Title
Two Dimensional Water Waves in Holomorphic Coordinates

Journal
Communications in Mathematical Physics, 346(2)

ISSN
0010-3616

Authors
Hunter, John K
Ifrim, Mihaela
Tataru, Daniel

Publication Date
2016-09-01

DOI
10.1007/s00220-016-2708-6

Peer reviewed

.

TWO DIMENSIONAL WATER WAVES IN HOLOMORPHIC COORDINATES

JOHN K. HUNTER, MIHAELA IFRIM, AND DANIEL TATARU

Abstract. This article is concerned with the infinite depth water wave equation in two space dimensions. We consider this problem expressed in position-velocity potential holomorphic coordinates. Viewing this problem as a quasilinear dispersive equation, we establish two results: (i) local well-posedness in Sobolev spaces, and (ii) almost global solutions for small localized data. Neither of these results are new; they have been recently obtained by Alazard-Burq-Zuily [1], respectively by Wu [23] using different coordinates and methods. Instead our goal is improve the understanding of this problem by providing a single setting for both problems, by proving sharper versions of the above results, as well as presenting new, simpler proofs. This article is self contained.

1. Introduction

We consider the two dimensional water wave equations with infinite depth with gravity but without surface tension. This is governed by the incompressible Euler's equations with boundary conditions on the water surface. Under the additional assumption that the flow is irrotational the fluid dynamics can be expressed in terms of a one-dimensional evolution of the water surface coupled with the trace of the velocity potential on the surface.

This problem was previously considered by several other authors. The local in time existence and uniqueness of solutions was proved in [15, 21, 22], both for finite and infinite depth. Later, Wu [23] proved almost global existence for small localized data. Very recently, global results for small localized data were independently obtained by Alazard & Delort [3] and by Ionescu & Pusateri [13]. Extensive work was also done on the same problem in three or higher space dimensions, and also on related problems with surface tension, vorticity, finite bottom, etc. Without being exhaustive, we list some of the more recent references [1, 2, 4, 5, 7, 8, 14, 16, 19, 25].

Our goal here is to revisit this problem and to provide a new, self-contained approach which, we hope, considerably simplifies and improves on many of the results mentioned above. Our analysis is based on the use of holomorphic coordinates, which are described below. Our results include:

(i) local well-posedness in Sobolev spaces, improving on previous regularity thresholds, e.g. in [1], up to the point where the transport vector field is no longer Lipschitz, and has merely a BMO derivative.

The first author was partially supported by the NSF under grant number DMS-1312342.

The second author was supported by the National Science Foundation under Grant No. 0932078 000, while the author was in residence at the Mathematical Science Research Institute in Berkeley, California, during the Fall semester 2013.

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

(ii) cubic life-span bounds for small data. These are related to the normal form method, but are instead proved by a modified energy method, inspired from the authors' previous article [10].

(iii) almost global well-posedness for small localized data, refining and simplifying Wu's approach in [23].

We consider both the case of the real line $\mathbb{R}$ and the periodic case $\mathbb{S}^1$. Our equations are expressed in coordinates (t, α) where α corresponds to the holomorphic parametrization of the water domain by the lower half-plane restricted to the real line. To write the equations we use the Hilbert transform H, as well as the operator

$$P = \frac{1}{2}(I - iH).$$

Note that P is a projector in $\mathbb{R}$ but not on $\mathbb{S}^1$.

Our variables (Z, Q) represent the position of the water surface, respectively the holomorphic extension of the velocity potential. These will be restricted to the closed subspace of holomorphic functions within various Sobolev spaces. Here we define holomorphic functions on $\mathbb{R}$ or on $\mathbb{S}^1$ as those whose Fourier transform is supported in $(-\infty, 0]$; equivalently, they admit a bounded holomorphic extension into the lower half-space. On $\mathbb{R}$ this can be described by the relation $Pf = f$, but on $\mathbb{S}^1$ we also need to make some adjustments for the constants.

There is a one dimensional degree of freedom in the choice of α, namely the horizontal translations. To fix this, in the real case we are considering waves which either decay at infinity,

$$\lim_{|\alpha| \to \infty} Z(\alpha) - \alpha = 0.$$

In the periodic case we instead assume that $Z(\alpha) - \alpha$ has period 2π and purely imaginary average. We can also harmlessly assume that Q has real average.

In position-velocity potential holomorphic coordinates the equations have the form

$$\begin{cases} Z_t + F Z_\alpha = 0, \\ Q_t + F Q_\alpha - i(Z - \alpha) + P\left[\dfrac{|Q_\alpha|^2}{J}\right] = 0, \end{cases}$$

where

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

For the derivation of the above equations, we refer the reader to *Appendix* A. In the real case these equations originate in [17]. The changes needed for the periodic case are also described in the same *Appendix* A. There are also other ways of expressing the equations, for instance in Cartesian coordinates using the Dirichlet to Neumann map associated to the water domain, see e.g. [1] . Here we prefer the holomorphic coordinates due to the simpler form of the equations; in particular, in these coordinates the Dirichlet to Neumann map is given in terms of the standard Hilbert transform.

It is convenient to work with a new variable, namely

$$W = Z - \alpha.$$

The equations become

$$
\begin{cases} W_t + F(1+W_\alpha) = 0, \\ Q_t + FQ_\alpha - iW + P\left[\dfrac{|Q_\alpha|^2}{J}\right] = 0, \end{cases} \tag{1.1}
$$

where

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

These equations are considered either in $\mathbb{R}\times\mathbb{R}$ or in $\mathbb{R}\times\mathbb{S}^1$.

As the system (1.1) is fully nonlinear, a standard procedure is to convert it into a quasilinear system by differentiating it. Observing that almost no undifferentiated functions appear in (1.1), one sees that by differentiation we get a self-contained first order quasilinear system for (W_α, Q_α). To write this system we introduce the auxiliary real function b, which we call the *advection velocity*, and is given by

$$
b = P\left[\frac{Q_\alpha}{J}\right] + \bar{P}\left[\frac{\bar{Q}_\alpha}{J}\right].
$$

The reason for this will be immediately apparent. Using b, the system (1.1) is written in the form

$$
\begin{cases} W_t + b(1+W_\alpha) = \dfrac{\bar{Q}_\alpha}{1+\bar{W}_\alpha}, \\ Q_t + bQ_\alpha - iW = \bar{P}\left[\dfrac{|Q_\alpha|^2}{J}\right], \end{cases}
$$

where the terms on the right are antiholomorphic and disappear when the equations are projected onto the holomorphic space. Differentiating with respect to α yields a system for (W_α, Q_α), namely

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

The terms on the right are mostly antiholomorphic and can be viewed as lower order when projected on the holomorphic functions. Examining the expression on the left one easily sees that the above first order system is degenerate, and has a double speed b. Then it is natural to diagonalize it. This is done using the operator

$$
\mathbf{A}(w,q) := (w, q - Rw), \qquad R := \frac{Q_\alpha}{1+W_\alpha}. \tag{1.2}
$$

The factor R above has an intrinsic meaning, namely it is the complex velocity on the water surface. We also remark that

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

Thus, the pair $(\mathbf{W}, R)$ diagonalizes the differentiated system. Indeed, a direct computation yields the self-contained system

$$\begin{cases} \mathbf{W}_t + b\mathbf{W}_\alpha + \dfrac{(1+\mathbf{W})R_\alpha}{1+\bar{\mathbf{W}}} = (1+\mathbf{W})M, \\ R_t + bR_\alpha = i\left(\dfrac{\mathbf{W}-a}{1+\mathbf{W}}\right), \end{cases} \tag{1.3}$$

where the real *frequency-shift* a is given by

$$a := i\left(\bar{P}\left[\bar{R}R_\alpha\right] - P\left[R\bar{R}_\alpha\right]\right), \tag{1.4}$$

and the auxiliary function M is given by

$$M := \frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}} - b_\alpha = \bar{P}[\bar{R}Y_\alpha - R_\alpha\bar{Y}] + P[R\bar{Y}_\alpha - \bar{R}_\alpha Y]. \tag{1.5}$$

The function Y above, given by

$$Y := \frac{\mathbf{W}}{1+\mathbf{W}},$$

is introduced in order to avoid rational expressions above and in many places in the sequel. The system (1.3) governs an evolution in the space of holomorphic functions, and will be used both directly and in its projected version.

Incidentally, we note that when expressed in terms of (Y, R) the water wave system becomes purely polynomial, see also [24],

$$\begin{cases} Y_t + bY_\alpha + |1-Y|^2 R_\alpha = (1-Y)M, \\ R_t + bR_\alpha - i(1+a)Y = -ia, \end{cases}$$

where M is as above, and

$$b = 2\Re(R - P(R\bar{Y})), \qquad a = 2\Re P(R\bar{R}_\alpha).$$

However, we do not take advantage of this formulation in the present article.

The functions b and a also play a fundamental role in the linearized equation which is computed in the next section, Section 2. The linearized variables are denoted by (w, q) and, after the diagonalization, $(w, r := q - Rw)$. The linearized equation, see (2.1), has the form

$$\begin{cases} (\partial_t + b\partial_\alpha)w + \dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w = \ (1+\mathbf{W})(P\bar{m} + \bar{P}m), \\ (\partial_t + b\partial_\alpha)r - i\dfrac{1+a}{1+\mathbf{W}}w = \ \bar{P}n - P\bar{n}, \end{cases} \tag{1.6}$$

where

$$m := \frac{r_\alpha + R_\alpha w}{J} + \frac{\bar{R}w_\alpha}{(1+\mathbf{W})^2}, \qquad n := \frac{\bar{R}(r_\alpha + R_\alpha w)}{1+\mathbf{W}}.$$

In particular, we remark that the linearization of the system (1.3) around the zero solution is

$$\begin{cases} w_t + r_\alpha = 0, \\ r_t - iw = 0. \end{cases} \tag{1.7}$$

The analysis of the linearized equation, carried out in Section 2, is a key component of this paper.

It is also useful to further differentiate (1.3), in order to obtain a system for $(\mathbf{W}_\alpha, R_\alpha)$:

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

In order to better compare this with the linearized system we introduce the modified variable $\mathbf{R} := R_\alpha(1+\mathbf{W})$ to get the system

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} = -b_\alpha \mathbf{W}_\alpha + \mathbf{R}\bar{Y}_\alpha + \mathbf{W}_\alpha M + (1+\mathbf{W})M_\alpha, \\ \mathbf{R}_t + b\mathbf{R}_\alpha = -\left(b_\alpha + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R} + i\left(\dfrac{(1+a)\mathbf{W}_\alpha}{1+\mathbf{W}} - a_\alpha\right) + \mathbf{R}M. \end{cases}$$

Expanding the b_α terms via (1.5) this yields

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}_\alpha = G_2, \\ \mathbf{R}_t + b\mathbf{R}_\alpha - i\dfrac{(1+a)\mathbf{W}_\alpha}{1+\mathbf{W}} = K_2, \end{cases} \tag{1.8}$$

where

$$\begin{cases} G_2 = \mathbf{R}\bar{Y}_\alpha - \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}_\alpha + 2M\mathbf{W}_\alpha + (1+\mathbf{W})M_\alpha, \\ K_2 = -2\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R} + 2M\mathbf{R} + (R_\alpha\bar{R}_\alpha - ia_\alpha). \end{cases}$$

Next, we define our function spaces. The system (1.7) is a well-posed linear evolution in the space $\dot{\mathcal{H}}_0$ of holomorphic functions endowed with the $L^2 \times \dot{H}^{\frac{1}{2}}$ norm. A conserved energy for this system is

$$E_0(w,r) = \int \frac{1}{2}|w|^2 + \frac{1}{2i}(r\bar{r}_\alpha - \bar{r}r_\alpha)d\alpha. \tag{1.9}$$

The nonlinear system (1.1) also admits a conserved energy, which has the form

$$E(W,Q) = \int \frac{1}{2}|W|^2 + \frac{1}{2i}(Q\bar{Q}_\alpha - \bar{Q}Q_\alpha) - \frac{1}{4}(\bar{W}^2 W_\alpha + W^2\bar{W}_\alpha)\, d\alpha. \tag{1.10}$$

As suggested by the above energy, our main function spaces for the differentiated water wave system (1.3) are the spaces $\dot{\mathcal{H}}_n$ endowed with the norm

$$\|(\mathbf{W},R)\|^2_{\dot{\mathcal{H}}_n} := \sum_{k=0}^{n} \|\partial_\alpha^k(\mathbf{W},R)\|^2_{L^2\times\dot{H}^{\frac{1}{2}}},$$

where $n \geq 1$. As an auxiliary step, we will also consider solutions $(\mathbf{W},R)$ in the smaller space

$$\mathcal{H}_n := H^n \times H^{n+\frac{1}{2}},$$

with $n \geq 2$.

To describe the lifespan of the solutions we define the control norms

$$A := \|\mathbf{W}\|_{L^\infty} + \|Y\|_{L^\infty} + \||D|^{\frac{1}{2}}R\|_{L^\infty\cap B^{0,\infty}_2}, \tag{1.11}$$

respectively

$$B := \| |D|^{\frac{1}{2}} \mathbf{W} \|_{BMO} + \| R_\alpha \|_{BMO}. \tag{1.12}$$

where $|D|$ represents the multiplier with symbol $|\xi|$. Here A is a scale invariant quantity, while B corresponds to the homogeneous $\dot{\mathcal{H}}_1$ norm of $(\mathbf{W}, R)$. We note that B and all but the Y component of A are controlled by the $\dot{\mathcal{H}}_1$ norm of the solution.

Now we are ready to state our main local well-posedness result:

Theorem 1. *Let $n \geq 1$. The system* (1.3) *is locally well-posed for data in $\dot{\mathcal{H}}_n(\mathbb{R})$ so that $|\mathbf{W} + 1| > c > 0$. Further, the solution can be continued for as long as A and B remain bounded. The same result holds in the periodic setting.*

In terms of Sobolev regularity of the data, this result improves the thresholds in earlier results of Wu [22, 23] and Alazard-Burq-Zuily [1]. However, a direct comparison is nontrivial due to the fact that the above two papers use different coordinate frames, namely Lagrangian, respectively Eulerian.

As an interesting side remark, the above result makes no requirement that the curve $\{Z(\alpha); \alpha \in \mathbb{R}\}$ determined by $\mathbf{W}$ be nonself-intersecting. If self-intersections occur then the physical interpretation is lost, but the well-posedness of the system (1.3) is not affected.

Our second goal in this article is to consider the question of obtaining improved lifespan bounds for the small data problem. Since the nonlinearities in our equations contain quadratic terms, the standard result is to obtain an $O(\epsilon^{-1})$ lifespan for smooth initial data of size ϵ. However, this problem has the additional feature that there exists a quadratic normal form transformation which eliminates the quadratic terms in the equation. In the setting of holomorphic coordinates considered in this paper, this is most readily seen at the level of the system (1.1). There, the quadratically nonlinear terms may be removed from the water-wave equations by the near-identity, normal form transformation

$$\tilde{W} = W - 2\mathfrak{M}_{\Re W} W_\alpha, \qquad \tilde{Q} = Q - 2\mathfrak{M}_{\Re W} R, \tag{1.13}$$

where the holomorphic multiplication operator $\mathfrak{M}_f$ is given by $\mathfrak{M}_f g = P[fg]$. For a more symmetric form of this transformation, one can replace R by Q_α. However, it is more convenient to use the diagonal variable R. For $(\tilde{W}, \tilde{Q})$ we have

Proposition 1.1. *The normal form variables (1.13) satisfy equations of the form*

$$\begin{cases} \tilde{W}_t + \tilde{Q}_\alpha = \tilde{G}, \\ \tilde{Q}_t - i\tilde{W} = \tilde{K}, \end{cases} \tag{1.14}$$

where $\tilde{G}$, $\tilde{K}$ are cubic (and higher order) functions of $(W, \mathbf{W}, R, \mathbf{W}_\alpha, R_\alpha)$, given by

$$\begin{cases} \tilde{G} = 2P[(F - R)_\alpha \Re W + \mathbf{W}_\alpha F \Re W + \mathbf{W} \Re(\mathbf{W} F) + F_\alpha \mathbf{W} \Re W] \\ \qquad\quad - P[\bar{\mathbf{W}} R \bar{Y} - \mathbf{W}(P[\bar{R} Y] + \bar{P}[R \bar{Y}])], \\ \tilde{K} = P\left[(\bar{F}(1 + \bar{\mathbf{W}}) - \bar{R}) R + 2iP\left[\dfrac{\mathbf{W}^2 + a}{1 + \mathbf{W}} \right] \cdot \Re W + 2P[bR_\alpha] \cdot \Re W \right]. \end{cases} \tag{1.15}$$

The proof is straightforward; one rewrites the system (1.1) in terms of the normal form variables $(\tilde{W}, \tilde{Q})$, (3.18). The original variables are (W, Q), but the derivatives of Q from the perturbative terms G and K are expressed in terms of R and eliminated. We also make use of the identity $P + \bar{P} = I$. The details are left for the reader. We note that the difference $R - F$ is quadratic,

$$R - F = P[R\bar{Y} - \bar{R}Y].$$

Heuristically, having cubic nonlinearities yields an improved $O(\epsilon^{-2})$ lifespan for initial data of size ϵ. However, implementing this idea directly is fraught with difficulties. To start with, while $\tilde{G}$, $\tilde{K}$ are cubic and higher order terms they also depend on higher-order derivatives of (W, Q); thus it is not possible to directly close energy estimates for the normal form variables $(\tilde{W}, \tilde{Q})$. This is related to the fact that the normal form transformation (1.13) is not invertible, and further to the fact that the system (1.1) is fully nonlinear, as opposed to semilinear.

There are at least two existing methods in the literature which attempt to address this difficulty. One such method, introduced by Wu [23], is based on the idea that any transformation which agrees quadratically with the above normal form transform will have the same effect as the normal form transform, but perhaps one can also choose such a transformation such that it is invertible. In Wu's work this transformation is an implicit change of coordinates, which is further followed by a secondary normal form transformation. A related example where an implicit change of coordinates is fully sufficient appears in the work [11] of the first two authors for the related Burgers-Hilbert problem.

A second method, which appears in the work of Shatah etc [18], is based on a mix of quadratic energy estimates for high derivatives of the solutions, combined with a normal form method for low derivatives. This works well for water waves in dimension three, but is not precise enough for the two dimensional problem.

In the present paper we propose an alternative approach for two dimensional water waves, which seems to be both simpler and more accurate. Precisely, rather than attempting to modify the equations using a normal form transform, we instead construct modified energy functionals which have cubic accuracy. A significant advantage of this idea is that it applies even for the leading order energy functionals, which to our knowledge is new. In a simpler setting, this method was first introduced by the authors in [10] in the context of the Burgers-Hilbert problem.

Our first result is translation invariant, and yields a cubic lifespan bound.

Theorem 2. *Let $\epsilon \ll 1$. Assume that the initial data for the equation* (1.3) *on either $\mathbb{R}$ or $\mathbb{S}^1$ satisfies*

$$\|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}_1} \leq \epsilon. \tag{1.16}$$

Then the solution exists on an ϵ^{-2} sized time interval $I_\epsilon = [0, T_\epsilon]$, and satisfies a similar bound. In addition, the estimates

$$\sup_{t \in I_\epsilon} \|(\mathbf{W}(t), R(t))\|_{\dot{\mathcal{H}}_n} \lesssim \|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}_n}, \qquad n \geq 2,$$

hold whenever the right hand side is finite.

Our second result assumes some additional localization for the initial data, and establishes almost global existence of solutions. This applies only for the problem on $\mathbb{R}$, and relies on the dispersive properties of the linear equation (1.7), whose solutions with localized data have $t^{-\frac{1}{2}}$ dispersive decay. To state the result we need to return to the original set of variables (W, Q).

We also take advantage of the scale invariance of the water wave equations. Precisely, it is invariant with respect to the scaling law

$$(W(t,\alpha), Q(t,\alpha)) \to (\lambda^{-2}W(\lambda t, \lambda^2\alpha), \lambda^{-3}Q(\lambda t, \lambda^2\alpha)).$$

This suggests that we should use the scaling vector field

$$S = t\partial_t + 2\alpha\partial_\alpha,$$

and its action on the pair (W, Q), namely

$$\mathbf{S}(W, Q) = ((S-2)W, (S-3)Q).$$

However, these are not the correct diagonal variables; to diagonalize we use the notations

$$(w, r) =: \mathbf{AS}(W, Q).$$

Then $(\mathbf{W}, R)$ solve the linearized equations 1.6 and define the weighted energy

$$\|(W,Q)(t)\|^2_{\mathcal{WH}} := \|(W,Q)(t)\|^2_{\dot{\mathcal{H}}_0} + \|(\mathbf{W},R)(t)\|^2_{\dot{\mathcal{H}}_5} + \|(w,r)(t)\|^2_{\dot{\mathcal{H}}_0}. \tag{1.17}$$

Then we have

Theorem 3. *There exists $c > 0$ so that for each initial data $(W(0), Q(0))$ for the system* (1.1) *satisfying*

$$\|(W,Q)(0)\|^2_{\mathcal{WH}} \leq \epsilon \ll 1, \tag{1.18}$$

the solution exists up to time $T_\epsilon = e^{c\epsilon^{-2}}$ and satisfies

$$\|(W,Q)(t)\|^2_{\mathcal{WH}} \lesssim \epsilon, \qquad |t| < T_\epsilon. \tag{1.19}$$

as well as

$$|W| + |W_\alpha| + ||D|^{\frac{1}{2}}W_\alpha| + |R| + |R_\alpha| \lesssim \frac{\epsilon}{\langle t\rangle^{\frac{1}{2}}}, \qquad |t| < T_\epsilon. \tag{1.20}$$

This lifespan bound was originally established by Wu [23]. Here, we prove the same result under less restrictive assumptions, and, hopefully, with a simpler proof. We should also mention here the recent work of Ionescu-Pusateri [12],[13] and Alazard-Delort [3], where global well-posedness is proved for small localized data. In a follow-up paper we provide a simplified proof of this result as well.

While our research for this paper was largely complete by the time [13] and [3] appeared, there is one idea from Ionescu and Pusateri's article [13] which we adopted here in order to shorten the exposition; this is the fact that in order to close the estimates it suffices to use a single iteration of the scaling vector field S. However, our implementation of this idea is different from [13], and also more efficient, in the sense that we use no higher derivatives of $\mathbf{S}(W, Q)$.

For the reminder of the introduction we provide a brief outline of the paper. The first step of the analysis is to study the linearization of the equation (1.1); this is done in Section 2. We begin with the diagonalisation of the linearized equations; this in turn leads to energy estimates, which are crucial in the proof of the local well-posedness result. The linearized energy functional is then refined so that cubically nonlinear estimates can be proved; this is essential in the proof of the improved lifespan result. We make no use of dispersive decay in this normal form analysis, so it works also for spatially periodic solutions. The low regularity threshold is reached by using various bilinear Coifman-Meyer type estimates, as well as multilinear versions thereof.

In Section 3 we consider the equations for higher order derivatives of the solution. The principal part of these equations is closely related to the linearized equations studied in the previous section. After some normalization, the quadratic bounds follow directly from the ones for the linearized equation. The emphasis there is again on obtaining cubically nonlinear estimates. The essential idea is to construct a modified energy functional with better estimates. Our modified energy essentially combines the linearized energy, for the leading part, with the cubic normal form energy for the lower order terms. This is similar to the approach in the paper [10] devoted to the Burgers-Hilbert problem.

Section 4 contains the proof of the local well-posedness result. We begin with more regular data, both in terms of low frequencies and in terms of high frequencies. For such data, a standard mollifier technique suffices in order to establish well-posedness. The rough $\dot{\mathcal{H}}_1$ solutions are obtained as uniform limits of smooth solutions by using the estimates for the linearized equation. The same construction yields their continuous dependence on data.

In Section 5 we prove the cubic lifespan bounds for small initial data in Theorem 1.

In Section 6 we provide the proof of the long time results. The cubic lifespan result is a straightforward consequence of the cubic energy estimates. The proof of the almost global result is slightly more involved, as it requires, as an intermediate step, to prove the $t^{-\frac{1}{2}}$ dispersive decay for a limited number of derivatives of $(\mathbf{W}, R)$. These bounds are obtained from the vector field energy estimates, essentially in an elliptic fashion via Sobolev type embeddings.

Appendix A includes, for reader's convenience, a complete derivation of the holomorphic water wave equations. Finally, Appendix B contains a collection of bilinear, multilinear and commutator estimates which are used at various places in the paper. We are grateful to Camil Muscalu for useful conversations pointing us in the right direction for this last section.

2. The linearized equation

In this section we derive the linearized water wave equations, and prove energy estimates for them. We do this in three stages. First we prove quadratic energy estimates in $\dot{\mathcal{H}}_0$, which apply for the large data problem. Then we prove cubic energy estimates in $\dot{\mathcal{H}}_0$ for the small data problem. Various bilinear, multilinear and commutator estimates which are used in this section are collected in Appendix B.

2.1. Computing the linearization.

The solutions for the linearized water wave equation around a solution (W, Q) are denoted by (w, q). However, it will be more convenient to immediately switch to diagonal variables (w, r), where

$$r := q - Rw.$$

The linearization of R is

$$\delta R = \frac{q_\alpha - R w_\alpha}{1+\mathbf{W}} = \frac{r_\alpha + R_\alpha w}{1+\mathbf{W}},$$

while the linearization of F can be expressed in the form

$$\delta F = P[m - \bar{m}],$$

where the auxiliary variable m corresponds to differentiating F with respect to the holomorphic variables,

$$m := \frac{q_\alpha - Rw_\alpha}{J} + \frac{\bar{R}w_\alpha}{(1+\mathbf{W})^2} = \frac{r_\alpha + R_\alpha w}{J} + \frac{\bar{R}w_\alpha}{(1+\mathbf{W})^2}.$$

Denoting also

$$n := \bar{R}\delta R = \frac{\bar{R}(r_\alpha + R_\alpha w)}{1+\mathbf{W}},$$

the linearized water wave equations take the form

$$\begin{cases} w_t + Fw_\alpha + (1+\mathbf{W})P[m - \bar{m}] = 0, \\ q_t + Fq_\alpha + Q_\alpha P[m-\bar{m}] - iw + P\left[n + \bar{n}\right] = 0. \end{cases}$$

Recalling that $b = F + \dfrac{\bar{R}}{1+\mathbf{W}}$, this becomes

$$\begin{cases} (\partial_t + b\partial_\alpha)w + (1+\mathbf{W})P[m-\bar{m}] = \dfrac{\bar{R}w_\alpha}{1+\mathbf{W}}, \\ (\partial_t + b\partial_\alpha)q + Q_\alpha P[m-\bar{m}] - iw + P\left[n+\bar{n}\right] = \dfrac{\bar{R}q_\alpha}{1+\mathbf{W}}. \end{cases}$$

Now, we can use the second equation in (1.3) to switch from q to r and obtain

$$\begin{cases} (\partial_t + b\partial_\alpha)w + (1+\mathbf{W})P[m-\bar{m}] = \dfrac{\bar{R}w_\alpha}{1+\mathbf{W}}, \\ (\partial_t + b\partial_\alpha)r - i\dfrac{1+a}{1+\mathbf{W}}w + P\left[n+\bar{n}\right] = \dfrac{\bar{R}(r_\alpha + R_\alpha w)}{1+\mathbf{W}}. \end{cases}$$

Terms like $\bar{P}m$, $\bar{P}n$ are lower order since the differentiated holomorphic variables have to be lower frequency. The same applies to their conjugates. Moving those terms to the right and taking advantage of algebraic cancellations we are left with

$$\begin{cases} (\partial_t + b\partial_\alpha)w + \dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w = \mathcal{G}(w,r), \\ (\partial_t + b\partial_\alpha)r - i\dfrac{1+a}{1+\mathbf{W}}w = \mathcal{K}(w,r), \end{cases} \tag{2.1}$$

where

$$\mathcal{G}(w,r) = \ (1+\mathbf{W})(P\bar{m} + \bar{P}m), \quad \mathcal{K}(w,r) = \ \bar{P}n - P\bar{n}.$$

We remark that while (w,r) are holomorphic, it is not directly obvious that the above evolution preserves the space of holomorphic states. To remedy this one can also project the linearized equations onto the space of holomorphic functions via the projection P. Then we obtain the equations

$$\begin{cases} (\partial_t + \mathfrak{M}_b\partial_\alpha)w + P\left[\dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = P\mathcal{G}(w,r), \\ (\partial_t + \mathfrak{M}_b\partial_\alpha)r - iP\left[\dfrac{1+a}{1+\mathbf{W}}w\right] = P\mathcal{K}(w,r). \end{cases} \tag{2.2}$$

Since the original set of equations (1.1) is fully holomorphic, it follows that the two sets of equations, (2.1) and (2.2), are algebraically equivalent.

In order to obtain cubic linearized energy estimates it is also of interest to separate the quadratic parts $\mathcal{G}^2$ and $\mathcal{K}^2$ of $\mathcal{G}$ and $\mathcal{K}$. These are split into quadratic and higher terms as shown below

$$\mathcal{G} = \mathcal{G}^{(2)} + \mathcal{G}^{(3+)}, \qquad \mathcal{K} = \mathcal{K}^{(2)} + \mathcal{K}^{(3+)}.$$

For the quadratic parts we have

$$P\mathcal{G}^{(2)}(w,r) = -P\left[\mathbf{W}\bar{r}_\alpha\right] + P\left[R\bar{w}_\alpha\right], \qquad P\mathcal{K}^{(2)}(w,r) = -P\left[R\bar{r}_\alpha\right],$$

with $\bar{P}\mathcal{G}^{(2)}(w,r) = \overline{P\mathcal{G}^{(2)}(w,r)}$ and $\bar{P}\mathcal{K}^{(2)}(w,r) = -\overline{P\mathcal{K}^{(2)}(w,r)}$. We can also rewrite the above expressions in a commutator form

$$P\mathcal{G}^{(2)}(w,r) = -\left[P,\mathbf{W}\right]\bar{r}_\alpha + \left[P,R\right]\bar{w}_\alpha, \quad P\mathcal{K}^{(2)}(w,r) = -\left[P,R\right]\bar{r}_\alpha. \tag{2.3}$$

The cubic terms have the form

$$\mathcal{G}^{(3+)}(w,r) = P\bar{m}^{(3+)} + \bar{P}m^{(3+)} + \mathbf{W}(P\bar{m} + \bar{P}m), \quad \mathcal{K}^{(3+)}(w,r) = \bar{P}n^{(3+)} - P\bar{n}^{(3+)}.$$

For the purpose of simplifying nonlinear estimates, it is convenient to express $\mathcal{G}^{(3)}$ and $\mathcal{K}^{(3)}$ in a polynomial fashion. This is done using the variable $Y = \dfrac{\mathbf{W}}{1+\mathbf{W}}$. Then we have

$$\begin{aligned}
\bar{P}m &= \bar{P}[w_\alpha(1-Y)^2\bar{R} - (r_\alpha + R_\alpha w)(1-Y)\bar{Y}],\\
\bar{P}m^{(3+)} &= \bar{P}[r_\alpha(\bar{\mathbf{W}} + Y)\bar{Y} - R_\alpha w(1-Y)\bar{Y} - w_\alpha(2Y - Y^2)\bar{R}],\\
\bar{P}n^{(3+)} &= \bar{P}[-r_\alpha Y\bar{R} + R_\alpha w(1-Y)\bar{R}].
\end{aligned}$$

2.2. **Quadratic estimates for large data.** Our goal here is to study the well-posedness of the system (2.2) in $L^2 \times \dot{H}^{\frac{1}{2}}$. We begin with a more general version of the system (2.2), namely

$$\begin{cases}
(\partial_t + \mathfrak{M}_b\partial_\alpha)w + P\left[\dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = G,\\
(\partial_t + \mathfrak{M}_b\partial_\alpha)r - iP\left[\dfrac{1+a}{1+\mathbf{W}}w\right] = K,
\end{cases} \tag{2.4}$$

and define the associated positive definite linear energy

$$E^{(2)}_{lin}(w,r) = \int_{\mathbf{R}} (1+a)|w|^2 + \Im(r\bar{r}_\alpha)d\alpha.$$

We remark that, by Proposition 2.6, a is nonnegative and bounded, therefore

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

Our first result uses the control parameters A and B defined in (1.11), (1.12):

Proposition 2.1. *a) The linear equation* (2.4) *is well-posed in* $\dot{\mathcal{H}}_0$*, and the following estimate holds:*

$$\frac{d}{dt}E^{(2)}_{lin}(w,r) = 2\Re\int_{\mathbf{R}} (1+a)\bar{w}\,G - i\bar{r}_\alpha\,K\ d\alpha + O_A(AB)E^{(2)}_{lin}(w,r). \tag{2.5}$$

b) The linearized equation (2.2) *is well-posed in* $L^2 \times \dot H^{\frac12}$*, and the following estimate holds:*

$$\frac{d}{dt}E^{(2)}_{lin}(w,r) \lesssim_A BE^{(2)}_{lin}(w,r). \tag{2.6}$$

Proof. a) A direct computation yields

$$\begin{aligned}\frac{d}{dt}\int(1+a)|w|^2 d\alpha = 2\Re\!\int (1+a)\bar w(\partial_t+\mathfrak{M}_b\partial_\alpha)w + a\bar w[b,P]w_\alpha\, d\alpha,\\ +\int [a_t + ((1+a)b)_\alpha]\,|w|^2\, d\alpha.\end{aligned}$$

A similar computation shows that

$$\frac{d}{dt}\int \Im(r\partial_\alpha \bar r)\, d\alpha = 2\Im\int(\partial_t+\mathfrak{M}_b\partial_\alpha)r\,\partial_\alpha\bar r\, d\alpha.$$

Adding the two and using the equations (2.4), the quadratic $\Re(w\bar r_\alpha)$ term cancels modulo another commutator term, and we obtain

$$\frac{d}{dt}E^{(2)}_{lin}(w,r) = 2\Re\int(1+a)\bar w\, G - i\bar r_\alpha\, K\, d\alpha + err_1, \tag{2.7}$$

where

$$\begin{aligned}err_1 = \int [a_t + ((1+a)b)_\alpha]\,|w|^2 d\alpha - 2\Re\int(1+a)\frac{R_\alpha}{1+\bar{\mathbf{W}}}|w|^2\, d\alpha\\ -2\Re\int +a\bar w\left(\left[\bar Y, P\right](r_\alpha + R_\alpha w) + [P,b]w_\alpha\right) d\alpha.\end{aligned}$$

Using the auxiliary function M in (1.5), we rewrite it as

$$err_1 = \int (a_t + ba_\alpha)\,|w|^2 + M(1+a)|w|^2\, d\alpha - 2\Re\int a\bar w\left(\left[\bar Y,P\right](r_\alpha+R_\alpha w) + [P,b]w_\alpha\right) d\alpha.$$

The error term is at least quartic. To conclude the proof of (2.5) it suffices to show that

$$|err_1| \lesssim ABE^{(2)}_{lin}(w,r). \tag{2.8}$$

For the first term, by Proposition 2.6 in the *Appendix* B, we have $|a_t + ba_\alpha| \lesssim AB$. For the second term we combine the pointwise bounds $|a| \lesssim A^2$ in Lemma 2.6 together with $\|M\|_{L^\infty} \lesssim AB$ in Lemma 2.8.

For the last term it remains to estimate the commutators in L^2. Two of them are obtained using Lemma 2.1,

$$\|\left[\bar Y,P\right] r_\alpha\|_{L^2} \lesssim \||D|^{\frac12}Y\|_{BMO}\|r\|_{\dot H^{\frac12}}, \qquad \|[P,b]w_\alpha\|_{L^2} \lesssim \|b_\alpha\|_{BMO}\|w\|_{L^2},$$

and suffice due to the bounds for b and Y in Lemmas 2.7,2.5. For the remaining piece we write $[\bar Y,P](R_\alpha w) = [\bar P, \bar P[\bar Y R_\alpha]]w]$ and use (B.7) to estimate

$$\|\bar P[\bar P[\bar Y R_\alpha]w]\|_{L^2} \lesssim \|w\|_{L^2}\|\bar P[\bar Y R_\alpha]\|_{BMO} \lesssim \|w\|_{L^2}\||D|^{\frac12}Y\|_{BMO}\||D|^{\frac12}R\|_{BMO},$$

where the bilinear bound in the second step follows after a bilinear Littlewood-Paley decomposition from (B.12) and (B.15).

b) To estimate the terms involving $\mathcal{G}$ and $\mathcal{K}$ we separate the quadratic and cubic parts. It suffices to show that the quadratic terms satisfy

$$\|\mathcal{G}^{(2)}(w,r)\|_{L^2} + \|\mathcal{K}^{(2)}(w,r)\|_{\dot H^{\frac12}} \lesssim_A B(\|w\|_{L^2} + \|r\|_{\dot H^{\frac12}}), \tag{2.9}$$

while the cubic and higher terms satisfy

$$\|\mathcal{G}^{(3+)}(w,r)\|_{L^2} + \|\mathcal{K}^{(3+)}(w,r)\|_{\dot H^{\frac12}} \lesssim_A AB(\|w\|_{L^2} + \|r\|_{\dot H^{\frac12}}). \tag{2.10}$$

In order to obtain the estimates claimed in (2.9),(2.10) we use the Coifman-Meyer [6] type commutator estimates described in the *Appendix* B, Lemma 2.1. Precisely, for the first term in $P\mathcal{G}^{(2)}(w,r)$ we use (B.10) with $s=0$, and $\sigma = \frac12$ to write

$$\| [P, \mathbf{W}]\, \bar r_\alpha\|_{L^2} \lesssim \||D|^{\frac12}\mathbf{W}\|_{BMO}\|r\|_{\dot H^{\frac12}}.$$

For the second term in $P\mathcal{G}^{(2)}(w,r)$ we use (B.10) with $s=0$ and $\sigma=1$ to obtain

$$\| [P, R]\, \bar w_\alpha\|_{L^2} \lesssim \|R_\alpha\|_{BMO}\|w\|_{L^2},$$

and for $P\mathcal{K}^{(2)}(w,r)$ we use (B.10) with $s=\frac12$, and $\sigma=\frac12$, and conclude that

$$\| [P, R]\, \bar r_\alpha\|_{\dot H^{\frac12}} \lesssim \|R_\alpha\|_{BMO}\|r\|_{\dot H^{\frac12}}.$$

The same estimate applies to the antiholomorphic parts of $\mathcal{G}^{(2)}$ and $\mathcal{K}^{(2)}$, and (2.3) follows.

For the cubic and higher parts of $\mathcal{G}$ and $\mathcal{K}$ we apply the same type of commutator estimates, as well as the BMO bounds in Proposition 2.2, as follows:

$$\|\bar P[r_\alpha(1-Y)\bar Y\bar{\mathbf{W}}]\|_{L^2} \lesssim \|r\|_{\dot H^{\frac12}}\|(1-Y)\bar Y\bar{\mathbf{W}}\|_{BMO^{\frac12}} \lesssim_A \|Y\|_{L^\infty}\|r\|_{\dot H^{\frac12}},$$

using (B.16) at the last step.

$$\|\bar P[w(1-Y)R_\alpha\bar Y]\|_{L^2} \lesssim \|w(1-Y)\|_{L^2}\|\bar P[R_\alpha\bar Y]\|_{BMO} \lesssim_A \|w\|_{L^2}\|R\|_{BMO^{\frac12}}\|Y\|_{BMO^{\frac12}}$$

using (B.12) and (B.15) at the last step.

$$\|\bar P[w_\alpha(2Y-Y^2)\bar R]\|_{L^2} \lesssim \|w\|_{L^2}\|\partial_\alpha\bar P[(2Y-Y^2)\bar R]\|_{BMO}, \lesssim_A \|w\|_{L^2}\|Y\|_{L^\infty}\|R\|_{BMO}$$

using (B.12), and (B.14) at the last step.

$$\||D|^{\frac12}\bar P[r_\alpha Y\bar R]\|_{L^2} \lesssim \|r\|_{\dot H^{\frac12}}\|\partial_\alpha\bar P[Y\bar R]\|_{L^2} \lesssim_A \|r\|_{\dot H^{\frac12}}\|Y\|_{L^\infty}\|R\|_{BMO},$$

again by (B.12) and (B.14). Finally,

$$\begin{aligned}\||D|^{\frac12}\bar P[w(1-Y)R_\alpha\bar R]\|_{L^2} &\lesssim \|w(1-Y)\|_{L^2}\||D|^{\frac12}\bar P[R_\alpha\bar R]\|_{BMO}\\ &\lesssim_A \|w\|_{L^2}\||D|^{\frac12}R\|_{BMO}\|R_\alpha\|_{BMO}\end{aligned}$$

follows using (B.12) and (B.15).

□

2.3. **Cubic estimates for small data.** For the small data problem it is of further interest to track the solution on larger time scales. For this we add to the equations the holomorphic quadratic parts $P\mathcal{G}^{(2)}$ and $P\mathcal{K}^{(2)}$ of $\mathcal{G}$ and $\mathcal{K}$ and consider the linear equations

$$\begin{cases} (\partial_t + \mathfrak{M}_b\partial_\alpha)w + P\left[\dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = -P\left[\mathbf{W}\bar r_\alpha - R\bar w_\alpha\right] + G,\\ (\partial_t + \mathfrak{M}_b\partial_\alpha)r - iP\left[\dfrac{1+a}{1+\mathbf{W}}w\right] = -P\left[R\bar r_\alpha\right] + K. \end{cases} \tag{2.11}$$

For this problem we add appropriate cubic terms and define the modified energy

$$E^{(3)}_{lin}(w,r) = \int_{\mathbf{R}} (1+a)|w|^2 + \Im(r\bar r_\alpha) + 2\Im(\bar R w r_\alpha) - 2\Re(\bar{\mathbf{W}}w^2)\, d\alpha.$$

Then we have:

Proposition 2.2. *Assume that $A \ll 1$. Then*

$$E^{(3)}_{lin}(w,r) = (1+O(A))E_0(w,r). \tag{2.12}$$

In addition, the following properties hold:

a) The solutions to (2.11) *satisfy*

$$\frac{d}{dt}E^{(3)}_{lin}(w,r) = 2\Re \int \left((1+a)\bar{w} - i\bar{R}r_\alpha - 2\bar{\mathbf{W}}w\right) G + i(\bar{r} - \bar{R}w)\, K_\alpha \, d\alpha \\ + O_A(AB)E^{(3)}_{lin}(w,r). \tag{2.13}$$

b) For solutions to the linearized equation (2.2) *we have:*

$$\frac{d}{dt}E^{(3)}_{lin}(w,r) \lesssim_A ABE^{(3)}_{lin}(w,r). \tag{2.14}$$

Proof. For (2.12) we need to estimate the added cubic terms in $E^{(3)}_{lin}(w,r)$. The second is trivially bounded, while the first is rewritten as

$$\Im \int w\bar{P}[\bar{R}r_\alpha]\, d\alpha.$$

By Lemma 2.1 we have $\|P[\bar{R}r_\alpha]\|_{L^2} \lesssim \||D|^{\frac{1}{2}}R\|_{BMO}\|r\|_{\dot{H}^{\frac{1}{2}}}$, hence (2.12) follows.

a) To prove the estimate (2.13) we compute the time derivative of the cubic component of the energy $E^3_{lin}(w,r)$, using the projected equations for w and r and the unprojected equations for R and $\mathbf{W}$:

$$\frac{d}{dt}\left(\Im\int \bar{R}wr_\alpha\, d\alpha - \Re\int \bar{\mathbf{W}}w^2 d\alpha\right) = \Im\int -i\bar{\mathbf{W}}wr_\alpha - \bar{R}r_\alpha r_\alpha + i\bar{R}ww_\alpha + \bar{R}r_\alpha G + \bar{R}wK_\alpha\, d\alpha \\ + \Re\int \bar{R}_\alpha w^2 + 2\bar{\mathbf{W}}wr_\alpha + 2\bar{\mathbf{W}}wF\, d\alpha + err_2,$$

where

$$\begin{aligned} err_2 = \Im\int\Bigg\{&\left(i\left(\frac{\bar{\mathbf{W}}^2+a}{1+\bar{\mathbf{W}}}\right) - b\bar{R}_\alpha\right)wr_\alpha - \bar{R}w\partial_\alpha\left(\mathfrak{M}_b r_\alpha - iP\left[\frac{a-\mathbf{W}}{1+\mathbf{W}}w\right] + P[R\bar{r}_\alpha]\right) \\ &-\bar{R}r_\alpha\left(\mathfrak{M}_b w_\alpha - P\left[\frac{\bar{\mathbf{W}}}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] + P[\mathbf{W}\bar{r}_\alpha - R\bar{w}_\alpha]\right)\Bigg\}\, d\alpha \\ + \Re\int\Bigg\{&w^2\left(b\bar{\mathbf{W}}_\alpha + \frac{\bar{\mathbf{W}}-\mathbf{W}}{1+\mathbf{W}}\bar{R}_\alpha - (1+\bar{\mathbf{W}})\bar{M}\right) \\ &+2\bar{\mathbf{W}}w\left(\mathfrak{M}_b w_\alpha - P\left[\frac{\bar{\mathbf{W}}}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] + P[\mathbf{W}\bar{r}_\alpha - R\bar{w}_\alpha]\right)\Bigg\}\, d\alpha. \end{aligned} \tag{2.15}$$

Adding this to (2.5) (but applied to solutions to (2.11)) we obtain

$$\frac{d}{dt}E^{(3)}_{lin}(w,r) = 2\Re\int\left((1+a)\bar{w} - i\bar{R}r_\alpha - 2\bar{\mathbf{W}}w\right) G + i(\bar{r}-\bar{R}w)\, K_\alpha\, d\alpha + err_1 + err_3, \tag{2.16}$$

where

$$err_3 = 2err_2 - 2\Re\int a\bar{w}P\left[\mathbf{W}\bar{r}_\alpha - R\bar{w}_\alpha\right]\, d\alpha.$$

Given the bound (2.8) for err_1, the proof of (2.13) is concluded if we show that

$$|err_3| \lesssim ABE_0(w,r). \tag{2.17}$$

Further, recalling the estimate (2.9), which in expanded form reads

$$\|P\left[\mathbf{W}\bar{r}_\alpha - R\bar{w}_\alpha\right]\|_{L^2} + \|P[R\bar{r}_\alpha]\|_{\dot{H}^{\frac{1}{2}}} \lesssim B\|(w,r)\|_{L^2\times \dot{H}^{\frac{1}{2}}}, \tag{2.18}$$

it suffices to estimate err_2,

$$|err_2| \lesssim ABE_0(w,r). \tag{2.19}$$

For the remainder of the proof we separately estimate several types of terms in err_2:

A. Terms involving b. Here, we use the bounds for b in Lemma 2.7, which give

$$\|b_\alpha\|_{BMO} \lesssim_A B, \qquad \||D|^{\frac{1}{2}}b\|_{BMO} \lesssim_A A.$$

We first collect all the terms that are contained in the first integral in err_2 and include b,

$$I_1 = \int -b\bar{R}_\alpha w r_\alpha - \bar{R}r_\alpha \mathfrak{M}_b w_\alpha - \bar{R}w\partial_\alpha(\mathfrak{M}_b r_\alpha)\, d\alpha.$$

We claim that

$$|I_1| \lesssim (\||D|^{\frac{1}{2}}R\|_{BMO}\|b_\alpha\|_{BMO} + \|R_\alpha\|_{BMO}\||D|^{\frac{1}{2}}b\|_{BMO})\|w\|_{L^2}\|r\|_{\dot{H}^{\frac{1}{2}}}. \tag{2.20}$$

Integrating by parts we get $I_1 = I_2 + I_3$, where

$$I_2 = \int \bar{R}_\alpha w \bar{P}[br_\alpha]\, d\alpha, \qquad I_3 = \int -\bar{R}r_\alpha \mathfrak{M}_b w_\alpha - \bar{R}w\partial_\alpha(\mathfrak{M}_b r_\alpha)\, d\alpha.$$

The first term on the right has a commutator structure and will be estimated separately later, see I_5 below. The bound for I_3 is proved in the appendix, see (B.35).

We next collect all the terms that are contained in the second integral in err_2 and include b, and rewrite them as

$$I_4 = \int w^2\partial_\alpha \overline{\mathfrak{M}_b \mathbf{W}} + 2\bar{\mathbf{W}}w\mathfrak{M}_b w_\alpha\, d\alpha = \int -2ww_\alpha b\bar{\mathbf{W}} + 2\bar{\mathbf{W}}w\mathfrak{M}_b w_\alpha\, d\alpha = \int -2\bar{\mathbf{W}}w\bar{P}[bw_\alpha]\, d\alpha.$$

The expression $\bar{P}[bw_\alpha]$ is bounded in L^2 using Lemma 2.1 to obtain

$$|I_4| \lesssim \|\mathbf{W}\|_{L^\infty}\|b_\alpha\|_{BMO}\|w\|_{L^2}^2.$$

B. Quadrilinear terms bounded via both $L^2 \cdot L^2$ and $\dot{H}^{\frac{1}{2}} \cdot \dot{H}^{-\frac{1}{2}}$ pairings. This includes the following expressions:

$$I_5 = \int \bar{R}_\alpha w\bar{P}[br_\alpha]\, d\alpha = \int \bar{P}[br_\alpha]P[\bar{R}_\alpha w]\, d\alpha,$$

$$I_6 = \int \bar{R}r_\alpha P\left[\frac{\bar{\mathbf{W}}}{1+\bar{\mathbf{W}}}r_\alpha\right]\, d\alpha = \int \bar{P}[\bar{R}r_\alpha]P\left[\bar{Y}r_\alpha\right]\, d\alpha,$$

$$I_7 = \int \bar{R}r_\alpha P\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}w\right]\, d\alpha = \int \bar{P}[\bar{R}r_\alpha]P\left[R_\alpha w(1-\bar{Y})\right]\, d\alpha,$$

$$I_8 = \int \bar{\mathbf{W}}wP\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}w\right]\, d\alpha = \int \bar{P}[\bar{\mathbf{W}}w]P\left[R_\alpha(1-\bar{Y})w\right]\, d\alpha,$$

$$I_9 = \int \bar{\mathbf{W}}wP\left[\frac{\bar{\mathbf{W}}}{1+\bar{\mathbf{W}}}r_\alpha\right]\, d\alpha = \int \bar{P}[\bar{\mathbf{W}}w]P\left[\bar{Y}r_\alpha\right]\, d\alpha.$$

The strategy here is to bound the first factor in both L^2 and $\dot H^{\frac12}$, and the second, partially in L^2 and partially in $\dot H^{-\frac12}$. For the first factor we have by Lemma 2.1:

$$\|\bar P[br_\alpha]\|_{L^2} + \|\bar P[\bar R r_\alpha]\|_{L^2} \lesssim (\||D|^{\frac12} b\|_{BMO} + \||D|^{\frac12} R\|_{BMO})\|r\|_{\dot H^{\frac12}} \lesssim A\|r\|_{\dot H^{\frac12}},$$
$$\|\bar P[br_\alpha]\|_{\dot H^{\frac12}} + \|\bar P[\bar R r_\alpha]\|_{\dot H^{\frac12}} \lesssim (\|b_\alpha\|_{BMO} + \|R_\alpha\|_{BMO})\|r\|_{\dot H^{\frac12}} \lesssim B\|r\|_{\dot H^{\frac12}},$$

as well as

$$\|\bar P[\bar{\mathbf W} w]\|_{L^2} + \||D|^{\frac12} \bar P[\bar R w]\|_{L^2} \lesssim (\|\mathbf W\|_{BMO} + \||D|^{\frac12} R\|_{BMO})\|w\|_{L^2} \lesssim A\|w\|_{L^2},$$
$$\|\bar P[\bar{\mathbf W} w]\|_{\dot H^{\frac12}} + \||D|^{\frac12} \bar P[\bar R w]\|_{\dot H^{\frac12}} \lesssim (\||D|^{\frac12} \mathbf W\|_{BMO} + \|R_\alpha\|_{BMO})\|w\|_{L^2} \lesssim B\|w\|_{L^2}.$$

We now consider the second factor in the above integrals. For $P[\bar R_\alpha w]$ we have

$$\|\sum_k P[\bar R_{k,\alpha} w_k]\|_{L^2} \lesssim \|R_\alpha\|_{BMO}\|w\|_{L^2}, \quad \|\sum_k P[\bar R_{<k,\alpha} w_k]\|_{\dot H^{-\frac12}} \lesssim \||D|^{\frac12} R\|_{BMO}\|w\|_{L^2}.$$

The same argument applies to $P\left[R_\alpha(1-\bar Y)w\right]$ once we use the decomposition

$$P\left[(1-\bar Y)R_\alpha w\right] = P\left[(1-\bar Y)\sum_{k\in\mathbf Z} R_{\alpha,\geq k} w_k\right] - P\left[\sum_{k\in\mathbf Z} \bar Y_k R_{\alpha,<k} w_k\right] + \sum_{k\in\mathbf Z}(1-\bar Y)_{<k} R_{\alpha,<k} w_k.$$

The first term is easily bounded in L^2 by Lemma 2.1. The second is also in L^2 using (B.14) for the product of the first two factors. Finally, the third is bounded in $\dot H^{-\frac12}$ by estimating $\|R_{\alpha,<k}\|_{L^\infty} \lesssim 2^{\frac k2} A$.

It remains to consider the expression

$$P\left[\bar Y r_\alpha\right] = P\left[\sum_k \bar Y_k r_{k,\alpha}\right] + \sum_k \bar Y_{<k} r_{k,\alpha}.$$

Here, the first term is estimated in L^2 using Lemma 2.1, while the second goes into $\dot H^{-\frac12}$.

C. Quadrilinear terms bounded via an $L^2\cdot L^2$ pairing. This includes the following expressions:

$$I_9 = \int \bar R w \partial_\alpha P\left[\frac{a-\mathbf W}{1+\mathbf W} w\right] d\alpha = -\int \partial_\alpha \bar P[\bar R w] P\left[(a(1-Y)-Y)w\right] d\alpha,$$
$$I_{10} = \int \bar R w \partial_\alpha P[R\bar r_\alpha]\, d\alpha = -\int \partial_\alpha \bar P[\bar R w]\partial_\alpha P[R\bar r_\alpha]\, d\alpha,$$
$$I_{11} = \int \bar R r_\alpha P[\mathbf W \bar r_\alpha - R\bar w_\alpha]\, d\alpha = \int \bar P[\bar R r_\alpha] P[\mathbf W\bar r_\alpha - R\bar w_\alpha] d\alpha,$$
$$I_{12} = \int \bar{\mathbf W} w P[\mathbf W \bar r_\alpha - R\bar w_\alpha]\, d\alpha = \int \bar P[\bar{\mathbf W} w] P[W\bar r_\alpha - R\bar w_\alpha]\, d\alpha.$$

In all cases both factors are estimated directly in L^2, using Lemma 2.1, see also (2.18).

D. Trilinear estimates. This includes the terms:

$$I_{13} = \int \frac{\bar{\mathbf{W}}^2 + a}{1+\bar{\mathbf{W}}} w r_\alpha \, d\alpha = \int w \bar{P}[\bar{P} f r_\alpha] \, d\alpha, \qquad f = \frac{\bar{\mathbf{W}}^2 + a}{1+\bar{\mathbf{W}}},$$

$$I_{14} = \int w^2 \bar{P} \left[\frac{\bar{\mathbf{W}} - \mathbf{W}}{1+\mathbf{W}} \bar{R}_\alpha \right] d\alpha = \int w \bar{P}[\bar{P} g w] \, d\alpha, \qquad g = \frac{\bar{\mathbf{W}} - \mathbf{W}}{1+\mathbf{W}} \bar{R}_\alpha,$$

$$I_{15} = \int w^2 \bar{P} \left[(1+\bar{\mathbf{W}}) M \right] d\alpha = \int w \bar{P}[\bar{P} h w] \, d\alpha, \qquad h = (1+\bar{\mathbf{W}}) M.$$

Using Lemma 2.1 we have

(2.21)

$$|I_{13}| \lesssim \| |D|^{\frac{1}{2}} \bar{P} f \|_{BMO} \|w\|_{L^2} \|r\|_{\dot{H}^{\frac{1}{2}}}, \ |I_{14}| \lesssim \|\bar{P} g\|_{BMO} \|w\|^2_{L^2}, \ |I_{15}| \lesssim \|\bar{P} h\|_{BMO} \|w\|^2_{L^2},$$

so it suffices to show that

$$\| |D|^{\frac{1}{2}} f \|_{BMO} + \|g\|_{BMO} + \|h\|_{BMO} \lesssim AB.$$

The f bound follows from the algebra property of $BMO^{\frac{1}{2}} \cap L^\infty$ in (B.16) in view of (B.24) and (B.25). The g bound is obtained by writing

$$\frac{\bar{\mathbf{W}} - \mathbf{W}}{1+\mathbf{W}} \bar{R}_\alpha = \sum_k P_{\leq k} \left(\frac{\bar{\mathbf{W}} - \mathbf{W}}{1+\mathbf{W}} \right) R_{k,\alpha} + \sum_k P_k \left(\frac{\bar{\mathbf{W}} - \mathbf{W}}{1+\mathbf{W}} \right) R_{<k,\alpha}.$$

For the first term we use (B.14), while for the second, (B.15). Finally, the h bound is trivial due to (B.32). The proof of (2.13) is concluded.

b) To prove the bound (2.14) it suffices to apply the estimate in (2.13) with

$$F = P\mathcal{F}^{3+}(w,r), \qquad G = P\mathcal{G}^{3+}(w,r).$$

Given the estimate (2.10) for the cubic components of $\mathcal{F}$ and $\mathcal{G}$ and the pointwise bound (B.24) for a, it remains to consider the terms

$$\int \bar{R} r_\alpha P\mathcal{F}^{(3+)} \, d\alpha, \qquad \int \bar{\mathbf{W}} w P\mathcal{F}^{(3+)} \, d\alpha, \qquad \int \bar{R} w) \, P\mathcal{G}^{(3+)} \, d\alpha.$$

For the first one we use the second part of (2.21) to get

$$\left| \int \bar{R} r_\alpha P\mathcal{F}^{(3+)} \, d\alpha \right| \lesssim \| |D|^{\frac{1}{2}} R \|_{L^\infty} \|r\|_{\dot{H}^{\frac{1}{2}}} \|\mathcal{F}^{(3+)}\|_{L^2} \lesssim AB \|(w,r)\|^2_{\dot{\mathcal{H}}_0} \tag{2.22}$$

The second one is directly estimated as

$$\left| \int \bar{R} w \partial_\alpha \mathcal{K} \, d\alpha \right| \lesssim \| |D|^{\frac{1}{2}} R \|_{L^\infty} \|w\|_{L^2} \|\mathcal{K}\|_{\dot{H}^{\frac{1}{2}}} \lesssim AB \|(w,r)\|^2_{\dot{\mathcal{H}}_0} \tag{2.23}$$

On the last term, using the first part of (2.21), we get

$$\left| \int \bar{R} w \partial_\alpha \mathcal{K} \, d\alpha \right| \lesssim \| |D|^{\frac{1}{2}} R \|_{L^\infty} \|w\|_{L^2} \|\mathcal{K}\|_{\dot{H}^{\frac{1}{2}}} \lesssim AB \|(w,r)\|^2_{\dot{\mathcal{H}}_0} \tag{2.24}$$

The proof of the proposition is concluded.

□

3. Higher order energy estimates

The main goal of this section is to establish two energy bounds for $(\mathbf{W}, R)$ and their higher derivatives. The first one is a quadratic bound which applies for all solutions. The second one is a cubic bound which only applies for small solutions. The large data result is as follows:

Proposition 3.1. *For any $n \geq 1$ there exists an energy functional $E^{n,(2)}$ with the following properties: (i) Norm equivalence:*

$$E^{n,(2)}(\mathbf{W}, R) \approx_A E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R),$$

(ii) Quadratic energy estimates for solutions to (1.3)*:*

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

The small data result is as follows:

Proposition 3.2. *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:

$$E^{n,(3)}(\mathbf{W}, R) = (1 + O(A))E_0(\partial^{n-1}\mathbf{W}, \partial^{n-1}R),$$

(ii) Cubic energy estimates:

$$\frac{d}{dt}E^{n,(3)}(\mathbf{W}, R) \lesssim_A ABE^{n,(3)}(\mathbf{W}, R).$$

We remark that the case $n = 1$ corresponds to bounds for $(\mathbf{W}, R)$. But these solve the linearized system (2.2), so the above results are consequences of Proposition 2.1 and Proposition 2.2. In the sequel we consider separately the cases $n = 2$ and $n \geq 3$.

3.1. **The case $n = 2$.** We use the system (1.8) for $(\mathbf{W}_\alpha, \mathbf{R} := R_\alpha(1 + \mathbf{W}))$, which for convenience we recall here:

$$\begin{cases} \mathbf{W}_{\alpha t} + b\mathbf{W}_{\alpha\alpha} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}_\alpha = \mathbf{R}\bar{Y}_\alpha - \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}_\alpha + 2M\mathbf{W}_\alpha + (1+\mathbf{W})M_\alpha, \\ \mathbf{R}_t + b\mathbf{R}_\alpha - i\dfrac{(1+a)\mathbf{W}_\alpha}{1+\mathbf{W}} = -2\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R} + 2M\mathbf{R} + (R_\alpha\bar{R}_\alpha - ia_\alpha). \end{cases}$$

Here we have isolated on the left the leading part of the linearized equation. We want to interpret the terms on the right as mostly perturbative, but also pay attention to the quadratic part. Thus, for bookkeeping purposes, we introduce two types of error terms, denoted $\mathbf{err}(L^2)$ and $\mathbf{err}(\dot{H}^{\frac{1}{2}})$, which correspond to the two equations. The bounds for these errors are in terms of the control variables A, B, as well as the L^2 type norm

$$\mathbf{N}_2 = \|(\mathbf{W}_\alpha, R_\alpha)\|_{L^2\times\dot{H}^{\frac{1}{2}}}.$$

By $\mathbf{err}(L^2)$ we denote terms G, which satisfy the estimates

$$\|PG\|_{L^2} \lesssim_A AB\mathbf{N}_2,$$

and

$$\text{either} \quad \|\bar{P}G\|_{L^2} \lesssim_A B\mathbf{N}_2 \quad \text{or} \quad \|\bar{P}G\|_{\dot{H}^{-\frac{1}{2}}} \lesssim_A A\mathbf{N}_2.$$

By $\mathbf{err}(\dot H^{\frac12})$ we denote terms K, which are at least cubic and which satisfy the estimates

$$\|PK\|_{\dot H^{\frac12}} \lesssim_A AB\mathbf{N}_2, \qquad \|PK\|_{L^2} \lesssim_A A^2\mathbf{N}_2,$$

and

$$\|\bar P K\|_{L^2} \lesssim_A A\mathbf{N}_2.$$

The use of the more relaxed quadratic control on the antiholomorphic terms, as opposed to the cubic control on the holomorphic terms, is motivated by the fact that the equations will eventually get projected on the holomorphic space, so the antiholomorphic components will have less of an impact. A key property of the space of errors is contained in the following

Lemma 3.3. *Let Φ be a function which satisfies*

$$\|\Phi\|_{L^\infty} \lesssim A, \qquad \||D|^{\frac12}\Phi\|_{BMO} \lesssim B. \tag{3.1}$$

Then, we have the multiplicative bounds

$$\Phi\cdot \mathbf{err}(L^2) = \mathbf{err}(L^2), \qquad \Phi\cdot \mathbf{err}(\dot H^{\frac12}) = \mathbf{err}(\dot H^{\frac12}), \tag{3.2}$$

$$\Phi\cdot P\mathbf{err}(L^2) = A\,\mathbf{err}(L^2), \qquad \Phi\cdot P\mathbf{err}(\dot H^{\frac12}) = A\,\mathbf{err}(\dot H^{\frac12}). \tag{3.3}$$

The proof of the lemma, based on Lemma 2.1, is relatively straightforward and is left for the reader. We will apply this lemma for Φ which are arbitrary smooth functions of $\mathbf{W}$ and $\bar{\mathbf{W}}$. Then the estimates (3.1) are consequences of our Moser estimates in (B.17).

We now expand some of the terms in the above system. For this we will use the following bounds for M, see (B.32) and (B.33):

$$\|M\|_{L^\infty} \lesssim AB, \qquad \|M\|_{\dot H^{\frac12}} \lesssim A\mathbf{N}_2. \tag{3.4}$$

First we note that

$$M\mathbf{W}_\alpha = \mathbf{err}(L^2), \qquad M\mathbf{R} = \mathbf{err}(\dot H^{\frac12}). \tag{3.5}$$

The first is straightforward in view of pointwise bound for M. For the second, by Lemma 3.3 we can replace $M\mathbf{R}$ by MR_α. After a Littlewood-Paley decomposition, the $\dot H^{\frac12}$ estimate for MR_α is a consequence of the pointwise bound in (3.4) for low-high and balanced interactions, and of the $\dot H^{\frac12}$ bound in (3.4) combined with Lemma 2.1 for the high-low interactions.

It remains to estimate MR_α in L^2. If the frequency of M is larger than or equal to the frequency of R_α, then we can use the $\dot H^{\frac12}$ bound for M. We are left with

$$\sum_k R_{k,\alpha}M_{<k} = \sum_k R_{k,\alpha}M(R_{<k},Y_{<k}) + \sum_k\sum_{j\geq k} R_{k,\alpha}P_{<k}M(R_j,Y_j).$$

For the first sum we use

$$\|M(R_{<k},Y_{<k})\|_{L^\infty} \lesssim 2^{\frac{k}{2}}A^2.$$

For the second we bound

$$\|\sum_k\sum_{j\geq k} R_{k,\alpha}P_{<k}M(R_j,Y_j)\|^2_{L^2} \lesssim \sum_{j\geq k} 2^{k-j}\||D|^{\frac12}R\|^4_{L^\infty}\|Y_{j,\alpha}\|^2 \lesssim A^2\mathbf{N}_2.$$

Next we consider $(1+\mathbf{W})M_\alpha$, for which we claim that

$$\begin{aligned} &M_\alpha = R_\alpha\bar Y_\alpha - \bar R_\alpha Y_\alpha + P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar R_{\alpha\alpha}\mathbf{W}] + \mathbf{err}(L^2),\\ &P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar R_{\alpha\alpha}\mathbf{W}] = A^{-1}\mathbf{err}(L^2). \end{aligned} \tag{3.6}$$

By Lemma 3.3, this shows that

$$(1+\mathbf{W})M_\alpha = R\bar{Y}_\alpha - \frac{\bar{R}_\alpha}{1+\mathbf{W}}W_\alpha + P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}\mathbf{W}] + \mathbf{err}(L^2).$$

To prove (3.6) we write

$$\begin{aligned} M_\alpha &= R_\alpha\bar{Y}_\alpha - \bar{R}_\alpha Y_\alpha + P[R\bar{Y}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}Y] + \bar{P}(g_1+2g_2) \\ &= R_\alpha\bar{Y}_\alpha - \bar{R}_\alpha Y_\alpha + P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}\mathbf{W}] - Pf + \bar{P}(g_1+2g_2), \end{aligned}$$

where

$$f = R(\bar{\mathbf{W}}\bar{Y})_{\alpha\alpha} - \bar{R}_{\alpha\alpha}(\mathbf{W}Y), \quad g_1 = \bar{R}Y_{\alpha\alpha} - R_{\alpha\alpha}\bar{Y}, \quad g_2 = \bar{R}_\alpha Y_\alpha - R_\alpha\bar{Y}_\alpha.$$

For f and g_1 we have L^2 bounds

$$\|Pf\|_{L^2} \lesssim AB\mathbf{N}_2, \qquad \|\bar{P}g_1\|_{L^2} \lesssim B\mathbf{N}_2,$$

which follow from commutator type bounds

$$\|P[R\bar{\Phi}_{\alpha\alpha}]\|_{L^2} \lesssim \|R_\alpha\|_{BMO}\|\Phi_\alpha\|_{L^2}, \quad \|P[\bar{R}_{\alpha\alpha}\Phi]\|_{L^2} \lesssim \|R_\alpha\|_{\dot{H}^{\frac12}}\||D|^{\frac12}\Phi\|_{BMO}, \tag{3.7}$$

derived from Lemma 2.1. For the first term in g_2 we have a similar L^2 bound, but for the second we split

$$\bar{P}[R_\alpha\bar{Y}_\alpha] = \bar{P}[\sum_k R_{k,\alpha}\bar{Y}_{k,\alpha}] + \sum_k R_{<k,\alpha}\bar{Y}_{k,\alpha}.$$

The first sum is bounded in L^2 using Lemma 2.1, but for the second we only get a $\dot{H}^{-\frac12}$ bound,

$$\|\sum_k R_{<k,\alpha}\bar{Y}_{k,\alpha}\|_{\dot{H}^{-\frac12}} \lesssim A\mathbf{N}_2.$$

Finally, we also claim that

$$ia_\alpha = R_\alpha\bar{R}_\alpha + P[R\bar{R}_{\alpha\alpha}] + \mathbf{err}(\dot{H}^{\frac12}), \qquad P[R\bar{R}_{\alpha\alpha}] = A^{-1}\mathbf{err}(\dot{H}^{\frac12}),$$

which is again a consequence of commutator type estimates for holomorphic V:

$$\|P[R\bar{V}_\alpha]\|_{\dot{H}^{\frac12}} \lesssim \|R_\alpha\|_{BMO}\|V\|_{\dot{H}^{\frac12}}, \qquad \|P[R\bar{V}_\alpha]\|_{L^2} \lesssim \||D|^{\frac12}R\|_{L^\infty}\|V\|_{\dot{H}^{\frac12}}. \tag{3.8}$$

Taking into account all of the above expansions, it follows that our system can be rewritten in the form

$$\begin{cases} (\partial_t + b\partial_\alpha)\mathbf{W}_\alpha + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha\mathbf{W}_\alpha}{1+\bar{\mathbf{W}}} = 2\mathbf{R}\bar{Y}_\alpha - \dfrac{2\bar{R}_\alpha\mathbf{W}_\alpha}{1+\mathbf{W}} + P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}\mathbf{W}] + \mathbf{err}(L^2), \\ (\partial_t + b\partial_\alpha)\mathbf{R} - i\dfrac{(1+a)\mathbf{W}_\alpha}{1+\mathbf{W}} = -2\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R} - P[\bar{R}_{\alpha\alpha}R] + \mathbf{err}(\dot{H}^{\frac12}). \end{cases}$$

One might wish to compare this system with the linearized system which was studied before. However, the terms on the right cannot be all bounded in $L^2 \times \dot{H}^{\frac12}$, even after applying the projection operator P. Precisely, the terms on the right which cannot be bounded directly in $L^2 \times \dot{H}^{\frac12}$ are $-2\dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}_\alpha$, respectively $-2\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{R}$.

But these terms can be eliminated by conjugation with respect to a real exponential weight $e^{2\phi}$, where $\phi = -2\Re \log(1+\mathbf{W})$. Then

$$\phi_\alpha = -2\Re \frac{\mathbf{W}_\alpha}{1+\mathbf{W}}, \qquad (\partial_t + b\partial_\alpha)\phi = 2\Re \frac{R_\alpha}{1+\bar{\mathbf{W}}} - 2M.$$

We denote the weighted variables by

$$w = e^{2\phi}\mathbf{W}_\alpha, \qquad r = e^{2\phi}\mathbf{R}.$$

Using (3.5) and Lemma 3.3 it follows that $Mw = \mathbf{err}(L^2)$, $Mr = \mathbf{err}(\dot{H}^{\frac{1}{2}})$. Then we get the equations

$$\begin{cases} w_t + bw_\alpha + \dfrac{r_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}} w = P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}\mathbf{W}] + \mathbf{err}(L^2), \\ r_t + br_\alpha - i\dfrac{(1+a)w}{1+\mathbf{W}} = -P[\bar{R}_{\alpha\alpha}R] + \mathbf{err}(\dot{H}^{\frac{1}{2}}). \end{cases}$$

We are not yet in a position to use our bounds for the linearized equation since w and r are not exactly holomorphic. We project onto the holomorphic space to write a system for the variables (Pw, Pr). At this point one may legitimately be concerned that restricting to the holomorphic part might remove a good portion of our variables. However, this is not the case:

Lemma 3.4. *The energy of (Pw, Pr) above is equivalent to the energy of $(\mathbf{W}_\alpha, R_\alpha)$*

$$\|(Pw, Pr)\|_{L^2\times \dot{H}^{\frac{1}{2}}} \approx_A \|(w,r)\|_{L^2\times \dot{H}^{\frac{1}{2}}} \approx_A \|(\mathbf{W}_\alpha, R_\alpha)\|_{L^2\times \dot{H}^{\frac{1}{2}}} = \mathbf{N}_2. \tag{3.9}$$

Proof. The estimate for w is easy. We trivially have $\|w\|_{L^2} \lesssim_A \|\mathbf{W}_\alpha\|_{L^2}$, while for the converse we write

$$\|\mathbf{W}_\alpha\|_{L^2}^2 \lesssim \int \bar{\mathbf{W}}_\alpha e^{\phi}\mathbf{W}_\alpha \, d\alpha = \int \bar{\mathbf{W}}_\alpha Pw \, d\alpha \lesssim \|\mathbf{W}_\alpha\|_{L^2}\|Pw\|_{L^2}.$$

To obtain the estimate for r we write

$$|D|^{\frac{1}{2}} P(e^{\phi}(1+\mathbf{W})R_\alpha) = e^{\phi}(1+\mathbf{W})|D|^{\frac{1}{2}} R_\alpha + [P|D|^{\frac{1}{2}}, e^{\phi}(1+\mathbf{W})]R_\alpha.$$

We bound all terms in L^2. The one on the left is $\|r\|_{\dot{H}^{\frac{1}{2}}}$, while the first one on the right is $\approx \|R_\alpha\|_{\dot{H}^{\frac{1}{2}}}$. It remains to bound the commutator on the right, for which we have, with $\Phi = e^{\phi}(1+\mathbf{W})$,

$$\|[P|D|^{\frac{1}{2}}, \Phi]R_\alpha\|_{L^2} \lesssim \||D|\Phi\|_{L^2}\||D|^{\frac{1}{2}}R\|_{BMO} \lesssim_A \|\mathbf{W}_\alpha\|_{L^2}.$$

□

Now, we write the system for (Pw, Pr):

$$\begin{cases} Pw_t + \mathfrak{M}_b Pw_\alpha + P\left[\dfrac{Pr_\alpha}{1+\bar{\mathbf{W}}}\right] + P\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}} Pw\right] = P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar{R}_{\alpha\alpha}\mathbf{W}] + G_2 + \mathbf{err}(L^2), \\ Pr_t + \mathfrak{M}_b Pr_\alpha - iP\left[\dfrac{(1+a)Pw}{1+\mathbf{W}}\right] = -P[\bar{R}_{\alpha\alpha}R] + K_2 + \mathbf{err}(\dot{H}^{\frac{1}{2}}), \end{cases}$$

where

$$G_2 = -P[b\bar{P}w_\alpha] - P\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}\bar{P}w\right], \qquad K_2 = P[b\bar{P}r_\alpha] + iP\left[\frac{(1+a)\bar{P}w}{1+\mathbf{W}}\right].$$

We claim that $G_2 = \mathbf{err}(L^2)$ and $K_2 = \mathbf{err}(\dot H^{\frac12})$. As in Lemma 3.4 we have

$$\|\bar P w\|_{L^2} + \|\bar P r\|_{\dot H^{\frac12}} \lesssim_A A\mathbf{N}_2.$$

Then, using the commutator bounds in Lemma 2.1, we estimate G_2 by

$$\|G_2\|_{L^2} \lesssim_A (\|b_\alpha\|_{BMO} + \|R_\alpha\|_{BMO})\|\bar P w\|_{L^2} \lesssim_A AB\mathbf{N}_2.$$

Similarly, we bound K_2 in $\dot H^{\frac12}$ by

$$\|K_2\|_{\dot H^{\frac12}} \lesssim_A \|b_\alpha\|_{BMO}\|\bar P r\|_{\dot H^{\frac12}} + \left\| |D|^{\frac12}\left(\frac{a-\mathbf{W}}{1+\mathbf{W}}\right)\right\|_{BMO}\|\bar P w\|_{L^2} \lesssim_A AB\mathbf{N}_2,$$

and in L^2 by

$$\|K_2\|_{L^2} \lesssim \||D|^{\frac12} b\|_{BMO}\|\bar P r\|_{\dot H^{\frac12}} + \left\|\frac{a-\mathbf{W}}{1+\mathbf{W}}\right\|_{L^\infty}\|\bar P w\|_{L^2} \lesssim_A A^2\mathbf{N}_2.$$

Finally, in view of the bilinear estimates (3.7), (3.8), we can replace $P[R\bar{\mathbf{W}}_{\alpha\alpha} - \bar R_{\alpha\alpha}\mathbf{W}]$ and $P[\bar R_{\alpha\alpha} R]$ by $P[R\bar P\bar w_\alpha - \mathbf{W}\bar P\bar r_\alpha]$, respectively $P[R\bar P\bar r_\alpha]$ modulo acceptable error terms.

Taking into account the discussion above, we obtain a system for (Pw, Pr) which is very much like the linearized system in the previous section:

$$\begin{cases} Pw_t + \mathfrak{M}_b Pw_\alpha + P\left[\dfrac{Pr_\alpha}{1+\bar{\mathbf{W}}}\right] + P\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}Pw\right] = P[R\bar P\bar w_\alpha - \mathbf{W}\bar P\bar r_\alpha] + \mathbf{err}(L^2), \\ Pr_t + \mathfrak{M}_b Pr_\alpha - iP\left[\dfrac{(1+a)Pw}{1+\mathbf{W}}\right] = -P[R\bar P\bar r_\alpha] + \mathbf{err}(\dot H^{\frac12}). \end{cases}$$

The results of Proposition 3.1 and Proposition 3.2 follow from the energy estimates for the linearized equation, namely part (a) of Propositions 2.1 2.2; further, if $n = 2$ then we can take

$$E^{n,(2)}(\mathbf{W}, R) = E^{(2)}_{lin}(Pw, Pr), \qquad E^{n,(3)}(\mathbf{W}, R) = E^{(3)}_{lin}(Pw, Pr).$$

3.2. **The case $n \geq 3$, large data.** We follow the same strategy as in the case $n = 2$ and derive the equations for $(\mathbf{W}^{(n-1)}, R^{(n-1)})$. We start again with the equations (1.3) and differentiate $n-1$ times. Compared with the case $n = 2$, we obtain many more terms. To separate them into leading order and lower order, we call lower order terms any terms which do not involve $\mathbf{W}^{(n-1)}$, $R^{(n-1)}$ or derivatives thereof. In the computation below we take care to separate all the leading order terms, as well as all the quadratic terms which are lower order. Toward that end we define again the notion of *error term.* Unlike in the case $n = 2$, here we also include lower order quadratic terms into the error. As before, we describe the error bounds in terms of the parameters A, B and

$$\mathbf{N}_n = \|(\mathbf{W}^{(n-1)}, R^{(n-1)})\|_{L^2\times \dot H^{\frac12}}. \tag{3.10}$$

The acceptable errors in the $\mathbf{W}^{(n-1)}$ equation are denoted by $\mathbf{err}(L^2)$ and are of two types, $\mathbf{err}(L^2)^{[2]}$ and $\mathbf{err}(L^2)^{[3]}$. $\mathbf{err}(L^2)^{[2]}$ consists of holomorphic quadratic lower order terms of the form

$$P[\mathbf{W}^{(j)}\mathbf{R}^{(n-j)}], \quad P[\bar{\mathbf{W}}^{(j)}\mathbf{R}^{(n-j)}], \quad P[\mathbf{W}^{(j)}\bar{\mathbf{R}}^{(n-j)}], \qquad 2 \leq j \leq n-2.$$

By interpolation and Hölder's inequality, terms G in $\mathbf{err}(L^2)^{[2]}$ satisfy the bound

$$\|G\|_{L^2} \lesssim B\mathbf{N}_n.$$

By $\mathbf{err}(L^2)^{[3]}$ we denote terms G which satisfy the estimates

$$\|PG\|_{L^2} \lesssim_A AB\mathbf{N}_n,$$

and

$$\text{either} \quad \|\bar{P}G\|_{L^2} \lesssim_A B\mathbf{N}_n \quad \text{or} \quad \|\bar{P}G\|_{\dot{H}^{-\frac{1}{2}}} \lesssim_A A\mathbf{N}_n.$$

The acceptable errors in the $R^{(n-1)}$ equation are denoted by $\mathbf{err}(\dot{H}^{\frac{1}{2}})$ and are also of two types, $\mathbf{err}(\dot{H}^{\frac{1}{2}})^{[2]}$ and $\mathbf{err}(\dot{H}^{\frac{1}{2}})^{[3]}$. $\mathbf{err}(\dot{H}^{\frac{1}{2}})^{[2]}$ consists of holomorphic quadratic lower order terms of the form

$$P[\mathbf{R}^{(j)}\mathbf{R}^{(n-j)}], \quad P[\bar{\mathbf{R}}^{(j)}\mathbf{R}^{(n-j)}], \qquad 2 \le j \le n-2,$$

and

$$P[\mathbf{W}^{(j)}\mathbf{W}^{(n-j)}], \quad P[\bar{\mathbf{W}}^{(j)}\mathbf{W}^{(n-j-1)}], \qquad 1 \le j \le n-1.$$

By interpolation and Hölder's inequality, terms K in $\mathbf{err}(\dot{H}^{\frac{1}{2}})^{[2]}$ satisfy the bound

$$\|K\|_{\dot{H}^{\frac{1}{2}}} \lesssim B\mathbf{N}_n, \qquad \|K\|_{L^2} \lesssim A\mathbf{N}_n.$$

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

$$\|PK\|_{\dot{H}^{\frac{1}{2}}} \lesssim_A AB\mathbf{N}_n, \qquad \|PG\|_{L^2} \lesssim_A A^2\mathbf{N}_n, \qquad \|\bar{P}G\|_{L^2} \lesssim_A A\mathbf{N}_n.$$

We begin by differentiating the terms in the $\mathbf{W}$ equation, where we expand using Leibnitz rule. For the b term we have

$$\begin{aligned}
\partial^{n-1}(b\mathbf{W}_\alpha) &= b\mathbf{W}_\alpha^{(n-1)} + (n-1)b_\alpha \mathbf{W}^{(n-1)} + b^{(n-1)}\mathbf{W}_\alpha + err_1, \\
&= b\mathbf{W}_\alpha^{(n-1)} + (n-1)\left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}}\right)\mathbf{W}^{(n-1)} + 2\mathbf{W}_\alpha \Re R^{(n-1)} + err_2.
\end{aligned}$$

Here err_1 only contains lower order terms, so by interpolation and Hölder's inequality we get[1] $err_1 = \mathbf{err}(L^2)$. The difference $err_2 - err_1$ is cubic,

$$err_2 = err_1 + (n-1)M\mathbf{W}^{(n-1)} + \mathbf{W}_\alpha(P[R^{(n-1)}\bar{Y}] + \bar{P}[\bar{R}^{(n-1)}Y]).$$

Using the L^∞ bound for M in (3.4), Sobolev embeddings and interpolation it is easily seen that $err_2 = \mathbf{err}(L^2)$.

A similar analysis leads to

$$\begin{aligned}
\partial^{n-1}\frac{(1+\mathbf{W})R_\alpha}{1+\bar{\mathbf{W}}} &= \frac{[(1+\mathbf{W})R^{(n-1)}]_\alpha}{1+\bar{\mathbf{W}}} + \frac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}^{(n-1)} - R_\alpha \bar{\mathbf{W}}^{(n-1)} \\
&\quad + R^{(n-1)}((n-2)\mathbf{W}_\alpha - (n-1)\bar{\mathbf{W}}_\alpha) + \mathbf{err}(L^2).
\end{aligned}$$

[1]Here we remark that all terms in the $\mathbf{W}^{(n-1)}$ equation have the same scaling; thus, whenever all the Sobolev exponents are within the lower order range, we are guaranteed to get the correct L^2 estimate after interpolation and Hölder's inequality. The same applies to all the terms in the $R^{(n-1)}$ equation.

In the M term we also bound lower order terms by Hölder's inequality and interpolation to obtain

$$\begin{aligned}\partial^{n-1}[(1+\mathbf{W})M] = \mathbf{err}(L^2) + R^{(n-1)}\bar{\mathbf{W}}_\alpha - \frac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} \\ + P\left[R\bar{\mathbf{W}}_\alpha^{(n-1)} - \bar{R}_\alpha^{(n-1)}\mathbf{W} + (n-1)(R_\alpha\bar{\mathbf{W}}^{(n-1)} - \bar{R}^{(n-1)}\mathbf{W}_\alpha)\right] \\ + \bar{P}[-R^{(n-1)}\bar{\mathbf{W}}_\alpha + \frac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} + \bar{R}^{(n-1)}\mathbf{W}_\alpha - \frac{R_\alpha(1+\mathbf{W})}{(1+\bar{\mathbf{W}})^2}\bar{\mathbf{W}}^{(n-1)} \\ + \bar{R}\mathbf{W}_\alpha^{(n-1)} - R_\alpha^{(n-1)}\bar{\mathbf{W}} + (n-1)(\bar{R}_\alpha\mathbf{W}^{(n-1)} - R^{(n-1)}\bar{\mathbf{W}}_\alpha)].\end{aligned}$$

Estimating also the quadratic $\bar{P}$ terms, the above relation takes the simpler form

$$\begin{aligned}\partial^{n-1}[(1+\mathbf{W})M] = R^{(n-1)}\bar{\mathbf{W}}_\alpha - \frac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} + P\left[R\bar{\mathbf{W}}_\alpha^{(n-1)} - \bar{R}_\alpha^{(n-1)}\mathbf{W}\right] \\ + (n-1)(R_\alpha\bar{\mathbf{W}}^{(n-1)} - \bar{R}^{(n-1)}\mathbf{W}_\alpha) + \mathbf{err}(L^2).\end{aligned}$$

Now we turn our attention to the R equation. We begin with

$$\begin{aligned}\partial^{n-1}(bR_\alpha) &= bR_\alpha^{(n-1)} + (n-1)b_\alpha R^{(n-1)} + b^{(n-1)}R_\alpha + err_3 \\ &= bR_\alpha^{(n-1)} + (n-1)\left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}}\right)R^{(n-1)} + \frac{R_\alpha}{1+\bar{\mathbf{W}}}R^{(n-1)} + \frac{R_\alpha}{1+\mathbf{W}}\bar{R}^{(n-1)} \\ &\quad + err_4,\end{aligned}$$

where we trivially have $err_3 = \mathbf{err}(\dot{H}^{\frac{1}{2}})$ as it contains only lower order terms, both quadratic and higher order. In addition, the difference is cubic, and is given by

$$err_4 - err_3 = (n-1)MR^{(n-1)} + R_\alpha\left(b^{(n-1)} - \frac{R^{(n-1)}}{1+\bar{\mathbf{W}}} - \frac{\bar{R}^{(n-1)}}{1+\mathbf{W}}\right).$$

We claim that this is also $\mathbf{err}(\dot{H}^{\frac{1}{2}})$. The L^2 bound follows trivially by interpolation and Hölder's inequality.

The $\dot{H}^{\frac{1}{2}}$ bound is also easy to obtain for the second term, where the unfavorable $R^{(n-1)}$ factors only appear with a convenient frequency balance as $R_\alpha(\bar{P}[R^{(n-1)}\bar{Y}] - P[R^{(n-1)}Y])$. Consider now the $\dot{H}^{\frac{1}{2}}$ bound for the first term. Since $M = O_{L^\infty}(AB)$, the nontrivial case is when M has the high frequency, where we need to estimate

$$\||D|^{\frac{1}{2}}\sum_k M_k R_{<k}^{(n-1)}\|_{L^2} \lesssim \||D|^{n-\frac{3}{2}}M\|_{L^2}\|DR\|_{BMO} \lesssim AB\mathbf{N}_n.$$

Here, we have used the bound (B.33).

For the remaining term in the R equation we write

$$\begin{aligned}\partial^{n-1}\frac{\mathbf{W}-a}{1+\mathbf{W}} &= \frac{(1+a)\mathbf{W}^{(n-1)}}{(1+\mathbf{W})^2} + \frac{a^{(n-1)}}{1+\mathbf{W}} + err_7 \\ &= \frac{(1+a)\mathbf{W}^{(n-1)}}{(1+\mathbf{W})^2} + \frac{i}{1+\mathbf{W}}\left(P[R\bar{R}_\alpha^{(n-1)} + (n-1)R_\alpha\bar{R}^{(n-1)} + \bar{R}_\alpha R^{(n-1)}]\right. \\ &\quad \left.-\bar{P}[\bar{R}R_\alpha^{(n-1)} + (n-1)\bar{R}_\alpha R^{(n-1)} + R_\alpha\bar{R}^{(n-1)}]\right) + err_8.\end{aligned}$$

Here err_7 contains lower order quadratic terms in $\mathbf{W}$ (without a) as well as cubic terms which can be easily estimated, so $err_7 = \mathbf{err}(\dot{H}^{\frac{1}{2}})$. The difference $err_8 - err_7$ only contains lower order terms so it also can be placed in $\mathbf{err}(\dot{H}^{\frac{1}{2}})$. Just as in the case of the $\mathbf{W}^{(n-1)}$ equation, quadratic $\bar{P}$ terms can also be placed in the error. Then the above relation becomes

$$\partial^{n-1}\frac{\mathbf{W}-a}{1+\mathbf{W}} = \frac{(1+a)\mathbf{W}^{(n-1)}}{(1+\mathbf{W})^2} + i\left(P[R\bar{R}_\alpha^{(n-1)}] + (n-1)R_\alpha\bar{R}^{(n-1)} + \frac{\bar{R}_\alpha R^{(n-1)}}{1+\mathbf{W}}\right) + \mathbf{err}(\dot{H}^{\frac{1}{2}}).$$

Combining the above computations we obtain the differentiated system

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{((1+\mathbf{W})R^{(n-1)})_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} = G, \\ R_t^{(n-1)} + bR_\alpha^{(n-1)} - i\left(\dfrac{(1+a)\mathbf{W}^{(n-1)}}{(1+\mathbf{W})^2}\right) = K, \end{cases}$$

where

$$\begin{aligned} G =& -n\frac{\bar{R}_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} - (n-1)\frac{R_\alpha}{1+\bar{\mathbf{W}}}\mathbf{W}^{(n-1)} + P[R\bar{\mathbf{W}}_\alpha^{(n-1)} - \mathbf{W}\bar{R}_\alpha^{(n-1)}] \\ &+ R^{(n-1)}(n\bar{\mathbf{W}}_\alpha - (n-1)\mathbf{W}_\alpha) + n(R_\alpha\bar{\mathbf{W}}^{(n-1)} - \mathbf{W}_\alpha\bar{R}^{(n-1)}) + \mathbf{err}(L^2), \\ K =& -n\left(\frac{R_\alpha}{1+\bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1+\mathbf{W}}\right)R^{(n-1)} - \left(P[R\bar{R}_\alpha^{(n-1)}] - nR_\alpha\bar{R}^{(n-1)}\right) + \mathbf{err}(\dot{H}^{\frac{1}{2}}). \end{aligned}$$

After the usual substitution $\mathbf{R} = (1+\mathbf{W})R^{(n-1)}$, we get

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{\mathbf{R}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} = G, \\ \mathbf{R}_t + b\mathbf{R}_\alpha - i\left(\dfrac{(1+a)\mathbf{W}^{(n-1)}}{1+\mathbf{W}}\right) = K_1, \end{cases}$$

where

$$K_1 = -(n+1)\frac{R_\alpha\mathbf{R}}{1+\bar{\mathbf{W}}} - n\frac{\bar{R}_\alpha\mathbf{R}}{1+\mathbf{W}} - P[R\bar{\mathbf{R}}_\alpha] - nR_\alpha\bar{\mathbf{R}} + \mathbf{err}(\dot{H}^{\frac{1}{2}}).$$

The more delicate terms here are the ones on the right where the leading order terms appear unconjugated. We would like to eliminate those with an exponential factor as in the $n=2$ case, but their coefficients on the right are not properly matched. To remedy that we take the additional step of the 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 exactly three terms, the contribution of the added quadratic correction is cubic and lower order, so we obtain

$$\begin{cases} \mathbf{W}_t^{(n-1)} + b\mathbf{W}_\alpha^{(n-1)} + \dfrac{\tilde{\mathbf{R}}_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}\mathbf{W}^{(n-1)} = -n\left(\dfrac{\bar{R}_\alpha}{1+\mathbf{W}} + \dfrac{R_\alpha}{1+\bar{\mathbf{W}}}\right)\mathbf{W}^{(n-1)} \\ \qquad + P[R\bar{\mathbf{W}}_\alpha^{(n-1)} - \mathbf{W}\bar{R}_\alpha^{(n-1)}] + n\tilde{\mathbf{R}}(\bar{\mathbf{W}}_\alpha + \mathbf{W}_\alpha) + n(R_\alpha\bar{\mathbf{W}}^{(n-1)} - \mathbf{W}_\alpha\bar{R}^{(n-1)}) + \mathbf{err}(L^2), \\ \tilde{\mathbf{R}}_t + b\tilde{\mathbf{R}}_\alpha - i\dfrac{(1+a)\mathbf{W}^{(n-1)}}{1+\mathbf{W}} = -n\left(\dfrac{R_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{\bar{R}_\alpha}{1+\mathbf{W}}\right)\tilde{\mathbf{R}} - P[R\bar{\tilde{\mathbf{R}}}_\alpha] - nR_\alpha\bar{\tilde{\mathbf{R}}} + \mathbf{err}(\dot{H}^{\frac{1}{2}}). \end{cases}$$

Now we can multiply by $e^{n\phi}$ where, as before, $\phi = -2\Re\log(1+\mathbf{W})$, in order to eliminate the unbounded terms on the right. We get an equation for $(w := e^{\phi}\mathbf{W}^{(n-1)}, r := e^{\phi}\tilde{\mathbf{R}})$:

$$\begin{cases} w_t + bw_\alpha + \dfrac{r_\alpha}{1+\bar{\mathbf{W}}} + \dfrac{R_\alpha}{1+\mathbf{W}}w = P[R\bar{\mathbf{W}}_\alpha^{(n-1)} - \mathbf{W}\bar{R}_\alpha^{(n-1)}] + n(R_\alpha\bar{w} - \mathbf{W}_\alpha\bar{r}) + \mathbf{err}(L^2), \\ r_t + br_\alpha - i\left(\dfrac{(1+a)w}{1+\mathbf{W}}\right) = -P[R\bar{R}_\alpha^{(n-1)}] - nR_\alpha\bar{r} + \mathbf{err}(\dot{H}^{\frac{1}{2}}). \end{cases}$$

As (w, r) are no longer holomorphic, we project and work with the projected variables. After some additional commutator estimates which are identical to those in the $n = 2$ case we obtain

$$\begin{cases} Pw_t + \mathfrak{M}_b Pw_\alpha + P\left[\dfrac{Pr_\alpha}{1+\bar{\mathbf{W}}}\right] + P\left[\dfrac{R_\alpha Pw}{1+\mathbf{W}}\right] = P[R\bar{P}\bar{w}_\alpha - \mathbf{W}\bar{P}\bar{r}_\alpha] \\ \qquad\qquad\qquad\qquad\qquad\qquad + nP[R_\alpha\bar{P}\bar{w} - \mathbf{W}_\alpha\bar{P}\bar{r}] + \mathbf{err}(L^2), \\ Pr_t + \mathfrak{M}_b Pr_\alpha - iP\left[\dfrac{(1+a)Pw}{1+\mathbf{W}}\right] = -P[R\bar{P}\bar{r}_\alpha] - nP[R_\alpha\bar{P}\bar{r}] + \mathbf{err}(\dot{H}^{\frac{1}{2}}). \end{cases} \tag{3.11}$$

Compared to the linearized equation in the previous section, here we have two additional terms that need to be estimated. We have

Lemma 3.5. *a) The energy of (Pw, Pr) above is equivalent to that of $(\mathbf{W}^{(n-1)}, R^{(n-1)})$,*

$$\|(Pw, Pr)\|_{L^2\times\dot{H}^{\frac{1}{2}}} \approx_A \|(w,r)\|_{L^2\times\dot{H}^{\frac{1}{2}}} \approx_A \mathbf{N}_n, \tag{3.12}$$

b) The additional error terms above are bounded,

$$\|(P[R_\alpha\bar{P}\bar{w} - \mathbf{W}_\alpha\bar{P}\bar{r}], P[R_\alpha\bar{P}\bar{r}])\|_{L^2\times\dot{H}^{\frac{1}{2}}} \lesssim_A B\mathbf{N}_n, \tag{3.13}$$

Proof. a) For w we argue as in the proof of Lemma 3.4 to get

$$\|Pw\|_{L^2} \approx_A \|w\|_{L^2} \approx_A \|\mathbf{W}^{(n-1)}\|_{L^2}.$$

For r we need, again as in the proof of Lemma 3.4, with $\Phi = e^{\phi}(1+\mathbf{W})$, to bound in L^2 the difference

$$\begin{aligned} K &= |D|^{\frac{1}{2}}Pr - \Phi|D|^{\frac{1}{2}}R^{(n-1)}, \\ &= [|D|^{\frac{1}{2}}P, \Phi]R^{(n-1)} + |D|^{\frac{1}{2}}P[e^{\phi}(-R_\alpha W^{(n-2)} + (2n-1)\mathbf{W}_\alpha R^{(n-2)})], \end{aligned}$$

as well as the similar difference but with all P omitted. It suffices to prove the estimate

$$\|K\|_{L^2} \lesssim_A \|\mathbf{W}^{(n-1)}\|_{L^2} + \|\mathbf{W}^{(n-1)}\|_{L^2}^{\frac{1}{n-1}}\|R^{(n-1)}\|_{\dot{H}^{\frac{1}{2}}}^{\frac{n-2}{n-1}},$$

which follows by standard multiplicative and commutator estimates.

b) For the first term $P[R_\alpha\bar{P}\bar{w}]$ we directly use the Coifman-Meyer type estimates in Lemma 2.1. For the second we bound $\mathbf{W}_\alpha$ in L^{4n-6} and r in $L^{\frac{2n-3}{n-2}}$ by Hölder's inequality and interpolation. For the third we have to bound $\||D|^{\frac{1}{2}}P[R_\alpha\bar{P}\bar{r}]\|_{L^2}$. For the balanced frequency interactions, by Coifman-Meyer it suffices to bound R_α in BMO and r in $\dot{H}^{\frac{1}{2}}$. For the high-low interactions, on the other hand, the half-derivative goes to R_α, and we need to bound $|D|^{\frac{1}{2}}R_\alpha$, and r in L^{4n-6}, respectively $L^{\frac{2n-3}{n-2}}$.

□

Given the above Lemma 3.5, the $n \geq 3$ case of the result in Proposition 3.1 is a direct consequence of our quadratic estimates for the linearized equation in Proposition 2.1(a).

The small data cubic energy estimates in Proposition 3.2 are proved in the next section. The key is to produce a modified cubic energy, whose leading part is given by

$$E^{n,(3)}_{high}(w,r) = \int (1+a)|w|^2 + \Im(\bar{r} r_\alpha) + 2n\Im(R_\alpha \bar{w}\bar{r}) + 2(\Im[\bar{R} w r_\alpha] - \Re[\bar{W}_\alpha w^2])\, d\alpha.$$

We claim that the evolution of this energy is governed by the following

Lemma 3.6. *Let (w,r) be defined as above. Then*

a) Assuming that $A \ll 1$, we have

$$E^{n,(3)}_{high}(Pw, Pr) \approx E_0(Pw, Pr) \approx \mathbf{N}_n, \tag{3.14}$$

b) The solutions (Pw, Pr) of (3.11) satisfy

$$\begin{aligned} \frac{d}{dt} E^{n,(3)}_{high}(Pw,Pr) = 2\int & \Re(\bar{w}\cdot \mathbf{err}(L^2)^{[2]}) - \Im(\bar{r}_\alpha \cdot \mathbf{err}(\dot{H}^{\frac12})^{[2]})\, d\alpha \\ & + O_A(AB\mathbf{N}_n). \end{aligned} \tag{3.15}$$

Further, the same relation holds if $(\bar{w},\bar{r})$ on the right are replaced by $(\bar{\mathbf{W}}^{(n-1)}, \bar{R}^{(n-1)})$.

Proof. a) Given the bounds already proved in Proposition 2.2 for the linearized equation, it suffices to estimate the additional term,

$$\left| \int R_\alpha \bar{w}\bar{r}\ d\alpha \right| \lesssim A\mathbf{N}_n.$$

For this we use interpolation to bound R_α, w and r in L^{4n-6}, L^2, respectively $L^{\frac{2n-3}{n-2}}$ in terms of A and $\mathbf{N}_n$.

b) Here, we begin with the cubic linearized energy, $E^{(3)}_{lin}$. According to the bound (2.13) in Proposition 2.2, we have

$$\begin{aligned} \frac{d}{dt} E^{(3)}_{lin}(Pw,Pr) = \int & 2\Re\left((nP[R_\alpha \bar{P}\bar{w} - \mathbf{W}_\alpha \bar{P}\bar{r}] + P\mathbf{err}(L^2))\cdot(\bar{w} - \bar{P}[\bar{R} r_\alpha] - \bar{P}[\bar{W}_\alpha w])\right) \\ & - 2\Im\left((-nP[R_\alpha \bar{P}\bar{r}] + P\mathbf{err}(\dot{H}^{\frac12}))\cdot(\bar{r}_\alpha + \bar{P}[\bar{R} w]_\alpha)\right)\, d\alpha \\ & + O_A\left(AB\|(Pw,Pr)\|^2_{L^2\times \dot{H}^{\frac12}}\right). \end{aligned}$$

By the Coifman-Meyer type estimates in Lemma 2.1 the following bounds hold:

$$\|\bar{P}[\bar{R} r_\alpha]\|_{L^2} + \|\bar{P}[\bar{W}_\alpha w]\|_{L^2} + \|\bar{P}[\bar{R} w]\|_{\dot{H}^{\frac12}} \lesssim A\|(w,r)\|_{L^2\times\dot{H}^{\frac12}}. \tag{3.16}$$

Combining this with (3.13) and with the bounds for the error terms we get

$$\begin{aligned} \frac{d}{dt} E^{(3)}_{lin}(Pw,Pr) \leq \int & 2\Re\left((nP[R_\alpha \bar{P}\bar{w} - \mathbf{W}_\alpha \bar{P}\bar{r}] + P\mathbf{err}(L^2)^{[2]})\cdot \bar{w}\right) \\ & - 2\Im\left((-nP[R_\alpha \bar{P}\bar{r}] + P\mathbf{err}(\dot{H}^{\frac12})^{[2]})\cdot \bar{r}_\alpha\right)\, d\alpha \\ & + O_A\left(AB\|(Pw,Pr)\|^2_{L^2\times\dot{H}^{\frac12}}\right), \end{aligned}$$

where the output from all error terms which are cubic and higher error terms is all included in the last RHS term.

It remains to consider the contribution of the extra term in $E^{n,(3)}_{high}$ and show that

$$\frac{d}{dt}\int \Im(R_\alpha \bar{P}\bar{w}\bar{P}\bar{r})\, d\alpha = \int \Re\left((R_\alpha \bar{P}\bar{w} - \mathbf{W}_\alpha \bar{P}\bar{r})\bar{P}\bar{w}\right) + \Im\left(R_\alpha \bar{P}\bar{r}\bar{P}\bar{r}_\alpha\right)\, d\alpha \\ + O_A\left(AB\|(Pw, Pr)\|^2_{L^2\times \dot{H}^{\frac{1}{2}}}\right). \tag{3.17}$$

Denote by G_n, respectively K_n the two right hand sides in (3.11). By the definition of error terms and by (3.13) they satisfy the bounds

$$\|(G_n, K_n)\|_{L^2\times \dot{H}^{\frac{1}{2}}} \lesssim_A B\mathbf{N}_n, \qquad \|K_n\|_{L^2} \lesssim_A A\mathbf{N}_n.$$

Then their contribution in the above time derivative is estimated

$$\left|\int \Im(R_\alpha \bar{P}\bar{G}_n\bar{P}\bar{r} + R_\alpha \bar{P}\bar{w}\bar{P}\bar{K}_n)\ d\alpha\right| = \left|\int \Im(R_\alpha \bar{P}\bar{F}_n\bar{P}\bar{r} + P[R_\alpha \bar{P}\bar{w}]\bar{P}\bar{K}_n)\ d\alpha\right| \lesssim_A AB\mathbf{N}_2,$$

by using Hölder's inequality for the first term and the Coifman-Meyer commutator estimate in Lemma 2.1 for the second.

The contributions of the b terms are collected together in the imaginary part of the expression

$$\begin{aligned} I &= \int \partial_\alpha(bR_\alpha)\bar{P}\bar{w}\bar{P}\bar{r} + R_\alpha\bar{P}(b\bar{P}\bar{w}_\alpha)\bar{P}\bar{r} + R_\alpha\bar{P}\bar{w}\bar{P}(b\bar{P}\bar{r}_\alpha)\, d\alpha \\ &= \int R_\alpha\left([b,P](\bar{P}\bar{w}_\alpha)\bar{P}\bar{r} + \bar{P}\bar{w}[b,P](\bar{P}\bar{r}_\alpha)\right)\, d\alpha. \end{aligned}$$

Since $\|b_\alpha\|_{BMO} \lesssim B$, we can bound using Lemma 2.1, and then use Hölder's inequality for all terms.

Next, we consider the remaining contribution of the time derivative of R_α, for which we use the equation (1.3). This is

$$\Im\int \bar{P}\bar{w}\bar{P}\bar{r}\partial_\alpha\left(\frac{\mathbf{W}-a}{1+\mathbf{W}}\right)\, d\alpha = \Re\int \bar{P}\bar{w}\bar{P}\bar{r}\mathbf{W}_\alpha\, d\alpha - \Re\int \bar{P}\bar{w}\bar{P}\bar{r}\partial_\alpha\left(\frac{\mathbf{W}^2+a}{1+\mathbf{W}}\right)\, d\alpha.$$

The first term on the right yields the second term on the right of (3.17), while the rest of the terms are directly bounded using Hölder's inequality.

It remains to consider the contribution of the remaining left hand side terms in (3.11). The expression $\dfrac{Pr_\alpha}{1+\bar{\mathbf{W}}}$ in the r equation yields the third term on the right of (3.17), plus the quartic term

$$\int \Im R_\alpha \bar{P}(\bar{P}\bar{r}_\alpha Y)\bar{P}\bar{r}\, d\alpha = \int \Im(R_\alpha[\bar{P}, Y](\bar{P}\bar{r}_\alpha)\bar{P}\bar{r} + R_\alpha Y\bar{P}\bar{r}_\alpha\bar{P}\bar{r})\, d\alpha.$$

In the first term we apply a commutator estimate and then Hölder's inequality, and in the second we use Hölder inequality directly.

The contribution of $P\left[\dfrac{R_\alpha}{1+\mathbf{W}}Pw\right]$ is purely a Hölder term. Finally, the contribution of $P\left[\dfrac{1+a}{1+\mathbf{W}}Pw\right]$ yields the first term on the right of (3.17), plus a Hölder quartic term.

□

3.3. **Normal form energy estimates:** $n \geq 3$ **, small data.** In this section, we construct an n-th order energy with cubic estimates. One ingredient for this is the high frequency cubic energy $E^{n,(3)}_{high}$ in Lemma 3.6. However, this does not suffice, as the right hand side of the energy relation (3.15) still contains lower order cubic terms. Here we use normal forms in order to add a lower order correction to $E^{n,(3)}_{high}$, which removes the above mentioned cubic terms. We recall that the normal form variables $(\tilde{W}, \tilde{Q})$ are given by

$$\begin{cases} \tilde{W} = W - 2\mathfrak{M}_{\Re W} W_\alpha, \\ \tilde{Q} = Q - 2\mathfrak{M}_{\Re W} R, \end{cases} \tag{3.18}$$

where $\mathfrak{M}_u F = P[uF]$. They solve an equation where all nonlinearities are cubic and higher,

$$\begin{cases} \tilde{W}_t + \tilde{Q}_\alpha = \tilde{G}, \\ \tilde{Q}_t - i\tilde{W} = \tilde{K}, \end{cases} \tag{3.19}$$

see Proposition 1.1.

The obvious energy functional associated to the normal form equations (1.14) is

$$E^n_{NF,0} = \int \left(|\tilde{W}^{(n)}|^2 + \Im[\tilde{Q}^{(n)} \bar{\tilde{Q}}^{(n)}_\alpha] \right) d\alpha.$$

In view of Proposition 1.1, this functional satisfies an energy equation of the form

$$\frac{d}{dt} E^n_{NF,0} = quartic + higher, \tag{3.20}$$

but it has several defects:

(1) It is expressed in terms of $Q^{(n)}$ rather than the natural variable $R^{(n-1)}$,
(2) It is not equivalent to the linear energy $E^{(2)}_{lin}(\mathbf{W}^{(n-1)}, R^{(n-1)})$,
(3) Its energy estimate has a loss of derivatives.

However, the last two issues concerning $E^n_{NF,0}$ arise at the level of quartic and higher order terms, and they are specific to the water wave problem. This motivates our strategy, which is modify $E^n_{NF,0}$ by quartic and higher terms to obtain a "good" energy $E^{n,(3)}$ without spoiling the cubic energy estimate (3.20).

We carry out this procedure in two steps: **(i)** we construct a modified normal form energy E^n_{NF} that depends on $(\mathbf{W}^{(n-1)}, R^{(n-1)})$ and is equivalent to the linearized energy $E^{(2)}_{lin}(\mathbf{W}^{(n-1)}, R^{(n-1)})$; this addresses the issues (1) and (2) above, but not (3); **(ii)** we separate the leading order part $E^n_{NF,high}$ and modify that to the correct high frequency expression $E^{n,(3)}_{high}$ defined in the previous section, which was inspired from the analysis of the linearized equation. This modification is needed due to the quasilinear nature of our problem. Thus, we obtain an energy $E^{n,(3)}$ with good, cubic estimates.

The first step described above is implemented in the following proposition:

Proposition 3.7. *There exists a modified normal form energy* E^n_{NF} *of the form*

$$
(3.21)\qquad
\begin{aligned}
E^n_{NF} &= E^n_{NF,high} + E^n_{NF,low},\\
E^n_{NF,high} &= \int (1-4n\Re\mathbf{W})\left(|\mathbf{W}^{(n)}|^2 + \Im[\tilde{\mathbf{R}}\bar{\tilde{\mathbf{R}}}_\alpha]\right) + 2n\Im[R_\alpha \bar{\mathbf{W}}^{(n-1)}\bar{\tilde{\mathbf{R}}}]\, d\alpha,\\
&\quad + 2\int \Im[\bar{R}\mathbf{W}^{(n-1)}\tilde{\mathbf{R}}_\alpha] - \Re[\bar{\mathbf{W}}(\mathbf{W}^{(n-1)})^2]\, d\alpha,\\
E^n_{NF,low} &= \Re\int\left(\sum_{j+k+l=2n-2} c_{jkl}\mathbf{W}^{(j)}\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)} + \sum_{j+k+l=2n-1} d_{jk_1l_1}\mathbf{W}^{(j)}R^{(k)}\bar{R}^{(l)}\right) d\alpha,
\end{aligned}
$$

such that

$$
(3.22)\qquad E^n_{NF} = E^n_{NF,0} + \text{(quartic and higher terms)},
$$

and

$$
(3.23)\qquad E^n_{NF,high} = [1+O(A)]E_0(\mathbf{W}^{(n-1)}, R^{(n-1)}), \qquad E^n_{NF,low} = O(A)E_0(\mathbf{W}^{(n-1)}, R^{(n-1)}).
$$

Moreover, the sums in (3.21) for $E^n_{NF,low}$ *contain only indices* (j,k,l) *with* $1 \leq j,k,l \leq n-1$.

Remark 3.8. *The normal form transformation is expressed at the level of* (W,Q) *variables, and cannot be easily switched to the level of* $(\mathbf{W},R)$*. For this reason, initially the computation of the normal form energy is done in terms of the original variables* (W,Q)*. The interesting fact in the above proposition is that in the end we are able express the energies in the convenient variables* $(\mathbf{W},R)$*.*

Proof. We start from the normal form energy $E^n_{NF,0}$ and express it in terms of $(\mathbf{W},R)$ and their derivatives. First, consider the term involving $\tilde{W}^{(n)}$. Using (3.18), we get that

$$
\int |\tilde{W}^{(n)}|^2\, d\alpha = \int |W^{(n)}|^2 - 4\Re\left[\bar{W}^{(n)}\partial^n_\alpha(\mathfrak{M}_{\Re W}W_\alpha)\right] + 4\left|\partial^n_\alpha(\mathfrak{M}_{\Re W}W_\alpha)\right|^2\, d\alpha.
$$

The higher-order derivatives $W^{(n+1)}$ cannot be removed from the last term, but it is quartic and therefore harmless. The cubic term also contain derivatives of order $n+1$, but as we show next they integrate out; as a result, the cubic energy is equivalent to the linear energy.

Moving the projection P across the inner product, we have

$$
\int \bar{W}^{(n)}\partial^n_\alpha(\mathfrak{M}_{\Re W}W_\alpha)\, d\alpha = \int \bar{W}^{(n)}\partial^n_\alpha(P[W_\alpha\Re W])\, d\alpha = \int \bar{W}^{(n)}\partial^n_\alpha(W_\alpha\Re W)\, d\alpha,
$$

which shows that

$$
\int |\tilde{W}^{(n)}|^2\, d\alpha = \int |W^{(n)}|^2 - 4\Re\left[\bar{W}^{(n)}\partial^n_\alpha(W_\alpha\Re W)\right]\, d\alpha + \text{quartic}.
$$

Thus, expanding derivatives, we get

$$\begin{aligned}\int |\tilde{W}^{(n)}|^2 \, d\alpha = \int & |W^{(n)}|^2 - 4\Re[\bar{W}^{(n)} W_\alpha^{(n)}]\Re W - 4n|W^{(n)}|^2 \Re W_\alpha \\ & - 4\sum_{j=2}^{n-1} \binom{n}{j} \Re\left[\bar{W}^{(n)} W^{(n-j+1)}\right] \Re W^{(j)} - 4\Re[W_\alpha \bar{W}^{(n)}]\Re W^{(n)} \, d\alpha \\ & + \text{quartic}.\end{aligned} \tag{3.24}$$

Integrating by parts in the cubic term that contain derivatives of W of the order $n+1$, we get

$$\int \Re[\bar{W}^{(n)} W_\alpha^{(n)}]\Re W \, d\alpha = -\frac{1}{2}\int (\Re W_\alpha)|W^{(n)}|^2 \, d\alpha.$$

In addition,

$$\int \Re[W_\alpha \bar{W}^{(n)}]\Re W^{(n)} \, d\alpha = \frac{1}{2}\int (\Re W_\alpha)|W^{(n)}|^2 + \Re[W_\alpha (\bar{W}^{(n)})^2] \, d\alpha.$$

It follows that

$$\begin{aligned}\int |\tilde{W}^{(n)}|^2 \, d\alpha = \int & (1 - 4n\Re W_\alpha)\, |W^{(n)}|^2 - 4\sum_{j=2}^{n-1} \binom{n}{j} \Re\left[\bar{W}^{(n)} W^{(n-j+1)}\right] \Re W^{(j)} \\ & - 2\Re[W_\alpha (\bar{W}^{(n)})^2] \, d\alpha + \text{quartic}.\end{aligned}$$

A similar, but longer, computation for the terms involving $\tilde{Q}$ yields

$$\begin{aligned}\int \Im[\tilde{Q}^{(n)} \bar{\tilde{Q}}_\alpha^{(n)}] \, d\alpha = \Im \int & (1 - 4n\Re W_\alpha) \left(Q^{(n)} - Q_\alpha W^{(n)}\right) \left(\bar{Q}^{(n)} - \bar{Q}_\alpha \bar{W}^{(n)}\right)_\alpha \\ & + 2n \left(Q_{\alpha\alpha} W^{(n)} \bar{Q}^{(n)} + Q_{\alpha\alpha} \bar{W}^{(n)} \bar{Q}^{(n)}\right) + 2\bar{Q}_\alpha W^{(n)} Q_\alpha^{(n)} \, d\alpha \\ & + 4\int \sum_{j=3}^{n-1} \binom{n+1}{j} \Im[\bar{Q}^{(n)} Q^{(n-j+2)}]\Re W^{(j)} \, d\alpha + \text{quartic}.\end{aligned}$$

Up to quartic corrections, we may replace $Q^{(n)}$ by $R^{(n-1)}$ and $Q^{(j)}$ by $R^{(j-1)}$ for $j \leq n$ in the cubic terms on the right-hand side of this equation. Further, we have

$$\begin{aligned}\partial_\alpha^n Q - R\partial_\alpha^n W &= (1 + W_\alpha)R^{(n-1)} + \sum_{j=1}^{n-2} \binom{n-1}{j} R^{(j)} W^{(n-j)} \\ &= \tilde{\mathbf{R}} + n(R_\alpha W^{(n-1)} - W_{\alpha\alpha} R^{(n-2)}) + \sum_{j=2}^{n-3} \binom{n-1}{j} R^{(j)} W^{(n-j)}.\end{aligned}$$

Thus, we obtain

(3.25)

$$\int \Im[\tilde{Q}^{(n)}\bar{\tilde{Q}}_\alpha^{(n)}]\, d\alpha = \Im \int (1-4n\Re W_\alpha)\, \tilde{\mathbf{R}}\bar{\tilde{\mathbf{R}}}_\alpha + 2nR_\alpha \bar{W}^{(n)}\bar{\tilde{\mathbf{R}}} + 2\bar{R}_\alpha W^{(n)} R_\alpha^{[n]} d\alpha$$

$$+ 2\Im\int n\bar{\tilde{\mathbf{R}}}(W^{(3)}R^{(n-2)} - R^{(2)}W^{(n-1)}) + \bar{\tilde{\mathbf{R}}}_\alpha \sum_{j=2}^{n-3} \binom{n-1}{j} R^{(j)}W^{(n-j)})d\alpha$$

$$+ 4\int \sum_{j=3}^{n-1} \binom{n+1}{j} \Im[\bar{R}^{(n-1)}R^{(n-j+1)}]\Re W^{(j)} \;+ \text{quartic.}$$

Adding (3.24) and (3.25), we find that

$$E^n_{NF,0} = E^n_{NF} + \text{quartic terms},$$

where E^n_{NF} is given by (3.21) with

$$E^n_{NF,low} = -4\int \sum_{j=2}^{n-1} \binom{n}{j} \Re\left[\bar{W}^{(n)}W^{(n-j+1)}\right] \Re W^{(j)}\, d\alpha$$

$$+ 4\int \sum_{j=3}^{n-1} \binom{n+1}{j} \Im[\bar{R}^{(n-1)}R^{n-j+1}]\Re W^{(j)}\, d\alpha$$

$$+ 2\Im\int n\bar{\tilde{\mathbf{R}}}(W^{(3)}R^{(n-2)} - R^{(2)}W^{(n-1)}) + \bar{\tilde{\mathbf{R}}}_\alpha \sum_{j=2}^{n-3} \binom{n-1}{j} R^{(j)}W^{(n-j)}\, d\alpha,$$

which, after we substitute W_α by $\mathbf{W}$, gives us an energy of the form stated in the proposition.

It remains to establish (3.22). The second estimate follows immediately from Hölder's inequality and interpolation. So does most of the first, except for two terms. By the Coifman-Meyer estimate in Lemma 2.1 we have

$$\int \bar{R}\mathbf{W}^{(n-1)}R_\alpha^{(n-1)} d\alpha = \int \mathbf{W}^{(n-1)}\bar{P}[\bar{R}R_\alpha^{(n-1)}]d\alpha = O(A)\|\mathbf{W}^{(n-1)}\|_{L^2}\|R^{(n-1)}\|_{\dot{H}^{\frac{1}{2}}}.$$

On the other hand, for the integral

$$\int \Re\mathbf{W}\,\Im[R^{(n-1)}\bar{R}_\alpha^{(n-1)}]d\alpha,$$

we do a Littlewood-Paley decomposition, using the $\dot{H}^{\frac{1}{2}}$ norm of $R^{(n-1)}$ if the two R frequencies are high, and interpolation and Hölder's inequality otherwise.

□

To get our final energy functionals $E^{n,(3)}$, we replace $E^n_{NF,high}$ in E^n_{NF} by its nonlinear version, $E^{n,(3)}_{high} := E^{n,(3)}_{high}(Pw, Pr)$. That is, we define

(3.26) $$E^{n,(3)} = E^n_{NF} - E^n_{NF,high} + E^{n,(3)}_{high} = E^n_{NF,low} + E^{n,(3)}_{high}.$$

Note that $E^{n,(3)}$ differs from E^n_{NF} only by a quartic term.

Now we proceed to prove Proposition 3.2. The norm equivalence is already known from (3.14) and (3.23), so we still need the energy estimate. First, we write

$$\frac{d}{dt}E^{n,(3)} = \frac{d}{dt}E^n_{NF} + \frac{d}{dt}\left(E^{n,(3)}_{high} - E^n_{NF,high}\right).$$

This equation shows that there are no cubic terms on the right-hand side, since the derivatives of E^n_{NF} and $E^{n,(3)}_{high} - E^n_{NF,high}$ contain only terms that are quartic or higher order.

Next, we write

$$\frac{d}{dt}E^{n,(3)} = \frac{d}{dt}E^n_{NF,low} + \frac{d}{dt}E^{n,(3)}_{high}.$$

Both expressions have cubic terms, but these cancel due to the prior computation. To make this cancellation precise, at this point we make the convention that all multilinear expansions are in terms of $\mathbf{W}$ and R. To make this cancellation explicit, we introduce a truncation operator Λ^4 that removes the cubic terms and retains everything which is quartic and higher.

Hence, we obtain

$$\frac{d}{dt}E^{n,(3)} = \Lambda^4\left(\frac{d}{dt}E^n_{NF,low}\right) + \Lambda^4\left(\frac{d}{dt}E^{n,(3)}_{high}\right). \tag{3.27}$$

It remains to prove the following estimates:

$$\left|\Lambda^4\left(\frac{d}{dt}E^n_{NF,low}\right)\right| \lesssim_A AB\mathbf{N}_n^2, \tag{3.28}$$

$$\left|\Lambda^4\left(\frac{d}{dt}E^{n,(3)}_{high}\right)\right| \lesssim_A AB\mathbf{N}_n^2. \tag{3.29}$$

The second bound follows directly from (3.15), so it remains to prove (3.28).

3.4. **Estimates for lower order terms: proof of (3.28).** We have two main types of energy terms (or their complex conjugates) to consider,

$$I_1 = \int \mathbf{W}^{(j)}\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)}d\alpha, \qquad j+k+l = 2n-2, \quad 1 \leq j,k,l \leq n,$$

$$I_2 = \int \mathbf{W}^{(j)}R^{(k)}\bar{R}^{(l)}d\alpha, \qquad j+k+l = 2n-1, \quad 1 \leq j,k,l \leq n.$$

To estimate their time derivatives it is easiest to use the unprojected form (1.3) of the equations for $\mathbf{W}$ and R, which for our purposes here we write in the form

$$\begin{cases} (\partial_t + b\partial_\alpha)\mathbf{W} = \ -b_\alpha(1+\mathbf{W}) - \bar{R}_\alpha := G, \\ (\partial_t + b\partial_\alpha)R = \ i\dfrac{\mathbf{W}-a}{1+\mathbf{W}} := K. \end{cases} \tag{3.30}$$

Of G and K we will only need their quadratic parts and higher,

$$G^{2+} = -b_\alpha\mathbf{W} + P(R\bar{Y}) + \bar{P}(\bar{R}Y), \qquad K^{2+} = -i\frac{\mathbf{W}^2+a}{1+\mathbf{W}}.$$

Then, we have

$$\Lambda^4\left(\frac{d}{dt}I_1\right) = \int \partial^{j-1}(-b\mathbf{W}_\alpha + G^{2+})\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)} + \mathbf{W}^{(j)}\partial^{k-1}(-b\mathbf{W}_\alpha + G^{2+})\bar{\mathbf{W}}^{(l)} + \mathbf{W}^{(j)}\mathbf{W}^{(k)}\partial^{l-1}(-b\bar{\mathbf{W}}_\alpha + \bar{G}^{2+})\, d\alpha$$

Distributing derivatives, we separate the terms with undifferentiated b as

$$\int -b\partial_\alpha(\mathbf{W}^{(j)}\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)})\, d\alpha = \int b_\alpha \mathbf{W}^{(j)}\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)}\, d\alpha,$$

therefore all terms involving b have the form

$$\int b^{(m)}\mathbf{W}^{(j)}\mathbf{W}^{(k)}\bar{\mathbf{W}}^{(l)}\, d\alpha, \qquad m+j+k+l = 2n-1, \quad 1 \le m \le n-1, \quad 1 \le j,k,l \le n-1,$$

which we can estimate by Hölder's inequality and interpolation, using Lemma 2.7, to get the b bounds

$$\||D|^{\frac{1}{2}}b\|_{BMO} \lesssim_A A, \qquad \||D|^{n-\frac{1}{2}}b\|_{L^2} \lesssim_A \mathbf{N}_n.$$

The remaining terms have the form

$$\int \partial^{j-1}P(R\bar{Y})W^{(k)}\bar{W}^{(l)}d\alpha.$$

These are again estimated by Hölder's inequality and interpolation, using the bounds proved in Lemma 2.7, which show that

$$\||D|^{\frac{1}{2}}P(R\bar{Y})\|_{BMO} \lesssim_A A^2, \qquad \||D|^{n-\frac{1}{2}}P(R\bar{Y})\|_{L^2} \lesssim_A A\mathbf{N}_2.$$

.

The argument for I_2 is similar, using the algebra property of $L^\infty \cap \dot{H}^s$, together with the L^∞ and $\dot{H}^{n-1}$ bound for a in Proposition 2.6 in order to show that

$$\|K^{2+}\|_{BMO} \lesssim_A A^2, \qquad \||D|^{n-1}K^{2+}\|_{L^2} \lesssim_A A\mathbf{N}_2.$$

.

4. Local well-posedness

As the water wave equations (1.1) are fully nonlinear, the standard strategy to prove well-posedness would be to differentiate the equations to turn them into a system of quasilinear equations for $(w,q) := (W_\alpha, Q_\alpha)$, and then apply an iteration scheme. The problem with a direct implementation of this idea is that the quasilinear problem is degenerate, and diagonalizing it requires using the exact equations; thus the diagonalization would fail in an approximation scheme.

To remedy this, we use the form (1.3) of the equations in terms of the diagonal variables $(\mathbf{W}, R)$ directly. Projecting those on the holomorphic space we obtain

$$\text{(4.1)} \qquad \begin{cases} (\partial_t + \mathfrak{M}_b\partial_\alpha)\mathbf{W} + P\left[\dfrac{1+\mathbf{W}}{1+\bar{\mathbf{W}}}R_\alpha\right] = K(\mathbf{W}, R), \\ (\partial_t + \mathfrak{M}_b\partial_\alpha)R - iP\left[\dfrac{(1+a)\mathbf{W}}{1+\mathbf{W}}\right] = K(\mathbf{W}, R), \end{cases}$$

where $K(\mathbf{W}, R) := P\,[(1+\mathbf{W})M]$ and $K(\mathbf{W}, R) := P\,[a]$.

We now turn to the business of solving the system (4.1). The state space for this will be the space $\dot{\mathcal{H}}_n$ endowed with the norm

$$\|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}_n} := \sum_{k=0}^{n} \|\partial_\alpha^k (\mathbf{W}, R)\|_{L^2 \times \dot{H}^{\frac{1}{2}}},$$

where $n \geq 1$. As a preliminary step, we will also consider solutions in the smaller space

$$\mathcal{H}_n = H^n \times H^{n+\frac{1}{2}}, \quad \text{with } n \geq 2.$$

We remark that, given a solution in $\dot{\mathcal{H}}_n$ for the above equation, we already know how to obtain uniform energy estimates for it for $n \geq 1$. The issue at hand is to convert those estimates into a well-posedness statement. We also remark that our energy estimates are expressed in terms of the control norms A and B. These are in turn mostly controlled using the $\dot{\mathcal{H}}_1$ norm of $(\mathbf{W}, R)$. The exception is the L^∞ bound for Y, which, as it turns out, can be bounded in terms of its initial data and the $\dot{\mathcal{H}}_1$ norm of $(\mathbf{W}, R)$.

To better understand the evolution of the $\dot{\mathcal{H}}_1$ norm of the solution it is convenient to use the language of frequency envelopes. We say that a sequence $c_k \in \ell^2$ is a $\dot{\mathcal{H}}_1$ frequency envelope for $(\mathbf{W}, R) \in \dot{\mathcal{H}}^1$ if (i) it is slowly varying, $c_j/c_k \leq 2^{-\delta|j-k|}$ with a small universal constant δ, and (ii) it bounds the dyadic norms of $(\mathbf{W}, R)$, namely $\|P_k(\mathbf{W}, R)\|_{\dot{\mathcal{H}}_1} \leq c_k$.

Our main result here is:

Proposition 4.1. *a) Let $n \geq 1$. Then the problem* (1.1) *is locally well-posed in for initial data $(\mathbf{W}, R)$ in $\dot{\mathcal{H}}_n$.*

b) (lifespan) There exists $T = T(\|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}_1}, \|Y\|_{L^\infty})$ so that the above solutions are well defined in $[0, T]$, with uniform bounds.

c) (frequency envelopes) Given a frequency envelope c_k for the initial data in $\dot{\mathcal{H}}_1$, a similar frequency envelope $C(\|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}_1}, \|Y\|_{L^\infty})c_k$ applies for the solutions in $[0, T]$.

Thereom 1 is a consequence of the above proposition. The statement about the persistence of solutions for as long as A, B remain bounded is a consequence of the energy estimates in Proposition 2.1 and Proposition 3.1, where the constants depend only on A and B.

We remark that the well-posedness result in part (a) carries different meanings depending on n. If $n \geq 2$, then we obtain existence and uniqueness in $C(\dot{\mathcal{H}}_n)$ together with continuous dependence on the data with respect to the stronger $\mathcal{H}_n$ topology. On the other hand if $n = 1$ then we produce rough solutions $C(\dot{\mathcal{H}}_1)$ as the unique strong limit of smooth solutions, with continuous dependence on the data with respect to the stronger $\mathcal{H}_1$ topology. The $\mathcal{H}_1$ continuous dependence is a standard consequence of the strong $\mathcal{H}_n$ continuous dependence on data together with the frequency envelope bounds. However, for $n = 1$, we do not establish a direct uniqueness result.

The proof proceeds in several steps:

4.1. **Existence of regular solutions.** Here we consider data $(\mathbf{W}, R)(0) \in \mathcal{H}_n$ with $n \geq 2$, and prove the existence of solutions in the same space. Our strategy is to obtain approximate solutions by solving the mollified system

$$
(4.2)\quad \begin{cases} (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})\mathbf{W} + P_{<N}P\left[\dfrac{1+P_{<N}\mathbf{W}}{1+P_{<N}\bar{\mathbf{W}}}P_{<N}R_\alpha\right] = P_{<N}G(P_{<N}\mathbf{W}, P_{<N}R), \\ (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})R - iP_{<N}P\left[\dfrac{(1+a_N)P_{<N}\mathbf{W}}{1+P_{<N}\mathbf{W}}\right] = P_{<N}K(P_{<N}\mathbf{W}, P_{<N}R), \end{cases}
$$

where $P_{<N}$ is a multiplier which selects frequencies less than N, and

$$
b_N = b(P_{<N}\mathbf{W}, P_{<N}R), \qquad a_N = a(P_{<N}R).
$$

For fixed N these equations form a system of ordinary differential equations in $\mathcal{H}_n$, which admits a local solution. We can consider it with a single data, or with a one parameter family of data. The latter will help with the dependence of data for our original equation.

We will prove uniform estimates for this evolution in $\mathcal{H}_n$, $n \geq 1$, and then obtain our solution (or one parameter family of solutions) as a weak limit on a subsequence as $N \to \infty$.

The (G, K) terms are Lipschitz, indeed C^1 from $\mathcal{H}_n$ to $\mathcal{H}_n$, therefore harmless. The $\mathcal{H}_{n-1}$ norm of $(\mathbf{W}, R)$ is estimated directly by time integration,

$$
(4.3)\qquad \frac{d}{dt}\|(\mathbf{W}, R)\|^2_{\mathcal{H}_{n-1}} \lesssim c(\|(\mathbf{W}, R)\|^2_{\mathcal{H}_n})\|(\mathbf{W}, R)\|^2_{\mathcal{H}_n}.
$$

It remains to estimate the $\mathcal{H}_0$ norm of $\partial^n_\alpha(\mathbf{W}, R)$. We differentiate the equations (4.2) n times. This yields

$$
(4.4)\quad \begin{cases} (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})\mathbf{W}^{(n)} + P_{<N}P\left[\dfrac{(1+P_{<N}\mathbf{W})}{1+P_{<N}\bar{\mathbf{W}}}\partial_\alpha P_{<N}R^{(n)}\right] = G_n, \\ (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})R^{(n)} - iP_{<N}P\left[\dfrac{(1+a_N)P_{<N}\mathbf{W}^{(n)}}{(1+P_{<N}\mathbf{W})^2}\right] = K_n, \end{cases}
$$

where all other terms, included in G_n and K_n, are estimated directly in $\mathcal{H}_0$ in terms of the $\mathcal{H}^n$ norm of $(\mathbf{W}, R)$. We observe that the fact that we work in $\mathcal{H}_n$ with $n \geq 2$ allows us to use pointwise bounds for R, R_α, b, b_α, and thus deal with a larger number of terms in this fashion.

To bring this to the standard form, where we can apply energy estimates previously obtained in Section 2, we make the substitution

$$
\mathbf{R}^{(n)} := R^{(n)}(1 + P_{<N}\mathbf{W}).
$$

Multiplying in the second equation by $(1 + P_{<N}\mathbf{W})$, all of the commutator terms are also perturbative, and we obtain the system

$$
\begin{cases} (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})\mathbf{W}^{(n)} + P_{<N}P\left[\dfrac{1}{1+P_{<N}\bar{\mathbf{W}}}P_{<N}\mathbf{R}^{(n)}_\alpha\right] = \mathbf{G}_n, \\ (\partial_t + P_{<N}\mathfrak{M}_{b_N}\partial_\alpha P_{<N})\mathbf{R}^{(n)} - iP_{<N}P\left[\dfrac{(1+a_N)P_{<N}\mathbf{W}^{(n)}}{(1+P_{<N}\mathbf{W})}\right] = \mathbf{K}_n, \end{cases}
$$

where $\mathbf{G}_n$ and $\mathbf{K}_n$ are appropriate replacements of the (perturbative) terms in (4.4), G_n and K_n respectively.

For this system we do energy estimates as before, with the energy functional

$$
E^n = \int (1 + a_N)|\mathbf{W}^{(n)}|^2 + Im(\mathbf{R}^{(n)}\partial_\alpha\mathbf{R}^{(n)}) + |\mathbf{R}^{(n)}|^2\, d\alpha.
$$

We obtain

$$\frac{dE^n}{dt} \lesssim c(\|(\mathbf{W},R)\|^2_{\mathcal{H}_n})\|(\mathbf{W},R)\|^2_{\mathcal{H}_n}.$$

We combine this with (4.3). Since

$$\|R\|_{\dot{H}^{\frac{1}{2}}} \lesssim \|\mathbf{R}\|_{\dot{H}^{\frac{1}{2}}}\|Y\|_{L^\infty} + \|\mathbf{R}\|_{L^2}\||D|^{\frac{1}{2}}Y\|_{BMO},$$

we have that

$$\|(\mathbf{W},R)\|^2_{\mathcal{H}_n} \lesssim c(\|(\mathbf{W},R)\|^2_{\mathcal{H}_{n-1}})\left(E^n + \|(\mathbf{W},R)\|^2_{\mathcal{H}_{n-1}}\right),$$

which leads to a bound for our approximate system which is uniform in N,

$$\|(\mathbf{W},R)(t)\|_{\mathcal{H}_n} \lesssim \|(\mathbf{W},R)(0)\|_{\mathcal{H}_n}, \quad 0 \leq t \leq T(\|(\mathbf{W},R)(0)\|_{\mathcal{H}_n}, \|Y(0)\|_{L^\infty}). \tag{4.5}$$

Similarly, one can consider a smooth family of data $(\mathbf{W}_h, R_h)$ in $\mathcal{H}_n$ for $h \in [0,1]$. Then the solutions depend smoothly on h, with a lifespan uniformly bounded from below. We consider the h derivatives $(\tilde{w},\tilde{r}) = \partial_h(\mathbf{W}_h, R_h)$. These solve the linearized equation, which when considered in $\mathcal{H}_{n-1}$, can be written in the same form as (4.4), with perturbative terms on the right. Thus, we obtain

$$\|(\tilde{w},\tilde{r})(t)\|_{\mathcal{H}_{n-1}} \lesssim \|(\tilde{w},\tilde{r})(0)\|_{\mathcal{H}_{n-1}}, \quad 0 \leq t \leq T(\|(\mathbf{W},R)(0)\|_{\mathcal{H}_n}, \|Y(0)\|_{L^\infty}). \tag{4.6}$$

In the same manner one can obtain estimates for the second order derivatives with respect to h in $\mathcal{H}_{n-2}$, *etc.* Passing to a weak limit on a subsequence as $N \to \infty$ we obtain a family of solutions $(\mathbf{W}_h, R_h)$ which is uniformly bounded in $\mathcal{H}_n$, with h derivatives uniformly bounded in $\mathcal{H}_{n-1}$, *etc.*

4.2. **Uniqueness of regular solutions.** In the previous subsection we have constructed $\mathcal{H}_n$ solutions for $n \geq 2$. Here we prove that these solutions are unique. For later use, we show that uniqueness holds in the larger class of $\dot{\mathcal{H}}_n$ solutions for $n \geq 2$.

Suppose we have two $\dot{\mathcal{H}}_2$ solutions $(\mathbf{W}_1, R_1)$ and $(\mathbf{W}_2, R_2)$ to (4.1). Subtracting the two sets of equations we obtain a system for the difference $(\tilde{w},\tilde{r})$, namely

$$\begin{cases} (\partial_t + \mathfrak{M}_{b_1}\partial_\alpha)\tilde{w} + P\left[\dfrac{1+\mathbf{W}_1}{1+\bar{\mathbf{W}}_1}\tilde{r}_\alpha\right] = \tilde{G}, \\ (\partial_t + \mathfrak{M}_{b_1}\partial_\alpha)\tilde{r} - iP\left[\dfrac{(1+a_1)\tilde{w}}{(1+\mathbf{W}_1)^2}\right] = \tilde{K}, \end{cases} \tag{4.7}$$

where

$$\begin{cases} \tilde{G} = G(\mathbf{W}_1, R_1) - G(\mathbf{W}_2, R_2) + \mathfrak{M}_{b_1-b_2}\partial_\alpha \mathbf{W}_2 + P\left[\left(\dfrac{1+\mathbf{W}_1}{1+\bar{\mathbf{W}}_1} - \dfrac{1+\mathbf{W}_2}{1+\bar{\mathbf{W}}_2}\right)\partial_\alpha R_2\right], \\ \tilde{K} = K(\mathbf{W}_1, R_1) - K(\mathbf{W}_2, R_2) + \mathfrak{M}_{b_1-b_2}\partial_\alpha R_2 + iP\left[\dfrac{(1+a_1)\tilde{w}^2}{(1+\mathbf{W}_1)^2(1+\mathbf{W}_2)} + \dfrac{(a_1-a_2)\mathbf{W}_2}{1+\mathbf{W}_2}\right] \end{cases}$$

With implicit constants depending on the $\dot{\mathcal{H}}_2$ solutions $(\mathbf{W}_1, R_1)$ and $(\mathbf{W}_2, R_2)$, we have

$$\|(\tilde{G},\tilde{K})\|_{\mathcal{H}_0} \lesssim \|(\tilde{w},\tilde{r})\|_{\mathcal{H}_0}.$$

Then we simultaneously do energy estimates for $(\tilde{w}, \tilde{r}(1+\mathbf{W}_1))$ in $\mathcal{H}^{\frac{1}{2}} = L^2 \times \dot{H}^{\frac{1}{2}}$ and for R in L^2, and then apply Gronwall's inequality to get $(\tilde{w},\tilde{r}) = (0,0)$.

4.3. $\dot{\mathcal{H}}_1$ **bounds.** The solutions produced above have a lifespan which depends on the $\mathcal{H}_n$ size of data. Here we prove that in effect the lifespan depends only on the $\dot{\mathcal{H}}_1$ size of data, and that we have uniform bounds for as long as the $\dot{\mathcal{H}}_1$ size of the solutions is controlled.

Precisely, suppose we have an $\mathcal{H}_n$ solution $(\mathbf{W}, R)$ which satisfies the bounds

$$\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}_1} < \mathcal{M}_0, \qquad \|Y(0)\|_{L^\infty} < \mathcal{K}_0.$$

Then we claim that there exists $T = T(\mathcal{M}_0, \mathcal{K}_0)$ so that the solution exists in $[0, T]$ and satisfies the bounds

$$\|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}_1)} < \mathcal{M}(\mathcal{M}_0, \mathcal{K}_0), \qquad \|Y\|_{L^\infty([0,T]\times\mathbf{R})} < \mathcal{K}(\mathcal{M}_0, \mathcal{K}_0), \tag{4.8}$$

as well as the $\mathcal{H}_n$ and $\dot{\mathcal{H}}_n$ bounds

$$\|(\mathbf{W}, R)\|_{L^\infty(0,T;\mathcal{H}_n)} \leq C(\mathcal{M}_0, \mathcal{K}_0)\|(\mathbf{W}, R)(0)\|_{\mathcal{H}_n},$$
$$\|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}_n)} \leq C(\mathcal{M}_0, \mathcal{K}_0)\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}_n}.$$

To prove this, we begin by making the bootstrap assumption

$$\|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}_1)} < 2\mathcal{M}, \qquad \|Y\|_{L^\infty([0,T]\times\mathbf{R})} < 2\mathcal{K}.$$

We will show that for a suitable choice $\mathcal{M}(\mathcal{M}_0, \mathcal{K}_0)$ and $\mathcal{K}(\mathcal{M}_0, \mathcal{K}_0)$, depending only on $\mathcal{M}_0$ and $\mathcal{K}_0$, we can improve this to (4.8), provided that $T < T(\mathcal{M}_0, \mathcal{K}_0)$.

We begin by applying the linearized energy estimates obtained in Proposition 2.1 to $(\mathbf{W}, R)$

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_0} \lesssim e^{Ct}\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}_0}, \qquad C = C(\mathcal{M}, \mathcal{K}). \tag{4.9}$$

Applying the energy estimates proven in Proposition 3.1 (ii) for the pair $(\mathbf{W}_\alpha, (1+\mathbf{W})R_\alpha)$ we get

$$\|(\mathbf{W}_\alpha, (1+\mathbf{W})R_\alpha)(t)\|_{\mathcal{H}_0} \lesssim e^{Ct}, \|(\mathbf{W}_\alpha, (1+\mathbf{W})R_\alpha)(0)\|_{\dot{\mathcal{H}}_0}. \tag{4.10}$$

To combine (4.9) and (4.10) we need to invert $1 + \mathbf{W}$. However a brute force argument introduces a constant which depends on both $\mathcal{K}$ and $\mathcal{M}$, which wreaks havoc with our bootstrap. Instead we do a more careful argument, using the pair of bounds

$$\begin{aligned} \|(1+\mathbf{W})R_\alpha\|_{\dot{H}^{\frac{1}{2}}} &\lesssim_{\mathcal{K}} \|R_\alpha\|_{\dot{H}^{\frac{1}{2}}} + \|\mathbf{W}_\alpha\|_{L^2}\||D|^{\frac{1}{2}}R\|_{L^\infty}, \\ \|R_\alpha\|_{\dot{H}^{\frac{1}{2}}} &\lesssim_{\mathcal{K}} \|(1+\mathbf{W})R_\alpha\|_{\dot{H}^{\frac{1}{2}}} + \|\mathbf{W}_\alpha\|_{L^2}\||D|^{\frac{1}{2}}R\|_{L^\infty}. \end{aligned} \tag{4.11}$$

Since

$$\||D|^{\frac{1}{2}}R\|_{L^\infty}^2 \lesssim_{\mathcal{K}} \|R\|_{\dot{H}^{\frac{1}{2}}}\|R_\alpha\|_{\dot{H}^{\frac{1}{2}}},$$

we obtain

$$\||D|^{\frac{1}{2}}R\|_{L^\infty}^2 \lesssim_{\mathcal{K}} \mathcal{M}_0^2 e^{2Ct}(1 + \||D|^{\frac{1}{2}}R\|_{L^\infty}),$$

so

$$\||D|^{\frac{1}{2}}R\|_{L^\infty} \leq C_0\mathcal{M}_0^2 e^{2Ct}, \qquad C_0 = C_0(\mathcal{K}).$$

Then it follows that

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_1} \leq C_0\mathcal{M}_0^3 e^{3Ct}. \tag{4.12}$$

Since $\mathcal{M}$ appears only in the exponent where it is controlled by choosing t small, the bound (4.12) suffices in order to bootstrap $\mathcal{M}$. It remains to recover the bootstrap assumption on $\|Y\|_{L^\infty}$. For this we use an estimate of the form

$$\|Y\|_{L^\infty}^2 \lesssim \|\mathbf{W}_\alpha\|_{L^2}\|\mathbf{W}(1+\mathbf{W})^{-3}\|_{L^2}.$$

The bound for the first factor is independent of $\mathcal{K}$. For the second we write the transport equation

$$(\partial_t + b\partial_\alpha)\frac{\mathbf{W}}{(1+\mathbf{W})^3} = \frac{3-2\mathbf{W}}{(1+\mathbf{W})^3}\left([P, W_\alpha]\frac{\bar{R}}{(1+\mathbf{W})^2} - P\left[\frac{R}{1+\bar{\mathbf{W}}}\right]_\alpha\right).$$

We can estimate the right hand side in L^2 with constants depending on $\mathcal{K}$. To bound $\dfrac{\mathbf{W}}{(1+\mathbf{W})^3}$ in L^2 we use an estimate of the form

$$\frac{d}{dt}\|u\|_{L^2}^2 = \int_{\mathbb{R}} b_\alpha |u|^2 + 2\Re(\partial_t + b\partial_\alpha)u\bar{u}\, d\alpha.$$

For the second term on the right we use the Cauchy-Schwarz inequality and for the first term we use a Littlewood-Paley trichotomy. When the frequency of b_α is strictly less than the frequencies of u and $\bar{u}$ then we can move half of derivative on either of u or $\bar{u}$; otherwise Coiman-Meyer type estimates apply, and we obtain

$$|\int_R b_\alpha |u|^2\, d\alpha| \lesssim \|b_\alpha\|_{BMO}\|u\|_{L^2}^2 + \||D|^{\frac{1}{2}}b\|_{BMO}\|u\|_{\dot{H}^{\frac{1}{2}}}\|u\|_{L^2}.$$

We conclude that

$$\frac{d}{dt}\|u\|_{L^2} \lesssim \|b_\alpha\|_{BMO}\|u\|_{L^2}^2 + \||D|^{\frac{1}{2}}b\|_{BMO}\|u\|_{\dot{H}^{\frac{1}{2}}}\|u\|_{L^2} + \|u\|_{L^2}\|(\partial_t + b\partial_\alpha)u\|_{L^2}. \tag{4.13}$$

We apply this estimate to $\dfrac{\mathbf{W}}{(1+\mathbf{W})^3}(t)$ to obtain

$$\|\frac{\mathbf{W}}{(1+\mathbf{W})^3}(t)\|_{L^2} \leq \|\frac{\mathbf{W}}{(1+\mathbf{W})^3}(0)\|_{L^2} + tC(\mathcal{K}, \mathcal{M}).$$

This leads to

$$\|Y\|_{L^\infty}^2 \lesssim \mathcal{M}_0\mathcal{K}_0^3 + tC(\mathcal{K}, \mathcal{M}).$$

Hence in order for our bootstrap argument to succeed we need to find $\mathcal{K}, \mathcal{M}$ and T so that

$$\mathcal{M} > 2C_0(\mathcal{K})\mathcal{M}_0^3 e^{C(\mathcal{K},\mathcal{M})T}, \qquad \mathcal{K}^2 > 2(\mathcal{M}_0\mathcal{K}_0^3 + tC(\mathcal{K}, \mathcal{M})).$$

This is easily achieved by succesively choosing

$$\mathcal{K}^2 = 10\mathcal{M}_0\mathcal{K}_0^3, \qquad \mathcal{M} = 10C_0(\mathcal{K})\mathcal{M}_0^3, \qquad T < C(\mathcal{K}, \mathcal{M})^{-1}.$$

Thus, the bootstrap is complete.

The next step is to show that we can propagate the full $\mathcal{H}_n$ norm given control of $\dot{\mathcal{H}}_1$ norm of the solution $(\mathbf{W}, R)$. For higher derivatives we can use Proposition 3.1 to obtain

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_n} \leq Ce^{Ct}\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_n}, \qquad C = C(\mathcal{K}, \mathcal{M}). \tag{4.14}$$

We also need to control the growth of the L^2 norm of R; for this we use equation 1.3 for which we can easily obtain L^2 bounds of the RHS. Applying (4.13) we obtain

$$\|R(t)\|_{L^2} \leq \|R(0)\|_{L^2} + tC(\mathcal{K}, \mathcal{M}).$$

The $\mathcal{H}_n$ bound shows that the solution can be continued for as long as it stays bounded in $\dot{\mathcal{H}}_1$, i.e., at least until time $T(\mathcal{K}_0, \mathcal{M}_0)$.

4.4. $\dot{\mathcal{H}}_n$ **solutions for** $n \geq 2$ **.** Our goal here is to obtain solutions for $\dot{\mathcal{H}}_n$ data. We already know that such solutions, if they exist, are unique. The idea is to approximate a $\dot{\mathcal{H}}_n$ data set $(\mathbf{W}, R)(0)$ with $\mathcal{H}_n$ data in the $\dot{\mathcal{H}}_n$ topology. As the uniform $\dot{\mathcal{H}}_n$ bounds hold uniformly for the approximating sequence, we would like to conclude that on a subsequence these approximate solutions converge weakly to the desired solution. The only difficulty with this plan is that the $\dot{\mathcal{H}}_n$ convergence does not guarantee uniform pointwise convergence for R. This is because the lowest Sobolev norm we control for R is the $\dot{H}^{\frac{1}{2}}$ norm, and that does not see constants.

To address the above difficulty, we take an approximating sequence of data $(\mathbf{W}_k, R_k)(0)$ which has the following two properties:

(i) $(\mathbf{W}_k, R_k)(0) \to (\mathbf{W}, R)(0)$ in $\dot{\mathcal{H}}_n$,

(ii) $R_k(0) \to R(0)$ uniformly on compact sets.

The second requirement effectively removes the Galilean invariance. It suffices to ask for pointwise convergence at a single point; in view of the known average growth rates for BMO functions, this implies the weighted uniform convergence

$$\| \log(2 + |\alpha|)^{-1} R_k(0) - R(0) \|_{L^\infty} \to 0.$$

We will use the second requirement (ii) to produce weighted uniform bounds for the R_k part of the solution. Starting from the uniform bound

$$\|(\mathbf{W}_k, R_k)\|_{\dot{\mathcal{H}}_n} \lesssim 1,$$

we estimate uniformly most of the terms in the R_k equation to obtain

$$\|(\partial_t + 2\Re R_k \partial_\alpha) R_k\|_{L^\infty} \lesssim 1.$$

This yields a uniform bound for R_k along the corresponding characteristic

$$\dot{\alpha}(t) = 2\Re R_k(\alpha), \qquad \alpha(0) = 0,$$

namely

$$|R_k(t, \alpha(t))| \lesssim 1.$$

This in turn shows that locally in time we have

$$|\alpha(t)| \lesssim 1,$$

which leads to the uniform bound

$$|R_k(t, 0)| \lesssim 1,$$

and further to the global bound

$$|R_k| \lesssim \log(2 + |\alpha|).$$

This in turn yields a similar bound for $\partial_t \mathbf{W}_k$ and $\partial_t R_k$, and suffices in order to insure local uniform convergence of $(\mathbf{W}_k, R_k)$ on a subsequence. Thus, the desired solution $(\mathbf{W}, R)$ is obtained in the limit.

4.5. **Rough solutions.** Here we construct solutions for data in $\dot{\mathcal{H}}_1$ as unique limits of smooth solutions. Given a $\dot{\mathcal{H}}_1$ initial data $(\mathbf{W}_0, R_0)$ as above we regularize it to produce smooth approximate data $(\mathbf{W}_0^k, R_0^k) = P_{<k}(\mathbf{W}_0, R_0)$. We denote the corresponding solutions by $(\mathbf{W}^k, R^k)$. By the previous analysis, these solutions exist on a k-independent time interval $[0, T]$ and satisfy uniform $\dot{\mathcal{H}}_1$ bounds. Further, they are smooth and have a smooth dependence on k.

Consider a frequency envelope c_k for the initial data $(\mathbf{W}_0, R_0)$ in $\dot{\mathcal{H}}_1$. Then for the regularized data we have

$$\|(\mathbf{W}_0^k, R_0^k)\|_{\mathcal{H}_n} \lesssim c_k 2^{(n-1)k}, \qquad n \geq 2.$$

Hence, in the time interval $[0, T]$ we also have the estimates

$$\|(\mathbf{W}^k, R^k)\|_{\mathcal{H}_n} \lesssim c_k 2^{(n-1)k}, \qquad n \geq 2. \tag{4.15}$$

We will use these for the high frequency part of the regularized solutions.

For the low frequency part, on the other hand, we view k as a continuous rather than a discrete parameter, differentiate $(\mathbf{W}^k, R^k)$ with respect to k and use the estimates for the linearized equation. One minor difficulty is that the linearized equation (2.1) arises from the linearization of the (W, Q) system in (1.1) rather than the differentiated $(\mathbf{W}, R)$ system in (1.3). Assuming that (W^k, Q^k) were also defined, we formally denote

$$(w^k, r^k) = (\partial_k W^k, \partial_k Q^k - R\partial_k W^k).$$

These would solve the linearized equation around the $(\mathbf{W}^k, R^k)$ solution. For our analysis we want to refer only to the differentiated variables, so we we compute

$$\begin{aligned} \partial_\alpha w^k &= \partial_k W^k, \\ \partial_\alpha r^k &= (1 + \mathbf{W}^k)\partial_k R^k - R_\alpha^k w^k. \end{aligned}$$

We take these formulas as the definition of (w^k, r^k), and observe that inverting the ∂_α operator is straightforward since the above multiplications involve only holomorphic factors therefore the products are at frequency 2^k and higher. To take advantage of the bounds in Proposition 2.1 for the linearized equation, we need a $\dot{\mathcal{H}}_0$ bound for $(w^k(0), r^k(0))$, namely

$$\|(w^k(0), r^k(0))\|_{\dot{\mathcal{H}}_0} \lesssim c_k 2^{-2k}. \tag{4.16}$$

The bound for $w^k(0)$ is straightforward, but some work is required for $r^k(0)$. This follows via the usual Littlewood-Paley trichotomy and Bernstein's inequality for the low frequency factor, with the twist that, since both factors are holomorphic, no high-high to low interactions occur.

In view of the uniform $\dot{\mathcal{H}}_1$ bound for (W^k, Q^k), Proposition 2.1 shows that in $[0, T]$ we have the uniform estimate

$$\|(w^k, r^k)\|_{\dot{\mathcal{H}}_0} \lesssim c_k 2^{-2k}. \tag{4.17}$$

Now, we return to $(\mathbf{W}^k, R^k)$ and claim the bound

$$\|P_{\leq k}(\partial_k \mathbf{W}^k, \partial_k R^k)\|_{\dot{\mathcal{H}}_0} \lesssim c_k 2^{-k}. \tag{4.18}$$

Again the $\mathbf{W}^k$ bound is straightforward. For $\partial_k R^k$ we write

$$\partial_k R^k = (1 - Y^k)(\partial_\alpha r^k + R_\alpha^k \partial_k W^k),$$

where again all factors are holomorphic. Then applying $P_{\leq k}$ restricts all frequencies to $\lesssim 2^k$, and the Littlewood-Paley trichotomy and Bernstein's inequality again apply.

Now we integrate (4.18) over unit k intervals and use it to estimate the differences. Combining the result with (4.15) we obtain

$$\begin{aligned} \|(\mathbf{W}_{k+1} - \mathbf{W}_k, R_{k+1} - R_k)\|_{\dot{\mathcal{H}}_0} &\lesssim c_k 2^{-k}, \\ \|\partial_\alpha^2(\mathbf{W}_{k+1} - \mathbf{W}_k, R_{k+1} - R_k)\|_{\dot{\mathcal{H}}_0} &\lesssim c_k 2^{k}. \end{aligned} \tag{4.19}$$

Summing up with respect to k it follows that the sequence $(\mathbf{W}^k, R^k)$ converges uniformly in $\dot{\mathcal{H}}_1$ to a solution $(\mathbf{W}, R)$, which also inherits the frequency envelope bounds from the data.

The frequency envelope bounds allow us to prove continuous dependence on the initial data in the $\mathcal{H}^1$ topology. This is standard, but we briefly outline the argument. Suppose that $(\mathbf{W}_j, R_j)(0) \in \dot{\mathcal{H}}_1$ and $(\mathbf{W}_j, R_j)(0) - (\mathbf{W}, R)(0) \to 0$ in $\mathcal{H}_1$. We consider the approximate solutions $((\mathbf{W}_j^k, R_j^k)$, respectively $(\mathbf{W}^k, R^k)$. According to our result for more regular solutions, we have

$$((\mathbf{W}_j^k, R_j^k) - (\mathbf{W}^k, R^k) \to 0 \qquad \text{in } \ \mathcal{H}_n. \tag{4.20}$$

On the other hand, from the $\mathcal{H}_1$ data convergence we get

$$(\mathbf{W}_j^k, R_j^k)(0) - (\mathbf{W}_j, R_j)(0) \to 0 \qquad \text{in } \ \mathcal{H}_1 \ \text{ uniformly in } \ j.$$

Then the above frequency envelope analysis, shows that

$$(\mathbf{W}_j^k, R_j^k) - (\mathbf{W}_j, R_j) \to 0 \qquad \text{in } \ \mathcal{H}_1 \quad \text{uniformly in } \ j.$$

Hence we can let k go to infinity in (4.20) and conclude that

$$((\mathbf{W}_j, R_j) - (\mathbf{W}, R) \to 0 \qquad \text{in } \ \mathcal{H}_1.$$

5. Enhanced cubic lifespan bounds

In this section we prove Theorem 2. Given initial data $(\mathbf{W}, R)$ for (1.3) satisfying

$$\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}_1} \leq \epsilon,$$

we consider the solutions on a time interval $[0, T]$ and seek to prove the estimate

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_1} \leq C\epsilon, \quad t \in [0, T], \tag{5.1}$$

provided that $T \ll e^{-2}$. In view of our local well-posedness result this shows that the solutions can be extended up to time $T_\epsilon = ce^{-2}$ concluding the proof of the theorem.

In order to prove (5.1) we can harmlessly make the bootstrap assumption

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}_1} \leq 2C\epsilon, \quad t \in [0, T]. \tag{5.2}$$

From (5.2) we obtain the bounds

$$A, B \lesssim C\epsilon.$$

Hence, by the energy estimates in Proposition 2.2 applied to $(\mathbf{W}, R)$, and those in Proposition 3.2, with $n = 2$, applied to $(\mathbf{W}_\alpha, R_\alpha)$ we obtain

$$\|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}_1)} \lesssim \|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}_1} + TAB\|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}_1)} \lesssim \epsilon + TC^3\epsilon^3.$$

Hence, the desired estimate (5.1) follows provided that $T \ll (C\epsilon)^{-2}$.

6. Pointwise decay and long time solutions

In this section we prove the almost global existence result in Theorem 3. This is achieved via a bootstrap argument for the energy norm $\|(W,Q)(t)\|_{\mathcal{WH}}$ defined in (1.17) as well as the control norms $A(t)$ and $B(t)$ in (1.11),(1.12). In order to have a more robust argument we will work with a stronger norm $\|(W,R)\|_X \gtrsim A(t)+B(t)$, namely

$$\|(W,R)\|_X = \|W\|_{L^\infty} + \|R\|_{L^\infty} + \||D|^{\frac{1}{2}} W_\alpha\|_{L^\infty} + \|R_\alpha\|_{L^\infty}$$

Then we will establish the energy estimates

$$\sup_{|t|\leq T_\epsilon} \|(W,Q)(t)\|_{\mathcal{WH}} \lesssim \epsilon, \tag{6.1}$$

as well as the pointwise bounds

$$\|(W,R)\|_X \lesssim \epsilon \langle t\rangle^{-\frac{1}{2}}, \qquad |t|\leq T_\epsilon, \tag{6.2}$$

for times T_ϵ satisfying

$$T_\epsilon \leq e^{c\epsilon^{-2}}, \qquad c\ll 1. \tag{6.3}$$

A continuity argument based on our local well-posedness results shows that it suffices to prove that (6.2) and (6.1) hold for all T_ϵ as in (6.3), given the bootstrap assumptions

$$\sup_{|t|\leq T_\epsilon} \|(W,Q)(t)\|_{\mathcal{WH}} \leq C\epsilon, \tag{6.4}$$

$$\|(W,R)\|_X \leq C\epsilon \langle t\rangle^{-\frac{1}{2}}, \qquad 0\leq t\leq T_\epsilon, \tag{6.5}$$

with a large constant C (independent of ϵ).

6.1. The energy estimates in (6.1).

Here we use the bootstrap assumption (6.5) in order to establish (6.4). The only role of (6.4) is to insure that a solution with appropriate regularity exists up to time T_ϵ. We summarize the result in the following

Proposition 6.1. *Assume that in a time interval* $[-T,T]$ *we have a solution* (W,Q) *to* (1.1) *which satisfies* (1.18) *and* (6.5). *Then we also have the energy estimate*

$$\|(W,Q)(t)\|^2_{\mathcal{WH}} \lesssim \epsilon \langle t\rangle^{C_1\epsilon^2}, \qquad t\in[-T,T] \tag{6.6}$$

for some $C_1 \gg C$.

Then the bound (6.1) holds with a constant independent of C for times as in (6.3) if we choose $c = C_1^{-1}$.

Proof. The energy bound for (W,Q) is a consequence of the conserved energy (1.10). The energy bounds for $(\mathbf{W},R)$ and $(w,r) := \mathbf{AS}(W,Q)$ follow by Gronwall's inequality from the cubic energy estimates for the linearized equation in Proposition 2.2; indeed, by our bootstrap assumption (6.5) we have $A(t),B(t)\leq C\epsilon\langle t\rangle^{-\frac{1}{2}}$, therefore,

$$\|(w,r)(t)\|_{\dot{\mathcal{H}}_0} \lesssim e^{\int_0^t C^2\epsilon^2\langle s\rangle^{-1} ds} \|(w,r)(0)\|_{\dot{\mathcal{H}}_0} \lesssim \epsilon e^{C^2\epsilon^2 \log t},$$

which suffices for T_ϵ as in (6.3). Finally, the bound for $\partial^k(\mathbf{W},R)$ with $1\leq k\leq 5$ follows also by Grönwall's inequality from the cubic energy estimates in Proposition 3.2. ∎

6.2. **The pointwise estimates.** Here we use the bootstrap assumption (6.4) in order to establish (6.2). To state the main result here we introduce the notation

$$\omega(t,\alpha) = \frac{1}{\langle t\rangle^{\frac{1}{22}}} + \frac{1}{(\langle \alpha\rangle/\langle t\rangle + \langle t\rangle/\langle \alpha\rangle)^{\frac12}} \lesssim 1. \tag{6.7}$$

Then we have:

Proposition 6.2. *Assume that* (6.6) *and* (6.5) *hold in some interval* $[-T,T]$. *Then we also have*

$$|W| + |R| + ||D|^{\frac12} W_\alpha| + |R_\alpha| \lesssim \epsilon \langle t\rangle^{-\frac12} \langle t\rangle^{C_1\epsilon^2} \omega(t,\alpha) \tag{6.8}$$

Then our pointwise bound (6.2) follows for times as in (6.3), and the proof of Theorem 3 is concluded.

We remark that the result we prove here is somewhat stronger than what we need. However, on one hand this is what follows from our analysis, and on the other hand this stronger result will come in handy when we prove the global result in a follow-up paper.

The rest of this section is devoted to the proof of the above proposition. We note that (6.3) plays no role in this argument.

In order to obtain pointwise bounds it is convenient to work with the normal form variables $(\tilde W, \tilde Q)$, given by (1.13). Then we prove several very simple Lemmas. The first one shows that we can harmlessly replace (W,R) by $(\tilde W, \tilde Q_\alpha)$ in the pointwise estimates.

Lemma 6.3. *Assume that* (6.6) *and* (6.5) *hold in some interval* $[-T,T]$. *Then*

$$\|(W - \tilde W, R - \tilde Q_\alpha)\|_X \lesssim \langle t\rangle^{-\frac18} \|(\tilde W, \tilde Q_\alpha)\|_X \tag{6.9}$$

Proof. It suffices to show that

$$\|(\tilde W - W, \tilde Q_\alpha - R)\|_X \lesssim \langle t\rangle^{-\frac18} \|(W,R)\|_X$$

For the $\tilde W$ bound we have $W - \tilde W = \mathfrak{M}_{\Re W} W_\alpha$ so we use Sobolev embeddings, product Sobolev bounds and interpolation to estimate

$$\begin{aligned}
\|\mathfrak{M}_{\Re W} W_\alpha\|_{L^\infty} + \||D|^{\frac32}(\mathfrak{M}_{\Re W} W_\alpha)\|_{L^\infty} &\lesssim \|\Re W W_\alpha\|_{L^4} + \|D^2(\Re W W_\alpha)\|_{L^4} \\
&\lesssim \|D^2 W\|_{L^3} \|W_\alpha\|_{L^\infty} + \|W\|_{L^\infty}(\|W_\alpha\|_{L^4} + \|D^3 W\|_{L^4}) \\
&\lesssim (\|W\|_{L^\infty} + \|W_\alpha\|_{L^\infty}) A(t)^{\frac12} \|W\|_{H^5}^{\frac12} \\
&\lesssim C^{\frac12} \epsilon^2 \langle t\rangle^{-\frac14 + C_1\epsilon^2} \|(W,R)\|_X.
\end{aligned}$$

which suffices since ϵ is small.

For the R bound we write

$$\tilde Q_\alpha - R = W_\alpha R - 2\partial_\alpha(\mathfrak{M}_{\Re W} R)$$

and a similar argument as above applies.

□

Our second lemma translates the energy bounds to $(\tilde W, \tilde Q)$:

Lemma 6.4. *Assume that* (6.6) *and* (6.5) *hold in some time interval* $[-T,T]$. *Then*

$$\|(\tilde W, \tilde Q)\|_{\dot{\mathcal H}_5} + \|\mathbf{S}(\tilde W, \tilde Q)\|_{\dot{\mathcal H}_0 + \dot{\mathcal H}_{-1}} \lesssim \epsilon \langle t\rangle^{C_1\epsilon^2}. \tag{6.10}$$

Proof. For $\tilde{W}$ we estimate the quadratic terms

$$\begin{aligned}\|\Re W W_\alpha\|_{L^2} + \|\partial^5(\Re W W_\alpha)\|_{L^2} &\lesssim \|W\|_{L^\infty}(\|W_\alpha\|_{L^2} + \|\partial^5 W_\alpha\|_{L^2}) + \|W_\alpha\|_{L^\infty}\|\partial^5 W\|_{L^2} \\ &\lesssim \epsilon(\|W_\alpha\|_{L^2} + \|\partial^6 W\|_{L^2}),\end{aligned}$$

By interpolation and Sobolev embeddings we can combine (6.6) and (6.5) to obtain the rough bound

$$\|W\|_{L^\infty} + \|R\|_{L^\infty} \lesssim \epsilon,$$

which we will use to supplement (6.5). For $\tilde{Q}$ we first bound the quadratic term in $\dot{H}^{\frac{1}{2}}$,

$$\|\Re W R\|_{\dot{H}^{\frac{1}{2}}} \lesssim \|W\|_{L^\infty}\|R\|_{\dot{H}^{\frac{1}{2}}} + \|W\|_{\dot{H}^{\frac{1}{2}}}\|R\|_{L^\infty} \lesssim \epsilon(\|R\|_{\dot{H}^{\frac{1}{2}}} + \|W\|_{\dot{H}^{\frac{1}{2}}}).$$

For higher derivatives we write

$$\tilde{Q}_\alpha = R(1 + W_\alpha) - 2\partial_\alpha(\mathfrak{M}_{\Re W} R)$$

and apply the same method.

The goal of the reminder of the proof is to prove that

$$\|S(\tilde{W}, \tilde{Q})\|_{\dot{\mathcal{H}}_0 + \dot{\mathcal{H}}_{-1}} \lesssim \epsilon\|S(W, R)\|_{\dot{\mathcal{H}}_0} \tag{6.11}$$

Recalling the notation $(w, r) = (SW, SQ - RSW)$, we first write $S\tilde{W}$ as

$$S\tilde{W} = w + 2P[\Re w W_\alpha - \Re W_\alpha w - 2\Re W W_\alpha + 2P\partial_\alpha[\mathfrak{M}_{2\Re W} w],$$

and use the L^2 bound on SW to estimate all but the last term in L^2, and the last term in $\dot{H}^{-1}$. Finally, for $S\tilde{Q}$ we have

$$S\tilde{Q} = SQ - \mathfrak{M}_{2\Re SW} R - \mathfrak{M}_{2\Re W} SR = r + Rw - 2P(\Re w R) - 2P\left[\frac{\Re W(r_\alpha + R_\alpha w)}{1 + W_\alpha}\right]$$

Here it suffices to estimate the contribution of w in L^2, while the contribution of r_α is estimated by

$$\|r_\alpha H\|_{H^{-\frac{1}{2}}} \lesssim \|r\|_{\dot{H}^{\frac{1}{2}}}(\|H\|_{L^\infty} + \||D|^{\frac{1}{2}} H\|_{BMO}), \qquad H := \frac{\Re W}{1 + W_\alpha},$$

where $\||D|^{\frac{1}{2}} H\|_{BMO}$ is estimated using (B.16), (B.17). The proof of (6.11) is concluded. □

The advantage of working with $(\tilde{W}, \tilde{Q})$ is that they solve an equation with a cubic nonlinear term, namely (1.14), where the nonlinearities $\tilde{G}$ and $\tilde{K}$ are given by (1.15). They involve second order derivatives of W and Q, which is why one cannot simply use the above equations as the main evolution.

Lemma 6.5. *Assume that* (6.1) *and* (6.5) *hold in some time interval* $[-T, T]$. *Then*

$$\|(\tilde{G}, \tilde{K})\|_{\dot{\mathcal{H}}_0} \lesssim \frac{\epsilon^3}{\langle t\rangle}\langle t\rangle^{C_1\epsilon^2}. \tag{6.12}$$

Proof. Given the expression above for $\tilde{G}$, it suffices to bound each factor in each term in suitable L^p norms, interpolating between the L^2 norms in (6.1) and the L^∞ norms in (6.5). For $\tilde{K}$ the argument is similar, but we also need to use Lemma 2.4 in the Appendix B in order to distribute the half derivative. □

Taking into account the correspondence, established in the last three lemmas, between the original variables (W, Q) and the normal form variables $(\tilde{W}, \tilde{Q})$, it follows that we can restate Proposition 6.2 in the following linear form:

Proposition 6.6. *Suppose* $(\tilde{W}, \tilde{Q})$ *solve* (1.14) *and that the following bounds hold at some time* t*:*

$$\|\mathbf{S}(\tilde{W}, \tilde{Q})\|_{\dot{\mathcal{H}}_0+\dot{\mathcal{H}}_{-1}} + \|(\tilde{W}, \tilde{Q})\|_{\dot{\mathcal{H}}_5} \lesssim 1,$$

$$\|(\tilde{G}, \tilde{K})\|_{\dot{\mathcal{H}}_0} \lesssim \langle t\rangle^{-1}.$$

Then

$$|\tilde{W}| + ||D|^{\frac{1}{2}}\tilde{Q}| + |D^2\tilde{W}| + ||D|^{\frac{5}{2}}\tilde{Q}| \lesssim \langle t\rangle^{-\frac{1}{2}}\omega(t, \alpha). \tag{6.13}$$

Combining the scaling bound with the equation (1.14) we are led to a system of the form

$$\begin{cases} 2\alpha\partial_\alpha \tilde{W} + t\partial_\alpha \tilde{Q} = \tilde{G}_1 := S\tilde{W} - \tilde{G}, \\ 2\alpha\partial_\alpha \tilde{Q} - it\tilde{W} = \tilde{K}_1 := S\tilde{Q} - \tilde{K}. \end{cases} \tag{6.14}$$

where

$$\|(\tilde{G}_1, \tilde{K}_1)\|_{\dot{\mathcal{H}}_0+\dot{\mathcal{H}}_{-1}} \lesssim 1. \tag{6.15}$$

From here on, all our analysis is at fixed t.

After the substitution

$$(w, r) = (\tilde{W}, |D|^{\frac{1}{2}}\tilde{Q}), \qquad (g, k) = (\tilde{G}_1, |D|^{\frac{1}{2}}\tilde{K}_1 + |D|^{\frac{1}{2}}\tilde{Q}),$$

the above system is written in a more symmetric form as

$$\begin{cases} 2\alpha\partial_\alpha w - it|D|^{\frac{1}{2}} r = g, \\ 2\alpha\partial_\alpha r - it|D|^{\frac{1}{2}} w = k. \end{cases} \tag{6.16}$$

For this it suffices to establish the following result:

Lemma 6.7. *The following pointwise bounds hold for solutions to* (6.16)*:*

$$|w| + |r| \lesssim |\alpha|^{-\frac{1}{2}}(\|(w, r)\|_{L^2} + \|(g, k)\|_{L^2}), \tag{6.17}$$

$$|w| + |r| \lesssim \langle t\rangle^{-\frac{1}{2}}\left(\frac{1}{\langle t\rangle^{\frac{1}{4}}} + \frac{1}{(\langle\alpha\rangle/\langle t\rangle + \langle t\rangle/\langle\alpha\rangle)^{\frac{1}{2}}}\right)(\|(w, r)\|_{H^2} + \|(g, k)\|_{H^{-1}}), \tag{6.18}$$

$$|\partial^k w| + |\partial^k r| \lesssim \langle t\rangle^{-\frac{1}{2}}\left(\frac{1}{\langle t\rangle^{\frac{1}{2(4k+5)}}} + \frac{1}{(\langle\alpha\rangle/\langle t\rangle + \langle t\rangle/\langle\alpha\rangle)^{\frac{1}{2}}}\right)(\|(w, r)\|_{H^{2k+2}} + \|(g, k)\|_{H^{-1}}). \tag{6.19}$$

The last part is applied with $k \leq \frac{3}{2}$, which justifies the exponent $\frac{1}{22}$ in the definition (6.7) of $\omega(t, \alpha)$.

Proof. Without any loss in generality we assume that $|t| \geq 1$. It is convenient to work with frequency localized versions of (6.16), at frequency -2^ℓ, with $\ell \in \mathbb{Z}$. The localized dyadic portions (w_ℓ, r_ℓ) solve similar equations with frequency localized right hand sides (g_ℓ, k_ℓ). Further, a straightforward commutator estimate shows that

$$\sum_{\ell\in\mathbb{Z}} \|(g_\ell, k_\ell)\|_{H^s}^2 \lesssim \|(w,r)\|_{H^s}^2 + \|(g,k)\|_{H^s}^2. \tag{6.20}$$

To prove (6.17) we observe that the system (6.16) is elliptic away from frequency $2^\ell \approx t^2\alpha^{-2}$ and degenerate at frequency zero. At frequencies less than α^{-1} our source for the pointwise estimate is Bernstein's inequality. Comparing the two frequencies yields the threshold $\alpha = t^2$, $2^\ell = t^{-2}$. Thus, we distinguish the following regions:

Case A: $2^\ell \leq t^{-2}$. We group all such frequencies together. We can harmlessly discard the $it|D|^{\frac{1}{2}}$ term from the equations and compute

$$\frac{d}{dt}|w(\alpha)|^2 = 2\Re(\bar{w}w_\alpha),$$

and similarly for r. Depending on the sign of α we integrate from either $+\infty$ or $-\infty$ and apply the Cauchy-Schwartz inequality to obtain

$$|w_{<t^{-2}}|^2 \lesssim |\alpha|^{-1}\|w_{<t^{-2}}\|_{L^2}\|\alpha w_{<t^{-2},\alpha}\|_{L^2}.$$

We remark that by Bernstein's inequality we also get

$$|w_{<t^{-2}}| \lesssim |t|^{-1}\|w_{<t^{-2}}\|_{L^2}.$$

It follows that

$$|w_{<t^{-2}}| \lesssim (t+|\alpha|)^{-1}.$$

Case B. $t^{-2} \lesssim 2^\ell$. Here we have three regions to consider:

B1. The outer region $|\alpha| \gg |t|2^{-\frac{\ell}{2}}$ where the problem is elliptic, with $\alpha\partial_\alpha$ as the dominant term.

B2. The inner region $|\alpha| \ll |t|2^{-\frac{\ell}{2}}$ where the problem is elliptic, with $it|D|^{\frac{1}{2}}$ as the dominant term.

B3. The intermediate region $|\alpha| \approx |t|2^{-\frac{\ell}{2}}$ where the problem is hyperbolic.

We consider three overlapping smooth positive cutoff functions χ_{out}^ℓ, χ_{med}^ℓ and χ_{in}^ℓ associated with the three regions. In order to keep the frequency localization we assume that all three cutoffs are localized at frequency $\ll 2^\ell$, at the expense of having tails which decay rapidly on the $2^{-\ell}$ scale. We remark that the three cutoffs begin to separate exactly at $2^\ell = t^{-2}$.

For the regions B1 and B3 we use elliptic estimates, while for B2 we use a propagation bound.

Using the frequency localized form of (6.16) we can bound

$$\|\chi_{out}^\ell \alpha\partial_\alpha(w_\ell, r_\ell)\|_{L^2} \lesssim |t|\|\chi_{out}^\ell |D|^{\frac{1}{2}}(r_\ell, w_\ell)\|_{L^2} + \|(g_\ell, k_\ell)\|_{L^2}..$$

After some commutations this gives

$$2^\ell\|\alpha\chi_{out}^\ell(w_\ell, r_\ell)\|_{L^2} \lesssim 2^{\frac{\ell}{2}}|t|\|\chi_{out}^\ell(r_\ell, w_\ell)\|_{L^2} + \|(g_\ell, k_\ell)\|_{L^2} + \|(w_\ell, r_\ell)\|_{L^2}.$$

Taking into account the localization of χ_{out}^ℓ, this yields

$$2^\ell\|\alpha\chi_{out}^\ell(w_\ell, r_\ell)\|_{L^2} \lesssim \|(g_\ell, k_\ell)\|_{L^2} + \|(w_\ell, r_\ell)\|_{L^2}. \tag{6.21}$$

By Bernstein's inequality this gives the pointwise bound

$$\chi^{\ell}_{out}|(w_\ell, r_\ell)| \lesssim 2^{-\frac{\ell}{2}}|\alpha|^{-1}\left(\|(g_\ell, k_\ell)\|_{L^2} + \|(w_\ell, r_\ell)\|_{L^2}\right). \tag{6.22}$$

A similar computation, but with the roles of the two terms on the left in (6.16) reversed gives

$$|t|2^{\frac{\ell}{2}}\|\chi^{\ell}_{in}(w_\ell, r_\ell)\|_{L^2} \lesssim \|(g_\ell, k_\ell)\|_{L^2} + \|(w_\ell, r_\ell)\|_{L^2}. \tag{6.23}$$

By Bernstein's inequality this gives the pointwise bound

$$\chi^{\ell}_{in}|(w_\ell, r_\ell)| \lesssim |t|^{-1}\left(\|(g_\ell, k_\ell)\|_{L^2} + \|(w_\ell, r_\ell)\|_{L^2}\right). \tag{6.24}$$

It remains to consider the intermediate region, where we produce instead a propagation estimate. Precisely, for $\chi^{\ell}_{med}(w_\ell, r_\ell)$ we estimate

$$\begin{aligned}\|(4\alpha^2\partial_\alpha - it^2)\chi^{\ell}_{med}w_\ell\|_{L^2} &\lesssim \|2\alpha(2\alpha\partial_\alpha\chi^{\ell}_{med}w_\ell - it|D|^{\frac{1}{2}}\chi^{\ell}_{med}r_\ell)\|_{L^2} \\ &\quad + \|t(\alpha|D|^{\frac{1}{2}}\chi^{\ell}_{med}r_\ell - it\chi^{\ell}_{med}w_\ell)\|_{L^2} \\ &\lesssim |t|2^{-\frac{\ell}{2}}(\|(g_\ell, k_\ell)\|_{L^2} + \|(r_\ell, w_\ell)\|_{L^2}),\end{aligned}$$

and similarly for r_ℓ. Applying

$$\frac{d}{dt}|u|^2 = 2\Re\left[(\partial_\alpha - i\frac{t^2}{\alpha^2})u\cdot\bar{u}\right]$$

for $u = \chi^{\ell}_{med}w_\ell$ and $u = \chi^{\ell}_{med}r_\ell$, integrating from infinity and using the Cauchy-Schwarz inequality yields

$$\chi^{\ell}_{med}|(w_\ell, r_\ell)|^2 \lesssim \alpha^{-2}|t|2^{-\frac{\ell}{2}}(\|(g_\ell, k_\ell)\|_{L^2} + \|(r_\ell, w_\ell)\|_{L^2})\|(r_\ell, w_\ell)\|_{L^2}.$$

Using this in the interesting region $|\alpha| \approx |t|\ell^{\frac{1}{2}}$ and the inner and outer estimates away from it we obtain

$$\chi^{\ell}_{med}|(w_\ell, r_\ell)| \lesssim |\alpha|^{-\frac{1}{2}}(\|(g_\ell, k_\ell)\|^{\frac{1}{2}}_{L^2}\|(r, w_\ell)\|^{\frac{1}{2}}_{L^2} + \|(r_\ell, w_\ell)\|_{L^2}). \tag{6.25}$$

Now, we prove the bounds in the Lemma by dyadic summation. There are several cases to consider:

Case 1: $t^{-2} < 2^\ell < 1$, where all three bounds coincide, and it suffices to prove (6.18). Assume that the two norms in the right hand side of (6.18) are ≤ 1. For α we have three cases:

(a) $|\alpha| > t^2$, where we are in case B1 for all ℓ. There we need only (6.22) to conclude that

$$|(w_{[t^{-2},1]}, r_{[t^{-2},1]})(\alpha)| \lesssim \sum_{2^\ell = t^{-2}}^{1} 2^{-\frac{\ell}{2}}|\alpha|^{-1} \approx |t||\alpha|^{-1}.$$

(b) $|t| < |\alpha| < t^2$, where we are successively in case B1, B2 and B3. There we use (6.22) (6.25) and (6.24) to conclude that

$$\begin{aligned}|(w_{[t^{-2},1]}, r_{[t^{-2},1]})(\alpha)| &\lesssim \sum_{2^\ell = t^{-2}}^{\alpha^{-2}t^2} |t|^{-1} + |\alpha|^{-\frac{1}{2}} + \sum_{2^\ell = \alpha^2 t^{-2}}^{1} 2^{-\frac{\ell}{2}}|\alpha|^{-1} \\ &\approx |t|^{-1}|\log(1 + t^2|\alpha|^{-1})| + |\alpha|^{-\frac{1}{2}} + |\alpha|^{-1} \lesssim |\alpha|^{-\frac{1}{2}}.\end{aligned}$$

(c) $|\alpha| < |t|$, where we are in case B3 for all ℓ. There we need only (6.24) to conclude that

$$|(w_{[t^{-2},1]}, r_{[t^{-2},1]})(\alpha)| \lesssim \sum_{2^\ell = t^{-2}}^{1} |t|^{-1} \lesssim t^{-1} |\log(|t|+2)|.$$

Case 2: $1 < 2^\ell$**.** Here we have two subcases:

(i) $|\alpha| > |t|$, where we are in case B1 for all ℓ. There we need only (6.22) to conclude that

$$|(w_{>1}, r_{>1})(\alpha)| \lesssim \sum_{2^\ell=1}^{\infty} 2^{-\frac{\ell}{2}} |\alpha|^{-1} (\|(w,r)\|_{L^2} + \|(g,k)\|_{L^2}) \approx |\alpha|^{-1} (\|(w,r)\|_{L^2} + \|(g,k)\|_{L^2}),$$

which suffices for (6.17). In order to also obtain (6.19) we also use the Bernstein bound

$$|(w_\ell, r_\ell)| \lesssim 2^{\frac{\ell}{2}} \|(w_\ell, r_\ell)\|_{L^2}. \tag{6.26}$$

Then we obtain

$$\begin{aligned} |\partial^k (w_{>1}, r_{>1})(\alpha)| &\lesssim \sum_{2^\ell=1}^{\infty} \min\{2^{(k+\frac{1}{2})\ell} |\alpha|^{-1} (\|(g,k)\|_{H^{-1}} + \|(w,r)\|_{H^{-1}}), 2^{-(k-\frac{3}{2})\ell} \|(w,r)\|_{\dot H^{2k+2}}\} \\ &\lesssim |\alpha|^{-\frac{1}{2}} (\|(g,k)\|_{H^{-1}} + \|(w,r)\|_{H^{-1}})^{\frac{1}{2}} \|(w,r)\|_{\dot H^{2k+2}}^{\frac{1}{2}}. \end{aligned}$$

(ii) $|\alpha| < |t|$, where we are successively in cases B1, B2 and B3. There we use (6.22) (6.25) and (6.24) to conclude that

$$\begin{aligned} |(w_{>1}, r_{>1})(\alpha)| &\lesssim \left(\sum_{2^\ell=1}^{\alpha^{-2}t^2} |t|^{-1} + |\alpha|^{-\frac{1}{2}} + \sum_{2^\ell = \alpha^{-2}t^2}^{\infty} 2^{-\frac{\ell}{2}} |\alpha|^{-1} \right) (\|(w,r)\|_{L^2} + \|(g,k)\|_{L^2}) \\ &= \left(|t|^{-1} |\log(1 + |t|/|\alpha|)| + |\alpha|^{-\frac{1}{2}} + |t|^{-1} \right) |\alpha|^{-\frac{1}{2}} (\|(w,r)\|_{L^2} + \|(g,k)\|_{L^2}) \\ &\lesssim |\alpha|^{-\frac{1}{2}} (\|(w,r)\|_{L^2} + \|(g,k)\|_{L^2}), \end{aligned}$$

which suffices for (6.17).

Finally, the bound (6.19) is obtained exactly as in (i) by combining the last computation with the trivial pointwise bound for $\partial^k(w_\ell, r_\ell)$ obtained from Bernstein's inequality (6.26). Separating the contributions from cases B1, B2 and B3 we obtain

$$|\partial^k (w_{>1}, r_{>1})(\alpha)| \lesssim I + II + III,$$

where by (6.22) and (6.26) we have

$$\begin{aligned} I &= \sum_{2^\ell=1}^{\alpha^{-2}t^2} \min\{2^{(k+1)\ell} |t|^{-1} (\|(w,r)\|_{H^{-1}} + \|(g,k)\|_{H^{-1}}), 2^{-(k+\frac{3}{2})\ell} \|(w,r)\|_{\dot H^{2k+2}}\} \\ &\lesssim |t|^{-\frac{1}{2}} |t|^{-\frac{1}{8k+10}} (\|(w,r)\|_{H^{-1}} + \|(g,k)\|_{H^{-1}} + \|(w,r)\|_{\dot H^{2k+2}}) \end{aligned}$$

by (6.25) we have

$$II = \alpha^{\frac{1}{2}} |t|^{-1} (\|(w,r)\|_{H^{-1}} + \|(g,k)\|_{H^{-1}}) \|(w,r)\|_{\dot H^{2k+2}}^{\frac{1}{2}},$$

and by (6.22) and (6.26) we have

$$III = \sum_{2^\ell=\alpha^{-2}t^2}^{\infty} \min\{2^{(k+\frac{1}{2})\ell}|\alpha|^{-1}(\|(w,r)\|_{H^{-1}} + \|(g,k)\|_{H^{-1}}), 2^{-(k+\frac{3}{2})\ell}\|(w,r)\|_{\dot H^{2k+2}}\}$$
$$\lesssim |\alpha|^{\frac{1}{2}}|t|^{-1}(\|(w,r)\|_{H^{-1}} + \|(g,k)\|_{H^{-1}})^{\frac{1}{2}}\|(w,r)\|^{\frac{1}{2}}_{\dot H^{2k+2}}$$

□

Appendix A. Holomorphic equations

In this section we give an alternative derivation for the evolution equations for water waves in conformal coordinates. They were first obtained in [17], and also later in [9, 24] but using a different set up. We use a holomorphic form of the equations, as in [9, 24], but we compactify the equations even more, as we will show below. We also express the normal derivative of the pressure on the boundary in terms of our variables.

We consider two-dimensional, irrotational gravity water waves in an inviscid, incompressible fluid of infinite depth. First, we discuss the localized case on $\mathbb{R}$ in which the waves decay at infinity. The spatially periodic case is almost identical, and we describe the appropriate modifications afterwards.

1.1. **Holomorphic coordinates.** Suppose that at time t the fluid occupies a spatial region $\Omega(t) \subset \mathbb{R}^2$ whose simple nondegenerate boundary $\Gamma(t) = \partial\Omega(t)$ approaches $y = 0$ at infinity. Then there is a unique conformal map $\mathcal{F}(t) : \mathbb{H} \to \Omega(t)$ from the lower half-plane

$$\mathbb{H} = \{\alpha + i\beta : \beta < 0\}$$

onto $\Omega(t)$, with $x = x(t, \alpha, \beta)$ and $y = y(t, \alpha, \beta)$, such that $z = x + iy$ satisfies

$$z - (\alpha + i\beta) \to 0 \qquad \text{as} \quad \alpha + i\beta \to \infty.$$

Since $\mathcal{F}(t)$ is conformal, we have $x_\alpha = y_\beta$, $x_\beta = -y_\alpha$. If $f(t, \cdot) : \Omega(t) \to \mathbb{C}$ is a time-dependent spatial function and $g(t, \cdot) = f(t, \cdot) \circ \mathcal{F}(t) : \mathbb{H} \to \mathbb{C}$ is the corresponding conformal function, then $g_t = f_t + x_t f_x + y_t f_y$, so

$$f_t = g_t - \frac{1}{j}(x_\alpha x_t + y_\alpha y_t)\, g_\alpha - \frac{1}{j}(x_\beta x_t + y_\beta y_t)\, g_\beta, \quad j = x_\alpha^2 + y_\alpha^2. \tag{A.1}$$

Also, if $f + ig : \mathbb{H} \to \mathbb{C}$ is a holomorphic function with boundary value $F + iG$ on the real axis $\beta = 0$ that vanishes at infinity, then $F = HG$ where the Hilbert transform H is defined by

$$Hf(\alpha) = \frac{1}{\pi}\text{p.v.}\int_{-\infty}^{\infty} \frac{f(\alpha')}{\alpha - \alpha'}\, d\alpha', \qquad He^{ik\alpha} = -i(\operatorname{sgn} k)e^{ik\alpha}.$$

We denote by $P = \frac{1}{2}(I - iH)$ the projection onto boundary values of functions that are holomorphic in the lower half-plane and vanish at infinity. That is, P projects functions onto their negative wavenumber components.

1.2. **Water waves in holomorphic coordinates.** Let $\phi : \Omega(t) \to \mathbb{R}$ be the spatial velocity potential of the fluid, chosen so that it vanishes at infinity, and $\psi = \phi \circ \mathcal{F} : \mathbb{H} \to \mathbb{R}$ the corresponding conformal velocity potential,

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

Then ψ is harmonic since ϕ is harmonic; we denote the conjugate function of $\psi(t, \alpha, \beta)$ by $\theta(t, \alpha, \beta)$. The velocity components of the fluid $(u, v) = (\phi_x, \phi_y)$ are given in terms of ψ by

$$u = \frac{1}{j}\left(x_\alpha \psi_\alpha + x_\beta \psi_\beta\right), \quad v = \frac{1}{j}\left(y_\alpha \psi_\alpha + y_\beta \psi_\beta\right). \tag{A.2}$$

The conformally parametrized equation of the free surface $\Gamma(t)$ is $x = X(t, \alpha)$, $y = Y(t, \alpha)$, where $X(t, \alpha) = x(t, \alpha, 0)$, $Y(t, \alpha) = y(t, \alpha, 0)$.

To avoid any confusions, we emphasize that the variable Y used through this section has a different meaning than elsewhere in the paper; it is the vertical component of the parametrized free surface $Z(t, \alpha)$.

Since $(x - \alpha) + i(y - \beta)$ is holomorphic in the lower half-plane and vanishes at infinity, we have

$$X = \alpha + HY, \qquad Y = -H(X - \alpha). \tag{A.3}$$

Let $\Psi(t, \alpha) = \psi(t, \alpha, 0)$ denote the boundary value of the conformal velocity potential and $\Theta(t, \alpha) = \theta(t, \alpha, 0)$ the conjugate function, where $\Theta = -H\Psi$ and

$$\psi_\beta|_{\beta=0} = H\Psi_\alpha = -\Theta_\alpha. \tag{A.4}$$

After these preliminaries, we transform the spatial boundary conditions for water-waves into conformal coordinates.

Kinematic BC. A spatial normal to the free surface Γ is $(-Y_\alpha, X_\alpha)$. The kinematic BC, that the normal component of the velocity of the free surface is equal to the normal component of the fluid velocity, is

$$(X_t, Y_t) \cdot (-Y_\alpha, X_\alpha) = (u, v) \cdot (-Y_\alpha, X_\alpha) \qquad \text{on } \Gamma(t).$$

Using (A.2) and (A.4) in this equation and simplifying the result, we get

$$X_\alpha Y_t - Y_\alpha X_t = -\Theta_\alpha. \tag{A.5}$$

In addition, the function z_t/z_α is holomorphic in $\mathbb{H}$ and decays at infinity, so the real part of its boundary value on the real axis is the Hilbert transform of its imaginary part. After the use of (A.5), this gives the equation

$$X_\alpha X_t + Y_\alpha Y_t = -JH\left[\frac{\Theta_\alpha}{J}\right]. \tag{A.6}$$

Solving (A.5)–(A.6) for X_t, Y_t, we get an expression for the velocity of a conformal point on the free surface

$$X_t = -H\left[\frac{\Theta_\alpha}{J}\right] X_\alpha + \frac{\Theta_\alpha}{J} Y_\alpha, \quad Y_t = -\frac{\Theta_\alpha}{J} X_\alpha - H\left[\frac{\Theta_\alpha}{J}\right] Y_\alpha. \tag{A.7}$$

Dynamic BC. Bernoulli's equation for the pressure p in the fluid, with gravitational acceleration $g = 1$, is

$$\phi_t + \frac{1}{2}|\nabla\phi|^2 + y + p = 0. \tag{A.8}$$

The arbitrary function of t that may appear in this equation is zero since we assume that ϕ vanishes at infinity and $p = 0$ on the free surface which approaches $y = 0$. The spatial form of the dynamic BC, without surface tension, is

$$\phi_t + \frac{1}{2}|\nabla\phi|^2 + y = 0 \qquad \text{on } \Gamma(t).$$

Using (A.1) to compute ϕ_t, evaluating the result at $\beta = 0$, and using (A.4)–(A.6), we find that

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

We also have

$$\frac{1}{2}\left.|\nabla\phi|^2\right|_{\beta=0} = \frac{1}{2J}\left(\Psi_\alpha^2 + \Theta_\alpha^2\right).$$

Hence, the dynamic BC in conformal variables is

$$\Psi_t + H\left[\frac{\Theta_\alpha}{J}\right]\Psi_\alpha + \frac{1}{2J}\left(\Psi_\alpha^2 - \Theta_\alpha^2\right) + Y = 0. \tag{A.9}$$

To put these equations in holomorphic form, we define

$$Z = X + iY, \quad Q = \Psi + i\Theta, \quad F = P\left[\frac{Q_\alpha - \bar{Q}_\alpha}{J}\right], \quad J = |Z_\alpha|^2. \tag{A.10}$$

Then Z, Q, F are the boundary values of functions that are holomorphic in the lower half-plane, and $P[Z-\alpha] = Z - \alpha$, $PQ = Q$. The kinematic BC (A.7) is equivalent to

$$Z_t + FZ_\alpha = 0. \tag{A.11}$$

Applying the holomorphic projection P to the dynamic BC (A.9), using Hilbert transform identities, and simplifying the result, we get that

$$Q_t + FQ_\alpha + P\left[\frac{|Q_\alpha|^2}{J}\right] = i\left(Z - \alpha\right), \qquad J = |Z_\alpha|^2. \tag{A.12}$$

Thus, the holomorphic equations are (A.11)–(A.12).

1.3. **The normal derivative of the pressure.** In this subsection, for comparison purposes, we compute the normal derivative of the pressure in terms of our variables. This played a role in the subject as the Taylor sign condition

$$\frac{\partial p}{\partial n}_{|\Gamma_t} < 0$$

was identified as necessary for the well-posedness of the water wave equation, see [20]. In our context this is automatically satisfied, see the discussion at the end of this section.

From Bernoulli's equation (A.8) we have that

$$-\frac{\partial p}{\partial n} = -\frac{1}{J}\,p_\beta|_{\beta=0} = \frac{1}{J}\,\partial_\beta\left.\left(\phi_t + \frac{1}{2}|\nabla\phi|^2 + y\right)\right|_{\beta=0}.$$

Converting β-derivatives to α-derivatives and using the evolution equations, we find that

$$\begin{aligned}\partial_\beta \phi_t|_{\beta=0} &= \partial_\alpha \left[-\Theta_t + \frac{\Psi_\alpha}{J}\left(X_\alpha Y_t - Y_\alpha X_t\right) + \frac{\Theta_\alpha}{J}\left(X_\alpha X_t + Y_\alpha Y_t\right)\right] \\ &= -\partial_\alpha \left[\Theta_t + \Theta_\alpha H\left(\frac{\Theta_\alpha}{J}\right) + \frac{\Psi_\alpha \Theta_\alpha}{J}\right] \\ &= -\partial_\alpha \left[H\left(\frac{\Psi_\alpha^2 + \Theta_\alpha^2}{2J}\right) + X - \alpha\right].\end{aligned}$$

Since $y_\beta|_{\beta=0} = X_\alpha$ and

$$|\nabla\phi|^2\big|_{\beta=0} = \frac{\Psi_\alpha^2 + \Theta_\alpha^2}{J},$$

we find that

$$-J\frac{\partial p}{\partial n} = 1 + \frac{1}{2}\left(\partial_\beta - H\partial_\alpha\right)|\nabla\phi|^2\big|_{\beta=0}.$$

To put this equation in holomorphic form, we introduce

$$r = \frac{q_\alpha}{z_\alpha}, \qquad \partial = \frac{1}{2}\left(\partial_\alpha - i\partial_\beta\right), \quad \bar\partial = \frac{1}{2}\left(\partial_\alpha + i\partial_\beta\right),$$

where $|\nabla\phi|^2 = r\bar r$ and $r|_{\beta=0} = R$ is defined in (1.2). Then, using the fact that $\partial r = r_\alpha$ and $\bar\partial r = 0$, we get that

$$-J\frac{\partial p}{\partial n} = 1 + i\left[\bar P(\bar R R_\alpha) - P(R\bar R_\alpha)\right] = 1 + a,$$

where a is defined in (1.4). Comparing this result with Wu [23], we see that up to Jacobian factors, our $1+a$ is proportional to her $\mathfrak{a}$. Moreover, as shown in [22] under the assumption of non-self intersecting boundary, we have $a \geq 0$. A shorter alternate proof of this fact is provided in our Lemma 2.6; further we impose no condition on the self intersections of the curve $Z(t,\alpha)$.

1.4. **The periodic case.** In the spatially periodic case, the map $\mathcal{F}(t)$ is uniquely determined by the requirement that the holomorphic function $z(t,\alpha,\beta) - (\alpha + i\beta)$ is a periodic function of α, whose real part approaches zero[2] as $\beta \to -\infty$. It follows that $\Re Z(t,\alpha) - \alpha$ has zero mean with respect to α, otherwise the holomorphic function would have nonzero limit.

In the original coordinates, the velocity field $u + iv$ is holomorphic, periodic and bounded. Thus it has a limit $u_0 + iv_0$ as $\beta \to \infty$. Further, this limit is independent of time. By extension, in the holomorphic coordinates Q_α also has a limit $u_0 + iv_0$ as $b \to \infty$. Thus, we can normalize Q by setting

$$Q = Q_0 + (u_0 + iv_0)(\alpha + i\beta) + c(t),$$

where Q_0 is periodic with average zero, and $c(t)$ is a real normalization constant needed for Bernoulli's law. One could continue the computations using u_0 and v_0 as constants of motion, but this is not needed because we can factor them out using a Galilean transformation. From here on we set them equal to zero.

The relation (A.5) rests unchanged,

$$X_\alpha Y_t - Y_\alpha X_t = -\Theta_\alpha. \tag{A.13}$$

[2]The imaginary part need not have zero mean even if the average height stays equal to zero.

In integrated form this expresses the conservation of mass. Consider now both terms divided by J, then this becomes

$$\Im\left(\frac{z_t}{z_\alpha}\right) = -\frac{\Theta_\alpha}{J} \qquad \text{on} \quad \beta = 0.$$

The function on the left is holomorphic and its real part has limit zero as β goes to infinity, so we can get its real part using the Hilbert transform. Hence (A.6) also holds, and (A.7) follows. Similarly, the derivation of (A.9) remainss unchanged.

To put the equations (A.7) and (A.9) in holomorphic form we keep the definition of the operator P as

$$P = \frac{1}{2}(I - iH)$$

even though it is no longer a projector, as it selects exactly half of the zero mode. With the same notations as in (A.10), the equations (A.11) and (A.12) remains unchanged.

Concerning the balance of averages in these two equations, we remark that in (A.11) the terms $\Re Z_t$ and F have purely imaginary averages while Z_α has average 1. The nontrivial average here is that of F, which contributes to the motion of nonzero frequencies. In the second equation (A.12), the real part of the average of Q is nonzero due to the integrating constant in Bernoulli's law; however this plays a trivial role, as it does not affect any of the remaining equations. Further, R has no zero modes.

All equations in the first section of the paper remain unchanged, most importantly the expressions for the frequency shift a, the advection velocity b and the auxiliary function M. Further, all estimates in Lemmas 2.6, 2.7, 2.8 remain unchanged; only the zero mode estimates need to be added, and those are straightforward. The normal form transformation also remains valid.

We next consider the linearized equations. The derivation of (2.1) is purely algebraic, so it stays unchanged. We remark that the average of w is purely imaginary, while the average of r is the same as the average of q and is purely real.

There is some choice to be made when writing the projected equations (2.2). The operator P defined as above is no longer a projector, so we can no longer use it directly. The new question that arises here is how we treat the zero modes. For that we introduce some variants of P which differ in how the zero modes are treated. Defining P_0 as the projection onto the zero modes, we define the projectors

$$P^\sharp = P - \frac{1}{2}P_0, \qquad P^r = P^\sharp + \Re P_0, \qquad P^i = P^\sharp + i\Im P_0,$$

and similarly $\bar{P}^\sharp$, $\bar{P}^r$ and $\bar{P}^i$. We have the relations

$$P = P^i + \bar{P}^r = P^r + \bar{P}^i = P^\sharp + \bar{P}^\sharp + P_0, \qquad P^i\bar{P}^r = P^r\bar{P}^i = 0, \qquad P^i = -iP^r i.$$

With these notations, it is natural to project the first equation using P^i, and the second using P^r. Thus instead of (2.2) we write

$$\text{(A.14)} \qquad \begin{cases} (\partial_t + P^i b\partial_\alpha)w + P^i\left[\dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P^i\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = P^i\mathcal{G}(w,r), \\ (\partial_t + P^r\partial_\alpha)r - iP^i\left[\dfrac{1+a}{1+\mathbf{W}}w\right] = P^r\mathcal{K}(w,r). \end{cases}$$

The quadratic part of $\mathcal{G}$ has real average, and $\mathcal{K}$ has imaginary average, both of which get projected out. So

$$P^i\mathcal{G}^{(2)} = P^\sharp[R\bar{w}_\alpha - \mathbf{W}\bar{r}_\alpha], \qquad P^i\mathcal{K}^{(2)} = -P^\sharp[R\bar{r}_\alpha].$$

After a similar modification in (2.4), the statement and the proof of Proposition 2.1 remain largely unchanged; the difference is that P gets replaced by P^i in err_1.

Moving on to the cubic estimates, the equation (2.11) is replaced by

$$\text{(A.15)} \qquad \begin{cases} (\partial_t + P^i b\partial_\alpha)w + P^i\left[\dfrac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P^i\left[\dfrac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = P^\sharp[R\bar{w}_\alpha - \mathbf{W}\bar{r}_\alpha] + G, \\ (\partial_t + P^r\partial_\alpha)r - iP^i\left[\dfrac{1+a}{1+\mathbf{W}}w\right] = -P^\sharp[R\bar{r}_\alpha] + K. \end{cases}$$

Also, the cubic energy needs to be modified. Precisely, the zero modes of w and r do not affect the quadratic terms on the right hand side above. Hence, the cubic energy correction should not involve these zero modes either,

$$\text{(A.16)} \qquad E^{(3)}_{lin}(w,r) = \int_{\mathbf{R}} (1+a)|w|^2 + \Im(r\bar{r}_\alpha) + 2\Im(\bar{R}w^\sharp r_\alpha) - 2\Re(\bar{\mathbf{W}}(w^\sharp)^2)\, d\alpha.$$

With this modification, the result in Proposition 2.2 remains valid, and the proof applies with minor modifications.

The higher energies in the periodic case are even more similar to the nonperiodic case, since differentiation eliminates the zero modes, and the frequency localization is achieved using only $P^\sharp$ and $\bar{P}^\sharp$.

An alternative to the above scheme is to select just the negative wave numbers in the linearized equation. Then we lose the evolution of the average of the imaginary part of w. This is not so significant since we have the conservation of mass relation

$$\int YX_\alpha d\alpha = const,$$

where the (time independent) constant on the right can be set arbitrarily (say to zero) by a vertical translation of the coordinates,

$$(Z,Q) \to (Z+ic, Q-ct).$$

This gives

$$\int Y d\alpha = i\int (\bar{Z}-\alpha)(Z_\alpha - 1)\, d\alpha + const,$$

where the average of Y plays no role on the right. This shows that also for the linearized equation, the average of w is determined by the initial data and the negative frequencies of w and W,

$$\int \Im w d\alpha = i\int \bar{w}W_\alpha + \bar{W}w_\alpha\, d\alpha + const.$$

Again, the constant can be removed as the pair $(i, t-iR)$ solves the linearized problem. It follows that the contribution of the average of w to the linearized equations can be viewed as cubic and higher.

Appendix B. Norms and multilinear estimates

Here we prove some of the estimates used in Section 2, and Section 3. We use a standard Littlewood-Paley decomposition in frequency

$$1 = \sum_{k \in \mathbf{Z}} P_k,$$

where the multipliers P_k have smooth symbols localized at frequency 2^k.

A good portion of our analysis happens at the level of homogeneous Sobolev spaces $\dot{H}^s$, whose norm is given by

$$\|f\|_{\dot{H}^s} \sim \|(\sum_k |2^{ks} P_k f|^2)^{1/2}\|_{L^2} = \|2^{ks} P_k f\|_{L^2_x \ell^2_k}.$$

We will also use the Littlewood-Paley square function and its restricted version,

$$S(f)(x) := \left(\sum_{k \in \mathbf{Z}} |P_k(f)(x)|^2 \right)^{\frac{1}{2}}, \qquad S_{>k}(u) = (\sum_{j>k} |P_j u|^2)^{\frac{1}{2}}.$$

The Littlewood-Paley inequality is recalled below

$$\|S(f)\|_{L^p(\mathbf{R})} \simeq_p \|f\|_{L^p(\mathbf{R})}, \qquad 1 < p < \infty. \tag{B.1}$$

By duality this also yields the estimate

$$\| \sum_{k \in \mathbf{Z}} P_k f_k \|_{L^p} \lesssim \| \sum_{k \in \mathbf{Z}} (|f_k|^2)^{1/2} \|_{L^p}, \qquad 1 < p < \infty. \tag{B.2}$$

The $p = 1$ version of the above estimate for the Hardy space H_1 is

$$\|f\|_{H_1} \simeq \|S(f)\|_{L^1_x \ell^2_k}, \tag{B.3}$$

which by duality implies the BMO bound

$$\| \sum_{k \in \mathbf{Z}} P_k f_k \|_{BMO} \lesssim \|S(f)\|_{L^\infty}. \tag{B.4}$$

The square function characterization of BMO is slightly different,

$$\|u\|^2_{BMO} \approx \sup_k \sup_{|Q| = 2^{-k}} 2^k \int_Q |S_{>k}(u)|^2 \, dx. \tag{B.5}$$

We will also need the maximal function bound

$$\|P_{<k} f\|_{L^2_x L^\infty_k} \lesssim \|f\|_{L^2}, \qquad 1 < p < \infty. \tag{B.6}$$

2.1. **Coifman-Meyer and and Moser type estimates.** In the context of bilinear estimates a standard tool is to consider a Littlewood-Paley paraproduct type decomposition of the product of two functions,

$$fg = \sum_{k \in \mathbb{Z}} f_{<k-4} g_k + \sum_{k \in \mathbb{Z}} f_k g_{<k-4} + \sum_{|k-l| \leq 4} f_k g_l := T_f g + T_g f + \Pi(f, g).$$

Here and below we use the notation $f_k = P_k f$, $f_{<k} = P_{<k} f$, etc. By a slight abuse of notation, in the sequel we will omit the frequency separation from our notations in bilinear

Littlewood-Paley decomposition; for instance instead of the above formula we will use the shorter expression

$$fg = \sum_{k\in\mathbb{Z}} f_{<k} g_k + \sum_{k\in\mathbb{Z}} f_k g_{<k} + \sum_{k\in\mathbb{Z}} f_k g_k.$$

Away from the exponents 1 and ∞ one has a full set of estimates

$$\|T_f g\|_{L^r} + \|\Pi(f,g)\|_{L^r} \lesssim \|f\|_{L^p}\|g\|_{L^q}, \qquad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \qquad 1 < p,q,r < \infty. \tag{B.7}$$

Corresponding to $q = \infty$ one also has a BMO estimate

$$\|T_f g\|_{L^p} + \|\Pi(f,g)\|_{L^p} \lesssim \|f\|_{L^p}\|g\|_{BMO}, \qquad 1 < p < \infty, \tag{B.8}$$

which in turn leads to the commutator bound

$$\|[P,g]f\|_{L^p} \lesssim \|f\|_{L^p}\|g\|_{BMO}, \qquad 1 < p < \infty. \tag{B.9}$$

For $p = 2$ we also need an extension of this, namely

Lemma 2.1. *The following commutator estimates hold:*

$$\||D|^s [P,R] |D|^\sigma w\|_{L^2} \lesssim \||D|^{\sigma+s} R\|_{BMO}\|w\|_{L^2}, \qquad \sigma \geq 0, \quad s \geq 0, \tag{B.10}$$

$$\||D|^s [P,R] |D|^\sigma w\|_{L^2} \lesssim \||D|^{\sigma+s} R\|_{L^2}\|w\|_{BMO}, \qquad \sigma > 0, \quad s \geq 0. \tag{B.11}$$

We remark that later this is applied to functions which are holomorphic/antiholomorphic, but that no such assumption is made above.

Proof. If $\sigma = s = 0$ then (B.10) is the classical commutator estimate of Coifman and Meyer (B.9), so we take $\sigma + s > 0$. We consider the usual paradifferential decomposition, and observe that the expression $[P,R]\,|D|^\sigma w$ vanishes if the frequency of w is much larger than the frequency of R. For the remaining frequency balances we discard P, and we are left with having to estimate the expressions

$$f_{hh} = \sum_k 2^{(\sigma+s)k}(2^{-k}|D|)^s(R_k w_k), \qquad f_{hl} = \sum_k 2^{(\sigma+s)k} R_k 2^{-\sigma k}|D|^\sigma w_{<k}.$$

In the term f_{hh} the σ derivatives are already moved to R, so this is bounded using (B.8) if $s = 0$, and directly if $s > 0$. For the remaining part we only need the infinity Besov norm of R_k, as

$$\begin{aligned}\|f_{hl}\|_{L^2}^2 &\lesssim \sum_k \|2^{(\sigma+s)k} R_k\|_{L^\infty}^2 \|2^{-\sigma k}|D|^\sigma w_{<k}\|_{L^2}^2 \lesssim \sup_k \|2^{(\sigma+s)k} R_k\|_{L^\infty}^2 \sum_k \|2^{-\sigma k}|D|^\sigma w_{<k}\|_{L^2}^2 \\ &\lesssim \||D|^{\sigma+s} R\|_{BMO}^2 \|w\|_{L^2}^2.\end{aligned}$$

The proof of (B.11) is similar. $\square$

Next we consider some similar product type estimates involving BMO and L^∞ norms. We define

$$\|u\|_{BMO^{\frac{1}{2}}} = \||D|^{\frac{1}{2}} u\|_{BMO}.$$

Then

Proposition 2.2. *a) The following estimates hold:*

$$\| \sum_k u_k v_k \|_{BMO} \lesssim \|u\|_{BMO} \|v\|_{BMO}, \tag{B.12}$$

$$\| \sum_k (2^{-k}|D|)^\sigma (u_k v_k) \|_{BMO} \lesssim \|u\|_{BMO} \|v\|_{\dot{B}^0_{\infty,\infty}}, \qquad \sigma > 0, \tag{B.13}$$

$$\| \sum_k u_{<k} v_k \|_{BMO} \lesssim \|u\|_{L^\infty} \|v\|_{BMO}, \tag{B.14}$$

$$\| \sum_k (2^{-k}|D|)^\sigma (u_{<k} v_k) \|_{BMO} \lesssim \|u\|_{\dot{B}^0_{\infty,\infty}} \|v\|_{BMO}, \qquad \sigma > 0. \tag{B.15}$$

b) The space $L^\infty \cap BMO^{\frac{1}{2}}$ is an algebra,

$$\|uv\|_{BMO^{\frac{1}{2}}} \lesssim \|u\|_{L^\infty} \|v\|_{BMO^{\frac{1}{2}}} + \|v\|_{L^\infty} \|u\|_{BMO^{\frac{1}{2}}}, \tag{B.16}$$

c) The following Moser estimate holds for a smooth function F:

$$\|F(u)\|_{BMO^{\frac{1}{2}}} \lesssim_{\|u\|_{L^\infty}} \|u\|_{BMO^{\frac{1}{2}}}, \tag{B.17}$$

Proof. a) For (B.12) we fix a cube Q, which by scaling can be taken to have size 1. Suppose first that $\sigma = 0$. For $k > 0$ we use the square function estimate,

$$\begin{aligned} \| \sum_{k>0} u_k v_k \|_{L^1(Q)} \ &\lesssim \ \|u_k v_k\|_{\ell^1_k L^1(Q)} \lesssim \|u_k\|_{\ell^2_k L^2(Q)} \|v_k\|_{\ell^2_k L^2(Q)} \lesssim \|S_{>0}(u)\|_{L^2(Q)} \|S_{>0}(v)\|_{L^2(Q)} \\ &\lesssim \|u\|_{BMO} \|v\|_{BMO}. \end{aligned}$$

For $k > 0$ we subtract the average and estimate the output in L^∞,

$$\| \sum_{k \leq 0} u_k v_k - (u_k v_k)_Q \|_{L^\infty(Q)} \lesssim \sum_{k<0} \|\partial_\alpha (u_k v_k)\|_{L^\infty} \lesssim \sum_{k<0} 2^k \|u\|_{BMO} \|v\|_{BMO}.$$

Adding the two we get

$$\| \sum_k u_k v_k - (u_k v_k)_Q \|_{L^1(Q)} \lesssim \|u\|_{BMO} \|v\|_{BMO},$$

and (B.12) follows.

The case $k \leq 0$ is similar in the proof of (B.13). For $k > 0$ we first eliminate directly the low frequency output,

$$\|P_{<0} \sum_{k>0} (2^{-k}|D|)^\sigma (u_k v_k) \|_{BMO} \lesssim \sum_{k>0} 2^{-\sigma k} \|u_k\|_{L^\infty} \|v_k\|_{L^\infty} \lesssim \|u\|_{BMO} \|v\|_{\dot{B}^0_{\infty,\infty}}.$$

For the high frequency output we consider a bump function χ_Q adapted to Q, which is localized at frequency less than 1, and thus does not change the frequency localization of the factors it multiplies in the sequel. Then, we have

$$\begin{aligned} \|P_{>0} \sum_{k>0} (2^{-k}|D|)^\sigma (u_k v_k) \|_{L^2(Q)} \ &\lesssim \|\chi_Q P_{>0} \sum_{k>0} (2^{-k}|D|)^\sigma (u_k v_k) \|_{L^2} \\ &\lesssim \|[\chi_Q, P_{>0} \sum_{k>0} (2^{-k}|D|)^\sigma] (u_k v_k) \|_{L^2} + \|P_{>0} \sum_{k>0} (2^{-k}|D|)^\sigma (\chi_Q u_k v_k) \|_{L^2}. \end{aligned}$$

For the commutator term we gain a small power of 2^k so L^∞ bounds suffice. The remaining term is bounded in L^2 in terms of the square function using orthogonality,

$$
\begin{aligned}
\|P_{>0}\sum_{k>0}(2^{-k}|D|)^\sigma(\chi_Q u_k v_k)\|_{L^2}^2 &\lesssim \sum_{k>0}\|\chi_Q u_k\|_{L^2}^2\|v_k\|_{L^\infty}^2 \lesssim \|\chi_Q S_{>0}(u)\|_{L^2}^2\|v\|_{\dot B^0_{\infty,\infty}} \\
&\lesssim \|u\|_{BMO}^2\|v\|_{\dot B^0_{\infty,\infty}}.
\end{aligned}
$$

The argument for (B.14) is similar, with the following modification in the case $k > 0$, which leads to L^2 rather than L^1 bounds:

$$
\begin{aligned}
\|\sum_{k>0} u_{<k}v_k\|_{L^2(Q)} &\lesssim \|\sum_{k>0}\chi_Q u_{<k}v_k\|_{L^2(Q)} \lesssim \|u\|_{L^\infty}(\sum_k \|\chi_Q v_k\|_{L^2(Q)}^2)^{\frac{1}{2}} \\
&\lesssim \|u\|_{L^\infty}\|\chi_Q S_{>0}(v)\|_{L^2(Q)}. \lesssim \|u\|_{L^\infty}\|v\|_{BMO}.
\end{aligned}
$$

Finally, the bound (B.15) is similar since

$$
\|(2^{-k}|D|)^\sigma u_{<k}\|_{L^\infty} \lesssim \|u\|_{\dot B^0_{\infty,\infty}}.
$$

b) With the same paradifferential decomposition as before we need to estimate the terms

$$
f_{hh} = \sum_k |D|^\sigma u_k v_k, \qquad f_{hl} = \sum_k 2^{\sigma k} u_k v_{<k}.
$$

For f_{hl} we use (B.14), while for f_{hh} we use (B.12).

c) We write

$$
\begin{aligned}
F(u) &= \int_{-\infty}^{\infty} u_k F'(u_{<k})\, dk \\
&= \int_{-\infty}^{\infty} u_k P_{<k}F'(u_{<k})\, dk + \int_{-\infty}^{\infty} u_k P_k F'(u_{<k}) dk + \int_{-\infty}^{\infty}\sum_{j>0} u_k P_{k+j}F'(u_{<k})\, dk.
\end{aligned}
$$

For $F'(u_{<k})$ we can use the chain rule to obtain the bound

$$
\|P_{k+j}F'(u_{<k})\|_{L^\infty} \lesssim 2^{-Nj}, \qquad j \geq 0.
$$

With $\sigma = \frac{1}{2}$ we estimate $|D|^\sigma F(u)$. The first term in $|D|^\sigma F(u)$ is

$$
f_1 = \int_{-\infty}^{\infty} (2^{\sigma k}u_k) P_{<k}F'(u_{<k})\, dk,
$$

and is estimated in BMO exactly as in the proof of (B.14).

The second term is

$$
f_2 = \int_{-\infty}^{\infty} (2^{-k}|D|)^\sigma (u_k P_k F'(u_{<k}))\, dk,
$$

and is estimated as in the proof of (B.13).

The last term is $\sum_{j>0} f_{3,j}$, where

$$
f_{3,j} = \int_{-\infty}^{\infty} 2^{\sigma k}u_k 2^{\sigma j} P_j F'(u_{<k})\, dk.
$$

The $k \leq 0$ case is easy; it follows using pointwise estimates. For fixed j and $k > 0$ we bound $f_{3,j}$ by

$$\|f_{3,j,>0}\|^2_{L^2(Q)} \lesssim \|\chi_Q f_{3,j,>0}\|^2_{L^2} \lesssim \int_{-\infty}^{\infty} \|\chi_Q 2^{\sigma k} u_k 2^{\sigma j} P_j F'(u_{<k})\|^2_{L^2}\, dk$$
$$\lesssim 2^{(\sigma-N)j} \|\chi_Q S_{>0}(|D|^\sigma u)\|^2_{L^2} \lesssim 2^{(\sigma-N)j} \|u\|^2_{BMO^\sigma}.$$

□

A more standard algebra estimate and the corresponding Moser bound is as follows:

Lemma 2.3. *Let $\sigma > 0$. Then $\dot{H}^\sigma \cap L^\infty$ is an algebra, and*

$$\|fg\|_{\dot{H}^\sigma} \lesssim \|f\|_{\dot{H}^\sigma}\|g\|_{L^\infty} + \|f\|_{L^\infty}\|g\|_{\dot{H}^\sigma}. \tag{B.18}$$

In addition, the following Moser estimate holds for a smooth function F:

$$\|F(u)\|_{\dot{H}^\sigma} \lesssim_{\|u\|_{L^\infty}} \|u\|_{\dot{H}^\sigma}. \tag{B.19}$$

We also need to consider some multilinear estimates. Our starting point is the bound

$$\|f_1 \cdots f_n\|_{L^r} \lesssim \prod_{j=\overline{1,n}} \|f_j\|_{L^{p_j}}, \qquad \frac{1}{r} = \sum \frac{1}{p_j}, \qquad 1 \leq r, p_j \leq \infty.$$

Adding derivatives, we need the following generalization:

Lemma 2.4. *The following estimate holds for $\sigma > 0$ and $1 < r, p_j^{(k)} \leq \infty$:*

$$\||D|^\sigma (f_1 \cdots f_n)\|_{L^r} \lesssim \sum_{k=1}^{n} \||D|^\sigma f_k\|_{L^{p_k^{(k)}}} \prod_{j \neq k} \|f_j\|_{L^{p_j^{(k)}}}, \qquad \frac{1}{r} = \sum \frac{1}{p_j^{(k)}}. \tag{B.20}$$

The same bound holds if for $L^{p_k^{(k)}}$ is replaced by BMO whenever $p_k^{(k)} = \infty$.

Proof. By induction it suffices to consider the case $n = 2$. After a Littlewood-Paley decomposition we place the derivatives on the highest frequency factor and apply either (B.7) or (B.8), or (B.12),(B.14). □

2.2. **Water-wave related bounds.** Here we consider estimates for objects related to the water wave equations, primarily the real phase shift a and advection velocity b. We recall that these are given by

$$a = 2\Im P[R\bar{R}_\alpha], \qquad b = 2\Re P\left[\frac{R}{1+\bar{\mathbf{W}}}\right] = 2\Re(R - P[R\bar{Y}]).$$

These are estimated in terms of the control parameters A and B defined in (1.11), (1.12), and in terms of the H^s Sobolev norms of $\mathbf{W}$ and R. In all nonlinear bounds the implicit constant is allowed to depend on A.

We begin with the auxiliary variable $Y = \dfrac{\mathbf{W}}{1+\mathbf{W}}$, which inherits its regularity from $\mathbf{W}$ due to (B.17) and (B.19):

Lemma 2.5. *The function Y satisfies the BMO bound*

$$\| |D|^{\frac{1}{2}} Y \|_{BMO} \lesssim_A B, \tag{B.21}$$

and the $\dot H^\sigma$ bound

$$\|Y\|_{\dot H^\sigma} \lesssim_A \|\mathbf{W}\|_{\dot H^\sigma}. \tag{B.22}$$

We continue with bounds for a. In particular the positivity of a is established, providing a short alternate proof to Wu's result in [22]:

Proposition 2.6. *Assume that $R \in \dot H^{\frac{1}{2}} \cap \dot H^{\frac{3}{2}}$. Then the real frequency-shift a is nonnegative and satisfies the BMO bound*

$$\|a\|_{BMO} \lesssim \|R\|^2_{BMO^{\frac{1}{2}}}, \tag{B.23}$$

and the uniform bound

$$\|a\|_{L^\infty} \lesssim \|R\|^2_{\dot B^{\frac{1}{2}}_{\infty,2}}. \tag{B.24}$$

Moreover,

$$\| |D|^{\frac{1}{2}} a\|_{BMO} \lesssim \|R_\alpha\|_{BMO} \| |D|^{\frac{1}{2}} R\|_{L^\infty}, \qquad \|a\|_{B^{\frac{1}{2},\infty}_2} \lesssim \|R_\alpha\|_{B^{\frac{1}{2},\infty}_2} \| |D|^{\frac{1}{2}} R\|_{L^\infty}, \tag{B.25}$$

$$\|(\partial_t + b\partial_\alpha) a\|_{L^\infty} \lesssim AB, \tag{B.26}$$

and

$$\|a\|_{\dot H^s} \lesssim \|R\|_{\dot H^{s+\frac{1}{2}}} \|R\|_{BMO^{\frac{1}{2}}}, \qquad s > 0. \tag{B.27}$$

Proof. We recall that $a = i\left(\bar P\left[\bar R R_\alpha\right] - P\left[R \bar R_\alpha\right]\right)$. Switching to the Fourier space, this leads to the representation

$$\hat a(\zeta) = \int_{\xi - \eta = \zeta} \min\{\xi, \eta\} 1_{\{\xi,\eta>0\}} \hat R(\xi) \hat{\bar R}(\eta) d\xi. \tag{B.28}$$

Here ξ and η are restricted to the positive real axis due to the fact that R is holomorphic.

To prove the positivity of a we represent the above kernel as

$$\min\{\xi, \eta\} 1_{\{\xi,\eta>0\}} = \int_{M>0} 1_{\{\xi > M\}} 1_{\{\eta > M\}}\, dM.$$

Inserting this in the previous representation of $\hat a$ and inverting the Fourier transform we obtain

$$a = \int |1_{|D|>M} R|^2 dM,$$

and the positivity follows.

To prove both the BMO bound and the pointwise bound for a we use a bilinear Littlewood-Paley decomposition,

$$a = \sum_k i\left(\bar R_k R_{\alpha,<k} - R_k \bar R_{\alpha,<k}\right) + i\left(\bar P\left[\bar R_k R_{\alpha,k}\right] - P\left[R_k \bar R_{\alpha,k}\right]\right). \tag{B.29}$$

To estimate the first term in BMO we use directly the bound (B.15) with $\sigma = 1$. To estimate it in L^∞ we use the Cauchy-Schwarz inequality,

$$\| \sum_k \bar{R}_k R_{\alpha,<k} \|_{L^\infty}^2 \lesssim (\sum_k 2^k \|R_k\|_{L^\infty}^2)(\sum_k 2^{-k} \|R_{<k}\|_{L^\infty}^2) \lesssim \|R\|^2_{B^{\frac{1}{2}}_{\infty,2}}.$$

For the second term in a we rewrite the symbol of the bilinear form as

$$\min\{\xi, \eta\} = \frac{1}{2}(\xi + \eta) - \frac{1}{2}|\xi - \eta|,$$

which allow us to rewrite it in the form

$$\frac{1}{2} P \sum_k i \left(\bar{R}_k R_{\alpha,k} - R_k \bar{R}_{\alpha,k} \right) - |D|(\bar{R}_k R_k).$$

Now the two terms are estimated in BMO using (B.12), respectively (B.13), and in L^∞ by the Cauchy-Schwarz inequality as above.

The proof of (B.25) is essentially identical to the proof of (B.23).

We continue with the proof of (B.26), where we begin with the decomposition in (B.29). For the first term in (B.29) we apply the time derivative to obtain the expression

$$A_1 = [b\partial_\alpha, \bar{P}_k] \bar{R} R_{\alpha,<k} + \bar{R}_k [b\partial_\alpha, P_{<k}\partial_\alpha] R + i P_k \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right) R_{\alpha,<k} + i \bar{R}_k P_{<k} \partial_\alpha \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right).$$

In the first term of A_1 we split the commutator according to the usual Littlewood-Paley trichotomy. We get several terms:

$$b_{<k,\alpha} \bar{R}_k R_{\alpha,<k} + b_k \bar{R}_{<k,\alpha} R_{\alpha,<k} + 2^m P_k(b_m \bar{R}_m) R_{\alpha,<k} + b_m \bar{R}_{k,\alpha} R_{\alpha,<k}.$$

In all cases we use the B norm to estimate the highest frequency term, and the A norm for the other two. The second term is similar; for comparison purposes we list the ensuing terms:

$$\bar{R}_k b_{<k,\alpha} R_{<k,\alpha} + 2^m \bar{R}_k P_{<k} \partial_\alpha (b_m R_m) + \bar{R}_k b_m R_{<k,\alpha\alpha}.$$

The third and fourth terms in A_1 require the bound

$$\| \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right) \|_{B^{\frac{1}{2},\infty}_{2}} \lesssim B,$$

which follows by combining the bounds (B.24) and (B.25) for a with the similar bounds for $\mathbf{W}$ and Y.

Now we consider the last term in (B.29). This has two components, one of the form $2^k R_k \bar{R}_k$ and the other of the form $|D|(R_k \bar{R}_k)$. The first component yields an output

$$A_2 = [b\partial_\alpha, \bar{P}_k] \bar{R} R_{\alpha,k} + \bar{R}_k [b\partial_\alpha, P_k \partial_\alpha] R + i P_k \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right) R_{\alpha,k} + i \bar{R}_k P_k \partial_\alpha \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right),$$

which is treated in exactly the same way as A_1.

The second component yields the slightly more involved output

$$\begin{aligned} A_3 = {} & b\partial_\alpha |D|(R_k \bar{R}_k) - |D|(P_k(bR_\alpha)\bar{R}_k) - |D|(R_k \overline{P_k(bR_\alpha)}) \\ & + i|D| \left(P_k \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right) R_k \right) + i|D| \left(\bar{R}_k P_k \partial_\alpha \left(\frac{\bar{\mathbf{W}} - a}{1 + \bar{\mathbf{W}}} \right) \right). \end{aligned}$$

The last two terms are no different from above, but in the first three there is a more delicate commutator estimate. We split $b = b_{<k} + b_{\geq k}$ and estimate the output of $b_{\geq k}$ directly for each term using the $s = \frac{1}{2}$ case of Lemma 2.7. The output of $b_{<k}$, on the other hand, is expressed as a commutator

$$[b_{<k}, |D|_{\leq k}]\partial_\alpha(R_k \bar{R}_k) + |D|\left([b_{<k}, P_k]R_\alpha \bar{R}_k\right) + |D|\left(R_k \overline{[b, P_k]R_\alpha}\right).$$

The last two terms are like $2^k|D|(b_{<k,\alpha}R_k\bar{R}_k)$ and can be estimated directly. For the first term we bound $\partial_\alpha(R_k\bar{R}_k)$ in L^∞ by $2^{-\frac{k}{2}}$, and then we need to show that

$$\|[b_{<k}, |D|_{\leq k}]\|_{L^\infty \to L^\infty} \lesssim 2^{\frac{k}{2}} \||D|^{\frac{1}{2}} b\|_{BMO}.$$

Indeed the kernel of $[b_{<k}, |D|_{\leq k}]$ is bound by

$$|b_{<k}(\alpha) - b_{<k}(\beta)| \frac{2^{2k}}{(1 + 2^k|\alpha - \beta|)^2} \lesssim \frac{2^{\frac{3}{2}k}}{(1 + 2^k|\alpha - \beta|)^{\frac{3}{2}}},$$

which integrates to $2^{\frac{k}{2}}$.

Finally, (B.27) is a direct consequence of the commutator estimates in Lemma 2.1. ☐

Next we consider b, for which we have the follwoing result

Lemma 2.7. *Let $s > 0$. Then the transport coefficient b satisfies*

(B.30) $$\||D|^s b\|_{BMO} \lesssim_A \||D|^s R\|_{BMO}, \qquad \||D|^s b\|_{L^2} \lesssim_A \||D|^s R\|_{L^2}.$$

In particular we have

(B.31) $$\||D|^{\frac{1}{2}} b\|_{BMO} \lesssim_A A, \qquad \|b_\alpha\|_{BMO} \lesssim_A B.$$

Proof. Recall that

$$b = \Re P[R(1 - \bar{Y})] = R - P(R\bar{Y}).$$

Hence, it remains to estimate $\partial_\alpha P(R\bar{Y})$. Consider first the BMO bound. As before, the role of P is to restrict the bilinear frequency interactions to high - low, in which case we can use the bound (B.14), and the high-high case, where (B.13) applies.

A direct argument, taking into account the same two cases, yields the L^2 bound. ☐

Next we consider the auxiliary expression M:

Lemma 2.8. *The function M satisfies the pointwise bound*

(B.32) $$\|M\|_{L^\infty} \lesssim_A AB,$$

as well as the Sobolev bounds

(B.33) $$\|M\|_{\dot{H}^{n-\frac{3}{2}}} \lesssim_A A\mathbf{N}_n.$$

Proof. For the pointwise bound we claim that

(B.34) $$\|M\|_{L^\infty} \lesssim \|R\|_{B_2^{\frac{3}{4},\infty}} \|Y\|_{B_2^{\frac{1}{4},\infty}}.$$

This suffices since each of the the right hand side factors is bounded by $\sqrt{AB}$ by interpolation. To achieve this we write M in two different ways,

$$M = \bar{P}[\bar{R}Y_\alpha - R_\alpha\bar{Y}] + P[R\bar{Y}_\alpha - \bar{R}_\alpha Y] = \partial_\alpha P_{<k+4}(\bar{P}[\bar{R}Y] + P[R\bar{Y}]) - (\bar{R}_\alpha Y + R_\alpha\bar{Y}).$$

We apply a bilinear Littlewood-Paley decomposition and use the first expression above for the high-low interactions, and the second for the high-high interactions, to write $M = M_1 + M_2$ where

$$M_1 = \sum_k [\bar{R}_k Y_{<k,\alpha} - R_{<k,\alpha}\bar{Y}_k] + [R_k \bar{Y}_{<k,\alpha} - \bar{R}_{<k,\alpha} Y_k],$$

$$M_2 = \sum_k \partial_\alpha(\bar{P}[\bar{R}_k Y_k] + P[R_k \bar{Y}_k]) - (\bar{R}_{k,\alpha} Y_k + R_{k,\alpha}\bar{Y}_k).$$

We estimate the terms in M_1 separately; we show the argument for the first:

$$\|\sum_k \bar{R}_k Y_{<k,\alpha}\|_{L^\infty} \lesssim \sum_{j\leq k} 2^{\frac{3}{4}(j-k)} \|R_k\|_{B_2^{\frac{3}{4},\infty}} \|Y_j\|_{B_2^{\frac{1}{4},\infty}} \lesssim \|R\|_{B_2^{\frac{3}{4},\infty}} \|Y\|_{B_2^{\frac{1}{4},\infty}}.$$

For the first term in M_2 we note that the multiplier $\partial_\alpha P_{<k+4} P$ has an $O(2^k)$ L^∞ bound. Hence, we can estimate

$$\|M_2\|_{L^\infty} \lesssim \sum_k 2^k \|R_k\|_{L^\infty} \|Y_k\|_{L^\infty} \lesssim \|R\|_{B_2^{\frac{3}{4},\infty}} \|Y\|_{B_2^{\frac{1}{4},\infty}}.$$

For the L^2 bound we consider again all terms in M_1 and M_2 separately. For an M_1 term we compute

$$\|\sum_k \bar{R}_k Y_{<k,\alpha}\|^2_{\dot{H}^{n-\frac{3}{2}}} \lesssim \sup_k 2^{-2k} \|Y_{<k,\alpha}\|^2_{L^\infty} \cdot \sum_k 2^{(2n-1)k} \|R_k\|^2_{L^2} \lesssim \|Y\|^2_{L^\infty} \|R\|^2_{\dot{H}^{\frac{n-1}{2}}}.$$

For M_2 we compute

$$\|M_2\|^2_{\dot{H}^{n-\frac{3}{2}}} \lesssim \sum_k 2^{(2n-1)k} \|Y_k R_k\|^2_{L^2} \lesssim \|Y\|^2_{L^\infty} \|R\|^2_{\dot{H}^{\frac{n-1}{2}}}.$$

□

Finally, we also need a quadrilinear bound related to the energy estimates:

Lemma 2.9. *The following estimate holds for holomorphic functions R, w and r :*
(B.35)

$$\left|\int \bar{R} r_\alpha \mathfrak{M}_b w_\alpha - \bar{R} w_\alpha \mathfrak{M}_b r_\alpha \, d\alpha\right| \lesssim (\||D|^{\frac{1}{2}} R\|_{BMO} \|b_\alpha\|_{BMO} + \|R_\alpha\|_{BMO} \||D|^{\frac{1}{2}} b\|_{BMO}) \|w\|_{L^2} \|r\|_{\dot{H}^{\frac{1}{2}}}$$

Proof. We denote by I_1 the integral on the left. In a first step we replace the holomorphic multiplication operator $\mathfrak{M}_{\bar{P}b}$ by the corresponding paraproduct operator

$$T_{\bar{P}b} f = \sum_k \bar{P} b_{<k} f_k.$$

Thus, I_1 is replaced by

$$I_1' = \int -\bar{R} r_\alpha T_{\bar{P}b} w_\alpha + \bar{R} w_\alpha T_{\bar{P}b} r_\alpha \, d\alpha.$$

To estimate the difference $I_1' - I_1$ we observe that for holomorphic f we have

$$\mathfrak{M}_{\bar{P}b} f = T_{\bar{P}b} f + P(\sum_k \bar{P} b_k f_k).$$

We use this for $f = w_\alpha$, respectively $f = r_\alpha$. Then

$$I_1 - I_1' = \int -\bar{P}[\bar{R} r_\alpha] P \left[\sum_k \bar{P} b_k w_{k,\alpha}\right] + \bar{P}[\bar{R} w_\alpha] P \left[\sum_k \bar{P} b_k r_{k,\alpha}\right] d\alpha.$$

Applying the bounds in Lemma 2.1 to estimate each of the four factors in L^2 we obtain

$$|I_1 - I_1'| \lesssim (\||D|^{\frac{1}{2}} R\|_{BMO} \|b_\alpha\|_{BMO} + \|R_\alpha\|_{BMO} \||D|^{\frac{1}{2}} b\|_{BMO}) \|w\|_{L^2} \|r\|_{\dot{H}^{\frac{1}{2}}}$$

as needed.

It remains to estimate the integral I_1'. We take a Littlewood-Paley decomposition and denote by k, j, l the frequencies of w, r, respectively b. After canceling the common terms we are left with

$$I_1' = \int \sum_{k \leq l < j} \bar{R}_j r_{j,\alpha} b_l w_{k,\alpha} \, d\alpha - \int \sum_{j \leq l < k} \bar{R}_k r_{j,\alpha} b_l w_{k,\alpha} \, d\alpha := I_2 - I_3.$$

The first sum I_2 can be estimated using only the infinity Besov norms for R_α and $|D|^{\frac{1}{2}} b$,

$$\begin{aligned} |I_2| &\lesssim \|R_\alpha\|_{BMO} \||D|^{\frac{1}{2}} b\|_{BMO} \sum_{k<j} (j-k) 2^{\frac{k-j}{2}} \|r_j\|_{\dot{H}^{\frac{1}{2}}} \|w_k\|_{L^2} \\ &\lesssim \|R_\alpha\|_{BMO} \||D|^{\frac{1}{2}} b\|_{BMO} \|r\|_{\dot{H}^{\frac{1}{2}}} \|w\|_{L^2}. \end{aligned}$$

The argument for I_3 is slightly more involved since we cannot gain rapid decay in $k - j$. Instead, we rewrite it as

$$I_3 = \int \sum_k \bar{R}_k w_{k,\alpha} \sum_{j \leq l} b_l r_{j,\alpha} \, d\alpha - \int \sum_{j,k \leq l} \bar{R}_k w_{k,\alpha} b_l r_{j,\alpha} \, d\alpha := I_3' - I_3''.$$

The first term has a product structure, and we can bound each factor in L^2 using Lemma 2.1 to obtain

$$|I_3'| \lesssim \|R_\alpha\|_{BMO} \||D|^{\frac{1}{2}} b\|_{BMO} \|r\|_{\dot{H}^{\frac{1}{2}}} \|w\|_{L^2}.$$

The second term is bounded in the same manner as I_2,

$$\begin{aligned} |I_3''| &\lesssim \||D|^{\frac{1}{2}} R\|_{BMO} \|b_\alpha\|_{BMO} \sum_{j,k \leq l} 2^{\frac{j-l}{2} + \frac{k-l}{2}} \|r_j\|_{\dot{H}^{\frac{1}{2}}} \|w_k\|_{L^2} \\ &\lesssim \||D|^{\frac{1}{2}} R\|_{BMO} \|b_\alpha\|_{BMO} \|r\|_{\dot{H}^{\frac{1}{2}}} \|w\|_{L^2}. \end{aligned}$$

Thus, the proof of (B.35) is concluded.

□

References

[1] T. Alazard, N. Burq and C. Zuily, On the Cauchy problem for gravity water waves, *Preprint*, 2012, http://arxiv.org/abs/1212.0626.

[2] T. Alazard, N. Burq and C. Zuily, On the water waves equations with surface tension, *Duke Math. J.*, 158, no. 3, 2011, 413-499.

[3] T. Alazard, J.M. Delort, Global solutions and asymptotic behavior for two dimensional gravity water waves, *Preprint*, 2013.

[4] H. Christianson, V. Hur, G. Staffilani, Local smoothing effects for the water wave problem with surface tension, *Comm. Partial Differential Equations*, 35, 2010, 12, 2195-2252.

[5] D. Christodoulou, H. Lindblad, On the motion of the free surface of a liquid, *Comm. Pure Appl. Math.*, 53(12), 2000, 1536-1602.
[6] R. R. Coifman and Y. Meyer, Le double commutateur, *Analyse Harmonique d'Orsay*, 1976
[7] D. Coutand, S. Shkoller, Wellposedness of the free-surface incompressible Euler equations with or without surface tension. *J. Am. Math. Soc.*, 20(3), 2007, 829-930.
[8] D. Lannes, Well-posedness of the water-wave equations. *J. Am. Math. Soc.*, 18, 2005, 605-654.
[9] A. I. Dyachenko, E. A. Kuznetsov, M. D. Spector, and V. E. Zakharov, Analytical description of the free surface dynamics of an ideal fluid (canonical formalism and conformal mapping),*Phys Lett A*, 221, 1996, 73-79.
[10] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong, Long time solutions for a Burgers-Hilbert equation via a modified energy method, *Proceedings of the AMS*, 2013
[11] J. K. Hunter, M. Ifrim, Enhanced lifespan of smooth solutions of a Burgers-Hilbert equation, *SIAM Journal on Mathematical Analysis*, Volume 44, Issue 3, 2012, 1279-2235.
[12] Alexandru D. Ionescu and Fabio Pusateri, Nonlinear fractional Schrödinger equations in one dimension. *J. Funct. Anal.*, 266(1):139–176, 2014.
[13] A. D. Ionescu and F. Pusateri, Global solutions for the gravity water waves system in 2D, *Preprint*, 2013,http://arxiv.org/pdf/1303.5357.pdf.
[14] H. Lindblad, Well-posedness for the motion of an incompressible liquid with free surface boundary, *Ann. Math.*, 162(1), 2005, 109-194.
[15] V. I. Nalimov, The Cauchy-Poisson problem. *Dyn. Splosh. Sredy*, 18, 1974, 104-210.
[16] M. Ogawa, A. Tani, Free boundary problem for an incompressible ideal fluid with surface tension, *Math. Models Methods Appl. Sci.*, 12(12), 2002, 1725-1740.
[17] L. V. Ovsjannikov, To the shallow water foundation, *Arch. Mech. Stos.*, 26, 1974, 407-422.
[18] J. Shatah, Normal forms and quadratic nonlinear Klein-Gordon equations, *Comm. Pure Appl. Math.*, 38, 1985, 685-696.
[19] J. Shatah, C. Zeng, Geometry and a priori estimates for free boundary problems of the Eulers equation, *Comm. Pure Appl. Math.*, 61(5), 2008, 698-744.
[20] G. I. Taylor, The instability of liquid surfaces when accelerated in a direction perpendicular to their-planes *I. Proc. R. Soc. Lond.*, A 201, 1950, 192-196.
[21] H. Yosihara, Gravity waves on the free surface of an incompressible perfect fluid of finite depth, *RIMS Kyoto*, 18, 1982, 49-96.
[22] S. Wu, Well-posedness in Sobolev spaces of the full water wave problem in 2-D, *Invent. Math.*, 130, 1997, 39-72.
[23] S. Wu, Almost global wellposedness of the 2-D full water wave problem, *Invent. Math.*, 177, 2009, 45-135.
[24] V. E. Zakharov, A. I. Dyachenko, and O. A. Vasilyev, New method for numerical simulation of a nonstationary potential flow of incompressible fluid with a free surface, *European Journal of Mechanics B/Fluids*, 21, 2002, 283-291.
[25] P. Zhang, Z. Zhang, On the free boundary problem of 3-D incompressible Euler equations, *Comm. Pure Appl. Math.*, 61(7), 2008, 877-940.

Department of Mathematics, University of California at Davis
E-mail address: hunter@math.davis.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