

Two-dimensional gravity waves at low regularity I: Energy estimates

Albert Ai, Mihaela Ifrim, and Daniel Tataru

Abstract. This article represents the first installment of a series of papers concerned with low regularity solutions for the water wave equations in two space dimensions. Our focus here is on sharp cubic energy estimates. Precisely, we introduce and develop the techniques to prove a new class of energy estimates, which we call *balanced cubic estimates*. This yields a key improvement over the earlier cubic estimates of Hunter–Ifrim–Tataru (2016), while preserving their scale-invariant character and their position-velocity potential holomorphic coordinate formulation. Even without using any Strichartz estimates, these results allow us to significantly lower the Sobolev regularity threshold for local well-posedness, drastically improving earlier results obtained by Alazard–Burq–Zuily (2014, 2018), Hunter–Ifrim–Tataru (2016) and Ai (2019).

Contents

1. Introduction	541
2. Norms and multilinear estimates	552
3. Water wave related bounds	558
4. The linearized equations	569
5. Bounds for the linearized equation	572
6. Energy estimates for the full equation	601
7. Local well-posedness	606
References	612

1. Introduction

We consider the two-dimensional water wave equations with infinite depth and with gravity but without surface tension. This is a free boundary problem which is governed by the incompressible Euler equations within the fluid domain, and with appropriate boundary conditions on the water surface, which is the free boundary. Under the additional assumption that the flow is irrotational, the two-dimensional dynamics can be expressed in terms

Mathematics Subject Classification 2020: 76B15 (primary); 35Q31 (secondary).

Keywords: water waves, normal forms, balanced energies.

of a one-dimensional evolution of the free boundary coupled with the trace of the velocity potential on the surface.

The choice of the parametrization of the free boundary plays an important role here, and can be viewed as a form of gauge freedom. Historically, there are three such choices of coordinates: The first two, namely the Eulerian and Lagrangian coordinates, arise in the broader context of fluid dynamics. The third employs the so-called conformal method, which is specific to two-dimensional irrotational flows; this leads to what we call the holomorphic coordinates, which play a key role in the present paper.

Our objective in this and subsequent papers is to improve, streamline, and simplify the analysis of the two-dimensional gravity wave equations. This is a challenging quasilinear, nonlocal, nondiagonal system. We aim to develop its analysis in multiple ways, including

- (1) proving better, scale-invariant energy estimates,
- (2) improving the existing results on long time solutions, and
- (3) refining the study of the dispersive properties and improving the low regularity theory.

In the present article we carry out the first step of this program, and obtain a new class of sharp, scale-invariant energy estimates, which we call *balanced cubic energy estimates*.

As a consequence of our estimates, we are able to drastically lower the regularity threshold below the prior results in [1, 6, 15]. This is despite the fact that here we are using no Strichartz estimates, whose investigation is left for a future installment; thus, this should be seen as an improvement over [15], with further improvements yet to come. On the other hand, for reference we note that the second step of our program, namely the application of these results to the study of long time solutions, has subsequently been carried out in [3].

We note that a family of cubic, scale-invariant energy estimates has already been obtained in [15]. The key improvement here is that our new estimates are *balanced*, as opposed to the unbalanced ones in [15]; this will be further explained below. We emphasize the fact that we also prove balanced cubic estimates for the linearized equation; this is essential for the local well-posedness result.

The proof of these new energy estimates brings together and refines the two main methods in the study of the long time dynamics for water waves, namely (i) the *modified energy method* of the last two authors and collaborators [15, 16], and (ii) the *paradiagonalization method* of Alazard–Delort [7]. This reflects our dual goal, which is both to prove the new energy estimates, and to reduce the study of the nonlinear equation to the corresponding linear paradifferential equation. The last part, which here could simply be seen as an element of the proof of the energy estimates, will become critical in the study of the dispersive properties in later work.

To conclude, we emphasize that the idea of balanced cubic energy estimates, first developed here, should not be seen as specific to water waves, or even to one-dimensional flows, but rather as a much more general principle. For instance, the reader may also look

at our subsequent work [4], where this idea is implemented for the hyperbolic minimal surface equation in all dimensions.

1.1. A brief history

The water wave equations in general, and the equations for gravity waves in particular, are fundamental in fluid dynamics and have received considerable attention over the years. While the following discussion unavoidably gives priority to the dynamical problem, we note that extensive work was also done toward understanding periodic and solitary waves. We refer the reader to [11, 14] for an overview of work in this direction.

The first steps toward understanding the local theory for this problem are due to Nalimov and Ovsjannikov; see [19, 20] and references therein. Ovsjannikov primarily considered solutions to shallow water equations in spaces of analytic functions. Nalimov instead worked on the infinite depth problem, which is also the one considered here, and proved the first small data result in Sobolev spaces. His approach was later extended to the finite depth problem by Yosihara [29], and to large data by Wu [26].

One key observation in the study of water waves was that one can fully describe the evolution in terms of the free boundary and the trace of the velocity potential on the free boundary. This is classical and possibly goes all the way back to Stokes. Further progress was made in the work of Zakharov [30], where the Hamiltonian structure of the problem was first uncovered; this eventually led to the current Eulerian formulation of the problem, which first appeared in [12]. More recently, the Eulerian setting was further exploited in the study of the local well-posedness; for further references we refer the reader to the more recent paper of Alazard–Burq–Zuily [5], as well as to Lannes’s book [17].

The well-posedness result in [5] is based purely on energy estimates. In the two-dimensional setting this was improved in [15] using refined energy estimates, which exploit more of the structure of the problem; this required the use of holomorphic coordinates, which are discussed in the next subsection. Further improvements to the local theory in [6] were based on Strichartz estimates, following the model developed initially for the quasilinear wave equation by Tataru [23–25] and Bahouri–Chemin [8, 9]. The Strichartz estimates, and implicitly the result of [6], were further improved by the first author in [2], in all dimensions, and even further in [1] in two dimensions. This latter work represents in some sense the water wave analogue of the results of Smith–Tataru [22] in the nonlinear wave equation context.

1.2. Water waves in holomorphic (conformal) coordinates

The conformal method first appeared in the study of the stationary problem for solitary waves; one such example is in the work of Levi-Civita [18], but possibly also earlier. Ovsjannikov’s work [21] is the first that we are aware of where the conformal method is used in the study of the evolution problem. The complete evolution equations restricted to the boundary were independently written by Wu [26] and Dyachenko–Kuznetsov–Spector–Zakharov [13] in slightly different forms. However, it was only in Wu’s paper [26] that

this formulation was fully exploited to prove local well-posedness in the large data problem. Finally, the formulation of gravity waves in holomorphic coordinates was revisited in [15], and it is that formulation that we will be using in the present paper.

Our results apply in 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 avoid distracting technicalities we will do the analysis for the real line, and refer the reader to the discussion in [15] concerning the (minor) changes in the periodic case.

To write the equations we use the Hilbert transform H , as well as the projection operator to negative frequencies,

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

Our main variables (Z, Q) are functions of (t, α) which can be described as follows:

- Z denotes the trace of the conformal map on the real line, which is viewed as the boundary of the lower half-space. In other words, Z represents the conformal parametrization of the free surface in complex notation.
- Q is the holomorphic velocity potential, defined as the trace on the real line of the holomorphic extension of the velocity potential ϕ , i.e. $\Re Q = \phi$.

Both Z and Q are complex-valued functions which satisfy an additional spectral condition. Precisely, both $Z - \alpha$ and Q will be restricted to the closed subspace of holomorphic functions within various Sobolev spaces. Here we define holomorphic functions on \mathbb{R} as those whose Fourier transform is supported in $(-\infty, 0]$; equivalently, they admit a bounded holomorphic extension into the lower half-space. This can be described by the relation $Pf = f$.

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 consider waves which decay at infinity,

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

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 water wave equations have the form

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

where the gravity is normalized to 1 and

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

For the derivation of the above equations, we refer the reader to the appendix of [15]. Slightly different forms of these equations appeared in Wu [26] and Dyachenko–Kuznetsov–Spector–Zakharov [13]. The system form of the equations, used here, is closer to [13], but is in complex rather than real form, as in Wu’s work [26].

It is convenient to remove the leading, linear part of Z and work with a new variable, namely

$$W = Z - \alpha.$$

Then the equations become

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

where

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

Note that J represents the Jacobian of the conformal change of coordinates.

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) . However, this system turns out to be degenerate hyperbolic at leading order, so it is far better to diagonalize it. This was done in [15] using the operator

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

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

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

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

$$\begin{cases} D_t \mathbf{W} + \frac{(1 + \mathbf{W})R_\alpha}{1 + \bar{\mathbf{W}}} = (1 + \mathbf{W})M, \\ D_t R = i\left(\frac{\mathbf{W} - a}{1 + \mathbf{W}}\right). \end{cases} \tag{1.2}$$

Here, the *material derivative* D_t is given by

$$D_t := \partial_t + b\partial_\alpha,$$

the *advection velocity* b is given by

$$b := \Re F = P\left[\frac{R}{1 + \bar{\mathbf{W}}}\right] + \bar{P}\left[\frac{\bar{R}}{1 + \mathbf{W}}\right],$$

the real *Taylor coefficient* $1 + a$ is given by

$$a := i(\bar{P}[\bar{R}R_\alpha] - P[R\bar{R}_\alpha])$$

(though we will often abuse terminology by referring to a as the Taylor coefficient), and the auxiliary function M , closely related to the material derivative of a , is given by

$$M := \frac{R_\alpha}{1 + \bar{\mathbf{W}}} + \frac{\bar{R}_\alpha}{1 + \mathbf{W}} - b_\alpha = \bar{P}[\bar{R}Y_\alpha - R_\alpha\bar{Y}] + P[R\bar{Y}_\alpha - \bar{R}_\alpha Y]. \tag{1.3}$$

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 (1.2) governs an evolution in the space of holomorphic functions, and will be used both directly and in its projected version.

Observe that using these definitions, (1.1) may be expressed as

$$\begin{cases} D_t W + b = \bar{R}, \\ D_t Q - iW = \bar{P}\left[\frac{|Q_\alpha|^2}{J}\right], \end{cases}$$

and (1.2) may be expressed as

$$\begin{cases} D_t Y + |1 - Y|^2 R_\alpha = (1 - Y)M, \\ D_t R - i(1 + a)Y = -ia. \end{cases} \tag{1.4}$$

The functions b and a also play fundamental roles in the linearized equations which are computed in Section 4. Indeed, these functions are essential to understanding the quasi-linear evolution, and appear in one form or another in all works on the subject; see e.g. [5, 17, 19, 27]. Our expressions here are similar in form to those in [27].

We denote the linearized variables around a solution (W, Q) of (1.1) by (w, q) and, after the diagonalization,

$$(w, r := q - R w).$$

The linearized equations (see (4.2)) have the form

$$\begin{cases} D_t w + \frac{1}{1 + \bar{\mathbf{W}}} r_\alpha + \frac{R_\alpha}{1 + \bar{\mathbf{W}}} w = (1 + \mathbf{W})(P\bar{m} + \bar{P}m), \\ D_t r - i\frac{1 + a}{1 + \mathbf{W}} w = \bar{P}n - P\bar{n}, \end{cases}$$

where

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

In particular, the linearization of the system (1.2) around the zero solution is

$$\begin{cases} w_t + r_\alpha = 0, \\ r_t - iw = 0. \end{cases} \tag{1.5}$$

An important role in our paper is played by the paradifferential counterpart of the linearized equations, which is the linear flow

$$\begin{cases} T_{D_t} w + T_{1-\bar{Y}} \partial_\alpha r + T_{(1-\bar{Y})R_\alpha} w = 0, \\ T_{D_t} r - iT_{1-Y} T_{1+a} w = 0, \end{cases} \tag{1.6}$$

where T_f denotes the standard paraproduct operator and the operator T_{D_t} , which we will call the *para-material derivative*, is given by

$$T_{D_t} := \partial_t + T_b \partial_\alpha.$$

In a similar vein, we will also consider the linearization of the equations (1.2) and its associated linear paradifferential flow, which we list here for reference in inhomogeneous form:

$$\begin{cases} T_{D_t} w + T_{b_\alpha} w + \partial_\alpha T_{1-\bar{Y}} T_{1+W_\alpha} r = G, \\ T_{D_t} r + T_{b_\alpha} r - iT_{(1-Y)^2} T_{1+a} w + T_M r = K. \end{cases} \tag{1.7}$$

The two sets of linearized equations are derived in Section 4, where we also describe the connection between them. The analysis of the linearized equations (4.4), carried out in Section 5, is a key component of this paper. One intermediate step in this analysis is the study of the above paradifferential equation (1.6).

The second paradifferential equation (1.7) will play a role in the study of the full differentiated equation (1.2). The two paradifferential flows above will be shown to be equivalent in a suitable sense.

1.3. Function spaces and control norms

The system (1.5) is a well-posed linear evolution in the space \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

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

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

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

As suggested by the above energy, our main function spaces for the differentiated water wave system (1.2) are the spaces \mathcal{H}^s endowed with the norm

$$\|(\mathbf{W}, R)\|_{\mathcal{H}^s}^2 := \|\langle D \rangle^s (\mathbf{W}, R)\|_{L^2 \times \dot{H}^{\frac{1}{2}}}^2,$$

where $s \in \mathbb{R}$. With this notation, the linear energy E_0 is equivalent to (the square of) the \mathcal{H}^0 norm.

Almost all the estimates in this paper are scale invariant, and to describe them it is very useful to also have homogeneous versions of the above spaces, namely the spaces $\dot{\mathcal{H}}^s$ endowed with the norm

$$\|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}^s}^2 := \||D|^s(\mathbf{W}, R)\|_{L^2 \times \dot{H}^{\frac{1}{2}}}^2.$$

We caution the reader that, in order to streamline the exposition here, our notation for the energy spaces differs slightly from the notation used in [15].

The energy estimates for the solutions in [15] were described in terms of the (time-dependent) control norms (A, B) , which are defined and redented here as

$$A_0 = A := \|\mathbf{W}\|_{L^\infty} + \|Y\|_{L^\infty} + \||D|^{\frac{1}{2}}R\|_{L^\infty \cap B_{\infty,2}^0}, \tag{1.8}$$

respectively

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

Instead, in this article the leading role will be played by an intermediate control norm interpolating between A_0 and $A_{\frac{1}{2}}$,

$$A_{\frac{1}{4}} := \|\mathbf{W}\|_{\dot{B}_{\infty,2}^{\frac{1}{4}}} + \|R\|_{\dot{B}_{\infty,2}^{\frac{3}{4}}}. \tag{1.9}$$

Here, the subscript of A represents the difference in terms of derivatives between our control norm and scaling. In particular, A_s corresponds to and is controlled by the homogeneous $\dot{\mathcal{H}}^{\frac{1}{2}+s}$ norm of (\mathbf{W}, R) , and A_0 is a scale-invariant quantity. Concerning $A_{\frac{1}{4}}$, we note the following inequality:

$$\||D|^{\frac{1}{4}}\mathbf{W}\|_{\text{BMO}} + \||D|^{\frac{3}{4}}R\|_{\text{BMO}} \lesssim A_{\frac{1}{4}}. \tag{1.10}$$

With one notable exception, in all of the paper we use only this slightly weaker BMO bound.

In addition to the pointwise scale-invariant norm measured by A , we will also need a stronger scale-invariant Sobolev control norm A^\sharp defined by

$$A^\sharp := \||D|^{\frac{1}{4}}\mathbf{W}\|_{L^4} + \||D|^{3/4}R\|_{L^4}. \tag{1.11}$$

A number of implicit constants in our estimates will depend on A^\sharp .

1.4. The main results

For comparison purposes, we begin by recalling the energy estimates in [15] for the differentiated system (1.2).

Theorem 1. *For each $n \geq 0$ there exists an energy functional E_n associated to the differentiated equation (1.2) with the following two properties:*

(i) *Energy equivalence:*

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

(ii) *Cubic energy bound:*

$$\frac{d}{dt} E_n(\mathbf{W}, R) \lesssim_A AB \|(\mathbf{W}, R)\|_{\mathcal{H}^n}^2.$$

The above energy bounds are already scale invariant, so one cannot hope to simply lower the regularity of any of the factors. However, the two factors in the product AB are unbalanced, with the second factor $B = A_{\frac{1}{2}}$ being the critical one in determining low regularity well-posedness thresholds.

A main idea in this article is that, by taking full advantage of the structure of the water wave system, one can better rebalance the product AB into $A_{\frac{1}{4}}$. Along the way, we also allow n (now redenoted by s) to be noninteger.

Theorem 2. *For each $s \geq 0$ there exists an energy functional E_s associated to the differentiated equation (1.2) with the following two properties:*

(i) *Energy equivalence if $A \ll 1$:*

$$E_s(\mathbf{W}, R) \approx \|(\mathbf{W}, R)\|_{\mathcal{H}^s}^2.$$

(ii) *Balanced cubic energy bound:*

$$\frac{d}{dt} E_s(\mathbf{W}, R) \lesssim_{A^\sharp} A_{\frac{1}{4}}^2 \|(\mathbf{W}, R)\|_{\mathcal{H}^s}^2.$$

One should think of the energies E_s as refined versions of the earlier energies E_n in Theorem 1. In particular, E_s would satisfy the conclusion of Theorem 1, but not vice versa.

In the same vein, in this article we also prove a similar balanced energy bound for the linearized equation (4.4) in the space $\mathcal{H}^{\frac{1}{4}}$; see Theorem 7 in Section 5. This improves a corresponding unbalanced bound in [15] (done there in \mathcal{H}^0). A small price to pay is the use of the stronger control norm A^\sharp instead of A in the implicit constant; this has no other impact in either this paper or the following installments of our work.

As a consequence of the above energy bound and of the similar bound for the linearized equation, we will prove the following low regularity well-posedness result.

Theorem 3. *Let $s \geq s_0 = 3/4$. The system (1.2) is locally well posed for all initial data (\mathbf{W}_0, R_0) in $\mathcal{H}^s(\mathbb{R})$ (or \mathbb{T}) so that $A(0) = A(\mathbf{W}_0, R_0)$ is small. Further, the solution can be continued for as long as A remains small and $A_{\frac{1}{4}} \in L_t^2$.*

We complete this theorem with two remarks which discuss different possible formulations of this result.

Remark 1.1. The smallness assumption on A is not essential here, but it does allow us to avoid a number of distracting technicalities and instead focus on the main new ideas of

the present paper. One should be able to replace this with a bound of the form $|1 + \mathbf{W}_0| > c > 0$, which in a nutshell prevents the free surface from developing corners. To see the potential effect of eliminating the smallness assumption on A we refer the reader to the local well-posedness argument in [15].

Remark 1.2. The $\dot{\mathcal{H}}^{\frac{3}{4}}$ regularity of the data and of the solution is needed in this paper in order to provide uniform control over $A_{\frac{1}{4}}$. On the other hand, the low frequency regularity \mathcal{H}^0 is actually never used, but simply propagated in time. With some extra work in the local well-posedness argument one could relax the low frequency assumption and prove well-posedness in the larger space $\dot{\mathcal{H}}^{\frac{1}{4}} \cap \dot{\mathcal{H}}^{\frac{3}{4}}$, and possibly even better than that.

It is revealing to compare this result with prior work. The direct comparison is with the results in [5] and [15], both of which are purely based on energy estimates. These results correspond to $s = 1 + \varepsilon$, respectively $s = 1$. Thus our improvement is by $1/4$ of a derivative, which is halfway to the absolute threshold $s = \frac{1}{2}$ given by scaling.

One could also compare with the results of [6] with $s = 1 - \frac{1}{24} + \varepsilon$ or [1] with $s = 1 - \frac{1}{8} + \varepsilon$, and our new result is still considerably better. However, such a comparison would be biased because both [6] and [1] use Strichartz, while here we do not. Strichartz-based improvements to the above result will be discussed in a subsequent paper, where another substantial improvement is obtained for the well-posedness range of s .

A limitation of our exposition is that we work with the infinite depth problem, whereas the local well-posedness threshold should not depend on that. Indeed, we expect that this result easily carries over to the finite depth case.

One interesting consequence of the above energy estimates and well-posedness result is that we can also prove a low regularity cubic lifespan bound.

Theorem 4. *Let $\varepsilon \ll 1$, $0 < \delta \leq \frac{1}{2}$, and $c_0 \ll 1$. Assume that the initial data for the equation (1.2) satisfies*

$$\|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}^{\frac{3}{4}}} \leq \varepsilon,$$

as well as

$$\|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}^{\frac{1}{2}-\delta}}^{\frac{1}{4}} \|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}^{\frac{3}{4}}}^{\delta} \leq c_0^{\delta+\frac{1}{4}}.$$

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

$$\sup_{t \in I_\varepsilon} \|(\mathbf{W}(t), R(t))\|_{\dot{\mathcal{H}}^s} \lesssim \|(\mathbf{W}(0), R(0))\|_{\dot{\mathcal{H}}^s}$$

hold for all $s \geq 0$ whenever the right-hand side is finite.

We note that in the current version we have slightly adjusted the statement of this theorem in order to make it fully scale invariant and to more accurately reflect the exact consequences of our main result in Theorem 2. This does not change the (very brief) proof in any significant way. This is in response to a discussion with S. Wu, in order to facilitate a direct comparison of this result with the more recent results in [28]. The constant c_0

is universal, and comes from the smallness requirement on A in Theorem 2 (see also Remark 1.1). The power of c_0 is simply used for homogeneity purposes. See also [10] for related results.

Another key point of this paper, which is less important for the results stated above but will be critical in our follow-up work, is that we establish a normal form based equivalence between our differentiated variables (\mathbf{W}, R) and normalized variables $(\tilde{\mathbf{W}}, \tilde{R})$, which in turn solve a paradifferential form of the linearized equations (4.4), with a favorable perturbative source term.

Theorem 5. *There exists a normal form transformation*

$$(\mathbf{W}, R) \rightarrow (\mathbf{W}_{\text{NF}}, R_{\text{NF}})$$

with the following properties:

- (i) **Regularity:** *The normal form transformation is smooth in \mathcal{H}^s for all $s > \frac{1}{2}$, with comparable control parameters $A, A^\sharp, A_{\frac{1}{4}}$ for (\mathbf{W}, R) and $(\mathbf{W}_{\text{NF}}, R_{\text{NF}})$.*
- (ii) **Uniform bounds:** *For each $s \geq 0$ we have*

$$\|(\mathbf{W}, R) - (\mathbf{W}_{\text{NF}}, R_{\text{NF}})\|_{\dot{\mathcal{H}}^s} \lesssim_A A \|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}^s}.$$

- (iii) **Normalized equation:** *The normal form variables $(\mathbf{W}_{\text{NF}}, R_{\text{NF}})$ solve an inhomogeneous paradifferential equation (1.7) with sources (G, K) satisfying*

$$\|(G, K)\|_{\dot{\mathcal{H}}^s} \lesssim_A A_{\frac{1}{4}}^2 \|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}^s}, \quad s \geq 0.$$

From a normal forms perspective, one might ask whether the results in this paper could be achieved directly using a normal form transformation. As far as the low frequency analysis is concerned, there is indeed a very simple normal form transformation which eliminates the quadratic terms in equation (1.1). As computed in [15], this has the form

$$\begin{cases} \tilde{W} &= W - 2P(\Re W W_\alpha), \\ \tilde{Q} &= Q - 2P(\Re W Q_\alpha). \end{cases} \tag{1.12}$$

Unfortunately, this normal form transformation is unbounded at high frequency, which simply reflects the quasilinear character of the problem.

To compound the difficulty, if one attempts to differentiate this in order to obtain a normal form transformation for the differentiated equation (1.2), this will contain inverse derivatives, so it will be unbounded also at low frequencies.

While we cannot use the above normal form directly, we will nevertheless rely on it as a guide when computing partial normal form transformations, both in its differentiated form for equation (1.2), and in linearized form for the linearized equation (4.4).

1.5. The structure of the paper

Much of the analysis in this paper is phrased in the language of paradifferential calculus. To prepare for this, we begin in Section 2 with a review of notation and of some of the

classical Coifman–Meyer paraproduct estimates, some further bounds from [15], as well as several new bounds.

This is immediately followed and applied in Section 3 by a series of bounds which are specific to the water wave system. We successively consider the Taylor coefficient a , the advection velocity b , as well as several other auxiliary functions.

Next we turn our attention to linearizations and para-linearizations. Our goal in Section 4 is to derive the linearizations of the original equations (1.1) and of the differentiated equations (1.2), as well as the connection between them. To each of these two linearized equations we associate the respective paradifferential flows.

After that, in Section 5, we study the well-posedness of the first linearized equations (4.4) in the space $\mathcal{H}^{\frac{1}{4}}$, as well as the corresponding balanced energy estimates. As an intermediate step, we prove a similar result for the corresponding paradifferential flow (1.6), and carry out a modified normal form type reduction.

In Section 6 we consider the energy estimates in Sobolev spaces $\dot{\mathcal{H}}^s$, $s \geq 0$, for the solutions to the nonlinear equations. The principal part of these equations is closely related to the associated paradifferential linearized equations (1.7). The bulk of the work is to show that there exists a modified normal form type reduction of the full equations (1.2) to the paradifferential ones (1.7). The latter is then shown to be equivalent to the paradifferential flow (1.6), studied in the previous section.

Section 7 contains the proof of the local well-posedness result in Theorem 3 for the equations (1.2), as well as the proof of the cubic lifespan bounds in Theorem 4. We begin with more regular data, e.g. \mathcal{H}^1 , for which well-posedness has already been proved in [15]. The rough $\mathcal{H}^{\frac{3}{4}}$ 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 the data.

2. Norms and multilinear estimates

Here we review some of the function spaces and estimates used later in the paper. We are primarily using the paradifferential calculus, and our analysis will require a heavy dose of paraproduct-type estimates. Many of these are relatively standard, like the classical Coifman–Meyer bounds and some of their generalizations. Several others are from [15]. To those we add a few more useful multilinear paraproduct-type bounds.

2.1. Function spaces

We use a standard Littlewood–Paley decomposition in frequency,

$$1 = \sum_{k \in \mathbb{Z}} P_k,$$

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

$$\|f\|_{\dot{H}^s} \sim \left\| \left(\sum_k |2^{ks} P_k f|^2 \right)^{1/2} \right\|_{L^2} = \|2^{ks} P_k f\|_{L^2_\alpha \ell_k^2}.$$

We will also use the Littlewood–Paley square function and its restricted version,

$$S(f)(\alpha) := \left(\sum_{k \in \mathbb{Z}} |P_k(f)(\alpha)|^2 \right)^{1/2}, \quad S_{>k}(f)(\alpha) = \left(\sum_{j>k} |P_j f|^2 \right)^{1/2}.$$

The Littlewood–Paley inequality is recalled below:

$$\|S(f)\|_{L^p(\mathbb{R})} \simeq_p \|f\|_{L^p(\mathbb{R})}, \quad 1 < p < \infty.$$

By duality this also yields the estimate

$$\left\| \sum_{k \in \mathbb{Z}} P_k f_k \right\|_{L^p} \lesssim \left\| \sum_{k \in \mathbb{Z}} (|f_k|^2)^{1/2} \right\|_{L^p}, \quad 1 < p < \infty,$$

for any sequence of functions $\{f_k\}_k \in L^p_\alpha \ell_k^2$.

The $p = 1$ version of the above estimate for the Hardy space H_1 is

$$\|f\|_{H_1} \simeq \|S(f)\|_{L^1_\alpha \ell_k^2},$$

which by duality implies the BMO bound

$$\left\| \sum_{k \in \mathbb{Z}} P_k f_k \right\|_{\text{BMO}} \lesssim \|S(f)\|_{L^\infty}.$$

The square function characterization of BMO is slightly different:

$$\|u\|_{\text{BMO}}^2 \approx \sup_k \sup_{|Q|=2^{-k}} 2^k \int_Q |S_{>k}(u)|^2 dx. \tag{2.1}$$

For real s we define the homogeneous spaces¹ BMO^s with norm

$$\|u\|_{\text{BMO}^s} = \| |D|^s u \|_{\text{BMO}}.$$

¹These are the same as the Triebel–Lizorkin spaces $F_{\infty,2}^s$.

2.2. Coifman–Meyer- 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,

$$fg = \sum_{k \in \mathbb{Z}} f_{<k-4} g_k + \sum_{k \in \mathbb{Z}} f_k g_{<k-4} + \sum_{|k-l| \leq 4} f_k g_l := T_f g + T_g f + \Pi(f, g).$$

We also define the restricted diagonal sum

$$\Pi_{\geq k}(f, g) = \sum_{j \geq k} f_j g_j.$$

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 notation 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

$$\|T_f g\|_{L^r} + \|\Pi(f, g)\|_{L^r} \lesssim \|f\|_{L^p} \|g\|_{L^q}, \quad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \quad 1 < p, q, r < \infty. \tag{2.2}$$

Corresponding to $q = \infty$ one also has a BMO estimate

$$\|T_f g\|_{L^p} + \|\Pi(f, g)\|_{L^p} \lesssim \|f\|_{L^p} \|g\|_{\text{BMO}}, \quad 1 < p < \infty, \tag{2.3}$$

while for the remaining product term we have the weaker bound

$$\|T_g f\|_{L^p} \lesssim \|f\|_{\dot{W}^{s,p}} \|g\|_{\text{BMO}^{-s}}, \quad 1 < p < \infty, \quad s > 0. \tag{2.4}$$

These in turn lead to the commutator bound

$$\|[P, g]f\|_{L^p} \lesssim \|f\|_{L^p} \|g\|_{\text{BMO}}, \quad 1 < p < \infty.$$

We also need an extension of this, namely the following lemma.

Lemma 2.1. *The following commutator estimates hold for $1 < p < \infty$:*

$$\begin{aligned} \| |D|^s [P, g] |D|^\sigma f \|_{L^p} &\lesssim \| |D|^{\sigma+s} g \|_{\text{BMO}} \|f\|_{L^p}, \quad \sigma \geq 0, \quad s \geq 0, \\ \| |D|^s [P, g] |D|^\sigma f \|_{L^p} &\lesssim \| |D|^{\sigma+s} g \|_{L^p} \|f\|_{\text{BMO}}, \quad \sigma > 0, \quad s \geq 0. \end{aligned}$$

We remark that later this is applied to functions which are holomorphic/antiholomorphic, but that no such assumption is made above.

Next we consider some similar product-type estimates involving BMO and L^∞ norms.

Proposition 2.2. *Let u, v be two functions in $L^\infty \cap \text{BMO}^s, s > 0$:*

(a) *The following estimates hold:*

$$\|\Pi(u, v)\|_{\text{BMO}} \lesssim \|u\|_{\text{BMO}} \|v\|_{\text{BMO}}, \tag{2.5}$$

$$\|P_{\leq k} \Pi_{\geq k}(u, v)\|_{L^\infty} \lesssim \|u\|_{\text{BMO}} \|v\|_{\text{BMO}}, \tag{2.6}$$

$$\|T_u v\|_{\text{BMO}} \lesssim \|u\|_{L^\infty} \|v\|_{\text{BMO}}, \tag{2.7}$$

$$\|T_u v\|_{\text{BMO}} \lesssim \|u\|_{\text{BMO}^{-\sigma}} \|v\|_{\text{BMO}^\sigma}, \quad \sigma > 0. \tag{2.8}$$

(b) *For $s > 0$ the space $L^\infty \cap \text{BMO}^s$ is an algebra,*

$$\|uv\|_{\text{BMO}^s} \lesssim \|u\|_{L^\infty} \|v\|_{\text{BMO}^s} + \|v\|_{L^\infty} \|u\|_{\text{BMO}^s}.$$

(c) *In addition, the following Moser estimate holds for a smooth function F vanishing at 0:*

$$\|F(u)\|_{\text{BMO}^s} \lesssim \|u\|_{L^\infty} \|u\|_{\text{BMO}^s}. \tag{2.9}$$

This proposition is a paradifferential reformulation of results from [15], with the exception of the estimate (2.6), which represents a mild strengthening of (2.5) in the hi \times hi \rightarrow low scenario. Its proof is a direct consequence of the square function characterization of BMO in (2.1), and is left for the reader.

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*

$$\|fg\|_{\dot{H}^\sigma} \lesssim \|f\|_{\dot{H}^\sigma} \|g\|_{L^\infty} + \|f\|_{L^\infty} \|g\|_{\dot{H}^\sigma}.$$

In addition, the following Moser estimate holds for a smooth function F vanishing at 0:

$$\|F(u)\|_{\dot{H}^\sigma} \lesssim \|u\|_{L^\infty} \|u\|_{\dot{H}^\sigma}. \tag{2.10}$$

2.3. Paraproduct estimates

Here we record the following para-commutator, paraproduct, and para-associativity lemmas.

Lemma 2.4 (Para-commutators). *Assume that $\gamma_1, \gamma_2 < 1$. Then we have*

$$\|T_f T_g - T_g T_f\|_{\dot{H}^s \rightarrow \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} g \|_{\text{BMO}}$$

and the L^4 -based version

$$\|T_f T_g - T_g T_f\|_{\dot{W}^{s,4} \rightarrow \dot{W}^{s+\gamma_1+\gamma_2,4}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} g \|_{\text{BMO}},$$

as well as the unbalanced version

$$\|T_f T_g - T_g T_f\|_{\text{BMO}^s \rightarrow \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \| |D|^{\gamma_1} f \|_{L^2} \| |D|^{\gamma_2} g \|_{\text{BMO}}.$$

We note that the BMO norms are not strictly speaking necessary; we only need the Besov norms $\dot{B}_{\infty,\infty}^\gamma$. The limiting case $\gamma_j = 1$ is more delicate, as there one would need instead $\dot{W}^{1,\infty}$; see also (2.11) below.

Proof of Lemma 2.4. By orthogonality we can fix the frequency of the argument u ; call it λ . If f is supported at frequency $\ll \lambda$, then $T_f u = f u$. Since multiplication is commutative, the only nontrivial contributions are those where at least one of the two frequencies of f and g is comparable to λ . We consider two cases:

- (i) Both f and g are at frequency $O(\lambda)$. Then we neglect the commutator and estimate the output directly using the $\dot{B}_{\infty,\infty}^{\gamma_j}$ norms.
- (ii) f is at frequency $O(\lambda)$, and g is at frequency $\ll \lambda$. Then the commutator becomes

$$[T_{f_\lambda}, T_{g_{\ll \lambda}}]u_\lambda = [T_{f_\lambda}, g_{\ll \lambda}]u_\lambda = \lambda^{-1}L(f_\lambda, \partial_\alpha g_{\ll \lambda}, u_\lambda),$$

where L denotes a translation-invariant trilinear form with integrable kernel. Hence we obtain

$$\|[T_{f_\lambda}, T_{g_{\ll \lambda}}]u_\lambda\|_{L^2} \lesssim \|f_\lambda\|_{L^\infty} \|\partial_\alpha g_{\ll \lambda}\|_{L^\infty} \|u_\lambda\|_{L^2}$$

and the desired conclusion easily follows. ■

Lemma 2.5 (Paraproducts). *Assume that $\gamma_1, \gamma_2 < 1, \gamma_1 + \gamma_2 \geq 0$. Then*

$$\|T_f T_g - T_{fg}\|_{\dot{H}^s \rightarrow \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} g \|_{\text{BMO}}$$

and the L^4 -based version

$$\|T_f T_g - T_{fg}\|_{\dot{W}^{s,4} \rightarrow \dot{W}^{s+\gamma_1+\gamma_2,4}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} g \|_{\text{BMO}}.$$

One may also ask whether this extends also to the case $\gamma_1 = \gamma_2 = 1$. As it turns out, there one would have to use L^∞ norms instead,

$$\|T_f T_g - T_{fg}\|_{\dot{H}^s \rightarrow \dot{H}^{s+2}} \lesssim \|\partial f\|_{L^\infty} \|\partial g\|_{L^\infty}. \tag{2.11}$$

This is however not needed in the present paper.

Proof of Lemma 2.5. Reasoning as in the previous lemma, we can fix the argument u 's frequency to λ , and full cancellation occurs unless at least one of the frequencies of f and g is $\gtrsim \lambda$. Thus we distinguish three cases, the first two of which are the same as before. Case (i) is identical, and also case (ii) in view of the relation

$$[T_{f_\lambda} g_{\ll \lambda} - T_{f_\lambda g_{\ll \lambda}}]u_\lambda = \lambda^{-1}L(f_\lambda, \partial_\alpha g_{\ll \lambda}, u_\lambda).$$

Thus we consider the remaining case:

(iii) The frequency of f is $\gg \lambda$. Then the frequency of g must be comparable. We consider the worst-case scenario, when $\gamma_1 + \gamma_2 = 0$. Then we have

$$(fg)_{<\lambda} = P_{<\lambda} \left(\sum_{\mu>\lambda} f_\mu g_\mu \right)$$

and we can use (2.6) to estimate

$$\|(fg)_{<\lambda}\|_{L^\infty} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} g \|_{\text{BMO}},$$

which suffices. ■

Lemma 2.6 (Para-associativity). *For $s + \gamma_2 \geq 0$, $s + \gamma_1 + \gamma_2 \geq 0$, and $\gamma_1 < 1$ we have*

$$\|T_f \Pi(v, u) - \Pi(v, T_f u)\|_{\dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} v \|_{\text{BMO}} \|u\|_{\dot{H}^s}$$

and the L^4 -based version

$$\|T_f \Pi(v, u) - \Pi(v, T_f u)\|_{\dot{W}^{s+\gamma_1+\gamma_2,4}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} v \|_{\text{BMO}} \|u\|_{\dot{W}^{s,4}}, \quad (2.12)$$

as well as a version where f is measured in an L^2 -based norm

$$\|T_f \Pi(v, u) - \Pi(v, T_f u)\|_{\dot{H}^{1/4}} \lesssim \|f\|_{\dot{H}^{1/4}} \| |D|^{-1/4} v \|_{\text{BMO}} \| |D|^{1/4} u \|_{\text{BMO}}. \quad (2.13)$$

We remark that one could easily expand the range of indices in (2.13), but for simplicity we have limited ourselves to this case, which is used in a single thread in this paper in the study of the linearized equation in Section 4.

Proof of Lemma 2.6. We remark that our convention in this paper will be that the Π operator also contains the projection P . This is immaterial for the purpose of this lemma: We can prove the result without P , and then harmlessly insert it, as it commutes with T_f . We denote by λ the joint frequencies of u and v . Then only frequencies $\leq \lambda$ of f will contribute. The output of frequencies $O(\lambda)$ in f yields no cancellations, but it is easy to bound the two terms separately using two Coifman–Meyer estimates. Thus, in what follows we consider only frequencies $\ll \lambda$ of f , in which case we have

$$(T_{f_{\ll\lambda}} u)_\lambda = (f_{\ll\lambda} u)_\lambda = f_{\ll\lambda} u_\lambda + \lambda^{-1} L(\partial_\alpha f_{\ll\lambda}, u_\lambda).$$

The second term is easy to estimate directly in L^2 ; this leads to the restriction $\gamma' < 1$. For the contribution of the first term we use cancellations to write

$$\sum_\lambda T_{f_{\ll\lambda}}(u_\lambda v_\lambda) - f_{\ll\lambda} u_\lambda v_\lambda = \sum_{\mu \ll \lambda} f_\mu P_{\lesssim \mu}(u_\lambda v_\lambda) = \sum_\lambda \Pi(f_{\ll\lambda}, u_\lambda v_\lambda) + T_{u_\lambda v_\lambda} f_{\ll\lambda}.$$

Harmlessly reinserting frequencies of size λ in f , we can replace the above expression by

$$\Pi(f, \Pi(u, v)) + T_{\Pi(u,v)} f,$$

at which point it suffices to apply the Coifman–Meyer bounds twice.

The proofs of the L^4 -type bounds (2.12) and of (2.13) are similar. ■

We finish with an alternative form of the para-associativity bound.

Lemma 2.7. *We have, for $\gamma_1, \gamma_2 < 1, s + \gamma_1 + \gamma_2 \geq 0, s + \gamma_2 \geq 0$, and $\bar{v} = \bar{P}\bar{v}$,*

$$\|T_f P(\bar{v}u) - P(\bar{v}T_f u)\|_{\dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} v \|_{\text{BMO}} \|u\|_{\dot{H}^s}$$

and the L^4 -based version

$$\|T_f P(\bar{v}u) - P(\bar{v}T_f u)\|_{\dot{W}^{s+\gamma_1+\gamma_2,4}} \lesssim \| |D|^{\gamma_1} f \|_{\text{BMO}} \| |D|^{\gamma_2} v \|_{\text{BMO}} \|u\|_{\dot{W}^{s,4}},$$

as well as

$$\|T_f P(\bar{v}u) - P(\bar{v}T_f u)\|_{\dot{H}^{1/4}} \lesssim \|f\|_{\dot{H}^{1/4}} \| |D|^{-1/4} v \|_{\text{BMO}} \| |D|^{1/4} u \|_{\text{BMO}}.$$

Proof. This is a consequence of Lemmas 2.4 and 2.6, also using the fact that P eliminates the case where \bar{v} is high frequency. ■

3. Water wave related bounds

Here we consider estimates for objects related to the water wave equations, primarily the Taylor coefficient a and advection velocity b . We recall that these are given by

$$a = 2\Im P[R\bar{R}_\alpha], \quad b = 2\Re P\left[\frac{R}{1+\bar{\mathbf{W}}}\right] = 2\Re(P[(1-\bar{Y})R]).$$

In addition to these, we will also consider several auxiliary functions, namely

$$Y = \frac{\mathbf{W}}{1+\mathbf{W}}, \quad X = T_{1-Y}W, \quad M = 2\Re P[R\bar{Y}_\alpha - \bar{R}_\alpha Y].$$

The auxiliary variable Y is often used as a coefficient in our computations. The function X is a convenient leading-order substitute for $\partial^{-1}Y$, and is very useful in some of our normal form computations. The function M is essentially the material derivative of $\ln(1+a)$.

We estimate these in terms of the control parameters A, A^\sharp , and $A_{\frac{1}{4}}$ defined in (1.8), (1.11), and (1.9), and in terms of the H^s Sobolev norms of \mathbf{W} and R .

3.1. L^2 and pointwise bounds

We begin with the auxiliary variable Y , which inherits its regularity from \mathbf{W} due to (2.9) and (2.10).

Lemma 3.1. *The function Y satisfies the BMO bound*

$$\| |D|^{\frac{1}{4}} Y \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}},$$

and the \dot{H}^σ bounds

$$\|Y\|_{\dot{H}^\sigma} \lesssim_A \|\mathbf{W}\|_{\dot{H}^\sigma}, \lesssim_A \|\mathbf{W}\|_{\dot{W}^{\sigma,4}}, \quad \sigma \geq 0.$$

We also have the following paradifferential identities relating W , X , and Y .

Lemma 3.2. *Let W , X , and Y be defined as above:*

(a) *The function Y satisfies*

$$T_{1+W_\alpha} Y = T_{1-Y} W_\alpha - \Pi(Y, W_\alpha).$$

(b) *We have the relations*

$$X_\alpha = T_{1-Y} W_\alpha + \text{Err}_X = T_{1+W_\alpha} Y + \text{Err}_X,$$

where Err_X denotes a varying error term satisfying

$$\|\text{Err}_X\|_{\dot{W}^{\frac{1}{2},4}} \lesssim_{A^\#} A^{\frac{1}{4}}, \quad \||D|^{\frac{1}{2}} \text{Err}_X\|_{\text{BMO}} \lesssim_A A^{\frac{2}{4}}.$$

Proof. (a) It suffices to write the paraproduct expansion

$$Y = (1 - Y)W_\alpha = T_{1-Y} W_\alpha - T_{W_\alpha} Y - \Pi(Y, W_\alpha).$$

(b) For the first identity, by Lemma 3.1, we have

$$-T_{Y_\alpha} W = \text{Err}_X.$$

Then apply (a) to obtain the second identity for Y , where we similarly have

$$\Pi(Y, W_\alpha) = \text{Err}_X. \quad \blacksquare$$

We continue with bounds for a .

Proposition 3.3. *The Taylor coefficient a is nonnegative and satisfies the BMO bound*

$$\|a\|_{\text{BMO}} \lesssim \|R\|_{\text{BMO}^{\frac{1}{2}}}^2, \tag{3.1}$$

and the uniform bound

$$\|a\|_{L^\infty} \lesssim \|R\|_{\dot{B}_{\infty,2}^{\frac{1}{2}}}^2.$$

Moreover,

$$\||D|^{\frac{1}{4}} a\|_{\text{BMO}} \lesssim A A^{\frac{1}{4}}, \quad \||D|^{\frac{1}{2}} a\|_{\text{BMO}} \lesssim A^{\frac{2}{4}}, \tag{3.2}$$

and

$$\|a\|_{\dot{H}^s} \lesssim A \|R\|_{\dot{H}^{s+\frac{1}{2}}}, \quad s \geq 0, \tag{3.3}$$

as well as

$$\||D|^{1/2} a\|_{L^4} \lesssim A^\# A^{\frac{1}{4}}, \quad \|a\|_{\dot{F}_{2,1}^{\frac{1}{2}}} \lesssim A^{\#2}. \tag{3.4}$$

Proof. The first part is from [15], so we only need to prove the bounds from (3.2) on. We recall that

$$a = i(\bar{P}[\bar{R}R_\alpha] - P[R\bar{R}_\alpha]).$$

The bounds of (3.2) are direct consequences of (2.5) and (2.8), by bounding the holomorphic and antiholomorphic parts of a separately, and writing for instance

$$P[R\bar{R}_\alpha] = T_{\bar{R}_\alpha} R + \Pi(\bar{R}_\alpha, R).$$

We next see that (3.3) and the first part of (3.4) are direct consequences of the commutator estimates in Lemma 2.1.

Finally, we consider the second part of (3.4). If the two input frequencies of R and the output frequency are all fixed and equal, then this is a straightforward $L^4 \times L^4 \rightarrow L^2$ bound. It remains to account for the dyadic summation, which is now dealt with in a Triebel–Lizorkin fashion. Precisely, for R we have the Triebel–Lizorkin interpretation of the control norm A^\sharp as $\dot{W}^{\frac{3}{4},4} = \dot{F}^{\frac{3}{4},2}_{4,2}$ with an ℓ^2 dyadic summation. Due to the off-diagonal decay in the bilinear estimate, this yields the corresponding ℓ^1 summability for the output. ■

Next we consider b , for which we have the following result.

Lemma 3.4. *Let $s > 0$. Then the transport coefficient b satisfies*

$$\| |D|^s b \|_{\text{BMO}} \lesssim_A \| |D|^s R \|_{\text{BMO}}, \quad \| |D|^s b \|_{L^4} \lesssim_A \| |D|^s R \|_{L^4}, \quad \| b \|_{\dot{H}^s} \lesssim_A \| R \|_{\dot{H}^s}. \quad (3.5)$$

In particular, we have

$$\| |D|^{\frac{1}{2}} b \|_{\text{BMO}} \lesssim_A A, \quad \| |D|^{\frac{3}{4}} b \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}, \quad \| |D|^{\frac{3}{4}} b \|_{L^4} \lesssim_A A^\sharp. \quad (3.6)$$

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 (2.7), and the high–high case, where (2.5) applies.

A direct argument, taking into account the same two cases, yields the L^2 and the L^4 bounds. ■

Next we consider the auxiliary expression M . For this we have the following bounds.

Lemma 3.5. *The function M satisfies the pointwise bound*

$$\| M \|_{L^\infty} \lesssim_A A_{\frac{1}{4}}^2, \quad (3.7)$$

as well as the Sobolev bounds

$$\| M \|_{\dot{H}^{s-\frac{1}{2}}} \lesssim_A A \| (\mathbf{W}, R) \|_{\mathcal{H}^s}, \quad s \geq 0.$$

We remark that this is the only place in the paper where the Besov norm in $A_{\frac{1}{4}}$ is used. If instead we were to use only the weaker control norm (1.10), then we would only get the slightly weaker bound

$$\|M\|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2.$$

Proof of Lemma 3.5. For the pointwise bound we claim that

$$\|M\|_{L^\infty} \lesssim \|R\|_{\dot{B}_{\infty,2}^{\frac{3}{4}}} \|Y\|_{\dot{B}_{\infty,2}^{\frac{1}{4}}}.$$

To achieve this we write M in two different ways:

$$\begin{aligned} M &= \bar{P}[\bar{R}Y_\alpha - R_\alpha\bar{Y}] + P[R\bar{Y}_\alpha - \bar{R}_\alpha Y] \\ &= \partial_\alpha(\bar{P}[\bar{R}Y] + P[R\bar{Y}]) - (\bar{R}_\alpha Y + R_\alpha\bar{Y}). \end{aligned}$$

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

$$\begin{aligned} M_1 &:= \sum_k [\bar{R}_k Y_{<k,\alpha} - R_{<k,\alpha} \bar{Y}_k] + [R_k \bar{Y}_{<k,\alpha} - \bar{R}_{<k,\alpha} Y_k], \\ M_2 &:= \sum_k \partial_\alpha(\bar{P}[\bar{R}_k Y_k] + P[R_k \bar{Y}_k]) - (\bar{R}_{k,\alpha} Y_k + R_{k,\alpha} \bar{Y}_k). \end{aligned}$$

We estimate the terms in M_1 separately; we show the argument for the first:

$$\left\| \sum_k \bar{R}_k Y_{<k,\alpha} \right\|_{L^\infty} \lesssim \sum_{j \leq k} 2^{\frac{3}{4}(j-k)} \|R_k\|_{L^\infty} \|Y_j\|_{L^\infty} \lesssim \|R\|_{\dot{B}_{\infty,2}^{\frac{3}{4}}} \|Y\|_{\dot{B}_{\infty,2}^{\frac{1}{4}}}.$$

For the first term in M_2 we note that the multiplier $\partial_\alpha P_{<k+4} P$ has an $O(2^k)$ bound in L^∞ . Hence, we can estimate

$$\|M_2\|_{L^\infty} \lesssim \sum_k 2^k \|R_k\|_{L^\infty} \|Y_k\|_{L^\infty} \lesssim \|R\|_{\dot{B}_{\infty,2}^{\frac{3}{4}}} \|Y\|_{\dot{B}_{\infty,2}^{\frac{1}{4}}}.$$

The \dot{H}^s bounds are a direct application of the Coifman–Meyer bounds in (2.3), using only A as in (1.8) for the lower frequency factor and the $\dot{\mathcal{H}}^s$ norm for the higher frequency factor. ■

3.2. Material and para-material derivative bounds

The material derivative $D_t = \partial_t + b\partial_\alpha$ plays a key role in the water wave equations. When applying this to products, we can use the well-known Leibniz rule. On the other hand, at the paradifferential level we replace the material derivative with the *para-material derivative*

$$T_{D_t} = \partial_t + T_b \partial_\alpha.$$

Then it becomes interesting to know whether we have an analogue of the Leibniz rule for paraproducts. We consider this in the next lemma, which we will refer to as the para-Leibniz rule. To state it we define four versions of para-Leibniz errors. The first two apply in the unbalanced case, namely

$$E_L^P(u, v) = T_{D_t} T_u v - T_{T_{D_t} u} v - T_u T_{D_t} v,$$

respectively

$$\tilde{E}_L^P(u, v) = T_{D_t} T_u v - T_{D_t u} v - T_u T_{D_t} v.$$

Correspondingly, in the balanced case we set

$$E_L^\pi(u, v) = T_{D_t} \Pi(u, v) - \Pi(T_{D_t} u, v) - \Pi(u, T_{D_t} v),$$

respectively

$$\tilde{E}_L^\pi(u, v) = T_{D_t} \Pi(u, v) - \Pi(D_t u, v) - \Pi(u, T_{D_t} v).$$

With this notation, our para-Leibniz rule reads as follows.

Lemma 3.6. *Let $E_L^P, \tilde{E}_L^P, E_L^\pi,$ and \tilde{E}_L^π be defined as above:*

- (a) *For the unbalanced Leibniz rule error $E_L^P(u, v)$ we have the bound*

$$\|E_L^P(u, v)\|_{\dot{H}^s} \lesssim A_{\frac{1}{4}} \|u\|_{\text{BMO}^{\frac{1}{4}-\sigma}} \|v\|_{\dot{H}^{s+\sigma}}, \quad \sigma > 0, \tag{3.8}$$

as well as

$$\|E_L^P(u, v)\|_{\dot{H}^s} \lesssim A_{\frac{1}{4}} \|u\|_{\dot{H}^{\frac{1}{4}-\sigma}} \|v\|_{\text{BMO}^{s+\sigma}}, \quad \sigma > 0, \tag{3.9}$$

whereas in the $\sigma = 0$ case the same bounds (3.8), (3.9) hold for $\tilde{E}_L^P(u, v)$ with $\sigma = 0$.

- (b) *For the balanced Leibniz rule error $E_L^\pi(u, v)$ we have the bound*

$$\|E_L^\pi(u, v)\|_{\dot{H}^s} \lesssim A_{\frac{1}{4}} \|u\|_{\text{BMO}^{\frac{1}{4}-\sigma}} \|v\|_{\dot{H}^{s+\sigma}}, \quad \sigma \in \mathbb{R}, s \geq 0. \tag{3.10}$$

In the $\sigma = 0$ case we also have the same bound for $\tilde{E}_L^\pi(u, v)$.

- (c) *We also have the L^4 versions of (3.8) and (3.10):*

$$\|E_L^P(u, v)\|_{\dot{W}^{s,4}} \lesssim A_{\frac{1}{4}} \|u\|_{\text{BMO}^{\frac{1}{4}-\sigma}} \|v\|_{\dot{W}^{s+\sigma,4}}, \quad \sigma > 0,$$

respectively

$$\|E_L^\pi(u, v)\|_{\dot{W}^{s,4}} \lesssim A_{\frac{1}{4}} \|u\|_{\text{BMO}^{\frac{1}{4}-\sigma}} \|v\|_{\dot{W}^{s+\sigma,4}}, \quad \sigma \in \mathbb{R}, s \geq 0.$$

Proof. (a) We begin with the case $\sigma > 0$, where

$$E_L^P(u, v) = T_b T_{\partial_\alpha u} v - T_{T_b \partial_\alpha u} v + [T_b, T_u] \partial_\alpha v.$$

In the first difference we have cancellation for the low frequencies of b , and we are left with a dyadic sum of the form

$$\sum_{j \leq l < k} \partial_\alpha u_j b_l v_k$$

where we can rebalance derivatives. Precisely, if we bound the first two terms in L^∞ and the third in L^2 then we have off-diagonal gain away from $j = l = k$. Hence it suffices to use the ∞ Besov norms for the first two factors.

We next consider the case $\sigma = 0$. We write

$$E_L^p(u, v) = T_b T_{\partial_\alpha u} v - T_{b \partial_\alpha u} v + [T_b, T_u] \partial_\alpha v.$$

For the first difference we use Lemma 2.5, while for the commutator term we use Lemma 2.4.

(b) Here we start from

$$E_L^\pi(u, v) = T_b \Pi(\partial_\alpha u, v) - \Pi(T_b \partial_\alpha u, v) + T_b \Pi(u, \partial_\alpha v) - \Pi(u, T_b \partial_\alpha v).$$

The two differences are identical, and in both cases we can use Lemma 2.6.

For the separate estimate in the case $\sigma = 0$, we bound

$$\Pi(T_{\partial_\alpha u} b, v) + \Pi(\Pi(\partial_\alpha u, b), v)$$

using ∞ Besov norms and (2.8).

(c) The proof of the L^4 -type bounds is similar, using the corresponding L^4 bounds in Lemma 2.6. ■

Finally, we rewrite the equations for (\mathbf{W}, R) in a paradifferential form, and thereby capture identities and estimates for their para-material derivatives. Here we also record the para-material derivatives of X, Y , and a .

Lemma 3.7. *We have the following para-material derivatives:*

(a) *Para-material derivative of W :*

$$T_{D_t} W = -T_{1+W_\alpha} P[(1 - \bar{Y})R] - P \Pi(W_\alpha, b).$$

(b) *Para-material derivative of (\mathbf{W}, R) :*

$$\begin{aligned} T_{D_t} W_\alpha &= -T_{b_\alpha} W_\alpha - P \partial_\alpha [T_{1+W_\alpha} [(1 - \bar{Y})R] + \Pi(W_\alpha, b)], \\ T_{D_t} R &= -T_{R_\alpha} b - \Pi(R_\alpha, b) + i(1 + a)Y - ia. \end{aligned} \tag{3.11}$$

(c) *Leading term of the para-material derivative of (\mathbf{W}, R) :*

$$\begin{aligned} T_{D_t} W_\alpha + T_{1+W_\alpha} T_{1-\bar{Y}} \partial_\alpha R &= G, \\ T_{D_t} R - iT_{1+a} Y &= K, \end{aligned} \tag{3.12}$$

where the source terms (G, K) in (3.12) satisfy the BMO bounds

$$\|G\|_{\text{BMO}} + \|K\|_{\text{BMO}^{\frac{1}{2}}} \lesssim_A A_{\frac{1}{4}}^2 \tag{3.13}$$

and

$$\|K\|_{\text{BMO}^{\frac{1}{4}}} \lesssim_A A_{\frac{1}{4}}. \tag{3.14}$$

Proof. For (a), we have

$$T_{D_t} W = -T_{1+W_\alpha} b - \Pi(W_\alpha, b) + \bar{R}.$$

Observing that we may apply a projection P , which freely passes over low frequency paraproducts as well as T_{D_t} , we obtain the identity.

We proceed with the equation and estimates for W_α . Differentiating the identity from (a),

$$T_{D_t} W_\alpha = -T_{b_\alpha} W_\alpha - P \partial_\alpha [T_{1+W_\alpha} [(1 - \bar{Y})R] + \Pi(W_\alpha, b)].$$

Notice that all terms on the right-hand side have a good balance of derivatives, with a derivative falling on a low frequency variable, with the exception of the case where ∂_α falls on (the high frequency) R . Thus, we rewrite the right-hand side as

$$-T_{1+W_\alpha} T_{1-\bar{Y}} R_\alpha + G,$$

where

$$\begin{aligned} G &= T_{1+W_\alpha} \Pi(\bar{Y}, R_\alpha) - T_{b_\alpha} W_\alpha \\ &\quad - P [T_{W_\alpha} [(1 - \bar{Y})R] + T_{1+W_\alpha} [\bar{Y}_\alpha R] + \partial_\alpha \Pi(W_\alpha, b)]. \end{aligned}$$

As observed, terms in G have a good balance of derivatives and thus satisfy the desired estimate. For instance, repeated application of Proposition 2.2 yields

$$\|T_{1+W_\alpha} \Pi(\bar{Y}, R_\alpha)\|_{\text{BMO}} \lesssim \|W_\alpha\|_{L^\infty} \| |D|^{1/4} \bar{Y} \|_{\text{BMO}} \| |D|^{3/4} R \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2.$$

We proceed with the equation for R . The formula in part (b) is straightforward. Then observe that the only term on the right-hand side with an unfavorable balance of derivatives is $i T_{1+a} Y$. For instance,

$$\|i |D|^{1/2} \Pi(a, Y)\|_{\text{BMO}} \lesssim \|Y\|_{L^\infty} \| |D|^{1/2} a \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2$$

using Lemma 3.3 for the estimate on a . The second BMO bound for K also follows by repeated application of Proposition 2.2. ■

Next we record the para-material derivatives of X and Y .

Lemma 3.8. *We have the following para-material derivatives:*

(a) *Leading term of the para-material derivative of Y :*

$$T_{D_t} Y = -T_{|1-Y|^2} R_\alpha + G =: G' + G, \tag{3.15}$$

where the source term G satisfies the pointwise bounds

$$\|G\|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2$$

and G' satisfies

$$\||D|^{-1/4} G'\|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}.$$

(b) *Para-material derivative of X :*

$$-T_{D_t} X = -T_{T_{1-\bar{Y}} R_\alpha} X + P[(1-\bar{Y})R] + P\Pi(X_\alpha, b) + E_1,$$

where, for $s + \frac{3}{4} \geq 0$,

$$\|E_1\|_{\dot{W}^{\frac{5}{4},4}} \lesssim_{A^\#} A_{\frac{1}{4}}^2, \quad \|E_1\|_{\dot{H}^{s+\frac{3}{4}}} \lesssim_A A_{\frac{1}{4}} \|\mathbf{W}\|_{\dot{H}^s}.$$

(c) *Leading term of the para-material derivative of X :*

$$T_{D_t} X + T_{1-\bar{Y}} R = E_2,$$

where

$$\||D|E_2\|_{\text{BMO}} \lesssim_{A^\#} A_{\frac{1}{4}}^2.$$

(d) *Leading term of the para-material derivative of X_α :*

$$T_{D_t} X_\alpha + T_{1-\bar{Y}} R_\alpha = E_3, \quad D_t X_\alpha + T_{1-\bar{Y}} R_\alpha = E_3,$$

where, for $s + \frac{3}{4} \geq 0$,

$$\|E_3\|_{\text{BMO}} \lesssim_{A^\#} A_{\frac{1}{4}}^2, \quad \|E_3\|_{\dot{H}^{s-\frac{1}{4}}} \lesssim_A A_{\frac{1}{4}} \|(\mathbf{W}, R)\|_{\dot{H}^s}.$$

Proof. (a) First we write down the identity

$$\begin{aligned} T_{D_t} Y + T_{|1-Y|^2} R_\alpha &= -(T_{R_\alpha} |1-Y|^2 + \Pi(R_\alpha, |1-Y|^2)) \\ &\quad - (T_{Y_\alpha} b + \Pi(Y_\alpha, b)) + (1-Y)M. \end{aligned}$$

Observe that the first three terms on the right-hand side have a good balance of derivatives, and thus satisfy the estimate for G . For instance,

$$\|\Pi(Y_\alpha, b)\|_{\text{BMO}} \lesssim \||D|^{1/4} Y\|_{\text{BMO}} \||D|^{3/4} b\|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2.$$

For the last term we use (3.7).

The estimate for G' is straightforward, using (2.8) again.

(b) First we apply the para-Leibniz rule in Lemma 3.6,

$$-T_{D_t} X = T_{D_t Y} W - T_{1-Y} T_{D_t} W + E_1.$$

When the material derivative falls on the low frequency Y , we use (1.4) to write

$$T_{D_t Y} W = T_{-|1-Y|^2 R_\alpha + (1-Y)M} W = -T_{T_{1-\bar{Y}} R_\alpha} X + E_1.$$

Here, for the term $T_{-|1-Y|^2 R_\alpha} W$, we use Lemma 2.5 to rewrite W in terms of X , and Proposition 2.2 to peel off perturbative components, where the (differentiated) R_α is the lowest frequency. For the term $T_{(1-Y)M} W$, on the other hand, we directly use the bound (3.7).

When the material derivative falls on the high frequency, we obtain

$$-T_{1-Y} T_{D_t} W = T_{1-Y} (T_{1+W_\alpha} P[(1-\bar{Y})R] + P \Pi(W_\alpha, b)).$$

Applying Lemmas 2.6 and 2.5, and Lemma 3.2 on the last term, yields (b).

(c) All terms from (b) have a good balance of derivatives and may be absorbed into E_2 , except

$$P((1-\bar{Y})R) = T_{1-\bar{Y}} R - \Pi(\bar{Y}, R),$$

the latter of which may also be absorbed into E_2 .

(d) This is a straightforward consequence of (b) and (c). ■

Lastly, we record the para-material derivative of a .

Lemma 3.9. *We have the leading terms of the material and para-material derivatives of a ,*

$$\begin{aligned} D_t a &= -(1+a)M + E, \\ T_{D_t} a &= -T_{1+a}M + E, \end{aligned} \tag{3.16}$$

where the errors E satisfy

$$\|E\|_{\dot{W}^{\frac{1}{4},4} \cap L^\infty} \lesssim_{A^\#} A_{\frac{1}{4}}^2.$$

In particular, we have

$$\|D_t a\|_{L^\infty} \lesssim_{A^\#} A_{\frac{1}{4}}^2. \tag{3.17}$$

Proof. We reduce the first relation in (3.16) to the second by replacing products with paraproducts. On the left, we need to estimate

$$\|\Pi(b, a_\alpha)\|_{\dot{W}^{\frac{1}{4},4} \cap L^\infty} + \|T_{a_\alpha} b\|_{\dot{W}^{\frac{1}{4},4} \cap L^\infty} \lesssim_{A^\#} A_{\frac{1}{4}}^2, \tag{3.18}$$

and on the right we need

$$\|\Pi(M, a)\|_{\dot{W}^{\frac{1}{4},4} \cap L^\infty} + \|T_M a\|_{\dot{W}^{\frac{1}{4},4} \cap L^\infty} \lesssim_{A^\#} A_{\frac{1}{4}}^2, \tag{3.19}$$

where, slightly abusing notation, Π does not include a projection P .

The $\dot{W}^{\frac{1}{4},4}$ part of (3.18) follows from (3.2) and the second part of (3.5) via (2.2) and (2.4). The L^∞ part of (3.18) follows from (3.6) and (3.1) and (3.2). Here we need to take advantage of the fact that we have off-diagonal decay, interpolating between the two pairs of bounds for both a and b in order to gain dyadic summability. The bound (3.19) is a consequence of the L^∞ bound (3.7) for M and the second part of (3.4) via the embedding $F_{2,1}^{1/2} \subset F_{\infty,1}^0$.

Now we turn our attention to the proof of second part of (3.16). We first give a quick proof for the $\dot{W}^{\frac{1}{4},4}$ bound. There we use Lemmas 3.6, 3.7, and the para-associativity lemma, Lemma 2.7, to expand the material derivative,

$$\begin{aligned} T_{D_t} P(\bar{R}_\alpha R) &= P[(\partial_\alpha T_{D_t} \bar{R})R] + P[\bar{R}_\alpha T_{D_t} R] + E \\ &= -iP[(\partial_\alpha T_{1+a} \bar{Y})R] + iP[\bar{R}_\alpha T_{1+a} Y] + E \\ &= iT_{1+a}(-\bar{Y}_\alpha R + \bar{R}_\alpha Y) + E, \end{aligned}$$

with errors E satisfying

$$\|E\|_{\dot{W}^{\frac{1}{4},4}} \lesssim_{A^\#} A^\# A_{\frac{1}{4}}^2.$$

We would like to repeat the same argument for the L^∞ bound, except that we have the issue that P is not bounded in L^∞ . To avoid this problem we use a trick we employed before, which is to take better advantage of the structure of both M and R . Precisely, we can write a in two ways,

$$a = -i(P[\bar{R}_\alpha R] - \bar{P}[R_\alpha \bar{R}]) = -i(\partial_\alpha P[\bar{R}R] - R_\alpha \bar{R}),$$

while M can be represented in a matching fashion as

$$M = P[R\bar{Y}_\alpha - \bar{R}_\alpha Y] + \bar{P}[\bar{R}Y_\alpha - R_\alpha \bar{Y}] = \partial_\alpha P[R\bar{Y} - Y\bar{R}] + \bar{R}Y_\alpha - R_\alpha \bar{Y}.$$

Here, the first representation is good for unbalanced frequencies, as it selects the correct frequency balance. The second, on the other hand, is good for balanced interactions, as it pairs each projection with a differentiation, therefore improving the kernel decay.

Correspondingly, we consider matching paradifferential decompositions of a and M , namely

$$a = a^p + a^\Pi, \quad M = M^p + M^\Pi,$$

where, using the two expressions above, we set

$$a^p = -i(T_{\bar{R}_\alpha} R - T_{R_\alpha} \bar{R}), \quad a^\Pi = -i(\partial_\alpha P \Pi(\bar{R}, R) - \Pi(R_\alpha, \bar{R})),$$

as well as the matching decomposition for M . We will now prove that the desired L^∞ bound holds separately for the pairs a^p, M^p respectively a^Π, M^Π .

(a) *The unbalanced case, a^p, M^p .* We consider the first terms in a^p, M^p , for which we reproduce the computation above:

$$\begin{aligned} T_{D_t} T_{\bar{R}_\alpha} R &= T_{\partial_\alpha T_{D_t} \bar{R}} R + T_{\bar{R}_\alpha} T_{D_t} R + E_1 \\ &= -i T_{\partial_\alpha T_{1+a} \bar{Y}} R + i T_{\bar{R}_\alpha} T_{1+a} Y + E_1 + E_2 \\ &= i T_{1+a} (-T_{\bar{Y}_\alpha} R + T_{\bar{R}_\alpha} Y) + E_1 + E_2 + E_3. \end{aligned}$$

We successively consider the three errors: E_1 is the para-associativity error, which can be represented as a sum of two translation-invariant trilinear forms

$$E_1 = L_{\text{lmh}}(\bar{R}_{\alpha\alpha}, b, R) + L_{\text{lmh}}(b_\alpha, R_\alpha, R),$$

with frequency localizations ordered as listed. Then we can apply the bound (2.8) twice to obtain

$$\|E_1\|_{\text{BMO}^{-\frac{1}{4}}} \lesssim_A A^2 A_{\frac{1}{4}}, \quad \|E_1\|_{\text{BMO}^{\frac{1}{4}}} \lesssim_A A^{\frac{3}{4}}, \tag{3.20}$$

after which the desired L^∞ bound follows by interpolation.

The error E_2 can be written in the form

$$E_2 = T_{\partial_\alpha \bar{\mathcal{K}}} R + T_{\bar{R}_\alpha} \mathcal{K},$$

where \mathcal{K} is from (3.12), which also satisfies a pair of unbalanced BMO bounds in (3.13) and (3.14). Then a single application of the bound (2.8) yields the same bound as for E_1 .

Finally, E_3 is also a para-associativity error, but better than E_1 because it is quartic, and it is treated in the same way.

(b) *The balanced case, a^Π, M^Π .* We follow the same outline as before, considering the first terms in a^Π, M^Π and the corresponding computation:

$$\begin{aligned} T_{D_t} \partial_\alpha P \Pi(\bar{R}, R) &= \partial_\alpha P \Pi(T_{D_t} \bar{R}, R) + \partial_\alpha P \Pi(\bar{R}, T_{D_t} R) + E_1 \\ &= -i \partial_\alpha P \Pi(T_{1+a} \bar{Y}, R) + i \partial_\alpha P \Pi(\bar{R}, T_{1+a} Y) + E_1 + E_2 \\ &= i T_{1+a} (-\partial_\alpha P \Pi(\bar{Y}, R) + \partial_\alpha P \Pi(\bar{R}, Y)) + E_1 + E_2 + E_3. \end{aligned}$$

As above, E_1 is the para-associativity error, which can be represented as a sum of trilinear forms

$$E_1 = L_{\text{lh}}(b_\alpha, \bar{R}_\alpha, R) + L_{\text{lh}}(b_\alpha, \bar{R}, R_\alpha),$$

where the first frequency is smaller and the last two are balanced. Then we obtain the same bound as in (3.20) by applying (2.8) once and (2.5) once.

The error E_2 can be written in the form

$$E_2 = \partial_\alpha P \Pi(\bar{\mathcal{K}}, R) + \partial_\alpha P \Pi(\bar{R}, \mathcal{K}),$$

which can be treated using (3.13), (3.14), and (2.5). Here, more care is needed for the second term in a^Π , which yields an error E_2 of the form

$$E_2 = \Pi(\bar{\mathcal{K}}, R_\alpha) + \Pi(\mathcal{K}, \bar{R}_\alpha).$$

This we can estimate by the Cauchy–Schwarz inequality,

$$\|E_2\|_{L^\infty} \lesssim \|\mathcal{K}\|_{\dot{B}_{\infty,2}^{\frac{3}{8}}} \|R\|_{\dot{B}_{\infty,2}^{-\frac{3}{8}}}.$$

Then we bound the two norms on the right by interpolation between (3.13), (3.14) for the first factor on the right, respectively A and $A_{\frac{1}{4}}$ for the second factor.

Finally, the error E_3 is similar to E_1 but better. ■

4. The linearized equations

The proofs of the main results in this paper take advantage of both the linearization of the original equation (1.1) for (W, Q) and the linearization of the differentiated equation (1.2) for (\mathbf{W}, R) . Both of these linearizations are used either directly or via their associated paradifferential flows. In this section we derive both of these linearized equations, as well as the relation between them. We also write down the corresponding paradifferential flows.

4.1. The linearization of the (W, Q) equation

We denote solutions for the linearized water wave equations (1.1) around a solution (W, Q) by (w, q) . However, it is more convenient to immediately switch to diagonal variables (w, r) , where

$$r := q - R w.$$

The system of linearized equations for (w, r) was computed in [15]. We review here the key points.

The linearization of R is

$$\delta R = \frac{q_\alpha - R w_\alpha}{1 + \mathbf{W}} = \frac{r_\alpha + R_\alpha w}{1 + \mathbf{W}}, \tag{4.1}$$

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,

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

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

$$\begin{cases} w_t + F w_\alpha + (1 + \mathbf{W}) P[m - \bar{m}] = 0, \\ q_t + F q_\alpha + Q_\alpha P[m - \bar{m}] - i w + P[n + \bar{n}] = 0. \end{cases}$$

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

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

Now, we can use the second equation in (1.2) to switch from q to r and obtain

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

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

$$\begin{cases} (\partial_t + b\partial_\alpha)w + \frac{1}{1+\bar{\mathbf{W}}}r_\alpha + \frac{R_\alpha}{1+\bar{\mathbf{W}}}w = \mathcal{G}_0(w, r), \\ (\partial_t + b\partial_\alpha)r - i\frac{1+a}{1+\mathbf{W}}w = \mathcal{K}_0(w, r), \end{cases} \tag{4.2}$$

where

$$\mathcal{G}_0(w, r) = (1 + \mathbf{W})(P\bar{m} + \bar{P}m), \quad \mathcal{K}_0(w, r) = \bar{P}n - P\bar{n}. \tag{4.3}$$

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

$$\begin{cases} \partial_t w + P[b\partial_\alpha w] + P\left[\frac{1}{1+\bar{\mathbf{W}}}r_\alpha\right] + P\left[\frac{R_\alpha}{1+\bar{\mathbf{W}}}w\right] = P\mathcal{G}_0(w, r), \\ \partial_t r + P[b\partial_\alpha r] - iP\left[\frac{1+a}{1+\mathbf{W}}w\right] = P\mathcal{K}_0(w, r). \end{cases} \tag{4.4}$$

Since the original set of equations (1.1) is fully holomorphic, it follows that the two sets of equations, (4.2) and (4.4), are algebraically equivalent. Later in this paper we will work with (4.4).

Based on the above form of the linearized equations, we can also write the corresponding paradifferential flow:

$$\begin{cases} T_{D_t}w + T_{1-\bar{Y}}\partial_\alpha r + T_{(1-\bar{Y})R_\alpha}w = 0, \\ T_{D_t}r - iT_{1-Y}T_{1+a}w = 0. \end{cases} \tag{4.5}$$

Here we use the Weyl quantization for the paraproducts in order to ensure self-adjointness. We remark that the right-hand side terms in (4.4) have no paradifferential component (i.e. components such that the (w, r) factor is the highest frequency).

The $\dot{\mathcal{H}}^{1/4}$ well-posedness of the linearized equations (4.4) will be studied in Section 5. The $\dot{\mathcal{H}}^s$ well-posedness of the paradifferential equations (4.5), studied in Sections 5.2, 5.3, will play a key role.

4.2. The linearization of the (W, R) equation

We denote solutions for the linearized water wave equations (1.2) around a solution (W, R) by (\hat{w}, \hat{r}) . Recalling (1.2),

$$\begin{cases} D_t W + \frac{(1+W)R_\alpha}{1+\bar{W}} = (1+W)M, \\ D_t R = i\left(\frac{W-a}{1+W}\right), \end{cases}$$

we begin by computing the linearizations of b ,

$$\delta b = 2\Re P[(1-\bar{Y})\hat{r} - (1-\bar{Y})^2 R\bar{\hat{w}}],$$

and of M ,

$$\delta M = 2\Re P[\hat{r}\bar{Y}_\alpha - 2R(1-\bar{Y})\bar{\hat{w}}\bar{Y}_\alpha + R(1-\bar{Y})^2\bar{\hat{w}}_\alpha - \bar{\hat{r}}_\alpha Y - \bar{R}_\alpha(1-Y)^2\bar{\hat{w}}].$$

Then we can write the linearized equation in the form

$$\begin{cases} D_t \hat{w} + (1-\bar{Y})(1+W_\alpha)\hat{r}_\alpha + (1-\bar{Y})R_\alpha \hat{w} \\ \quad = M\hat{w} + (1+W_\alpha)\delta M - W_{\alpha\alpha}\delta b + (1-\bar{Y})^2(1+W_\alpha)R_\alpha \bar{\hat{w}}, \\ D_t \hat{r} - i(1+a)(1-Y)^2\hat{w} \\ \quad = -R_\alpha\delta b - (1-Y)[\bar{P}(\hat{r}R_\alpha) + \bar{P}(\bar{R}\hat{r}_\alpha) - P(\hat{r}\bar{R}_\alpha) - P(R\bar{\hat{r}}_\alpha)]. \end{cases} \tag{4.6}$$

Next we identify the corresponding paradifferential equation, which contains only paraproduct-type contributions where \hat{w} and \hat{r} terms are the highest frequency. Thus we can exclude all $(\bar{\hat{w}}, \bar{\hat{r}})$ terms, as well as all terms where \hat{w} and \hat{r} are inside a \bar{P} projection. Using the relation (1.3) to group terms, we are left with

$$\begin{cases} T_{D_t}\hat{w} + T_{b_\alpha}\hat{w} + T_{1-\bar{Y}}T_{1+W_\alpha}\hat{r}_\alpha - T_{\bar{Y}_\alpha}T_{1+W_\alpha}\hat{r} + T_{1-\bar{Y}}T_{W_{\alpha\alpha}}\hat{r} = 0, \\ T_{D_t}\hat{r} + T_{b_\alpha}\hat{r} - iT_{(1-Y)^2}T_{1+a}\hat{w} + T_M\hat{r} = 0. \end{cases} \tag{4.7}$$

This linearized equation will play a key role in the study of the differentiated nonlinear flow (1.2).

4.3. The relation between the two linearized flows

To each solution (w, r) to the linearized equation (4.4) corresponds a solution (\hat{w}, \hat{r}) to the linearized differentiated equation (4.6). In view of (4.1), the connection between the two is given by the relation

$$(\hat{w}, \hat{r}) = \left(w_\alpha, \frac{r_\alpha + R_\alpha w}{1+W}\right). \tag{4.8}$$

In Section 6 we will use a paradifferential version of this relation in order to transfer the $\dot{\mathcal{H}}^s$ well-posedness results from the paradifferential equation (4.5) to (4.7).

5. Bounds for the linearized equation

A key result of [15] asserts that the linearized equation (4.4) is well posed in \mathcal{H}^0 for as long as A and $\int B$ remain finite.

Theorem 6 ([15]). *The linearized equation (4.4) is well posed in \mathcal{H}^0 . Furthermore, there exists an energy functional $E_{\text{lin}}(w, r)$ so that we have*

(a) *norm equivalence:*

$$E_{\text{lin}}(w, r) \approx_A \|(w, r)\|_{\dot{\mathcal{H}}^0}^2,$$

(b) *energy estimates:*

$$\frac{d}{dt} E_{\text{lin}}(w, r) \lesssim_A AB \|(w, r)\|_{\dot{\mathcal{H}}^0}^2.$$

We remark that this bound is scale invariant, and thus no improvement is possible where one lowers the regularity measured by either factor A or factor B . Instead, our goal in this section is to improve this result by replacing the product AB by a more balanced version, namely $A_{\frac{1}{4}}^2$. We remark that $A_{\frac{1}{4}}^2 \lesssim AB$, but the two terms scale in the same way. The price to pay will be a shift in our function spaces, from \mathcal{H}^0 to $\dot{\mathcal{H}}^{\frac{1}{4}}$.

Theorem 7. *Assume $A \lesssim 1$ and $A_{\frac{1}{4}} \in L^2$. Then the linearized equation (4.4) is well posed in $\dot{\mathcal{H}}^{\frac{1}{4}}$. Furthermore, there exists an energy functional $E_{\text{lin}}^{\frac{1}{4}}(w, r)$ so that we have*

(a) *norm equivalence:*

$$E_{\text{lin}}^{\frac{1}{4}}(w, r) \approx_{A^\#} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}}^2,$$

(b) *energy estimates:*

$$\frac{d}{dt} E_{\text{lin}}^{\frac{1}{4}}(w, r) \lesssim_{A^\#} A_{\frac{1}{4}}^2 \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}}^2.$$

In the remainder of the section, we prove Theorem 7.

5.1. Reduction to the paradifferential equation

A key step in the proof of Theorem 7 is to reduce to the linearized paradifferential approximation (4.5) of the full linearized evolution (4.2). Recall that this has the form

$$\begin{cases} T_{D_t} w + T_{1-\bar{Y}} \partial_\alpha r + T_{(1-\bar{Y})R_\alpha} w = 0, \\ T_{D_t} r - iT_{1-Y} T_{1+a} w = 0. \end{cases}$$

Recall that throughout, we fix a self-adjoint quantization for T . For instance, we may use the Weyl quantization, or simply the average

$$\frac{1}{2}(T + T^*).$$

The first step in the proof of Theorem 7 is to show that the conclusion of the theorem holds for the evolution (4.5).

Proposition 5.1. *Assume $A \lesssim 1$ and $A_{\frac{1}{4}} \in L^2$. Then the linearized paradifferential equation (4.5) is well posed in \mathcal{H}^s for every $s \in \mathbb{R}$. Furthermore, there exist energy functionals $E_{\text{lin}}^{s,\text{para}}(w, r)$ so that we have*

(a) *norm equivalence:*

$$E_{\text{lin}}^{s,\text{para}}(w, r) \approx_{A^\#} \|(w, r)\|_{\mathcal{H}^s}^2,$$

(b) *energy estimates:*

$$\frac{d}{dt} E_{\text{lin}}^{s,\text{para}}(w, r) \lesssim_{A^\#} A_{\frac{1}{4}}^2 \|(w, r)\|_{\mathcal{H}^s}^2.$$

Here, the regularity index s plays no role. We will prove this result first in the easier special case $s = 0$ in Section 5.2, and then for all s in Section 5.3.

We remark that this result also allows us to treat linear perturbations of (4.5) of the form

$$\begin{cases} T_{D_t} w + T_{1-\bar{Y}} \partial_\alpha r + T_{(1-\bar{Y})R_\alpha} w = G, \\ T_{D_t} r - iT_{1-Y} T_{1+a} w = K, \end{cases} \tag{5.1}$$

provided that the source terms satisfy the bounds

$$\|(G, K)\|_{\dot{\mathcal{H}}^s} \lesssim_A A_{1/4}^2 \|(w, r)\|_{\dot{\mathcal{H}}^s}, \tag{5.2}$$

for some s . Then, by the variation of parameters formula, the conclusion of the last proposition also applies to the equations (5.1) for the same s .

Once we have dealt with the paradifferential equations, our next objective is to reduce the analysis of the full linearized equations to the paradifferential case. Toward this goal, we follow the clue above and rewrite the equations (4.2) as a paradifferential evolution with a source term,

$$\begin{cases} T_{D_t} w + T_{1-\bar{Y}} \partial_\alpha r + T_{(1-\bar{Y})R_\alpha} w = \mathcal{G}^\#(w, r), \\ T_{D_t} r - iT_{1-Y} T_{1+a} w = \mathcal{K}^\#(w, r), \end{cases} \tag{5.3}$$

where the source terms $(\mathcal{G}^\#, \mathcal{K}^\#)$ have the form

$$\mathcal{G}^\# = P(\mathcal{G}_0 + \mathcal{G}_1), \quad \mathcal{K}^\# = P(\mathcal{K}_0 + \mathcal{K}_1)$$

with $(\mathcal{G}_0, \mathcal{K}_0)$ as in (4.3) and

$$\begin{aligned} \mathcal{G}_1 &= -(T_{w_\alpha} b + \Pi(w_\alpha, b)) + (T_{r_\alpha} \bar{Y} + \Pi(r_\alpha, \bar{Y})) \\ &\quad - (T_w(1-\bar{Y})R_\alpha + \Pi(w, (1-\bar{Y})R_\alpha)), \\ \mathcal{K}_1 &= -(T_{r_\alpha} b + \Pi(r_\alpha, b)) \\ &\quad + i(T_{1-Y} T_w a + T_{1-Y} \Pi(w, a) - T_{(1+a)w} Y - \Pi((1+a)w, Y)) \end{aligned}$$

denote the paradifferential truncations. Henceforth, it will be convenient to include an implicit projection P in Π , so that $\Pi = P\Pi$.

Unfortunately, our source terms $(\mathcal{G}^\#, \mathcal{K}^\#)$ do not satisfy the cubic bounds (5.2) for any s , both because they contain quadratic contributions and because of unbalanced cubic contributions. However, this still allows us to perturbatively discard portions of $(\mathcal{G}^\#, \mathcal{K}^\#)$ which do satisfy (5.2).

To deal with the unfavorable part of $(\mathcal{G}^\#, \mathcal{K}^\#)$ we will rely on a more accurate paradifferential version of normal form analysis. We also optimize the choice of s . Our result reads as follows.

Proposition 5.2. *Given (w, r) satisfying (5.3), there exist modified normal form linear variables $(w_{\text{NF}}, r_{\text{NF}})$ satisfying (5.1) such that we have*

(i) *invertibility:*

$$\|(w_{\text{NF}}, r_{\text{NF}}) - (w, r)\|_{\dot{H}^{\frac{1}{4}}} \lesssim_A A^\# \|(w, r)\|_{\dot{H}^{\frac{1}{4}}},$$

(ii) *perturbative source term:*

$$\|(G, K)\|_{\dot{H}^{\frac{1}{4}}} \lesssim_{A^\#} A^{\frac{2}{4}} \|(w, r)\|_{\dot{H}^{\frac{1}{4}}}. \tag{5.4}$$

To facilitate the normal form analysis in the proof of the above proposition, we separately state a pair of weaker quadratic bounds for $(\mathcal{E}_0, \mathcal{K}_0)$ and most of $(\mathcal{E}_1, \mathcal{K}_1)$, which will suffice in order to evaluate their impact as source terms for the material derivative of (w, r) . In the case of $(\mathcal{E}_0, \mathcal{K}_0)$ it will be useful to simultaneously extract the perturbative part of $(\mathcal{E}_0, \mathcal{K}_0)$, in preparation for the normal form analysis on the unfavorable part.

It will be convenient to define the unfavorable part in terms of the following variant of w with a low frequency coefficient,

$$x = T_{1-Y}w.$$

Lemma 5.3. *The source terms $(\mathcal{E}_0, \mathcal{K}_0)$ satisfy the bound*

$$\|(\mathcal{E}_0, \mathcal{K}_0)\|_{\mathcal{H}^0} \lesssim_{A^\#} A^{\frac{1}{4}} \|(w, r)\|_{\dot{H}^{\frac{1}{4}}}. \tag{5.5}$$

Furthermore, they have the representation

$$P\mathcal{E}_0 = \mathcal{E}_{0,0} + G, \quad P\mathcal{K}_0 = \mathcal{K}_{0,0} + K,$$

where (G, K) satisfy the bounds (5.4), whereas $(\mathcal{E}_{0,0}, \mathcal{K}_{0,0})$ are given by

$$\begin{aligned} \mathcal{E}_{0,0}(w, r) &= -P[T_{1+W_\alpha}(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{X}}\bar{R}_\alpha)Y] + P[T_{1+W_\alpha}T_{1-\bar{Y}}\bar{x}_\alpha R], \\ \mathcal{K}_{0,0}(w, r) &= -P[(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{X}}\bar{R}_\alpha)R]. \end{aligned}$$

We note that (G, K) also satisfy (5.5), but we will not prove it as this is not needed in the sequel.

We now turn our attention to $(\mathcal{E}_1, \mathcal{K}_1)$, where we have the following lemma.

Lemma 5.4. *The source terms $(\mathcal{G}_1, \mathcal{K}_1)$ have the representation*

$$\begin{aligned} P \mathcal{G}_1 &= -T_w T_{1-\bar{Y}} R_\alpha + G, \\ P \mathcal{K}_1 &= -iT_w T_{1+a} Y + K, \end{aligned}$$

where (G, K) satisfy the quadratic bound (5.5).

Using these lemmas, we will obtain the following formula for the para-material derivatives of (w, r) :

$$\begin{cases} T_{D_t} w = -T_{1-\bar{Y}}(r_\alpha + T_w R_\alpha) + G, \\ T_{D_t} r = iT_{1+a}(T_{1-Y} w - T_w Y) + K, \end{cases} \tag{5.6}$$

where (G, K) satisfy the quadratic bound (5.5). We will also obtain a similar formula for x .

We will first prove Lemmas 5.3, 5.4, and the para-material derivative formula (5.6) in Section 5.4. Next, Proposition 5.2 is proved in two steps in Sections 5.5 and 5.6, where we construct normal form corrections for the terms $(\mathcal{G}_1, \mathcal{K}_1)$ and $(\mathcal{G}_0, \mathcal{K}_0)$ respectively.

5.2. \mathcal{H}^0 bounds for the paradifferential equation

In this section we consider the \mathcal{H}^0 well-posedness of (4.5), which we reproduce here for convenience with general source term (G, K) :

$$\begin{cases} T_{D_t} w + T_{1-\bar{Y}} r_\alpha + T_{(1-\bar{Y})R_\alpha} w = G, \\ T_{D_t} r - iT_{1-Y} T_{1+a} w = K. \end{cases} \tag{5.7}$$

In the study of (4.4) in [15], the following associated positive definite linear energy was used:

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

Here we slightly alter this by considering its paradifferential version,

$$E_{\text{lin}}^{0,\text{para}}(w, r) = \int_{\mathbb{R}} T_{1+a} w \cdot \bar{w} + \Im(r\bar{r}_\alpha) d\alpha.$$

With this notation, we will prove the following bounds.

Proposition 5.5. *The linear equation (5.7) is well posed in \mathcal{H}^0 , and the following estimate holds:*

$$\frac{d}{dt} E_{\text{lin}}^{0,\text{para}}(w, r) = 2\Re \int_{\mathbb{R}} T_{1+a} \bar{w} G - i\bar{r}_\alpha K d\alpha + O_{A^\#}(A_4^2) E_{\text{lin}}^{(2)}(w, r).$$

Proof. A direct computation yields

$$\begin{aligned} \frac{d}{dt} \int T_{1+a} \bar{w} \cdot w + \Im(r\bar{r}_\alpha) d\alpha &= 2\Re \int T_{1+a} \bar{w} \cdot w_t d\alpha + 2\Im \int \bar{r}_\alpha r_t d\alpha \\ &\quad + \int T_{a_t} \bar{w} \cdot w d\alpha. \end{aligned}$$

Next we augment the time derivatives w_t and r_t on the right-hand side into material derivatives. Integrating by parts, write

$$\begin{aligned} 2\Re \int T_{1+a} \bar{w} \cdot T_b \partial_\alpha w \, d\alpha &= - \int T_{((1+a)b)_\alpha} \bar{w} \cdot w \, d\alpha \\ &\quad - \int (T_{b_\alpha} T_a + T_b T_{a_\alpha} - T_{(ab)_\alpha}) \bar{w} \cdot w \, d\alpha \\ &\quad - \int (T_b T_a - T_a T_b) \bar{w}_\alpha \cdot w \, d\alpha, \end{aligned}$$

and similarly,

$$2\Im \int \bar{r}_\alpha \cdot T_b \partial_\alpha r \, d\alpha = 0.$$

Adding the above three identities and using the equations (5.7), the quadratic $\Re(w \bar{r}_\alpha)$ term cancels:

$$\begin{aligned} \frac{d}{dt} E_{\text{lin}}^{0,\text{para}}(w, r) &= 2\Re \int (1 + T_a) \bar{w} \cdot G - i \bar{r}_\alpha K \, d\alpha \\ &\quad - 2\Re \int (1 + T_a) \bar{w} \cdot T_{\frac{R_\alpha}{1+W}} w \, d\alpha \\ &\quad + \int (T_{b_\alpha} T_a + T_b T_{a_\alpha} - T_{(ab)_\alpha}) \bar{w} \cdot w \, d\alpha \\ &\quad + \int (T_b T_a - T_a T_b) \bar{w}_\alpha \cdot w \, d\alpha \\ &\quad + \int T_{a_t + b a_\alpha + (1+a) b_\alpha} \bar{w} \cdot w \, d\alpha. \end{aligned}$$

We can write two of the error terms (precisely, the second and fifth lines on the right-hand side) in terms of the auxiliary function M defined in (1.3), modulo another paraproduct error term:

$$\frac{d}{dt} E_{\text{lin}}^{0,\text{para}}(w, r) = 2\Re \int (1 + T_a) \bar{w} \cdot G - i \bar{r}_\alpha K \, d\alpha + \text{Err}_M + \text{Err}_{\text{com}},$$

where

$$\begin{aligned} \text{Err}_M &= \int T_{D_t a} \bar{w} \cdot w - (1 + T_a) \bar{w} \cdot T_M w \, d\alpha, \\ \text{Err}_{\text{com}} &= \int (T_{b_\alpha} T_a + T_b T_{a_\alpha} - T_{(ab)_\alpha}) \bar{w} \cdot w \, d\alpha + \int (T_b T_a - T_a T_b) \bar{w}_\alpha \cdot w \, d\alpha \\ &\quad + \int (T_a b_\alpha - T_a T_{b_\alpha}) \bar{w} \cdot w \, d\alpha. \end{aligned}$$

For the first term in Err_M , by Lemma 3.9, we have

$$\|D_t a\|_{L^\infty} \lesssim_{A^\#} A_{\frac{1}{4}}^2.$$

For the second term we combine the pointwise bounds $|a| \lesssim A^2$ in Proposition 3.3 together with

$$\|M\|_{L^\infty} \lesssim_A A_{\frac{1}{4}}^2$$

in Lemma 3.5.

All terms in Err_{com} may be estimated using the product and commutator estimates, Lemmas 2.4 and 2.5. ■

5.3. $\dot{\mathcal{H}}^s$ bounds for the paradifferential equation

Our aim here is to prove $\dot{\mathcal{H}}^s$ well-posedness for the linearized paradifferential equation (4.5), and to identify a suitable energy functional. Our main result is as follows.

Proposition 5.6. *Let $s \in \mathbb{R}$. Given (w, r) solving the homogeneous linearized equation (4.5), there exist linearized, normalized variables $(\tilde{w}^s, \tilde{r}^s)$ solving*

$$\begin{cases} T_{D_t} \tilde{w}^s + T_{1-\bar{Y}} \partial_\alpha \tilde{r}^s + T_{(1-\bar{Y})R_\alpha} \tilde{w}^s = G_s, \\ T_{D_t} \tilde{r}^s - iT_{1-\bar{Y}} T_{1+a} \tilde{w}^s = K_s, \end{cases}$$

and such that

$$\|(\tilde{w}^s, \tilde{r}^s) - D^s(w, r)\|_{\mathcal{H}^0} \lesssim_A A \| (w, r) \|_{\dot{\mathcal{H}}^s}$$

and

$$\|(G_s, K_s)\|_{\mathcal{H}^0} \lesssim_A A_{\frac{1}{4}}^2 \| (w, r) \|_{\dot{\mathcal{H}}^s}. \tag{5.8}$$

The $\dot{\mathcal{H}}^s$ well-posedness for (5.7) follows by a fixed point argument by applying Proposition 5.5 to $(\tilde{w}^s, \tilde{r}^s)$ given by the above proposition. The $\dot{\mathcal{H}}^s$ energy in Proposition 5.1 will then be given by

$$E_{\text{lin}}^{s,\text{para}}(w, r) = E_{\text{lin}}^{0,\text{para}}(\tilde{w}^s, \tilde{r}^s).$$

Proof of Proposition 5.6. The natural approach to this problem is to consider the equations for the variables

$$(w^s, r^s) = (D^s w, D^s r).$$

This can be written in the form

$$\begin{cases} T_{D_t} w^s + T_{1-\bar{Y}} \partial_\alpha r^s + T_{(1-\bar{Y})R_\alpha} w^s = \mathcal{G}_0^s(w^s, r^s), \\ T_{D_t} r^s - iT_{1-\bar{Y}} T_{1+a} w^s = \mathcal{K}_0^s(w^s, r^s), \end{cases} \tag{5.9}$$

where the source terms $(\mathcal{G}_0^s, \mathcal{K}_0^s)$ have the form

$$\begin{aligned} \mathcal{G}_0^s(w^s, r^s) &= L(b_\alpha, w^s) - L(\bar{Y}_\alpha, r^s) + L([(1-\bar{Y})R_\alpha]_\alpha, \partial_\alpha^{-1} w^s), \\ \mathcal{K}_0^s(w^s, r^s) &= L(b_\alpha, r^s) + iL(Y_\alpha, \partial_\alpha^{-1} T_{1+a} w^s). \end{aligned} \tag{5.10}$$

Here L denotes the order-zero paradifferential bilinear form

$$L(f_\alpha, u) = -[D^s, T_f] \partial_\alpha D^{-s} u$$

which has leading term

$$L(g, u) = -sT_g u.$$

We need to study the evolution (5.9) in \mathcal{H}^0 . The difficulty is that the source terms $(\mathcal{E}_0^s, \mathcal{K}_0^s)$ cannot be treated perturbatively, both because they have a quadratic component and because their cubic and higher parts are unbalanced.²

To facilitate the next discussion we note that we can organize the source terms in $(\mathcal{E}, \mathcal{K})$ by order, i.e. by the relative Sobolev regularity of the leading factors (w, r) . For instance, in (5.10) the first terms on the right are leading order 0, the second are order $-\frac{1}{2}$, and the third are order -1 . The lower the order the better; this is most readily seen in the normal form analysis.

The natural attempt would be to try to apply a normal form transformation in order to eliminate all the bad terms, and replace them by perturbative terms. Such a strategy would work for terms of order $\leq -\frac{1}{2}$. Precisely, all nonperturbative terms of order $\leq -\frac{1}{2}$ can be viewed as quadratic expressions, possibly with a lower frequency coefficient. Then the normal form computation works directly, and applies to the quadratic part, while simply retaining the low frequency coefficient.

On the other hand, for the zero-order terms the above procedure no longer works. A normal form would indeed replace the quadratic terms by cubic ones, but those will still be partly unbalanced. One could reiterate and eventually obtain an infinite series of corrections. Instead of this, we will borrow an idea from [15], where it was observed that there is an appropriate exponential conjugation (which can be thought of as the sum of the infinite series of corrections alluded to above) which does the job modulo “lower-order terms”, which can be dealt with directly by normal forms.

The conjugation introduced in [15], only for integer s , was purely algebraic, by a factor

$$\Phi = \phi(W_\alpha) = J^{-s}.$$

Here we will also allow for a noninteger s , and use the conjugation in a paradifferential fashion. We denote the conjugated variables by

$$(\tilde{w}_1^s, \tilde{r}_1^s) = (T_\Phi w^s, T_\Phi r^s)$$

and compute the equations for $(\tilde{w}_1^s, \tilde{r}_1^s)$. Along the way we will cheerfully discard good cubic terms which can be placed in (G_s, K_s) . By good terms, we mean terms which satisfy the good bounds (5.8). From here on, we will simply denote by (G_s, K_s) any such expressions.

We first use the para-Leibniz rule in Lemma 3.6 to write

$$T_{D_t} \tilde{w}_1^s = T_\Phi T_{D_t} w^s + T_{D_t, \Phi} w^s + G_s,$$

²This would be less of an issue if all we wanted was an AB bound.

where we have used the bound

$$\|D^{\frac{1}{4}}\Phi\|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}.$$

We compute

$$\begin{aligned} D_t\Phi &= \partial_{W_\alpha}\phi \cdot D_tW_\alpha + \partial_{\bar{W}_\alpha}\phi \cdot D_t\bar{W}_\alpha \\ &= s\Phi[(1-\bar{Y})R_\alpha + (1-Y)\bar{R}_\alpha - 2M] = s\Phi(b_\alpha - M). \end{aligned}$$

We may discard the M term by Lemma 3.5, arriving at

$$T_{D_t}\tilde{w}_1^s = T_\Phi T_{D_t}w^s + sT_{\Phi b_\alpha}w^s + G_s.$$

Finally, applying Lemma 2.5 for the Φb_α paraproduct we conclude that

$$T_{D_t}\tilde{w}_1^s = T_\Phi T_{D_t}w^s + sT_{b_\alpha}\tilde{w}_1^s + G_s.$$

Similarly we have

$$T_{D_t}\tilde{r}_1^s = T_\Phi T_{D_t}r^s + sT_{b_\alpha}\tilde{r}_1^s + K_s.$$

For the other terms on the left-hand side of (5.9), again using Lemmas 2.4, 2.5, we have

$$\begin{aligned} T_{1-\bar{Y}}\partial_\alpha\tilde{r}_1^s &= T_\Phi T_{1-\bar{Y}}\partial_\alpha r^s - sT_{1-\bar{Y}}T_{(1-Y)W_{\alpha\alpha}+(1-\bar{Y})\bar{W}_{\alpha\alpha}}\tilde{r}_1^s + G_s \\ &= T_\Phi T_{1-\bar{Y}}\partial_\alpha r^s - s(T_{1-\bar{Y}}T_{(1+W_\alpha)Y_\alpha} + T_{\bar{Y}_\alpha})\tilde{r}_1^s + G_s, \end{aligned}$$

and

$$T_{(1-\bar{Y})R_\alpha}\tilde{w}_1^s = T_\Phi T_{(1-\bar{Y})R_\alpha}w^s + G_s$$

in the first equation, and

$$T_{1-Y}T_{1+a}\tilde{w}_1^s = T_\Phi T_{1-Y}T_{1+a}w^s + K_s$$

in the second. Finally, by Lemma 2.4, T_Φ commutes with the coefficients on the right-hand side of (5.9) modulo good errors, so for $(\tilde{w}_1^s, \tilde{r}_1^s)$ we obtain the system

$$\begin{cases} T_{D_t}\tilde{w}_1^s + T_{1-\bar{Y}}\partial_\alpha\tilde{r}_1^s + T_{(1-\bar{Y})R_\alpha}\tilde{w}_1^s = \mathcal{G}_1^s + G_s, \\ T_{D_t}\tilde{r}_1^s - iT_{1-Y}T_{1+a}\tilde{w}_1^s = \mathcal{K}_1^s + K_s, \end{cases} \quad (5.11)$$

with the nonperturbative source terms

$$\begin{aligned} \mathcal{G}_1^s &= L(b_\alpha, \tilde{w}_1^s) + sT_{b_\alpha}\tilde{w}_1^s - L(\bar{Y}_\alpha, \tilde{r}_1^s) - sT_{\bar{Y}_\alpha}\tilde{r}_1^s + L([(1-\bar{Y})R_\alpha]_\alpha, \partial_\alpha^{-1}\tilde{w}_1^s) \\ &\quad - sT_{1-\bar{Y}}T_{(1+W_\alpha)Y_\alpha}\tilde{r}_1^s, \\ \mathcal{K}_1^s &= L(b_\alpha, \tilde{r}_1^s) + sT_{b_\alpha}\tilde{r}_1^s + iL(Y_\alpha, \partial_\alpha^{-1}T_{1+a}\tilde{w}_1^s). \end{aligned}$$

The main point here is that there is cancellation in the first two leading terms in both G_1 and K_1 , leaving us with only terms of order $-\frac{1}{2}$ or lower. Precisely, we may write

$$\begin{aligned} \mathcal{G}_1^s &= L_1(b_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{w}_1^s) - L_1(\bar{Y}_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) + L([(1 - \bar{Y})R_\alpha]_\alpha, \partial_\alpha^{-1} \tilde{w}_1^s) \\ &\quad - sT_{1-\bar{Y}}T_{(1+W_\alpha)Y_\alpha} \tilde{r}_1^s, \\ \mathcal{K}_1^s &= L_1(b_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) + iL(Y_\alpha, \partial_\alpha^{-1} T_{1+a} \tilde{w}_1^s), \end{aligned}$$

where L_1 denotes

$$L_1(f_{\alpha\alpha}, \partial_\alpha^{-1} u) = L(f_\alpha, u) + sT_{f_\alpha} u$$

which, like L , has leading term

$$L_1(g, u) = -s(s - 1)T_g u.$$

We also observe that $(\mathcal{G}_1^s, \mathcal{K}_1^s)$ satisfy bilinear estimates of the form³

$$\| |D|^{-1/4} (\mathcal{G}_1^s, \mathcal{K}_1^s) \|_{\mathcal{X}^0} \lesssim_A A_{\frac{1}{4}} \| (w, r) \|_{\mathcal{Y}^s} \tag{5.12}$$

which will suffice to control the secondary impact of the source terms in the normal form analysis below.

We proceed with the normal form corrections. We recall that their purpose is twofold:

- (i) to turn bilinear source terms into trilinear ones, and
- (ii) to replace trilinear and higher unbalanced interactions with balanced ones.

To achieve the first purpose it suffices to consider classical normal forms for bilinear interactions. However, for the second goal more care is needed.

Precisely, we first identify the leading-order bilinear interaction, and think of the remaining factors (if any) as low frequency paradifferential coefficients. For the leading-order part we compute its associated bilinear normal form correction. But after this we follow up with a second step, which is to add low frequency coefficients in order to fully match the low frequency coefficients in the source terms.

One favorable feature in this process is that the secondary outcome of the source terms $(\mathcal{G}_1^s, \mathcal{K}_1^s)$ (as they appear in the material derivative of the normal form corrections) is perturbative. This allows us to divide and conquer, i.e. to compute the normal form corrections separately for different parts of the source terms.

We begin with a correction consisting of L_1 bilinear forms, which will have the twofold effect on $(\mathcal{G}_1^s, \mathcal{K}_1^s)$ of cancelling the L_1 terms, and exchanging the L terms with their leading terms. This is easily computed to have the form

$$\begin{cases} \tilde{w}_2^s &= \partial_\alpha T_{1-Y} L_1(W_{\alpha\alpha} \partial_\alpha^{-2} \tilde{w}_1^s) + T_{1-\bar{Y}} L_1(\bar{W}_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{w}_1^s) =: \tilde{w}_{2,1}^s + \tilde{w}_{2,2}^s, \\ \tilde{r}_2^s &= T_{1-Y} L_1(W_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) + T_{1-\bar{Y}} L_1(\bar{W}_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) =: \tilde{r}_{2,1}^s + \tilde{r}_{2,2}^s, \end{cases}$$

³Similar bounds are also satisfied by (G_s, K_s) , but will not be needed in the sequel because we have (5.8).

where the L_1 part is the leading bilinear expression and the preceding paraproducts should be thought of as coefficients. We claim that this correction has the following effect when inserted in (5.9):

$$\begin{cases} T_{D_t} \tilde{w}_2^s + T_{1-\bar{Y}} \partial_\alpha \tilde{r}_2^s + T_{(1-\bar{Y})R_\alpha} \tilde{w}_2^s = \mathcal{G}_2^s - \mathcal{G}_1^s + G_s, \\ T_{D_t} \tilde{r}_2^s - iT_{1-Y} T_{1+a} \tilde{w}_2^s = \mathcal{K}_2^s - \mathcal{K}_1^s + K_s, \end{cases} \quad (5.13)$$

where

$$\begin{aligned} \mathcal{G}_2^s &= -sT_{1-\bar{Y}} T_{(1+W_\alpha)Y_\alpha} \tilde{r}_1^s - sT_{(1-\bar{Y})R_{\alpha\alpha}} \partial_\alpha^{-1} \tilde{w}_1^s, \\ \mathcal{K}_2^s &= -isT_{Y_\alpha} \partial_\alpha^{-1} T_{1+a} \tilde{w}_1^s. \end{aligned}$$

To prove this we begin with the first equation of the system, and only the contributions of the bilinear holomorphic correction $\tilde{w}_{2,1}^s$. When applying T_{D_t} , we commute it with differentiation while using Lemma 3.6 to distribute it to (para)products. The commutators with ∂_α involve T_{b_α} , and their contributions can all be placed in G_s using (3.5). Then we have

$$\begin{aligned} T_{D_t} \tilde{w}_{2,1}^s &= -\partial_\alpha T_{D_t} Y L_1(W_{\alpha\alpha}, \partial_\alpha^{-2} \tilde{w}_1^s) + \partial_\alpha T_{1-Y} L_1(\partial_\alpha T_{D_t} W_\alpha, \partial_\alpha^{-2} \tilde{w}_1^s) \\ &\quad + \partial_\alpha T_{1-Y} L_1(W_{\alpha\alpha}, \partial_\alpha^{-2} T_{D_t} \tilde{w}_1^s) + G_s. \end{aligned}$$

The first term on the right is perturbative in view of (3.15). For $T_{D_t} W_\alpha$ we use equation (3.12) along with the source term bound (3.13), whereas for the source term arising from $T_{D_t} \tilde{w}_1^s$ we use (5.11) with the source term bounds (5.12) and (5.8). These bounds allow us to estimate the corresponding L_1 contributions by taking advantage of the fact that L_1 has a paraproduct structure. Then we arrive at

$$\begin{aligned} T_{D_t} \tilde{w}_{2,1}^s &= -\partial_\alpha T_{1-Y} L_1(T_{1+W_\alpha} T_{1-\bar{Y}} R_{\alpha\alpha}, \partial_\alpha^{-2} \tilde{w}_1^s) \\ &\quad - \partial_\alpha T_{1-Y} L_1(W_{\alpha\alpha}, \partial_\alpha^{-2} T_{1-\bar{Y}} \partial_\alpha \tilde{r}_1^s) + G_s \\ &= -L_1(T_{1-\bar{Y}} R_{\alpha\alpha\alpha}, \partial_\alpha^{-2} \tilde{w}_1^s) - L_1(T_{1-\bar{Y}} R_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{w}_1^s) \\ &\quad - \partial_\alpha T_{1-Y} L_1(W_{\alpha\alpha}, T_{1-\bar{Y}} \partial_\alpha^{-1} \tilde{r}_1^s) + G_s. \end{aligned}$$

For the remaining holomorphic terms on the left-hand side of the first equation in (5.11) we similarly have

$$\begin{cases} T_{1-\bar{Y}} \partial_\alpha \tilde{r}_{2,1}^s = \partial_\alpha T_{1-Y} L_1(W_{\alpha\alpha}, T_{1-\bar{Y}} \partial_\alpha^{-1} \tilde{r}_1^s) + G_s, \\ T_{(1-\bar{Y})R_\alpha} \tilde{w}_{2,1}^s = G_s. \end{cases}$$

Throughout the above computation we have made repeated use of two facts in the course of collecting error terms into G_s . First, we repeatedly apply the product and commutator estimates, Lemmas 2.4 and 2.5. Second, we observe that when derivatives fall on the coefficients, we have a good balance of derivatives. More generally, any cubic paradifferential terms where the lowest frequency variable is differentiated to an order higher (in fact

strictly higher due to integer mismatch) than $A_{\frac{1}{4}}$ may be absorbed into G_s (this property is combined with the low order $\leq -\frac{1}{2}$ of the high frequency variable to pigeonhole sufficient derivatives onto the middle frequency variable; we then obtain two contributions of $A_{\frac{1}{4}}$ from the low and middle frequencies, possibly after redistributing derivatives from the lowest frequency variable upward).

We similarly consider the contribution of the bilinear corrections with one antiholomorphic input (staying with the first equation of the system):

$$\begin{aligned} T_{D_t} \tilde{w}_{2,2}^s &= -L_1(T_{1-Y} \bar{R}_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{w}_1^s) - L_1(\bar{Y}_\alpha, \tilde{r}_1^s) + G_s, \\ T_{1-\bar{Y}} \partial_\alpha \tilde{r}_{2,2}^s &= L_1(\bar{Y}_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) + L_1(\bar{Y}_\alpha, \tilde{r}_1^s) + G_s, \\ T_{(1-\bar{Y})R_\alpha} \tilde{w}_{2,2}^s &= G_s. \end{aligned}$$

Here we used the same two facts as above in collecting G_s . In addition, we also used Lemma 3.2 to exchange instances of W_α with Y .

After collecting these contributions, observing cancellations, and extracting perturbative paradifferential terms from \mathcal{G}_1^s into G_s using the two facts discussed above, we see that all L_1 terms of \mathcal{G}_1^s cancel (note that we need to expand b into its real and imaginary parts for both an R and \bar{R} component) and all L terms are exchanged with their leading-order terms, thus proving our claim for the first equation in (5.13).

We repeat this computation for the second equation of the system (5.13) (here it is convenient to compute the entire correction at once):

$$\begin{aligned} T_{D_t} \tilde{r}_2^s &= L_1(2\Re(T_{1-Y} \partial_\alpha T_{D_t} W_\alpha), \partial_\alpha^{-1} \tilde{r}_1^s) + L_1(2\Re(T_{1-Y} W_{\alpha\alpha}), \partial_\alpha^{-1} T_{D_t} \tilde{r}_1^s) + K_s \\ &= -L_1(2\Re(T_{1-Y} T_{1+W_\alpha} T_{1-\bar{Y}} R_{\alpha\alpha}), \partial_\alpha^{-1} \tilde{r}_1^s) \\ &\quad + L_1(2\Re(T_{1-Y} W_{\alpha\alpha}), \partial_\alpha^{-1} i T_{1-Y} T_{1+a} \tilde{w}_1^s) + K_s \\ &= -L_1(b_{\alpha\alpha}, \partial_\alpha^{-1} \tilde{r}_1^s) + i T_{1-Y} L_1(2\Re(T_{1-Y} W_{\alpha\alpha}), \partial_\alpha^{-1} T_{1+a} \tilde{w}_1^s) + K_s, \end{aligned}$$

as well as

$$\begin{aligned} -i T_{1-Y} T_{1+a} \tilde{w}_{1,2}^s &= -i T_{1-Y} \partial_\alpha L_1(T_{1-Y} W_{\alpha\alpha}, \partial_\alpha^{-2} T_{1+a} \tilde{w}_1^s) + K_s, \\ -i T_{1-Y} T_{1+a} \tilde{w}_{2,2}^s &= -i T_{1-Y} L_1(T_{1-\bar{Y}} \bar{W}_{\alpha\alpha}, \partial_\alpha^{-1} T_{1+a} \tilde{w}_1^s) + K_s. \end{aligned}$$

Thus, the L_1 and L terms in \mathcal{K}_1^s are cancelled or exchanged for their leading terms, respectively. This proves our claim for the second equation in (5.13).

We are now left to deal with the source terms $(\mathcal{G}_2^s, \mathcal{K}_2^s)$. We address these remaining source terms with a second normal form correction. Write

$$\begin{aligned} \tilde{w}_3^s &= -s T_{(1+W_\alpha)Y_\alpha} \partial_\alpha^{-1} \tilde{w}_1^s, \\ \tilde{r}_3^s &= 0. \end{aligned}$$

This is easily checked to cancel $(\mathcal{G}_2^s, \mathcal{K}_2^s)$, when inserted into (5.9). Then setting

$$\begin{aligned} \tilde{w}^s &= \tilde{w}_1^s + \tilde{w}_2^s + \tilde{w}_3^s, \\ \tilde{r}^s &= \tilde{r}_1^s + \tilde{r}_2^s + \tilde{r}_3^s \end{aligned}$$

concludes the proof of Proposition 5.6. ■

5.4. The quadratic bound for the paradifferential source term

Here we prove Lemmas 5.3, 5.4. In addition, in Lemmas 5.7, 5.8 we derive formulas for the para-material derivatives of w , r , and x .

Proof of Lemma 5.3. We recall our objective, which is to isolate the nonperturbative part of $(\mathcal{E}_0, \mathcal{K}_0)$ with bilinear estimates, provided in Lemma 5.3. For convenience, we recall that

$$\begin{aligned} \mathcal{E}_0(w, r) &= (1 + W_\alpha)(P\bar{m} + \bar{P}m), \\ \mathcal{K}_0(w, r) &= \bar{P}n - P\bar{n}, \end{aligned}$$

where

$$m = |1 - Y|^2(r_\alpha + R_\alpha w) + (1 - Y)^2 \bar{R}w_\alpha, \quad n = (1 - Y)\bar{R}(r_\alpha + R_\alpha w).$$

We consider \mathcal{E}_0 first. We may absorb

$$P((1 + W_\alpha)\bar{P}m) = P(W_\alpha \bar{P}m)$$

into G . To see this, we begin by applying Coifman–Meyer bounds to write

$$\|P(W_\alpha \bar{P}m)\|_{\dot{H}^{\frac{1}{4}}} \lesssim \|D^{\frac{1}{4}} W_\alpha\|_{\text{BMO}} \|\bar{P}m\|_{L^2}.$$

Then it remains to show that

$$\|\bar{P}m\|_{L^2} \lesssim_A A_{\frac{1}{4}} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}}. \tag{5.14}$$

Incidentally, this estimate also works toward proving that \mathcal{E}_0 satisfies the bound in (5.5) (we will prove (5.14) on each term of \mathcal{E}_0).

To prove (5.14) we observe that the differentiation in the terms of m is always applied to one of the holomorphic variables, which has to be at or below the frequency of the undifferentiated antiholomorphic variable. For instance, a typical term is

$$\|\bar{P}(\bar{Y}r_\alpha)\|_{L^2} \lesssim \|r\|_{\dot{H}^{\frac{3}{4}}} \| |D|^{1/4} \bar{Y} \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}} \|r\|_{\dot{H}^{\frac{3}{4}}}.$$

(Here, also note that in the case of $\Pi(\bar{Y}, r_\alpha)$, derivatives cannot be shifted to the high frequency W_α . We have addressed this with our choice of function space $\dot{\mathcal{H}}^{\frac{1}{4}}$.) We also note the slightly different case where w is not differentiated, e.g.

$$\begin{aligned} \|\bar{P}(\bar{Y}R_\alpha w)\|_{L^2} &\lesssim \|w\|_{L^4} \|\bar{P}(\bar{Y}R_\alpha)\|_{L^4} \lesssim \|w\|_{\dot{H}^{\frac{1}{4}}} \|D^{\frac{1}{4}} Y\|_{\text{BMO}} \| |D|^{3/4} R \|_{L^4} \\ &\lesssim_A A^\sharp A_{\frac{1}{4}} \|w\|_{\dot{H}^{\frac{1}{4}}}. \end{aligned}$$

All other terms are similar to one of the two considered above.

Similarly, we have

$$(1 + W_\alpha)P\bar{m} = T_{1+W_\alpha}P\bar{m} + G,$$

where the G term is perturbative. Then expand $P\bar{m}$, with terms

$$P[|1 - Y|^2(\bar{r}_\alpha + \bar{R}_\alpha\bar{w})] = P[(1 - \bar{Y})(\bar{r}_\alpha + \bar{R}_\alpha\bar{w})Y], \quad P[(1 - \bar{Y})^2\bar{w}_\alpha R].$$

We can peel off frequency components, using the observation that cubic terms where the lowest frequency variable is differentiated may be absorbed into G , similar to the above estimates. In particular, here Y and R respectively cannot be the lowest frequency due to P , and the cases where \bar{r}_α , \bar{R}_α , and \bar{w}_α are the lowest frequency, respectively, are perturbative. The remaining cases form

$$\begin{aligned} P[T_{1-\bar{Y}}(\bar{r}_\alpha + T_{\bar{w}}\bar{R}_\alpha)Y] &= P[(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{x}}\bar{R}_\alpha)Y] + G, \\ P[T_{(1-\bar{Y})^2}\bar{w}_\alpha R] &= P[T_{1-\bar{Y}}\bar{x}_\alpha R] + G. \end{aligned}$$

After applying para-associativity Lemma 2.7 to factor in T_{1+W_α} , we obtain $\mathcal{G}_{0,0}$.

The analysis for $P\mathcal{K}_0 = P\bar{n}$ is similar, based on proving the bound

$$\|P\bar{n}\|_{\dot{H}^{\frac{1}{2}}} \lesssim_A A_{\frac{1}{4}} \|(w, r)\|_{\dot{H}^{\frac{1}{4}}}.$$

■

Proof of Lemma 5.4. We recall that

$$\begin{aligned} \mathcal{G}_1 &= -(T_{w_\alpha}b + \Pi(w_\alpha, b)) + (T_{r_\alpha}\bar{Y} + \Pi(r_\alpha, \bar{Y})) \\ &\quad - (T_w(1 - \bar{Y})R_\alpha + \Pi(w, (1 - \bar{Y})R_\alpha)), \\ \mathcal{K}_1 &= -(T_{r_\alpha}b + \Pi(r_\alpha, b)) \\ &\quad + i(T_{1-Y}T_w a + T_{1-Y}\Pi(w, a) - T_{(1+a)w}Y - \Pi((1 + a)w, Y)). \end{aligned}$$

Putting aside the nonperturbative terms on the right-hand side of Lemma 5.4, and those antiholomorphic terms eliminated by P , we observe that it remains to bound

$$-(T_{w_\alpha}b + \Pi(w_\alpha, b)) + \Pi(r_\alpha, \bar{Y}) - (T_w\Pi(-\bar{Y}, R_\alpha) + \Pi(w, (1 - \bar{Y})R_\alpha)),$$

respectively

$$-(T_{r_\alpha}b + \Pi(r_\alpha, b)) + i(T_{1-Y}T_w a + T_{1-Y}\Pi(w, a) - \Pi((1 + a)w, Y)).$$

We consider terms from \mathcal{G}_1 first. We observe that all the terms are balanced, with derivatives on low frequencies, so that we have the desired estimate. For instance,

$$\|T_{w_\alpha}b\|_{L^2} \lesssim \|w\|_{\dot{H}^{1/4}} \| |D|^{3/4}b \|_{\text{BMO}} \lesssim_A \|w\|_{\dot{H}^{1/4}} \| |D|^{3/4}R \|_{\text{BMO}}.$$

A similar analysis holds for most of the terms from \mathcal{K}_1 , except

$$T_{1-Y}T_w a + T_{1-Y}\Pi(w, a).$$

For these terms, we use

$$\|T_{1-Y}T_w a\|_{\dot{H}^{1/2}} \lesssim \|Y\|_{L^\infty} \|w\|_{L^4} \| |D|^{1/2}a \|_{L^4} \lesssim_{A^\#} A_{\frac{1}{4}} \|w\|_{\dot{H}^{1/4}},$$

using here Lemma 3.3 to estimate a .

■

Putting these lemmas together, we obtain the following formulas, with estimates, for the para-material derivatives of (w, r) :

Lemma 5.7. *Given (w, r) satisfying (5.3), we have the representations*

$$\begin{cases} T_{D_t} w = -T_{1-\bar{Y}}(r_\alpha + T_w R_\alpha) + G_2 =: G_1 + G_2, \\ T_{D_t} r = iT_{1+a}(T_{1-Y} w - T_w Y) + K_2 =: K_1 + K_2, \end{cases}$$

and likewise with $D_t(w, r)$ in place of $T_{D_t}(w, r)$, where (G_2, K_2) satisfies the quadratic bound

$$\|(G_2, K_2)\|_{\mathcal{H}^0} \lesssim_{A^\#} A_{\frac{1}{4}} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}},$$

and (G_1, K_1) satisfies the linear bound

$$\||D|^{-1/4}(G_1, K_1)\|_{\mathcal{H}^0} \lesssim_{A^\#} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}}.$$

Proof. Apply Lemmas 5.3 and 5.4. Then observe that

$$\|T_{(1-\bar{Y})R_\alpha} w\|_{L^2} \lesssim \|\bar{Y}\|_{L^\infty} \||D|^{3/4} R\|_{\text{BMO}} \|w\|_{\dot{H}^{1/4}},$$

so that this remaining term from the left-hand side of (5.3) may be absorbed in G_2 . Note that the para-commutator of Lemma 2.4 allows us to reorder paraproducts freely.

To estimate (G_1, K_1) , the terms

$$-T_{1-\bar{Y}} r_\alpha, \quad iT_{1+a} T_{1-Y} w$$

are straightforward, with a sufficient balance of derivatives on (w, r) already. For the remaining two terms, also apply Sobolev embedding. For instance,

$$\||D|^{-1/4} T_{1-\bar{Y}} T_w R_\alpha\|_{L^2} \lesssim \|\bar{Y}\|_{L^\infty} \|w\|_{L^4} \||D|^{3/4} R\|_{L^4} \lesssim_A A A^\# \|w\|_{\dot{H}^{1/4}}. \quad \blacksquare$$

We likewise have an expansion for the para-material derivative of x .

Lemma 5.8. *Given (w, r) satisfying (5.3), we have the representation*

$$T_{D_t} x = -T_{1-\bar{Y}}(T_{1-Y} r_\alpha + T_x R_\alpha) + G_2 =: G_1 + G_2,$$

and likewise with $D_t x$ in place of $T_{D_t} x$, where G_2 satisfies the quadratic bound

$$\|G_2\|_{L^2} \lesssim_{A^\#} A_{\frac{1}{4}} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}},$$

and G_1 satisfies the linear bound

$$\||D|^{-1/4} G_1\|_{L^2} \lesssim_{A^\#} \|(w, r)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}}.$$

Even though this is not needed later, we remark that G_2 also satisfies the last bound.

Proof of Lemma 5.8. Using the para-Leibniz rule Lemma 3.6, we may write (collecting all terms satisfying the appropriate estimate into G_2)

$$T_{D_t}x = T_{-T_{D_t}Y}w + T_{1-Y}T_{D_t}w + G_2.$$

To the first term, we apply Lemma 3.8 to see that it is perturbative. Apply Lemma 5.7 to the second term, to obtain (rewriting using the para-commutator Lemma 2.4, and dropping perturbative components as usual)

$$-T_{1-Y}T_{1-\bar{Y}}(r_\alpha + T_w R_\alpha) + G_2 = -T_{1-\bar{Y}}(T_{1-Y}r_\alpha + T_x R_\alpha) + G_2. \quad \blacksquare$$

5.5. Normal form analysis for $(\mathcal{G}_1, \mathcal{K}_1)$

In this section we introduce normal form corrections to (w, r) which will cancel the leading-order contributions of the paradifferential truncations $(\mathcal{G}_1, \mathcal{K}_1)$. Precisely, we will show the following.

Proposition 5.9. *Assume that (w, r) solve (5.3). Then there exists a linear normal form correction*

$$(\tilde{w}, \tilde{r})(t) = \text{NF}((w, r)(t))$$

solving an equation of the form

$$\begin{cases} T_{D_t}\tilde{w} + T_{1-\bar{Y}}\partial_\alpha\tilde{r} + T_{(1-\bar{Y})R_\alpha}\tilde{w} = -P\mathcal{G}_1(w, r) + G, \\ T_{D_t}\tilde{r} - iT_{1-Y}T_{1+a}\tilde{w} = -P\mathcal{K}_1(w, r) + K, \end{cases}$$

with the following properties:

(i) *Quadratic correction bound:*

$$\|\text{NF}(w, r)\|_{\dot{\mathcal{H}}^{1/4}} \lesssim_A A^\# \|(w, r)\|_{\dot{\mathcal{H}}^{1/4}}. \quad (5.15)$$

(ii) *Secondary correction bound:*

$$\|\text{NF}(g, k)\|_{\dot{\mathcal{H}}^{1/4}} \lesssim_A A_{\frac{1}{4}} \|(g, k)\|_{\mathcal{H}^0}. \quad (5.16)$$

(iii) *Cubic error bound:*

$$\|(G, K)\|_{\dot{\mathcal{H}}^{1/4}} \lesssim_{A^\#} A_{\frac{1}{4}}^2 \|(w, r)\|_{\dot{\mathcal{H}}^{1/4}}. \quad (5.17)$$

The motivation for the secondary bound is that the normal form corrections will be constructed not in terms of solutions for the homogeneous problem (4.5), but in terms of solutions for the inhomogeneous problem (5.3) (which recall originally takes the form (4.2)). As such, we need to be able to estimate the effect of source terms via (ii). For this purpose we use the \mathcal{H}^0 norm, which is suitable in conjunction with the source terms (G_2, K_2) of Lemma 5.7.

We will build (\tilde{w}, \tilde{r}) in three steps. In the first step, we apply a correction to generate the quadratic paradifferential truncations $(\mathcal{G}_1, \mathcal{K}_1)$ except temporarily putting aside the cubic truncation $T_w a$ corresponding to the Taylor coefficient a , which is itself quadratic.

By viewing the linearized variable w as a low frequency coefficient, $T_w a$ may be viewed as quadratic in the nonlinear variable R . More generally, the correction of the first step misses unbalanced cubic source terms which are quadratic in the nonlinear variables (W, R) with low frequency w coefficients. We address most of these source terms via a correction in the second step, other than those quadratic in R , which we resolve in a third step.

Before defining the corrections (\tilde{w}, \tilde{r}) , we enumerate notation that we will use in this subsection. Corresponding to the three steps above, we will define

$$(\tilde{w}, \tilde{r}) = (\tilde{w}_1, \tilde{r}_1) + (\tilde{w}_2, \tilde{r}_2) + (\tilde{w}_3, \tilde{r}_3),$$

where $(\tilde{w}_i, \tilde{r}_i)$ will be defined below. We denote the source terms of the paradifferential $(\tilde{w}_i, \tilde{r}_i)$ equation by $(\tilde{\mathcal{G}}_i, \tilde{\mathcal{K}}_i)$, i.e.

$$\begin{cases} T_{D_t} \tilde{w}_i + T_{1-\bar{Y}} \partial_\alpha \tilde{r}_i + T_{(1-\bar{Y})R_\alpha} \tilde{w}_i = \tilde{\mathcal{G}}_i, \\ T_{D_t} \tilde{r}_i - iT_{1-Y} T_{1+a} \tilde{w}_i = \tilde{\mathcal{K}}_i. \end{cases} \tag{5.18}$$

Throughout, our normal form corrections reflect two cases:

- (i) high–low interactions,
- (ii) balanced interactions with low frequency output.

Regarding the second case, it is convenient to let $\Pi = P \Pi$ throughout, thus including an implicit projection P .

Lastly, it will be convenient to state our corrections in terms of the following intermediate between W and $\partial_\alpha^{-1} Y$:

$$X := T_{1-Y} W.$$

We proceed to define our first set of corrections,

$$\begin{aligned} \tilde{w}_1 &= -\partial_\alpha (T_w X + \Pi(w, X)) - \Pi(w_\alpha, \bar{X}), \\ \tilde{r}_1 &= -(T_{r_\alpha} X + \Pi(r_\alpha, X)) - \Pi(r_\alpha, \bar{X}). \end{aligned}$$

We remark that at the quadratic level, this correction is nothing but the corresponding (i.e. non-paradifferential) holomorphic (i.e. linear in (w, r)) part of the linearization of the normal form transformation (1.12).

We first verify that this correction satisfies (5.15). Indeed, by Sobolev embeddings and paraproduct bounds,

$$\|\partial_\alpha (T_w X)\|_{\dot{H}^{1/4}} \lesssim \|w\|_{L^4} \| |D|^{1/4} X_\alpha \|_{L^4} \lesssim_A \|w\|_{\dot{H}^{1/4}} \|W_\alpha\|_{\dot{W}^{1/4,4}}.$$

Similarly,

$$\|T_{r_\alpha} X\|_{\dot{H}^{3/4}} \lesssim \| |D|^{-1/4} r_\alpha \|_{L^2} \|X_\alpha\|_{\text{BMO}} \lesssim_A \|r\|_{\dot{H}^{3/4}} \|W_\alpha\|_{L^\infty}.$$

The balanced frequency terms are straightforward; for instance,

$$\|\partial_\alpha \Pi(w, X)\|_{\dot{H}^{1/4}} \lesssim_A \|w\|_{\dot{H}^{1/4}} \|W_\alpha\|_{L^\infty}.$$

It is also straightforward to verify that this correction satisfies the secondary bound (5.16); for instance,

$$\|\partial_\alpha(T_g X)\|_{\dot{H}^{1/4}} \lesssim \| |D|^{1/4} W_\alpha \|_{\text{BMO}} \|g\|_{L^2}.$$

Next we turn to the source terms, and prove the following.

Lemma 5.10. *We have the representation*

$$\tilde{\mathcal{G}}_1 = -P\mathcal{G}_1 + \mathcal{G}_{1,1} + G, \quad \tilde{\mathcal{K}}_1 = -P\mathcal{K}_1 + \mathcal{K}_{1,1} + K$$

where (G, K) satisfy the bound (5.17), whereas $(\mathcal{G}_{1,1}, \mathcal{K}_{1,1})$ are given by

$$\begin{aligned} \mathcal{G}_{1,1}(w, r) &= T_w(T_{1-\bar{Y}}\partial_\alpha^2 \Pi(X, R) + \Pi(Y_\alpha, \bar{R}) + T_{1-\bar{Y}}\Pi(R_{\alpha\alpha}, \bar{X}) - PM), \\ \mathcal{K}_{1,1}(w, r) &= iT_w(T_{1+a}\partial_\alpha \Pi(Y, X) + T_{1+a}\Pi(Y_\alpha, \bar{X}) + T_{1-Y}Pa). \end{aligned}$$

In other words, our first normal form correction \tilde{w}_1 does not quite allow us to dispense with the paradifferential truncations $(\mathcal{G}_1, \mathcal{K}_1)$. Instead, it allows us to replace them with the milder source term $(\mathcal{G}_{1,1}, \mathcal{K}_{1,1})$ (modulo perturbative errors (G, K)). These residual source terms are milder in that they are cubic, though unbalanced and hence not quite perturbative.

Throughout the computations below, we use the observation that cubic and higher-order terms such that the lowest frequency variable is “fully differentiated” may be absorbed into G or K , where

$$(|D|^{5/4}W, |D|^{3/4}R) \in \text{BMO}, \quad (|D|^{1/4}w, |D|^{3/4}r) \in L^2.$$

For instance, for the following typical term appearing in the first equation, the lowest frequency variable is either Y or r , and both are fully differentiated (and in fact over-differentiated):

$$\|T_{Y_\alpha} T_{r_\alpha} W\|_{\dot{H}^{\frac{1}{4}}} \lesssim \| |D|^{-3/4} Y_\alpha \|_{\text{BMO}} \| |D|^{-1/4} r_\alpha \|_{L^2} \| |D|^{5/4} W \|_{\text{BMO}} \lesssim_A A_{\frac{1}{4}}^2 \|r\|_{\dot{H}^{\frac{3}{4}}},$$

which verifies (5.17).

For quartic and higher-order terms, terms such that the lowest frequency variable is a linearized variable w (even with no derivatives) may also be absorbed into G or K , by measuring w in L^4 and applying a Sobolev embedding. For instance, for the following typical term appearing in the first equation, the lowest frequency variable may also be w :

$$\begin{aligned} \|T_{Y_\alpha R_\alpha} T_w W\|_{\dot{H}^{\frac{1}{4}}} &\lesssim \| |D|^{-3/4} Y_\alpha \|_{\text{BMO}} \| |D|^{-1/4} R_\alpha \|_{\text{BMO}} \|w\|_{L^4} \|D^{\frac{1}{4}} W_\alpha\|_{L^4} \\ &\lesssim_{A^\#} A_{\frac{1}{4}}^2 \|w\|_{\dot{H}^{\frac{1}{4}}}, \end{aligned}$$

which verifies (5.17).

Proof of Lemma 5.10. We compute the contributions from $(\tilde{w}_1, \tilde{r}_1)$ when inserted in the left-hand side of (5.18), separating the output into the three components as in the statement of the lemma. It is instructive to compare the leading-order bilinear interactions with the normal form computations in the proof of Proposition 5.6, though the situation here is more delicate as we cannot apply an exponential conjugation to lower the order of our source terms and corrections.

(1) We begin by computing the first equation in (5.18), starting with the high–low corrections in $(\tilde{w}_1, \tilde{r}_1)$. After commuting D_t with ∂_α ,

$$-T_{D_t} \partial_\alpha T_w X = -\partial_\alpha T_{D_t} T_w X + T_{b_\alpha} \partial_\alpha T_w X = -\partial_\alpha T_{D_t} T_w X + T_{b_\alpha} T_w X_\alpha + G,$$

we have the following four contributions to the first equation, using the para-Leibniz rule in Lemma 3.6 (precisely the $\sigma = 0$ case of the bound (3.9)), then part (b) of Lemma 3.8 for the para-material derivative of X , and Lemma 5.7 for the para-material derivative of (w, r) (combined with the secondary bound (5.16) on $(\tilde{w}_1, \tilde{r}_1)$ to justify absorbing the source terms into (G, K)):

$$\begin{aligned} -\partial_\alpha T_{D_t} T_w X &= -\partial_\alpha (T_{D_t} w X + T_w T_{D_t} X) + G \\ &= \partial_\alpha (T_{T_{(1-\bar{Y})}(r_\alpha + T_w R_\alpha)} X + T_w (-T_{T_{1-\bar{Y}} R_\alpha} X + P((1-\bar{Y})R) \\ &\quad + \Pi(X_\alpha, b))) + G \\ &= \partial_\alpha T_{T_{1-\bar{Y}} r_\alpha} X + \partial_\alpha T_w P[(1-\bar{Y})R] + T_w \partial_\alpha \Pi(X_\alpha, b) + G, \\ T_{b_\alpha} T_w X_\alpha &= T_w T_{T_{1-\bar{Y}} R_\alpha + T_{1-\bar{Y}} \bar{R}_\alpha} X_\alpha + G, \\ -T_{1-\bar{Y}} \partial_\alpha T_{r_\alpha} X &= -\partial_\alpha T_{T_{1-\bar{Y}} r_\alpha} X + G, \\ -T_{(1-\bar{Y}) R_\alpha} \partial_\alpha T_w X &= -T_w T_{T_{1-\bar{Y}} R_\alpha} X_\alpha + G. \end{aligned}$$

Consider the middle term of the first contribution. Applying the product rule, we have

$$\partial_\alpha T_w P[(1-\bar{Y})R] = T_{w_\alpha} P[(1-\bar{Y})R] + T_w P[(1-\bar{Y})R_\alpha] - T_w P[\bar{Y}_\alpha R].$$

The first two terms (note $T_{w_\alpha} P[(1-\bar{Y})R] = P T_{w_\alpha} b$) are two of the high–low paradifferential truncations in \mathcal{E}_1 .

Summing all remaining nonperturbative terms and observing cancellations, we obtain the secondary error terms

$$T_w P[-\bar{Y}_\alpha R + \partial_\alpha \Pi(X_\alpha, b) + T_{T_{1-\bar{Y}} \bar{R}_\alpha} X_\alpha] \tag{5.19}$$

which we collect in $\mathcal{E}_{1,1}$.

(2) Continuing with the first equation in (5.18), we have analogous computations for the frequency-balanced corrections with holomorphic variables, with the following differences:

- (i) Several terms have better derivative balance and may be absorbed into G directly. Indeed, all of the errors corresponding to those in (5.19) are perturbative.

(ii) On the other hand, the following cancellation no longer applies, as the right-hand side is in fact perturbative:

$$\partial_\alpha \Pi(T_{1-\bar{Y}} T_w R_\alpha, X) \neq \partial_\alpha \Pi(w, T_{T_{1-\bar{Y}} R_\alpha} X) = G.$$

As a result, we obtain the desired balanced interaction truncations analogous to before, but with a different residual for $\mathcal{G}_{1,1}$,

$$T_w T_{1-\bar{Y}} \partial_\alpha \Pi(R_\alpha, X), \tag{5.20}$$

where we have used Lemma 2.6 to factor out the coefficients.

(3) Lastly for the first equation in (5.18), we have similar computations for the frequency-balanced corrections with one antiholomorphic variable. The difference from (2) is that we include only the correction where ∂_α falls on w . Using Lemmas 3.6, 3.8 we have

$$\begin{aligned} -T_{D_t} \Pi(w_\alpha, \bar{X}) &= \Pi(\partial_\alpha T_{D_t} w, \bar{X}) + \Pi(w_\alpha, T_{D_t} \bar{X}) + G \\ &= T_{1-\bar{Y}} \Pi(r_{\alpha\alpha} + T_w R_{\alpha\alpha}, \bar{X}) + \Pi(w_\alpha, T_{1-Y} \bar{R}) + G, \\ -T_{1-\bar{Y}} \partial_\alpha \Pi(r_\alpha, \bar{X}) &= -T_{1-\bar{Y}} \Pi(r_{\alpha\alpha}, \bar{X}) - \Pi(r_\alpha, T_{1-\bar{Y}} \bar{X}_\alpha) + G. \end{aligned}$$

The second term of the first contribution,

$$\Pi(w_\alpha, T_{1-Y} \bar{R}) = \Pi(w_\alpha, P((1-Y)\bar{R})) + G,$$

is the other desired balanced interaction truncation in \mathcal{G}_1 from $\Pi(w_\alpha, b)$.

Consider the second term of the second contribution. Applying Lemma 3.2 to express X in terms of Y , we have

$$-\Pi(r_\alpha, T_{1-\bar{Y}} \bar{X}_\alpha) = -\Pi(r_\alpha, \bar{Y}) + G.$$

This is the last desired paradifferential truncation in \mathcal{G}_1 .

After cancellations, the only residual term for $\mathcal{G}_{1,1}$ is

$$T_{1-\bar{Y}} \Pi(T_w R_{\alpha\alpha}, \bar{X}) = T_w T_{1-\bar{Y}} \Pi(R_{\alpha\alpha}, \bar{X}) + G, \tag{5.21}$$

where we have used Lemma 2.6 to factor out the coefficients.

Summing (5.19), (5.20), and (5.21), we have the residual terms

$$-T_w P[\bar{Y}_\alpha R + \partial_\alpha \Pi(X_\alpha, b) + T_{T_{1-Y} \bar{R}_\alpha} X_\alpha + \partial_\alpha T_{1-\bar{Y}} \Pi(R_\alpha, X) + T_{1-\bar{Y}} \Pi(R_{\alpha\alpha}, \bar{X})].$$

Applying perturbative modifications, and in particular Lemma 3.2, we obtain exactly $\mathcal{G}_{1,1}$ modulo perturbative terms G .

(4) Next we compute the second equation in (5.18). We begin with the high–low corrections. Applying the bound (3.9) in Lemma 3.6 and part (c) of Lemma 3.8 for the first

identity and Lemma 3.2 for the second, we have

$$\begin{aligned}
 -T_{D_t} T_{r_\alpha} X &= T_{\partial_\alpha T_{D_t} r} X + T_{r_\alpha} T_{D_t} X + K \\
 &= -T_{iT_{1+a}T_{1-Y}w_\alpha - T_{1+a}T_w Y_\alpha} X + T_{r_\alpha} T_{1-\bar{Y}} R + K \\
 &= -iT_{1+a}T_{((1-Y)w)_\alpha} X + T_{r_\alpha} T_{1-\bar{Y}} R + K, \\
 iT_{1-Y}T_{1+a}\partial_\alpha T_w X &= iT_{1-Y}T_{1+a}T_{((1-Y)w)_\alpha} W + iT_{1-Y}T_{1+a}T_w T_{1-Y}W_\alpha + K \\
 &= iT_{1+a}T_{((1-Y)w)_\alpha} X + iT_{1+a}T_w Y + iT_{1-Y}T_{1+a}T_w \Pi(Y, W_\alpha) + K.
 \end{aligned}$$

After summing the two contributions with cancellations and applying Lemma 3.2, we obtain

$$T_{r_\alpha} T_{1-\bar{Y}} R + iT_{1+a}T_w Y + iT_{1+a}T_w \Pi(Y, X_\alpha) + K.$$

The first two terms (note that $T_{r_\alpha} T_{1-\bar{Y}} R = P T_{r_\alpha} b + K$) are two of the high–low paradiﬀerential truncations in \mathcal{K}_1 , leaving to $\mathcal{K}_{1,1}$ only the third term,

$$iT_{1+a}T_w \Pi(Y, X_\alpha). \tag{5.22}$$

(5) Continuing with the second equation in (5.18), we have analogous computations for the frequency-balanced corrections with holomorphic variables, with the following diﬀerences, similar to step (2):

- (i) The error term in (5.22) becomes

$$iT_{1+a}\Pi(w, \Pi(Y, X_\alpha)),$$

which may be absorbed into K directly.

- (ii) On the other hand, the following cancellation no longer applies, as the right-hand side is in fact perturbative:

$$iT_{1+a}\Pi(Y_\alpha w, X) \neq iT_{1-Y}T_{1+a}\Pi(w, T_{Y_\alpha} W) = K.$$

As a result, we obtain the desired balanced interaction truncations analogous to before, but with a diﬀerent residual for $\mathcal{K}_{1,1}$,

$$iT_{1+a}T_w \Pi(Y_\alpha, X). \tag{5.23}$$

(6) Lastly, for the second equation in (5.18), we have similar computations for the frequency-balanced corrections with one antiholomorphic variable. The only diﬀerence from (5) is that we lack the term analogous to

$$iT_{1-Y}T_{1+a}\Pi(w, X_\alpha)$$

(contributed by \tilde{w}_1), which would be

$$iT_{1-Y}T_{1+a}\Pi(w, \bar{X}_\alpha),$$

but accordingly, this is absent as a paradifferential truncation in \mathcal{K}_1 . We thus obtain only the residual term analogous to that in (5):

$$iT_{1+a}T_w\Pi(Y_\alpha, \bar{X}). \tag{5.24}$$

Summing (5.22), (5.23), and (5.24), and observing that in \mathcal{K}_1 , only the high–low paradifferential error from the Taylor coefficient a remains, we obtain $\mathcal{K}_{1,1}$. ■

We proceed with the second step in the construction of the normal form correction, which cancels those errors in $(\mathcal{G}_{1,1}, \mathcal{K}_{1,1})$ which can be seen as quadratic in the nonlinear variables (W, R) and with w as a low frequency coefficient. The corresponding correction addresses the quadratic part, while w is simply retained as a coefficient. We define this correction as⁴

$$\begin{aligned} \tilde{w}_2 &= T_w\left(-T_{\frac{1}{1+a}}Pa + \frac{1}{2}\partial_\alpha^2\Pi(X, X) + \Pi(X_{\alpha\alpha}, \bar{X})\right), \\ \tilde{r}_2 &= 0. \end{aligned}$$

We first verify that this correction satisfies (5.15). By Sobolev embedding and Proposition 3.3,

$$\|T_wT_{\frac{1}{1+a}}a\|_{\dot{H}^{\frac{1}{4}}} \lesssim \|w\|_{L^4}(1 + \|a\|_{L^\infty})\| |D|^{1/4}a\|_{L^4} \lesssim_A A^{\#2}\|w\|_{\dot{H}^{1/4}}.$$

The balanced terms are then easily estimated.

As with $(\tilde{w}_1, \tilde{r}_1)$, the secondary bound (5.16) is also easily verified. For instance,

$$\|T_fT_{\frac{1}{1+a}}a\|_{\dot{H}^{\frac{1}{4}}} \lesssim \|f\|_{L^2}(1 + \|a\|_{L^\infty})\| |D|^{1/4}a\|_{\text{BMO}} \lesssim_A AA_{\frac{1}{4}}\|f\|_{L^2}.$$

Next we turn our attention to the source terms.

Lemma 5.11. *We have the representation*

$$\tilde{\mathcal{G}}_2 = -\mathcal{G}_{1,1} + \mathcal{G}_{1,2} + G, \quad \tilde{\mathcal{K}}_2 = -\mathcal{K}_{1,1} + \mathcal{K}_{1,2} + K,$$

where (G, K) satisfy the bound (5.17), whereas $(\mathcal{G}_{1,2}, \mathcal{K}_{1,2})$ are given by

$$\mathcal{G}_{1,2}(w, r) = 0, \quad \mathcal{K}_{1,2}(w, r) = 2iT_xPa.$$

We see this nearly dispenses with $(\mathcal{G}_{1,1}, \mathcal{K}_{1,1})$, other than a residual Taylor coefficient truncation in $\mathcal{K}_{1,2}$, which we will leave to the third step of our normal form construction.

⁴This is obtained again by a standard normal form analysis, but taking care to also match the low frequency coefficients.

Proof of Lemma 5.11. We begin by considering the Taylor coefficient correction term in \tilde{w}_2 , starting with the contribution to the first equation of (5.18). This is done as in the proof of the previous lemma, using Lemma 3.6 to distribute the para-material derivative:

$$-T_{D_t} T_w T_{\frac{1}{1+a}} a = T_{D_t} w T_{\frac{1}{1+a}} a + T_w T_{D_t} T_{\frac{1}{1+a}} a + T_w T_{\frac{1}{1+a}} T_{D_t} a + G.$$

The first term goes into G using the second estimate of Lemma 5.7 for the material derivative of w along with the bounds on a in Lemma 3.3:

$$\|T_{D_t} w T_{\frac{1}{1+a}} a\|_{\dot{H}^{\frac{1}{4}}} \lesssim_A \|T_{D_t} w\|_{\dot{H}^{-\frac{1}{4}}} \| |D|^{\frac{1}{2}} a \|_{\text{BMO}} \lesssim_{A^\#} A_{\frac{1}{4}}^2 \|(w, r)\|_{\dot{H}^{\frac{1}{4}}}.$$

The second term is also perturbative, as we can use (3.2) to complete T_{D_t} to D_t and then finish with (3.17). Lastly, we have by (3.16),

$$T_{D_t} P a = -T_{1+a} P M + E,$$

where E is perturbative via the error bound

$$\|E\|_{\dot{W}^{\frac{1}{4},4}} \lesssim_{A^\#} A_{\frac{1}{4}}^2.$$

Summarizing, we have proved that

$$-T_{D_t} T_w T_{\frac{1}{1+a}} P a = T_w P M + G$$

which matches the last two terms of $\mathcal{E}_{1,1}$. To the second equation, this correction term generates

$$i T_{1-Y} T_{1+a} T_w T_{\frac{1}{1+a}} P a = i T_x P a + K.$$

This contributes to $\mathcal{K}_{1,2}$ a second copy of the Taylor coefficient term, where we have rewritten

$$i T_w T_{1-Y} P a = i T_x P a + K,$$

perturbatively peeling off the case where w is lowest frequency, as this component is quartic.

We proceed with the remaining correction terms, which are somewhat simpler:

(1) We compute the first equation in (5.18), starting with the term with holomorphic variables and using Lemmas 3.6 and 3.8:

$$\frac{1}{2} T_{D_t} T_w \partial_\alpha^2 \Pi(X, X) = -T_w T_{1-\bar{Y}} \partial_\alpha^2 \Pi(X, R) + G$$

which is the first term of $\mathcal{E}_{1,1}$.

(2) Continuing with the first equation, we address the term with one antiholomorphic variable in the same manner:

$$T_{D_t} T_w \Pi(X_{\alpha\alpha}, \bar{X}) = -T_w (T_{1-\bar{Y}} \Pi(R_{\alpha\alpha}, \bar{X}) + T_{1-Y} \Pi(X_{\alpha\alpha}, \bar{R})) + G.$$

These are the remaining terms of $\mathcal{E}_{1,1}$, using Lemma 3.2 on the second term.

(3) We compute the second equation in (5.18), starting with the term with holomorphic variables:

$$-\frac{1}{2}iT_{1-Y}T_{1+a}T_w\partial_\alpha^2\Pi(X, X) = -iT_{1-Y}T_{1+a}T_w\partial_\alpha\Pi(X_\alpha, X) + K,$$

which is the first term of $\mathcal{K}_{1,1}$, after applying Lemma 3.2.

(4) Continuing with the second equation, we address the term with one antiholomorphic variable:

$$-iT_{1-Y}T_{1+a}T_w\Pi(X_\alpha, \bar{X}) = -iT_{1+a}T_w\Pi(Y_\alpha, \bar{X}) + G.$$

This is the remaining term of $\mathcal{K}_{1,1}$. ■

It remains to cancel the paradifferential truncation from the Taylor coefficient. We define the correction

$$\begin{aligned} \tilde{w}_3 &= -P((T_w\bar{X}_\alpha)X_\alpha) - T_{\frac{1}{1+a}}T_wPa, \\ \tilde{r}_3 &= -P((T_w\bar{X}_\alpha)R). \end{aligned}$$

As with $(\tilde{w}_2, \tilde{r}_2)$, it is straightforward to verify that this correction satisfies (5.15) and (5.16).

Lemma 5.12. *We have*

$$\tilde{\mathcal{E}}_3 = -\mathcal{E}_{1,2} + G = G, \quad \tilde{\mathcal{K}}_2 = -\mathcal{K}_{1,2} + K = -2iT_xPa + K,$$

where (G, K) satisfy the bound (5.17).

Proof. (1) We begin by checking that no nonperturbative errors are contributed to the first equation in (5.18). For this we repeatedly use Lemma 3.6, noting that the terms involving the para-material derivative T_{D_t} of the “coefficients” w and $\frac{1}{1+a}$, as well as those involving $[T_{D_t}, \partial_\alpha]$, are all perturbative. We also repeatedly apply the para-associativity lemma, Lemma 2.7, and use (3.16) to compute the para-material derivative of a :

$$\begin{aligned} -T_{D_t}P((T_w\bar{X}_\alpha)X_\alpha) &= P[(T_w\partial_\alpha T_{D_t}\bar{X})X_\alpha] + P[(T_w\bar{X}_\alpha)\partial_\alpha T_{D_t}X] + G \\ &= P[(T_wT_{1-Y}\bar{R}_\alpha)X_\alpha] + P[(T_w\bar{X}_\alpha)T_{1-\bar{Y}}R_\alpha] + G \\ &= T_wP[\bar{R}_\alpha Y] + P[(T_w\bar{X}_\alpha)T_{1-\bar{Y}}R_\alpha] + G, \\ -T_{D_t}T_{\frac{1}{1+a}}T_wPa &= -T_{\frac{1}{1+a}}T_wT_{D_t}Pa + G = T_wPM + G \\ &= T_wP[R\bar{Y}_\alpha - \bar{R}_\alpha Y] + G, \\ -T_{1-\bar{Y}}\partial_\alpha((T_w\bar{X}_\alpha)R) &= -T_{1-\bar{Y}}[(T_w\bar{X}_\alpha\alpha)R] - T_{1-\bar{Y}}[(T_w\bar{X}_\alpha)R_\alpha] + G, \\ T_{(1-\bar{Y})R_\alpha}(-T_w\bar{X}_\alpha)X_\alpha - T_{\frac{1}{1+a}}T_wPa &= G. \end{aligned}$$

Summing and using Lemma 3.2, all the terms cancel up to G errors.

(2) For the second equation of (5.18), we compute

$$\begin{aligned}
 -T_{D_t} P[(T_w \bar{X}_\alpha)R] &= P[(T_w T_{1-Y} \bar{R}_\alpha)R] \\
 &\quad - iP[(T_w \bar{X}_\alpha)T_{1+a}Y] + K, \\
 -iT_{1-Y}T_{1+a}(-P[(T_w \bar{X}_\alpha)X_\alpha] - T_{\frac{1}{1+a}}T_w Pa) &= iT_{1+a}P[(T_w \bar{X}_\alpha)Y] + iT_x Pa + K.
 \end{aligned}$$

Summing, we see that two terms cancel, leaving two copies of the desired Taylor coefficient term. ■

5.6. Normal form analysis for $(\mathcal{G}_0, \mathcal{K}_0)$

In this section we introduce normal form corrections to (w, r) which will cancel the leading-order contributions of the source terms $(\mathcal{G}_0, \mathcal{K}_0)$ of the linearized equations, defined in (4.3). Precisely, we will show the following.

Proposition 5.13. *Assume that (w, r) solve (5.3). Then there exists a linear normal form correction*

$$(\tilde{w}, \tilde{r})(t) = \text{NF}((w, r)(t))$$

solving an equation of the form

$$\begin{cases} T_{D_t} \tilde{w} + T_{1-\bar{Y}} \partial_\alpha \tilde{r} + T_{(1-\bar{Y})R_\alpha} \tilde{w} = -P \mathcal{G}_0 + G, \\ T_{D_t} \tilde{r} - iT_{1-Y}T_{1+a} \tilde{w} = -P \mathcal{K}_0 + K, \end{cases}$$

with the following properties:

(i) *Quadratic correction bound:*

$$\| \text{NF}(w, r) \|_{\dot{\mathcal{H}}^{1/4}} \lesssim_A A^\sharp \| (w, r) \|_{\dot{\mathcal{H}}^{1/4}}.$$

(ii) *Secondary correction bound:*

$$\| \text{NF}(g, k) \|_{\dot{\mathcal{H}}^{1/4}} \lesssim_A A_{\frac{1}{4}} \| (g, k) \|_{\dot{\mathcal{H}}^0}.$$

(iii) *Cubic error bound:*

$$\| (G, K) \|_{\dot{\mathcal{H}}^{1/4}} \lesssim_{A^\sharp} A_{\frac{1}{4}}^2 \| (w, r) \|_{\dot{\mathcal{H}}^{1/4}}.$$

Similar to the normal form correction for the paradifferential truncations $(\mathcal{G}_1, \mathcal{K}_1)$, we will build the correction in three steps, denoting the output of the three steps by, respectively,

$$(\tilde{w}, \tilde{r}) = (\tilde{w}_1, \tilde{r}_1) + (\tilde{w}_2, \tilde{r}_2) + (\tilde{w}_3, \tilde{r}_3),$$

along with their source terms $(\mathcal{G}_i, \mathcal{K}_i)$ in (5.18).

In the first step, we construct a correction to generate the quadratic terms, which are linear in the linearized variables (w, r) . Second, we construct a correction to generate the terms which are quadratic in the nonlinear variables (W, R) and their conjugates (with \bar{w}

appearing as a low frequency coefficient). Lastly, we have a correction to address terms quadratic specifically in R .

We define our first set of corrections, which rectify the quadratic source terms which are linear in (w, r) (unlike the last two corrections, which are primarily quadratic in (W, R) and with (w, r) playing the role of coefficients):

$$\begin{aligned} \tilde{w}_1 &= -P(\bar{x}W_\alpha), \\ \tilde{r}_1 &= -P(\bar{x}T_{1+W_\alpha}R). \end{aligned}$$

We remark that at the quadratic level, this correction is nothing but the corresponding (i.e. non-paradifferential) antiholomorphic (i.e. linear in (\bar{w}, \bar{r})) part of the linearization of the normal form transformation (1.12).

These corrections satisfy the bound (5.15), as is easily verified with Sobolev embedding, as well as (5.16), similar to \tilde{w}_1 in Section 5.4.

We will prove the following, where we recall that $(\mathcal{G}_{0,0}, \mathcal{K}_{0,0})$ is defined in Lemma 5.3.

Lemma 5.14. *We have the representation*

$$\tilde{\mathcal{G}}_1 = -\mathcal{G}_{0,0} + \mathcal{G}_{0,1} + G, \quad \tilde{\mathcal{K}}_1 = -\mathcal{K}_{0,0} + \mathcal{K}_{0,1} + K,$$

where (G, K) satisfy the bound (5.17), whereas $(\mathcal{G}_{0,1}, \mathcal{K}_{0,1})$ are given by

$$\begin{aligned} \mathcal{G}_{0,1}(w, r) &= T_{\bar{x}}(\Pi(W_{\alpha\alpha}, b) + \Pi(T_{1-\bar{Y}}W_\alpha - T_{1+W_\alpha}\bar{Y}, R_\alpha) - T_{1+W_\alpha}PM), \\ \mathcal{K}_{0,1}(w, r) &= T_{\bar{x}}(iPa + 2T_{1+W_\alpha}T_{1-\bar{Y}}TR_\alpha R + \Pi(T_{1+W_\alpha}R_\alpha, b) + iT_{1+a}\Pi(Y, W_\alpha)). \end{aligned}$$

Proof. (1) We compute the first equation in (5.18). For this we use Lemma 3.6 to distribute the para-material derivative and then Lemma 5.8 to capture the leading part of $D_t\bar{x}$. On the other hand, for $T_{D_t}W_\alpha$ we need to use the full expression (3.11) of Lemma 3.7. We will also use Lemma 3.2 to relate Y and W_α paraproducts:

$$\begin{aligned} -T_{D_t}P(\bar{x}W_\alpha) &= -P[(D_t\bar{x})W_\alpha] - P[\bar{x}T_{D_t}W_\alpha] + G \\ &= P[T_{1-Y}(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{x}}\bar{R}_\alpha)W_\alpha] \\ &\quad + P[\bar{x}(\partial_\alpha T_{1+W_\alpha}P[(1-\bar{Y})R] + T_{b_\alpha}W_\alpha + \partial_\alpha \Pi(W_\alpha, b))] + G \\ &= P[T_{1+W_\alpha}(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{x}}\bar{R}_\alpha)Y] \\ &\quad + P[\bar{x}(TW_{\alpha\alpha}T_{1-\bar{Y}}R + T_{1+W_\alpha}P[-\bar{Y}_\alpha R] \\ &\quad\quad + T_{1+W_\alpha}P[(1-\bar{Y})R_\alpha])] \\ &\quad + P[\bar{x}(T_{b_\alpha}W_\alpha + \Pi(W_{\alpha\alpha}, b) + \Pi(W_\alpha, b_\alpha))] + G, \\ -T_{1-\bar{Y}}\partial_\alpha P(\bar{x}T_{1+W_\alpha}R) &= -P[T_{1+W_\alpha}T_{1-\bar{Y}}\bar{x}_\alpha R] \\ &\quad - P[\bar{x}TW_{\alpha\alpha}T_{1-\bar{Y}}R] \\ &\quad - P[\bar{x}T_{1+W_\alpha}T_{1-\bar{Y}}R_\alpha] + G, \\ -T_{(1-\bar{Y})R_\alpha}P[\bar{x}W_\alpha] &= -P[\bar{x}T_{(1-\bar{Y})R_\alpha}W_\alpha] + G. \end{aligned}$$

Summing and observing cancellations, we obtain the following contributions to the source terms in the first equation in (5.18):

$$\begin{aligned} & P[T_{1-\bar{Y}}T_{1+W_\alpha}(\bar{r}_\alpha + T_{\bar{w}}\bar{R}_\alpha)Y] - P[T_{1+W_\alpha}T_{1-\bar{Y}}\bar{x}_\alpha R] \\ & + P[\bar{x}(-T_{1+W_\alpha}P[\bar{Y}_\alpha R] - T_{1+W_\alpha}\Pi(\bar{Y}, R_\alpha))] \\ & + P[\bar{x}(T_{(1-Y)\bar{R}_\alpha}W_\alpha + \Pi(W_{\alpha\alpha}, b) + \Pi(W_\alpha, b_\alpha))] + G. \end{aligned}$$

The first two terms form $\mathcal{E}_{0,0}$, as desired. The remaining terms are precisely $\mathcal{E}_{0,1}$, after perturbative rearrangements (noting that the case where \bar{x} is balanced with the holomorphic terms is perturbative) and applying Lemma 3.2.

(2) We compute the second equation similarly, by applying Lemma 3.6 to distribute T_{D_t} and Lemma 5.8 for the leading part of $D_t\bar{x}$. We use (3.11) of Lemma 3.7 for $T_{D_t}R$, but (3.12) suffices for $T_{D_t}W_\alpha$. Thus we obtain

$$\begin{aligned} -T_{D_t}P[\bar{x}T_{1+W_\alpha}R] &= -P[(D_t\bar{x})T_{1+W_\alpha}R] - P[\bar{x}T_{D_t,W_\alpha}R] - P[\bar{x}T_{1+W_\alpha}T_{D_t}R] + K \\ &= P[T_{1-Y}(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{x}}\bar{R}_\alpha)T_{1+W_\alpha}R] + P[\bar{x}T_{T_{1-\bar{Y}}T_{1+W_\alpha}R_\alpha}R] \\ &\quad + P[\bar{x}T_{1+W_\alpha}(i(1+a)(1-Y) + T_{R_\alpha}b + \Pi(R_\alpha, b))] + K, \\ iT_{1-Y}T_{1+a}P[\bar{x}W_\alpha] &= iT_{1+a}P[\bar{x}(T_{1+W_\alpha}Y + \Pi(Y, W_\alpha))] + K. \end{aligned}$$

Summing and peeling away perturbative terms using para-commutator and paraproduct estimates, we obtain the following contributions to the source term in the second equation:

$$\begin{aligned} & P[(T_{1-\bar{Y}}\bar{r}_\alpha + T_{\bar{x}}\bar{R}_\alpha)R] \\ & + P[\bar{x}(ia + 2T_{1+W_\alpha}T_{1-\bar{Y}}T_{R_\alpha}R + \Pi(T_{1+W_\alpha}R_\alpha, b) + iT_{1+a}\Pi(Y, W_\alpha))] \\ & + K. \end{aligned}$$

The first term matches $\mathcal{K}_{0,0}$, as desired. The remaining terms are precisely $\mathcal{K}_{0,1}$ (again noting that the case where \bar{x} is balanced with the holomorphic terms is perturbative). ■

In the second step, we introduce the corrections which generate the source terms which are quadratic in the nonlinear variables appearing in $(\mathcal{E}_{0,1}, \mathcal{K}_{0,1})$:

$$\begin{aligned} \tilde{w}_2 &= T_{\bar{x}}(\Pi(W_{\alpha\alpha}, \Re X) - \Pi(W_\alpha, \Re X_\alpha) - \Pi(W_\alpha, X_\alpha) - T_{\frac{1}{1+a}}T_{1+W_\alpha}Pa), \\ \tilde{r}_2 &= T_{\bar{x}}(\Pi(T_{1+W_\alpha}R_\alpha, \Re X) - \Pi(\Re T_{1+W_\alpha}R, X_\alpha)). \end{aligned}$$

Here we have treated $T_{\bar{x}}$ as a low frequency coefficient which simply carries over to the corrections. These corrections satisfy the bound (5.15), as is easily verified with Sobolev embeddings, as well as (5.16), which is similar to the case of \tilde{w}_2 in Section 5.4. The source terms generated by these corrections in (5.18) are as follows.

Lemma 5.15. *We have the representation*

$$\tilde{\mathcal{E}}_2 = -\mathcal{E}_{0,1} + \mathcal{E}_{0,2} + G, \quad \tilde{\mathcal{K}}_1 = -\mathcal{K}_{0,1} + \mathcal{K}_{0,2} + K,$$

where (G, K) satisfy the bound (5.17), whereas $(\mathcal{G}_{0,2}, \mathcal{K}_{0,2})$ are given by

$$\begin{aligned} \mathcal{G}_{0,2}(w, r) &= 0, \\ \mathcal{K}_{0,2}(w, r) &= T_{\bar{x}}(2iPa + 2T_{1+W_\alpha}T_{1-\bar{Y}}T_{R_\alpha}R + \frac{1}{2}\partial_\alpha\Pi(T_{1+W_\alpha}R, b)). \end{aligned}$$

Proof. We begin by considering the Taylor coefficient correction in \tilde{w}_2 . Its contribution to the source term $\tilde{\mathcal{G}}_2$ of the first equation in (5.18) is computed using Lemma 3.6, and then (3.16) to compute the para-material derivative of a . Here, as well as in all computations from here on, the para-material derivative $T_{D_t}\bar{x}$ of \bar{x} , given by Lemma 5.8, plays a perturbative role. Note that the only nonperturbative case is when the derivative falls on the high frequency a :

$$-T_{D_t}T_{\bar{x}}T_{\frac{1}{1+a}}T_{1+W_\alpha}Pa = T_{\bar{x}}T_{1+W_\alpha}PM + G,$$

matching the last two terms in $\mathcal{G}_{0,1}$.

On the other hand, its contribution to the second equation is

$$-iT_{1-Y}T_{1+a}(-T_{\bar{x}}T_{\frac{1}{1+a}}T_{1+W_\alpha}Pa) = iT_{\bar{x}}Pa.$$

This is the extra Taylor coefficient term $iT_{\bar{x}}Pa$ in $\mathcal{K}_{0,2}$.

It is now convenient to compute the derivatives of the term

$$-T_{\bar{x}}\Pi(W_\alpha, X_\alpha)$$

in \tilde{w}_2 . We have, in the first equation,

$$T_{D_t}T_{\bar{x}}\Pi(W_\alpha, X_\alpha) = -2T_{\bar{x}}T_{1-\bar{Y}}\Pi(W_\alpha, R_\alpha) + G,$$

and in the second equation,

$$-iT_{1-Y}T_{1+a}T_{\bar{x}}\Pi(W_\alpha, X_\alpha) = -iT_{\bar{x}}T_{1+a}\Pi(W_\alpha, Y) + G.$$

Thus, after (G, K) -perturbative modifications, it remains to match

$$\begin{aligned} \mathcal{G}'_{0,1}(w, r) &= T_{\bar{x}}(\Pi(W_{\alpha\alpha}, 2\Re T_{1-\bar{Y}}R) - T_{1+W_\alpha}T_{1-\bar{Y}}\Pi(2\Re X_\alpha, R_\alpha)), \\ \mathcal{K}'_{0,1}(w, r) &= T_{\bar{x}}(ia + 2T_{1+W_\alpha}T_{1-\bar{Y}}T_{R_\alpha}R + \Pi(T_{1+W_\alpha}R_\alpha, 2\Re T_{1-\bar{Y}}R)), \end{aligned} \tag{5.25}$$

instead of $(\mathcal{G}_{0,1}, \mathcal{K}_{0,1})$, with correction terms

$$\begin{aligned} \tilde{w}'_2 &= T_{\bar{x}}(\Pi(W_{\alpha\alpha}, \Re X) - \Pi(W_\alpha, \Re X_\alpha)), \\ \tilde{r}'_2 &= T_{\bar{x}}(\Pi(T_{1+W_\alpha}R_\alpha, \Re X) - \Pi(\Re T_{1+W_\alpha}R, X_\alpha)). \end{aligned}$$

We substitute this in the para-linearized equation, and apply our para-Leibniz rule in Lemma 3.6. Whenever the para-material derivative or a regular derivative applies to the coefficient $T_{\bar{x}}$ (or any other low frequency coefficient) we get perturbative contributions which we can neglect. This greatly simplifies the computations below.

(1) We compute the contributions in the first equation in (5.18):

$$\begin{aligned} T_{D_t} \tilde{w}'_2 &= T_{\bar{x}}(-\Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_{\alpha\alpha}, \Re X) - \Pi(W_{\alpha\alpha}, \Re T_{1-\bar{Y}} R) \\ &\quad + \Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_\alpha, \Re X_\alpha) + \Pi(W_\alpha, \Re T_{1-\bar{Y}} R_\alpha)) + G, \\ T_{1-\bar{Y}} \partial_\alpha \tilde{r}'_2 &= T_{\bar{x}}(\Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_{\alpha\alpha}, \Re X) + \Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_\alpha, \Re X_\alpha) \\ &\quad - \Pi(T_{1-\bar{Y}} \Re T_{1+W_\alpha} R_\alpha, X_\alpha) - \Pi(T_{1-\bar{Y}} \Re T_{1+W_\alpha} R, X_{\alpha\alpha})) + G \\ &= T_{\bar{x}}(\Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_{\alpha\alpha}, \Re X) + \Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_\alpha, \Re X_\alpha) \\ &\quad - \Pi(\Re T_{1-\bar{Y}} R_\alpha, W_\alpha) - \Pi(\Re T_{1-\bar{Y}} R, W_{\alpha\alpha})) + G. \end{aligned}$$

Summing the two contributions and observing cancellations, we obtain the source terms

$$T_{\bar{x}}(\Pi(T_{1-\bar{Y}} T_{1+W_\alpha} R_\alpha, 2\Re X_\alpha) - \Pi(W_{\alpha\alpha}, 2\Re T_{1-\bar{Y}} R)) + G$$

which match the remaining terms in $\mathcal{E}_{0,1}$, recorded in $\mathcal{E}'_{0,1}$, (5.25).

(2) We compute the second equation in (5.18):

$$\begin{aligned} T_{D_t} \tilde{r}'_2 &= T_{\bar{x}}(iT_{1+a} \Pi(T_{1+W_\alpha} Y_\alpha, \Re X) - \Pi(T_{1+W_\alpha} R_\alpha, \Re T_{1-\bar{Y}} R) \\ &\quad - iT_{1+a} \Pi(\Re T_{1+W_\alpha} Y, X_\alpha) + \Pi(T_{1+W_\alpha} R, \Re T_{1-\bar{Y}} R_\alpha)) + K, \\ -iT_{1-\bar{Y}} T_{1+a} \tilde{w}'_2 &= -iT_{\bar{x}} T_{1+a} \Pi(T_{1+W_\alpha} Y_\alpha, \Re X) + iT_{\bar{x}} T_{1+a} \Pi(T_{1+W_\alpha} Y, \Re X_\alpha) + K, \\ &= -iT_{\bar{x}} T_{1+a} \Pi(T_{1+W_\alpha} Y_\alpha, \Re X) + iT_{\bar{x}} T_{1+a} \Pi(\Re T_{1+W_\alpha} Y, X_\alpha) + K. \end{aligned}$$

Summing the two contributions and observing cancellations, we obtain the source terms

$$\frac{1}{2} T_{\bar{x}} (\Pi(R, \bar{R}_\alpha) - \Pi(R_\alpha, \bar{R})) + K.$$

Summing this with the remaining terms in $\mathcal{K}_{0,1}$, recorded in $\mathcal{K}'_{0,1}$, (5.25), we obtain $\mathcal{K}_{0,2}$. ■

It remains to build a correction to match $\mathcal{K}_{0,2}$, consisting of errors in the second equation which are quadratic in R and with \bar{x} coefficients. Define

$$\begin{aligned} \tilde{w}_3 &= iT_{1+W_\alpha} (P[T_{\bar{x}} \bar{X}_\alpha X_\alpha] - iT_{\frac{1}{1+a}} P[T_{\bar{x}} \bar{R}_\alpha R]) \\ &\quad + \frac{1}{2} T_{\bar{x}} \Re T_{1+W_\alpha} (\Pi(\bar{X}_\alpha, X_\alpha) - iT_{\frac{1}{1+a}} \Pi(\bar{R}_\alpha, R)) \\ &\quad + \partial_\alpha \left(T_{T_{\bar{x}} W_\alpha} X + \frac{1}{2} T_{\bar{x}} \Pi(W_\alpha, X) \right) =: \tilde{w}_{3,1} + \tilde{w}_{3,2} + \tilde{w}_{3,3} + \tilde{w}_{3,4}, \\ \tilde{r}_3 &= iT_{1+W_\alpha} P[T_{\bar{x}} \bar{X}_\alpha R] \\ &\quad + \frac{1}{2} T_{\bar{x}} \Re T_{1+W_\alpha} \Pi(\bar{X}_\alpha, R) \\ &\quad + T_{T_{\bar{x}} W_\alpha} R + T_{T_{\bar{x}} R_\alpha} W \\ &\quad + \frac{1}{2} T_{\bar{x}} \partial_\alpha \Pi(W, R) =: \tilde{r}_{3,1} + \tilde{r}_{3,2} + \tilde{r}_{3,3} + \tilde{r}_{3,4}. \end{aligned}$$

As usual, it is easy to verify (5.15) and (5.16).

Lemma 5.16. *We have*

$$\tilde{\mathcal{G}}_3 = -\mathcal{E}_{0,2} + G = G, \quad \tilde{\mathcal{K}}_2 = -\mathcal{K}_{0,2} + K,$$

where (G, K) satisfy the bound (5.17).

Proof. It remains to match

$$\begin{aligned} \mathcal{K}_{0,2}(w, r) &= T_{\bar{x}} \left(2ia + 2T_{1+W_\alpha} T_{1-\bar{Y}} T_{R_\alpha} R + \frac{1}{2} \partial_\alpha \Pi(T_{1+W_\alpha} R, b) \right) \\ &= T_{\bar{x}} \left(2ia + \frac{1}{2} \partial_\alpha \Pi(T_{1+W_\alpha} R, T_{1-Y} \bar{R}) \right. \\ &\quad \left. + 2T_{1+W_\alpha} T_{1-\bar{Y}} T_{R_\alpha} R + \frac{1}{2} \partial_\alpha \Pi(T_{1+W_\alpha} R, T_{1-\bar{Y}} R) \right) + G \\ &=: \mathcal{K}_{0,2,1} + \mathcal{K}_{0,2,2} + \mathcal{K}_{0,2,3} + \mathcal{K}_{0,2,4}. \end{aligned}$$

The Taylor coefficient term, and more generally the quadratic $\bar{R}_\alpha R$ and $\bar{R} R_\alpha$ terms collected in $\mathcal{K}_{0,2,1}$, $\mathcal{K}_{0,2,2}$, are corrected by $(\tilde{w}_{3,1}, \tilde{r}_{3,1})$, $(\tilde{w}_{3,2}, \tilde{r}_{3,2})$, respectively. The computations are easily adapted from the Taylor coefficient corrections of Lemma 5.12, so we omit them here.

The remaining terms in $\mathcal{K}_{0,2,3}$, $\mathcal{K}_{0,2,4}$ are of the form $R_\alpha R$, and will be corrected by $(\tilde{w}_{3,3}, \tilde{r}_{3,3})$, $(\tilde{w}_{3,4}, \tilde{r}_{3,4})$, respectively. In the following, we demonstrate only the computations for the high–low interactions $\mathcal{K}_{0,2,3}$, since the frequency-balanced interactions in $\mathcal{K}_{0,2,4}$ are similar. Thus, we are reduced to correcting

$$\mathcal{K}_{0,2,3} = 2T_{\bar{x}} T_{1+W_\alpha} T_{1-\bar{Y}} T_{R_\alpha} R$$

with the corrections

$$\begin{aligned} \tilde{w}_{3,3} &= \partial_\alpha T_{T_{\bar{x}} W_\alpha} X, \\ \tilde{r}_{3,3} &= T_{T_{\bar{x}} W_\alpha} R + T_{T_{\bar{x}} R_\alpha} W. \end{aligned}$$

(1) We begin with the first equation, using as usual the para-Leibniz rule and the appropriate para-material derivative formulas, and observing that cases where the derivative falls on \bar{x} are perturbative:

$$\begin{aligned} T_{D_t} \tilde{w}_{3,3} &= -\partial_\alpha (T_{T_{\bar{x}}(1-\bar{Y}) R_\alpha} W + T_{T_{\bar{x}} W_\alpha} T_{1-\bar{Y}} R) + G, \\ T_{1-\bar{Y}} \partial_\alpha \tilde{r}_{3,3} &= \partial_\alpha (T_{1-\bar{Y}} T_{T_{\bar{x}} R_\alpha} W + T_{1-\bar{Y}} T_{T_{\bar{x}} W_\alpha} R) + G, \end{aligned}$$

so that we have no nonperturbative new contributions.

(2) For the second equation,

$$\begin{aligned} T_{D_t} \tilde{r}_{3,3} &= -T_{T_{\bar{x}} T_{1-\bar{Y}} T_{1+W_\alpha} R_\alpha} R + iT_{T_{\bar{x}} W_\alpha} T_{1+a} Y \\ &\quad + iT_{T_{\bar{x}} T_{1+a} Y_\alpha} W - T_{T_{\bar{x}} R_\alpha} T_{1-\bar{Y}} T_{1+W_\alpha} R + G, \\ -iT_{1-Y} T_{1+a} \tilde{w}_{3,3} &= -iT_{1+a} (T_{T_{\bar{x}} W_\alpha} Y + T_{T_{\bar{x}} Y_\alpha} W) + G. \end{aligned}$$

We see the remaining terms, after perturbative modifications and cancellations, match $\mathcal{K}_{0,2,3}$. ■

6. Energy estimates for the full equation

The main goal of this section is to prove energy estimates for the (\mathbf{W}, R) equation in $\dot{\mathcal{H}}^s$ for $s \geq 0$ with an $A_{\frac{1}{4}}^2$ constant, as stated in Theorem 2.

We are not only interested in having a good energy estimate, but also in understanding the structure of the differentiated evolution (1.2). Thus, as part of the proof, we will also prove the renormalization result in Theorem 5. This will be useful in the next paper in our series, where we prove Strichartz estimates at low regularity.

We remark that, since (\mathbf{W}, R) solves the linearized equation, the bounds for the linearized equation in the previous section yield the case $s = \frac{1}{4}$. We are interested in all s but particularly the case $s \leq \frac{3}{4}$. The estimates for the linearized equation will not work for $s \neq \frac{1}{4}$, but here we can take advantage of the fact that $(w, r) = (\mathbf{W}, R)$ and balance things better.

6.1. The main steps of the argument

To begin the analysis, it is useful to recast (1.2) in a paradifferential form with a source term, based on the paradifferential equation (4.7). Within the source term, we will peel off perturbative contributions, denoted by (G, K) for the remainder of the section. These are nonlinear expressions in (\mathbf{W}, R) with the property that they satisfy favorable balanced cubic bounds of the form

$$\|(G, K)\|_{\dot{\mathcal{H}}^s} \lesssim_A A_{\frac{1}{4}}^2 \|(\mathbf{W}, R)\|_{\dot{\mathcal{H}}^s}. \tag{6.1}$$

The remaining source terms are of two types:

- (a) quadratic terms in (\mathbf{W}, R) ;
- (b) unbalanced cubic terms, which can also be viewed as a quadratic expression in (\mathbf{W}, R) but with a lower frequency coefficient which depends only on an undifferentiated \mathbf{W} .

Implicitly, a similar analysis was carried out in [15] by further differentiating the (\mathbf{W}, R) equation. The difference is that in [15] all of the cubic source terms went into the perturbative box, whereas here we make a finer distinction.

We state the outcome of this computation in the following.

Lemma 6.1. *Equation (1.2) can be rewritten as a paradifferential equation for the variables $(\hat{w}, \hat{r}) = (\mathbf{W}, R)$, of the form*

$$\begin{cases} T_{D_t} \hat{w} + T_{b_\alpha} \hat{w} + \partial_\alpha T_{1-\bar{Y}} T_{1+W_\alpha} \hat{r} = \mathcal{G}(\mathbf{W}, R) + G, \\ T_{D_t} \hat{r} + T_{b_\alpha} \hat{r} - iT_{(1-Y)^2} T_{1+a} \hat{w} + T_M \hat{r} = \mathcal{K}(\mathbf{W}, R) + K, \end{cases} \tag{6.2}$$

where

$$\begin{aligned} \mathcal{G} &= -\partial_\alpha [\Pi(\mathbf{W}, T_{1-\bar{Y}} R) + \Pi(\mathbf{W}, T_{1-Y} \bar{R}) + \Pi(\bar{Y}, T_{1+W} R)], \\ \mathcal{K} &= -T_{1-\bar{Y}} \Pi(R_\alpha, R) - T_{1-Y} \Pi(R_\alpha, \bar{R}) - iT_{1-Y} T_{1+a} \Pi(Y, \mathbf{W}) \\ &\quad - T_{1-Y} \Pi(\bar{R}_\alpha, R). \end{aligned}$$

As in the previous section, here we include an implicit projection P in Π . We note that in \mathcal{E} one can equivalently take all the paraproducts out in front. The above lemma is proved in Section 6.2.

We will divide the analysis of equation (6.2) into two steps. We will first prove an \mathcal{H}^s well-posedness result for the paradifferential equation (4.7). Fortunately, we do not need to do this from scratch, and instead we can rely on a similar result proved in Proposition 5.1 in the course of the analysis of the linearized equation. For this reason, we state the outcome of this step as follows.

Proposition 6.2. *For the paradifferential equation (6.2) above, we have that*

- (a) *there exists a bounded linear transformation (independent of s) reducing (4.7) to the linearized paradifferential equation (4.5). Precisely, given (\hat{w}, \hat{r}) satisfying (4.7), there exists (w, r) satisfying (5.1) such that for any $s \in \mathbb{R}$ we have*

- (i) *invertibility:*

$$\|(w_\alpha, r_\alpha) - (\hat{w}, \hat{r})\|_{\mathcal{H}^s} \lesssim_A A \|(w, r)\|_{\mathcal{H}^s},$$

- (ii) *perturbative source term:*

$$\|(G, K)\|_{\mathcal{H}^s} \lesssim_A A \frac{1}{4} \|(w, r)\|_{\mathcal{H}^s}.$$

- (b) *As a consequence, Proposition 5.1 holds with (4.7) in place of (4.5).*

This result is proved in Section 6.3; in particular, it yields a good \mathcal{H}^s energy functional for (4.7).

Our second step is to interpret the differentiated system (1.2), expressed in the form (6.2), perturbatively based on the paradifferential equation (4.7). However, the source terms $(\mathcal{G}, \mathcal{K})$ are not directly perturbative. Instead, we will show that there exists a favorable normal form transformation which largely eliminates the source terms.

Proposition 6.3. *Given (\mathbf{W}, R) satisfying (1.2), there exist modified normal form variables $(\mathbf{W}_{\text{NF}}, R_{\text{NF}})$ satisfying an equation of the form*

$$\begin{cases} T_{D_t} \mathbf{W}_{\text{NF}} + T_{b_\alpha} \mathbf{W}_{\text{NF}} + \partial_\alpha T_{1-\bar{Y}} T_{1+W_\alpha} R_{\text{NF}} = \tilde{\mathcal{E}}(\mathbf{W}, R), \\ T_{D_t} R_{\text{NF}} + T_{b_\alpha} R_{\text{NF}} - iT_{(1-Y)^2} T_{1+a} \mathbf{W}_{\text{NF}} + T_M R_{\text{NF}} = \tilde{\mathcal{K}}(\mathbf{W}, R), \end{cases}$$

such that for any $s \geq 0$ we have

- (i) *invertibility:*

$$\|(\mathbf{W}_{\text{NF}}, R_{\text{NF}}) - (\mathbf{W}, R)\|_{\mathcal{H}^s} \lesssim_A A \|(\mathbf{W}, R)\|_{\mathcal{H}^s},$$

- (ii) *perturbative source term:*

$$\|\tilde{\mathcal{E}}(\mathbf{W}, R), \tilde{\mathcal{K}}(\mathbf{W}, R)\|_{\mathcal{H}^s} \lesssim_A A \frac{1}{4} \|(\mathbf{W}, R)\|_{\mathcal{H}^s}.$$

This result is proved in Section 6.4.

6.2. The paradifferential form of the differentiated equation

Our goal here is to prove Lemma 6.1. We successively consider all terms in the two equations in (1.2), which for convenience we recall here, written in a slightly more convenient algebraic fashion

$$\begin{cases} D_t \mathbf{W} + (1 - \bar{Y})(1 + \mathbf{W})R_\alpha = (1 + \mathbf{W})M, \\ D_t R = i(1 - (1 - Y)(1 + a)). \end{cases}$$

We will use this equation in its projected form.

Along the way, we place perturbative terms in (G, K) . As a general rule, all terms we place in (G, K) are cubic and balanced, in the sense that the two highest frequencies are comparable and the low frequency variable is fully differentiated.

(1) For the first term in the first equation, expanding b and discarding balanced terms, we have

$$\begin{aligned} PD_t \mathbf{W} &= T_{D_t} \mathbf{W} + T_{W_\alpha} P b + \Pi(\mathbf{W}_\alpha, b) \\ &= T_{D_t} \mathbf{W} + T_{W_\alpha} T_{1-\bar{Y}} R + \Pi(\mathbf{W}_\alpha, T_{1-\bar{Y}} R + T_{1-Y} \bar{R}) + G. \end{aligned}$$

For the second term in the first equation we use a paraproduct expansion discarding all balanced terms,

$$\begin{aligned} P[(1 + \mathbf{W})(1 - \bar{Y})R_\alpha] &= T_{1+\mathbf{W}} T_{1-\bar{Y}} R_\alpha - P T_{1+\mathbf{W}} T_{R_\alpha} \bar{Y} + T_{1-\bar{Y}} T_{R_\alpha} \mathbf{W} \\ &\quad - T_{1+\mathbf{W}} \Pi(\bar{Y}, R_\alpha) + T_{1-\bar{Y}} \Pi(\mathbf{W}, R_\alpha) + G. \end{aligned}$$

Finally, for the source term in the first equation we use (1.3) and Lemma 3.5 to get

$$\begin{aligned} P[(1 + \mathbf{W})M] &= T_{1+\mathbf{W}} P M + T_M \mathbf{W} + G = T_{1+\mathbf{W}} P [\bar{Y}_\alpha R - \bar{R}_\alpha Y] + T_M \mathbf{W} + G \\ &= T_{1+\mathbf{W}} (T_{\bar{Y}_\alpha} R - T_{\bar{R}_\alpha} Y) + T_{1+\mathbf{W}} [\Pi(\bar{Y}_\alpha, R) - \Pi(\bar{R}_\alpha, Y)] + G. \end{aligned}$$

We combine these terms, using Lemmas 3.2 and 2.5 to write

$$T_{1-\bar{Y}} T_{R_\alpha} \mathbf{W} + T_{1+\mathbf{W}} T_{\bar{R}_\alpha} Y = T_{1-\bar{Y}} T_{R_\alpha} \mathbf{W} + T_{1-Y} T_{\bar{R}_\alpha} \mathbf{W} - T_M \mathbf{W} + G = T_{b_\alpha} \mathbf{W} + G,$$

in order to obtain the first equation in (6.2).

(2) We now consider the two terms in the second equation, which we expand in a similar fashion:

$$\begin{aligned} PD_t R &= T_{D_t} R + T_{R_\alpha} P b + \Pi(R_\alpha, b) \\ &= T_{D_t} R + T_{R_\alpha} T_{1-\bar{Y}} R + \Pi(R_\alpha, T_{1-\bar{Y}} R + T_{1-Y} \bar{R}) + K \end{aligned}$$

respectively, using Lemma 3.3,

$$\begin{aligned} P \left[\frac{\mathbf{W} - a}{1 + \mathbf{W}} \right] &= P [1 - (1 + a)(1 - Y)] \\ &= T_{1+a} Y - T_{1-Y} P a + \Pi(a, Y) \\ &= T_{1+a} Y - i T_{1-Y} [\bar{R}_\alpha R] + K \\ &= T_{1+a} Y - i T_{1-Y} T_{\bar{R}_\alpha} R - i T_{1-Y} \Pi(\bar{R}_\alpha, R) + K. \end{aligned}$$

Just as in the case of the first equation, we recombine

$$T_{R_\alpha} T_{1-\bar{Y}} R + T_{1-Y} T_{\bar{R}_\alpha} R = T_{b_\alpha+M} R + K,$$

and use Lemma 3.2 to write

$$T_{1+a} Y = T_{(1-Y)^2} T_{1+a} \mathbf{W} - T_{1-Y} T_{1+a} \Pi(Y, \mathbf{W}) + K,$$

obtaining the second equation in (6.2).

6.3. Well-posedness for the paradifferential flow

In this section we prove the paradifferential well-posedness result in Proposition 6.2. One could also do this directly, but instead it is more efficient to transfer this result from a similar result in Proposition 5.1.

Consider a solution (\hat{w}, \hat{r}) to (4.7). Motivated by (4.8), define

$$(w, r) = (\partial_\alpha^{-1} \hat{w}, \partial_\alpha^{-1} T_{1+\mathbf{W}} \hat{r} - \partial_\alpha^{-1} T_{R_\alpha} \partial_\alpha^{-1} \hat{w}).$$

It is easily seen that we have the norm equivalence

$$\|(w, r)\|_{\dot{\mathcal{H}}^{s+1}} \lesssim_A \|(\hat{w}, \hat{r})\|_{\dot{\mathcal{H}}^s}.$$

We will prove that the equation for (w, r) is a perturbation of equation (4.5), which will allow us to establish an equivalence between the $\dot{\mathcal{H}}^s$ well-posedness of (4.7) and the $\dot{\mathcal{H}}^{s+1}$ well-posedness of (4.5).

We insert (w, r) into (4.5) and compute the corresponding source terms. For the first equation of the system for (w, r) we have full cancellation,

$$\partial_\alpha(T_{D_t} w + T_{1-\bar{Y}} \partial_\alpha r + T_{1-\bar{Y}} T_{R_\alpha} w) = T_{D_t} \hat{w} + T_{b_\alpha} \hat{w} + \partial_\alpha T_{1-\bar{Y}} T_{1+W_\alpha} \hat{r} = 0.$$

For the second equation we use (1.2) and (4.7) to compute material derivatives, Lemma 3.6 to distribute para-material derivatives, and Lemma 2.4 to commute paraproducts. Denoting by K perturbative terms, i.e. which satisfy

$$\|K\|_{\dot{H}^{s+\frac{1}{2}}} \lesssim_A A_{\frac{1}{4}}^2 \|(w, r)\|_{\dot{\mathcal{H}}^{s+1}},$$

we successively evaluate the expression

$$\mathcal{K} = \partial_\alpha(T_{D_t} r - iT_{1-Y} T_{1+a} w)$$

as follows:

$$\begin{aligned} \mathcal{K} &= T_{D_t} (T_{1+\mathbf{W}} \hat{r} - T_{R_\alpha} \partial_\alpha^{-1} \hat{w}) + T_{b_\alpha} (T_{1+\mathbf{W}} \hat{r} - T_{R_\alpha} \partial_\alpha^{-1} \hat{w}) - i \partial_\alpha T_{1-Y} T_{1+a} w \\ &= (T_{D_t} \mathbf{W} \hat{r} + T_{1+\mathbf{W}} T_{D_t} \hat{r} - T_{\partial_\alpha D_t R} w - T_{R_\alpha} T_{D_t} \partial_\alpha^{-1} \hat{w}) + T_{1+\mathbf{W}} T_{b_\alpha} \hat{r} \\ &\quad + iT_{1+a} T_{Y_\alpha} w - iT_{1-Y} T_{1+a} w_\alpha + K \end{aligned}$$

$$\begin{aligned}
 &= T_{1+\mathbf{W}}(T_{D_t} + T_{b_\alpha})\hat{r} - T_{R_\alpha}\partial^{-1}(T_{D_t} + T_{b_\alpha})\hat{w} + T_{D_t}\mathbf{W}\hat{r} - T_{\partial_\alpha D_t R}w \\
 &\quad + iT_{1+a}TY_\alpha w - iT_{1-Y}T_{1+a}w_\alpha + K \\
 &= iT_{1+\mathbf{W}}T_{(1+a)(1-Y)^2}\hat{w} - T_{1+\mathbf{W}}TM\hat{r} + T_{R_\alpha}T_{1-\bar{Y}}T_{1+\mathbf{W}}\hat{r} - T_{(1+\mathbf{W})(1-\bar{Y})R_\alpha}\hat{r} \\
 &\quad + T_{(1+\mathbf{W})M}\hat{r} + iT_{\partial_\alpha(1+a)(1-Y)}w + iT_{1+a}TY_\alpha w - iT_{1-Y}T_{1+a}\hat{w} + K \\
 &= K.
 \end{aligned}$$

Here we have harmlessly replaced the para-material derivatives of \mathbf{W} and R_α with their full material derivatives, and also we have discarded the $\partial_\alpha a$ term using (3.2). Finally, at the last step we have manipulated paraproducts using Lemmas 2.4, 2.5, 2.6.

Thus, the $\dot{\mathcal{H}}^{s+1}$ well-posedness of (4.5) given by Proposition 5.1 implies the $\dot{\mathcal{H}}^s$ well-posedness of (4.7).

6.4. The normal form transformation

In this section we construct the normal form transformation in Proposition 6.3. We recall the nonperturbative part of the source term,

$$\begin{aligned}
 \mathcal{G} &= -\partial_\alpha[\Pi(\mathbf{W}, T_{1-\bar{Y}}R) + \Pi(\mathbf{W}, T_{1-Y}\bar{R}) + \Pi(\bar{Y}, T_{1+\mathbf{W}}R)], \\
 \mathcal{K} &= -T_{1-\bar{Y}}\Pi(R_\alpha, R) - T_{1-Y}\Pi(R_\alpha, \bar{R}) - iT_{1-Y}T_{1+a}\Pi(Y, \mathbf{W}) - T_{1-Y}\Pi(\bar{R}_\alpha, R),
 \end{aligned}$$

where one could harmlessly commute outside the paraproducts in the first expression.

Then we define the correction

$$\begin{aligned}
 \tilde{\mathbf{W}} &= -\partial_\alpha\Pi(\mathbf{W}, 2\Re X), \\
 \tilde{R} &= -\Pi(R_\alpha, 2\Re X) - \Pi(T_{1-\bar{Y}}\bar{\mathbf{W}}, R).
 \end{aligned}$$

For a partial verification of these expressions, we remark that the quadratic part of this correction coincides with the balanced part of the derivative of the correction in (1.12), after switching to the good variables (\mathbf{W}, R) .

We now insert these corrections into equation (4.7), and verify that the generated source terms cancel $(\mathcal{G}, \mathcal{K})$ modulo perturbative terms (i.e. which satisfy (6.1)).

Note that for both equations, the contribution of the T_{b_α} term is perturbative. Also, in the first equation, the undifferentiated \hat{r} terms are perturbative. Also recall that we use Lemma 3.6 to distribute the para-material derivative, and Lemma 3.8 to compute the para-material derivative of X .

We compute the contributions to the first equation of (4.7):

$$\begin{aligned}
 -T_{D_t}\partial_\alpha\Pi(\mathbf{W}, 2\Re X) &= \partial_\alpha\Pi(T_{(1-\bar{Y})(1+\mathbf{W})}R_\alpha, 2\Re X) \\
 &\quad + \Pi(\mathbf{W}, T_{1-\bar{Y}}R + T_{1-Y}\bar{R}) + G, \\
 -T_{1-\bar{Y}}T_{1+\mathbf{W}}\partial_\alpha\Pi(R_\alpha, 2\Re X) &= -\partial_\alpha\Pi(T_{(1-\bar{Y})(1+\mathbf{W})}R_\alpha, 2\Re X) + G, \\
 -T_{1-\bar{Y}}T_{1+\mathbf{W}}\partial_\alpha\Pi(T_{1-\bar{Y}}\bar{\mathbf{W}}, R) &= -\partial_\alpha\Pi(\bar{Y}, T_{1+\mathbf{W}}R) + G.
 \end{aligned}$$

After cancellations, only the second terms of the first contribution and the last contribution remain, modulo perturbative G terms, matching \mathcal{G} .

Similarly, for the second equation of (4.7) we have

$$\begin{aligned} -T_{D_t} \Pi(R_\alpha, 2\Re X) &= -i \Pi(T_{1+a} Y_\alpha, 2\Re X) \\ &\quad + \Pi(R_\alpha, (T_{1-\bar{Y}} R + T_{1-Y} \bar{R})) + K, \\ -T_{D_t} \Pi(T_{1-\bar{Y}} \bar{W}, R) &= T_{1-Y} \Pi(\bar{R}_\alpha, R) - iT_{1+a} \Pi(T_{1-\bar{Y}} \bar{W}, Y) + K, \\ iT_{(1-Y)^2} T_{1+a} \partial_\alpha \Pi(W, 2\Re X) &= i \Pi(T_{1+a} Y_\alpha, 2\Re X) \\ &\quad + iT_{1+a} (T_{1-Y} \Pi(W, Y) + \Pi(Y, \bar{X}_\alpha)) + K, \\ T_M \tilde{R} &= K. \end{aligned}$$

After cancellations, we match \mathcal{K} modulo perturbative K terms. Setting

$$(\mathbf{W}_{\text{NF}}, R_{\text{NF}}) = (\mathbf{W} + \tilde{W}, R + \tilde{R})$$

yields the desired normal form variables.

7. Local well-posedness

In this section we prove the low regularity local well-posedness result stated in Theorem 3. The result from our previous paper [15] asserts that local well-posedness holds for more regular data. We use those more regular solutions to construct the rough solutions by a scheme which is similar to the one used there. Precisely, we truncate the data in frequency and then move through a continuous family of solutions, thereby estimating only a solution for a linearized equation at each step. We implement this strategy directly on the equations (1.2) which are in terms of the diagonal variables (\mathbf{W}, R) . We begin with the main well-posedness result in [15].

Theorem 8. *Let $n \geq 1$.*

- (a) *(Well-posedness) The problem (1.1) is locally well posed for initial data (\mathbf{W}_0, R_0) in \mathcal{H}^n .*
- (b) *(Lifespan) There exists $T = T(\|\mathbf{W}_0, R_0\|_{\mathcal{H}^1}, \|Y_0\|_{L^\infty})$ so that the above solutions are well defined in $[0, T]$, with uniform bounds.*

Our goal will be to obtain an \mathcal{H}^s version of the above theorem with $s = \frac{3}{4}$. We recall 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 $C(\mathcal{H}^n)$ together with continuous dependence on the data. On the other hand, if $n = 1$ then we only produce rough solutions $C(\mathcal{H}^1)$ as the unique strong limit of smooth solutions, again with continuous dependence on the data; however, for $n = 1$ we do not establish a direct uniqueness result.

7.1. $\dot{\mathcal{H}}^s$ bounds for regular solutions

The solutions in the last theorem have a lifespan which depends on the \mathcal{H}^1 size of data. Here we prove that in effect the lifespan depends only on the $\mathcal{H}^{\frac{3}{4}}$ size of data, and that we have uniform bounds for as long as the $\mathcal{H}^{\frac{3}{4}}$ size of the solutions is controlled.

Precisely, suppose that for some $n \geq 2$ we have an \mathcal{H}^n solution (\mathbf{W}, R) which satisfies the bounds

$$\|(\mathbf{W}, R)(0)\|_{\mathcal{H}^{\frac{3}{4}}} < \mathcal{M}_0 \ll 1.$$

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

$$\|(\mathbf{W}, R)\|_{L^\infty(0, T; \mathcal{H}^{\frac{3}{4}})} < \mathcal{M}(\mathcal{M}_0), \tag{7.1}$$

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

$$\begin{aligned} \|(\mathbf{W}, R)\|_{L^\infty(0, T; \mathcal{H}^n)} &\leq C(\mathcal{M}_0)\|(\mathbf{W}, R)(0)\|_{\mathcal{H}^n}, \\ \|(\mathbf{W}, R)\|_{L^\infty(0, T; \dot{\mathcal{H}}^n)} &\leq C(\mathcal{M}_0)\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}^n}. \end{aligned}$$

To prove this, we begin by making the bootstrap assumption

$$\|(\mathbf{W}, R)\|_{L^\infty(0, T; \mathcal{H}^{\frac{3}{4}})} < 2\mathcal{M}. \tag{7.2}$$

We will show that for a suitable choice $\mathcal{M}(\mathcal{M}_0)$ depending only on \mathcal{M}_0 , we can improve this to (7.1), provided that $T < T(\mathcal{M}_0)$.

From the bootstrap assumption (7.2) we bound the control parameter

$$A_{\frac{1}{4}} \lesssim \mathcal{M}.$$

By applying the energy estimates obtained in Theorem 2, and Gronwall’s inequality to (\mathbf{W}, R) ,

$$\|(\mathbf{W}, R)(t)\|_{\mathcal{H}^{\frac{3}{4}}} \lesssim e^{Ct} \|(\mathbf{W}, R)(0)\|_{\mathcal{H}^{\frac{3}{4}}} \lesssim e^{Ct} \mathcal{M}_0, \quad C = C(\mathcal{M}).$$

It is here that our smallness assumption on A is used, as it guarantees that our energies E_s are comparable with the $\dot{\mathcal{H}}^s$ norm of the solution.

The above estimate allows us to first choose \mathcal{M} large enough and then T such that the bound in (7.1) holds for $t \in [0, T]$.

At the same time, applying the energy estimates proven in Theorem 2 and Gronwall’s inequality for higher norms of the pair (\mathbf{W}, R) we get

$$\|(\mathbf{W}, R)(t)\|_{\mathcal{H}^n} \lesssim e^{Ct} \|(\mathbf{W}, R)(0)\|_{\mathcal{H}^n},$$

which allow us to propagate (\mathbf{W}, R) as classical solutions up to time T using Theorem 8.

7.2. Rough solutions

Here we construct solutions for data in $\mathcal{H}^{\frac{3}{4}}$ as unique limits of smooth solutions. Given an $\mathcal{H}^{\frac{3}{4}}$ 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 $\mathcal{H}^{\frac{3}{4}}$ bounds. Further, they are smooth and have a smooth dependence on k .

To better understand the evolution of the $\mathcal{H}^{\frac{3}{4}}$ norm of the solution it is convenient to use the language of frequency envelopes.

Definition 7.1. A sequence $c_k \in \ell^2$ is called an \mathcal{H}^s frequency envelope for $(\mathbf{W}, R) \in \mathcal{H}^s$ 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 $\|P_k(\mathbf{W}, R)\|_{\mathcal{H}^s} \leq c_k$.

Consider a frequency envelope c_k for the initial data (\mathbf{W}_0, R_0) in $\mathcal{H}^{\frac{3}{4}}$. Then for the regularized data we have

$$\|(\mathbf{W}_0^k, R_0^k)\|_{\mathcal{H}^n} \lesssim c_k 2^{(n-\frac{3}{4})k}, \quad n \geq 1.$$

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

$$\|(\mathbf{W}^k, R^k)\|_{\mathcal{H}^n} \lesssim c_k 2^{(n-\frac{3}{4})k}, \quad n \geq 1. \tag{7.3}$$

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 (4.2) arises from the linearization of the (W, Q) system in (1.1) rather than the differentiated (\mathbf{W}, R) system in (1.2). Assuming that (W^k, Q^k) were also defined, we formally denote

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

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 compute

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

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 and therefore the products are at frequency 2^k and higher. To take advantage of the bounds in Theorem 7 for the linearized equation, we need a $\mathcal{H}^{\frac{1}{4}}$ bound for $(w^k(0), r^k(0))$, namely

$$\|(w^k(0), r^k(0))\|_{\mathcal{H}^{\frac{1}{4}}} \lesssim c_k 2^{-\frac{3}{2}k}.$$

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 $\mathcal{H}^{\frac{3}{4}}$ bound for (\mathbf{W}^k, R^k) , Theorem 7 shows that in $[0, T]$ we have the uniform estimate

$$\|(w^k, r^k)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}} \lesssim c_k 2^{-\frac{3}{2}k}.$$

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

$$\|P_{\leq k}(\partial_k \mathbf{W}^k, \partial_k R^k)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}} \lesssim c_k 2^{-\frac{1}{2}k}. \tag{7.4}$$

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 (7.4) over unit k intervals and use it to estimate the differences. Combining the result with (7.3) we obtain

$$\begin{aligned} \|(\mathbf{W}^{k+1} - \mathbf{W}^k, R^{k+1} - R^k)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}} &\lesssim c_k 2^{-\frac{1}{2}k}, \\ \|\partial_\alpha^2(\mathbf{W}^{k+1} - \mathbf{W}^k, R^{k+1} - R^k)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}} &\lesssim c_k 2^{\frac{3}{2}k}. \end{aligned} \tag{7.5}$$

Summing with respect to k it follows that the sequence (\mathbf{W}^k, R^k) converges uniformly in $\dot{\mathcal{H}}^{\frac{3}{4}} \cap \dot{\mathcal{H}}^{\frac{1}{4}}$ in the time interval $[0, T]$ to a solution (\mathbf{W}, R) . As the sequence (\mathbf{W}^k, R^k) is uniformly bounded in $\mathcal{H}^{\frac{3}{4}}$, it follows that we also have $(\mathbf{W}, R) \in \mathcal{H}^{\frac{3}{4}}$. Furthermore, it is easily seen that the solution (\mathbf{W}, R) inherits the frequency envelope bounds from the data.

For the continuous dependence on the initial data argument, we need to show that the sequence of solutions (\mathbf{W}^k, R^k) converges to (\mathbf{W}, R) in the nonhomogeneous $\mathcal{H}^{\frac{3}{4}}$ topology. For the computations below, we note that the solutions (\mathbf{W}^k, R^k) and (\mathbf{W}, R) have uniformly bounded control parameters $A, A_{\frac{1}{4}}$, and A^\sharp , so we will use this notation without any k indices. In addition, we introduce

$$A^{\sharp\sharp} = \|(\mathbf{W}_0, R_0)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}},$$

which uniformly controls the $\dot{\mathcal{H}}^{\frac{3}{4}}$ norm of all the solutions (\mathbf{W}^k, R^k) and (\mathbf{W}, R) . We also introduce the notation

$$\delta \mathbf{W}^k := \mathbf{W}^k - \mathbf{W}, \quad \delta R^k := R^k - R.$$

From (7.5) we get the $\dot{\mathcal{H}}^{\frac{1}{4}}$ convergence of the approximating sequence

$$\|(\delta \mathbf{W}^k, \delta R^k)\|_{\dot{\mathcal{H}}^{\frac{1}{4}}} \lesssim 2^{-\frac{k}{2}}. \tag{7.6}$$

We will use this, together with the equations (1.2) for (\mathbf{W}^k, R^k) and (\mathbf{W}, R) , in order to supplement this with a low frequency bound

$$\|(\delta\mathbf{W}^k, \delta R^k)\|_{\dot{\mathcal{H}}^0} \lesssim_{A^\#, A^{\#\#}, T} 2^{-\frac{k}{4}}. \tag{7.7}$$

By construction this holds at the initial time, so we need to propagate it.

We first claim that from (7.6) we obtain the following difference bounds:

$$\|\delta Y^k\|_{\dot{H}^{\frac{1}{4}}} + \|\delta a^k\|_{\dot{H}^{\frac{1}{4}}} + \|\delta b^k\|_{\dot{H}^{\frac{1}{4}}} + \|\delta M^k\|_{\dot{H}^{-\frac{1}{4}}} \lesssim_{A^\#} 2^{-\frac{k}{2}}, \tag{7.8}$$

where a^k and b^k are the corresponding a and b where (\mathbf{W}, R) is replaced with (\mathbf{W}^k, R^k) . These are all relatively straightforward balanced L^2 bounds, whose proofs are left for the reader. The bounds for $\delta\mathbf{W}^k$ and δb^k are too low in regularity for our purposes, but we can interpolate with the uniform $\dot{\mathcal{H}}^{\frac{3}{4}}$ bounds to get

$$\|\delta b^k\|_{\dot{H}^{\frac{1}{2}} \cap L^\infty} + \|\delta\mathbf{W}^k\|_{\dot{H}^{\frac{1}{2}} \cap L^\infty} \lesssim_{A^\#, A^{\#\#}} 2^{-\frac{k}{4}}. \tag{7.9}$$

Subtracting the equations (1.2) for (\mathbf{W}^k, R^k) and (\mathbf{W}, R) we obtain

$$\begin{cases} D_t \delta \mathbf{W} = \delta b^k \mathbf{W}_\alpha^k - (1 + \mathbf{W})(1 - \bar{Y}) \delta R_\alpha^k + (1 - \bar{Y}) R_\alpha^k \delta \mathbf{W}^k - (1 + \mathbf{W}^k) R_\alpha^k \delta \bar{Y}^k \\ \quad + (1 + \mathbf{W}^k) \delta M^k + M \delta \mathbf{W}^k, \\ D_t \delta R^k = \delta b^k R_\alpha^k - i(1 - Y) \delta a^k + i(1 + a^k) \delta Y^k. \end{cases}$$

Here we use the relations (7.8), (7.9) to estimate all the terms on the right in $\dot{\mathcal{H}}^{-\frac{1}{4}}$, obtaining

$$\|D_t \delta \mathbf{W}^k\|_{\dot{H}^{-\frac{1}{4}}} + \|D_t \delta R^k\|_{\dot{H}^{\frac{1}{4}}} \lesssim_{A^\#, A^{\#\#}} 2^{-\frac{k}{4}}$$

with $D_t = \partial_t + b \partial_\alpha$. In the last bound we can harmlessly convert D_t to T_{D_t} modulo errors that can be placed on the right,

$$\|T_{D_t} \delta \mathbf{W}^k\|_{\dot{H}^{-\frac{1}{4}}} + \|T_{D_t} \delta R^k\|_{\dot{H}^{\frac{1}{4}}} \lesssim_{A^\#, A^{\#\#}} 2^{-\frac{k}{4}}.$$

These two bounds allow us to compute energy estimates for $\delta\mathbf{W}_k$ and δR_k as follows:

$$\begin{aligned} \frac{d}{dt} \|\delta\mathbf{W}^k\|_{L^2}^2 &= \int T_{b_\alpha} \delta \mathbf{W}^k \cdot \delta \bar{\mathbf{W}}^k \, d\alpha + 2\Re \int T_{D_t} \delta \mathbf{W}^k \cdot \delta \bar{\mathbf{W}}^k \, d\alpha \\ &\lesssim_A \|\delta\mathbf{W}^k\|_{\dot{H}^{\frac{1}{4}}}^2 + \|\delta\mathbf{W}^k\|_{\dot{H}^{\frac{1}{4}}} \|T_{D_t} \delta \mathbf{W}^k\|_{\dot{H}^{-\frac{1}{4}}} \\ &\lesssim_{A^\#, A^{\#\#}} 2^{-\frac{k}{4}}, \end{aligned}$$

and a nearly identical computation for δR^k . This yields the desired bound (7.7) for the low frequencies of the difference.

Now we return to the last step of the proof of the well-posedness result, which is the continuous dependence on the initial data in the $\dot{\mathcal{H}}^{\frac{3}{4}}$ topology. The frequency envelope

bounds are very useful in this proof. This is standard, but we briefly outline the argument. Suppose that $(\mathbf{W}_j, R_j)(0) \in \mathcal{H}^{\frac{3}{4}}$ and $(\mathbf{W}_j, R_j)(0) - (\mathbf{W}, R)(0) \rightarrow 0$ in $\mathcal{H}^{\frac{3}{4}}$. 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

$$(\mathbf{W}_j^k, R_j^k) - (\mathbf{W}^k, R^k) \rightarrow 0 \quad \text{in } \mathcal{H}^n. \tag{7.10}$$

On the other hand, from the $\mathcal{H}^{\frac{3}{4}}$ data convergence we get

$$(\mathbf{W}_j^k, R_j^k)(0) - (\mathbf{W}_j, R_j)(0) \rightarrow 0 \quad \text{in } \mathcal{H}^{\frac{3}{4}} \text{ uniformly in } j.$$

Then the above frequency envelope analysis, shows that

$$(\mathbf{W}_j^k, R_j^k) - (\mathbf{W}_j, R_j) \rightarrow 0 \quad \text{in } \mathcal{H}^{\frac{3}{4}} \text{ uniformly in } j.$$

Hence we can let k go to infinity in (7.10) and conclude that

$$(\mathbf{W}_j, R_j) - (\mathbf{W}, R) \rightarrow 0 \quad \text{in } \mathcal{H}^{\frac{3}{4}}.$$

7.3. Enhanced cubic lifespan bounds

In this section we prove Theorem 4. Given initial data (\mathbf{W}, R) for (1.2) satisfying the conditions in the theorem, namely

$$\|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}} \leq \varepsilon, \quad \|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}^{\frac{1}{2}-\delta}}^{\frac{1}{4}} \|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}}^{\delta} \leq c_0^{\frac{1}{4}+\delta},$$

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

$$\begin{aligned} \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}} &\leq C\varepsilon, \\ \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{1}{2}-\delta}}^{\frac{1}{4}} \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}}^{\delta} &\leq (Cc_0)^{\frac{1}{4}+\delta}, \quad t \in [0, T], \end{aligned} \tag{7.11}$$

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

In order to prove (7.11) we can harmlessly make the bootstrap assumption

$$\begin{aligned} \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}} &\leq 2C\varepsilon, \\ \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{1}{2}-\delta}}^{\frac{1}{4}} \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}}^{\delta} &\leq (2Cc_0)^{\frac{1}{4}+\delta}, \quad t \in [0, T]. \end{aligned} \tag{7.12}$$

By Sobolev embeddings and interpolation, from (7.12) we obtain the bounds

$$A \lesssim Cc_0 \ll 1, \quad A_{\frac{1}{4}} \lesssim C\varepsilon.$$

Hence, by the energy estimates in Theorem 2 applied to (\mathbf{W}, R) with $s = \frac{3}{4}$, we obtain

$$\begin{aligned} \|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}^{\frac{3}{4}})} &\lesssim \|(\mathbf{W}, R)(0)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}} + TA_{\frac{1}{4}}^2 \|(\mathbf{W}, R)\|_{L^\infty(0,T;\dot{\mathcal{H}}^{\frac{3}{4}})} \\ &\lesssim \varepsilon + TC^3\varepsilon^3. \end{aligned}$$

Combining this with the same energy estimates as in Theorem 2 but for $s = \frac{1}{2} - \delta$, we also get

$$\|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{1}{2}-\delta}}^{\frac{1}{4}} \|(\mathbf{W}, R)(t)\|_{\dot{\mathcal{H}}^{\frac{3}{4}}}^{\delta} \lesssim (c_0(1 + T\varepsilon^2))^{\frac{1}{4}+\delta}.$$

Hence, the desired estimate (7.11) follows provided that $T \ll (C\varepsilon)^{-2}$.

Funding. The first author was partially supported by the Henry Luce Foundation, as well as by the Simons Foundation. The second author was supported by a Luce Assistant Professorship, by the Sloan Foundation, and by an NSF CAREER grant DMS-1845037. The last author was supported by the NSF grant DMS-1800294, as well as by a Simons Investigator grant from the Simons Foundation.

References

- [1] A. Ai, [Low regularity solutions for gravity water waves](#). *Water Waves* **1** (2019), no. 1, 145–215 Zbl 1451.35125 MR 4161284
- [2] A. Ai, [Low regularity solutions for gravity water waves II: The 2D case](#). *Ann. PDE* **6** (2020), no. 1, article no. 4 Zbl 1462.35275 MR 4098033
- [3] A. Ai, M. Ifrim, and D. Tataru, [Two-dimensional gravity waves at low regularity II: Global solutions](#). *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **39** (2022), no. 4, 819–884 Zbl 1524.76051 MR 4483135
- [4] A. Ai, M. Ifrim, and D. Tataru, [The time-like minimal surface equation in Minkowski space: Low regularity solutions](#). *Invent. Math.* **235** (2024), no. 3, 745–891 Zbl 1537.35245 MR 4701880
- [5] T. Alazard, N. Burq, and C. Zuily, [On the Cauchy problem for gravity water waves](#). *Invent. Math.* **198** (2014), no. 1, 71–163 Zbl 1308.35195 MR 3260858
- [6] T. Alazard, N. Burq, and C. Zuily, [Strichartz estimates and the Cauchy problem for the gravity water waves equations](#). *Mem. Amer. Math. Soc.* **256** (2018), no. 1229, v+108 Zbl 1447.35001 MR 3852259
- [7] T. Alazard and J.-M. Delort, [Global solutions and asymptotic behavior for two dimensional gravity water waves](#). *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 5, 1149–1238 Zbl 1347.35198 MR 3429478
- [8] H. Bahouri and J.-Y. Chemin, [Équations d’ondes quasi-linéaires et effet dispersif](#). *C. R. Acad. Sci. Paris Sér. I Math.* **329** (1999), no. 2, 117–120 Zbl 0932.35147 MR 1710506
- [9] H. Bahouri and J.-Y. Chemin, [Équations d’ondes quasilineaires et estimations de Strichartz](#). *Amer. J. Math.* **121** (1999), no. 6, 1337–1377 Zbl 0952.35073 MR 1719798
- [10] M. Berti, R. Feola, and F. Pusateri, [Birkhoff normal form for gravity water waves](#). *Water Waves* **3** (2021), no. 1, 117–126 Zbl 1477.35156 MR 4246390
- [11] A. Constantin, [Nonlinear water waves with applications to wave-current interactions and tsunamis](#). CBMS-NSF Regional Conf. Ser. in Appl. Math. 81, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011 Zbl 1266.76002 MR 2867413
- [12] W. Craig, C. Sulem, and P.-L. Sulem, [Nonlinear modulation of gravity waves: A rigorous approach](#). *Nonlinearity* **5** (1992), no. 2, 497–522 Zbl 0742.76012 MR 1158383

- [13] A. I. Dyachenko, E. A. Kuznetsov, M. 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), no. 1-2, 73–79
- [14] M. D. Groves, [Steady water waves](#). *J. Nonlinear Math. Phys.* **11** (2004), no. 4, 435–460
Zbl 1064.35144 MR 2097656
- [15] J. K. Hunter, M. Ifrim, and D. Tataru, [Two dimensional water waves in holomorphic coordinates](#). *Comm. Math. Phys.* **346** (2016), no. 2, 483–552 Zbl 1358.35121 MR 3535894
- [16] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong, [Long time solutions for a Burgers–Hilbert equation via a modified energy method](#). *Proc. Amer. Math. Soc.* **143** (2015), no. 8, 3407–3412
Zbl 1320.35071 MR 3348783
- [17] D. Lannes, *The water waves problem*. Math. Surveys Monogr. 188, American Mathematical Society, Providence, RI, 2013 Zbl 1410.35003 MR 3060183
- [18] T. Levi-Civita, [Détermination rigoureuse des ondes permanentes d’ampleur finie](#). *Math. Ann.* **93** (1925), no. 1, 264–314 Zbl 51.0671.06 MR 1512238
- [19] V. I. Nalimov, The Cauchy–Poisson problem. *Dinamika Splošn. Sredy* (1974), no. 18, Dinamika Židkost. so Svobod. Granicami, 104–210, 254 MR 0609882
- [20] M. Ogawa and A. Tani, [Free boundary problem for an incompressible ideal fluid with surface tension](#). *Math. Models Methods Appl. Sci.* **12** (2002), no. 12, 1725–1740 Zbl 1023.76007
MR 1946720
- [21] L. V. Ovsjannikov, To the shallow water theory foundation. *Arch. Mech. (Arch. Mech. Stos.)* **26** (1974), 407–422 Zbl 0283.76012 MR 0347218
- [22] H. F. Smith and D. Tataru, [Sharp local well-posedness results for the nonlinear wave equation](#). *Ann. of Math. (2)* **162** (2005), no. 1, 291–366 Zbl 1098.35113 MR 2178963
- [23] D. Tataru, [Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation](#). *Amer. J. Math.* **122** (2000), no. 2, 349–376 Zbl 0959.35125 MR 1749052
- [24] D. Tataru, [Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II](#). *Amer. J. Math.* **123** (2001), no. 3, 385–423 Zbl 0988.35037 MR 1833146
- [25] D. Tataru, [Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III](#). *J. Amer. Math. Soc.* **15** (2002), no. 2, 419–442 Zbl 0990.35027 MR 1887639
- [26] S. Wu, [Well-posedness in Sobolev spaces of the full water wave problem in 2-D](#). *Invent. Math.* **130** (1997), no. 1, 39–72 Zbl 0892.76009 MR 1471885
- [27] S. Wu, [Almost global wellposedness of the 2-D full water wave problem](#). *Invent. Math.* **177** (2009), no. 1, 45–135 Zbl 1181.35205 MR 2507638
- [28] S. Wu, The quartic integrability and long time existence of steep water waves in 2D. [v1] 2020, [v2] 2021, arXiv:2010.09117v2
- [29] H. Yosihara, [Gravity waves on the free surface of an incompressible perfect fluid of finite depth](#). *Publ. Res. Inst. Math. Sci.* **18** (1982), no. 1, 49–96 Zbl 0493.76018 MR 0660822
- [30] V. E. Zakharov, [Stability of periodic waves of finite amplitude on the surface of a deep fluid](#). *J. Appl. Mech. Tech. Phys.* **9** (1968), no. 2, 190–194

Received 23 March 2024; accepted 12 September 2024.

Albert Ai

Department of Mathematics, University of Wisconsin, Madison, 480 Lincoln Drive, Madison, WI 53706, USA; aai@math.wisc.edu

Mihaela Ifrim

Department of Mathematics, University of Wisconsin, Madison, 480 Lincoln Drive, Madison, WI 53706, USA; ifrim@wisc.edu

Daniel Tataru

Department of Mathematics, University of California, Berkeley, Evans Hall, Berkeley, CA 94720, USA; tataru@math.berkeley.edu