W, Q)(0)$. This linearly guarantees convergence

$$|(W_n, Q_n)(0) - (W, Q)(0)| \to 0 \qquad \text{in } H^2,$$

which also shows that $\mathbf{W}_n(0) \to \mathbf{W}(0)$ uniformly, and also

$$(\mathbf{W}_n, R_n) \to (\mathbf{W}, R) \qquad \text{in } H^1.$$

Now we turn our attention to the key point, which is to improve this last convergence to $\mathcal{H}^1$. We will in effect do slightly better than that, and for this we need to work with slowly varying frequency envelopes. Precisely, we have the following:

Lemma 9.1. *Let* $\{c_n\}_{n\geq 0}$ *be a slowly varying frequency envelope for* $(\mathbf{W}, R)(0)$ *in* $\mathcal{H}^1$*. Then we have the estimate*

$$2^{\frac{3}{2}n}\|(\mathbf{W}_n, R_n)(0) - P_{<n}(\mathbf{W}, R)(0)\|_{\mathcal{H}} + 2^{-n}\|(\mathbf{W}_n, R_n)(0)\|_{\mathcal{H}^2} \lesssim_A c_n.$$

We note that this lemma not only shows that $(\mathbf{W}_n, R_n)(0) \to (\mathbf{W}, R)(0)$ in $\mathcal{H}^1$, but also that they share the common c_n frequency envelope.

Proof. We drop the "(0)" notation for this proof. Only the R part of the bounds is nontrivial. Expressing all in terms of R and W, for the first expression above we need to bound in $H^{\frac{1}{2}}$ the difference

$$\frac{1}{1+P_{<n}\mathbf{W}}P_{<n}[R(1+\mathbf{W})] - P_{<n}R = \frac{1}{1+P_{<n}\mathbf{W}}(P_{<n}[R\mathbf{W}] - P_{<n}RP_{<n}\mathbf{W}).$$

We will bound the last difference in L^2 using the usual paradifferential decomposition. We can express it as

$$P_{<N}[R\mathbf{W}] - P_{<n}RP_{<n}\mathbf{W} = \Pi(P_nR, P_n\mathbf{W}) + [P_{<n}, R_{<n-4}]P_n\mathbf{W} + [P_{<n}, \mathbf{W}_{<n-4}]P_nR.$$

Estimating the high frequency factors in L^2 and the low frequency factors in L^∞ we obtain

$$\|P_{<n}[R\mathbf{W}] - P_{<n}RP_{<n}\mathbf{W}\|_{L^2} \lesssim A2^{-\frac{3}{2}n}c_n.$$

The R_n bound in the second expression above is easier and is left for the reader. □

Once we have uniform bounds for $(\mathbf{W}_n, R_n)(0)$ in $\mathcal{H}^1$, by the previous step it follows that the corresponding solutions $(\mathbf{W}_n, R_n)$ have a uniform life-span, with uniform bounds. Our next goal is to show that the frequency envelope bounds are inherited also by the solutions.

Lemma 9.2. *Let* (W_n, Q_n) *be the solutions associated to the initial data as above. Then we have the estimates*

$$\|(W_{n+1}, Q_{n+1}) - (W_n, Q_n)\|_{\mathfrak{H}} \lesssim_{A,B} 2^{-2n} \tag{9.5}$$

$$\|(\mathbf{W}_n, R_n)\|_{\mathcal{H}^2} \lesssim_{A,B} 2^n c_n, \tag{9.6}$$

respectively

$$\|(\mathbf{W}_{n+1}, R_{n+1}) - (\mathbf{W}_n, R_n)\|_{\mathcal{H}} \lesssim_{A,B} 2^{-n}c_n. \tag{9.7}$$

Proof. Given the $\mathcal{H}^2$ bound for the initial data $(W_n, Q_n)(0)$ in the previous lemma, the bound (9.6) is a direct consequence of our higher order energy bounds.

For (9.5) we will use instead the linearized equation. Precisely, we now interpret n as a continuous parameter. Then the functions

$$(w, q) = \frac{d}{dn}(W_n, Q_n)$$

solve the linearized equation, and have initial data $(w, q)(0) = P_n(W, Q)(0)$ localized at frequency 2^n. Considering now the diagonalized variables $(w, r) = (w, q + R\mathcal{T}^2 w)$, an argument similar to the proof of Lemma 9.1 shows that their data satisfies

$$\|(w, r)(0)\|_{\mathcal{H}} \lesssim 2^{-2n}c_n. \tag{9.8}$$

Applying the bounds for the linearized equation Theorem 5.1 we extend the estimate (9.8) along the flow,

$$\|(w, r)\|_{\mathcal{H}} \lesssim 2^{-2n}c_n \tag{9.9}$$

For the estimate (9.7) we first bound the high frequency part using the $\mathcal{H}^2$ bound (9.6). Precisely,

$$\|P_{>n+1}(\mathbf{W}_{n+1}, R_{n+1})\|_{\mathcal{H}} + \|P_{>n}(\mathbf{W}_n, R_n)\|_{\mathcal{H}} \lesssim 2^{-n}c_n,$$

where the constant is independent of n.

To bound the low frequency part we define

$$(w_1, r_1) = \frac{d}{dn}(\mathbf{W}_n, R_n).$$

We observe that in terms of (w, r) we have

$$(w_1, r_1) = (w_\alpha, r_\alpha + R_n(1 + \mathcal{T}^2)w_\alpha + (R_n)_\alpha \mathcal{T}^2 w).$$

A simple application of the usual Littlewood-Paley trichotomy then yields the estimate

$$\|P_{<n}(w_1, r_1)\|_{\mathcal{H}} \lesssim_A 2^n \|(w, r)\|_{\mathcal{H}}. \tag{9.10}$$

From the estimate (9.9) we gain an $\mathfrak{H}$ bound for (w, q), which integrated between $[n, n+1]$ yields (9.5). On the other hand, we have

$$\frac{d}{dn} P_{<n}(\mathbf{W}_n, R_n) = P_{<n}(w_1, r_1) + P_n(\mathbf{W}_n, R_n),$$

where the second term may again be bounded using (9.6). Integrating (9.10) we obtain (9.7). ☐

The bounds in the last lemma insure not only that the sequence (W_n, Q_n) converges strongly to a solution (W, Q) in the sense that

$$(W_n, Q_n) \to (W, Q) \qquad \text{uniformly in } H^2$$
$$(\mathbf{W}_n, R_n) \to (\mathbf{W}, R) \qquad \text{uniformly in } \mathcal{H}^1$$

but also that $(\mathbf{W}, R)$ inherits the same frequency envelope $\{c_n\}$ in $\mathcal{H}^1$.

Once we have constructed the rough solutions (W, Q) as the unique limit of the regularized problems, the frequency envelope bounds easily lead to continuous dependence with respect to data. This is a standard argument; for which we refer the reader to [8].

☐

Proof of Theorem 2, outline. Using the spatial scaling , it suffices to assume that $h = 1$. Given the initial data $(W, Q)(0)$ for (1.6) satisfying

$$g^{-1}\|(W, Q)(0)\|_{\mathcal{H}} + g^{-1}\|(\mathbf{W}, R)(0)\|_{\mathcal{H}} + \|(\mathbf{W}_\alpha, R_\alpha)(0)\|_{\mathcal{H}} \le \epsilon,$$

we consider the solutions on a time interval $[0, T]$ and seek to prove the estimate

$$g^{-1}\|(W, Q)(t)\|_{\mathcal{H}} + g^{-1}\|(\mathbf{W}, R)(t)\|_{\mathcal{H}} + \|(\mathbf{W}_\alpha, R_\alpha)(t)\|_{\mathcal{H}} \le C\epsilon, \quad t \in [0,\ T], \tag{9.11}$$

provided that T is much smaller than ϵ^{-2}. In view of our local well-posedness result this shows that the solutions can be extended up to time $T_\epsilon = C\epsilon^{-2}$, concluding the proof of the theorem.

In order to prove (9.11) we use a bootstrap argument; we make the bootstrap assumption

$$g^{-1}\|(W, Q)(t)\|_{\mathcal{H}} + g^{-1}\|(\mathbf{W}, R)(t)\|_{\mathcal{H}} + \|(\mathbf{W}_\alpha, R_\alpha)(t)\|_{\mathcal{H}} \le 2C\epsilon, \quad t \in [0,\ T]. \tag{9.12}$$

From (9.12), and by Sobolev embedding theorem, (1.15) and (1.16), our control norms A and B satisfy

$$A, B \lesssim C\epsilon.$$

To bound (W, Q) in time we directly use the conserved energy $\mathcal{E}$. Using the expression (1.7) for $\mathcal{E}$ we see that

$$\mathcal{E} = (1 + O(A))E_0(W, Q).$$

Hence, using the bootstrap assumption (9.12) we obtain

$$\|(W, Q)\|_{\mathcal{H}} \lesssim g(\epsilon + C\epsilon^2).$$

The bound for $(\mathbf{W}, R)$ can be obtained from the cubic energy estimates already established for the differentiated equation in Proposition 8.2. To obtain such a bound we first need to recall that the cubic energy estimate in there is in terms of the control norm $\mathbf{N}_1$, which is now taken uniformly in time. Explicitly, we integrate (8.2) in time

$$E^{1,(3)}(\mathbf{W}, R)(t) \lesssim E^{1,(3)}(\mathbf{W}, R)(0) + TAB\mathbf{N}_1^2, \tag{9.13}$$

and use (8.1) to obtain

$$E_0(\mathbf{W}, R)(t) \lesssim E_0(\mathbf{W}, R)(0) + TAB\mathbf{N}_1^2 + A\mathbf{N}_1^2. \tag{9.14}$$

We further need control of $\mathbf{N}_1$ norm, and this follows from

$$\mathbf{N}_1^2 \lesssim_A E_0(W, Q) + E_0(\mathbf{W}, R), \tag{9.15}$$

where the first term on the right is needed in order to account for the low frequencies in $(\mathbf{W}, R)$. Thus, we arrive at

$$\begin{aligned}\|(\mathbf{W}, R)\|^2_{L^\infty(0,T),\mathcal{H}} \lesssim & E_0(\mathbf{W}, R)(0) + TAB \sup_{t\in[0,T]} (E_0(W, Q)(t) + E_0(\mathbf{W}, R)(t)) \\ & + A \sup_{t\in[0,T]} ((E_0(W, Q)(t) + E_0(\mathbf{W}, R)(t)),\end{aligned}$$

and using the bootstrap assumptions (9.12) we get

$$\|(\mathbf{W}, R)\|_{\mathcal{H}} \lesssim g(\epsilon + TC^2\epsilon^3 + C\epsilon^2).$$

The bound $(\mathbf{W}_\alpha, R_\alpha)$ is obtained in the same way as above

$$\begin{aligned}\|(\mathbf{W}_\alpha, R_\alpha)\|^2_{L^\infty(0,T),\mathcal{H}} \lesssim & E_0(\mathbf{W}, R)(0) \\ & + TAB \sup_{t\in[0,T]} (E_0(W, Q)(t) + E_0(\mathbf{W}, R)(t) + E_0(\mathbf{W}_\alpha, R_\alpha)(t)) \\ & + A \sup_{t\in[0,T]} ((E_0(W, Q)(t) + E_0(\mathbf{W}, R)(t) + E_0(\mathbf{W}_\alpha, R_\alpha)(t)),\end{aligned}$$

and using the bootstrap assumptions (9.12) we get

$$\|(\mathbf{W}_\alpha, R_\alpha)\|_{\mathcal{H}} \lesssim \epsilon + TC^2\epsilon^3 + C\epsilon^2.$$

Hence, the estimate in (9.11) follows provided that $C \gg 1$ and $T \ll C^{-1}\epsilon^{-2}$. Similar bootstrap argument applies for higher derivatives. $\square$

Appendix A. Multilinear estimates

A.1. **Some harmonic analysis results.** In this section we collect a number of elementary estimates that will allow us to adapt the estimates established in infinite depth case [8] to the finite depth setting.

We take an inhomogeneous Littlewood-Paley decomposition $I = S_0 + \sum_{j\geq 1} P_j$ and denote

$$f_0 = S_0 f, \qquad f_j = P_j f, \quad j \geq 1.$$

We define the inhomogeneous Besov space $B^{s,p}_q$ with norm

$$\|f\|^q_{B^{s,p}_q} = \sum_{j\geq 0} \|\langle D\rangle^s f_j\|^q_{L^p},$$

with the usual modification when $q = \infty$. We also define the inhomogeneous space bmo of functions of bounded mean oscillation with norm

$$\|f\|_{\mathrm{bmo}} = \|f\|_{\mathrm{BMO}} + \|f_0\|_{L^\infty},$$

where

$$\|f\|_{\mathrm{BMO}} = \sup_Q \frac{1}{|Q|} \int_Q |f - f_Q| \, d\alpha, \qquad f_Q = \int_Q f \, d\alpha,$$

and the supremum is taken over all intervals $Q \subset \mathbb{R}$. We recall that $B^{0,\infty}_2 \subset \mathrm{bmo} \subset B^{0,\infty}_\infty$. We define the corresponding bmo-Sobolev spaces by

$$\|u\|_{\mathrm{bmo}^s} = \|\langle D\rangle^s u\|_{\mathrm{bmo}}.$$

We define the paraproduct operators

$$T_f g = \sum_{j>4} f_{<j-4} g_j, \qquad \Pi[f,g] = \sum_{\substack{|j-k|\leq 4\\ j,k\geq 0}} f_j g_k,$$

and the associated product decomposition

$$fg = T_f g + T_g f + \Pi[f,g].$$

We then have the following estimates (see for example [8, Propositions 2.2, 2.6]):

Lemma A.1 (Paraproduct bounds).

a) Coifman-Meyer paraproduct estimates. For $1 < p < \infty$ and $s, \sigma \geq 0$,

$$\begin{aligned} \|\langle D\rangle^s T_{\langle D\rangle^\sigma u} f\|_{L^p} &\lesssim \|f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{L^p}, \\ \|\langle D\rangle^s \Pi[f, \langle D\rangle^\sigma u]\|_{L^p} &\lesssim \|f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{L^p}. \end{aligned} \tag{A.1}$$

b) Besov endpoint estimates. For $s \geq 0$,

$$\begin{aligned} \|\langle D\rangle^s T_{\langle D\rangle^\sigma u} f\|_{L^\infty} &\lesssim \|f\|_{B^{s+\sigma,\infty}_2} \|u\|_{B^{0,\infty}_2}, \qquad \sigma > 0 \\ \|\langle D\rangle^s \Pi[f, \langle D\rangle^\sigma u]\|_{L^\infty} &\lesssim \|f\|_{B^{s+\sigma,\infty}_2} \|u\|_{B^{0,\infty}_2}, \qquad \sigma \geq 0. \end{aligned} \tag{A.2}$$

c) BMO *endpoint estimates. For $s \geq 0$,*

$$\begin{aligned} \|\langle D\rangle^s T_{\langle D\rangle^\sigma u} f\|_{\mathrm{bmo}} &\lesssim \|f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{\mathrm{bmo}}, \qquad \sigma > 0 \\ \|\langle D\rangle^s \Pi[f, \langle D\rangle^\sigma u]\|_{\mathrm{bmo}} &\lesssim \|f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{\mathrm{bmo}}, \qquad \sigma \geq 0. \end{aligned} \tag{A.3}$$

The following bounds, which are direct consequences of the classical Coifman-Meyer estimates, are closely related:

Lemma A.2 (Commutator bounds)**.**
a) Let $\mathcal{M} \in S^1$ be a smooth Fourier multiplier with principal symbol homogeneous of order 1. Then for $1 < p < \infty$ we have the estimate

$$\|[\mathcal{M}, f]u\|_{L^p} \lesssim \|f_\alpha\|_{L^\infty} \|u\|_{L^p}. \tag{A.4}$$

b) Let $\mathcal{M} \in S^s$ be a smooth Fourier multiplier with principal symbol homogeneous of order s with $0 \leq s < 1$. Then for $1 < p < \infty$ we have the estimate

$$\|[\mathcal{M}, f]u\|_{L^p} \lesssim \|\mathcal{T} f\|_{\mathrm{bmo}^s} \|u\|_{L^p}. \tag{A.5}$$

We also need the following more involved estimate:

Lemma A.3. *The following double commutator bound holds:*

$$\|[[\mathcal{T}L, b], L]\|_{L^2 \to L^2} \lesssim \|b_\alpha\|_{\mathrm{bmo}}. \tag{A.6}$$

Proof. We consider the paradifferential decomposition of the multiplication by b. For the map

$$u \to T_u b,$$

we have the bounds

$$\|T_u b\|_{H^1} \lesssim \|u\|_{L^2} \|b_\alpha\|_{\mathrm{bmo}}, \qquad \|T_u b\|_{L^2} \lesssim \|u\|_{H^{-1}} \|b_\alpha\|_{\mathrm{bmo}},$$

which follow from the first estimate in (A.1). Thus we can neglect the commutator structure.
Similarly, for the map

$$u \to \Pi[u, b],$$

we have the bounds

$$\|\Pi[u, b]\|_{H^1} \lesssim \|u\|_{L^2} \|b_\alpha\|_{\mathrm{bmo}}, \qquad \|\Pi[u, b]\|_{L^2} \lesssim \|u\|_{H^{-1}} \|b_\alpha\|_{\mathrm{bmo}},$$

from the second estimate in (A.1), and again we can neglect the commutator structure.
It remains to consider the contribution of T_b. For this we write

$$[T_b, \mathcal{T}L]u = \sum_k [b_{<k-4}, \mathcal{T}L]u_k = \sum_k 2^{-\frac{k}{2}} B_k(\partial_\alpha b_{<k-4}, u_k),$$

where B_k are translation invariant bilinear operators with uniformly integrable kernels. Commuting again we have

$$[[T_b, \mathcal{T}L], L]u = \sum_k \sum_k 2^{-\frac{k}{2}} B_k([\partial_\alpha b_{<k-4}, L], u_k) = \sum_k 2^{-k} C_k(\partial_\alpha^2 b_{<k-4}, u_k),$$

where again C_k are translation invariant bilinear operators with uniformly integrable kernels. Then we can bound

$$\|[[T_b, \mathcal{T}L], L]u\|_{L^2} \lesssim \sum_k 2^{-k} \|\partial_\alpha^2 b_{<k-4}\|_{L^\infty} \|u_k\|_{L^2} \lesssim \|b_\alpha\|_{B^{0,\infty}_\infty} \|u\|_{L^2},$$

which suffices. □

We will make use of the following estimates for rapidly decaying Fourier multipliers:

Lemma A.4. *Let S be a Fourier multiplier with Schwartz symbol. Then for all real s, σ, $1 \leq p \leq \infty$ and $N \geq 0$ we have the estimate*

$$\|\langle D\rangle^s S T_f \langle D\rangle^\sigma u\|_{L^p} + \|\langle D\rangle^s T_f S \langle D\rangle^\sigma u\|_{L^p} \lesssim_N \|f\|_{B^{-N,\infty}_\infty} \|u\|_{L^p}. \tag{A.7}$$

Further, we have the commutator estimate

$$\|\langle D\rangle^s [S, T_f] \langle D\rangle^\sigma u\|_{L^p} \lesssim_N \|\mathcal{T} f\|_{B^{-N,\infty}_\infty} \|u\|_{L^p}. \tag{A.8}$$

Proof. This is a standard argument based on the classical Littlewood-Paley trichotomy. Due to the frequency localization of the paraproduct operator T_f, the rapid decay in the symbol of S transfers to both the input u, the factor f and to the output. This directly leads to the derivative gains in the Lemma. □

The next result serves to bound commutators with the Tilbert transform $\mathcal{T}$:

Lemma A.5. *Let $\mathcal{M}$ be a Fourier multiplier whose symbol $m(x)$ is bounded with $m'(\xi)$ in the Schwartz class. Then for $1 < p < \infty$ and $s \geq 0$ we have the commutator estimates*

$$\begin{aligned} \|\langle D\rangle^s [\mathcal{M}, f] \langle D\rangle^\sigma u\|_{L^2} &\lesssim \|\mathcal{T} f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{L^2}, & \sigma \geq 0 \\ \|\langle D\rangle^s [\mathcal{M}, f] \langle D\rangle^\sigma u\|_{L^\infty} &\lesssim \|\mathcal{T} f\|_{B^{s+\sigma,\infty}_2} \|u\|_{B^{0,\infty}_2}, & \sigma > 0 \\ \|\langle D\rangle^s [\mathcal{M}, f] \langle D\rangle^\sigma u\|_{\mathrm{bmo}} &\lesssim \|\mathcal{T} f\|_{\mathrm{bmo}^{s+\sigma}} \|u\|_{\mathrm{bmo}}, & \sigma > 0. \end{aligned} \tag{A.9}$$

Proof. By hypothesis we can split the multiplier $\mathcal{M}$ as

$$\mathcal{M} = m(\infty) P_{>10} + m(-\infty) P_{<-10} + S,$$

where $P_{>10}$ and $P_{<-10}$ are multipliers whose symbols are smooth cutoff functions selecting the indicated frequency regions, and S has Schwartz kernel. For the commutator with $P_{>10}$ (and similarly with $P_{<-10}$) we have

$$\langle D\rangle^s [P_{>10}, T_f] \langle D\rangle^\sigma u_{>20} \equiv 0, \qquad \langle D\rangle^s [P_{>10}, f_0] \langle D\rangle^\sigma u_0 \equiv 0.$$

The estimates then follow from Lemma A.1 as in the infinite depth case [8].

For the second term we write

$$\begin{aligned} \langle D\rangle^s [S, f] \langle D\rangle^\sigma u = {} & \langle D\rangle^s [S, T_f] \langle D\rangle^\sigma u + \langle D\rangle^s [S, f_0] \langle D\rangle^\sigma u_{\leq 4} + \langle D\rangle^s S T_{\langle D\rangle^\sigma u} f \\ & - \langle D\rangle^s T_{S\langle D\rangle^\sigma u} f + \langle D\rangle^s S \Pi[f_{\geq 1}, \langle D\rangle^\sigma u] - \langle D\rangle^s \Pi[f_{\geq 1}, S\langle D\rangle^\sigma u]. \end{aligned}$$

The first and second terms term may be estimated using (A.8). The remaining terms may be estimated using Lemma A.1. □

Finally we recall two Moser estimates, the first of which is classical, and the second from [8].

Lemma A.6. *Let F be a smooth function such that $F(0) = 0$ then for $s \geq 0$ and $u \in L^\infty \cap H^s$ we have the Moser estimate*

$$\|F(u)\|_{H^s} \lesssim_{\|u\|_{L^\infty}} \|u\|_{H^s}. \tag{A.10}$$

Similarly, for $u \in \mathrm{bmo}^s$ we have

$$\|F(u)\|_{\mathrm{bmo}^s} \lesssim_{\|u\|_{L^\infty}} \|u\|_{\mathrm{bmo}^s}. \tag{A.11}$$

A.2. **Holomorphic functions on the strip.** We recall the projection to holomorphic functions is given by

$$\mathbf{P}u = \frac{1}{2}\left[(1-i\mathcal{T})\operatorname{Re} u + i(1+i\mathcal{T}^{-1})\operatorname{Im} u\right] = \frac{1}{4}\left[(2-i\mathcal{T}+i\mathcal{T}^{-1})u - i(\mathcal{T}+\mathcal{T}^{-1})\bar{u}\right].$$

As a consequence,

$$\operatorname{Re}\mathbf{P}u = \frac{1}{2}\left[\operatorname{Re} u - \mathcal{T}^{-1}\operatorname{Im} u\right] = \frac{1}{4}\left[(1+i\mathcal{T}^{-1})u + (1-i\mathcal{T}^{-1})\bar{u}\right],$$
$$\operatorname{Im}\mathbf{P}u = -\frac{1}{2}\left[\mathcal{T}\operatorname{Re} u - \operatorname{Im} u\right] = \frac{1}{4i}\left[(1-i\mathcal{T})u - (1+i\mathcal{T})\bar{u}\right].$$

We also recall the definition of the inner product, which is given by

$$\begin{aligned}\langle u, v\rangle &= \int \mathcal{T}\operatorname{Re} u\cdot\mathcal{T}\operatorname{Re} v + \operatorname{Im} u\cdot\operatorname{Im} v\, d\alpha\\ &= \frac{1}{2}\operatorname{Re}\int (\mathcal{T}u\cdot\mathcal{T}\bar{v} + u\cdot\bar{v}) + (\mathcal{T}u\cdot\mathcal{T}v - u\cdot v)\ d\alpha.\end{aligned}$$

It is useful to understand the adjoints of multiplication operators with respect to this inner product:

Lemma A.7. *Let f be a complex-valued function. With respect the inner product $\langle\cdot,\cdot\rangle$ the adjoint of the operator $\mathfrak{M}_f$ is*

$$\mathfrak{M}_f^* u = \mathcal{T}^{-1}(\mathbf{P}-\bar{\mathbf{P}})\left[\bar{f}\mathcal{T}\mathbf{P}[u]\right]. \tag{A.12}$$

Proof. Using that $\mathcal{T}\operatorname{Re}\mathbf{P}[v] = -\operatorname{Im}\mathbf{P}[v]$ and that $\mathcal{T}$ is skew-symmetric we may write the inner product as

$$\begin{aligned}\langle\mathfrak{M}_f u, v\rangle &= \int \operatorname{Re}[fu]\cdot\mathcal{T}\operatorname{Im}\mathbf{P}[v] - \operatorname{Im}[fu]\cdot\mathcal{T}\operatorname{Re}\mathbf{P}[v]\\ &= -\int \mathcal{T}\operatorname{Re} u\cdot\mathcal{T}^{-1}\operatorname{Im}(\bar{f}\mathcal{T}\mathbf{P}[v]) + \operatorname{Im} u\cdot\operatorname{Re}(\bar{f}\mathcal{T}\mathbf{P}[v]).\end{aligned}$$

As a consequence we have

$$\mathfrak{M}_f^* v = -\mathcal{T}^{-2}\operatorname{Im}(\bar{f}\mathcal{T}\mathbf{P}[v]) - i\operatorname{Re}(\bar{f}\mathcal{T}\mathbf{P}[v]).$$

Comparing this to the expression for $\mathbf{P}$ we obtain the formula (A.12). ☐

The following immediate consequence of the above Lemma is very handy to use:

Corollary A.8. *If u and v are holomorphic functions in $\mathfrak{H}$ then we have*

$$\langle f\mathcal{T}u, v\rangle = -\langle u, \bar{f}\mathcal{T}v\rangle. \tag{A.13}$$

As our function spaces $\mathfrak{H}$, $\mathcal{H}$ lose a derivative at low frequency in the real component, for a space X of complex-valued functions we define the norm

$$\|f\|_{\mathfrak{H}X}^2 = \|\mathcal{T}\operatorname{Re} f\|_X^2 + \|\operatorname{Im} f\|_X^2,$$

with the shorthand $\mathfrak{H} = \mathfrak{H}L^2$. We will frequently use the following estimate for the commutator with the projection to holomorphic functions:

Lemma A.9. *For $s \geq 0$ we have the estimates*

$$\begin{aligned} \|\langle D\rangle^s[\mathbf{P}, f]\langle D\rangle^\sigma \mathcal{T} g\|_{\mathfrak{H}} &\lesssim \|f\|_{\mathfrak{H}\mathrm{bmo}^{s+\sigma}}\|g\|_{\mathfrak{H}}, \qquad &\sigma \geq 0 \\ \|\langle D\rangle^s[\mathbf{P}, f]\langle D\rangle^\sigma \mathcal{T} g\|_{\mathfrak{H}L^\infty} &\lesssim \|f\|_{\mathfrak{H}B_2^{s+\sigma,\infty}}\|g\|_{\mathfrak{H}B_2^{0,\infty}}, \qquad &\sigma > 0. \end{aligned} \tag{A.14}$$

Proof. We may write the real and imaginary parts of the commutator as

$$\begin{aligned} \mathcal{T}\,\mathrm{Re}[\mathbf{P}, f]\langle D\rangle^\sigma \mathcal{T} g &= \frac{1}{2}[\mathcal{T}, \mathrm{Re}\, f]\langle D\rangle^\sigma \,\mathrm{Im}\, g - \frac{1}{2}[\mathcal{T}, \mathrm{Im}\, f]\langle D\rangle^\sigma \mathcal{T}^2\,\mathrm{Re}\, g \\ &\quad - \frac{1}{2}\,\mathrm{Im}\, f(1+\mathcal{T}^2)\langle D\rangle^\sigma \mathcal{T}\,\mathrm{Re}\, g, \\ \mathrm{Im}[\mathbf{P}, f]\langle D\rangle^\sigma \mathcal{T} g &= -\frac{1}{2}[\mathcal{T}, \mathrm{Re}\, f]\langle D\rangle^\sigma \mathcal{T}\,\mathrm{Re}\, g + \frac{1}{2}[\mathcal{T}, \mathrm{Im}\, f]\langle D\rangle^\sigma \mathcal{T}\,\mathrm{Im}\, g \\ &\quad + \frac{1}{2}\,\mathrm{Im}\, f\langle D\rangle^\sigma(1+\mathcal{T}^2)\,\mathrm{Im}\, g. \end{aligned}$$

The estimates then follow from the commutator estimate A.9 and the paraproduct estimates (A.1), (A.2) and (A.7), using that the operator $1+\mathcal{T}^2$ has Schwartz symbol. □

Finally we prove the following lemma that allows us to estimate the product of two holomorphic functions in negative Sobolev spaces:

Lemma A.10. *If f, g are holomorphic then for $s > 0$ and $2 \leq p, q \leq \infty$ satisfying $\frac{1}{p} + \frac{1}{q} = \frac{1}{2}$ we have the estimate*

$$\|fg\|_{H^{-s}} \lesssim \|f\|_{L^p}\|g\|_{B_2^{-s,q}}. \tag{A.15}$$

Proof. For each $j \geq 0$ we decompose

$$\|P_j[fg]\|_{L^2} = \|P_j[fg_{\leq j+10}]\|_{L^2} + \|P_j[fg_{>j+10}]\|_{L^2}.$$

The first term may now be estimated using dyadic decomposition. For the second term both f, g must be localized at comparable dyadic frequencies $\gg 2^j$. In particular, one term must be localized at negative wavenumbers and the other at positive wavenumbers. However, as both terms are holomorphic we may harmlessly apply the projection $\mathbf{P}$ to each term, which is rapidly decaying on positive wavenumbers. □

A.3. **Water wave related bounds.** We begin with the following result for the function Y which follows directly from [8, Lemma 2.5] and Moser type estimates:

Lemma A.11. *The function $Y = \dfrac{\mathbf{W}}{1+\mathbf{W}}$ satisfies the bounds*

$$\|Y\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim_A g^{-\frac{1}{2}} B, \tag{A.16}$$

respectively

$$\|Y\|_{H^{n-1}} \lesssim_A g^{-\frac{1}{2}}\mathbf{N}_n, \qquad n \geq 1. \tag{A.17}$$

Next we consider the advection velocity b:

Lemma A.12. *The the advection velocity b satisfies the estimates*

$$\|\mathcal{T}b\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim_A g^{\frac{1}{2}}A, \qquad \|\mathcal{T}b\|_{\mathrm{bmo}^1} \lesssim_A B. \tag{A.18}$$

respectively

$$\|\mathcal{T}b\|_{H^{n-\frac{1}{2}}} \lesssim_A \mathbf{N}_n, \qquad n \geq 1. \tag{A.19}$$

Proof. We write $b = b_1 + b_2$ where

$$b_1 = 2\operatorname{Re} R, \qquad b_2 = -2\operatorname{Re}\mathbf{P}[R\bar{Y}].$$

For b_1 we have the estimate

$$\|\mathcal{T}b_1\|_{\mathrm{bmo}^s} \leq \|R\|_{\mathrm{bmo}^s}.$$

For b_2 we may write $\mathbf{P}[R\bar{Y}] = [\mathbf{P}, R]\bar{Y}$ so

$$\|\mathcal{T}b_2\|_{\mathrm{bmo}^s} = \|\mathcal{T}\operatorname{Re}[\mathbf{P}, R]\bar{Y}\|_{\mathrm{bmo}^s}.$$

As Y is antiholomorphic we have

$$\mathcal{T}\operatorname{Re}[\mathbf{P}, R]\bar{Y} = \frac{1}{2}[\mathcal{T}, \operatorname{Re} R]\operatorname{Re}\bar{Y} - \frac{1}{2}[\mathcal{T}, \operatorname{Im} R]\operatorname{Im}\bar{Y} - \frac{1}{2}\operatorname{Im} R(1+\mathcal{T}^2)\operatorname{Re}\bar{Y}.$$

For the first two terms we may use the commutator estimate (A.9) to obtain

$$\|[\mathcal{T}, \operatorname{Re} R]\operatorname{Re}\bar{Y}\|_{\mathrm{bmo}^s} + \|[\mathcal{T}, \operatorname{Im} R]\operatorname{Im}\bar{Y}\|_{\mathrm{bmo}^s} \lesssim A\|R\|_{\mathrm{bmo}^s}.$$

For the final term we simply use that $1+\mathcal{T}^2$ has Schwartz symbol to estimate

$$\|\operatorname{Im} R(1+\mathcal{T}^2)\operatorname{Re}\bar{Y}\|_{\mathrm{bmo}^s} \lesssim A\|R\|_{\mathrm{bmo}^s}.$$

The proof of the L^2-type bound follows in a similar manner. $\square$

Next we prove a number of estimates for the real frequency shift $\mathfrak{a}$. Our estimates are similar to [8, Proposition 2.6] although the present case is slightly more involved due to the different projector $\mathbf{P}$, as well as the extra term in $\mathfrak{a}$.

Lemma A.13. *The following bounds hold for the frequency shift $\mathfrak{a}$:*

$$\|\mathfrak{a}\|_{L^\infty} \lesssim_A gA, \qquad \|\mathfrak{a}\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim_A g^{\frac{1}{2}}B, \tag{A.20}$$

$$\|\mathfrak{a}\|_{H^{n-1}} \lesssim_A g^{\frac{1}{2}}\mathbf{N}_n \tag{A.21}$$

$$\|\mathfrak{a}_t + b\mathfrak{a}_\alpha + g(1+\mathcal{T}^2)\operatorname{Re} R_\alpha\|_{L^\infty} \lesssim gAB. \tag{A.22}$$

Proof. We recall that $\mathfrak{a} = a + a_1$ where

$$a = 2\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha], \qquad a_1 = g(1+\mathcal{T}^2)\operatorname{Re}\mathbf{W}.$$

We will prove the bounds in the Lemma separately for a and for a_1.

1. L^∞, $\mathrm{bmo}^{\frac{1}{2}}$ and H^{n-1} bounds. For a_1 we use that $1+\mathcal{T}^2$ has Schwartz symbol to obtain

$$\|a_1\|_{L^\infty} \lesssim g\|\mathbf{W}\|_{L^\infty}, \qquad \|a_1\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim g\|\mathbf{W}\|_{\mathrm{bmo}^{\frac{1}{2}}}, \qquad \|a_1\|_{H^{n-1}} \lesssim g\|\mathbf{W}\|_{H^{n-1}}.$$

For a we use that $\mathbf{P}\bar{R}_\alpha = 0$ to write $\operatorname{Im}\mathbf{P}[R\bar{R}_\alpha] = \operatorname{Im}[\mathbf{P}, R]\bar{R}_\alpha$. We then apply the commutator estimate (A.14) to obtain

$$\|a\|_{L^\infty} \lesssim \|R\|^2_{B_2^{\frac{1}{2},\infty}}, \qquad \|a\|_{\mathrm{bmo}^{\frac{1}{2}}} \lesssim \|\langle D\rangle^{\frac{1}{2}}a\|_{L^\infty} \lesssim \|R\|^2_{B_2^{\frac{3}{4},\infty}},$$

and for the second of these we apply the interpolation estimate

$$\|R\|^2_{B^{\frac{3}{4},\infty}_2} \lesssim \|\langle D\rangle^{\frac{1}{2}} R\|_{L^\infty} \|R\|_{\mathrm{bmo}^1}.$$

For the H^{n-1} estimate we first differentiate

$$\partial^{n-1} a = 2\,\mathrm{Im} \sum_{k=0}^{n-1} \mathbf{P}[R^{(k)} \bar{R}^{(n-k)}].$$

If $k \geq 1$ then we estimate by interpolation and if $k = 0$ then we apply the commutator bound (A.14).

2. Transport equation bounds. For a_1 we calculate

$$(\partial_t + b\partial_\alpha) a_1 + g(1+\mathcal{T}^2)\,\mathrm{Re}\, R_\alpha = g(1+\mathcal{T}^2)\,\mathrm{Re}\,[\mathbf{W}_t + b\mathbf{W}_\alpha + \mathbf{R}_\alpha] - ig[\mathcal{T}, b]\mathbf{W}_\alpha.$$

The first term may be bounded using Lemmas A.1, A.4, the estimate (A.24) for M and that $1+\mathcal{T}^2$ has Schwartz symbol. For the second term we apply the commutator estimate (A.9) to obtain

$$\|g[\mathcal{T}, b]\mathbf{W}_\alpha\|_{L^\infty} \lesssim g\|\mathcal{T} b\|_{B^{\frac{3}{4},\infty}_2} \|\mathbf{W}\|_{B^{\frac{1}{4},\infty}_2}.$$

By interpolation,

$$\|\mathcal{T} b\|_{B^{\frac{3}{4},\infty}_2} \lesssim \|\mathcal{T} b\|^{\frac{1}{2}}_{\mathrm{bmo}^{\frac{1}{2}}} \|\mathcal{T} b\|^{\frac{1}{2}}_{\mathrm{bmo}^1}, \qquad \|\mathbf{W}\|_{B^{\frac{1}{4},\infty}_2} \lesssim \|\mathbf{W}\|^{\frac{1}{2}}_{L^\infty} \|\mathbf{W}\|^{\frac{1}{2}}_{\mathrm{bmo}^{\frac{1}{2}}}.$$

and we may then apply the estimate (A.18) for b.

For a we have

$$\begin{aligned}(\partial_t + b\partial_\alpha) a = {} & 2\,\mathrm{Im}[\mathbf{P}, \mathbf{P}\,[R_t + bR_\alpha]]\bar{R}_\alpha + 2\,\mathrm{Im}[\mathbf{P}, R]\partial_\alpha \bar{\mathbf{P}}\left[\bar{R}_t + b\bar{R}_\alpha\right] \\ & + 2\,\mathrm{Im}\left(b\partial_\alpha \mathbf{P}[R\bar{R}_\alpha] - \mathbf{P}\left[bR_\alpha \bar{R}_\alpha\right] - \mathbf{P}\left[R\partial_\alpha \bar{\mathbf{P}}(b\bar{R}_\alpha)\right]\right).\end{aligned}$$

For the first two terms we apply the commutator estimate (A.14) to obtain

$$\begin{aligned}&\left\|2\,\mathrm{Im}[\mathbf{P}, \mathbf{P}\,[R_t + bR_\alpha]]\bar{R}_\alpha\right\|_{L^\infty} + \left\|2\,\mathrm{Im}[\mathbf{P}, R]\partial_\alpha \bar{\mathbf{P}}\left[\bar{R}_t + b\bar{R}_\alpha\right]\right\|_{L^\infty} \\ &\quad \lesssim \|\,\mathrm{Im}\,\mathbf{P}\,[R_t + bR_\alpha]\,\|_{B^{\frac{1}{4},\infty}_2} \|R\|_{B^{\frac{3}{4},\infty}_2}.\end{aligned}$$

We observe that

$$\mathrm{Im}\,\mathbf{P}\,[R_t + bR_\alpha] = \frac{1}{2}(g+\mathfrak{a})\,\mathrm{Re}\, Y - \frac{1}{2}\mathfrak{a} + \frac{1}{2}\mathcal{T}\left[(g+\mathfrak{a})\,\mathrm{Im}\, Y\right],$$

and hence

$$\|\,\mathrm{Im}\,\mathbf{P}\,[R_t + bR_\alpha]\,\|_{B^{\frac{1}{4},\infty}_2} \lesssim \|\mathfrak{a}\|_{B^{\frac{1}{4},\infty}_2}(1 + \|Y\|_{L^\infty}) + (g + \|\mathfrak{a}\|_{L^\infty})\|Y\|_{B^{\frac{1}{4},\infty}_2}.$$

and the estimate follows from interpolation and the estimates (A.20) for $\mathfrak{a}$ and (A.16) for Y.

For the final term appearing in $a_t + ba_\alpha$ we must ensure that b does not appear undifferentiated at low frequency. We start by dividing up dyadically according to the frequency of the holomorphic term R:

$$b\partial_\alpha \mathbf{P}[R\bar{R}_\alpha] - \mathbf{P}\left[bR_\alpha \bar{R}_\alpha\right] - \mathbf{P}\left[R\partial_\alpha \bar{\mathbf{P}}(b\bar{R}_\alpha)\right] = \sum_{j\geq 0} f_j,$$

where

$$f_j = b\partial_\alpha \mathbf{P}[R_j \bar{R}_\alpha] - \mathbf{P}\left[bR_{\alpha,j}\bar{R}_\alpha\right] - \mathbf{P}\left[R_j \partial_\alpha \bar{\mathbf{P}}(b\bar{R}_\alpha)\right].$$

We then decompose each $f_j = f_j^{high} + f_j^{low}$ according to the frequency balance of b and R,

$$f_j^{high} = b_{>j}\partial_\alpha \mathbf{P}[R_j\bar{R}_\alpha] - \mathbf{P}\left[b_{>j}R_{\alpha,j}\bar{R}_\alpha\right] - \mathbf{P}\left[R_j\partial_\alpha\bar{\mathbf{P}}(b_{>j}\bar{R}_\alpha)\right],$$
$$f_j^{low} = b_{\leq j}\partial_\alpha \mathbf{P}[R_j\bar{R}_\alpha] - \mathbf{P}\left[b_{\leq j}R_{\alpha,j}\bar{R}_\alpha\right] - \mathbf{P}\left[R_j\partial_\alpha\bar{\mathbf{P}}(b_{\leq j}\bar{R}_\alpha)\right].$$

When b is at high frequency we write

$$f_j^{high} = b_{>j}\partial_\alpha[\mathbf{P}, R_j]\bar{R}_\alpha - [\mathbf{P}, b_{>j}R_{\alpha,j}]\bar{R}_\alpha - [\mathbf{P}, R_j]\partial_\alpha\bar{\mathbf{P}}(b_{>j}\bar{R}_\alpha).$$

Taking the imaginary part and applying the commutator estimate (A.14) we obtain

$$\|\operatorname{Im} f_j^{high}\|_{L^\infty} \lesssim 2^j \|b_{>j}\|_{B_2^{\frac{1}{4},\infty}} \|R_j\|_{L^\infty} \|R\|_{B_2^{\frac{3}{4},\infty}}.$$

Summing over $j \geq 0$ we obtain

$$\sum_{j\geq 0} \|\operatorname{Im} f_j^{high}\|_{L^\infty} \lesssim \|b_{>0}\|_{B_2^{\frac{3}{4},\infty}} \|R\|_{B_2^{\frac{1}{2},\infty}} \|R\|_{B_2^{\frac{3}{4},\infty}},$$

and the estimate follows from interpolation and the estimate (A.18) for b.

When b is at low frequency we write

$$f_j^{low} = \partial_\alpha[b_{\leq j}, \mathbf{P}](R_j\bar{R}_\alpha) - b_{\leq j,\alpha}[\mathbf{P}, R_j]\bar{R}_\alpha + \mathbf{P}[R_j\partial_\alpha[\mathbf{P}, b_{\leq j}]\bar{R}_\alpha].$$

Again we apply the commutator estimate (A.14), using that $b_{\leq j,\alpha}$ is real-valued, to obtain

$$\|\operatorname{Im} f_j^{low}\|_{L^\infty} \lesssim 2^{\frac{3}{8}j} \|\mathcal{T} b_{\leq j}\|_{B_2^{\frac{7}{8},\infty}} \|R_j\|_{L^\infty} \|R\|_{B_2^{\frac{3}{4},\infty}}.$$

Summing over $j \geq 0$ we obtain

$$\sum_{j\geq 0} \|\operatorname{Im} f_j^{low}\|_{L^\infty} \lesssim \|\mathcal{T} b\|_{B_2^{\frac{3}{4},\infty}} \|R\|_{B_2^{\frac{1}{2},\infty}} \|R\|_{B_2^{\frac{3}{4},\infty}} \lesssim_A AB,$$

which completes the proof of (A.22). □

We now estimate some of the secondary auxiliary functions d and M:

Lemma A.14. *We have the estimate*

$$\|d\|_{\mathrm{bmo}} \lesssim_A B. \tag{A.23}$$

Proof. We recall that

$$d = R_\alpha(1 - \bar{Y}).$$

As a consequence it suffices to show that

$$\|R_\alpha\bar{Y}\|_{\mathrm{bmo}} \lesssim AB.$$

Decomposing using paraproducts we have

$$R_\alpha\bar{Y} = T_{R_\alpha}\bar{Y} + T_{\bar{Y}}R_\alpha + \Pi[R_\alpha, \bar{Y}].$$

We then use (A.3) to estimate

$$\|T_{R_\alpha}\bar{Y}\|_{\mathrm{bmo}} \lesssim \|\langle D\rangle^{\frac{1}{2}} R\|_{L^\infty} \|Y\|_{\mathrm{bmo}^{\frac{1}{2}}}, \qquad \|\Pi[R_\alpha, \bar{Y}]\|_{\mathrm{bmo}} \lesssim \|R_\alpha\|_{\mathrm{bmo}} \|Y\|_{\mathrm{bmo}}.$$

For the remaining term we are unable to use (A.3), but we can obtain a similar estimate by relaxing bmo to L^∞ for the low frequency term (see [8, Proposition 2.2]),

$$\|T_{\bar{Y}}R_\alpha\|_{\mathrm{bmo}} \lesssim \|Y\|_{L^\infty} \|R_\alpha\|_{\mathrm{bmo}}.$$

The estimate (A.23) then follows.

□

Lemma A.15. *The function M satisfies the pointwise bounds*

$$\|M\|_{L^\infty} \lesssim AB, \tag{A.24}$$

as well as the Sobolev bounds for $n \geq 1$

$$\|M\|_{H^{k-\frac{3}{2}}} \lesssim_A A\mathbf{N}_k, \qquad \|M\|_{H^{k-1}} \lesssim_A g^{-\frac{1}{2}} B\mathbf{N}_k. \tag{A.25}$$

Proof. We start with the proof of (A.24). We first decompose $M = M_0 + M_{\geq 1}$ into a low and high frequency part.

For the high frequency part we first write M in the form

$$M = 2\operatorname{Re}[\mathbf{P}, R]\bar{Y}_\alpha - 2\operatorname{Re}[\mathbf{P}, Y]\bar{R}_\alpha,$$

and then use (A.14) to obtain

$$\|M_{\geq 1}\|_{L^\infty} \lesssim \|\mathcal{T} M\|_{L^\infty} \lesssim \|R\|_{B_2^{\frac{3}{4},\infty}} \|Y\|_{B_2^{\frac{1}{4},\infty}} \lesssim AB.$$

For the low frequency part we face an additional difficulty compared to the infinite depth case, which is due to the low frequenct unboundedness of the projector $\mathbf{P}$. To address this we observe that M has a certain null structure, by writing

$$M = 2\operatorname{Re}\mathbf{P}[R\bar{Y}_\alpha - Y\bar{R}_\alpha] = \operatorname{Re}[R\bar{Y}_\alpha - Y\bar{R}_\alpha] - \mathcal{T}^{-1}\partial_\alpha \operatorname{Im}(R\bar{Y}).$$

Applying the projection S_0 we obtain

$$\|M_0\|_{L^\infty} \lesssim \|\Pi[R, \bar{Y}_\alpha]\|_{L^\infty} + \|\Pi[Y, \bar{R}_\alpha]\|_{L^\infty} + \|\Pi[R, \bar{Y}]\|_{L^\infty}.$$

We may then estimate these terms using (A.2) to complete the proof of (A.24). The proof of (A.25) is similar. □

Lemma A.16. *The following estimates hold:*

$$\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)\mathbf{W}\|_{B^{0,\infty}_\infty} + g^{-\frac{1}{2}}\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)R\|_{B_\infty^{\frac{1}{2},\infty}} \lesssim_A AB, \tag{A.26}$$

respectively the L^2 bounds

$$g^{-\frac{1}{2}}\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)R\|_{H^{n-1}} \lesssim_A A\mathbf{N}_n, \qquad n \geq 1, \tag{A.27}$$

and

$$\|\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)\mathbf{W}\|_{H^{n-\frac{3}{2}}} \lesssim_A A\mathbf{N}_n, \qquad n \geq 2. \tag{A.28}$$

If instead $n = 1$ then for each k there is a decomposition

$$P_{<k}\Lambda^{\geq 2}(\partial_t + T_b\partial_\alpha)\mathbf{W} = F_k^1 + F_k^2,$$

so that

$$\|F_k^1\|_{L^2} \lesssim_A B\mathbf{N}_1, \qquad \|F_k^2\|_{L^2} \lesssim_A 2^{\frac{k}{2}} A\mathbf{N}_1. \tag{A.29}$$

Proof. We recall the equations for $(\mathbf{W}, R)$:

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \dfrac{1+\mathbf{W}}{1+\bar{\mathbf{W}}} R_\alpha = (1+\mathbf{W})M \\ R_t + bR_\alpha = i\dfrac{g\mathbf{W} - \mathfrak{a}}{1+\mathbf{W}}. \end{cases}$$

We begin with the pointwise bounds. For the M term we use (A.24). Next we estimate

$$\left\| \frac{\mathbf{W} - \bar{\mathbf{W}}}{1+\bar{\mathbf{W}}} R_\alpha \right\|_{\text{bmo}} \lesssim \|R_\alpha\|_{\text{bmo}} \left\| \frac{\mathbf{W} - \bar{\mathbf{W}}}{1+\bar{\mathbf{W}}} \right\|_{L^\infty} + \|R\|_{B_2^{\frac{3}{4},\infty}} \left\| \frac{\mathbf{W} - \bar{\mathbf{W}}}{1+\bar{\mathbf{W}}} \right\|_{B_2^{\frac{1}{4},\infty}} \lesssim AB,$$

which is akin to the bmo bound for d. For the Y term in the second equation we use (A.16) as well as the algebra property for $\text{bmo}^{\frac{1}{2}}$. The same applies for the a term in combination with (A.20).

It remains to bound the b terms, where we carefully note that no low frequencies of b are included here. Then using (A.18) we have

$$\|(b - T_b)\mathbf{W}_\alpha\|_{L^\infty} \lesssim \|\mathcal{T}b\|_{B_2^{\frac{3}{4},\infty}} \|\mathbf{W}\|_{B_2^{\frac{1}{4},\infty}} \lesssim AB,$$

respectively

$$\|(b - T_b)R_\alpha\|_{\text{bmo}^{\frac{1}{2}}} \lesssim \|\mathcal{T}b\|_{B_2^{\frac{3}{4},\infty}} \|R\|_{B_2^{\frac{3}{4},\infty}} \lesssim AB.$$

Next we consider the L^2 bounds. For the M term we use a standard Littlewood-Paley decomposition together with (A.24) and (A.25). For the a term we similarly use (A.20) and (A.21). For the b paradifferential remainder we use (A.18) and (A.19). The other terms follow in standard bilinear fashion.

In the case $n = 1$ the same method applies once we have produced a convenient decomposition of $(\partial_t + T_b\partial_\alpha)\mathbf{W}$. Precisely, all contributions go to F_k^1 except for those arising from the terms $(b_{<k+4} - T_b)\mathbf{W}_\alpha$, respectively $\dfrac{1+\mathbf{W}}{1+\bar{\mathbf{W}}} R_{<k+4,\alpha}$.

∎

Acknowledgments

This material is based in part upon work supported by the NSF under grant DMS-1440140 while the authors were in residence at the MSRI in Berkeley, California during the Fall 2015 semester. The last author was also supported by the NSF grant DMS-1266182. In addition, all three authors were supported in part by the Simons Foundation.

References

[1] T. Alazard, N. Burq, and C. Zuily. On the Cauchy problem for gravity water waves. *Invent. Math.*, 198(1):71–163, 2014.

[2] B. Alvarez-Samaniego and D. Lannes. Large time existence for 3D water-waves and asymptotics. *Invent. Math.*, 171(3):485–541, 2008.

[3] W. Choi and R. Camassa. Exact evolution equations for surface waves. *Journal of engineering mechanics*, 125(7):756–760, 1999.

[4] A. Constantin. *Nonlinear water waves with applications to wave-current interactions and tsunamis*, volume 81 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[5] W. Craig, C. Sulem, and P.-L. Sulem. Nonlinear modulation of gravity waves: a rigorous approach. *Nonlinearity*, 5(2):497–522, 1992.
[6] A. Dyachenko, E. Kuznetsov, M. Spector, and V. Zakharov. Analytical description of the free surface dynamics of an ideal fluid (canonical formalism and conformal mapping). *Physics Letters A*, 221(12):73 – 79, 1996.
[7] A. I. Dyachenko, V. E. Zakharov, and E. A. Kuznetsov. Nonlinear dynamics of the free surface of an ideal fluid. *Plasma Physics Reports*, 22:829–840, Oct. 1996.
[8] J. Hunter, M. Ifrim, and D. Tataru. Two dimensional water waves in holomorphic coordinates. *Comm. Math. Phys.,* to appear*, ArXiv eprints*, Jan. 2014.
[9] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong. Long time solutions for a Burgers-Hilbert equation via a modified energy method. *Proc. Amer. Math. Soc.*, 143(8):3407–3412, 2015.
[10] M. Ifrim and D. Tataru. The lifespan of small data solutions in two dimensional capillary water waves. *Arch. Rational Mech. Anal. ,* to appear*, ArXiv e-prints*, June 2014.
[11] M. Ifrim and D. Tataru. Two dimensional gravity water waves with constant vorticity: I. Cubic lifespan. *ArXiv e-prints*, Oct. 2015.
[12] D. Lannes. *The water waves problem*, volume 188 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2013. Mathematical analysis and asymptotics.
[13] Y. A. Li, J. M. Hyman, and W. Choi. A numerical study of the exact evolution equations for surface waves in water of finite depth. *Stud. Appl. Math.*, 113(3):303–324, 2004.
[14] V. I. Nalimov. The Cauchy-Poisson problem. *Dinamika Splošn. Sredy*, (Vyp. 18 Dinamika Zidkost. so Svobod. Granicami):104–210, 254, 1974.
[15] L. V. Ovsjannikov. To the shallow water theory foundation. *Arch. Mech. (Arch. Mech. Stos.)*, 26:407–422, 1974. Papers presented at the Eleventh Symposium on Advanced Problems and Methods in Fluid Mechanics, Kamienny Potok, 1973.
[16] J. Shatah. Normal forms and quadratic nonlinear Klein-Gordon equations. *Comm. Pure Appl. Math.*, 38(5):685–696, 1985.
[17] X. Wang. Global solution for the 3D gravity water waves system above a flat bottom. *ArXiv e-prints*, Aug. 2015.
[18] X. Wang. On 3D water waves system above a flat bottom. *ArXiv e-prints*, Aug. 2015.
[19] S. Wu. Well-posedness in Sobolev spaces of the full water wave problem in 2-D. *Invent. Math.*, 130(1):39–72, 1997.
[20] H. Yosihara. Gravity waves on the free surface of an incompressible perfect fluid of finite depth. *Publ. Res. Inst. Math. Sci.*, 18(1):49–96, 1982.
[21] V. E. Zakharov. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Journal of Applied Mechanics and Technical Physics*, 9(2):190–194, 1968.

Courant Institute of Mathematical Sciences, New York University
E-mail address: benjamin.harrop-griffiths@cims.nyu.edu

Department of Mathematics, University of California at Berkeley
E-mail address: ifrim@math.berkeley.edu

Department of Mathematics, University of California at Berkeley
E-mail address: tataru@math.berkeley.edu