INVISCID DAMPING AND THE ASYMPTOTIC STABILITY OF PLANAR SHEAR FLOWS IN THE 2D EULER EQUATIONS

by JACOB BEDROSSIAN* *and* NADER MASMOUDI

ABSTRACT

We prove asymptotic stability of shear flows close to the planar Couette flow in the 2D inviscid Euler equations on $\mathbf{T} \times \mathbf{R}$. That is, given an initial perturbation of the Couette flow small in a suitable regularity class, specifically Gevrey space of class smaller than 2, the velocity converges strongly in L^2 to a shear flow which is also close to the Couette flow. The vorticity is asymptotically driven to small scales by a linear evolution and weakly converges as $t \to \pm\infty$. The strong convergence of the velocity field is sometimes referred to as *inviscid damping*, due to the relationship with Landau damping in the Vlasov equations. This convergence was formally derived at the linear level by Kelvin in 1887 and it occurs at an algebraic rate first computed by Orr in 1907; our work appears to be the first rigorous confirmation of this behavior on the nonlinear level.

CONTENTS

1. Introduction . . . 195
2. Proof of Theorem 1 . . . 204
3. Growth mechanism and construction of A . . . 219
4. Elliptic estimates . . . 236
5. Transport . . . 247
6. Reaction . . . 252
7. Remainder . . . 268
8. Coordinate system controls . . . 269
9. Conclusion . . . 285
Acknowledgements . . . 286
Appendix A . . . 286
References . . . 297

1. Introduction

We consider the 2D Euler system in the vorticity formulation with a background shear flow:

$$\textbf{(1.1)}\qquad \begin{cases} \omega_t + y\partial_x\omega + U\cdot\nabla\omega = 0, \\ U = \nabla^{\perp}(\Delta)^{-1}\omega, \qquad \omega(t=0) = \omega_{in}. \end{cases}$$

Here, $(x, y) \in \mathbf{T} \times \mathbf{R}$, $\nabla^{\perp} = (-\partial_y, \partial_x)$ and (U, ω) are periodic in the x variable with period normalized to 2π. The physical velocity is $(y, 0) + U$ where $U = (U^x, U^y)$ denotes the velocity perturbation and the total vorticity is $-1 + \omega$. We denote the streamfunction by $\psi = \Delta^{-1}\omega$. The velocity itself satisfies the momentum equation

$$\textbf{(1.2)}\qquad \begin{cases} U_t + y\partial_x U + (U^y, 0) + U\cdot\nabla U = -\nabla P, \\ \nabla\cdot U = 0, \end{cases}$$

J. Bedrossian was partially supported by NSF Postdoctoral Fellowship in Mathematical Sciences, DMS-1103765.
N. Masmoudi was partially supported by NSF grant DMS-1211806.

* The current email and address are jacob@cscamm.umd.edu, Department of Mathematics and the Center for Scientific Computation and Mathematical Modeling, University of Maryland, College Park, MD, USA.

DOI 10.1007/s10240-015-0070-4

where P denotes the pressure. Linearizing the vorticity equation (1.1) yields the linear evolution

$$\textbf{(1.3)} \qquad \begin{cases} \omega_t + y\partial_x \omega = 0, \\ U = \nabla^{\perp}(\Delta)^{-1}\omega, \qquad \omega(t=0) = \omega_{in}. \end{cases}$$

In this work, we are interested in the long time behavior of (1.1) for small initial perturbations ω_{in}. In particular, we show that all sufficiently small perturbations in a suitable regularity class undergo 'inviscid damping' and satisfy $(y, 0) + U(t, x, y) \to (y + u_\infty(y), 0)$ as $t \to \infty$ for some $u_\infty(y)$ determined by the evolution.

The field of hydrodynamic stability started in the nineteenth century with Stokes, Helmholtz, Reynolds, Rayleigh, Kelvin, Orr, Sommerfeld and many others. Rayleigh [76] studied the linear stability and instability of planar inviscid shear flows using what is now referred to as the *normal mode method*. Such a method yields *spectral instability* or *spectral stability* depending on whether or not an unstable eigenvalue exists. In that work, Rayleigh proves the famous inflection point theorem which gives a necessary condition for spectral instability. At around the same time, Kelvin [47] constructed exact solutions to the linearized problem around the Couette flow (which are actually solutions of the nonlinear problem). This was the first attempt to solve the initial value problem of the linearized problem which was later developed further in [22, 66, 74].

Even to the present day, the methods used and the conclusions of these works are debated both on physical and mathematical grounds [93]. Experimental realizations of Couette and similar spectrally stable flows show instability and transition to turbulence for sufficiently high Reynolds numbers [15, 62, 75, 78, 86]. The paradox that Couette flow is known to be spectrally stable for all Reynolds numbers in contradiction with instabilities observed in experiments is now often referred to as the 'Sommerfeld paradox', or 'turbulence paradox'. However, experiments are ultimately inconclusive (mathematically) since many factors are notoriously difficult to control, such as imperfections in the walls and viscous boundary layers. Moreover, the early mathematical works did not fully treat 3D effects, which are by now known to have a major impact on the behavior [31, 87, 88, 93]. In fact, our work, together with our follow up work with V. Vicol [13], shows that the 'subcritical transition' observed in 3D flows does not occur in 2D for sufficiently regular perturbations.

Of course, from a mathematical point of view, the notion of linear stability was not completely precise in the early works: was it enough that the linear operator has no growing mode (*spectral stability*) or should one consider general initial perturbation and study the time evolution under the linear equation using, for instance, a Laplace transform in time (see [18, 22, 66]). These early works also pre-dated the notion of Lyapunov stability and Sobolev spaces; indeed, the stability of (1.3) depends heavily on the norm chosen. A more variational approach, which is based on the conserved quantities and uses the notion of Energy-Casimir, was introduced by Arnold [1], and yields Lyapunov stability for a class of shear flows (which does not include Couette flow). We also refer to [44, 53] for the use of the variational approach in the Vlasov case. Recently there were many mathematical studies of stability and instability of various flows (see for instance [10, 34, 43, 57]). We also refer to the following textbooks on the topic of hydrodynamic stability and instability [30, 56, 93].

There were many attempts in the literature to find an explanation to the Sommerfeld paradox (see [55] and the references therein). The first attempt might be due to Orr [74] in 1907, whose work plays a central role in ours. We give a more detailed discussion of the linear behavior in Section 2.1 below, but Orr's observation can be summarized in modern terminology (and adapted to our infinite-in-y setting) as follows. Given a disturbance in the vorticity, the linear evolution under (1.3) is simply advected by the background shear flow: $\omega(t, x, y) = \omega_{in}(x - ty, y)$. If one changes coordinates to $z = x - ty$ then the stream-function $\phi(t, z, y)$ in these variables solves $\partial_{zz}\phi + (\partial_y - t\partial_z)^2\phi = \omega_{in}$. On the Fourier side, $(z, y) \to (k, \eta) \in \mathbf{Z} \times \mathbf{R}$,

$$\hat{\phi}(t, k, \eta) = -\frac{\hat{\omega}_{in}(k, \eta)}{k^2 + |\eta - kt|^2}. \tag{1.4}$$

From (1.4), Orr made two important observations, together known now as the *Orr mechanism*. Firstly, if $\eta, k > 0$ and η is very large relative to k, then the stream-function amplifies by a factor $\frac{\eta^2}{k^2}$ at a *critical time* given by $t_c = \frac{\eta}{k}$. These modes correspond to waves tilted against the shear which are being advected to larger length-scales (lower frequencies). Orr suggested that this transient growth is a possible explanation for the observed *practical instability* or at least as a reason to question the validity of the linear approximation. Moreover, this shows that the Couette flow is linearly unstable (in the sense of Lyapunov) in the kinetic energy norm. On the other hand, Orr states in [74, ART. 12] that "the motion is stable, for the most general disturbance, if sufficiently small". Orr does not precise the meaning of *sufficiently small* but concludes in this case that "the y velocity-component eventually diminishes indefinitely as t^{-2}, and, the x component of the relative velocity as t^{-1}". In fact, on the linear level, it is not about smallness but about regularity. Indeed, rigorous proof of the stability and decay on the linear level requires the use of a stronger norm on the initial data than on the evolution, as already noticed in Case [22] and Marcus and Press [66] where this linear stability and decay are proved.

Physically, the decay predicted by the Orr mechanism can be understood as the transfer of enstrophy to small scales (which yields the decay of the velocity by the Biot-Savart law) and the transient growth can be understood as the time-reversed phenomenon: the transfer of enstrophy from small scales to large scales and hence the growth of the velocity (see also [17, 61] for further discussion). The transfer to small scales by mixing is now considered a fundamental mechanism intimately connected with the stability of coherent structures and the theory of 2D turbulence [38, 48]. However, to our knowledge, our work is the first mathematically rigorous study of this mechanism in the full 2D Euler equations. We refer to [42, 50, 83, 93] for the most recent developments. Mathematically, one can also explain the transient growth by the non-normality of the linearized operator (also an insight first due to Orr). See for example [87], where the implications of this are studied in terms of the spectra and pseudospectra of the linearized Couette and Poiseuille flows. Indeed, the fact that for non-normal operators the ϵ-pseudospectrum can be very different from the spectrum can be seen as another explanation of the transient linear growth [77, 88]. See also [81] for further information.

In 1946, Landau [51] predicted rapid decay of the electric field in hot plasmas perturbed from homogeneous equilibrium by solving the linearized Vlasov equation with a Laplace transform. Now referred to as *Landau damping*, this somewhat controversial prediction of collisionless relaxation in a time-reversible physical model was confirmed by experiments much later in [64] and is now a well-accepted, ubiquitous phenomenon in plasma physics [79]. In [90], van Kampen showed that one way to interpret this mechanism was through the transfer of information to small scales in velocity space; a scenario completely consistent with time-reversibility and conservation of entropy. In this scenario, the free-streaming of particles creates rapid oscillations of the distribution function which are averaged away by the non-local Coulomb interactions (see also [21, 28]). The fundamental stabilizing mechanism in this picture is the *phase-mixing* due to particle streaming. The gap between the linear and nonlinear theory of Landau damping was only bridged recently by the ground-breaking work of Mouhot and Villani, who showed that the phase-mixing indeed persists in the nonlinear Vlasov equations for small perturbations [70] (see also [12, 19, 46]).

The algebraic decay of the velocity field for solutions to (1.3) predicted by Orr can be most readily understood as a consequence of *vorticity mixing* driven by the shear flow, and hence can be considered as a hydrodynamic analogue of Landau damping, a viewpoint furthered by many authors [5, 16, 18, 80]. Hence the origin of the term *inviscid damping*. The first, and most fundamental, difference between (1.3) and the linearized Vlasov equations is the fact that the velocity field induced by mean-zero solutions to (1.3) in general does not converge back to the Couette flow, but in fact converges to a different nearby shear flow, whereas the electric field in the linearized Vlasov equations converges to zero. This 'quasi-linearity' will be a major difficulty in studying inviscid damping on the nonlinear level. Another key difference is that unlike in the Vlasov equations, the decay of the velocity field in (1.3) cannot generally be better than the algebraic rate predicted by Orr, which is not even integrable for the x component of the velocity; to contrast, in the Vlasov equations the decay is exponential for analytic perturbations.

It is well-known that the nonlinearity can change the picture dramatically. A clear example of this are the results of Lin and Zeng [58] who prove that there exists non-trivial periodic solutions to the vorticity equation (1.1) which are arbitrarily close to the Couette flow in H^s for $s < 3/2$. They have also proved the corresponding, and related, result for the Vlasov equations [59]. In our setting, the primary interest is to rule out the possibility that weakly nonlinear effects create a self-sustaining process and push the solution out of the linear regime. The idea that the interaction between nonlinear effects and non-normal transient growth can lead to instabilities is classical in fluid mechanics (see e.g. [87]). The basic mechanism suggested in [87] is that nonlinear effects can repeatedly excite growing modes and precipitate a sustained cascade or so-called 'nonlinear bootstrap', studied further in the fluid mechanics context in, for example, [2, 91, 92]. Actually, this effect is very similar to what is at work behind *plasma echos* in the Vlasov equations, first captured experimentally in [65]. This phenomenon is referred to as an 'echo' because the measurable result of nonlinear effects can occur long after the event. Very similar echos have been studied and observed in 2D Euler, both numerically [91, 92] and experimentally [94, 95] (interestingly, non-neutral plasmas in certain settings make excellent realizations of 2D Euler).

The plasma echos play a pivotal role in the work of Mouhot and Villani on Landau damping [70]. Although our approach to this challenge is quite different, one of the main difficulties we face is to precisely understand the weakly nonlinear effects at work; sometimes called *nonlinear transient growth* [92]. We will need a more precise alternative to the moment estimates of [70] which is tailored to the specific structure of 2D Euler; what we call the "toy model" (see Section 9 for a detailed discussion about the relationship of our work to [70]). The toy model, formally derived in Section 3.1.1, provides mode-by-mode upper bounds on the 'worst possible' growth of high frequencies that the weakly nonlinear effects can produce. The model is not just a heuristic and in fact plays a key role in our work: it is used in the construction of a norm specially designed to match the evolution of (1.1); this norm is the subject of Section 3. We remark that our model has not appeared in the literature before to our knowledge, however related models have been studied in [91, 92].

The mixing phenomenon behind the inviscid damping also appears in many other fluid models, for example, more general shear profiles [6, 16], stratified shear flows [20, 63] and 2D Euler with the β-plane approximation to the Coriolis force [17, 89]. A particularly fundamental setting is the 'axisymmetrization' of vortices in 2D Euler which has important implications for the meta-stability of coherent vortex structures in atmosphere and ocean dynamics (see e.g. [11, 38, 80, 94, 95] for a small piece of the extensive literature). Actually, this stability problem was mentioned by Rayleigh [76] and was considered by Orr as well [74]. Interestingly, it is also relevant to the stability of charged particle beams in cyclotrons [23].

In general, phase-mixing, or 'continuum damping', can be directly associated with the continuous spectrum of the linearized operator and is a phenomenon shared by a number of infinite-dimensional Hamiltonian systems, for example the damping of MHD waves [85], the Caldeira-Legget model from quantum mechanics [45] and synchronization models in biology [84]. See the series of works [5, 6, 8, 68, 69] which draws a connection between the van Kampen generalized eigenfunctions and the normal form transform to write the linearized 2D Euler and Vlasov-Poisson equations as a continuum of decoupled harmonic oscillators. See also [7] and the references therein for a recent survey which contains other examples and discusses some connections between these various models.

Phase mixing also shares certain similarities with scattering in the theory of dispersive wave equations (see for instance [36, 52, 60]) as already pointed out in [19, 28]. In both cases the long time behavior is governed by a linear operator, or a modified version of it due to long range interactions [39, 71] (something like this occurs in our Theorem 1). Unlike dissipative equations, the final linear evolution is usually chosen by the entire nonlinear dynamics and cannot be completely characterized by the relevant conservation laws. Also in both cases, the phenomena can be related to the continuous spectrum in the linear problem; for example, the RAGE theorem applies equally well to transport equations as to dispersive equations [27]. However, there are also clear differences since in dispersive wave equations, the dispersion uses the fact that different wave packets travel with different group velocities to yield decay of the L^∞ norm and hence nonlinear terms often become weaker. Normally, this decay costs *spatial localization* rather than *regularity*. In the inviscid damping (and Landau damping), the decay is due to the combination of the mixing which sends the information into high frequencies and the application of the inverse Laplacian (or any operator of negative order), which averages out the small scales. That is, dispersion transfers information to infinity in space whereas mixing transfers information to infinity in frequency.

1.1. *Statement and discussion*

In this section we state our nonlinear stability result and a few immediate corollaries. The key aspects of the proof are discussed after the statement.

The data will be chosen in a Gevrey space of class $1/s$ for $s > 1/2$ [37]; the origin of this restriction is (mathematically) natural and arises from the weakly nonlinear effects, discussed further in Section 3. We note that the analogous space for the Vlasov equations with Coulomb/Newton interaction is Gevrey-3 (e.g. $s = 1/3$) [12, 70]. It is worth noting that unlike, for example, [35] where the Gevrey regularity is required due to the linear growth of high frequencies, here (and [70]) the Gevrey regularity is required because of a potential *nonlinear* frequency cascade.

Our main result is

Theorem **1.** — *For all* $1/2 < s \leq 1$, $\lambda_0 > \lambda' > 0$ *there exists an* $\epsilon_0 = \epsilon_0(\lambda_0, \lambda', s) \leq 1/2$ *such that for all* $\epsilon \leq \epsilon_0$ *if* ω_{in} *satisfies* $\int \omega_{in} dxdy = 0$, $\int |y\omega_{in}(x, y)| dxdy < \epsilon$ *and*

$$\|\omega_{in}\|_{\mathcal{G}^{\lambda_0}}^2 = \sum_k \int \left|\hat{\omega}_{in}(k, \eta)\right|^2 e^{2\lambda_0 |k,\eta|^s} d\eta \leq \epsilon^2,$$

then there exists f_∞ *with* $\int f_\infty dxdy = 0$ *and* $\|f_\infty\|_{\mathcal{G}^{\lambda'}} \lesssim \epsilon$ *such that*

$$\left\|\omega\big(t, x + ty + \Phi(t, y), y\big) - f_\infty(x, y)\right\|_{\mathcal{G}^{\lambda'}} \lesssim \frac{\epsilon^2}{\langle t\rangle}, \tag{1.5}$$

where $\Phi(t,y)$ *is given explicitly by*

$$\Phi(t,y) = \frac{1}{2\pi}\int_0^t \int_{\mathbf{T}} U^x(\tau,x,y)\,dx\,d\tau = u_\infty(y)t + O\big(\epsilon^2\big), \tag{1.6}$$

with $u_\infty = \partial_y \partial_{yy}^{-1} \frac{1}{2\pi}\int_{\mathbf{T}} f_\infty(x,y)\,dx$. *Moreover, the velocity field* U *satisfies*

$$\left\| \frac{1}{2\pi}\int U^x(t,x,\cdot)\,dx - u_\infty \right\|_{\mathcal{G}^{\lambda'}} \lesssim \frac{\epsilon^2}{\langle t\rangle^2}, \tag{1.7a}$$

$$\left\| U^x(t) - \frac{1}{2\pi}\int U^x(t,x,\cdot)\,dx \right\|_{L^2} \lesssim \frac{\epsilon}{\langle t\rangle}, \tag{1.7b}$$

$$\left\| U^y(t) \right\|_{L^2} \lesssim \frac{\epsilon}{\langle t\rangle^2}. \tag{1.7c}$$

Remark **1.** — Of course, by time-reversibility, Theorem 1 is also true $t \to -\infty$ for some $f_{-\infty}$ and $u_{-\infty}$ (which will generally not be equal to their $+\infty$ counterparts). Also, due to the Hamiltonian structure of (1.1) (see e.g. [1, 68]), one could only hope to prove asymptotic stability in a norm weaker than the norm in which the initial data is given. This is an important theme underlying our work, and the works of [19, 46, 70], which is that *decay costs regularity*.

Remark **2.** — From the proof of Theorem 1, it is clear that $\|\omega_{in} - f_\infty\|_{\mathcal{G}^{\lambda'}} \lesssim \epsilon^2$, as the effect of the nonlinear evolution is one order weaker than that of the linear evolution.

Remark **3.** — Notice the surprisingly rapid convergence in (1.7a) (it is of course matched by a similar rapid convergence of the x-averaged vorticity). This arises from a subtle cancellation between the oscillations of ω and U^y upon taking x averages; indeed it was previously believed that the convergence should be $O(t^{-1})$ and that (1.6) involved a logarithmic correction. The origin of the rapid convergence rate can be best understood from studying the linearized problem (1.3), a computation that we carry out in Section A.4.

Remark **4.** — The proof of Theorem 1 implies that if ω_{in} is compactly supported then $\omega(t)$ remains supported in a strip $(x,y) \in \mathbf{T} \times [-R, R]$ for some $R > 0$ for all time.

Remark **5.** — The primary difficulty in treating more general shear flows is on the weakly nonlinear level (in contrast to the Vlasov case), which would most clearly manifest in Section 4. More information on this difficulty, along with other related open problems, is discussed in Section 9.

Remark **6.** — Both Orr and Kelvin (and many others) expressed doubt that the inviscid problem was stable unless the set of permissible data was of a certain type, suggesting that for *general data* the stability restriction would diminish with the inverse Reynolds number. To reconcile this viewpoint with Theorem 1, with V. Vicol, we have recently proved in [13] that for high (but finite) Reynolds number flows, an analogous result to Theorem 1 holds with initial data $\omega_{in}^{R} + \omega_{in}^{\nu}$ where ω_{in}^{R} has Gevrey-$\frac{1}{s}$ regularity uniformly in the Reynolds number and ω_{in}^{ν} has L^2 regularity with norm small with respect to the inverse Reynolds number. In addition to showing that the qualitative behavior predicted in Theorem 1 can be obtained via an inviscid limit, [13] also shows that the mixing greatly enhances the effect of the dissipation in modes which depend on x.

Remark **7.** — The spatial localization $\int |y\omega_{in}(x,y)|dxdy < \epsilon$ is only used to assert that the velocity U^x is in L^2 and to ensure the coordinate transformations used in the proof are not too drastic. This assumption can be relaxed to $\int |y|^\alpha |\omega_{in}(x,y)|dxdy < \epsilon$ for any $\alpha > 1/2$. It might be possible to treat more general cases with $U^x \notin L^2$ with some technical enhancements, as $U^x \in L^2$ does not play an important role in the proof.

The proof of Theorem 1 (actually Remark 2) provides the following corollaries.

Corollary **1.** — *There exists an open set of smooth solutions to* (1.1) *for which* $\{\omega(t)\}_{t\in\mathbf{R}}$ *is not pre-compact in* L^2 *as* $t \to \pm\infty$. *In particular,* $\omega(t) \rightharpoonup \omega_\infty = \frac{1}{2\pi}\int_{\mathbf{T}} f_\infty(x,y)dx$ *and in general* $\|\omega_\infty\|_2 < \|\omega(t)\|_2$.

This shows the existence of solutions for which enstrophy is lost to high frequencies in the limit $t \to \infty$, which to our knowledge was not previously known for 2D Euler in any setting. See [42, 50, 83] for further discussions on the physical interest of this fact and the potential relationship with 2D turbulence. A related corollary is the following which shows the linear growth of Sobolev norms as a direct consequence of the mixing. Compare with the construction of Denisov [29] which yields super-linear growth of the gradient.

Corollary **2.** — *There exists an open set of smooth solutions to* (1.1) *for which* $\|\langle\nabla\rangle^N \omega(t)\|_2 \approx \langle t\rangle^N$ *for all* $N \in [0,\infty)$.

Let us now outline the main new steps in the proof of Theorem 1. First, we provide a (well chosen) change of variable that adapts to the solution as it evolves and yields a new 'relative' velocity which is time-integrable while keeping the Orr critical times as in (1.4). This change of variables allows us to work on a quantity $f(t)$ which has a strong limit as t goes to infinity. This is related to the notion of "profile" used in dispersive wave equations (see [36] for instance) as well as the notion of "gliding regularity" in [70]. However, here it is important that the coordinate transformation depends on the solution, a source of large technical difficulty and an expression of the 'quasi-linearity' alluded to above.

A second new idea is the use of a special norm that loses regularity in a very precise way adapted to the Orr critical times and the associated nonlinear effect. The construction of this norm is based on the so-called "toy model" which mimics the worse possible growth of high frequencies (derived in Section 3.1.1). This special norm allows us to control the nonlinear growth due to the resonances at the critical times. However, this comes with a big danger: energy estimates and cancellations tend to dislike 'unbalanced' norms, namely norms that assign different regularities to different frequencies (see for instance [67] for a similar problem). In particular, by design, our norm is not an algebra. This is one of the main technical problems that we have to overcome, and here the decay of the velocity is crucial.

In the course of the proof, we need to gain regularity from inverting the Laplacian to get the streamfunction from the vorticity; indeed the ellipticity is the origin of the decay. However, in the new variables the Laplacian is transformed to a weakly elliptic operator with coefficients that depend on the solution. This additional nonlinearity presents huge difficulties due to the limited regularity of the coefficients (relative to what is desired). This has similarities with elliptic estimates in domains with limited regularity used for water waves (see for instance [82, Appendix A]). Here, the interplay between regularity and decay will be crucial to ensure that the final estimate holds. As in (1.4), the loss of ellipticity is an expression of the Orr critical times. It will be important for our work that the norm derived from the toy model precisely 'matches' the loss of ellipticity.

Related to the issue of inverting the Laplacian in the new variables is the final technical ingredient in our proof, which is the need to obtain a variety of precise controls on the evolving coordinate system (see Proposition 2.5 below). This will require us to quantify the convergence of the background shear flow in several ways. In particular, we will need to carefully estimate how the modes that depend on x force those that do not and in fact, this forcing loses a derivative (see the last term in (8.9)). However, the estimates turn out to be possible precisely under the assumption of Gevrey class with $s \geq 1/2$ (see the discussion after Proposition 2.5). Second to the toy model, here seems to be next most fundamental use of the regularity $s \geq 1/2$.

1.2. *Notation and conventions*

See Section A.1 for the Fourier analysis conventions we are taking. A convention we generally use is to denote the discrete x (or z) frequencies as subscripts. By convention we always use Greek letters such as η and ξ to denote frequencies in the y or v direction and lowercase Latin characters commonly used as indices such as k and l to denote frequencies in the x or z direction (which are discrete). Another convention we use is to denote K, M, N as dyadic integers $\mathrm{K}, \mathrm{M}, \mathrm{N} \in \mathbf{D}$ where

$$\mathbf{D} = \left\{ \frac{1}{2}, 1, 2, \ldots, 2^j, \ldots \right\}.$$

When a sum is written with indices K, M, M′, N or N′ it will always be over a subset of $\mathbf{D}$. This will be useful when defining Littlewood-Paley projections and paraproduct decompositions, see Section A.1. Given a function $m \in \mathrm{L}^\infty$, we define the Fourier multiplier

$m(\nabla)f$ by

$$\left(\widehat{m(\nabla)f}\right)_k(\eta) = m\big((ik, i\eta)\big)\hat{f}_k(\eta).$$

We use the notation $f \lesssim g$ when there exists a constant $C > 0$ independent of the parameters of interest such that $f \leq Cg$ (we analogously $f \gtrsim g$ define). Similarly, we use the notation $f \approx g$ when there exists $C > 0$ such that $C^{-1}g \leq f \leq Cg$. We sometimes use the notation $f \lesssim_\alpha g$ if we want to emphasize that the implicit constant depends on some parameter α. We will denote the l^1 vector norm $|k, \eta| = |k| + |\eta|$, which by convention is the norm taken in our work. Similarly, given a scalar or vector in $\mathbf{R}^n$ we denote

$$\langle v \rangle = \left(1 + |v|^2\right)^{1/2}.$$

We use a similar notation to denote the x (or z) average of a function: $\langle f \rangle = \frac{1}{2\pi}\int f(x,y)dx = f_0$. We also frequently use the notation $P_{\neq 0}f = f - f_0$. We denote the standard L^p norms by $\|f\|_{L^p}$. We make common use of the Gevrey-$\frac{1}{s}$ norm with Sobolev correction defined by

$$\|f\|^2_{\mathcal{G}^{\lambda,\sigma;s}} = \sum_k \int \left|\hat{f}_k(\eta)\right|^2 e^{2\lambda|k,\eta|^s} \langle k, \eta \rangle^{2\sigma} d\eta.$$

Since most of the paper we are taking s as a fixed constant, it is normally omitted. We refer to this norm as the $\mathcal{G}^{\lambda,\sigma;s}$ norm and occasionally refer to the space of functions

$$\mathcal{G}^{\lambda,\sigma;s} = \left\{f \in L^2 : \|f\|_{\mathcal{G}^{\lambda,\sigma}} < \infty\right\}.$$

See Section A.2 for a discussion of the basic properties of this norm and some related useful inequalities.

For $\eta \geq 0$, we define $E(\eta) \in \mathbf{Z}$ to be the integer part. We define for $\eta \in \mathbf{R}$ and $1 \leq |k| \leq E(\sqrt{|\eta|})$ with $\eta k \geq 0$, $t_{k,\eta} = |\frac{\eta}{k}| - \frac{|\eta|}{2|k|(|k|+1)} = \frac{|\eta|}{|k|+1} + \frac{|\eta|}{2|k|(|k|+1)}$ and $t_{0,\eta} = 2|\eta|$ and the critical intervals

$$I_{k,\eta} = \begin{cases} [t_{|k|,\eta}, t_{|k|-1,\eta}] & \text{if } \eta k \geq 0 \text{ and } 1 \leq |k| \leq E(\sqrt{|\eta|}), \\ \emptyset & \text{otherwise.} \end{cases}$$

For minor technical reasons, we define a slightly restricted subset as the *resonant intervals*

$$\mathbf{I}_{k,\eta} = \begin{cases} I_{k,\eta} & 2\sqrt{|\eta|} \leq t_{k,\eta}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Note this is the same as putting a slightly more stringent requirement on k: $k \gtrsim \frac{1}{2}\sqrt{|\eta|}$.

2. Proof of Theorem 1

We next give the proof of Theorem 1, stating the primary steps as propositions which are proved in subsequent sections.

2.1. *Linearized behavior and main challenges*

Before beginning the proof of Theorem 1, we discuss the linearized behavior in more detail and mention some of the main challenges that must be overcome for a nonlinear result. The linearized problem is stated above in (1.3); here we denote the streamfunction

$$\Delta\psi = \omega,$$

which satisfies $U = \nabla^{\perp}\psi$. As mentioned in Section 1, the solution to the linear problem and its Fourier transform are given by

$$\begin{aligned} \omega(t,x,y) &= \omega_{in}(x-ty,y) \\ \widehat{\omega}(t,k,\eta) &= \widehat{\omega_{in}}(k,\eta+kt). \end{aligned} \tag{2.1}$$

From (2.1) we can see the transfer of enstrophy to high frequencies, which for each k is linear in time. Since Δ^{-1} gains two derivatives, this transfer of information to high frequencies will cause the t^{-2} decay of $P_{\neq 0}\psi$ in L^2. Though not necessary to fully understand the linear problem, we begin by making the following change of coordinates (used also by Lord Kelvin [47] and Orr [74]):

$$z = x - ty \tag{2.2a}$$

$$f(t,z,y) = \omega(t,z+ty,y) = \omega(t,x,y) \tag{2.2b}$$

$$\phi(t,z,y) = \psi(t,z+ty,y) = \psi(t,x,y). \tag{2.2c}$$

From (1.3) we have

$$\partial_t f = 0 \tag{2.3a}$$

$$\partial_{zz}\phi + (\partial_y - t\partial_x)^2\phi = f, \tag{2.3b}$$

which re-derives (1.4)

$$\hat{\phi}(t,k,\eta) = -\frac{\hat{f}(t,k,\eta)}{k^2+|\eta-kt|^2} = -\frac{\hat{\omega}_{in}(k,\eta)}{k^2+|\eta-kt|^2}. \tag{2.4}$$

From (2.4) we derive the fundamental decay-by-mixing estimate: for any $\sigma \in [0,\infty)$ and $\beta \in [0,2]$,

$$\|P_{\neq 0}\phi\|_{H^\sigma} \lesssim \frac{1}{\langle t\rangle^{\beta}}\|f\|_{H^{\sigma+\beta}} = \frac{1}{\langle t\rangle^{\beta}}\|\omega_{in}\|_{H^{\sigma+\beta}}, \tag{2.5}$$

where we are using H^σ to denote the L^2 Sobolev norm of order σ. Due to

$$U^x(t,x,y) = -\partial_y\psi(t,x,y) = -\partial_y\big(\phi(t,x-ty,y)\big)$$

$$= \big((\partial_y - t\partial_x)\phi\big)(t, x - ty, y)$$

$$U^y(t,x,y) = \partial_x \psi(t,x,y) = \partial_x\big(\phi(t, x - ty, y)\big) = (\partial_x\phi)(t, x-ty, y),$$

we get the inviscid damping predicted by Orr in [74] and observed in Theorem 1 in (1.7)

$$\textbf{(2.6a)} \qquad \left\| P_{\neq 0} U^x \right\|_{L^2} \lesssim \langle t\rangle \|\nabla\phi\|_{L^2} \lesssim \langle t\rangle^{-1} \|\omega_{in}\|_{H^3}$$

$$\textbf{(2.6b)} \qquad \left\| U^y \right\|_{L^2} \lesssim \|\partial_x\phi\|_{L^2} \lesssim \langle t\rangle^{-2} \|\omega_{in}\|_{H^3}.$$

This shows that on the linear level, we have the convergence $(y + U^x, U^y) \to (y + \langle U^x_{in}\rangle(y), 0)$ in L^2 as time goes to infinity. Hence, the velocity field converges strongly back to a shear flow but not back to the Couette flow. As discussed in the introduction, Orr had a second observation from (2.4), which is that modes with $\eta k > 0$ undergo first a transient growth in ϕ before decaying. Physically, enstrophy in these modes is first *un-mixed* to larger scales, prompting growth in ϕ, before subsequently being mixed to smaller scales and yielding the eventual decay of ϕ.

A natural first attempt at proving something like Theorem 1 is to again make the change of variables (2.2) and derive the analogue of (2.3) from (1.1):

$$\textbf{(2.7a)} \qquad \partial_t f + \nabla^\perp_{z,y}\phi \cdot \nabla_{z,y} f = 0$$

$$\textbf{(2.7b)} \qquad \partial_{zz}\phi + (\partial_y - t\partial_x)^2\phi = f,$$

which again implies

$$\textbf{(2.8)} \qquad \hat{\phi}(t,k,\eta) = -\frac{\hat{f}(t,k,\eta)}{k^2 + |\eta - kt|^2}.$$

From (2.8) we see that (2.6) will follow as soon as we can get uniform control on f, the solution to (2.7), in H^3. However, there are several major reasons why this is extremely difficult. First, the contribution from $-\langle\partial_y\phi\rangle_z$ will not decay – this is the part of the shear flow that comes from f. Indeed, the coordinate transformation $z = x - ty$ is made assuming that ω will be mixed by the Couette flow as $t \to \infty$ but this is incorrect: it will be mixed by a shear flow which is in general $O(\epsilon)$ away from the Couette flow, where ϵ is the size of f. This in turn causes a growth of derivatives like $O(\epsilon t)$ (per derivative) on f and no convergence. This suggests that (2.2) is not the right way to study the nonlinear problem and in Theorem 1, this manifests in the presence of Φ in (1.5). The second major issue with (2.7a) is a bit more subtle and is connected to the echo phenomenon. Indeed, imagine trying to get an estimate on f in H^σ:

$$\frac{1}{2}\frac{d}{dt}\int \left|\langle\nabla\rangle^\sigma f\right|^2 dzdy = -\int \langle\nabla\rangle^\sigma f \langle\nabla\rangle^\sigma \left(\nabla^\perp\phi\cdot\nabla f\right) dzdy.$$

Even ignoring the contribution from $-\langle \partial_y \phi \rangle_z$, since f does not decay, in order to integrate the nonlinearity we will need $\|\nabla^{\perp} P_{\neq 0}\phi\|_{H^\sigma}$ to be integrable in time. From (2.5) we see that without any detailed analysis, this would already require a uniform bound on $\|f\|_{H^{\sigma+\beta}}$ for some $\beta > 2$ (since we already lose a derivative from the $\nabla^{\perp}$). Even with analytic regularity we could only hope to absorb a loss of one derivative by paying regularity along time, and hence we are quite far from closing an estimate (see e.g. classical Cauchy-Kovalevskaya-type arguments like [72, 73] and more recent, relevant variants [49, 54]). One can directly relate this loss of regularity in the energy estimate with the nonlinear effect of the Orr mechanism, which gives rise to the echo phenomenon observed in 2D Euler [68, 91, 95] (and plasmas [12, 65, 70]). In Section 3 below, we analyze the effect of this mechanism in more detail and in particular, show that this could potentially cause a frequency cascade and lead to the loss of Gevrey-2 regularity on f as $t \to \infty$, which is the origin of the regularity requirement in Theorem 1. Although Gevrey-2 is an infinite regularity class, note that this is far less pessimistic than what a naive argument based only on (2.6) suggests and, in particular, is weaker than analytic regularity in a qualitatively meaningful way.

Hence, we have two main challenges to overcome. The first is to choose a coordinate system that is properly adapted to the shear flow which is mixing the solution. Note that this shear flow is changing in time and cannot be determined directly from the initial data. We carry this out in Section 2.2 below. The next step is to get global-in-time, uniform regularity estimates on the resulting f, which is carried out in Section 2.3. To do this we will design a special norm with which to measure the solution that accounts for the nonlinear Orr mechanism. This norm is constructed and analyzed in Section 3. The remainder of the paper is devoted to proving the energy estimates set up in Section 2.3.

2.2. *Coordinate transform*

The original equations in vorticity form are (1.1), and we are trying essentially to prove that

$$\omega(t,x,y) \to f_\infty\big(x - ty - u_\infty(y)t, y\big),$$

as $t \to \infty$, where $u_\infty(y)$ is the correction to the shear flow determined by f_∞. From the initial data alone, there is no simple way to determine u_∞; it is chosen by the nonlinear evolution. In order to deal with this lack of information about how the final state evolves we choose a coordinate system which adapts to the solution and converges to the expected form as $t \to \infty$. The change of coordinates used is $(t,x,y) \to (t,z,v)$, where

$$z(t,x,y) = x - tv \tag{2.9a}$$

$$v(t,y) = y + \frac{1}{t}\int_0^t \langle U^x \rangle(\tau, y)\, d\tau, \tag{2.9b}$$

where we recall $\langle w \rangle$ denotes the average of w in the x variable (or equivalently in the z variable), namely $\langle w \rangle = \frac{1}{2\pi}\int_{\mathbf{T}} w dx$. The reason for the change $y \to v$ is not immediately clear, however v is named as such since it is an approximation for the background shear flow. If the velocity field in the integrand were constant in time, then we are simply transforming the y variables so that the shear appears linear. It will turn out that this choice of v ensures that the Biot-Savart law is in a form amenable to Fourier analysis in the variables (z, v); in particular, even when the shear is time-varying we may still study the *Orr critical times* as was explained in (1.4). In this light, the motivation for the shift in z is clear: as suggested by the discussion in Section 2.1, we are eliminating the contribution of $\langle U^x \rangle$ and following the flow in the horizontal variable to guarantee compactness.

Define $f(t, z, v) = \omega(t, x, y)$ and $\phi(t, z, v) = \psi(t, x, y)$, hence

$$\partial_t \omega = \partial_t f + \partial_t z \partial_z f + \partial_t v \partial_v f, \qquad \partial_x \omega = \partial_z f, \qquad \partial_y \omega = \partial_y v(\partial_v f - t\partial_z f),$$

where

$$\begin{aligned}
\partial_t z &= -y - \left\langle U^x \right\rangle(t,y) \\
\partial_t v &= \frac{1}{t}\left[\left\langle U^x \right\rangle(t,y) - \frac{1}{t}\int_0^t \left\langle U^x \right\rangle(s,y) ds\right] \\
\partial_y v &= 1 - \frac{1}{t}\int_0^t \langle \omega \rangle(s,y) ds \\
\partial_{yy} v &= -\frac{1}{t}\int_0^t \partial_y \langle \omega \rangle (s,y) ds.
\end{aligned}$$

Expressing $[\partial_t v](t, v) = \partial_t v(t,y)$, $v'(t, v) = \partial_y v(t,y)$ and $v''(t, v) = \partial_{yy} v(t,y)$, we get the following evolution equation for f,

$$\begin{aligned}
\partial_t f + [\partial_t v]\partial_v f + \partial_t z \partial_z f = -y\partial_z f + v'[\partial_v \phi + \partial_z \phi \partial_v z - \partial_z \phi \partial_v z]\partial_z f \\
- v'\partial_z \phi \partial_v f.
\end{aligned}$$

Using the definition of $\partial_t z$ and the Biot-Savart law to transform $\langle U^x \rangle$ to $-v'\partial_v \langle \phi \rangle$ in the new variables, this becomes

$$\partial_t f - \left(v'\partial_v\left(\phi - \langle \phi \rangle\right)\right)\partial_z f + \left([\partial_t v] + v'\partial_z \phi\right)\partial_v f = 0.$$

The Biot-Savart law also gets transformed into:

$$f = \partial_{zz}\phi + \left(v'\right)^2(\partial_v - t\partial_z)^2\phi + v''(\partial_v - t\partial_z)\phi = \Delta_t \phi. \tag{2.10}$$

The original 2D Euler system (1.1) is expressed as

$$\begin{cases}
\partial_t f + u \cdot \nabla_{z,v} f = 0, \\
u = (0, [\partial_t v]) + v'\nabla^{\perp}_{z,v} P_{\neq 0}\phi, \\
\phi = \Delta_t^{-1}[f].
\end{cases} \tag{2.11}$$

It what follows we will write $\nabla_{z,v} = \nabla$ and specify when other variables are used. Next we transform the momentum equation to allow us to express $[\partial_t v]$ in a form amenable to estimates. Denoting $\tilde{u}(t,z,v) = U^x(t,x,y)$ and $p(t,z,v) = P(t,x,y)$ we have by the same derivation on f,

$$\partial_t \tilde{u} + [\partial_t v]\partial_v \tilde{u} + \partial_z P_{\neq 0}\phi + v' \nabla^\perp P_{\neq 0}\phi \cdot \nabla \tilde{u} = -\partial_z p.$$

Taking averages in z we isolate the zero mode of the velocity field,

$$\partial_t \tilde{u}_0 + [\partial_t v]\partial_v \tilde{u}_0 + v' \left\langle \nabla^\perp P_{\neq 0}\phi \cdot \nabla \tilde{u} \right\rangle = 0. \tag{2.12}$$

Finally, one can express v' and $[\partial_t v]$ as solutions to a system of PDE in the (t,v) variables coupled to (2.11) (see Section 8.1 below for a detailed derivation):

$$\partial_t \big(t (v' - 1) \big) + [\partial_t v] t \partial_v v' = -f_0 \tag{2.13a}$$

$$\partial_t [\partial_t v] + \frac{2}{t}[\partial_t v] + [\partial_t v]\partial_v [\partial_t v] = -\frac{v'}{t}\left\langle \nabla^\perp P_{\neq 0}\phi \cdot \nabla \tilde{u} \right\rangle \tag{2.13b}$$

$$v''(t,v) = v'(t,v)\partial_v v'(t,v). \tag{2.13c}$$

Note that to leading order in ϵ, one can express $v' - 1$ as a time average of $-f_0$. Note also that we have a simple expression for $\partial_v \tilde{u}_0$ from the Biot-Savart law:

$$\partial_v \tilde{u}_0(t,v) = \frac{1}{v'(t,v)}\partial_y U_0^x(t,y) = -\frac{1}{v'(t,v)}\omega_0(t,y) = -\frac{1}{v'(t,v)} f_0(t,v). \tag{2.14}$$

Given a priori estimates on the system (2.11), (2.13), we can recover estimates on the original system (1.1) by the inverse function theorem as long as $v' - 1$ remains sufficiently small (see Section 2.4). Compared to the original system (1.1), the system (2.11), (2.13) appears much more complicated and nonlinear. Indeed, u is not divergence free and the dependence of ϕ on f through Δ_t is significantly more subtle than in the original variables. The main advantage of (2.11) is that u formally has an integrable decay, indeed, we will see that if one is willing to pay four derivatives, the decay rate is formally $O(t^{-2}\log t)$ (the decay we deduce is not quite as sharp).

2.3. *Main energy estimate*

In light of the previous section, our goal is to control solutions to (2.11) uniformly in a suitable norm as $t \to \infty$. The key idea we use for this is the carefully designed time-dependent norm written as

$$\left\| A(t,\nabla) f \right\|_2^2 = \sum_k \int_\eta \left| A_k(t,\eta) \hat{f}_k(t,\eta) \right|^2 d\eta.$$

The multiplier A has several components,

$$\mathrm{A}_k(t,\eta) = e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \mathrm{J}_k(t,\eta).$$

The index $\lambda(t)$ is the bulk Gevrey$-\frac{1}{s}$ regularity and will be chosen to satisfy

$$\lambda(t) = \frac{3}{4}\lambda_0 + \frac{1}{4}\lambda', \quad t \le 1 \tag{2.15a}$$

$$\dot\lambda(t) = -\frac{\delta_\lambda}{\langle t\rangle^{2\tilde q}}\big(1+\lambda(t)\big), \quad t > 1 \tag{2.15b}$$

where $\delta_\lambda \approx \lambda_0 - \lambda'$ is a small parameter that ensures $\lambda(t) > \lambda_0/2 + \lambda'/2$ and $1/2 < \tilde q \le s/8 + 7/16$ is a parameter chosen by the proof. The reason for (2.15a) is to account for the behavior of the solution on the time-interval [0, 1]; see Lemma 2.1 for this relatively minor detail. The use of a time-varying index of regularity is classical, for example the Cauchy-Kovalevskaya local existence theorem of Nirenberg [72, 73]. For more directly relevant works which use norms of this type, see [24, 25, 33, 49, 54, 70]. Let us remark here that to study analytic data, $s = 1$, we would need to add an additional Gevrey-$\frac{1}{\beta}$ correction to A with $1/2 < \beta < 1$ as an intermediate regularity so that we may take advantage of certain beneficial properties of Gevrey spaces; see for example Lemma A.3. In this case, the analytic regularity would simply be propagated more or less passively through the proof. Using the same idea, we may assume without loss of generality that s is close to $1/2$ (say $s < 2/3$), which simplifies some of the technical details but is not essential. The Sobolev correction with $\sigma > 12$ fixed is included mostly for technical convenience so we may easily quantify loss of derivatives without disturbing the index of regularity. We will also use the slightly stronger multiplier $\mathrm{A}^{\mathrm{R}}(t,\eta)$ that satisfies $\mathrm{A}^{\mathrm{R}}(t,\eta) \ge \mathrm{A}_0(t,\eta)$ to control the coefficients v' and v''; see (3.10) below for the definition.

The main multiplier for dealing with the Orr mechanism and the associated nonlinear growth is

$$\mathrm{J}_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)} + e^{\mu|k|^{1/2}}, \tag{2.16}$$

where $w_k(t,\eta)$ is constructed in Section 3 and describes the expected 'worst-case' growth due to nonlinear interactions at the critical times. What will be important is that J imposes more regularity on modes which satisfy $t \sim \frac{\eta}{k}$ (the 'resonant modes') than those that do not (the 'non-resonant modes'). The multiplier J replaces growth in time by controlled loss of regularity and is reminiscent of the notion of losing regularity estimates used in [3, 26]. One of the main differences is that here we have to be more precise in the sense that the loss of regularity occurs for different frequencies during different time intervals.

With this special norm, we can define our main energy:

$$\mathrm{E}(t) = \frac{1}{2}\big\|\mathrm{A}(t)f(t)\big\|_2^2 + \mathrm{E}_v(t), \tag{2.17}$$

where, for some constants K_v, K_D depending only on s, λ, λ' fixed by the proof,

$$E_v(t) = \langle t \rangle^{2+2s} \left\| \frac{A}{\langle \partial_v \rangle^s} v' \partial_v [\partial_t v](t) \right\|_2^2 + \langle t \rangle^{4-K_D \epsilon} \left\| [\partial_t v](t) \right\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 + \frac{1}{K_v} \left\| A^R \left(v' - 1 \right)(t) \right\|_2^2. \tag{2.18}$$

In a sense, there are two coupled energy estimates: the one on Af and the one on E_v. The latter quantity is encoding information about the coordinate system, or equivalently, the evolution of the background shear flow. It turns out $v'\partial_v[\partial_t v]$ is a physical quantity that measures the convergence of the x-averaged vorticity to its time average (see (8.5) in Section 8.1) and satisfies a useful PDE (see (8.9) in Section 8.2). It will be convenient to get two separate estimates on $[\partial_t v]$ as opposed to just one ($[\partial_t v]$ is essentially measuring how rapidly the x-averaged velocity is converging to its time average).

By the well-posedness theory for 2D Euler in Gevrey spaces [9, 32, 33, 49, 54] we may safely ignore the time interval (say) $[0, 1]$ by further restricting the size of the initial data. That is, we have the following lemma; see Section A.3 for a sketch of the proof.

Lemma **2.1.** — *For all $\epsilon > 0$, $s > 1/2$ and $\lambda_0 > \lambda' > 0$, there exists an $\epsilon' > 0$ such that if $\|\omega_{in}\|_{\mathcal{G}^{\lambda_0}} < \epsilon'$ and $\int |y\omega_{in}| dxdy < \epsilon'$, then* $\sup_{t \in [0,1]} \|f(t)\|_{\mathcal{G}^{3\lambda_0/4+\lambda'/4,\sigma}} < \epsilon$, $E(1) < \epsilon^2$, $\sup_{t\in[0,1]} \|1 - v'(t)\|_\infty < 6/10$.

The goal is next to prove by a continuity argument that this energy $E(t)$ (together with some related quantities) is uniformly bounded for all time if ϵ is sufficiently small. We define the following controls referred to in the sequel as the bootstrap hypotheses for $t \geq 1$,

(B1) $E(t) \leq 4\epsilon^2$;
(B2) $\|v' - 1\|_\infty \leq \frac{3}{4}$
(B3) 'CK' integral estimates (for 'Cauchy-Kovalevskaya'):

$$\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + K_v^{-1}\left(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau) \right) + K_v^{-1} \sum_{i=1}^{2} \left(CCK_w^i(\tau) + CCK_\lambda^i(\tau) \right) \Big] d\tau \leq 8\epsilon^2.$$

The CK terms above that appear without the K_v^{-1} prefactor arise from the time derivatives of $A(t)$ and are naturally controlled by the energy estimates we are making. The others are related quantities that are controlled separately in Proposition 2.5 below. These both will be defined below when discussing the energy estimates.

By Lemma 2.1, for the rest of the proof we may focus on times $t \geq 1$. Let I_E be the connected set of times $t \geq 1$ such that the bootstrap hypotheses (B1–B3) are all satisfied. We will work on regularized solutions for which we know $E(t)$ takes values continuously in time, and hence I_E is a closed interval $[1, T^\star]$ with $T^\star > 1$. The bootstrap is complete if we show that I_E is also open, which is the purpose of the following proposition, the proof of which constitutes the majority of this work.

Proposition **2.1** *(Bootstrap). — There exists an* $\epsilon_0 \in (0, 1/2)$ *depending only on* λ, λ', s *and* σ *such that if* $\epsilon < \epsilon_0$*, and on* $[1, T^\star]$ *the bootstrap hypotheses* (B1)–(B3) *hold, then for* $\forall t \in [1, T^\star]$,

1. $E(t) < 2\epsilon^2$,
2. $\|1 - v'\|_\infty < \frac{5}{8}$,
3. *and the CK controls satisfy:*

$$\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + K_v^{-1}\big(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\big) + K_v^{-1}\sum_{i=1}^{2}\big(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\big)\Big] d\tau \leq 6\epsilon^2,$$

from which it follows that $T^\star = +\infty$.

The remainder of the section is devoted to the proof of Proposition 2.1, the primary step being to show that on $[1, T^\star]$, we have

$$E(t) + \frac{1}{2}\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + K_v^{-1}\big(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\big) + K_v^{-1}\sum_{i=1}^{2}\big(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\big)\Big] d\tau \leq E(1) + K\epsilon^3 \tag{2.19}$$

for some constant K which is independent of ϵ and $T^\star$. If ϵ is sufficiently small then (2.19) implies Proposition 2.1. Indeed, the control $\|1 - v'\| < 5/8$ is an immediate consequence of (B1) by Sobolev embedding for ϵ sufficiently small.

To prove (2.19), it is natural to compute the time evolution of $E(t)$.

$$\frac{d}{dt}E(t) = \frac{1}{2}\frac{d}{dt}\int |Af|^2 dx + \frac{d}{dt}E_v(t).$$

The first contribution is of the form

$$\frac{1}{2}\frac{d}{dt}\int |\mathrm{A}f|^2 dx = -\mathrm{CK}_\lambda - \mathrm{CK}_w - \int \mathrm{A}f\,\mathrm{A}(u\cdot\nabla f)dx, \tag{2.20}$$

where the CK stands for 'Cauchy-Kovalevskaya' since these three terms arise from the progressive weakening of the norm in time, and are expressed as

$$\mathrm{CK}_\lambda = -\dot{\lambda}(t)\left\| |\nabla|^{s/2}\mathrm{A}f\right\|_2^2 \tag{2.21a}$$

$$\mathrm{CK}_w = \sum_k \int \frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)} e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)}\mathrm{A}_k(t,\eta)\left|\hat{f}_k(t,\eta)\right|^2 d\eta. \tag{2.21b}$$

In what follows we define

$$\tilde{\mathrm{J}}_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)}, \tag{2.22a}$$

$$\tilde{\mathrm{A}}_k(t,\eta) = e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \tilde{\mathrm{J}}_k(t,\eta). \tag{2.22b}$$

Note that $\tilde{\mathrm{A}} \leq \mathrm{A}$ and if $|k| \leq |\eta|$ then $\mathrm{A} \lesssim \tilde{\mathrm{A}}$.

Strictly speaking, equality (2.20) is not rigorous since it involves a derivative of $\mathrm{A}f$, which is not a priori well-defined. To make this calculation rigorous, we have first to approximate the initial data of (1.1) by (for instance) analytic initial data and use that the global solutions of (1.1) stay analytic for all time (see [9, 32, 33]). Hence, we can perform all calculations on these solutions with regularized initial data and then perform a passage to the limit to infer that (2.19) still holds. Similarly, the bootstrap is performed on these regularized solutions for which $\mathrm{E}(t)$ takes values continuously in time.

To treat the main term in (2.20), begin by integrating by parts, as in the techniques [33, 49, 54]

$$\int \mathrm{A}f\,\mathrm{A}(u\cdot\nabla f)dx = -\frac{1}{2}\int \nabla\cdot u|\mathrm{A}f|^2 dx + \int \mathrm{A}f\left[\mathrm{A}(u\cdot\nabla f) - u\cdot\nabla\mathrm{A}f\right]dx. \tag{2.23}$$

Notice that the relative velocity is not divergence free:

$$\nabla\cdot u = \partial_v[\partial_t v] + \partial_v v'\partial_z\phi.$$

The first term is controlled by the bootstrap hypothesis (B1). For the second term we use the 'lossy' elliptic estimate, Lemma 4.1, which shows that under the bootstrap hypotheses we have

$$\left\|\mathrm{P}_{\neq 0}\phi(t)\right\|_{\mathcal{G}^{\lambda(t),\sigma-3}} \lesssim \frac{\epsilon}{\langle t\rangle^2}. \tag{2.24}$$

Therefore, by Sobolev embedding, $\sigma > 5$ and the bootstrap hypotheses,

$$\left|\int \nabla \cdot u |\mathrm{A}f|^2 dx\right| \leq \|\nabla u\|_\infty \|\mathrm{A}f\|_2^2 \lesssim \frac{\epsilon}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}} \|\mathrm{A}f\|_2^2 \lesssim \frac{\epsilon^3}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}. \quad (\mathbf{2.25})$$

To handle the commutator, $\int \mathrm{A}f[\mathrm{A}(u\cdot\nabla f) - u\cdot\nabla \mathrm{A}f]dx$, we use a paraproduct decomposition (see e.g. [4, 14]). Precisely, we define three main contributions: *transport*, *reaction* and *remainder*:

$$\int \mathrm{A}f\left[\mathrm{A}(u\cdot\nabla f) - u\cdot\nabla \mathrm{A}f\right]dx = \frac{1}{2\pi}\sum_{\mathrm{N}\geq 8}\mathrm{T_N} + \frac{1}{2\pi}\sum_{\mathrm{N}\geq 8}\mathrm{R_N} + \frac{1}{2\pi}\mathcal{R}, \quad (\mathbf{2.26})$$

where (the factors of 2π are for future notational convenience)

$$\mathrm{T_N} = 2\pi\int \mathrm{A}f\left[\mathrm{A}(u_{<\mathrm{N}/8}\cdot\nabla f_{\mathrm{N}}) - u_{<\mathrm{N}/8}\cdot\nabla \mathrm{A}f_{\mathrm{N}}\right]dx$$

$$\mathrm{R_N} = 2\pi\int \mathrm{A}f\left[\mathrm{A}(u_{\mathrm{N}}\cdot\nabla f_{<\mathrm{N}/8}) - u_{\mathrm{N}}\cdot\nabla \mathrm{A}f_{<\mathrm{N}/8}\right]dx$$

$$\mathcal{R} = 2\pi\sum_{\mathrm{N}\in\mathbf{D}}\sum_{\frac{1}{8}\mathrm{N}\leq\mathrm{N}'\leq 8\mathrm{N}}\int \mathrm{A}f\left[\mathrm{A}(u_{\mathrm{N}}\cdot\nabla f_{\mathrm{N}'}) - u_{\mathrm{N}}\cdot\nabla \mathrm{A}f_{\mathrm{N}'}\right]dx.$$

Here $\mathrm{N}\in\mathbf{D} = \{\frac{1}{2}, 1, 2, 4, \ldots, 2^j, \ldots\}$ and g_{N} denotes the N-th Littlewood-Paley projection and $g_{<\mathrm{N}}$ means the Littlewood-Paley projection onto frequencies less than N (see Section A.1 for the Fourier analysis conventions we are taking). Formally, the paraproduct decomposition (2.26) represents a kind of 'linearization' for the evolution of higher frequencies around the lower frequencies. The terminology 'reaction' is borrowed from Mouhot and Villani [70] (see Section 9 for more information).

Controlling the transport contribution is the subject of Section 5, in which we prove:

Proposition **2.2** *(Transport). — Under the bootstrap hypotheses,*

$$\sum_{\mathrm{N}\geq 8}|\mathrm{T_N}| \lesssim \epsilon\mathrm{CK}_\lambda + \epsilon\mathrm{CK}_w + \frac{\epsilon^3}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}.$$

The proof of Proposition 2.2 uses ideas from the works of [33, 49, 54]. Since the velocity u is restricted to 'low frequency', we will have the available regularity required to apply (2.24). However, the methods of [33, 49, 54] do not adapt immediately since $\mathrm{J}_k(t,\eta)$ is imposing slightly different regularities to certain frequencies, which is problematic. Physically speaking, we need to ensure that resonant frequencies do not incur a very large growth due to nonlinear interactions with non-resonant frequencies (which are permitted to be slightly larger than the resonant frequencies). Controlling this imbalance is why CK_w appears in Proposition 2.2.

Controlling the reaction contribution in (2.26) is the subject of Section 6. Here we cannot apply (2.24), as an estimate on this term requires u in the highest norm on which we have control, and hence we have no regularity to spare. Physically, here in the reaction term is where the dangerous nonlinear effects are expressed and a great deal of precision is required to control them. In Section 6 we prove

Proposition **2.3** *(Reaction). — Under the bootstrap hypotheses,*

$$\sum_{N\geq 8}|R_N| \lesssim \epsilon CK_\lambda + \epsilon CK_w + \frac{\epsilon^3}{\langle t\rangle^{2-K_D\epsilon/2}} + \epsilon CK_\lambda^{v,1} + \epsilon CK_w^{v,1} + \epsilon\left\|\left\langle\frac{\partial_v}{t\partial_z}\right\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\right\|_2^2. \tag{2.27}$$

The $CK^{v,1}$ terms are defined below in (2.31). The first step to controlling the term in (2.27) involving ϕ is Proposition 2.4, proved in Section 4.2. This proposition treats Δ_t as a perturbation of $\partial_{zz}+(\partial_v-t\partial_z)^2$ and passes the multipliers in the last term of (2.27) onto f *and* the coefficients of Δ_t. Physically, these latter contributions are indicating the nonlinear interactions between the higher modes of f and the coefficients v', v'' (which involve time-averages of f_0 (2.13)).

Proposition **2.4** *(Precision elliptic control). — Under the bootstrap hypotheses,*

$$\left\|\left\langle\frac{\partial_v}{t\partial_z}\right\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\right\|_2^2 \lesssim CK_\lambda + CK_w + \epsilon^2\sum_{i=1}^{2} CCK_\lambda^i + CCK_w^i, \tag{2.28}$$

where the 'coefficient Cauchy-Kovalevskaya' terms are given by

$$CCK_\lambda^1 = -\dot{\lambda}(t)\left\||\partial_v|^{s/2}A^R\left(1-(v')^2\right)\right\|_2^2, \tag{2.29a}$$

$$CCK_w^1 = \left\|\sqrt{\frac{\partial_t w}{w}}A^R\left(1-(v')^2\right)\right\|_2^2, \tag{2.29b}$$

$$CCK_\lambda^2 = -\dot{\lambda}(t)\left\||\partial_v|^{s/2}\frac{A^R}{\langle\partial_v\rangle}v''\right\|_2^2, \tag{2.29c}$$

$$CCK_w^2 = \left\|\sqrt{\frac{\partial_t w}{w}}\frac{A^R}{\langle\partial_v\rangle}v''\right\|_2^2. \tag{2.29d}$$

The next step in the bootstrap is to provide good estimates on the coordinate system and the associated CK and CCK terms, a procedure that is detailed in Section 8. The following proposition provides controls on $v'-1$, the CCK terms arising in (2.29), the pair $[\partial_t v]$, $v'\partial_v[\partial_t v]$ and finally all of the $\mathrm{CK}^{v,i}$ terms. The norm defined by $\mathrm{A}^{\mathrm{R}}(t)$ is stronger than that defined by $\mathrm{A}(t)$, which we use to measure f. It turns out that we will be able to propagate this stronger regularity on $v'-1$ due to a time-averaging effect, derived via energy estimates on (2.13). By contrast, $[\partial_t v]$ is expected basically to have the regularity of $\tilde{u}_0$ and hence even (2.30b) has s fewer derivatives than expected. On the other hand, it has a significant amount of time decay, which near critical times can be converted into regularity.

Proposition **2.5** *(Coordinate system controls). — Under the bootstrap hypotheses, for ϵ sufficiently small and K_v sufficiently large there is a $\mathrm{K}>0$ such that*

$$\textbf{(2.30a)}\qquad \left\|\mathrm{A}^{\mathrm{R}}\big(v'-1\big)(t)\right\|_2^2+\frac{1}{2}\int_1^t\sum_{i=1}^{2}\mathrm{CCK}_w^i(\tau)d\tau+\frac{1}{2}\int_1^t\sum_{i=1}^{2}\mathrm{CCK}_\lambda^i(\tau)d\tau \le \frac{1}{2}\mathrm{K}_v\epsilon^2$$

$$\textbf{(2.30b)}\qquad \langle t\rangle^{2+2s}\left\|\frac{\mathrm{A}}{\langle\partial_v\rangle^s}v'\partial_v[\partial_t v]\right\|_2^2+\frac{1}{2}\int_1^t \mathrm{CK}_\lambda^{v,2}(\tau)+\mathrm{CK}_w^{v,2}(\tau)d\tau \le \epsilon^2+\mathrm{K}\epsilon^3$$

$$\textbf{(2.30c)}\qquad \langle t\rangle^{4-\mathrm{K}_{\mathrm{D}}\epsilon}\left\|[\partial_t v]\right\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2\le\epsilon^2+\mathrm{K}\epsilon^3$$

$$\textbf{(2.30d)}\qquad \int_1^t \mathrm{CK}_\lambda^{v,1}(\tau)+\mathrm{CK}_w^{v,1}(\tau)d\tau\le\frac{1}{2}\mathrm{K}_v\epsilon^2,$$

where the $\mathrm{CK}^{v,i}$ terms are given by

$$\textbf{(2.31a)}\qquad \mathrm{CK}_w^{v,2}(\tau)=\langle\tau\rangle^{2+2s}\left\|\sqrt{\frac{\partial_t w}{w}}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}v'\partial_v[\partial_t v](\tau)\right\|_2^2$$

$$\textbf{(2.31b)}\qquad \mathrm{CK}_\lambda^{v,2}(\tau)=\langle\tau\rangle^{2+2s}\big(-\dot{\lambda}(\tau)\big)\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}v'\partial_v[\partial_t v](\tau)\right\|_2^2$$

$$\textbf{(2.31c)}\qquad \mathrm{CK}_w^{v,1}(\tau)=\langle\tau\rangle^{2+2s}\left\|\sqrt{\frac{\partial_t w}{w}}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}[\partial_t v](\tau)\right\|_2^2$$

$$\textbf{(2.31d)}\qquad \mathrm{CK}_\lambda^{v,1}(\tau)=\langle\tau\rangle^{2+2s}\big(-\dot{\lambda}(\tau)\big)\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}[\partial_t v](\tau)\right\|_2^2.$$

Note that neither (2.30b) nor (2.30c) controls the other: at higher frequencies the former is stronger than the latter and at lower frequencies the opposite is true. One of the advantages of this scheme is that $v'\partial_v[\partial_t v]$ satisfies an equation that is simpler than $[\partial_t v]$ and so is easier to get good estimates on. Both (2.30b) and (2.30c) are linked to the convergence of the background shear flow; in particular, they rule out that the background flow oscillates or wanders due to nonlinear effects.

Finally we need to control the remainder term in (2.26). This is straightforward and is detailed in Section 7. There we prove

Proposition **2.6** *(Remainders). — Under the bootstrap hypotheses,*

$$\mathcal{R} \lesssim \frac{\epsilon^3}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}.$$

Collecting Propositions 2.2, 2.3, 2.4, 2.5, 2.6 with (2.26) and (2.25), we have finally (2.19) for ϵ sufficiently small with constants independent of both ϵ and $\mathrm{T}^\star$; hence for ϵ sufficiently small we may propagate the bootstrap control and prove Proposition 2.1.

2.4. *Conclusion of proof*

By Proposition 2.1 we have a global uniform bound on $\mathrm{E}(t)$, and therefore the uniform bounds

$$\begin{aligned}\|f(t)\|^2_{\mathcal{G}^{\lambda(t),\sigma}} + \langle t\rangle^4\|\mathrm{P}_{\neq 0}\phi\|^2_{\mathcal{G}^{\lambda(t),\sigma-3}} + \langle t\rangle^{4-\mathrm{K_D}\epsilon}\left\|[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda(t),\sigma-6}} \\ + \mathrm{K}_v^{-1}\left\|v'-1\right\|^2_{\mathcal{G}^{\lambda(t),\sigma}} \lesssim \epsilon^2.\end{aligned} \tag{2.32}$$

Define $\lambda_\infty = \lim_{t\to\infty}\lambda(t)$. By the method of characteristics and Sobolev embedding, Lemma 4.1 and (2.32) also imply the spatial localization $\int |vf(t,z,v)|dzdv \lesssim \epsilon$. Together with (2.32) it follows that

$$\int\left|\frac{v}{v'(t,v)}f(t,z,v)\right|dzdv \lesssim \epsilon,$$

which implies $\widehat{(v')^{-1}f}$ is Lipschitz continuous. Since $\int \omega dvdz = \int (v')^{-1}fdvdz = 0$, it follows that (using also the algebra property Lemma A.3),

$$\begin{aligned}&\int e^{2\lambda_\infty|\nabla|^s}\langle\nabla\rangle^{2\sigma}\left|\tilde{u}_0(t,v)\right|^2dv \\ &\quad\approx \int \frac{e^{2\lambda_\infty|\xi|^s}\langle\xi\rangle^{2\sigma}}{|\xi|^2}\left|\widehat{\left(v'\right)^{-1}f}(t,\xi)\right|^2d\xi \\ &\quad\lesssim \sup_{|\xi|\le 1}\left|\nabla_\xi\widehat{\left(v'\right)^{-1}f}(t,\xi)\right|^2 + \int_{|\xi|\ge 1} e^{2\lambda_\infty|\xi|^s}\langle\xi\rangle^{2\sigma-2}\left|\widehat{\left(v'\right)^{-1}f}(t,\xi)\right|^2d\xi\end{aligned} \tag{2.33}$$

$$\lesssim \epsilon^2 + \|f\|_{\mathcal{G}^{\lambda_\infty,\sigma-1}}\left(1 + \left\|v' - 1\right\|_{\mathcal{G}^{\lambda_\infty,\sigma-1}}\right)$$
$$\lesssim \epsilon^2.$$

Hence,

$$\|\tilde{u}_0\|_{\mathcal{G}^{\lambda(t),\sigma}} \lesssim \epsilon. \tag{2.34}$$

In order to deduce the convergence expressed in (1.5) and (1.7) we now undo the change of coordinates in v, switching to the more physically natural coordinates (z,y). Writing $h(t,z,y) = f(t,z,v) = \omega(t,x,y)$ and $\tilde{\psi}(t,z,y) = \phi(t,z,v)$ one derives from (1.1) as in Section 2.2,

$$\partial_t h + \nabla^\perp_{z,y} P_{\neq 0}\tilde{\psi} \cdot \nabla_{z,y} h = 0. \tag{2.35}$$

We remark that these coordinates are not the same as (2.2): indeed, $\tilde{\psi}$ and h do not satisfy (2.7b). In order to quantify the Gevrey regularity of $h(t,z,y) = f(t,z,v(t,y))$ we apply the composition inequality Lemma A.4 (with (A.15)) and hence we must show $v(t,y) - y$ is small in a suitable Gevrey class. Actually what we have a priori control on from (2.32) is $v'(t,v(t,y)) = \partial_y v(t,y)$. From the C^∞ inverse function theorem we may solve for $y = y(t,v)$, which implies the spatial control $\int |yh(t,z,y)|dzdy \lesssim \epsilon$ (and Remark 4). From (2.34) and (2.9), we also get at least the L^2 estimates $\|v(t,y) - y\|_2 + \|y(t,v) - v\|_2 \lesssim \epsilon$. To get control on the Gevrey regularity, we apply

$$\partial_v y(t,v) = \frac{1}{v'(t,v)} = \sum_{n=0}^{\infty}\left(1 - v'(t,v)\right)^n,$$

which implies together with Lemma A.3 and ϵ sufficiently small, for any $\lambda'_\infty \in (\lambda', \lambda_\infty)$:

$$\left\|y(t,v) - v\right\|_{\mathcal{G}^{\lambda'_\infty}} \lesssim \epsilon.$$

Therefore, by a Gevrey inverse function theorem (Lemma A.5 with $\alpha(v) = v - y(t,v)$ and $\beta(y) = v(t,y) - y$, which solves $\beta(y) = \alpha(y + \beta(y))$) followed by (A.15), for any $\lambda''_\infty \in (\lambda', \lambda'_\infty)$, we may choose ϵ sufficiently small such that $\|v(t,y) - y\|_{\mathcal{G}^{\lambda''_\infty}} \lesssim \epsilon$. Together with Lemma A.4, (A.15) and (2.32) this implies for any $\lambda'''_\infty \in (\lambda', \lambda''_\infty)$ (adjusting ϵ if necessary),

$$\left\|h(t)\right\|_{\mathcal{G}^{\lambda'''_\infty}} + \langle t\rangle^2 \left\|P_{\neq 0}\tilde{\psi}(t)\right\|_{\mathcal{G}^{\lambda'''_\infty}} \lesssim \epsilon. \tag{2.36}$$

Therefore by integrating (2.35), we define f_∞ by the absolutely convergent integral

$$f_\infty = h(1) - \int_1^\infty \nabla^\perp_{z,y} P_{\neq 0}\tilde{\psi}(s) \cdot \nabla_{z,y} h(s) ds. \tag{2.37}$$

More precisely, since $\lambda' < \lambda'''_\infty$, Minkowski, (A.10) and (A.12) imply

$$\left\|h(t) - f_\infty\right\|_{\mathcal{G}^{\lambda'}} = \left\|\int_t^\infty \nabla^\perp_{z,y} P_{\neq 0}\tilde{\psi}(\tau) \cdot \nabla_{z,y} h(\tau) d\tau\right\|_{\mathcal{G}^{\lambda'}} \lesssim \frac{\epsilon^2}{\langle t\rangle}. \tag{2.38}$$

By the definition of z and the phase Φ (1.6), this is equivalent to (1.5). Shifting in z, (1.7b) and (1.7c) follow from the decay estimates on ψ.

Finally, an argument similar to that used to derive (2.38) can be applied to $U_0^x(t,y) = \tilde{u}_0(t,v)$, which satisfies the following from (1.2) denoting $\tilde{U}(t,z,y) = \tilde{u}(t,z,v) = U^x(t,x,y)$:

$$\partial_t U_0^x + \left\langle \nabla_{z,y}^{\perp} P_{\neq 0}\tilde{\psi} \cdot \nabla_{z,y}\tilde{U} \right\rangle = 0. \tag{2.39}$$

By (8.39) below, it follows that $\|P_{\neq 0}\nabla_{z,y}\tilde{U}(t)\|_{\mathcal{G}^{\lambda'''_\infty}} \lesssim \epsilon\langle t\rangle^{-1}$ by the argument used to deduce (2.36). This, together with (2.36), implies (1.7a) by integrating (2.39) as in (2.38) (see Section A.4 for more discussion).

3. Growth mechanism and construction of A

3.1. *Construction of w*

3.1.1. *Formal derivation of toy model*

From Section 2, we see that the basic challenge to the proof of Theorem 1 is controlling the regularity of solutions to (2.11). Since we must pay regularity to deduce decay on the velocity u, it is natural to consider the frequency interactions in the product $u \cdot \nabla f$ with the frequencies of u much larger than f, which corresponds to the "reaction" term in (2.26) above. This leads us to study a simpler model

$$\partial_t f = -u \cdot \nabla f_{lo}, \tag{3.1}$$

where f_{lo} is a given function that we think of as much smoother than f. As we see from (2.11), u consists of several terms, however let us focus on the term we think should be the worst and also ignore the v', further reducing to the linear problem:

$$\partial_t f = \partial_v P_{\neq 0}\phi \partial_z f_{lo}.$$

This contribution was chosen as Δ_t loses ellipticity in v, not z. Suppose that instead of $f = \Delta_t\phi$, we had $f = \partial_{zz}\phi + (\partial_y - t\partial_z)^2\phi$ as in (1.4), then on the Fourier side:

$$\partial_t \hat{f}(t,k,\eta) = \frac{1}{2\pi}\sum_{l\neq 0}\int_\xi \frac{\xi(k-l)}{l^2 + |\xi - lt|^2}\hat{f}(l,\xi)\hat{f}_{lo}(t,k-l,\eta-\xi)d\xi.$$

Since f_{lo} weakens interactions between well-separated frequencies, let us consider a discrete model with η as a fixed parameter:

$$\partial_t \hat{f}(t,k,\eta) = \frac{1}{2\pi}\sum_{l\neq 0}\frac{\eta(k-l)}{l^2 + |\eta - lt|^2}\hat{f}(l,\eta)f_{lo}(t,k-l,0). \tag{3.2}$$

As time advances this system of ODEs will go through resonances or "critical times" given by $t = \frac{\eta}{k}$, at which time the k mode strongly forces the others. If $|\eta| k^{-2} \ll 1$ then the critical time does not have a serious detriment and so focus on the case $|\eta| k^{-2} > 1$. The scenario we are most concerned with is a high-to-low cascade in which the k mode has a strong effect at time η/k that excites the $k-1$ mode which has a strong effect at time $\eta/(k-1)$ that excites the $k-2$ mode and so on. The echoes observed experimentally in [94, 95] arise from an effect similar to this [91, 92]. Now focus near one critical time η/k on a time interval of length roughly η/k^2 and consider the interaction between the mode k and a nearby mode l with $l \neq k$. If one takes absolute values and retains only the leading order terms, then this reduces to the much simpler system of two ODEs (thinking of $f_{lo} = O(\kappa)$) which we refer to as the *toy model*:

$$\partial_t f_{\mathrm{R}} = \kappa \frac{k^2}{|\eta|} f_{\mathrm{NR}}, \tag{3.3a}$$

$$\partial_t f_{\mathrm{NR}} = \kappa \frac{|\eta|}{k^2 + |\eta - kt|^2} f_{\mathrm{R}}, \tag{3.3b}$$

where we think of f_{R} as being the evolution of the k mode and f_{NR} being the evolution of a nearby mode l with $l \neq k$. The factor $k^2/|\eta|$ in the ODE for f_{R} is an upper bound on the strongest interaction a non-resonant mode, for example the $k-1$ mode, can have with the resonant mode. Obviously (3.3) represents a major simplification compared to (3.1), however it will be sufficient to prove Theorem 1. See Section 9 for a discussion and speculation on whether it is possible to improve Theorem 1 by using a model closer to (3.1).

Remark **1.** — The toy model dropped several terms, one of which being a weak self-interaction. For example, one could replace the second equation in (3.3) by

$$\partial_\tau f_{\mathrm{NR}} = \kappa \frac{|\eta|}{k^2 + |\eta - kt|^2} f_{\mathrm{R}} + \kappa \frac{k^2}{|\eta|} f_{\mathrm{NR}}.$$

However, this does not significantly change the worse possible growth predicted by the model and is not necessary for the proof of Theorem 1. Actually, the proof of Theorem 1 strongly suggests that the most substantial simplifications in the derivation of (3.3) was the replacement of Δ_t by $\partial_z^2 + (\partial_v - t\partial_z)^2$.

3.1.2. *Construction of w*

For simplicity of notation in this section we usually take $\eta, k > 0$ but the work applies equally well to $\eta, k < 0$ ($w(t, \eta)$ will depend only on $|\eta|$). Note that modes where $\eta k < 0$ do not have resonances for positive times. Keeping with the intuition from the derivation of (3.3), in this section we will think of η as a fixed parameter and time varying. Here we will be concerned with critical times for which we have $\eta/k^2 \geq 1$. Accordingly, in this section we will use $I_{k,\eta}$ (see Section 1.2) to denote any resonant interval with $\eta/k^2 \geq 1$, in contrast to subsequent sections where this notation is restricted further.

A key feature of our methods is how the toy model is used to construct a norm which precisely matches the estimated worst-case behavior that the reaction terms create, done by choosing $w_k(t,\eta)$ to be an approximate solution to (3.3). First we have the following (easy to check) Proposition.

Proposition **3.1.** — *Let* $\tau = t - \frac{\eta}{k}$ *and consider the solution* $(f_{\mathrm{R}}(\tau), f_{\mathrm{NR}}(\tau))$ *to* (3.3) *with* $f_{\mathrm{R}}(-\frac{\eta}{k^2}) = f_{\mathrm{NR}}(-\frac{\eta}{k^2}) = 1$. *There exists a constant* C *such that for all* $\kappa < 1/2$ *and* $\frac{\eta}{k^2} \geq 1$,

$$f_{\mathrm{R}}(\tau) \leq \mathrm{C}\left(\frac{k^2}{\eta}\left(1+|\tau|\right)\right)^{-\mathrm{C}\kappa}, \quad -\frac{\eta}{k^2} \leq \tau \leq 0,$$

$$f_{\mathrm{NR}}(\tau) \leq \mathrm{C}\left(\frac{k^2}{\eta}\left(1+|\tau|\right)\right)^{-\mathrm{C}\kappa-1}, \quad -\frac{\eta}{k^2} \leq \tau \leq 0,$$

$$f_{\mathrm{R}}(\tau) \leq \mathrm{C}\left(\frac{\eta}{k^2}\right)^{\mathrm{C}\kappa}\left(1+|\tau|\right)^{1+\mathrm{C}\kappa}, \quad 0 \leq \tau \leq \frac{\eta}{k^2},$$

$$f_{\mathrm{NR}}(\tau) \leq \mathrm{C}\left(\frac{\eta}{k^2}\right)^{\mathrm{C}\kappa+1}\left(1+|\tau|\right)^{\mathrm{C}\kappa}, \quad 0 \leq \tau \leq \frac{\eta}{k^2}.$$

For the remainder of the paper we fix κ *such that* $3/2 < (1+2\mathrm{C}\kappa) < 10$.

Remark **2.** — It is important to notice that over the whole interval $[-\frac{\eta}{k^2}, \frac{\eta}{k^2}]$, both f_{R} and f_{NR} are at most amplified by roughly the same factor $\mathrm{C}(\frac{\eta}{k^2})^{1+2\mathrm{C}\kappa}$. Over the interval $[-\frac{\eta}{k^2}, 0]$, f_{NR} is amplified at most by $\mathrm{C}(\frac{\eta}{k^2})^{1+\mathrm{C}\kappa}$ and f_{R} is amplified at most by $\mathrm{C}(\frac{\eta}{k^2})^{\mathrm{C}\kappa}$. Whereas, over the interval $[0, \frac{\eta}{k^2}]$, f_{NR} is amplified at most by $\mathrm{C}(\frac{\eta}{k^2})^{\mathrm{C}\kappa}$ and f_{R} is amplified at most by $\mathrm{C}(\frac{\eta}{k^2})^{1+\mathrm{C}\kappa}$. Near the critical time, the imbalance between f_{NR} and f_{R} is the largest—in particular, the resonant mode f_{R} is a factor of $\frac{\eta}{k^2}$ less than f_{NR} at this time. However by the end of the interval, the total growth of the resonant and non-resonant modes are comparable. The fact that f_{R} and f_{NR} are amplified the same over that interval will simplify the construction of w; specifically, we will be able to assign w_{R} and w_{NR} below to agree at the end points of the critical interval.

On each interval $I_{k,\eta}$, growth of the resonant mode (k,η) will be modeled by w_{R} and the rest of the modes (which are non-resonant) will be modeled by w_{NR}. By Proposition 3.1, we will be able to choose w such that the total growth of w_{R} and w_{NR} exactly agree.

The construction is done backward in time, starting with $k=1$. For $t \in \mathrm{I}_{k,\eta}$ and $\tau = t - \frac{\eta}{k}$, we will choose $(w_{\mathrm{NR}}, w_{\mathrm{R}})$ such that over the interval $\mathrm{I}_{k,\eta}$ they approximately satisfy (3.3):

$$\begin{aligned} \partial_\tau w_{\mathrm{R}} &\approx \kappa \frac{k^2}{\eta} w_{\mathrm{NR}}, \\ \partial_\tau w_{\mathrm{NR}} &\approx \kappa \frac{\eta}{k^2(1+\tau^2)} w_{\mathrm{R}}. \end{aligned} \tag{3.4}$$

We first construct the non-resonant component w_{NR} and then explain how we should modify it over each interval $\mathrm{I}_{k,\eta}$ to construct w_{R}.

Let w_{NR} be a non-decreasing function of time with $w_{\mathrm{NR}}(t,\eta) = 1$ for $t \geq 2\eta$. For definiteness, we remark here that for $|\eta| < 1$, $w_{\mathrm{NR}}(t,\eta) = 1$, which will be a consequence of the ensuing the definition. Hence we may safely assume $|\eta| > 1$ for the duration of the section. For $k \geq 1$, we assume that $w_{\mathrm{NR}}(t_{k-1,\eta}, \eta)$ was computed. To compute w_{NR} on the interval $\mathrm{I}_{k,\eta}$, we use the growth predicted by Proposition 3.1: for $k = 1, 2, 3, \ldots, \mathrm{E}(\sqrt{\eta})$, we define

$$w_{\mathrm{NR}}(t,\eta) = \left(\frac{k^2}{\eta}\left[1 + b_{k,\eta}\left|t - \frac{\eta}{k}\right|\right]\right)^{\mathrm{C}\kappa} w_{\mathrm{NR}}(t_{k-1,\eta}, \eta), \quad \forall t \in \mathrm{I}^{\mathrm{R}}_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right], \tag{3.5a}$$

$$w_{\mathrm{NR}}(t,\eta) = \left(1 + a_{k,\eta}\left|t - \frac{\eta}{k}\right|\right)^{-1-\mathrm{C}\kappa} w_{\mathrm{NR}}\left(\frac{\eta}{k}, \eta\right), \quad \forall t \in \mathrm{I}^{\mathrm{L}}_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right]. \tag{3.5b}$$

The constant $b_{k,\eta}$ is chosen to ensure that $\frac{k^2}{\eta}[1 + b_{k,\eta}|t_{k-1,\eta} - \frac{\eta}{k}|] = 1$, hence for $k \geq 2$, we have

$$b_{k,\eta} = \frac{2(k-1)}{k}\left(1 - \frac{k^2}{\eta}\right) \tag{3.6}$$

and $b_{1,\eta} = 1 - 1/\eta$. Similarly, $a_{k,\eta}$ is chosen to ensure $\frac{k^2}{\eta}[1 + a_{k,\eta}|t_{k,\eta} - \frac{\eta}{k}|] = 1$, which implies

$$a_{k,\eta} = \frac{2(k+1)}{k}\left(1 - \frac{k^2}{\eta}\right). \tag{3.7}$$

Hence, $w_{\mathrm{NR}}(\frac{\eta}{k}, \eta) = w_{\mathrm{NR}}(t_{k-1,\eta}, \eta)(\frac{k^2}{\eta})^{\mathrm{C}\kappa}$ and $w_{\mathrm{NR}}(t_{k,\eta}, \eta) = w_{\mathrm{NR}}(t_{k-1,\eta}, \eta)(\frac{k^2}{\eta})^{1+2\mathrm{C}\kappa}$. The choice of $a_{k,\eta}$ and $b_{k,\eta}$ was made to ensure that the ratio between $w_{\mathrm{NR}}(t_{k,\eta}, \eta)$ and

$w_{\mathrm{NR}}(t_{k-1,\eta}, \eta)$ is exactly $(\frac{k^2}{\eta})^{1+2C\kappa}$. Finally, we take w_{NR} to be constant on the interval $[0, t_{\mathrm{E}(\sqrt{\eta}),\eta}]$, namely $w_{\mathrm{NR}}(t, \eta) = w(t_{\mathrm{E}(\sqrt{\eta}),\eta}, \eta)$ for $t \in [0, t_{\mathrm{E}(\sqrt{\eta}),\eta}]$. Note that we always have $0 \leq b_{k,\eta} < 1$ and $0 \leq a_{k,\eta} < 4$, but that $a_{k,\eta}$ and $b_{k,\eta}$ approach 0 when k approaches $\mathrm{E}(\sqrt{\eta})$. This will present minor technical difficulties in the sequel since this implies that $\partial_t w$ vanishes near this time and hence a loss of the lower bounds in (3.4).

On each interval $\mathrm{I}_{k,\eta}$, we define $w_{\mathrm{R}}(t, \eta)$ by

$$w_{\mathrm{R}}(t, \eta) = \frac{k^2}{\eta}\left(1 + b_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{\mathrm{NR}}(t, \eta), \quad \forall t \in \mathrm{I}^{\mathrm{R}}_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right], \tag{3.8a}$$

$$w_{\mathrm{R}}(t, \eta) = \frac{k^2}{\eta}\left(1 + a_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{\mathrm{NR}}(t, \eta), \quad \forall t \in \mathrm{I}^{\mathrm{L}}_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right]. \tag{3.8b}$$

Due to the choice of $b_{k,\eta}$ and $a_{k,\eta}$, we get that $w_{\mathrm{R}}(t_{k,\eta}, \eta) = w_{\mathrm{NR}}(t_{k,\eta}, \eta)$ and $w_{\mathrm{R}}(\frac{\eta}{k}, \eta) = \frac{k^2}{\eta} w_{\mathrm{NR}}(\frac{\eta}{k}, \eta)$.

To define the full $w_k(t, \eta)$, we then have

$$w_k(t, \eta) = \begin{cases} w_k(t_{\mathrm{E}(\sqrt{\eta}),\eta}, \eta) & t < t_{\mathrm{E}(\sqrt{\eta}),\eta} \\ w_{\mathrm{NR}}(t, \eta) & t \in [t_{\mathrm{E}(\sqrt{\eta}),\eta}, 2\eta] \setminus \mathrm{I}_{k,\eta} \\ w_{\mathrm{R}}(t, \eta) & t \in \mathrm{I}_{k,\eta} \\ 1 & t \geq 2\eta. \end{cases} \tag{3.9}$$

Since w_{R} and w_{NR} agree at the end-points of $\mathrm{I}_{k,\eta}$, $w_k(t, \eta)$ is Lipschitz continuous in time. This completes the construction of w which appears in the J defined above (2.16).

We also define $\mathrm{J}^{\mathrm{R}}(t, \eta)$ and $\mathrm{A}^{\mathrm{R}}(t, \eta)$ to assign resonant regularity at *every* critical time:

$$\begin{aligned} \mathrm{J}^{\mathrm{R}}(t, \eta) &= \begin{cases} e^{\mu|\eta|^{1/2}} w_{\mathrm{R}}^{-1}(t_{\mathrm{E}(\sqrt{\eta}),\eta}, \eta) & t < t_{\mathrm{E}(\sqrt{\eta}),\eta} \\ e^{\mu|\eta|^{1/2}} w_{\mathrm{R}}^{-1}(t, \eta) & t \in [t_{\mathrm{E}(\sqrt{\eta}),\eta}, 2\eta] \\ e^{\mu|\eta|^{1/2}} & t \geq 2\eta, \end{cases} \\ \mathrm{A}^{\mathrm{R}}(t, k, \eta) &= e^{\lambda(t)|\eta|^s} \langle \eta \rangle^{\sigma} \mathrm{J}^{\mathrm{R}}(t, \eta). \end{aligned} \tag{3.10}$$

We can easily see from (3.8) that $\mathrm{A}^{\mathrm{R}}(t, \eta) \geq \mathrm{A}_0(t, \eta)$ and since the zero frequency is always non-resonant from (3.9), we see that near the critical times, A^{R} can be as much as a factor of $|\eta|$ larger.

3.1.3. *Total growth of* $w_k(t, \eta)$

The following lemma shows that the toy model predicts a growth of high frequencies which amounts to a loss of Gevrey-2 regularity, which is the primary origin of the restriction $s > 1/2$ in Theorem 1. C. Mouhot and C. Villani have informed the authors that a heuristic similar to that used in Section 7 of [70] can also be used to predict the same Gevrey-2 regularity requirement, though note that the primary purpose of the toy model (3.3) is to provide precise mode-by-mode information about how this loss can occur.

Lemma **3.1** *(Growth of w). — For $\eta > 1$, we have for $\mu = 4(1 + 2C\kappa)$,*

$$\frac{1}{w_k(0,\eta)} = \frac{1}{w_k(t_{E(\sqrt{\eta}),\eta},\eta)} \sim \frac{1}{\eta^{\mu/8}} e^{\frac{\mu}{2}\sqrt{\eta}}. \tag{3.11}$$

Here $\sim$ is in the sense of asymptotic expansion.

Proof. — Counting the growth over each interval implied by (3.9) gives the exact formula:

$$\frac{1}{w_k(0,\eta)} = \left(\frac{\eta}{N^2}\right)^c \left(\frac{\eta}{(N-1)^2}\right)^c \cdots \left(\frac{\eta}{1^2}\right)^c = \left[\frac{\eta^N}{(N!)^2}\right]^c,$$

where $c = 1 + 2C\kappa$. Recall Stirling's formula $N! \sim \sqrt{2\pi N}(N/e)^N$, which implies

$$\left(w_k(0,\eta)\right)^{-1/c} \sim \frac{\eta^N}{(2\pi N)(N/e)^{2N}} \sim \frac{1}{2\pi\sqrt{\eta}} e^{2\sqrt{\eta}} \left[\frac{\sqrt{\eta}}{N} e^{2N-2\sqrt{\eta}} \left(\frac{\eta}{N^2}\right)^N\right]$$

and the result follows since the term between [..] is ≈ 1 by $|N - \sqrt{\eta}| \leq 1$. □

3.2. *Properties of w and J*

In this section we prove some of the important and useful properties of J and w. This section is fundamental to our work but at times the proofs are tediously combinatorial.

The following trichotomy expresses the well-separation of critical times and is used several times in the sequel.

Lemma **3.2.** *— Let ξ, η be such that there exists some $\alpha \geq 1$ with $\frac{1}{\alpha}|\xi| \leq |\eta| \leq \alpha|\xi|$ and let k, n be such that $t \in I_{k,\eta}$ and $t \in I_{n,\xi}$ (note that $k \approx n$). Then at least one of following holds:*

(a) *$k = n$ (almost same interval);*
(b) *$|t - \frac{\eta}{k}| \geq \frac{1}{10\alpha}\frac{|\eta|}{k^2}$ and $|t - \frac{\xi}{n}| \geq \frac{1}{10\alpha}\frac{|\xi|}{n^2}$ (far from resonance);*
(c) *$|\eta - \xi| \gtrsim_\alpha \frac{|\eta|}{|n|}$ (well-separated).*

Proof. — To see that $k \approx n$ note

$$\frac{|k|}{|n|} = \frac{|tk|}{|tn|} = \frac{|\eta|}{|\xi|}\frac{|tk|}{|\eta|}\frac{|\xi|}{|tn|} \approx_\alpha 1. \tag{3.12}$$

If $k = n$ then there is nothing to prove. Suppose now both (a) and (b) are false, which means one of the two inequalities in (b) fails. Without loss of generality suppose $|t - \frac{\xi}{n}| < \frac{1}{10\alpha}\frac{|\xi|}{n^2}$. Then,

$$\left|\frac{\eta}{n}-\frac{\xi}{n}\right| \geq \left|t-\frac{\eta}{n}\right|-\left|t-\frac{\xi}{n}\right| \geq \frac{|\eta|}{2n(n+1)}-\frac{1}{10\alpha}\frac{|\xi|}{n^2} \gtrsim \frac{|\eta|}{n^2},$$

where we also used $k \neq n$. This proves (c). □

From the definition of w, for $t \in \mathrm{I}_{k,\eta}$ and $t > 2\sqrt{|\eta|}$, we have for $\tau = t - \frac{\eta}{k}$:

$$\begin{aligned}\partial_\tau w_{\mathrm{R}} &\approx \kappa \frac{k^2}{|\eta|} w_{\mathrm{NR}},\\ \partial_\tau w_{\mathrm{NR}} &\approx \kappa \frac{|\eta|}{k^2(1+\tau^2)} w_{\mathrm{R}}.\end{aligned} \tag{3.13}$$

Moreover, we also have the following:

Lemma **3.3.** — *For $t \in \mathrm{I}_{k,\eta}$ and $t > 2\sqrt{|\eta|}$, we have the following with $\tau = t - \frac{\eta}{k}$:*

$$\frac{\partial_t w_{\mathrm{NR}}(t,\eta)}{w_{\mathrm{NR}}(t,\eta)} \approx \frac{1}{1+|\tau|} \approx \frac{\partial_t w_{\mathrm{R}}(t,\eta)}{w_{\mathrm{R}}(t,\eta)}. \tag{3.14}$$

The following two lemmas are more substantial and show that although the toy model neglected interactions in η and ξ, $w(t,\eta)$ with $w(t,\xi)$ can still be compared effectively.

Lemma **3.4.**

(i) *For $t \geq 1$, and k, l, η, ξ such that* $\max(2\sqrt{|\xi|}, \sqrt{|\eta|}) < t < 2\min(|\xi|, |\eta|)$,

$$\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)} \frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)} \lesssim \langle \eta - \xi \rangle. \tag{3.15}$$

(ii) *For all $t \geq 1$, and k, l, η, ξ, such that for some $\alpha \geq 1$, $\frac{1}{\alpha}|\xi| \leq |\eta| \leq \alpha|\xi|$,*

$$\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}} \lesssim_\alpha \left[\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}} + \frac{|\eta|^{s/2}}{\langle t \rangle^s}\right] \langle \eta - \xi \rangle. \tag{3.16}$$

Remark **8.** — Notice the requirement that $t > 2\sqrt{|\xi|}$ in (3.15) and $t > 2\sqrt{|\eta|}$ in (3.14). This is due to the fact that $\partial_t w(t,\xi) \to 0$ as $t \searrow \mathrm{E}(\sqrt{|\xi|})$, and hence we do not have the lower bounds. The upper bounds still hold. The convenience of (3.16) is that it accounts for this detail and is useful in allowing us to seamlessly treat the endpoint case $t \approx \sqrt{\eta}$.

Proof of Lemma 3.4. — We first prove (3.15). By (3.14), k and l do not play a role and are omitted for the duration of the proof. Without loss of generality, we may restrict to

$\eta\xi \geq 0$ and $\eta/2 < \xi < 2\eta$ as otherwise by (3.14)

$$\left|\frac{\partial_t w(t,\eta)}{w(t,\eta)}\frac{w(t,\xi)}{\partial_t w(t,\xi)}\right| \lesssim \langle \xi \rangle \lesssim \langle \eta - \xi \rangle.$$

Let j and n be such that $t \in \mathrm{I}_{j,\xi}$ and $t \in \mathrm{I}_{n,\eta}$. Notice that like in (3.12) above, $j \approx n$. From (3.14),

$$\left|\frac{\partial_t w(t,\eta)}{w(t,\eta)}\frac{w(t,\xi)}{\partial_t w(t,\xi)}\right| \lesssim \frac{1+|t-\frac{\xi}{j}|}{1+|t-\frac{\eta}{n}|}.$$

In the case $n = j$, (3.15) follows from the inequality: for $a, b \geq 0$,

$$\frac{1+a}{1+b} \leq 1 + |a-b|. \tag{3.17}$$

In the case $n \neq j$, (3.15) follows from Lemma 3.2. Indeed, if (b) holds then since $n \approx j$ and $\eta \approx \xi$:

$$\left|\frac{\partial_t w(t,\eta)}{w(t,\eta)}\frac{w(t,\xi)}{\partial_t w(t,\xi)}\right| \lesssim \frac{1+|\frac{\xi}{j^2}|}{1+|\frac{\eta}{n^2}|} \lesssim 1.$$

If (c) holds then it follows that

$$\left|\frac{\partial_t w(t,\eta)}{w(t,\eta)}\frac{w(t,\xi)}{\partial_t w(t,\xi)}\right| \lesssim 1 + \left|\frac{\xi}{j^2}\right| \lesssim \langle \eta - \xi \rangle,$$

which finishes the proof of (3.15).

Next we prove (3.16). First, there is nothing to prove unless $\mathrm{E}(\sqrt{|\xi|}) \leq t \leq 2|\xi|$, so assume this is the case. If $2\sqrt{|\eta|} < t < 2|\eta|$ then (3.16) is a consequence of (3.15). If $t \leq 2\sqrt{|\eta|}$ then (3.16) follows from

$$\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}} \lesssim 1 \lesssim \frac{|\eta|^{s/2}}{\langle t \rangle^s}.$$

Lastly, consider $t \geq 2|\eta|$. If $|t - |\xi|| < \frac{1}{2\alpha}|\xi|$ then we have

$$|\eta - \xi| \geq t - |\eta| + |\xi| - t > |\eta| - \frac{1}{2\alpha}|\xi| \geq \frac{1}{2\alpha}|\xi|.$$

If instead $|t - |\xi|| \geq \frac{1}{2\alpha}|\xi|$ then by (3.14),

$$\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}} \lesssim \frac{\alpha^{1/2}}{\sqrt{|\xi|}}.$$

Hence, in both cases it follows that

$$\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}} \lesssim_\alpha \frac{\langle \eta - \xi \rangle^{1/2}}{\langle \xi \rangle^{1/2}}.$$

Since $2|\eta| \le t \le 2|\xi| \le 2\alpha|\eta|$, $t \ge 1$ and $s - 1/2 \le s/2$,

$$\frac{1}{\langle \xi \rangle^{1/2}} \lesssim_\alpha \frac{|\eta|^{s/2}}{\langle t \rangle^s},$$

which completes the proof of (3.16). □

Lemma **3.5.** — *For all t, η, ξ, we have*

$$\frac{w_{\mathrm{NR}}(t,\xi)}{w_{\mathrm{NR}}(t,\eta)} \lesssim e^{\mu|\eta-\xi|^{1/2}}. \tag{3.18}$$

Proof of Lemma 3.5. — For the proof of Lemma 3.5, we use $w(t,\eta) = w_{\mathrm{NR}}(t,\eta)$ and $w(t,\xi) = w_{\mathrm{NR}}(t,\xi)$ as there is no possible ambiguity. Switching the roles of ξ and η, we may assume without loss of generality that $|\xi| \le |\eta|$ and prove instead of (3.18) that

$$e^{-\mu|\eta-\xi|^{1/2}} \lesssim \frac{w_{\mathrm{NR}}(t,\xi)}{w_{\mathrm{NR}}(t,\eta)} \lesssim e^{\mu|\eta-\xi|^{1/2}}. \tag{3.19}$$

If $|\xi| < |\eta|/2$, then (3.19) is clear since by Lemma 3.1,

$$e^{-\frac{\mu}{2}\sqrt{\xi}} \le w(t,\xi) \le 1,$$

$|\xi| \le |\eta - \xi|$ and $|\eta| \le 2|\eta - \xi|$. Hence, in the sequel we may assume that $\eta, \xi \ge 0$ and $\eta/2 \le \xi \le \eta$.

First, if $t \ge 2\eta$, then $w(t,\xi) = w(t,\eta) = 1$ so there is nothing to prove.

If $t \le \min(t_{\mathrm{E}(\sqrt{\xi}),\xi}, t_{\mathrm{E}(\sqrt{\eta}),\eta})$, then by Lemma 3.1, (3.19) follows by:

$$\frac{w(t,\xi)}{w(t,\eta)} = \frac{w(0,\xi)}{w(0,\eta)} \approx \left(\frac{\xi}{\eta}\right)^{\mu/8} e^{\frac{\mu}{2}(\sqrt{\eta}-\sqrt{\xi})}.$$

If $2\xi \le t \le 2\eta$, then, since $w(t,\eta)$ is non-decreasing in time and $w(t,\xi)$ is constant for $t \ge 2\xi$,

$$1 \le \frac{w(t,\xi)}{w(t,\eta)} \le \frac{w(2\xi,\xi)}{w(2\xi,\eta)}.$$

If $t_{\mathrm{E}(\sqrt{\xi}),\xi} \le t \le t_{\mathrm{E}(\sqrt{\eta}),\eta}$ then by similar logic,

$$\frac{w(0,\xi)}{w(0,\eta)} \le \frac{w(t,\xi)}{w(t,\eta)} \le \frac{w(t_{\mathrm{E}(\sqrt{\eta}),\eta},\xi)}{w(t_{\mathrm{E}(\sqrt{\eta}),\eta},\eta)},$$

and if $t_{\mathrm{E}(\sqrt{\eta}),\eta} \leq t \leq t_{\mathrm{E}(\sqrt{\xi}),\xi}$ (note that this can occur even if $\eta \geq \xi$), then

$$\frac{w(t_{\mathrm{E}(\sqrt{\xi}),\xi}, \xi)}{w(t_{\mathrm{E}(\sqrt{\xi}),\xi}, \eta)} \leq \frac{w(t, \xi)}{w(t, \eta)} \leq \frac{w(0, \xi)}{w(0, \eta)}.$$

Hence, (3.19) is reduced to the case where $\max(t_{\mathrm{E}(\sqrt{\xi}),\xi}, t_{\mathrm{E}(\sqrt{\eta}),\eta}) \leq t \leq 2\xi$. Let j and n be such that $t \in \mathrm{I}_{n,\eta}$ and $t \in \mathrm{I}_{j,\xi}$. Arguing as in (3.12), we see that $n \approx j \leq n$. We consider three cases.

Case $j = n$*:* Assume first that $t \in \mathrm{I}^{\mathrm{R}}_{n,\eta} \cap \mathrm{I}^{\mathrm{R}}_{n,\xi}$, hence denoting $c = 1 + 2\mathrm{C}\kappa$ by the definition (3.5),

$$w(t, \eta) = \left(\frac{1^2}{\eta}\right)^c \left(\frac{2^2}{\eta}\right)^c \dots \left(\frac{(n-1)^2}{\eta}\right)^c \left(\frac{n^2}{\eta}\left[1 + b_{n,\eta}\left|t - \frac{\eta}{n}\right|\right]\right)^{\mathrm{C}\kappa}, \tag{3.20}$$

$$w(t, \xi) = \left(\frac{1^2}{\xi}\right)^c \left(\frac{2^2}{\xi}\right)^c \dots \left(\frac{(n-1)^2}{\xi}\right)^c \left(\frac{n^2}{\xi}\left[1 + b_{n,\xi}\left|t - \frac{\xi}{n}\right|\right]\right)^{\mathrm{C}\kappa}, \quad \text{and} \tag{3.21}$$

$$\frac{w(t, \xi)}{w(t, \eta)} = \left(\frac{\eta}{\xi}\right)^{c(n-1)+\mathrm{C}\kappa} \left(\frac{1 + b_{n,\xi}|t - \frac{\xi}{n}|}{1 + b_{n,\eta}|t - \frac{\eta}{n}|}\right)^{\mathrm{C}\kappa} = \mathrm{F}_1(\mathrm{F}_2)^{\mathrm{C}\kappa}. \tag{3.22}$$

The first factor on the right-hand side of (3.22) satisfies:

$$1 \leq \mathrm{F}_1 = \left(\frac{\eta}{\xi}\right)^{c(n-1)+\mathrm{C}\kappa} \lesssim \left(1 + \frac{\eta - \xi}{\xi}\right)^{c\sqrt{\xi}} \lesssim e^{c\frac{\eta-\xi}{\sqrt{\xi}}} \leq e^{c\sqrt{\eta-\xi}}. \tag{3.23}$$

For the second factor in (3.22), from (3.17),

$$\begin{aligned} \max\left(\mathrm{F}_2, \frac{1}{\mathrm{F}_2}\right) &\leq 1 + \left|b_{n,\xi}\left|t - \frac{\xi}{n}\right| - b_{n,\eta}\left|t - \frac{\eta}{n}\right|\right| \\ &\lesssim 1 + b_{n,\xi}\left|\frac{\eta - \xi}{n}\right| + |b_{n,\eta} - b_{n,\xi}|\frac{\eta}{n^2} \\ &\lesssim 1 + |\eta - \xi| + \left|\frac{1}{\xi} - \frac{1}{\eta}\right|\eta \lesssim \langle \eta - \xi \rangle. \end{aligned}$$

The case where $t \in \mathrm{I}^{\mathrm{L}}_{n,\eta} \cap \mathrm{I}^{\mathrm{L}}_{n,\xi}$ can be treated in the same way.

Assume now that $t \in \mathrm{I}^{\mathrm{L}}_{n,\eta} \cap \mathrm{I}^{\mathrm{R}}_{n,\xi}$, hence $w(t, \xi)$ is given by (3.21) and $w(t, \eta)$ by (from (3.5)),

$$\begin{aligned} w(t, \eta) = &\left(\frac{1^2}{\eta}\right)^c \left(\frac{2^2}{\eta}\right)^c \dots \left(\frac{(n-1)^2}{\eta}\right)^c \left(\frac{n^2}{\eta}\right)^{\mathrm{C}\kappa} \\ &\times \left(\left[1 + a_{n,\eta}\left|t - \frac{\eta}{n}\right|\right]\right)^{-1-\mathrm{C}\kappa} \quad \text{and} \end{aligned} \tag{3.24}$$

$$\frac{w(t,\xi)}{w(t,\eta)} = \left(\frac{\eta}{\xi}\right)^{c(n-1)+C\kappa} \left(1 + b_{n,\xi}\left|t - \frac{\xi}{n}\right|\right)^{C\kappa} \left(\left[1 + a_{n,\eta}\left|t - \frac{\eta}{n}\right|\right]\right)^{1+C\kappa}.$$

Using that $\frac{\xi}{n} \le t \le \frac{\eta}{n}$ and that $b_{n,\xi}, a_{n,\eta} < 4$, we get from (3.23) and (A.12),

$$1 \le \frac{w(t,\xi)}{w(t,\eta)} \lesssim e^{c\sqrt{\eta-\xi}}\left(1 + 4|\eta - \xi|\right)^{1+2C\kappa} \lesssim e^{2c\sqrt{\eta-\xi}}.$$

Case $j = n-1$: If $t \in I^{L}_{n,\eta}$ then $t_{n-1,\xi} \le \frac{\eta}{n}$. If $t \in I^{R}_{n-1,\xi}$, then $\frac{\xi}{n-1} < t_{n-1,\eta}$. In either one of these cases, we deduce that $\frac{\xi}{n^2} \lesssim \frac{\eta-\xi}{n}$ and we conclude in a similar way to (3.29) below.

Next assume that $t \in I^{R}_{n,\eta} \cap I^{L}_{n-1,\xi}$, which implies

$$w(t,\eta) = \left(\frac{1^2}{\eta}\right)^{c}\left(\frac{2^2}{\eta}\right)^{c}\cdots\left(\frac{(n-1)^2}{\eta}\right)^{c}\left(\frac{n^2}{\eta}\left[1 + b_{n,\eta}\left|t - \frac{\eta}{n}\right|\right]\right)^{C\kappa}, \tag{3.25}$$

$$w(t,\xi) = \left(\frac{1^2}{\xi}\right)^{c}\left(\frac{2^2}{\xi}\right)^{c}\cdots\left(\frac{(n-1)^2}{\xi}\right)^{C\kappa} \times \left(1 + a_{n-1,\xi}\left|t - \frac{\xi}{n-1}\right|\right)^{-1-C\kappa}, \quad \text{and} \tag{3.26}$$

$$\frac{w(t,\xi)}{w(t,\eta)} = \left(\frac{\eta}{\xi}\right)^{c(n-2)+C\kappa}\left(\frac{(n-1)^2}{\eta}\left[1 + a_{n-1,\xi}\left|t - \frac{\xi}{n-1}\right|\right]\right)^{-1-C\kappa} \times \left(\frac{n^2}{\eta}\left[1 + b_{n,\eta}\left|t - \frac{\eta}{n}\right|\right]\right)^{-C\kappa}. \tag{3.27}$$

The result now follows from Lemma 3.2: if (b) holds then (3.27) and $\eta \approx \xi$ imply

$$\frac{w(t,\xi)}{w(t,\eta)} \approx \left(\frac{\eta}{\xi}\right)^{c(n-2)+C\kappa} \tag{3.28}$$

and we conclude as (3.23). If Lemma 3.2 (c) holds then

$$\left(\frac{\eta}{\xi}\right)^{c(n-2)+C\kappa} \lesssim \frac{w(t,\xi)}{w(t,\eta)} \lesssim \left(\frac{\eta}{\xi}\right)^{c(n-2)+C\kappa}\langle \eta - \xi\rangle^{1+2C\kappa}, \tag{3.29}$$

and we again apply (3.23) and (A.12) to deduce (3.19).

Case $j < n-1$: In this case, it is easy to see that $\frac{\xi}{n^2} \lesssim \frac{\eta-\xi}{n}$ and we may conclude in a similar way to (3.29). □

A consequence of Lemma 3.5 is the following, which allows to easily exchange $J_k(\eta)$ for $J_l(\xi)$.

Lemma **3.6.** — *In general we have*

$$\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim \frac{|\eta|}{k^2(1+|t-\frac{\eta}{k}|)} e^{9\mu|k-l,\eta-\xi|^{1/2}}. \tag{3.30}$$

If any one of the following holds: ($t \notin \mathbf{I}_{k,\eta}$) *or* ($k = l$) *or* ($t \in \mathbf{I}_{k,\eta}$, $t \notin \mathbf{I}_{k,\xi}$ *and* $\frac{1}{\alpha}|\xi| \leq |\eta| \leq \alpha|\xi|$ *for some* $\alpha \geq 1$) *then we have the improved estimate*

$$\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim e^{10\mu|k-l,\eta-\xi|^{1/2}}. \tag{3.31}$$

Finally if $t \in \mathbf{I}_{l,\xi}$, $t \notin \mathbf{I}_{k,\eta}$ *and* $\frac{1}{\alpha}|\xi| \leq |\eta| \leq \alpha|\xi|$ *for some* $\alpha > 0$ *then*

$$\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim \frac{l^2(1+|t-\frac{\xi}{l}|)}{|\xi|} e^{11\mu|k-l,\eta-\xi|^{1/2}}. \tag{3.32}$$

Remark **9.** — The leading factors in (3.30) and (3.32) both are due to ratios of w_{R} and w_{NR}. Moreover, in the case $t \in \mathbf{I}_{k,\eta} \cap \mathbf{I}_{k,\xi}$, $k \neq l$, the only case where (3.30) is needed, we also have $|\eta| \approx |\xi|$ and hence from (3.30), the definition (3.8), and the proof of (3.14) it follows that

$$\begin{aligned} \frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} &\lesssim \frac{|\eta|}{k^2(1+|t-\frac{\eta}{k}|)} e^{9\mu|\eta-\xi|^{1/2}} \\ &\lesssim \frac{|\xi|}{k^2(1+|t-\frac{\xi}{k}|)} e^{10\mu|\eta-\xi|^{1/2}} \\ &\lesssim \frac{w_{\mathrm{NR}}(t,\xi)}{w_{\mathrm{R}}(t,\xi)} e^{20\mu|\eta-\xi|^{1/2}}. \end{aligned} \tag{3.33}$$

A version often used is if $t \in \mathbf{I}_{k,\eta} \cap \mathbf{I}_{k,\xi}$, $k \neq l$, then by (3.30), (3.14), Lemma 3.4 and (A.12),

$$\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim \frac{|\eta|}{k^2} \sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}} \sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}} e^{20\mu|k-l,\eta-\xi|^{1/2}}. \tag{3.34}$$

Remark **10.** — Note the appearance of $\mathbf{I}_{k,\eta}$ as opposed to $\mathrm{I}_{k,\eta}$. Each are defined in Section 1.2. The use of $\mathbf{I}$ is to rule out the end case $t \approx \sqrt{|\eta|}$, for example, we see that (3.31) holds if $t \approx \sqrt{|\eta|}$ even if $t \in \mathrm{I}_{k,\eta}$ and hence inequalities like (3.34) will not be necessary.

Proof. — For $a, b, c, d > 0$, note the basic inequality:

$$\frac{a+b}{c+d} \leq \frac{a}{c} + \frac{b}{d}, \tag{3.35}$$

which implies

$$\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \leq \frac{w_l(t,\xi)}{w_k(t,\eta)} e^{\mu|\eta-\xi|^{1/2}} + e^{\mu|k-l|^{1/2}}.$$

Hence, if $t \in \mathbf{I}_{k,\eta}$ then (3.30) follows from (3.18), the definition of w_{R}, (3.8) and the definition of $w_k(t,\eta)$, (3.9). If $t \in \mathrm{I}_{k,\eta}$ and $t \approx \sqrt{|\eta|}$ then since $k^2 \approx |\eta|$, then (3.31) holds as well from the definition of w_{R}. If $t \notin \mathbf{I}_{k,\eta}$ then (3.31) follows from (3.18). Now consider the remaining cases.

Proof of (3.31) *when* $t \in \mathbf{I}_{k,\eta}$, $t \notin \mathbf{I}_{k,\xi}$ *and* $\frac{1}{\alpha}|\xi| \leq |\eta| \leq \alpha|\xi|$ *for some* $\alpha \geq 1$*:* In this case, (3.31) follows from (3.30) together with Lemma 3.2 (b) or (c) and (A.12). Indeed, (a) is ruled out by $t \in \mathbf{I}_{k,\eta}$, $t \notin \mathbf{I}_{k,\xi}$; if (b) holds then there is no loss and if (c) holds then (A.12) can be used to absorb the loss.

Proof of (3.31) *when* $t \in \mathbf{I}_{k,\eta}$, $t \in \mathbf{I}_{k,\xi}$ *and* $k = l$*:* Inequality (3.31) follows from (3.18) and the definition of $w_k(t,\eta)$, (3.9).

Proof of (3.32)*:* First consider the case that $t \in \mathbf{I}_{l,\eta}$, which also implies $k \neq l$ and $\eta \approx \xi$. By (3.8), (3.9) and (3.18),

$$\begin{aligned}\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} &\leq \frac{w_l(t,\xi)}{w_k(t,\eta)} e^{\mu|\eta-\xi|^{1/2}} + w_l(t,\xi) e^{\mu|k|^{1/2}-\mu|\xi|^{1/2}} \\ &\lesssim \frac{w_{\mathrm{NR}}(t,\xi)}{w_{\mathrm{R}}(t,\xi)} e^{10\mu|\eta-\xi|^{1/2}} + e^{\mu|k|^{1/2}-\mu|\xi|^{1/2}}.\end{aligned}$$

Since $|k|^2 \leq \frac{1}{4}|\eta|$ then (A.11) and (A.12) deals with the second term and (3.32) follows.

In the case $t \notin \mathbf{I}_{l,\eta}$, (3.32) follows from Lemma 3.2. Indeed, if (b) holds then

$$1 \lesssim_\delta \frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)},$$

and (3.32) follows from (3.31) whereas if (c) holds then (3.32) follows from (3.31) and (A.12). ☐

The following variant of the previous lemmas is used in Section 5 to recover 1/2 derivatives.

Lemma **3.7.** — *Let* $t \leq \frac{1}{2}\min(\sqrt{|\eta|}, \sqrt{|\xi|})$. *Then,*

$$\left|\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} - 1\right| \lesssim \frac{\langle \eta-\xi, k-l\rangle}{\sqrt{|\xi|+|\eta|+|k|+|l|}} e^{11\mu|k-l,\eta-\xi|^{1/2}}. \tag{3.36}$$

Proof. — Due to the assumption on t, we have that $\mathrm{J}_k(t,\eta) = \mathrm{J}_k(0,\eta)$ and $\mathrm{J}_l(t,\xi) = \mathrm{J}_l(0,\xi)$. If $|\xi|^{1/2} + |\eta|^{1/2} + |k|^{1/2} + |l|^{1/2} \lesssim |\xi-\eta| + |k-l|$, then (3.36) follows from

(3.31). Similarly, we may restrict to $|\xi|^{1/2}+|\eta|^{1/2}+|k|^{1/2}+|l|^{1/2}\gtrsim 1$, as otherwise (3.36) is weaker than (3.31).

From now on, we assume that

$$|\xi-\eta|+|k-l|\leq \frac{1}{100}\left(|\xi|^{1/2}+|\eta|^{1/2}+|k|^{1/2}+|l|^{1/2}\right). \tag{3.37}$$

Case 1: $\frac{1}{10}(|k|+|l|)\leq|\xi|+|\eta|\leq 10(|k|+|l|)$*:* In this case, recalling the definition (2.22a)

$$\left|\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)}-1\right|\leq\frac{|\tilde{\mathrm{J}}_k(\eta)-\tilde{\mathrm{J}}_l(\xi)|}{\tilde{\mathrm{J}}_l(\xi)+e^{\mu|l|^{1/2}}}+\frac{|e^{\mu|k|^{1/2}}-e^{\mu|l|^{1/2}}|}{e^{\mu|l|^{1/2}}+\tilde{\mathrm{J}}_l(\xi)}. \tag{3.38}$$

The first term on the right-hand side of (3.38) is controlled by

$$\frac{|\tilde{\mathrm{J}}_k(\eta)-\tilde{\mathrm{J}}_l(\xi)|}{\tilde{\mathrm{J}}_l(\xi)}\leq\frac{w(0,\xi)}{w(0,\eta)}\left|e^{\mu(|\eta|^{1/2}-|\xi|^{1/2})}-1\right|+\left|\frac{w(0,\xi)}{w(0,\eta)}-1\right|.$$

To control the first term, we use $|e^x-1|\leq xe^x$, and our assumption that $|k,l|\approx|\eta,\xi|$ to deduce

$$\begin{aligned}\left|e^{\mu(|\eta|^{1/2}-|\xi|^{1/2})}-1\right|&\leq\mu\left||\eta|^{1/2}-|\xi|^{1/2}\right|e^{\mu(|\eta|^{1/2}-|\xi|^{1/2})}\\&\lesssim\frac{|\eta-\xi|}{|\xi|^{1/2}+|\eta|^{1/2}}e^{\mu(|\eta|^{1/2}-|\xi|^{1/2})}\\&\lesssim\frac{|\eta-\xi|}{|\xi|^{1/2}+|\eta|^{1/2}}e^{\mu|\eta-\xi|^{1/2}}\\&\lesssim\frac{\langle\eta-\xi,k-l\rangle}{\sqrt{|\xi|+|\eta|+|k|+|l|}}e^{\mu|\eta-\xi|^{1/2}},\end{aligned}$$

which suffices together with (3.18).

To control the second term, we notice that the condition (3.37) (together with our assumption $|k|+|l|\approx|\eta|+|\xi|$) implies that $|\mathrm{E}(\sqrt{|\eta|})-\mathrm{E}(\sqrt{|\xi|})|\leq 1$. We first look at the case $\mathrm{E}(\sqrt{|\eta|})=\mathrm{E}(\sqrt{|\xi|})$. Using the inequality: for all a, b, $|a|<1$ and $b>1$,

$$\left(1+\frac{a}{b^2}\right)^b-1\leq e\frac{|a|}{b},$$

we have (denoting $c=1+2\mathrm{C}\kappa$),

$$\left|\frac{w(0,\xi)}{w(0,\eta)}-1\right|=\left|\left(\frac{|\eta|}{|\xi|}\right)^{c\mathrm{E}(\sqrt{|\eta|})}-1\right|\lesssim\frac{|\eta-\xi|}{\sqrt{|\xi|}}.$$

If $\mathrm{E}(\sqrt{|\eta|}) = \mathrm{E}(\sqrt{|\xi|}) + 1$, then $\sqrt{|\xi|} < \mathrm{E}(\sqrt{|\eta|}) \leq \sqrt{|\eta|}$ and

$$\left|\frac{w(0,\xi)}{w(0,\eta)} - 1\right| = \left|\left(\frac{|\eta|}{|\xi|}\right)^{c\mathrm{E}(\sqrt{|\xi|})}\left(\frac{|\eta|}{\mathrm{E}(\sqrt{|\eta|})^2}\right)^{c} - 1\right|$$
$$\lesssim \frac{|\eta-\xi|}{\sqrt{|\xi|}} + \left|\left(\frac{|\eta|}{\mathrm{E}(\sqrt{|\eta|})^2}\right)^{c} - 1\right|$$

and we conclude since

$$\left|\left(\frac{|\eta|}{\mathrm{E}(\sqrt{|\eta|})^2}\right)^{c} - 1\right| \lesssim \frac{\langle \eta-\xi\rangle}{|\xi|}.$$

The case $\mathrm{E}(\sqrt{|\eta|}) = \mathrm{E}(\sqrt{|\xi|}) - 1$ is treated in the same way.

The second term on the right-hand side of (3.38) is controlled by (A.6) and $|e^x - 1| \leq xe^x$,

$$\left|e^{\mu(|k|^{1/2}-|l|^{1/2})} - 1\right| \lesssim \mu\frac{|k-l|}{|k|^{1/2}+|l|^{1/2}}e^{\mu|k-l|^{1/2}}$$
$$\lesssim \frac{\langle \eta-\xi, k-l\rangle}{\sqrt{|\xi|+|\eta|+|k|+|l|}}e^{\mu|k-l|^{1/2}}.$$

Case 2: $|\xi| + |\eta| \geq 10(|k|+|l|)$: Here we can treat the first term on the right-hand side of (3.38) as above and use $|\xi| \geq 4(|k|+|l|)$ together with (A.12) to treat the second term.

Case 3: $|k|+|l| \geq 10(|\xi|+|\eta|)$: Here we can treat the second term on the right-hand side of (3.38) as above and use $|l| \geq 4(|\xi|+|\eta|)$ together with (A.12) to treat the first term. □

3.3. *Product lemma and other basic properties of* A

Unlike the $\mathcal{G}^{\lambda,\sigma}$ norm (see Section A.2), the norm defined by A is *not* an algebra due to the discrepancy between resonant and non-resonant modes which is as large as an entire derivative near the critical times. However, A does define an algebra when restricted to the zero mode, as the zero mode is never resonant. Although A defines an algebra on the zero modes, the multipliers that appear in the CK terms do not, hence more generally, we have the following product lemma.

Lemma **3.8** *(Product lemma). — For some $c \in (0,1)$, all $\sigma > 1$, all $\beta > -\sigma + 1$ and $\alpha \geq 0$, the following inequalities hold for all f, g which depend only on v,*

$$\left\| |\partial_v|^\alpha \langle\partial_v\rangle^\beta \mathrm{A}(fg)\right\|_2 \lesssim \|f\|_{\mathcal{G}^{c\lambda,\sigma}}\left\||\partial_v|^\alpha\langle\partial_v\rangle^\beta \mathrm{A}g\right\|_2 + \|g\|_{\mathcal{G}^{c\lambda,\sigma}}\left\||\partial_v|^\alpha\langle\partial_v\rangle^\beta \mathrm{A}f\right\|_2 \tag{3.39a}$$

$$\left\| \sqrt{\frac{\partial_t w}{w}} \langle \partial_v \rangle^\beta \mathrm{A}(fg) \right\|_2 \lesssim \|g\|_{\mathcal{G}^{c\lambda,\sigma}} \left\| \left(\sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \langle \partial_v \rangle^\beta \mathrm{A} f \right\|_2 + \|f\|_{\mathcal{G}^{c\lambda,\sigma}} \left\| \left(\sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \langle \partial_v \rangle^\beta \mathrm{A} g \right\|_2. \tag{3.39b}$$

We also have for $\beta > -\sigma + 1$ *the algebra property,*

$$\left\| \langle \partial_v \rangle^\beta \mathrm{A}(fg) \right\|_2 \lesssim \left\| \langle \partial_v \rangle^\beta \mathrm{A} f \right\|_2 \left\| \langle \partial_v \rangle^\beta \mathrm{A} g \right\|_2. \tag{3.40}$$

Moreover, (3.39) *and* (3.40) *both hold for* A *replaced by* A^{R}.

Remark **11.** — Writing $(v')^2 - 1 = (v'-1)^2 + 2(v'-1)$ and $v'' = \partial_v(v'-1) + (v'-1)\partial_v(v'-1)$ (recall (2.13c)), by the bootstrap hypotheses on $v'-1$ combined with (3.40) we have,

$$\left\| \mathrm{A}^{\mathrm{R}} \left(1 - (v')^2\right) \right\|_2 \lesssim \left\| \mathrm{A}^{\mathrm{R}} \left(1 - v'\right) \right\|_2 + \left\| \mathrm{A}^{\mathrm{R}} \left(1 - v'\right) \right\|_2^2 \lesssim \epsilon \tag{3.41a}$$

$$\left\| \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v \rangle} v'' \right\|_2 = \left\| \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v \rangle} \left(v' \partial_v v'\right) \right\|_2 \lesssim \left\| \mathrm{A}^{\mathrm{R}} \left(1 - v'\right) \right\|_2 + \left\| \mathrm{A}^{\mathrm{R}} \left(1 - v'\right) \right\|_2^2 \lesssim \epsilon. \tag{3.41b}$$

Proof of Lemma 3.8. — The proof of (3.39a) follows from Lemmas 3.5 and 3.4 combined with a paraproduct decomposition; the argument is similar to many used in the sequel so is omitted.

Let us now focus on (3.39b) which is more intricate. Let us also just treat the case $\beta = 0$; actually any $\beta > 1 - \sigma$ can be treated by adjusting c accordingly and using (A.12). Moreover, we focus on $t \geq 1$; the case $t \in (0, 1)$ is easier and is not actually necessary for the proof of Theorem 1. Write

$$\mathcal{M}(t, \xi) = \sqrt{\frac{\partial_t w(t, \xi)}{w(t, \xi)}} \mathrm{A}_0(t, \xi),$$

and decompose fg with a paraproduct (see Section A.1):

$$fg = \mathrm{T}_f g + \mathrm{T}_g f + \mathcal{R}(f, g).$$

Consider the first term on the Fourier side:

$$\widehat{\mathcal{M}\mathrm{T}_f g}(\xi) = \frac{1}{2\pi} \sum_{\mathrm{M} \geq 8} \mathcal{M}(t, \xi) \int_{\xi'} \hat{g}(\xi')_{\mathrm{M}} \hat{f}\left(\xi - \xi'\right)_{<\mathrm{M}/8} d\xi'.$$

The goal is to use (3.16) to pass $\mathcal{M}$ onto g. On the support of the integrand (see Section A.1),

$$\left| |\xi| - \left|\xi'\right| \right| \leq \left|\xi - \xi'\right| \leq 3/2 \left(\frac{\mathrm{M}}{16}\right) \leq 6 \left|\xi'\right| / 32, \tag{3.42}$$

and hence $26|\xi'|/32 \le |\xi| \le 38|\xi'|/32$. Therefore (A.7) implies for some $c' \in (0,1)$ that $|\xi|^s \le |\xi'|^s + c'|\xi - \xi'|^s$, and hence by (3.42),

$$\left|\widehat{\mathcal{M}\mathrm{T}_f g}(\xi)\right| \lesssim \sum_{\mathrm{M}\ge 8} \int_{\xi'} \sqrt{\frac{\partial_t w(t,\xi)}{w(t,\xi)}} \langle \xi' \rangle^{\sigma} e^{\lambda|\xi'|^s} \mathrm{J}_0(\xi) \left|\hat{g}(\xi')_{\mathrm{M}}\right| e^{c'\lambda|\xi-\xi'|^s} \times \left|\hat{f}(\xi-\xi')_{<\mathrm{M}/8}\right| d\xi'.$$

By (3.31) it follows

$$\left|\widehat{\mathcal{M}\mathrm{T}_f g}(\xi)\right| \lesssim \sum_{\mathrm{M}\ge 8} \int_{\xi'} \sqrt{\frac{\partial_t w(t,\xi)}{w(t,\xi)}} \mathrm{A}_0(\xi') \left|\hat{g}(\xi')_{\mathrm{M}}\right| e^{10\mu|\eta-\xi|^{1/2} + c'\lambda|\xi-\xi'|^s} \times \left|\hat{f}(\xi-\xi')_{<\mathrm{M}/8}\right| d\xi'.$$

Then by (3.16), (A.11) and (A.12), for any $c \in (c', 1)$,

$$\left|\widehat{\mathcal{M}\mathrm{T}_f g}(\xi)\right| \lesssim \sum_{\mathrm{M}\ge 8} \int_{\xi'} \left[\sqrt{\frac{\partial_t w(t,\xi')}{w(t,\xi')}} + \frac{|\xi'|^{s/2}}{\langle t \rangle^s}\right] \mathrm{A}_0(\xi') \left|\hat{g}(\xi')_{\mathrm{M}}\right| e^{c\lambda|\xi-\xi'|^s} \times \left|\hat{f}(\xi-\xi')_{<\mathrm{M}/8}\right| d\xi'.$$

Since $|\xi| \approx \mathrm{M}$ by (3.42), (A.2) and (A.3) imply

$$\begin{aligned} \|\mathcal{M}\mathrm{T}_f g\|_2^2 &\lesssim \sum_{\mathrm{M}\ge 8} \left\| \left(\sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \mathrm{A} g_{\mathrm{M}} \right\|_2^2 \|f_{<\mathrm{M}/8}\|_{\mathcal{G}^{c\lambda,\sigma}}^2 \\ &\lesssim \left\| \left(\sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \mathrm{A} g \right\|_2^2 \|f\|_{\mathcal{G}^{c\lambda,\sigma}}^2. \end{aligned}$$

The contribution from $\mathcal{M}\mathrm{T}_g f$ is analogous and yields the other term in (3.39b). Let us now turn to the remainder,

$$\widehat{\mathcal{M}\mathcal{R}(f,g)}(\xi) = \frac{1}{2\pi} \sum_{\mathrm{M}\in\mathbf{D}} \sum_{\mathrm{M}/8 \le \mathrm{M}' \le 8\mathrm{M}} \int_{\xi'} \mathcal{M}(t,\xi) g(\xi')_{\mathrm{M}} f(\xi-\xi')_{\mathrm{M}'} d\xi'.$$

On the support of the integrand (see Section A.1) $\frac{1}{24}|\xi - \xi'| \le |\xi'| \le 24|\xi - \xi'|$, hence (A.8) implies for some c', $|\xi|^s \le c'|\xi'|^s + c'|\xi - \xi'|^s$, which gives

$$\left|\mathcal{M}\mathcal{R}(f,g)(\xi)\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}} \sum_{\mathrm{M}/8 \le \mathrm{M}' \le 8\mathrm{M}} \int_{\xi'} \sqrt{\frac{\partial_t w(t,\xi)}{w(t,\xi)}} \langle \xi \rangle^{\sigma} e^{c'\lambda|\xi'|^s} \mathrm{J}_0(\xi) \times \left|g(\xi')_{\mathrm{M}}\right| e^{c'\lambda|\xi-\xi'|^s} \left|f(\xi-\xi')_{\mathrm{M}'}\right| d\xi'.$$

Since $1 \leq t \leq 2|\xi|$ on the support of the integrand, by Lemma 3.1, (A.11) and (A.12),

$$\begin{aligned}\left|\mathcal{MR}(f,g)(\xi)\right| &\lesssim \sum_{M\in\mathbf{D}}\sum_{M/8\leq M'\leq 8M}\int_{\xi'}\frac{|\xi|^{s/2}}{\langle t\rangle^{s}}\langle\xi\rangle^{\sigma+s/2}e^{c'\lambda|\xi'|^{s}}J_0(\xi)\\ &\quad\times\left|g(\xi')_{M}\right|e^{c'\lambda|\xi-\xi'|^{s}}\left|f(\xi-\xi')_{M'}\right|d\xi'\\ &\lesssim \sum_{M\in\mathbf{D}}\sum_{M/8\leq M'\leq 8M}\int_{\xi'}\left(\frac{|\xi'|^{s/2}}{\langle t\rangle^{s}}+\frac{|\xi-\xi'|^{s/2}}{\langle t\rangle^{s}}\right)e^{c\lambda|\xi'|^{s}}\frac{1}{\langle\xi'\rangle}\\ &\quad\times\left|g(\xi')_{M}\right|e^{c\lambda|\xi-\xi'|^{s}}\left|f(\xi-\xi')_{M'}\right|d\xi',\end{aligned}$$

for any $c \in (c', 1)$. Therefore (A.3) and implies

$$\begin{aligned}\left\|\mathcal{MR}(f,g)(\xi)\right\|_{2} &\lesssim \sum_{M\in\mathbf{D}}\sum_{M/8\leq M'\leq 8M}\left\|\frac{|\partial_v|^{s/2}}{\langle t\rangle^{s}}g_{M}\right\|_{\mathcal{G}^{c\lambda,\sigma-1}}\|f_{M'}\|_{\mathcal{G}^{c\lambda}}\\ &\quad+\|g_{M}\|_{\mathcal{G}^{c\lambda,\sigma-1}}\left\|\frac{|\partial_v|^{s/2}}{\langle t\rangle^{s}}f_{M'}\right\|_{\mathcal{G}^{c\lambda}}\\ &\lesssim\left(\sum_{M\in\mathbf{D}}\left\|\frac{|\partial_v|^{s/2}}{\langle t\rangle^{s}}g_{M}\right\|^{2}_{\mathcal{G}^{c\lambda,\sigma}}\right)^{1/2}\|f\|_{\mathcal{G}^{c\lambda}}\\ &\quad+\left(\sum_{M\in\mathbf{D}}\|g_{M}\|^{2}_{\mathcal{G}^{c\lambda,\sigma}}\right)^{1/2}\left\|\frac{|\partial_v|^{s/2}}{\langle t\rangle^{s}}f\right\|_{\mathcal{G}^{c\lambda}},\end{aligned}$$

which by (A.2), proves (3.39b).

The proof in the case with A replaced by A^{R} proceeds the same. Indeed, from Lemma 3.6 we see that it is not a matter of w_{NR} vs w_{R}, it is only a matter of having either one or other, but not both. □

4. Elliptic estimates

The purpose of this section is to provide a thorough analysis of Δ_t. In particular, in this section we prove Proposition 2.4.

4.1. *Lossy estimate*

The following is the fundamental estimate on ϕ which allows to trade the regularity of f in a high norm for decay of the streamfunction in a slightly lower norm; the analogue of (2.5). This estimate is clear for the elliptic operator that arises from the linearized problem, which we denote by

$$\Delta_L = \partial_z^2 + (\partial_v - t\partial_z)^2. \tag{4.1}$$

As can be easily seen from examining Δ_L^{-1} (e.g. (1.4)), we cannot expect to gain $O(t^{-2})$ decay without paying two derivatives. Notice that since the coefficient v'' effectively contains a derivative on f (see (2.13)), the estimate below loses *three* derivatives. This loss can be treated with more precision, which is necessary in Section 4.2.

Due to the 'lossy' nature of the lemma, this can only be used when ϕ is being measured in a low norm, however, this occurs in many places in the proof, most notably the treatment of transport in Section 5, the treatment of $[\partial_t v]$ and $v'\partial_v[\partial_t v]$ in Proposition 2.5 and even the proof of the more precise elliptic estimate in Section 4.2.

Lemma **4.1** *(Lossy elliptic estimate). — Under the bootstrap hypotheses, for ϵ sufficiently small,*

$$\left\| P_{\neq 0}\phi(t) \right\|_{\mathcal{G}^{\lambda(t),\sigma-3}} \lesssim \frac{\| f(t) \|_{\mathcal{G}^{\lambda(t),\sigma-1}}}{1+t^2}. \tag{4.2}$$

Proof. — Omitting the time-dependence in λ, ϕ and f, first note

$$\begin{aligned} \| P_{\neq 0}\phi \|^2_{\mathcal{G}^{\lambda,\sigma-3}} &= \sum_{k\neq 0}\int_\eta e^{2\lambda|k,\eta|^s}\langle k,\eta\rangle^{2\sigma-6}\left|\hat\phi(k,\eta)\right|^2 d\eta \\ &= \sum_{k\neq 0}\int_\eta e^{2\lambda|(k,\eta)|^s}\frac{\langle k,\eta\rangle^{2\sigma-2}}{\langle k,\eta\rangle^4(k^2+|\eta-kt|^2)^2} \\ &\qquad \times \left(k^2+|\eta-kt|^2\right)^2\left|\hat\phi(k,\eta)\right|^2 d\eta \\ &\lesssim \frac{1}{\langle t\rangle^4}\|\Delta_L P_{\neq 0}\phi\|^2_{\mathcal{G}^{\lambda,\sigma-1}}. \end{aligned} \tag{4.3}$$

We write Δ_t as a perturbation of Δ_L via (recall the definitions (2.10), (4.1)),

$$\Delta_L P_{\neq 0}\phi = P_{\neq 0} f + \left(1-(v')^2\right)(\partial_v - t\partial_z)^2 P_{\neq 0}\phi - v''(\partial_v - t\partial_z)P_{\neq 0}\phi.$$

By the algebra inequality (A.10),

$$\begin{aligned} \|\Delta_L P_{\neq 0}\phi\|_{\mathcal{G}^{\lambda,\sigma-1}} &\lesssim \|f\|_{\mathcal{G}^{\lambda,\sigma-1}} + \left\|1-(v')^2\right\|_{\mathcal{G}^{\lambda,\sigma-1}}\|\Delta_L P_{\neq 0}\phi\|_{\mathcal{G}^{\lambda,\sigma-1}} \\ &\quad + \left\|v''\right\|_{\mathcal{G}^{\lambda,\sigma-1}}\left\|(\partial_y - t\partial_z)P_{\neq 0}\phi\right\|_{\mathcal{G}^{\lambda,\sigma-1}}. \end{aligned}$$

Therefore, (3.41) implies,

$$\|\Delta_L P_{\neq 0}\phi\|_{\mathcal{G}^{\lambda,\sigma-1}} \lesssim \|f\|_{\mathcal{G}^{\lambda,\sigma-1}} + \epsilon\|\Delta_L P_{\neq 0}\phi\|_{\mathcal{G}^{\lambda,\sigma-1}}.$$

Together with (4.3), this implies the a priori estimate (4.2) provided that ϵ is sufficiently small. □

4.2. *Precision elliptic control*

Now we turn to the proof of Proposition 2.4, announced in Section 2.3. This is the main elliptic estimate which underlies the treatment of reaction in Section 6 and the key estimates of Section 8. If Δ_t were simply Δ_L then the estimate would be trivial. However, the coefficients depend on the vorticity, which both couples all of the frequencies in the v direction together and introduces the potential for losing regularity (it is key that the coefficients only depend on v). The simplest effect one can see is the appearance of the CK multipliers on the coefficients collected in (2.29), which occur when 'derivatives' taken on the LHS of (2.28) land on the coefficients of Δ_t. Notice that these 'CCK' terms contain the more dangerous *resonant* regularity (see (3.10)). This effect is controlled in Proposition 2.5. The other effect one sees is the $\langle \partial_v (\partial_z t)^{-1}\rangle^{-1}$ on the LHS, which is a precise way of treating the loss due to the fact that v'' effectively contains a derivative on f.

Proof of Proposition 2.4. — Since the coefficients only depend on v, $\Delta_t\phi = f$ decouples mode-by-mode in the z frequencies. Hence, we essentially prove a mode-by-mode analogue of (2.28) and then sum.

As in Lemma 4.1 write (recall (2.10), (4.1)),

$$\Delta_L \phi = f + \left(1-(v')^2\right)(\partial_v - t\partial_z)^2\phi - v''(\partial_v - t\partial_z)\phi.$$

Define the multipliers

$$\mathcal{M}_1(t,l,\xi) = \left\langle \frac{\xi}{lt}\right\rangle^{-1} \frac{|l,\xi|^{s/2}}{\langle t\rangle^s} \mathrm{AP}_{\neq 0}$$
$$\mathcal{M}_2(t,l,\xi) = \left\langle \frac{\xi}{lt}\right\rangle^{-1} \sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}.$$

Clearly,

$$\sum_{i=1,2} \|\mathcal{M}_i f\|_2^2 \lesssim \frac{1}{\langle t\rangle^{2s}} \left\| |\nabla|^{s/2} \mathrm{P}_{\neq 0} \mathrm{A} f\right\|_2^2 + \left\| \sqrt{\frac{\partial_t w}{w}} \mathrm{P}_{\neq 0}\tilde{\mathrm{A}} f\right\|_2^2,$$

and hence the proposition would be trivial if $\Delta_L\phi = f$. Define,

$$\mathrm{T}^1 = \left(1-(v')^2\right)(\partial_v - t\partial_z)^2\phi$$
$$\mathrm{T}^2 = -v''(\partial_v - t\partial_z)\phi$$

and divide each via a paraproduct decomposition in the v variable only

$$
\begin{aligned}
\mathrm{T}^1 &= \sum_{\mathrm{M}\geq 8} \left(1-(v')^2\right)_{\mathrm{M}} (\partial_v - t\partial_z)^2 \phi_{<\mathrm{M}/8} \\
&\quad + \sum_{\mathrm{M}\geq 8} \left(1-(v')^2\right)_{<\mathrm{M}/8} (\partial_v - t\partial_z)^2 \phi_{\mathrm{M}} \\
&\quad + \sum_{\mathrm{M}\in\mathbf{D}} \sum_{\frac{1}{8}\mathrm{M}\leq \mathrm{M}'\leq 8\mathrm{M}} \left(1-(v')^2\right)_{\mathrm{M}'} (\partial_v - t\partial_z)^2 \phi_{\mathrm{M}} \\
&= \mathrm{T}^1_{\mathrm{HL}} + \mathrm{T}^1_{\mathrm{LH}} + \mathrm{T}^1_{\mathcal{R}} \\
\mathrm{T}^2 &= -\sum_{\mathrm{M}\geq 8} (v'')_{\mathrm{M}} (\partial_v - t\partial_z) \phi_{<\mathrm{M}/8} - \sum_{\mathrm{M}\geq 8} (v'')_{<\mathrm{M}/8} (\partial_v - t\partial_z) \phi_{\mathrm{M}} \\
&\quad - \sum_{\mathrm{M}\in\mathbf{D}} \sum_{\frac{1}{8}\mathrm{M}\leq \mathrm{M}'\leq 8\mathrm{M}} (v'')_{\mathrm{M}'} (\partial_v - t\partial_z) \phi_{\mathrm{M}} \\
&= \mathrm{T}^2_{\mathrm{HL}} + \mathrm{T}^2_{\mathrm{LH}} + \mathrm{T}^2_{\mathcal{R}}.
\end{aligned}
$$

The basic idea is to treat the HL terms by passing $\mathcal{M}_i$ onto the coefficients and to treat the LH terms by passing the $\mathcal{M}_i$ onto ϕ and using ϵ sufficiently small to absorb these terms on the left-hand side of (2.28). Each step has several complications, dealt with and discussed below.

4.2.1. *Low-high interactions*

Since $\mathrm{T}^1_{\mathrm{LH}}$ contains more derivatives on ϕ than $\mathrm{T}^2_{\mathrm{LH}}$, the former is strictly harder so we treat only $\mathrm{T}^1_{\mathrm{LH}}$. In what follows we use the shorthand

$$
\mathrm{G}(\xi) = \widehat{\left(1-(v')^2\right)}(\xi).
$$

Writing $\mathrm{T}^1_{\mathrm{LH}}$ on the frequency side with this convention gives

$$
\begin{aligned}
\widehat{\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}}(l,\xi) = -\frac{1}{2\pi} \sum_{\mathrm{M}\geq 8} \int_{\xi'} \mathcal{M}_1(t,l,\xi) \mathrm{G}(\xi-\xi')_{<\mathrm{M}/8} \\
\times \left(\xi' - lt\right)^2 \hat{\phi}_l(\xi')_{\mathrm{M}} d\xi'.
\end{aligned}
$$

On the support of the integrand (see Section A.1),

$$
\left||l,\xi| - |l,\xi'|\right| \leq |\xi-\xi'| \leq 3/2\left(\frac{\mathrm{M}}{16}\right) \leq 6|\xi'|/32 \leq 6|l,\xi'|/32, \tag{4.4}
$$

and hence $26|l,\xi'|/32 \leq |l,\xi| \leq 38|l,\xi'|/32$ and $26|\xi'|/32 \leq |\xi| \leq 38|\xi'|/32$. Therefore, (A.7) implies that there exists a $c\in(0,1)$ such that

$$
e^{\lambda|l,\xi|^s} \leq e^{\lambda|l,\xi'|^s + c\lambda|\xi-\xi'|^s}.
$$

Therefore (also using $|\xi| \approx |\xi'|$):

$$\left|\widehat{\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\geq 8} \int_{\xi'} \left\langle \frac{\xi'}{lt} \right\rangle^{-1} \frac{|l,\xi'|^{s/2}}{\langle t\rangle^s} \langle l,\xi'\rangle^\sigma \mathrm{J}_l(\xi)$$
$$\times e^{c\lambda|\xi-\xi'|^s} \left|\mathrm{G}(\xi-\xi')_{<\mathrm{M}/8}\right| \left|\xi'-lt\right|^2 \left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right| e^{\lambda|l,\xi'|^s} d\xi'.$$

The goal is now to pass the multiplier $\mathcal{M}_1$ onto ϕ. It is important here that we are not comparing different modes in z, and hence (3.31) applies. Hence by (A.11) (since $c < 1$),

$$\left|\widehat{\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\geq 8} \int_{\xi'} \left\langle \frac{\xi'}{lt} \right\rangle^{-1} \frac{|l,\xi'|^{s/2}}{\langle t\rangle^s} e^{\lambda|\xi-\xi'|^s} \left|\mathrm{G}(\xi-\xi')_{<\mathrm{M}/8}\right| \left|\xi'-lt\right|^2$$
$$\times \mathrm{A}_l(\xi') \left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right| d\xi'.$$

Then (A.3), (A.2) and (3.41a) imply

$$\begin{aligned}
\left\|\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}\right\|_2^2 &= \sum_{l\neq 0} \left\|\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}(l)\right\|_2^2 \\
&\lesssim \sum_{l\neq 0} \sum_{\mathrm{M}\geq 8} \left\|\left(1-(v')^2\right)_{<\mathrm{M}/8}\right\|^2_{\mathcal{G}^{\lambda,2}} \left\|\mathcal{M}_1 \Delta_{\mathrm{L}} \mathrm{P}_{\neq 0}(\phi_l)_{\mathrm{M}}\right\|_2^2 \\
&\lesssim \epsilon^2 \|\mathcal{M}_1 \Delta_{\mathrm{L}} \mathrm{P}_{\neq 0} \phi\|_2^2,
\end{aligned}$$

completing the treatment of $\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}$, as this can be absorbed by the LHS of (2.28).

Now turn to $\mathcal{M}_2 \mathrm{T}^1_{\mathrm{LH}}$, which by the frequency localization on the support of the integrand (4.4) together with (A.7), is bounded by

$$\left|\widehat{\mathcal{M}_2 \mathrm{T}^1_{\mathrm{LH}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\geq 8} \int_{\xi'} \left\langle \frac{\xi'}{lt} \right\rangle^{-1} \sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}} \langle l,\xi'\rangle^\sigma \tilde{\mathrm{J}}_l(\xi)$$
$$\times e^{c\lambda|\xi-\xi'|^s} \mathrm{G}(\xi-\xi')_{<\mathrm{M}/8} \left|\xi'-lt\right|^2 \left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right| e^{\lambda|l,\xi'|^s} d\xi'.$$

Now, by (3.16). it follows that a proof similar to that used to treat $\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}$ implies

$$\left\|\mathcal{M}_2 \mathrm{T}^1_{\mathrm{LH}}\right\|_2^2 \lesssim \epsilon^2 \|\mathcal{M}_1 \Delta_{\mathrm{L}} \phi\|_2^2 + \epsilon^2 \|\mathcal{M}_2 \Delta_{\mathrm{L}} \phi\|_2^2,$$

which completes the treatment of $\|\mathcal{M}_2 \mathrm{T}^1_{\mathrm{LH}}\|_2^2$.

4.2.2. *High-low interactions*

Consider first $\mathcal{M}_1 \mathrm{T}^2_{\mathrm{HL}}$. The notation is deceptive: the frequency in z could be very large and hence more 'derivatives' are appearing on ϕ and we will be in a situation like the LH terms. Hence we break into two cases:

$$\begin{aligned}\widehat{\mathcal{M}_1 \mathrm{T}^2_{\mathrm{HL}}}(l,\xi) &= -\frac{1}{2\pi}\sum_{\mathrm{M}\geq 8}\int_{\xi'}[\mathbf{1}_{|l|\geq\frac{1}{16}|\xi|}+\mathbf{1}_{|l|<\frac{1}{16}|\xi|}]\mathcal{M}_1(t,l,\xi)\widehat{v''}(\xi-\xi')_{\mathrm{M}} \\ &\quad \times i(\xi'-lt)\hat{\phi}_l(\xi')_{<\mathrm{M}/8}d\xi' \\ &= \widehat{\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}}(l,\xi)+\widehat{\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}}(l,\xi).\end{aligned}$$

First consider $\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}$, which we treat as a Low-High term. On the support of the integrand, we claim that there is some $c\in(0,1)$ such that,

$$|l,\xi|^s \leq |l,\xi'|^s + c|\xi-\xi'|^s. \tag{4.5}$$

To see this, one can consider separately the cases $\frac{1}{16}|\xi|\leq|l|\leq 16|\xi|$ and $|l|>16|\xi|$, applying (A.8) and (A.7) respectively. It follows that

$$\begin{aligned}\left|\widehat{\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}}(l,\xi)\right| &\lesssim \sum_{\mathrm{M}\geq 8}\int_{\xi'}\mathbf{1}_{|l|>\frac{1}{16}|\xi|}\frac{|l|^{s/2}}{\langle t\rangle^s}\langle l\rangle^\sigma \mathrm{J}_l(\xi)e^{c\lambda|\xi-\xi'|^s}\left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right| \\ &\quad\times|\xi'-lt|\left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right|e^{\lambda|l,\xi'|^s}d\xi',\end{aligned}$$

where we also used $|\xi|\lesssim|lt|$ to remove the leading factor $\langle\xi/lt\rangle^{-1}$. As in the treatment of $\mathcal{M}_1 \mathrm{T}^1_{\mathrm{LH}}$, we apply (3.31) (as we are not comparing different z modes) and (A.11) to deduce

$$\begin{aligned}\left|\widehat{\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}}(l,\xi)\right| &\lesssim \sum_{\mathrm{M}\geq 8}\int_{\xi'}\mathbf{1}_{|l|>\frac{1}{16}|\xi|}\frac{|l|^{s/2}}{\langle t\rangle^s}e^{\lambda|\xi-\xi'|^s}\left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right||\xi'-lt|\mathrm{A}_l(\xi') \\ &\quad\times\left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right|d\xi'.\end{aligned}$$

Since $|\xi'-lt|\leq|l|^2+|\xi'-lt|^2$, applying (A.3), (A.2) and (3.41),

$$\begin{aligned}\left\|\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}\right\|_2^2 = \sum_{l\neq 0}\left\|\mathcal{M}_1 \mathrm{T}^{2,z}_{\mathrm{HL}}(l)\right\|_2^2 &\lesssim \sum_{\mathrm{M}\geq 8}\left\|v''_{\mathrm{M}}\right\|^2_{\mathcal{G}^{\lambda,3}}\|\mathcal{M}_1\Delta_{\mathrm{L}}\mathrm{P}_{\neq 0}\phi\|_2^2 \\ &\lesssim \epsilon^2\|\mathcal{M}_1\Delta_{\mathrm{L}}\mathrm{P}_{\neq 0}\phi\|_2^2.\end{aligned} \tag{4.6}$$

Next consider $\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}$. On the support of the integrand, the 'derivatives' are all landing on v'' and since this function essentially contains a derivative of f it is here where we need the $\langle\xi/lt\rangle^{-1}$ (see (2.13)). In this case, using $|\xi'|\leq\frac{3}{16}|\xi-\xi'|$ analogous to (4.4),

$$\begin{aligned}\left||\xi-\xi'|-|l,\xi|\right| \leq |l,\xi'| \leq |\xi|/16+|\xi'| &\leq |\xi-\xi'|/16+17|\xi'|/16 \\ &\leq \frac{67}{256}|\xi-\xi'|.\end{aligned} \tag{4.7}$$

Therefore by (A.7), there exists some $c \in (0, 1)$ such that

$$\left|\widehat{\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}}(l, \xi)\right| \lesssim \sum_{\mathrm{M} \geq 8} \int_{\xi'} \mathbf{1}_{|l| < \frac{1}{16}|\xi|} \left\langle \frac{\xi}{lt} \right\rangle^{-1} \frac{|l, \xi|^{s/2}}{\langle t \rangle^s} \langle l, \xi \rangle^\sigma \mathrm{J}_l(\xi) e^{\lambda|\xi - \xi'|^s} \times \left|\widehat{v''}(\xi - \xi')_{\mathrm{M}}\right| \left|\xi' - lt\right| \left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right| e^{c\lambda|l,\xi'|^s} d\xi'.$$

Since we will pass $\mathcal{M}_1$ onto v'', Lemma 3.6 could imply a loss. However, from Lemma 3.6 and the definitions (3.10), (3.8), we see that on the support of the integrand we have (using that $|\xi| \gtrsim |l|$):

$$\mathrm{J}_l(\xi) \lesssim \mathrm{J}^{\mathrm{R}}(\xi - \xi') e^{20\mu|\xi'|^{1/2}}. \tag{4.8}$$

Since $c < 1$ and $s > 1/2$, then (4.8) and (A.11) (also $|l| \lesssim |\xi| \approx |\xi - \xi'|$) imply

$$\left|\widehat{\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}}(l, \xi)\right| \lesssim \sum_{\mathrm{M} \geq 8} \int_{\xi'} \left\langle \frac{\xi - \xi'}{lt} \right\rangle^{-1} \frac{|\xi - \xi'|^{s/2}}{\langle t \rangle^s} \mathrm{A}^{\mathrm{R}}(\xi - \xi') \left|\widehat{v''}(\xi - \xi')_{\mathrm{M}}\right| \times \left|\xi' - lt\right| \left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right| e^{\lambda|l,\xi'|^s} d\xi'.$$

Therefore,

$$\begin{aligned}\left|\widehat{\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}}(l, \xi)\right| &\lesssim \sum_{\mathrm{M} \geq 8} \int_{\xi'} \frac{1}{\langle \frac{\xi - \xi'}{lt} \rangle \langle lt \rangle} \frac{|\xi - \xi'|^{s/2}}{\langle t \rangle^s} \mathrm{A}^{\mathrm{R}}(\xi - \xi') \left|\widehat{v''}(\xi - \xi')_{\mathrm{M}}\right| \langle lt \rangle \\ &\quad \times \left|\xi' - lt\right| \left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right| e^{\lambda|l,\xi'|^s} d\xi' \\ &\lesssim \sum_{\mathrm{M} \geq 8} \int_{\xi'} \frac{|\xi - \xi'|^{s/2}}{\langle t \rangle^s} \frac{\mathrm{A}^{\mathrm{R}}(\xi - \xi')}{\langle \xi - \xi' \rangle} \left|\widehat{v''}(\xi - \xi')_{\mathrm{M}}\right| \langle lt \rangle \left|\xi' - lt\right| \\ &\quad \times \left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right| e^{\lambda|l,\xi'|^s} d\xi'.\end{aligned}$$

Finally, by (A.3) ($\sigma > 6$), (A.2) Lemma 4.1 and the bootstrap hypotheses we have

$$\begin{aligned}\left\|\mathcal{M}_1 \mathrm{T}^{2,v}_{\mathrm{HL}}\right\|_2^2 &\lesssim \sum_{\mathrm{M} \geq 8} \frac{1}{\langle t \rangle^{2s}} \left\| |\partial_v|^{s/2} \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v \rangle} v''_{\mathrm{M}} \right\|_2^2 \left\| t^2 \phi_{<\mathrm{M}/8} \right\|^2_{\mathcal{G}^{\lambda,\sigma-3}} \\ &\lesssim \frac{\epsilon^2}{\langle t \rangle^{2s}} \left\| |\partial_v|^{s/2} \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v \rangle} v'' \right\|_2^2 \lesssim \epsilon^2 \mathrm{CCK}^2_\lambda.\end{aligned} \tag{4.9}$$

This is sufficient to treat $\mathcal{M}_1 \mathrm{T}^2_{\mathrm{HL}}$.

The corresponding argument to show

$$\left\|\mathcal{M}_1 \mathrm{T}^1_{\mathrm{HL}}\right\|_2^2 \lesssim \epsilon^2 \|\mathcal{M}_1 \Delta_{\mathrm{L}} \phi\|_2^2 + \epsilon^2 \mathrm{CCK}^1_\lambda$$

is similar and hence omitted. Note that in the case of $\mathrm{T}^1_{\mathrm{HL}}$ no derivative needs to be recovered on the coefficients (and would be impossible due to lack of time-decay). This completes the treatment of the High-Low terms involving $\mathcal{M}_1$.

The argument for $\mathcal{M}_2$ is only slightly different. We treat $\mathcal{M}_2\mathrm{T}^2_{\mathrm{HL}}$; the case $\mathcal{M}_2\mathrm{T}^1_{\mathrm{HL}}$ is analogous. As in $\mathcal{M}_1\mathrm{T}^2_{\mathrm{HL}}$, we divide into separate cases:

$$\begin{aligned}\widehat{\mathcal{M}_2\mathrm{T}^2_{\mathrm{HL}}}(l,\xi) &= -\frac{1}{2\pi}\sum_{\mathrm{M}\geq 8}\int_{\xi'}[\mathbf{1}_{|l|\geq\frac{1}{16}|\xi|}+\mathbf{1}_{|l|<\frac{1}{16}|\xi|}]\mathcal{M}_2(t,l,\xi)\widehat{v''}(\xi-\xi')_{\mathrm{M}}\\ &\quad\times i(\xi'-lt)\hat{\phi}_l(\xi')_{<\mathrm{M}/8}d\xi'\\ &=\widehat{\mathcal{M}_2\mathrm{T}^{2,z}_{\mathrm{HL}}}(l,\xi)+\widehat{\mathcal{M}_2\mathrm{T}^{2,v}_{\mathrm{HL}}}(l,\xi).\end{aligned}$$

First consider $\mathcal{M}_2\mathrm{T}^{2,z}_{\mathrm{HL}}$, which like $\mathcal{M}_1\mathrm{T}^{2,z}_{\mathrm{HL}}$, we treat as a Low-High term. As there, (4.5) applies on the support of the integrand and hence by (3.31) (since we are not comparing different z frequencies) and (A.11):

$$\begin{aligned}\left|\widehat{\mathcal{M}_2\mathrm{T}^{2,z}_{\mathrm{HL}}}(l,\xi)\right| &\lesssim \sum_{\mathrm{M}\geq 8}\int_{\xi'}\mathbf{1}_{|l|\geq\frac{1}{16}|\xi|}\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}}e^{\lambda|\xi-\xi'|^s}\left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right|\\ &\quad\times\left|\xi'-lt\right|\tilde{\mathrm{A}}_l(\xi')\left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right|d\xi'.\end{aligned}$$

As in the treatment of $\mathcal{M}_2\mathrm{T}^1_{\mathrm{LH}}$ above in Section 4.2.1, we apply (3.16), and then a proof similar to that used to treat $\mathcal{M}_1\mathrm{T}^{2,z}_{\mathrm{HL}}$ in (4.6) implies

$$\left\|\mathcal{M}_2\mathrm{T}^{2,z}_{\mathrm{HL}}\right\|_2^2\lesssim\epsilon^2\|\mathcal{M}_2\Delta_{\mathrm{L}}\phi\|_2^2+\epsilon^2\|\mathcal{M}_1\Delta_{\mathrm{L}}\phi\|_2^2, \tag{4.10}$$

which completes the treatment of $\mathcal{M}_2\mathrm{T}^{2,z}_{\mathrm{HL}}$.

Now turn to $\mathcal{M}_2\mathrm{T}^{2,v}_{\mathrm{HL}}$. As in the treatment of $\mathcal{M}_1\mathrm{T}^{2,v}_{\mathrm{HL}}$, (4.7) holds on the support of the integrand, and hence so does (4.8) with $\mathrm{J}_l(\xi)$ replaced by $\tilde{\mathrm{J}}_l(\xi)$ (recall (2.22a)). Therefore, by (A.7) for some $c\in(0,1)$ followed by (4.8) and (A.11) implies

$$\begin{aligned}\left|\widehat{\mathcal{M}_2\mathrm{T}^{2,v}_{\mathrm{HL}}}(l,\xi)\right| &\lesssim \sum_{\mathrm{M}\geq 8}\int_{\xi'}\mathbf{1}_{|l|<\frac{1}{16}|\xi|}\left\langle\frac{\xi}{lt}\right\rangle^{-1}\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}}\langle\xi-\xi'\rangle^{\sigma}\tilde{\mathrm{J}}_l(\xi)e^{\lambda|\xi-\xi'|^s}\\ &\quad\times\left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right|\left|\xi'-lt\right|\left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right|e^{c\lambda|l,\xi'|^s}d\xi'\\ &\lesssim\sum_{\mathrm{M}\geq 8}\int_{\xi'}\frac{\mathbf{1}_{|l|<\frac{1}{16}|\xi|}}{\langle\frac{\xi}{lt}\rangle|lt|}\sqrt{\frac{\partial_t w_l(\xi)}{w_l(\xi)}}\mathrm{A}^{\mathrm{R}}(\xi-\xi')\left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right|\\ &\quad\times|lt|\left|\xi'-lt\right|\left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right|e^{\lambda|l,\xi'|^s}d\xi'.\end{aligned}$$

Then by $|l| \lesssim |\xi| \approx |\xi - \xi'|$ and (3.16) we have

$$\left|\widehat{\mathcal{M}_2 \mathrm{T}^{2,v}_{\mathrm{HL}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\geq 8} \int_{\xi'} \left(\sqrt{\frac{\partial_t w_0(\xi-\xi')}{w_0(\xi-\xi')}} + \frac{|\xi-\xi'|^{s/2}}{\langle t\rangle^s} \right) \frac{\mathrm{A}^{\mathrm{R}}(\xi-\xi')}{\langle \xi-\xi'\rangle} \\ \times \left|\widehat{v''}(\xi-\xi')_{\mathrm{M}}\right| |lt| \left|\xi' - lt\right| \langle \xi'\rangle \left|\hat{\phi}_l(\xi')_{<\mathrm{M}/8}\right| e^{\lambda|l,\xi'|^s} d\xi'.$$

Hence, by (A.3) ($\sigma > 7$), (A.2), Lemma 4.1 and the bootstrap hypotheses we have

$$\left\|\mathcal{M}_2 \mathrm{T}^{2,v}_{\mathrm{HL}}\right\|_2^2 \lesssim \sum_{\mathrm{M}\geq 8} \left(\frac{1}{\langle t\rangle^{2s}} \left\| |\partial_v|^{s/2} \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v\rangle} v''_{\mathrm{M}} \right\|_2^2 + \left\| \sqrt{\frac{\partial_t w}{w}} \frac{\mathrm{A}^{\mathrm{R}}}{\langle \partial_v\rangle} v''_{\mathrm{M}} \right\|_2^2 \right) \\ \times \left\| t^2 \phi_{<\mathrm{M}/8} \right\|^2_{\mathcal{G}^{\lambda,\sigma-3}} \\ \lesssim \epsilon^2 \mathrm{CCK}^2_\lambda + \epsilon^2 \mathrm{CCK}^2_w.$$

Together with (4.10), this completes the treatment of $\mathcal{M}_2 \mathrm{T}^2_{\mathrm{HL}}$. The treatment of $\mathcal{M}_2 \mathrm{T}^1_{\mathrm{HL}}$ is analogous and omitted.

4.2.3. *Remainders*

The last terms to consider are $\mathrm{T}^1_{\mathcal{R}}$ and $\mathrm{T}^2_{\mathcal{R}}$. In these terms powers of ∂_v can be split evenly between the two factors. However, the same is not true of l. For this reason, we treat both remainders as Low-High terms. The difference between T^1 and T^2 here is insignificant since it is straightforward to gain a power of $\langle \partial_v\rangle^{-1}$ for v''. Hence we focus only on $\mathrm{T}^1_{\mathcal{R}}$.

Begin with $\mathcal{M}_1 \mathrm{T}^1_{\mathcal{R}}$ and divide into two cases based on the relative size of ξ' and l,

$$\left|\widehat{\mathcal{M}_1 \mathrm{T}^1_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}} \sum_{\mathrm{M}/8\leq \mathrm{M}'\leq 8\mathrm{M}} \int [\mathbf{1}_{|l|>100|\xi'|} + \mathbf{1}_{|l|\leq 100|\xi'|}] \mathcal{M}_1(t,l,\xi) \\ \times \left|\mathrm{G}(\xi-\xi')_{\mathrm{M}'}\right| \left|\xi' - lt\right|^2 \left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right| d\xi' \\ = \left|\widehat{\mathcal{M}_1 \mathrm{T}^{1,z}_{\mathcal{R}}}(l,\xi)\right| + \left|\widehat{\mathcal{M}_1 \mathrm{T}^{1,v}_{\mathcal{R}}}(l,\xi)\right|.$$

Consider first $\mathcal{M}_1 \mathrm{T}^{1,z}_{\mathcal{R}}$. Since on the support of the integrand,

$$\left||l,\xi| - |l,\xi'|\right| \leq |\xi-\xi'| \leq \frac{3\mathrm{M}'}{2} \leq 12\mathrm{M} \leq 24|\xi'| \leq \frac{24}{100}|l,\xi'|, \tag{4.11}$$

inequality (A.7) implies,

$$\left|\mathcal{M}_1\widehat{\mathrm{T}^{1,z}_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}/8\le\mathrm{M}'\le 8\mathrm{M}}\int \mathbf{1}_{|l|>100|\xi'|}\left\langle\frac{\xi}{lt}\right\rangle^{-1}\frac{|l,\xi'|^{s/2}}{\langle t\rangle^s}\langle l,\xi'\rangle^{\sigma}\mathrm{J}_l(\xi)$$
$$\times e^{c\lambda|\xi-\xi'|^s}\left|\mathrm{G}(\xi-\xi')_{\mathrm{M}'}\right|\left|\xi'-lt\right|^2\left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right|e^{\lambda|l,\xi'|^s}d\xi'.$$

By $\frac{1}{24}|\xi'|\le|\xi-\xi'|\le 24|\xi'|$ we have the rough bound from (3.11),

$$\frac{\mathrm{J}_l(\xi)}{\mathrm{J}_l(\xi')}\lesssim e^{\frac{3\mu}{2}|\xi|^{1/2}}\lesssim e^{50\mu|\xi-\xi'|^{1/2}}. \tag{4.12}$$

Therefore, (A.11) implies

$$\left|\widehat{\mathcal{M}_1\mathrm{T}^{1,z}_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}/8\le\mathrm{M}'\le 8\mathrm{M}}\int \mathbf{1}_{|l|>100|\xi'|}\left\langle\frac{\xi}{lt}\right\rangle^{-1}\frac{|l,\xi|^{s/2}}{\langle t\rangle^s}$$
$$\times e^{\lambda|\xi-\xi'|^s}\left|\mathrm{G}(\xi-\xi')_{\mathrm{M}'}\right|\left|\xi'-lt\right|^2\mathrm{A}_l(\xi')\left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right|d\xi'.$$

Since $t\ge 1$, and $\langle\xi'/lt\rangle\approx 1$ on the support of the integrand, by (4.11),

$$\left|\widehat{\mathcal{M}_1\mathrm{T}^{1,z}_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}/8\le\mathrm{M}'\le 8\mathrm{M}}\int \mathbf{1}_{|l|>100|\xi'|}\mathcal{M}_1\left(t,l,\xi'\right)e^{\lambda|\xi-\xi'|^s}$$
$$\times\left|\mathrm{G}(\xi-\xi')_{\mathrm{M}'}\right|\left|\xi'-lt\right|^2\left|\hat{\phi}_l(\xi')_{\mathrm{M}}\right|d\xi'.$$

Taking the L^2 norm in ξ, applying (A.3) and (A.2) (note the M' sum only contains 7 terms),

$$\begin{aligned}\left\|\mathcal{M}_1\mathrm{T}^{1,z}_{\mathcal{R}}(l)\right\|_2 &\lesssim \sum_{\mathrm{M}'\in\mathbf{D}}\left\|\left(1-(v')^2\right)_{\mathrm{M}'}\right\|_{\mathcal{G}^{\lambda,2}}\sum_{\mathrm{M}'/8\le\mathrm{M}\le 8\mathrm{M}'}\left\|\mathcal{M}_1\Delta_{\mathrm{L}}(\phi_l)_{\mathrm{M}}\right\|_2\\ &\lesssim \sum_{\mathrm{M}'\in\mathbf{D}}\left\|\left(1-(v')^2\right)_{\mathrm{M}'}\right\|_{\mathcal{G}^{\lambda,2}}\left\|\mathcal{M}_1\Delta_{\mathrm{L}}(\phi_l)_{\sim\mathrm{M}'}\right\|_2\\ &\lesssim\left(\sum_{\mathrm{M}'\in\mathbf{D}}\left\|\left(1-(v')^2\right)_{\mathrm{M}'}\right\|^2_{\mathcal{G}^{\lambda,2}}\right)^{1/2}\\ &\quad\times\left(\sum_{\mathrm{M}'\in\mathbf{D}}\left\|\mathcal{M}_1\Delta_{\mathrm{L}}(\phi_l)_{\sim\mathrm{M}'}\right\|_2^2\right)^{1/2}\\ &\lesssim\epsilon\|\mathcal{M}_1\Delta_{\mathrm{L}}\phi_l\|_2,\end{aligned}$$

where the last line followed from (3.41a). Taking squares and summing over $l\neq 0$ implies

$$\left\|\mathcal{M}_1\mathrm{T}^{1,z}_{\mathcal{R}}\right\|_2\lesssim\epsilon^2\|\mathcal{M}_1\Delta_{\mathrm{L}}\mathrm{P}_{\neq 0}\phi\|_2^2. \tag{4.13}$$

Turn now to $\mathcal{M}_1 \mathrm{T}^{1,v}_{\mathcal{R}}$. On the support of the integrand in this case,

$$\begin{aligned} \left|\xi - \xi'\right| &\leq 24\left|\xi'\right| \leq 24\left|l, \xi'\right| \\ \left|l, \xi'\right| &\leq 101\left|\xi'\right| \leq 2424\left|\xi - \xi'\right|. \end{aligned}$$

Therefore by (A.8) there exists a $c \in (0,1)$ such that

$$|l, \xi|^s \leq \left|\left|l, \xi'\right| + \left|\xi - \xi'\right|\right|^s \leq c\left|l, \xi'\right|^s + c\left|\xi - \xi'\right|^s.$$

Hence,

$$\begin{aligned} \left|\widehat{\mathcal{M}_1 \mathrm{T}^{1,v}_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M} \in \mathbf{D}} \sum_{\mathrm{M}/8 \leq \mathrm{M}' \leq 8\mathrm{M}} \int \mathbf{1}_{|l| \leq 100|\xi'|} \left\langle \frac{\xi}{lt} \right\rangle^{-1} \frac{|l,\xi|^{s/2}}{\langle t \rangle^s} \left\langle l, \xi' \right\rangle^{\sigma/2} \\ \times \left\langle \xi - \xi' \right\rangle^{\sigma/2} \mathrm{J}_l(\xi) e^{c\lambda|\xi-\xi'|^s} \left|\mathrm{G}\left(\xi - \xi'\right)_{\mathrm{M}'}\right| \left|\xi' - lt\right|^2 \\ \times \left|\hat{\phi}_l\left(\xi'\right)_{\mathrm{M}}\right| e^{c\lambda|l,\xi'|^s} d\xi'. \end{aligned}$$

Using again (4.12) and (A.11) with $|l, \xi'| \approx |\xi - \xi'|$, and $\langle \frac{\xi'}{lt} \rangle \langle \frac{\xi}{lt} \rangle^{-1} \lesssim \langle \xi' \rangle$ implies

$$\begin{aligned} \left|\widehat{\mathcal{M}_1 \mathrm{T}^{1,v}_{\mathcal{R}}}(l,\xi)\right| \lesssim \sum_{\mathrm{M} \in \mathbf{D}} \sum_{\mathrm{M}/8 \leq \mathrm{M}' \leq 8\mathrm{M}} \int \mathbf{1}_{|l| \leq 100|\xi'|} \left\langle \frac{\xi'}{lt} \right\rangle^{-1} \frac{|l,\xi'|^{s/2}}{\langle t \rangle^s} e^{\lambda|\xi-\xi'|^s} \\ \times \left\langle \xi - \xi' \right\rangle^{\sigma/2+1} \left|\mathrm{G}\left(\xi - \xi'\right)_{\mathrm{M}'}\right| \left|\xi' - lt\right|^2 \mathrm{A}_l\left(\xi'\right) \\ \times \left|\hat{\phi}_l\left(\xi'\right)_{\mathrm{M}}\right| d\xi'. \end{aligned}$$

An argument similar to that used to complete the proof of $\|\mathcal{M}_1 \mathrm{T}^{1,z}_{\mathcal{R}}\|_2$ in (4.13) implies

$$\left\|\mathcal{M}_1 \mathrm{T}^{1,v}_{\mathcal{R}}\right\|_2^2 \lesssim \epsilon^2 \left\|\mathcal{M}_1 \Delta_{\mathrm{L}} \phi\right\|_2^2.$$

This completes the treatment of $\mathcal{M}_1 \mathrm{T}^1_{\mathcal{R}}$; as discussed above, $\mathcal{M}_1 \mathrm{T}^2_{\mathcal{R}}$ is treated in a similar manner and is hence omitted. Combining this argument with those used to treat $\mathcal{M}_2 \mathrm{T}^1_{\mathrm{LH}}$ we may also easily treat $\mathcal{M}_2 \mathrm{T}^1_{\mathcal{R}}$ and $\mathcal{M}_2 \mathrm{T}^2_{\mathcal{R}}$. The proof is omitted and the result is

$$\left\|\mathcal{M}_2 \mathrm{T}^1_{\mathcal{R}}\right\|_2^2 + \left\|\mathcal{M}_2 \mathrm{T}^2_{\mathcal{R}}\right\|_2^2 \lesssim \epsilon^2 \left\|\mathcal{M}_1 \Delta_{\mathrm{L}} \phi\right\|_2^2 + \epsilon^2 \left\|\mathcal{M}_2 \Delta_{\mathrm{L}} \phi\right\|_2^2.$$

This completes the treatment of the remainder terms and hence the proof of Proposition 2.4. $\square$

5. Transport

In this section we prove Proposition 2.2. As discussed above, we adapt methods similar to [33, 49, 54] for this purpose. This adaptation is not completely straightforward since $J_k(\eta)$ assigns slightly different regularities to modes which are near the critical time. Dealing with this will require special attention and all of the available time decay from the velocity field.

In the methods of [33, 49, 54] the goal is to gain $1-s$ derivatives from the difference $A_k(\eta) - A_l(\xi)$, and hence be able to absorb the leading contributions of T_N with CK_λ. Decompose this difference:

$$\begin{aligned} A_k(\eta) - A_l(\xi) &= A_l(\xi)\left[e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} - 1\right] \\ &\quad + A_l(\xi)e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s}\left[\frac{J_k(\eta)}{J_l(\xi)} - 1\right]\frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma} \\ &\quad + A_l(\xi)e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s}\left[\frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma} - 1\right]. \end{aligned}$$

In what follows we write

$$\begin{aligned} T_N &= i\sum_{k,l}\int_{\eta,\xi} A_k(\eta)\bar{\hat{f}}_k(\eta)\hat{u}_{k-l}(\eta-\xi)_{<N/8}\cdot(l,\xi)A_l(\xi)\hat{f}_l(\xi)_N \\ &\quad \times\left[e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} - 1\right]d\eta d\xi \\ &\quad + i\sum_{k,l}\int_{\eta,\xi} A_k(\eta)\bar{\hat{f}}_k(\eta)\hat{u}_{k-l}(\eta-\xi)_{<N/8}\cdot(l,\xi)A_l(\xi)\hat{f}_l(\xi)_N e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} \\ &\quad \times\left[\frac{J_k(\eta)}{J_l(\xi)} - 1\right]\frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma}d\eta d\xi \\ &\quad + i\sum_{k,l}\int_{\eta,\xi} A_k(\eta)\bar{\hat{f}}_k(\eta)\hat{u}_{k-l}(\eta-\xi)_{<N/8}\cdot(l,\xi)A_l(\xi)\hat{f}_l(\xi)_N e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} \\ &\quad \times\left[\frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma} - 1\right]d\eta d\xi \\ &= T_{N;1} + T_{N;2} + T_{N;3}. \end{aligned}$$

5.1. *Term* $T_{N;1}$*: exponential regularity*

First we treat $T_{N;1}$ for which the methods of [33, 49, 54] easily adapt. By $|e^x - 1| \le xe^x$ and (A.6),

$$|T_{N;1}| \leq \sum_{k,l} \int_{\eta,\xi} \left|A\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| |l,\xi| A_l(\xi) \left|\hat{f}_l(\xi)_N\right| \times \lambda \left| |k,\eta|^s - |l,\xi|^s \right| e^{\lambda ||k,\eta|^s - |l,\xi|^s|} d\eta d\xi$$
$$\lesssim \lambda \sum_{k,l} \int_{\eta,\xi} \left|A\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| |l,\xi| A_l(\xi) \left|\hat{f}_l(\xi)_N\right| \times \frac{||k,\eta|-|l,\xi||}{|k,\eta|^{1-s}+|l,\xi|^{1-s}} e^{\lambda ||k,\eta|^s - |l,\xi|^s|} d\eta d\xi.$$

Since on the support of the integrand (see Section A.1),

$$\left| |k,\eta| - |l,\xi| \right| \leq |k-l,\eta-\xi| \leq \frac{6}{32}|l,\xi|, \tag{5.1a}$$

$$(26/32)|l,\xi| \leq |k,\eta| \leq (38/32)|l,\xi|, \tag{5.1b}$$

inequalities (A.7) and (A.12) imply, for some $c \in (0,1)$,

$$|T_{N;1}| \lesssim \lambda \sum_{k,l} \int_{\eta,\xi} \left|A\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| |l,\xi|^{s/2} |k,\eta|^{s/2} A_l(\xi) \times \left|\hat{f}_l(\xi)_N\right| e^{c\lambda|k-l,\eta-\xi|^s} d\eta d\xi.$$

Hence (A.4) implies (since $\sigma > 6$),

$$|T_{N;1}| \lesssim \lambda \left\| |\nabla|^{s/2} A f_{\sim N} \right\|_2 \left\| |\nabla|^{s/2} A f_N \right\|_2 \|u_{<N/8}\|_{\mathcal{G}^{\lambda;\sigma-4}}.$$

Therefore by Lemma 4.1 and the bootstrap hypotheses,

$$|T_{N;1}| \lesssim \epsilon \frac{\lambda}{\langle t \rangle^{2-K_D\epsilon/2}} \left\| |\nabla|^{s/2} A f_{\sim N} \right\|_2 \left\| |\nabla|^{s/2} A f_N \right\|_2. \tag{5.2}$$

5.2. *Term* $T_{N;2}$*: effect of* J

The most difficult of the three terms in T_N is $T_{N;2}$ since J is sensitive to where it is being evaluated in (t,k,η). We divide the integral as follows

$$T_{N;2} = i \sum_{k,l} \int_{\eta,\xi} [\chi^S + \chi^L] A\bar{\hat{f}}_k(\eta) \hat{u}_{k-l}(\eta-\xi)_{<N/8} \cdot (l,\xi) A_l(\xi) \hat{f}_l(\xi)_N \times e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} \left[\frac{J_k(\eta)}{J_l(\xi)} - 1 \right] \frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma} d\eta d\xi$$
$$= T^S_{N;2} + T^L_{N;2},$$

where $\chi^S = \mathbf{1}_{t \leq \frac{1}{2}\min(\sqrt{|\xi|},\sqrt{|\eta|})}$ and $\chi^L = 1 - \chi^S$.

Focus first on $T^{S}_{N;2}$. In this term we apply Lemma 3.7 to gain 1/2 derivatives. Indeed, on the support of the integrand, (5.1) holds and hence by (A.7) and Lemma 3.7 we deduce

$$\left|T^{S}_{N;2}\right| \lesssim \sum_{k,l} \int_{\eta,\xi} \chi^{S} \left|A\hat{f}_k(\eta)\right| |l,\xi|^{1/2} \left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| A_l(\xi) \left|\hat{f}_l(\xi)_N\right| \times \langle k-l,\eta-\xi\rangle e^{c\lambda|k-l,\eta-\xi|^s} d\xi\, d\eta.$$

Since $c < 1$ it follows by (5.1), (A.11) and (A.12) that

$$\begin{aligned}\left|T^{S}_{N;2}\right| &\lesssim \sum_{k,l} \int_{\eta,\xi} \chi^{S} \left|A\hat{f}_k(\eta)\right| |l,\xi|^{1/2} \left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| A_l(\xi) \left|\hat{f}_l(\xi)_N\right| \\ &\quad \times e^{\lambda|k-l,\eta-\xi|^s} d\xi\, d\eta \\ &\lesssim \sum_{k,l} \int_{\eta,\xi} \chi^{S} \left|A\hat{f}_k(\eta)\right| \left(1 + |l,\xi|^{s/2}|k,\eta|^{s/2}\right) \left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| \\ &\quad \times A_l(\xi) \left|\hat{f}_l(\xi)_N\right| e^{\lambda|k-l,\eta-\xi|^s} d\xi\, d\eta.\end{aligned}$$

Hence by (A.4) followed by the bootstrap hypotheses and Lemma 4.1,

$$\begin{aligned}\left|T^{S}_{N;2}\right| &\lesssim \|u_{<N/8}\|_{\mathcal{G}^{\lambda,\sigma-4}} \left\| |\nabla|^{s/2} A f_{\sim N}\right\|_2 \left\| |\nabla|^{s/2} A f_N \right\|_2 \\ &\quad + \|u_{<N/8}\|_{\mathcal{G}^{\lambda,\sigma-4}} \|A f_{\sim N}\|_2 \|A f_N\|_2 \\ &\lesssim \frac{\epsilon}{\langle t\rangle^{2-K_D\epsilon/2}} \left\| |\nabla|^{s/2} A f_{\sim N}\right\|_2^2 + \frac{\epsilon}{\langle t\rangle^{2-K_D\epsilon/2}} \|A f_{\sim N}\|_2^2.\end{aligned} \tag{5.3}$$

Now focus on the more difficult $T^{L}_{N;2}$, where the resonant and non-resonant modes are being assigned slightly different regularities. There is a potential problem if Lemma 3.6 incurs a loss. Hence we divide into the two natural cases:

$$\begin{aligned}T^{L}_{N;2} &= i \sum_{k,l} \int_{\eta,\xi} \chi^{L} A_k(\eta) \bar{\hat{f}}_k(\eta) \hat{u}_{k-l}(\eta-\xi)_{<N/8} \cdot (l,\xi) \left[\chi^{D} + \chi^{*}\right] A_l(\xi) \\ &\quad \times \hat{f}_l(\xi)_N e^{\lambda|k,\eta|^s - \lambda|l,\xi|^s} \left[\frac{J_k(\eta)}{J_l(\xi)} - 1\right] d\eta\, d\xi \\ &= T^{D}_{N;2} + T^{*}_{N;2},\end{aligned}$$

where $\chi^{D} = \mathbf{1}_{t\in \mathbf{I}_{k,\eta}} \mathbf{1}_{t\in \mathbf{I}_{k,\xi}} \mathbf{1}_{k\neq l}$ and $\chi^{*} = 1 - \chi^{D}$ ('D' for 'difficult').

First focus on $T^{D}_{N;2}$ which is expected to be challenging. We will throw away any possible gain from the -1 and apply (3.34); the time decay will make this unimportant. By (5.1) and (A.7) (note also that on the support of the integrand, $|\eta| \gtrsim 1$),

$$\left|\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\left|\mathrm{A}\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right||l,\xi|\chi^{\mathrm{D}}\mathrm{A}_l(\xi)\left|\hat{f}_l(\xi)_{\mathrm{N}}\right|$$
$$\times e^{c\lambda|k-l,\eta-\xi|^s}\frac{|\eta|}{k^2}\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}}e^{20\mu|k-l,\eta-\xi|^{1/2}}d\eta d\xi.$$

Applying (A.11) implies

$$\left|\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\left|\mathrm{A}\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|\chi^{\mathrm{D}}\mathrm{A}_l(\xi)\left|\hat{f}_l(\xi)_{\mathrm{N}}\right||l,\xi|$$
$$\times e^{\lambda|k-l,\eta-\xi|^s}\frac{|\eta|}{k^2}\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}}d\eta d\xi.$$

On the support of the integrand $\mathrm{A}_l(\xi)\lesssim\tilde{\mathrm{A}}_l(\xi)$ and $\mathrm{A}_k(\eta)\lesssim\tilde{\mathrm{A}}_k(\eta)$ and since $1\leq t\approx\frac{\eta}{k}$,

$$\left|\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\left|\tilde{\mathrm{A}}\hat{f}_k(\eta)\right|t^2\left|\hat{u}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|\chi^{\mathrm{D}}\tilde{\mathrm{A}}_l(\xi)\left|\hat{f}_l(\xi)_{\mathrm{N}}\right|$$
$$\times\frac{|l,\xi|}{|\eta|}e^{\lambda|k-l,\eta-\xi|^s}\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}}d\eta d\xi.$$

By the definition of χ^{D} and (5.1), we have $|l,\xi|\lesssim|\eta|$, hence (A.4) implies:

$$\left|\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}\right| \lesssim t^2\|\mathrm{P}_{\neq 0}u_{<\mathrm{N}/8}\|_{\mathcal{G}^{\lambda,\sigma-4}}\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\sim\mathrm{N}}\right\|_2\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\mathrm{N}}\right\|_2,$$

where note that since $k\neq l$ by the definition of χ^{D}, we may restrict to non-zero modes in u, crucial to get the full $O(t^{-2})$ decay. Therefore by Lemma 4.1 and the bootstrap hypotheses,

$$\left|\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}\right| \lesssim \epsilon\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\sim\mathrm{N}}\right\|_2\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\mathrm{N}}\right\|_2. \tag{5.4}$$

This completes the treatment of $\mathrm{T}_{\mathrm{N};2}^{\mathrm{D}}$.

It remains to treat $\mathrm{T}_{\mathrm{N};2}^{*}$. We divide into two cases based on the relative size of $|l|$ and $|\xi|$:

$$\mathrm{T}_{\mathrm{N};2}^{*} = i\sum_{k,l}\int_{\eta,\xi}\chi^{\mathrm{L}}\mathrm{A}_k(\eta)\bar{\hat{f}}_k(\eta)\hat{u}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\cdot(l,\xi)\chi^{*}$$
$$\times\left[\mathbf{1}_{|l|>100|\xi|}+\mathbf{1}_{|l|\leq 100|\xi|}\right]\mathrm{A}_l(\xi)\hat{f}_l(\xi)_{\mathrm{N}}$$

$$\times e^{\lambda|k,\eta|^s-\lambda|l,\xi|^s}\left[\frac{J_k(\eta)}{J_l(\xi)}-1\right]d\eta d\xi$$
$$= T^{*,z}_{N;2}+T^{*,v}_{N;2}.$$

First consider $T^{*,z}_{N;2}$. On the support of the integrand, note $|\eta| < \frac{313}{1000}|l|$ and hence by (A.6), (3.11) and (A.12),

$$\begin{aligned}\left|\frac{J_k(\eta)}{J_l(\xi)}-1\right| &= \left|\frac{w_k(t,\eta)^{-1}e^{\mu|\eta|^{1/2}}+e^{\mu|k|^{1/2}}}{w_l(t,\xi)^{-1}e^{\mu|\xi|^{1/2}}+e^{\mu|l|^{1/2}}}-1\right| \\ &\lesssim e^{\frac{3}{2}\mu|\eta|^{1/2}-\mu|l|^{1/2}}+\left|e^{\mu|k|^{1/2}-\mu|l|^{1/2}}-1\right| \\ &\lesssim_\mu \frac{1}{|l|^{1/2}}+\frac{|k-l|}{|k|^{1/2}+|l|^{1/2}}e^{\mu|k-l|^{1/2}}.\end{aligned}$$

Therefore by (A.11), (A.12) and $|l,\xi|\lesssim|l|$,

$$\begin{aligned}\left|T^{*,z}_{N;2}\right| &\lesssim \sum_{k,l}\int_{\eta,\xi}\chi^L\chi^*\left|A\hat{f}_k(\eta)\right|\mathbf{1}_{|l|>100|\xi|}|l|^{1/2}\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| \\ &\quad\times A_l(\xi)\left|\hat{f}_l(\xi)_N\right|e^{\lambda|k-l,\eta-\xi|^s}d\eta d\xi \\ &\lesssim \sum_{k,l}\int_{\eta,\xi}\chi^L\chi^*\left|A\hat{f}_k(\eta)\right|\mathbf{1}_{|l|>100|\xi|}|l|^s\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right| \\ &\quad\times A_l(\xi)\left|\hat{f}_l(\xi)_N\right|e^{\lambda|k-l,\eta-\xi|^s}d\eta d\xi.\end{aligned}$$

Applying (A.4), the bootstrap hypotheses and Lemma 4.1 as in the treatment of $T^S_{N;2}$,

$$\left|T^{*,z}_{N;2}\right|\lesssim\frac{\epsilon}{\langle t\rangle^{2-K_D\epsilon/2}}\left\||\nabla|^{s/2}Af_{\sim N}\right\|_2^2. \tag{5.5}$$

Turn now to $T^{*,v}_{N;2}$. Note that on the support of the integral, $|\eta|\approx|\xi|$. By definition of χ^*, we may apply (3.31), which together with (A.7) and (A.11) implies

$$\begin{aligned}\left|T^{*,v}_{N;2}\right| &\lesssim \sum_{k,l}\int_{\eta,\xi}\chi^L\chi^*\mathbf{1}_{|l|\le100|\xi|}\left|A\hat{f}_k(\eta)\right|\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right||l,\xi| \\ &\quad\times A_l(\xi)\left|\hat{f}_l(\xi)_N\right|e^{\lambda|k-l,\eta-\xi|^s}d\eta d\xi.\end{aligned}$$

Since $|l,\xi|\lesssim|\xi|\lesssim t^2$, we have $|l,\xi|\lesssim|k,\eta|^{s/2}|l,\xi|^{s/2}t^{2-2s}$ and therefore,

$$\begin{aligned}\left|T^{*,v}_{N;2}\right| &\lesssim \sum_{k,l}\int_{\eta,\xi}\chi^*\chi^L|k,\eta|^{s/2}\left|A\hat{f}_k(\eta)\right|t^{2-2s}\left|\hat{u}_{k-l}(\eta-\xi)_{<N/8}\right||l,\xi|^{s/2} \\ &\quad\times A_l(\xi)\left|\hat{f}_l(\xi)_N\right|e^{\lambda|k-l,\eta-\xi|^s}d\eta d\xi.\end{aligned}$$

By (A.4), Lemma 4.1 and the bootstrap hypotheses,

$$\left|\mathrm{T}^{*,v}_{\mathrm{N};2}\right| \lesssim \frac{\epsilon}{\langle t\rangle^{2s-\mathrm{K_D}\epsilon/2}}\left\| |\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_2^2, \tag{5.6}$$

where note $2s - \mathrm{K_D}\epsilon/2 \geq s + 1/2$ for ϵ sufficiently small.

Combining (5.3), (5.4), (5.5) and (5.6) we have for ϵ sufficiently small,

$$|\mathrm{T}_{\mathrm{N};2}| \lesssim \frac{\epsilon}{\langle t\rangle^{s+1/2}}\left\| |\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_2^2 + \epsilon\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\sim\mathrm{N}}\right\|_2^2 + \frac{\epsilon}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}\|\mathrm{A}f_{\sim\mathrm{N}}\|_2^2, \tag{5.7}$$

which completes the treatment of $\mathrm{T}_{\mathrm{N};2}$.

5.3. *Term* $\mathrm{T}_{\mathrm{N};3}$*: Sobolev correction*

Next, turn to $\mathrm{T}_{\mathrm{N};3}$ which is the easiest to treat. By the mean value theorem and (5.1),

$$\left|\frac{\langle k,\eta\rangle^\sigma}{\langle l,\xi\rangle^\sigma} - 1\right| \lesssim \frac{|k-l,\eta-\xi|}{\langle l,\xi\rangle},$$

which implies arguments similar to those applied above can deduce

$$|\mathrm{T}_{\mathrm{N};3}| \lesssim \frac{\epsilon}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}\|\mathrm{A}f_{\sim\mathrm{N}}\|_2\|\mathrm{A}f_{\mathrm{N}}\|_2. \tag{5.8}$$

Indeed, putting (5.2), (5.7) and (5.8) together with (A.2) proves Proposition 2.2.

6. Reaction

Focus first on an individual frequency shell and divide each one into several natural pieces

$$\mathrm{R_N} = \mathrm{R}^1_{\mathrm{N}} + \mathrm{R}^{\epsilon,1}_{\mathrm{N}} + \mathrm{R}^2_{\mathrm{N}} + \mathrm{R}^3_{\mathrm{N}}$$

where

$$\mathrm{R}^1_{\mathrm{N}} = \sum_{k,l\neq 0}\int_{\eta,\xi} \mathrm{A}\bar{\hat{f}}_k(\eta)\mathrm{A}_k(\eta)(\eta l - \xi k)\hat{\phi}_l(\xi)_{\mathrm{N}}\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\,d\eta\, d\xi$$

$$\mathrm{R}^{\epsilon,1}_{\mathrm{N}} = -\sum_{k,l\neq 0}\int_{\eta,\xi} \mathrm{A}\bar{\hat{f}}_k(\eta)\mathrm{A}_k(\eta)\left[\left(1-\widehat{v'\right)\nabla^\perp\phi_l}\right](\xi)_{\mathrm{N}} \cdot \widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\,d\eta\, d\xi$$

$$R_N^2 = \sum_k \int_{\eta,\xi} A\bar{\hat{f}}_k(\eta) A_k(\eta) \widehat{[\partial_t v]}(\xi)_N \widehat{\partial_v f}_k(\eta-\xi)_{<N/8} d\eta d\xi$$

$$R_N^3 = -\sum_{k,l} \int_{\eta,\xi} A\bar{\hat{f}}_k(\eta) A_{k-l}(\eta-\xi) \hat{u}_l(\xi)_N \widehat{\nabla f}_{k-l}(\eta-\xi)_{<N/8} d\eta d\xi.$$

6.1. *Main contribution*

The main contribution comes from R_N^1. We subdivide this integral depending on whether or not (l, ξ) and/or (k, η) are resonant as each combination requires a slightly different treatment. Define the partition:

$$\begin{aligned} 1 &= \mathbf{1}_{t\notin \mathbf{I}_{k,\eta}, t\notin \mathbf{I}_{l,\xi}} + \mathbf{1}_{t\notin \mathbf{I}_{k,\eta}, t\in \mathbf{I}_{l,\xi}} + \mathbf{1}_{t\in \mathbf{I}_{k,\eta}, t\notin \mathbf{I}_{l,\xi}} + \mathbf{1}_{t\in \mathbf{I}_{k,\eta}, t\in \mathbf{I}_{l,\xi}} \\ &= \chi^{NR,NR} + \chi^{NR,R} + \chi^{R,NR} + \chi^{R,R}, \end{aligned}$$

where the NR and R denotes 'non-resonant' and 'resonant' respectively referring to (k, η) and (l, ξ). Correspondingly, denote

$$\begin{aligned} R_N^1 &= \sum_{k,l\neq 0} \int_{\eta,\xi} [\chi^{NR,NR} + \chi^{NR,R} + \chi^{R,NR} + \chi^{R,R}] \\ &\quad \times A\bar{\hat{f}}_k(\eta) A_k(\eta)(\eta l - \xi k)\hat{\phi}_l(\xi)_N \hat{f}_{k-l}(\eta-\xi)_{<N/8} d\eta d\xi \\ &= R_N^{NR,NR} + R_N^{NR,R} + R_N^{R,NR} + R_N^{R,R}. \end{aligned}$$

6.1.1. *Treatment of* $R_N^{NR,NR}$

Since on the support of the integrand of R_N^1,

$$\big||l,\xi| - |k,\eta|\big| \leq |k-l, \eta-\xi| \leq \frac{6}{32}|l,\xi|, \tag{6.1}$$

it follows from (A.7) that for some $c \in (0, 1)$,

$$\begin{aligned} \left|R_N^{NR,NR}\right| &\leq \sum_{k,l\neq 0} \int_{\eta,\xi} \left|\chi^{NR,NR} A\bar{\hat{f}}_k(\eta) J_k(\eta) e^{\lambda|l,\xi|^s} e^{c\lambda|k-l,\eta-\xi|^s}\right| \\ &\quad \times \langle k,\eta\rangle^\sigma |\eta l - \xi k| \left|\hat{\phi}_l(\xi)_N \hat{f}_{k-l}(\eta-\xi)_{<N/8}\right| d\eta d\xi. \end{aligned}$$

Moreover, by (3.31), (6.1) (which implies $|k,\eta| \approx |l,\xi|$) and (A.11),

$$\begin{aligned} \left|R_N^{NR,NR}\right| &\lesssim \sum_{k,l\neq 0} \int_{\eta,\xi} \chi^{NR,NR} \left|A\hat{f}_k(\eta)\right| e^{\lambda|k-l,\eta-\xi|^s} |\eta l - \xi k| \\ &\quad \times A_l(\xi) \left|\hat{\phi}_l(\xi)_N \hat{f}_{k-l}(\eta-\xi)_{<N/8}\right| d\eta d\xi. \end{aligned}$$

Again by (6.1), $|\eta l - \xi k| \lesssim |l,\xi|^{1-s/2}|k,\eta|^{s/2}|k-l,\eta-\xi|$ which implies by (A.4),

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,NR}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi} |k,\eta|^{s/2}\chi^{\mathrm{NR,NR}}\left|\mathrm{A}\hat{f}_k(\eta)\right||l,\xi|^{1-s/2}\mathrm{A}_l(\xi)\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\right| \times |k-l,\eta-\xi|\left|\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|e^{\lambda|k-l,\eta-\xi|^s}d\eta d\xi \tag{6.2}$$

$$\lesssim \left\||\nabla|^{s/2}\mathrm{A}f_{\sim \mathrm{N}}\right\|_2\left\||\nabla|^{1-s/2}\mathrm{AP}_{\neq 0}\chi^{\mathrm{NR}}\phi_{\mathrm{N}}\right\|_2\|f\|_{\mathcal{G}^{\lambda,\sigma}}$$

$$\lesssim \epsilon\left\||\nabla|^{s/2}\mathrm{A}f_{\sim \mathrm{N}}\right\|_2\left\||\nabla|^{1-s/2}\mathrm{AP}_{\neq 0}\chi^{\mathrm{NR}}\phi_{\mathrm{N}}\right\|_2, \tag{6.3}$$

where the last line followed from the bootstrap hypotheses. Here we are denoting $\chi^{\mathrm{NR}}f$ the multiplier $\widehat{\chi^{\mathrm{NR}}f}(t,l,\xi) = \mathbf{1}_{t\notin \mathbf{I}_{l,\xi}}\hat{f}_l(t,\xi)$.

6.1.2. *Treatment of* $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR}}$

Next we turn to $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR}}$ which is one of the terms w was designed to treat. Physically, it describes the action of the non-resonant modes on the resonant modes. By (6.1) and (A.7), for some $c\in(0,1)$,

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\chi^{\mathrm{R,NR}}\left|\mathrm{A}\hat{f}_k(\eta)\right|\mathrm{J}_k(\eta)e^{\lambda|l,\xi|^s}e^{c\lambda|k-l,\eta-\xi|^s}\langle l,\xi\rangle^\sigma|l,\xi| \times\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.$$

Consider separately the following cases:

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\chi^{\mathrm{R,NR}}[\mathbf{1}_{t\in\mathbf{I}_{k,\xi}}+\mathbf{1}_{t\notin\mathbf{I}_{k,\xi}}]\left|\mathrm{A}\hat{f}_k(\eta)\right|\mathrm{J}_k(\eta) \times e^{\lambda|l,\xi|^s}e^{c\lambda|k-l,\eta-\xi|^s}\langle l,\xi\rangle^\sigma|l,\xi|\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi$$

$$= \mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;D}} + \mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;*}}.$$

The toy model is adapted to treat $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;D}}$, so consider this first. Note that on the support of the integrand in this case, we have $|\eta|\approx|\xi|$. Therefore, applying (3.34) and (A.11),

$$\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;D}} \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\chi^{\mathrm{R,NR}}\mathbf{1}_{t\in\mathbf{I}_{k,\xi}}\left|\mathrm{A}\hat{f}_k(\eta)\right|\mathrm{J}_l(\xi)\frac{|\eta|}{k^2}\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}} \times e^{\lambda|l,\xi|^s}e^{\lambda|k-l,\eta-\xi|^s}\langle l,\xi\rangle^\sigma|l,\xi|\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.$$

Since $l^2k^{-2}\leq\langle l-k\rangle^2$, $|\eta|\approx|\xi|$ and $|l|<\frac{1}{4}|\xi|$ (hence $\mathrm{J}_l(\xi)\lesssim\tilde{\mathrm{J}}_l(\xi)$),

$$\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;D}} \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi} \chi^{\mathrm{R,NR}} \mathbf{1}_{t\in \mathbf{I}_{k,\xi}} \left|\tilde{\mathrm{A}}\hat{f}_k(\eta)\right| \tilde{\mathrm{J}}_l(\xi) \frac{|\xi|^2}{l^2\langle k-l\rangle^2}$$
$$\times \sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{\partial_t w_l(t,\xi)}{w_l(t,\xi)}} e^{\lambda|l,\xi|^s} e^{\lambda|k-l,\eta-\xi|^s}\langle l,\xi\rangle^\sigma$$
$$\times \left|\hat{\phi}_l(\xi)_{\mathrm{N}}\langle k-l\rangle^4 \widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right| d\eta d\xi.$$

Applying $|k-t^{-1}\eta|\leq 1$, (A.4) and the bootstrap hypotheses (denoting $\chi^r(t,\eta) = \mathbf{1}_{2\sqrt{|\eta|}\leq t\leq 2|\eta|}$):

$$\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;D}} \lesssim \epsilon \left\|\sqrt{\frac{\partial_t w}{w}}\mathrm{A}f_{\sim \mathrm{N}}\right\|_2 \left\|\sqrt{\frac{\partial_t w}{w}} \frac{|\partial_v|^2}{\partial_z^2\langle t^{-1}\partial_v - \partial_z\rangle^2}\chi^r\chi^{\mathrm{NR}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi_{\mathrm{N}}\right\|_2. \tag{6.4}$$

Turn now to $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;*}}$. In this case we may apply (3.31) (as opposed to (3.30)) and hence we can exchange $\mathrm{J}_k(\eta)$ for $\mathrm{J}_l(\xi)$ without incurring a major cost. Therefore, by applying the same argument used to treat $\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,NR}}$ we deduce:

$$\mathrm{R}_{\mathrm{N}}^{\mathrm{R,NR;*}} \lesssim \epsilon \left\||\nabla|^{s/2}\mathrm{A}f_{\sim \mathrm{N}}\right\|_2 \left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{A}\mathrm{P}_{\neq 0}\phi_{\mathrm{N}}\right\|_2. \tag{6.5}$$

6.1.3. *Treatment of* $\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,R}}$

The next term we treat is $\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,R}}$, in which case (k,η) is non-resonant and (l,ξ) is resonant. It follows that $4|l|^2\leq|\xi|$ and since $\mathrm{N}\geq 8$, (6.1) implies $\mathrm{N}/4\leq|\xi|\leq 3\mathrm{N}/2$ and $|\eta|\approx|\xi|$. By (3.16),

$$1 \lesssim \sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}}\left[\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}} + \frac{|k,\eta|^{s/2}}{\langle t\rangle^s}\right]\langle \eta-\xi\rangle. \tag{6.6}$$

Applying (6.6), (3.32), (A.7) (using (6.1)) (A.12) and (A.11) ($s>1/2$),

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,R}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\chi^{\mathrm{NR,R}}\left[\sqrt{\frac{\partial_t w_k(\eta)}{w_k(\eta)}} + \frac{|k,\eta|^{s/2}}{\langle t\rangle^s}\right]$$
$$\times \left|\mathrm{A}\hat{f}_k(\eta)\right|\mathrm{J}_l(\xi)\frac{w_{\mathrm{R}}(\xi)}{w_{\mathrm{NR}}(\xi)}\sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}}$$
$$\times e^{\lambda|l,\xi|^s}e^{\lambda|k-l,\eta-\xi|^s}\langle l,\xi\rangle^\sigma |l,\xi|\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\widehat{\nabla f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.$$

On the support of the integrand, it follows from (6.1) that $|k|<|\eta|$ and hence $\mathrm{A}_k(\eta)\lesssim \tilde{\mathrm{A}}_k(\eta)$. Similarly, $\mathrm{A}_l(\xi)\lesssim\tilde{\mathrm{A}}_l(\xi)$. Therefore (A.4) and the bootstrap hypotheses imply

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{NR,R}}\right| \lesssim \epsilon\left(\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}} f_{\sim \mathrm{N}}\right\|_2 + \frac{1}{\langle t\rangle^s}\left\||\nabla|^{s/2}\mathrm{A} f_{\sim \mathrm{N}}\right\|_2\right) \times \left\|\sqrt{\frac{w}{\partial_t w}}|\nabla|\frac{w_{\mathrm{R}}}{w_{\mathrm{NR}}}\chi^{\mathrm{R}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2. \tag{6.7}$$

6.1.4. *Treatment of* $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,R}}$

In this case both (k,η) and (l,ξ) are resonant, an interaction that was neglected in the derivation of the toy model. We claim that on the support of the integrand of $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,R}}$:

$$|\eta l - \xi k|\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim \sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\left[|l,\xi|\frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)} + |l|\right] \times \sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}} e^{12\mu|k-l,\eta-\xi|^{1/2}}. \tag{6.8}$$

Indeed, if $k = l$ then (3.31) and Lemma 3.4 imply,

$$|l||\eta-\xi|\frac{\mathrm{J}_l(\eta)}{\mathrm{J}_l(\xi)} \lesssim |l|\langle\eta-\xi\rangle^2\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}}e^{10\mu|k-l,\eta-\xi|^{1/2}},$$

from which (6.8) follows by (A.12). If $k \neq l$ then as in the proof of (3.32) we apply Lemma 3.2. If Lemma 3.2 (b) holds then by Lemma 3.4, (3.31) (note that on the support of the integrand $|\eta| \approx |\xi|$ by (6.1) with $k^2 < \frac{1}{4}|\eta|$, $l^2 < \frac{1}{4}|\xi|$) and the definitions (3.5), (3.8):

$$|\eta l - \xi k|\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} \lesssim |l,\xi|\langle k-l,\eta-\xi\rangle^2\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}} \times \frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)}e^{10\mu|k-l,\eta-\xi|^{1/2}},$$

which again implies (6.8) by (A.12). Finally, if Lemma 3.2 (c) holds then by Lemma 3.4, (3.31) and (A.12),

$$\begin{aligned}|\eta l - \xi k|\frac{\mathrm{J}_k(\eta)}{\mathrm{J}_l(\xi)} &\lesssim |l,\xi||k-l,\eta-\xi|e^{9\mu|k-l,\eta-\xi|^{1/2}} \\ &\lesssim |l|\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}}e^{10\mu|k-l,\eta-\xi|^{1/2}},\end{aligned}$$

which proves (6.8) in the remaining case. Applying (6.8), (A.7) (using (6.1)) and (A.11) implies

$$\begin{aligned}\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{R,R}}\right| \lesssim \sum_{k,l\neq 0}\int_{\eta,\xi}\chi^{\mathrm{R,R}}\sqrt{\frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)}}\left|\mathrm{A}\hat{f}_k(\eta)\right| \\ \times\left[|l,\xi|\frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)}+|l|\right]\sqrt{\frac{w_l(t,\xi)}{\partial_t w_l(t,\xi)}}\mathrm{A}_l(\xi)\left|\hat{\phi}_l(\xi)_{\mathrm{N}}\right|e^{\lambda|k-l,\eta-\xi|^s} \\ \times\left|\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.\end{aligned}$$

Since (k,η) and (l,ξ) are both resonant, $\mathrm{A}_k(\eta)\lesssim\tilde{\mathrm{A}}_k(\eta)$ and $\mathrm{A}_l(\xi)\lesssim\tilde{\mathrm{A}}_l(\xi)$. Then by (A.4) and the bootstrap hypotheses,

$$\left|\mathrm{R}_{\mathrm{N}}^{\mathrm{R,R}}\right|\lesssim\epsilon\left\|\sqrt{\frac{\partial_t w}{w}}\tilde{\mathrm{A}}f_{\sim\mathrm{N}}\right\|_2\left(\left\|\sqrt{\frac{w}{\partial_t w}}|\nabla|\frac{w_{\mathrm{R}}}{w_{\mathrm{NR}}}\chi^{\mathrm{R}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2 + \left\|\sqrt{\frac{w}{\partial_t w}}|\partial_z|\chi^{\mathrm{R}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2\right), \tag{6.9}$$

which completes the treatment of $\mathrm{R}_{\mathrm{N}}^{\mathrm{R,R}}$.

6.1.5. *Contribution to (2.27)*

Combining (6.3), (6.4), (6.5), (6.7), (6.9) and Cauchy-Schwarz we deduce,

$$\begin{aligned}\left|\mathrm{R}_{\mathrm{N}}^{1}\right| \lesssim \frac{\epsilon}{\langle t\rangle^{2s}}\left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_2^2+\epsilon\left\|\sqrt{\frac{\partial_t w}{w}}\mathrm{A}f_{\sim\mathrm{N}}\right\|_2^2 \\ +\epsilon\langle t\rangle^{2s}\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{A}\mathrm{P}_{\neq 0}\phi_{\mathrm{N}}\right\|_2^2 \\ +\epsilon\left\|\sqrt{\frac{\partial_t w}{w}}\frac{|\partial_v|^2}{\partial_z^2\langle t^{-1}\partial_v-\partial_z\rangle^2}\chi^r\chi^{\mathrm{NR}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2^2 \\ +\epsilon\left\|\sqrt{\frac{w}{\partial_t w}}|\nabla|\frac{w_{\mathrm{R}}}{w_{\mathrm{NR}}}\chi^{\mathrm{R}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2^2+\epsilon\left\|\sqrt{\frac{w}{\partial_t w}}|\partial_z|\chi^{\mathrm{R}}\tilde{\mathrm{A}}\phi_{\mathrm{N}}\right\|_2^2,\end{aligned} \tag{6.10}$$

where $\chi^r(t,\eta)=\mathbf{1}_{2\sqrt{|\eta|}\leq t\leq 2|\eta|}$. The treatment of $\mathrm{R}_{\mathrm{N}}^1$ will then be complete once we have the following lemma to relate the latter four terms to those in (2.27), on which Proposition 2.4 can be applied. The primary complication in (6.11) below is the leading factor $\langle\partial_v(t\partial_z)^{-1}\rangle^{-1}$ which will require some additional care to include. Recall the presence of this multiplier arises from the ∂_v in the expression of v'' (see (2.13c)) which appears as a coefficient in Δ_t; see Section 4.2 for more information. This lemma expresses something important: that the multipliers coming from w exactly 'match' the loss of ellipticity in Δ_{L}.

Lemma **6.1.** — *Under the bootstrap hypotheses,*

$$\langle t\rangle^{2s}\left\| |\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{AP}_{\neq 0}\phi\right\|_2^2 \lesssim \left\| \left\langle \frac{\partial_v}{t\partial_z}\right\rangle^{-1} \frac{|\nabla|^{s/2}}{\langle t\rangle^s}\Delta_{\mathrm{L}}\mathrm{AP}_{\neq 0}\phi\right\|_2^2 \tag{6.11a}$$

$$\left\| \sqrt{\frac{\partial_t w}{w}} \frac{|\partial_v|^2}{\partial_z^2\langle t^{-1}\partial_v - \partial_z\rangle^2}\chi^r\chi^{\mathrm{NR}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2 \lesssim \left\| \left\langle \frac{\partial_v}{t\partial_z}\right\rangle^{-1}\sqrt{\frac{\partial_t w}{w}}\Delta_{\mathrm{L}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2 \tag{6.11b}$$

$$\left\| \sqrt{\frac{w}{\partial_t w}}|\nabla|\frac{w_{\mathrm{R}}}{w_{\mathrm{NR}}}\chi^{\mathrm{R}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2 \lesssim \left\| \left\langle \frac{\partial_v}{t\partial_z}\right\rangle^{-1}\sqrt{\frac{\partial_t w}{w}}\Delta_{\mathrm{L}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2 \tag{6.11c}$$

$$\left\| \sqrt{\frac{w}{\partial_t w}}|\partial_z|\chi^{\mathrm{R}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2 \lesssim \left\| \left\langle \frac{\partial_v}{t\partial_z}\right\rangle^{-1}\sqrt{\frac{\partial_t w}{w}}\Delta_{\mathrm{L}}\tilde{\mathrm{A}}\mathrm{P}_{\neq 0}\phi\right\|_2^2. \tag{6.11d}$$

Proof. — Note that (6.11) is only a statement about the Fourier multipliers and has nothing really to do with $\mathrm{A}\phi$. Indeed, (6.11a) follows from the pointwise inequality: for all $t \geq 1$ $l \neq 0$ and $\xi \in \mathbf{R}$,

$$\langle t\rangle^s |l,\xi|^{1-s/2}\mathbf{1}_{t\notin \mathbf{I}_{l,\xi}} \lesssim \left\langle \frac{\xi}{lt}\right\rangle^{-1}\left(l^2 + |\xi - lt|^2\right)\frac{|l,\xi|^{s/2}}{\langle t\rangle^s}. \tag{6.12}$$

Proof of (6.12)*:* Consider the case $\frac{1}{2}|lt| \leq |\xi| \leq 2|lt|$. By the presence of $\mathbf{1}_{t\notin \mathbf{I}_{l,\xi}}$, either $|\xi - tl| \gtrsim |\xi/l| \approx t$ (if $t^2 \gtrsim |\xi|$) or $l^2 \gtrsim t$ (if $t^2 \lesssim |\xi|$). In either case,

$$\langle t\rangle^{2s}|l,\xi|^{1-s} \lesssim |l|^{1-s}t^{1+s} \leq l^2 + t^2 \lesssim l^2 + |\xi - lt|^2,$$

which implies (6.12). Next consider the case $|\xi| < |lt|/2$, which implies

$$\langle t\rangle^{2s}|l,\xi|^{1-s} \lesssim t^{2s}|lt|^{1-s} \lesssim |lt|^{1+s} \lesssim l^2 + |\xi - lt|^2,$$

which again implies (6.12). Finally consider the case $|\xi| \geq 2|lt|$, in which the leading $\langle \xi/lt\rangle^{-1}$ plays a role. In this case (note since $t \geq 1$, $|\xi| \geq 2|l|$),

$$\langle t\rangle^{2s}|l,\xi|^{1-s} \lesssim |\xi|^{2-s}\langle t\rangle^{2s-1}\frac{|lt|}{|\xi|} \lesssim |\xi|^{1+s}\frac{|lt|}{|\xi|} \lesssim \left(l^2 + |\xi - lt|^2\right)\frac{|lt|}{|\xi|},$$

which implies (6.12). As all cases have been covered, this proves (6.11a).

Proof of (6.11b)*:* Inequality (6.11b) follows from the pointwise inequality: Let j be such that $t \in \mathbf{I}_{j,\xi}$, then for all $t \geq 1$, $l \neq 0$ and $\xi \in \mathbf{R}$ with $2\sqrt{|\xi|} \leq t \leq 2|\xi|$:

$$\textbf{(6.13)} \qquad \frac{|\xi|^2}{|l|^2\langle j-l\rangle^2}\mathbf{1}_{t\notin\mathbf{I}_{l,\xi}} \lesssim \left\langle\frac{\xi}{lt}\right\rangle^{-1}\left(l^2+|\xi-lt|^2\right).$$

Since $t \notin \mathbf{I}_{l,\xi}$, $|\xi|^2 \lesssim l^2|\xi - lt|^2$. If $2|lt| \geq |\xi|$ then the factor $\langle\frac{\xi}{lt}\rangle^{-1}$ does not play a role and (6.13) follows immediately. Next consider the case $2|lt| \leq |\xi|$, which implies $|\xi - lt| \gtrsim |\xi|$. In this case, since $\xi \approx jt$ and $|j| \leq |l||l-j|$,

$$\begin{aligned}\frac{|\xi|^2}{|l|^2\langle j-l\rangle^2} &\lesssim \frac{|\xi|}{\langle j-l\rangle^2|lt|}\left|\frac{lt}{\xi}\right||\xi-lt|^2 \lesssim \frac{|j|}{|l|\langle j-l\rangle^2}\left|\frac{lt}{\xi}\right||\xi-lt|^2\\ &\lesssim \left|\frac{lt}{\xi}\right||\xi-lt|^2.\end{aligned}$$

This verifies (6.13) in every case and hence (6.11b).

Proof of (6.11c)*:* By the definitions of w_{R} and w_{NR} (3.5), (3.8) and (3.14),

$$\begin{aligned}\sqrt{\frac{w(t,\xi)}{\partial_t w(t,\xi)}}|l,\xi|\frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)}\mathbf{1}_{t\in\mathbf{I}_{l,\xi}} &\lesssim |\xi|\left(1+\left|t-\frac{\xi}{l}\right|\right)^{1/2}\frac{w_{\mathrm{R}}(t,\xi)}{w_{\mathrm{NR}}(t,\xi)}\mathbf{1}_{t\in\mathbf{I}_{l,\xi}}\\ &\lesssim \left(l^2+|\xi-tl|^2\right)\sqrt{\frac{\partial_t w(t,\xi)}{w(t,\xi)}},\end{aligned}$$

which proves (6.11c).

Proof of (6.11d)*:* Similarly, by (3.14)

$$\begin{aligned}\sqrt{\frac{w(t,\xi)}{\partial_t w(t,\xi)}}|l|\mathbf{1}_{t\in\mathbf{I}_{l,\xi}} &\lesssim |l|\sqrt{1+\left|t-\frac{\xi}{l}\right|}\mathbf{1}_{t\in\mathbf{I}_{l,\xi}}\\ &\lesssim |l|\left(1+\left|t-\frac{\xi}{l}\right|\right)\sqrt{\frac{\partial_t w(t,\xi)}{w(t,\xi)}},\end{aligned}$$

which proves (6.11d). □

Finally, Lemma 6.11 completes the treatment of $\mathrm{R}_{\mathrm{N}}^1$; in particular, after summing in N, (A.2), (6.10) and Lemma 6.11 yield only terms appearing on the RHS of (2.27).

6.2. *Corrections*

6.2.1. *Term* $\mathrm{R}_{\mathrm{N}}^{\epsilon,1}$*:* $\mathrm{O}(\epsilon)$ *correction to* $\mathrm{R}_{\mathrm{N}}^1$

In this section we treat $\mathrm{R}_{\mathrm{N}}^{\epsilon,1}$ which is higher order in ϵ than $\mathrm{R}_{\mathrm{N}}^1$. We expand $(1 - v')\phi_l$ with a paraproduct *only in* v:

$$
\begin{aligned}
\mathrm{R}_{\mathrm{N}}^{\epsilon,1} = & -\frac{1}{2\pi}\sum_{\mathrm{M}\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \mathrm{A}\bar{\hat{f}}_k(\eta)\mathrm{A}_k(\eta)\big((\eta-\xi)l-\xi'(k-l)\big)\rho_{\mathrm{N}}(l,\xi) \\
& \times \Big[\widehat{(1-v')}(\xi'-\xi)\Big]_{<\mathrm{M}/8}\widehat{\phi}_l(\xi')_{\mathrm{M}}\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}d\eta d\xi d\xi' \\
& -\frac{1}{2\pi}\sum_{\mathrm{M}\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \mathrm{A}\bar{\hat{f}}_k(\eta)\mathrm{A}_k(\eta)\big((\eta-\xi)l-\xi'(k-l)\big)\rho_{\mathrm{N}}(l,\xi) \\
& \times \Big[\widehat{(1-v')}(\xi'-\xi)\Big]_{\mathrm{M}}\widehat{\phi}_l(\xi')_{<\mathrm{M}/8}\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}d\eta d\xi d\xi' \\
& -\frac{1}{2\pi}\sum_{\mathrm{M}\in\mathbf{D}}\sum_{\frac{1}{8}\mathrm{M}\leq\mathrm{M}'\leq 8\mathrm{M}}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \mathrm{A}\bar{\hat{f}}_k(\eta)\mathrm{A}_k(\eta)\big((\eta-\xi)l-\xi'(k-l)\big) \\
& \times \rho_{\mathrm{N}}(l,\xi)\Big[\widehat{(1-v')}(\xi'-\xi)\Big]_{\mathrm{M}'}\widehat{\phi}_l(\xi')_{\mathrm{M}}\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}d\eta d\xi d\xi' \\
= & \ \mathrm{R}_{\mathrm{N;LH}}^{\epsilon,1}+\mathrm{R}_{\mathrm{N;HL}}^{\epsilon,1}+\mathrm{R}_{\mathrm{N;HH}}^{\epsilon,1}.
\end{aligned}
$$

We recall that ρ_{N} denotes the Littlewood-Paley cut-off to the N-th dyadic shell in $\mathbf{Z}\times\mathbf{R}$; see (A.1). The intuition is as follows: $\mathrm{R}_{\mathrm{N;LH}}^{\epsilon,1}$ can be treated in a manner very similar to $\mathrm{R}_{\mathrm{N}}^{1}$ as here $(1-v')$ appears essentially as part of the background and $\mathrm{R}_{\mathrm{N;HL}}^{\epsilon,1}$ should be manageable since $(1-v')$ is controlled by the bootstrap hypotheses and ϕ_l provides decay in time.

Begin first with $\mathrm{R}_{\mathrm{N;LH}}^{\epsilon,1}$. On the support of the integrand,

$$\big||k,\eta|-|l,\xi|\big| \leq |k-l,\eta-\xi| \leq \frac{3}{16}|l,\xi|, \tag{6.14a}$$

$$\big|\left|l,\xi'\right|-|l,\xi|\big| \leq \left|\xi'-\xi\right| \leq \frac{3}{16}\left|\xi'\right| \leq \frac{3}{16}\left|l,\xi'\right|, \tag{6.14b}$$

and hence by two applications of (A.7), there is some $c\in(0,1)$ such that

$$e^{\lambda|k,\eta|^s} \leq e^{\lambda|l,\xi'|^s+c\lambda|k-l,\eta-\xi|^s+c\lambda|\xi'-\xi|^s}.$$

Therefore, (using that $|k,\eta|\approx|l,\xi'|$ from (6.14)),

$$
\begin{aligned}
\left|\mathrm{R}_{\mathrm{N;LH}}^{\epsilon,1}\right| \lesssim & \sum_{\mathrm{M}\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'}\left|\mathrm{A}\hat{f}_k(\eta)\right|\left|(\eta-\xi)l-\xi'(k-l)\right|\rho_{\mathrm{N}}(l,\xi)\mathrm{J}_k(\eta)\langle l,\xi'\rangle^{\sigma} \\
& \times e^{\lambda|l,\xi'|^s}\left|\widehat{\phi}_l(\xi')_{\mathrm{M}}\right|e^{c\lambda|k-l,\eta-\xi|^s+c\lambda|\xi'-\xi|^s}\left|\widehat{(1-v')}(\xi'-\xi)_{<\mathrm{M}/8}\right| \\
& \times\left|\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi d\xi'.
\end{aligned}
$$

From here we may proceed analogous to the treatment of $\mathrm{R}_{\mathrm{N}}^{1}$ with (l,ξ') playing the role of (l,ξ) and using (A.5) (instead of (A.4)) together with the bootstrap hypotheses to deal with the low-frequency factors. We omit the details and simply conclude that the result is analogous to (6.10), except with an additional power of ϵ.

Turn now to $R^{\epsilon,1}_{N;HL}$. Similar to what occurs in Section 4.2.2, l could be large relative to $\xi-\xi'$ hence more 'derivatives' are appearing on ϕ than $(1-v')$ and we are again in a situation similar to $R^{\epsilon,1}_{N;LH}$. As in Section 4.2.2 we divide the integral based on the relative size of l and ξ:

$$\begin{aligned}R^{\epsilon,1}_{N;HL} &= -\frac{1}{2\pi}\sum_{M\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} A\bar{\hat{f}}_k(\eta)[\mathbf{1}_{16|l|\geq|\xi|}+\mathbf{1}_{16|l|<|\xi|}]\\ &\quad\times A_k(\eta)\big((\eta-\xi)l-\xi'(k-l)\big)\rho_N(l,\xi)\\ &\quad\times \widehat{(1-v')}(\xi'-\xi)_M\hat{\phi}_l(\xi')_{<M/8}\hat{f}_{k-l}(\eta-\xi)_{<N/8}\,d\eta\, d\xi\, d\xi'\\ &= R^{\epsilon,1;z}_{N;HL}+R^{\epsilon,1;v}_{N;HL}.\end{aligned}$$

First consider $R^{\epsilon,1;z}_{N;HL}$, where on the support of the integrand, $16|l|\geq|\xi|$.

$$\big||k,\eta|-|l,\xi|\big|\leq|k-l,\eta-\xi|\leq 3|l,\xi|/16, \tag{6.15a}$$

$$\big||l,\xi|-|l,\xi'|\big|\leq|\xi-\xi'|\leq 38|\xi|/32\lesssim|l|. \tag{6.15b}$$

If $|l|\geq 16|\xi|$, then in fact $38|\xi|/32<|l|/4$, therefore by applying twice (A.7), for some $c\in(0,1)$,

$$e^{\lambda|k,\eta|^s}\leq e^{\lambda|l,\xi|^s+c\lambda|k-l,\eta-\xi|^s}\leq e^{\lambda|l,\xi'|^s+c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}.$$

Alternatively, if $\frac{1}{16}|\xi|\leq|l|\leq 16|\xi|$ then $|\xi-\xi'|\approx|l,\xi|$ and hence (A.7) and (A.8) imply for some (different) $c\in(0,1)$ we have,

$$e^{\lambda|k,\eta|^s}\leq e^{\lambda|l,\xi|^s+c\lambda|k-l,\eta-\xi|^s}\leq e^{c\lambda|l,\xi'|^s+c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}.$$

In both cases, it follows that (using also $\langle k,\eta\rangle\approx\langle l,\xi\rangle\lesssim\langle l\rangle$ from (6.15)),

$$\begin{aligned}\left|R^{\epsilon,1;z}_{N;HL}\right| &\lesssim \sum_{M\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'}\mathbf{1}_{16|l|\geq|\xi|}\left|A\hat{f}_k(\eta)\right|\left|(\eta-\xi)l-\xi'(k-l)\right|\\ &\quad\times\rho_N(l,\xi)J_k(\eta)\langle l\rangle^{\sigma}e^{\lambda|l,\xi'|^s}\left|\hat{\phi}_l(\xi')_{<M/8}\right|e^{c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}\\ &\quad\times\left|\widehat{(1-v')}(\xi'-\xi)_M\right|\left|\hat{f}_{k-l}(\eta-\xi)_{<N/8}\right|d\eta\, d\xi\, d\xi'.\end{aligned}$$

For minor technical convenience, divide into low and high frequencies: for some $M_0\geq 8$,

$$\left|R^{\epsilon,1;z}_{N;HL}\right|\lesssim\left(\sum_{M\leq M_0}+\sum_{M\geq M_0}\right)\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'}\mathbf{1}_{16|l|\geq|\xi|}\left|A\hat{f}_k(\eta)\right|$$

$$\times \left|(\eta-\xi)l-\xi'(k-l)\right|\rho_N(l,\xi)J_k(\eta)\langle l\rangle^\sigma e^{\lambda|l,\xi'|^s}\left|\widehat{\phi}_l(\xi')_{<M/8}\right|$$
$$\times e^{c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}\left|\widehat{(1-v')}(\xi'-\xi)_M\right|$$
$$\times \left|\hat{f}_{k-l}(\eta-\xi)_{<N/8}\right|d\eta d\xi d\xi'$$
$$= R^{\epsilon,1;z}_{N;HL;L} + R^{\epsilon,1;z}_{N;HL;H}.$$

Consider next $R^{\epsilon,1;z}_{N;HL;H}$. By choosing M_0 sufficiently large (relative to our $O(1)$ arithmetic conventions), on the support of the integrand (k,η) and (l,ξ') are both non-resonant by (6.15). Therefore, by (3.31) followed by (A.11),

$$R^{\epsilon,1;z}_{N;HL;H} \lesssim \sum_{M\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \mathbf{1}_{16|l|\geq|\xi|}\left|A\hat{f}_k(\eta)\right|\left|(\eta-\xi)l-\xi'(k-l)\right|$$
$$\times \rho_N(l,\xi)J_l(\xi')\langle l\rangle^\sigma e^{\lambda|l,\xi'|^s}\mathbf{1}_{t\notin \mathbf{I}_{l,\xi'}}\left|\widehat{\phi}_l(\xi')_{<M/8}\right|e^{\lambda|\xi-\xi'|^s+\lambda|k-l,\eta-\xi|^s}$$
$$\times \left|\widehat{(1-v')}(\xi'-\xi)_M\right|\left|\hat{f}_{k-l}(\eta-\xi)_{<N/8}\right|d\eta d\xi d\xi'.$$

Since l cannot be zero and by (6.15), we have $|l,\xi|\lesssim |l|^{1-s/2}|k,\eta|^{s/2}$, which implies

$$R^{\epsilon,1;z}_{N;HL;H} \lesssim \sum_{M\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \mathbf{1}_{16|l|\geq|\xi|}|k,\eta|^{s/2}\left|A\hat{f}_k(\eta)\right|\rho_N(l,\xi)|l|^{1-s/2}J_l(\xi')\langle l\rangle^\sigma$$
$$\times e^{\lambda|l,\xi'|^s}\mathbf{1}_{t\notin \mathbf{I}_{l,\xi'}}\left|\widehat{\phi}_l(\xi')_{<M/8}\right|e^{\lambda|\xi-\xi'|^s+\lambda|k-l,\eta-\xi|^s}\left|\widehat{(1-v')}(\xi'-\xi)_M\right.$$
$$\times\left||k-l,\eta-\xi|\left|\hat{f}_{k-l}(\eta-\xi)_{<N/8}\right|d\eta d\xi d\xi'.$$

Therefore, since $|k,\eta|\approx|l,\xi|\approx|l,\xi'|$ from (6.15), by (A.5) and (A.2) (denoting $\chi^{NR}_l(t,\xi) = \mathbf{1}_{t\notin \mathbf{I}_{l,\xi}}$),

$$R^{\epsilon,1;z}_{N;HL;H} \lesssim \sum_{M\geq 8}\left\||\nabla|^{s/2}Af_{\sim N}\right\|_2\left\||\nabla|^{1-s/2}\chi^{NR}AP_{\neq 0}\phi_{\sim N}\right\|_2$$
$$\times \left\|(1-v')_M\right\|_{\mathcal{G}^{\lambda,\sigma-1}}\|f_{<N/8}\|_{\mathcal{G}^{\lambda,\sigma}}$$
$$\lesssim \sum_{M\geq 8}\left\||\nabla|^{s/2}Af_{\sim N}\right\|_2\left\||\nabla|^{1-s/2}\chi^{NR}AP_{\neq 0}\phi_{\sim N}\right\|_2$$
$$\times \frac{1}{M}\left\|(1-v')_M\right\|_{\mathcal{G}^{\lambda,\sigma}}\|f_{<N/8}\|_{\mathcal{G}^{\lambda,\sigma}}$$
$$\lesssim \left\||\nabla|^{s/2}Af_{\sim N}\right\|_2\left\||\nabla|^{1-s/2}\chi^{NR}AP_{\neq 0}\phi_{\sim N}\right\|_2\|f_{<N/8}\|_{\mathcal{G}^{\lambda,\sigma}}$$
$$\times \left(\sum_{M\geq 8}\left\|(1-v')_M\right\|^2_{\mathcal{G}^{\lambda,\sigma}}\right)^{1/2},$$

where the last line followed by Cauchy-Schwarz. By (A.1) and the bootstrap hypotheses,

$$\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HL;H}} \lesssim \epsilon^2 \left\| |\nabla|^{s/2} \mathrm{A} f_{\sim \mathrm{N}} \right\|_2 \left\| |\nabla|^{1-s/2} \chi^{\mathrm{NR}} \mathrm{AP}_{\neq 0} \phi_{\sim \mathrm{N}} \right\|_2. \tag{6.16}$$

The treatment of $\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HL;L}}$ is straightforward by a similar argument. Indeed, on the support of the integrand, if $t \in \mathbf{1}_{l,\xi'}$ then necessarily $|l| + |\xi'| \lesssim 2^{\mathrm{M}_0} \approx 1$. Hence for resonant frequencies we may simply use Lemma 4.1 to handle ϕ (as the restriction to low frequencies allows us to gain regularity on ϕ). For non-resonant contributions to the integral we use the same method as that used on $\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HL;H}}$. We omit the details and state the result

$$\sum_{\mathrm{N}\geq 8} \mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HL;L}} \lesssim \frac{\epsilon^2}{\langle t\rangle^{2s}} \left\| |\nabla|^{s/2} \mathrm{A} f \right\|_2^2 + \epsilon^2 \langle t\rangle^{2s} \left\| |\nabla|^{1-s/2} \chi^{\mathrm{NR}} \mathrm{AP}_{\neq 0} \phi \right\|_2^2 + \frac{\epsilon^4}{\langle t\rangle^2}, \tag{6.17}$$

completing the treatment of $\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HL}}$.

Next we turn to $\mathrm{R}^{\epsilon,1;v}_{\mathrm{N;HL}}$, in which case we can consider all of the ‘derivatives’ to be landing on $1 - v'$. On the support of the integrand,

$$\big| |k,\eta| - |l,\xi| \big| \leq |k-l, \eta-\xi| \leq 3|l,\xi|/16, \tag{6.18a}$$

$$\begin{aligned} \big| |\xi-\xi'| - |l,\xi| \big| &\leq |l,\xi'| \leq |\xi|/16 + |\xi'| \leq |\xi-\xi'|/16 + 17|\xi'|/16 \\ &\leq 67|\xi-\xi'|/100, \end{aligned} \tag{6.18b}$$

which implies by two applications of (A.7) there exists some $c \in (0,1)$ such that

$$e^{\lambda|k,\eta|^s} \leq e^{\lambda|l,\xi|^s + c\lambda|k-l,\eta-\xi|^s} \leq e^{\lambda|\xi-\xi'|^s + c\lambda|l,\xi'|^s + c\lambda|k-l,\eta-\xi|^s}.$$

Therefore, since also $|l,\xi| \approx |\xi-\xi'|$ by (6.18),

$$\begin{aligned} \left|\mathrm{R}^{\epsilon,1;v}_{\mathrm{N;HL}}\right| \lesssim \sum_{\mathrm{M}\geq 8} \sum_{k,l\neq 0} \int_{\eta,\xi,\xi'} & \left|\mathrm{A}\hat{f}_k(\eta)\right| \mathbf{1}_{16|l|\leq|\xi|} \mathrm{J}_k(\eta) \langle \xi-\xi' \rangle^\sigma \rho_{\mathrm{N}}(l,\xi) e^{\lambda|\xi-\xi'|^s} \\ & \times \left| \widehat{(1-v')}(\xi'-\xi)_{\mathrm{M}} \right| |k-l,\eta-\xi| \left| l,\xi' \right| e^{c\lambda|l,\xi'|^s + c\lambda|k-l,\eta-\xi|^s} \\ & \times \left| \hat{\phi}_l(\xi')_{<\mathrm{M}/8} \hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8} \right| d\eta d\xi d\xi'. \end{aligned}$$

We will now use the following analogue of (4.8), which applies on the support of the integrand due to the frequency localizations (6.18):

$$\mathrm{J}_k(\eta) \lesssim \mathrm{J}^{\mathrm{R}}(\xi-\xi') e^{20\mu|\xi'|^{1/2} + 20\mu|\eta-\xi|^{1/2}}. \tag{6.19}$$

Applying this together with (A.11) and (A.12) implies

$$\left|R^{\epsilon,1}_{N;HL}\right| \lesssim \sum_{M\geq 8}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \left|A\hat{f}_k(\eta)\right| A^R(\xi-\xi')\left|\left(\widehat{1-v'}\right)(\xi'-\xi)_M\right|$$
$$\times \rho_N(l,\xi)\mathbf{1}_{16|l|\leq|\xi|}e^{\lambda|l,\xi'|^s+\lambda|k-l,\eta-\xi|^s}$$
$$\times\left|\widehat{\phi}_l(\xi')_{<M/8}\hat{f}_{k-l}(\eta-\xi)_{<N/8}\right|d\eta d\xi d\xi'.$$

Since $|l,\xi|\approx|\xi-\xi'|$ by (6.18), the sum only includes boundedly many terms. Therefore, by (A.5), Lemma 4.1 and the bootstrap hypotheses,

$$\textbf{(6.20)}\qquad \left|R^{\epsilon,1}_{N;HL}\right| \lesssim \|Af_{\sim N}\|_2\left\|A^R\left(1-v'\right)_{\sim N}\right\|_2\|f\|_{\mathcal{G}^{\lambda,\sigma-4}}\|P_{\neq}\phi\|_{\mathcal{G}^{\lambda,\sigma-4}}$$
$$\lesssim \frac{\epsilon}{\langle t\rangle^2}\|Af_{\sim N}\|_2^2+\frac{\epsilon^3}{\langle t\rangle^2}\left\|A^R\left(1-v'\right)_{\sim N}\right\|_2^2,$$

which completes our treatment of $R^{\epsilon,1}_{N;HL}$.

We turn to the remainder term $R^{\epsilon,1}_{N;HH}$. Analogous to Section 4.2.3, there is a problem when l is large compared to ξ. The situation here is only slightly more subtle. We divide into two cases:

$$R^{\epsilon,1}_{N;HH} = -\frac{1}{2\pi}\sum_{M\in\mathbf{D}}\sum_{\frac{1}{8}M\leq M'\leq 8M}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} A\bar{\hat{f}}_k(\eta)[\mathbf{1}_{|l|>100|\xi'|}+\mathbf{1}_{|l|\leq 100|\xi'|}]$$
$$\times A_k(\eta)\big((\eta-\xi)l-\xi'(k-l)\big)\rho_N(l,\xi)\left(\widehat{1-v'}\right)(\xi'-\xi)_{M'}$$
$$\times\widehat{\phi}_l(\xi')_M\hat{f}_{k-l}(\eta-\xi)_{<N/8}d\eta d\xi d\xi'$$
$$= R^{\epsilon,1;z}_{N;HH}+R^{\epsilon,1;v}_{N;HH}.$$

First consider $R^{\epsilon,1;z}_{N;HH}$. On the support of the integrand we have

$$\textbf{(6.21a)}\qquad \big||k,\eta|-|l,\xi|\big|\leq|k-l,\eta-\xi|\leq 6|l,\xi|/32,$$

$$\textbf{(6.21b)}\qquad \big||l,\xi|-|l,\xi'|\big|\leq|\xi-\xi'|\leq 24|\xi'|\leq\frac{24}{100}|l,\xi'|.$$

Therefore by two applications of (A.7),

$$e^{\lambda|k,\eta|^s}\leq e^{\lambda|l,\xi|^s+c\lambda|k-l,\eta-\xi|^s}\leq e^{\lambda|l,\xi'|^s+c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}.$$

By $|l|>100|\xi'|$ and (6.21), it follows that $|\xi|\leq\frac{2524}{10000}|l|$ and hence (l,ξ) cannot be resonant and $|\eta|\leq 1.531|k|$, which implies by $N\geq 8$ that (k,η) cannot be resonant. Also using

$$\big|(\eta-\xi)l-\xi'(k-l)\big|\leq\big|l,\xi'\big||k-l,\eta-\xi|,$$

we have (3.31), (A.11) and (A.12) together imply

$$\left|\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HH}}\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}'\approx\mathrm{M}}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \left|\mathrm{A}\hat{f}_k(\eta)\right|\mathbf{1}_{|l|>100|\xi'|}\rho_{\mathrm{N}}(l,\xi)\mathrm{A}_l\big(\xi'\big)\left|l,\xi'\right| \times \left|\widehat{\phi}_l\big(\xi'\big)_{\mathrm{M}}\right|e^{\lambda|\xi-\xi'|^s}\left|\widehat{\big(1-v'\big)}\big(\xi'-\xi\big)_{\mathrm{M}'}\right| \times e^{\lambda|k-l,\eta-\xi|^s}\left|\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi d\xi'.$$

By (A.5), $|l,\xi'|\approx \mathrm{N}$ (by (6.21)) and (A.2),

$$\left|\mathrm{R}^{\epsilon,1;z}_{\mathrm{N;HH}}\right| \lesssim \left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|\|f_{<\mathrm{N}/8}\|_{\mathcal{G}^{\lambda,\sigma}}\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{A}\phi_{\sim\mathrm{N}}\right\|_2 \times\sum_{\mathrm{M}'\in\mathbf{D}}\left\|\big(1-v'\big)_{\mathrm{M}'}\right\|_{\mathcal{G}^{\lambda,\sigma-1}}$$
$$\lesssim \left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|\|f_{<\mathrm{N}/8}\|_{\mathcal{G}^{\lambda,\sigma}}\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{A}\phi_{\sim\mathrm{N}}\right\|_2 \times\left(\sum_{\mathrm{M}'\in\mathbf{D}}\left\|\big(1-v'\big)_{\mathrm{M}'}\right\|^2_{\mathcal{G}^{\lambda,\sigma}}\right)^{1/2}$$
$$\lesssim \epsilon^2\left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_2\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{AP}_{\neq 0}\phi_{\sim\mathrm{N}}\right\|_2, \tag{6.22}$$

where the last line followed from the bootstrap hypotheses.

Turn to $\mathrm{R}^{\epsilon,1;v}_{\mathrm{N;HH}}$. On the support of the integrand there holds

$$\big||k,\eta|-|l,\xi|\big| \le |k-l,\eta-\xi| \le 3|l,\xi|/16, \tag{6.23a}$$

$$\left|\xi-\xi'\right| \le 24\left|\xi'\right| \le 24\left|l,\xi'\right| \tag{6.23b}$$

$$\left|l,\xi'\right| \le 101\left|\xi'\right| \le 2424\left|\xi-\xi'\right|, \tag{6.23c}$$

and hence by (A.7) followed by (A.8) for some $c\in(0,1)$,

$$e^{\lambda|k,\eta|^s} \le e^{\lambda|l,\xi|^s+c\lambda|k-l,\eta-\xi|^s} \le e^{c\lambda|l,\xi'|^s+c\lambda|\xi-\xi'|^s+c\lambda|k-l,\eta-\xi|^s}.$$

Notice that here $\mathrm{N}\lesssim|\xi,l|\lesssim\mathrm{M}$. By Lemma 3.1 and (6.23),

$$\mathrm{J}_k(\eta) \lesssim e^{2\mu|k,\eta|^{1/2}} \lesssim e^{2\mu|k-l,\eta-\xi|^{1/2}+2\mu|l,\xi'|^{1/2}+2\mu|\xi'-\xi|^{1/2}}.$$

The previous two estimates together with (A.11) and (A.12) imply

$$\left|\mathrm{R}^{\epsilon,1;v}_{\mathrm{N;HH}}\right| \lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}'\approx\mathrm{M}}\sum_{k,l\neq 0}\int_{\eta,\xi,\xi'} \left|\mathrm{A}\hat{f}_k(\eta)\right|\mathbf{1}_{|l|\le 100|\xi'|}\rho_{\mathrm{N}}(l,\xi) \times e^{\lambda|\xi-\xi'|^s}\left|\widehat{\big(1-v'\big)}\big(\xi'-\xi\big)_{\mathrm{M}'}\right|e^{\lambda|l,\xi'|^s}\left|\widehat{\phi}_l\big(\xi'\big)_{\mathrm{M}}\right|e^{\lambda|k-l,\eta-\xi|^s} \times\left|\hat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi d\xi'.$$

By applying (A.5), Lemma 4.1 and the bootstrap hypotheses similar to above,

$$\begin{aligned}\left|\mathrm{R}^{\epsilon,1;v}_{\mathrm{N;HH}}\right| &\lesssim \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}'\approx\mathrm{M}} \|\mathrm{A}f_{\sim\mathrm{N}}\|_2 \|\mathrm{P}_{\neq 0}\phi_{\mathrm{M}}\|_{\mathcal{G}^{\lambda,\sigma-3}} \left\|\left(1-v'\right)_{\mathrm{M}'}\right\|_{\mathcal{G}^{\lambda,\sigma-2}} \|f_{<\mathrm{N}/8}\|_{\mathcal{G}^{\lambda,\sigma}} \\ &\lesssim \epsilon \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}'\approx\mathrm{M}} \|\mathrm{A}f_{\sim\mathrm{N}}\|_2 \frac{1}{\mathrm{M}^2} \|\mathrm{P}_{\neq 0}\phi_{\mathrm{M}}\|_{\mathcal{G}^{\lambda,\sigma-3}} \left\|\left(1-v'\right)_{\mathrm{M}'}\right\|_{\mathcal{G}^{\lambda,\sigma}} \\ &\lesssim \frac{\epsilon^3}{\mathrm{N}\langle t\rangle^2}\|\mathrm{A}f_{\sim\mathrm{N}}\|_2,\end{aligned}$$

where the last two lines followed from $\mathrm{N} \lesssim \mathrm{M} \approx \mathrm{M}'$, Cauchy-Schwarz and (A.2). Together with (6.22), (6.20), (6.16) and (6.17), this completes the treatment of $\mathrm{R}^{\epsilon,1}_{\mathrm{N}}$ after summing N and applying (A.2), hence reducing to terms appearing on the RHS of (2.27).

6.2.2. *Term $\mathrm{R}^3_{\mathrm{N}}$: remainder from commutator*

Now we treat $\mathrm{R}^3_{\mathrm{N}}$, which arose from the integration by parts intended for treating transport. By (A.4) (for $\sigma > 6$), the bootstrap hypotheses,

$$\begin{aligned}\left|\mathrm{R}^3_{\mathrm{N}}\right| \lesssim & \sum_{k,l}\int_{\eta,\xi} \left|\mathrm{A}\hat{f}_k(\eta)\right| |k-l,\eta-\xi| \left|\hat{u}_l(\xi)_{\mathrm{N}}\right| \mathrm{A}_{k-l}(\eta-\xi) \\ &\times \left|\widehat{f}_{k-l}(\eta-\xi)_{<\mathrm{N}/8}\right| d\eta d\xi \\ \lesssim & \|\mathrm{A}f_{\sim\mathrm{N}}\|_2 \|u_{\mathrm{N}}\|_{\mathrm{H}^{\sigma-4}} \|\mathrm{A}f_{<\mathrm{N}/8}\|_2 \\ \lesssim & \frac{\epsilon}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}\|\mathrm{A}f_{\sim\mathrm{N}}\|_2^2 + \epsilon\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}\|u_{\mathrm{N}}\|^2_{\mathrm{H}^{\sigma-4}},\end{aligned}$$

where we also used that $|k-l,\eta-\xi| \lesssim |l,\xi|$ on the support of the integrand. By (A.2), Lemma 4.1 and the bootstrap hypotheses,

$$\sum_{\mathrm{N}\geq 8}\left|\mathrm{R}^3_{\mathrm{N}}\right| \lesssim \frac{\epsilon^3}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}.$$

This completes the treatment of $\mathrm{R}^3_{\mathrm{N}}$, as this appears on the RHS of (2.27).

6.3. *Zero mode reaction term*

Next we turn to $\mathrm{R}^2_{\mathrm{N}}$, which is the part of the reaction term involving $[\partial_t v]$. Here we need to make sure that assigning slightly less regularity to $[\partial_t v]$ than f is consistent with $\mathrm{R}^2_{\mathrm{N}}$. Also note that $[\partial_t v]$ has non-resonant regularity, but forces resonant frequencies here, expressed in the loss that Lemma 3.6 could incur. Write $\mathrm{R}^2_{\mathrm{N}}$ on the frequency-side and divide into the two natural cases

$$\begin{aligned}\mathrm{R}_{\mathrm{N}}^{2} &= -\sum_{k}\int_{\eta,\xi} \mathrm{A}\bar{\hat{f}}_{k}(\eta)\left[\chi^{\mathrm{D}}+\chi^{*}\right]\mathrm{A}_{k}(\eta)\widehat{[\partial_{t}v]}(\xi)_{\mathrm{N}}i(\eta-\xi) \\ &\quad \times \hat{f}_{k}(\eta-\xi)_{<\mathrm{N}/8}d\eta d\xi \\ &= \mathrm{R}_{\mathrm{N}}^{2;\mathrm{D}}+\mathrm{R}_{\mathrm{N}}^{2;*},\end{aligned}$$

where $\chi^{\mathrm{D}} = \mathbf{1}_{t\in\mathbf{I}_{k,\eta}}\mathbf{1}_{t\in\mathbf{I}_{k,\xi}}$ and $\chi^{*} = 1-\chi^{\mathrm{D}}$. Next note that on the support of the integrand we have

$$|k|+|\eta-\xi| \leq 3\mathrm{N}/32 \leq 3|\xi|/16 \tag{6.24}$$

which implies $|k| \lesssim |\eta| \approx |\xi|$. Also note $|\xi| \gtrsim \mathrm{N}$.

First treat the term $\mathrm{R}_{\mathrm{N}}^{2;*}$. By (A.7), (3.31), (A.11) and (A.12),

$$\left|\mathrm{R}_{\mathrm{N}}^{2;*}\right| \lesssim \sum_{k}\int_{\eta,\xi} \chi^{*}\left|\mathrm{A}\hat{f}_{k}(\eta)\right|\mathrm{A}_{0}(\xi)\left|\widehat{[\partial_{t}v]}(\xi)_{\mathrm{N}}\right|e^{\lambda|k,\eta-\xi|^{s}}\left|\hat{f}_{k}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.$$

To deal with the norm imbalance between $[\partial_{t}v]$ and f, by $|\xi| \gtrsim 1$ and $|\eta| \approx |\xi|$,

$$\begin{aligned}\left|\mathrm{R}_{\mathrm{N}}^{2;*}\right| &\lesssim \sum_{k}\int_{\eta,\xi} \chi^{*}\left|\mathrm{A}\hat{f}_{k}(\eta)\right|\frac{|\eta|^{s/2}|\xi|^{s/2}}{\langle\xi\rangle^{s}}\mathrm{A}_{0}(\xi)\left|\widehat{[\partial_{t}v]}(\xi)_{\mathrm{N}}\right| \\ &\quad \times e^{\lambda|k,\eta-\xi|^{s}}\left|\hat{f}_{k}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.\end{aligned}$$

It follows from (A.4) and the bootstrap hypotheses that

$$\begin{aligned}\left|\mathrm{R}_{\mathrm{N}}^{2;*}\right| &\lesssim \epsilon\left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_{2}\left\||\partial_{v}|^{s/2}\frac{\mathrm{A}}{\langle\partial_{v}\rangle^{s}}[\partial_{t}v]_{\mathrm{N}}\right\|_{2} \\ &\lesssim \frac{\epsilon}{\langle t\rangle^{2s}}\left\||\nabla|^{s/2}\mathrm{A}f_{\sim\mathrm{N}}\right\|_{2}^{2}+\epsilon\langle t\rangle^{2s}\left\||\partial_{v}|^{s/2}\frac{\mathrm{A}}{\langle\partial_{v}\rangle^{s}}[\partial_{t}v]_{\mathrm{N}}\right\|_{2}^{2}.\end{aligned} \tag{6.25}$$

The former is absorbed by CK_{λ} and the latter is controlled by $\mathrm{CK}_{\lambda}^{v,1}$.

Next turn to $\mathrm{R}_{\mathrm{N}}^{2;\mathrm{D}}$. Applying (A.7) and (6.24), there exists a $c \in (0,1)$ such that,

$$\begin{aligned}\left|\mathrm{R}_{\mathrm{N}}^{2;\mathrm{D}}\right| &\lesssim \sum_{k}\int_{\eta,\xi} \chi^{\mathrm{D}}\left|\mathrm{A}\hat{f}_{k}(\eta)\right|\mathrm{A}_{0}(\xi)\frac{\mathrm{J}_{k}(\eta)}{\mathrm{J}_{0}(\xi)}\left|\widehat{[\partial_{t}v]}(\xi)_{\mathrm{N}}\right|e^{c\lambda|k,\eta-\xi|^{s}}|\eta-\xi| \\ &\quad \times \left|\hat{f}_{k}(\eta-\xi)_{<\mathrm{N}/8}\right|d\eta d\xi.\end{aligned}$$

Since $t \in \mathbf{I}_{k,\eta}$, $\mathrm{A}_{k}(\eta) \lesssim \tilde{\mathrm{A}}_{k}(\eta)$. Moreover, by (3.34), (A.11) and (A.12),

$$\left|R_N^{2;D}\right| \lesssim \sum_k \int_{\eta,\xi} \chi^D \left|\tilde{A}\hat{f}_k(\eta)\right| A_0(\xi) \left|\widehat{[\partial_t v]}(\xi)_N\right| \frac{|\eta|}{k^2} \sqrt{\frac{\partial_t w_k(\eta)}{w_k(\eta)}} \sqrt{\frac{\partial_t w_0(\xi)}{w_0(\xi)}} \times e^{\lambda|k,\eta-\xi|^s} \left|\hat{f}_k(\eta-\xi)_{<N/8}\right| d\eta d\xi.$$

Since $t \approx \frac{\eta}{k}$, by (6.24), (A.4) and the bootstrap hypotheses,

$$\begin{aligned}
(\mathbf{6.26}) \qquad \left|R_N^{2;D}\right| &\lesssim \sum_k \int_{\eta,\xi} \chi^D \left|\tilde{A}\hat{f}_k(\eta)\right| A_0(\xi) \left|\widehat{[\partial_t v]}(\xi)_N\right| \frac{t^{1+s}}{|k|^{1-s}\langle\xi\rangle^s} \sqrt{\frac{\partial_t w_k(\eta)}{w_k(\eta)}} \\
&\quad \times \sqrt{\frac{\partial_t w_0(\xi)}{w_0(\xi)}} e^{\lambda|k,\eta-\xi|^s} \left|\hat{f}_k(\eta-\xi)_{<N/8}\right| d\eta d\xi \\
&\lesssim \epsilon t^{1+s} \left\| \sqrt{\frac{\partial_t w}{w}} \tilde{A} f_{\sim N} \right\|_2 \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle\partial_v\rangle^s} [\partial_t v]_N \right\|_2 \\
&\lesssim \epsilon \left\| \sqrt{\frac{\partial_t w}{w}} \tilde{A} f_{\sim N} \right\|_2^2 + \epsilon t^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle\partial_v\rangle^s} [\partial_t v]_N \right\|_2^2.
\end{aligned}$$

The first term is absorbed by CK_w and the latter is controlled by $CK_w^{v,1}$. This completes the treatment of R_N^2. Combining the results of (6.25), (6.26) with Sections 6.2.1 and 6.2.2 and applying (A.2) completes the treatment of R_N, proving Proposition 2.3.

7. Remainder

In this section we prove Proposition 2.6. The commutator cannot gain us anything so we may as well treat each term separately,

$$\begin{aligned}
\mathcal{R} &= 2\pi \sum_{N\in\mathbf{D}} \sum_{\frac{1}{8}N\le N'\le 8N} \int Af\left[A(u_N \nabla f_{N'})\right] dx dv \\
&\quad - 2\pi \sum_{N\in\mathbf{D}} \sum_{\frac{1}{8}N\le N'\le 8N} \int Af u_N \nabla A f_{N'} dx dv \\
&= \mathcal{R}_a + \mathcal{R}_b.
\end{aligned}$$

Consider first $\mathcal{R}_a$, written on the Fourier side:

$$\mathcal{R}_a = \sum_{N\in\mathbf{D}} \sum_{N'\approx N} \sum_{k,l} \int_{\eta,\xi} A\bar{\hat{f}}_k(\eta) A_k(\eta) \hat{u}_l(\xi)_N \cdot \widehat{\nabla f}_{k-l}(\eta-\xi)_{N'} d\xi d\eta.$$

On the support of the integrand, $|l,\xi| \approx |k-l,\eta-\xi|$ and hence by (A.8) for some $c \in (0,1)$,

$$|k,\eta|^s \leq c|k-l,\eta-\xi|^s + c|l,\xi|^s.$$

Hence,

$$|\mathcal{R}_a| \lesssim \sum_{N\in\mathbf{D}} \sum_{N'\approx N} \sum_{k,l} \int_{\eta,\xi} \left|A\hat{f}_k(\eta)\right| J_k(\eta) \langle l,\xi\rangle^{\sigma/2+1} e^{c\lambda|l,\xi|^s} \left|\hat{u}_l(\xi)_N\right| \\ \times \langle k-l,\eta-\xi\rangle^{\sigma/2-1} e^{c\lambda|k-l,\eta-\xi|^s} \left|\widehat{\nabla f}_{k-l}(\eta-\xi)_{N'}\right| d\xi\, d\eta.$$

By Lemma 3.1, (A.11) and (A.12) (since $c<1$ and $s>1/2$),

$$|\mathcal{R}_a| \lesssim \sum_{N\in\mathbf{D}} \sum_{N'\approx N} \sum_{k,l} \int_{\eta,\xi} \left|A\hat{f}_k(\eta)\right| e^{\lambda|l,\xi|^s} \left|\hat{u}_l(\xi)_N\right| \\ \times e^{\lambda|k-l,\eta-\xi|^s} \left|\widehat{\nabla f}_{k-l}(\eta-\xi)_{N'}\right| d\xi\, d\eta.$$

Hence by (A.4), Lemma 4.1 and the bootstrap hypotheses,

$$|\mathcal{R}_a| \lesssim \sum_{N\in\mathbf{D}} \sum_{N'\approx N} \|Af\|_2 \|u_N\|_{\mathcal{G}^{\lambda,\sigma-4}} \|f_{N'}\|_{\mathcal{G}^{\lambda,\sigma-1}} \\ \lesssim \sum_{N'\in\mathbf{D}} \frac{\epsilon}{t^{2-K_D\epsilon/2}N'} \|Af\|_2 \|f_{N'}\|_{\mathcal{G}^{\lambda,\sigma}} \lesssim \frac{\epsilon^3}{t^{2-K_D\epsilon/2}}.$$

This completes the treatment of $\mathcal{R}_a$. The term $\mathcal{R}_b$ is similar and hence omitted. This completes the proof of Proposition 2.6.

8. Coordinate system controls

In this section we detail the controls on (2.13) and prove Proposition 2.5.

8.1. *Derivation of (2.13)*

As in Section 2.2, denote $v'(t,v(t,y)) = \partial_y v(t,y)$, $v''(t,v(t,y)) = \partial_{yy} v(t,y)$ and $[\partial_t v](t,v(t,y)) = \partial_t v(t,y)$. Since by (2.9b), $v(t,y)$ satisfies

$$\frac{d}{dt}\big(t\big(v_y(t,y)-1\big)\big) = -\omega_0(t,y),$$

we can directly derive (2.13a) via the chain rule. Similarly, we may also derive (2.13c) via the chain rule using the definitions of v' and v''.

Deriving (2.13b) requires more work. Notice that

$$\frac{d}{dt}\big(t\big(v(t,y)-y\big)\big)=U_0^x(t,y), \tag{8.1}$$

and denote $C(t,v(t,y))=v(t,y)-y$. From (8.1) we get (recalling the definitions $[\partial_t v](t,v(t,y))=\partial_t v(t,y)$ and $\tilde{u}_0(t,v(t,y))=U_0^x(t,y)$):

$$\begin{aligned}\partial_t v(t,y)&=\frac{1}{t}U_0^x(t,y)-\frac{1}{t}\big(v(t,y)-y\big)\\ [\partial_t v](t,v)&=\frac{1}{t}\tilde{u}_0(t,v)-\frac{1}{t}C(t,v).\end{aligned} \tag{8.2}$$

Via the chain rule,

$$\partial_t C\big(t,v(t,y)\big)=\partial_t v(t,y)-\partial_v C\big(t,v(t,y)\big)\partial_t v(t,y). \tag{8.3}$$

Differentiating (8.2) in time and using (8.3), (8.2) and (2.12) eliminates C entirely and derives (2.13b).

Moreover,

$$v'\big(t,v(t,y)\big)-1=\partial_y v(t,y)-1=\partial_v C\big(t,v(t,y)\big)\partial_y v(t,y),$$

which in particular implies

$$\partial_v C(t,v)=\frac{v'(t,v)-1}{v'(t,v)}. \tag{8.4}$$

Finally, notice that (8.2) together with (8.4) and (2.14) implies

$$v'\partial_v[\partial_t v](t,v)=-\frac{1}{t}f_0(t,v)-\frac{1}{t}\big(v'-1\big)(t,v). \tag{8.5}$$

Remark **12.** — From (2.13a), (2.13c) and the bootstrap hypotheses, one can show with relative ease that

$$\big\|A\big(1-v'\big)(t)\big\|_2+\big\|A\big(\big(v'\big)^2-1\big)(t)\big\|_2+\big\|\langle\partial_v\rangle^{-1}Av''(t)\big\|_2\lesssim\epsilon^2. \tag{8.6}$$

The estimates on $(1-v')$ follow from energy estimates on (2.13a) performed with the unknown $t(v'(t,v)-1)$ in a manner that is very similar to, but easier than, techniques applied in Sections 5, 6.3 and 7. Similar control on $(v')^2-1$ and v'' then follows from (3.40) (recalling (2.13c)).

The estimates (8.6) suffice for most purposes, however they do not obtain control on the CCK terms in (2.29). This is because the estimates just described are insensitive to whether the background shear flow is converging and so cannot ensure that the CCK integrals are convergent. Hence, in order to do better, in Section 8.2 below we make estimates that imply the convergence of the shear flow.

8.2. *Proof of Proposition 2.5*

The purpose of this section is to prove Proposition 2.5, announced in Section 2.3. In order to get good control on (2.29), we will rearrange (2.13a) in the following manner using also (8.5) (see also Remark 12). For notational convenience, denote $h(t, v) = v'(t, v) - 1$ and write

$$\partial_t h + [\partial_t v]\partial_v h = \frac{1}{t}(-f_0 - h) = v'\partial_v[\partial_t v]. \tag{8.7}$$

For notational convenience we will denote

$$\bar{h}(t, v) = v'\partial_v[\partial_t v] = \frac{1}{t}(-f_0 - h). \tag{8.8}$$

The decay of $\bar{h}$ quantifies how rapidly h is converging to $-f_0$; one can also see this is as a measure of how rapidly the x-averaged vorticity is converging. The overline does not refer to complex conjugation but rather to emphasize that $\bar{h}$ is a measure of how close h is to $-f_0$. From (8.7) and (2.11) we derive,

$$\partial_t \bar{h} = -\frac{\bar{h}}{t} - \frac{1}{t}(\partial_t f_0 + \partial_t h) = -\frac{2}{t}\bar{h} - [\partial_t v]\partial_v \bar{h} + \frac{1}{t}\left\langle v'\nabla^\perp P_{\neq 0}\phi \cdot \nabla f\right\rangle. \tag{8.9}$$

The crucial step of the proof of Proposition 2.5 is the decay estimate on $\bar{h}$ given in (2.30b). The primary challenge to proving (2.30b) is controlling the last term in (8.9), which is the transfer of information to the zero modes by nonlinear interactions of non-zero modes (see (8.10) below). Since it is most crucial, it is natural to begin there.

Proof of (2.30b)*:* From (8.9) we have

$$\begin{aligned}\frac{d}{dt}\left(\langle t\rangle^{2+2s}\left\|\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2\right) &= -(2-2s)t\langle t\rangle^{2s}\left\|\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2 - CK_\lambda^{v,2} - CK_w^{v,2} \\ &\quad - 2\langle t\rangle^{2+2s}\int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left([\partial_t v]\partial_v\bar{h}\right)dv \\ &\quad + 2t^{-1}\langle t\rangle^{2+2s}\int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left\langle v'\nabla^\perp P_{\neq 0}\phi\cdot\nabla f\right\rangle dv \\ &= -CK_L^{v,2} - CK_\lambda^{v,2} - CK_w^{v,2} + \mathcal{T}^h + F.\end{aligned} \tag{8.10}$$

The term $\mathcal{T}^h$ is the same nonlinearity that occurs in $[\partial_t v]\partial_v f$ we can treat $\mathcal{T}^h$ in a manner similar to (2.23) but with u replaced just with $[\partial_t v]$, f replaced by $\bar{h}$, A replaced with $\langle\partial_v\rangle^{-s}$A and an additional t^{2+2s} out front (balanced by the decay of $\bar{h}$). We omit the details and conclude by the methods of Sections 5, 6.3 and 7 (except neither the zero mode reaction nor the transport have 'D' contributions) and the bootstrap hypotheses that

$$\left|\mathcal{T}^h\right| \lesssim \epsilon CK_\lambda^{v,2} + \epsilon CK_\lambda^{v,1} + \epsilon\langle t\rangle^{2s+K_D\epsilon/2}\left\|\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2. \tag{8.11}$$

We apply the bootstrap control on $CK_{\lambda}^{v,1}$ and absorb the rest by $CK_{\lambda}^{v,2}$ and $CK_{L}^{v,2}$ in (8.10).

The main challenge is in the F term (for ‘forcing’). This term describes the nonlinear feedback of the non-zero frequencies onto the zero frequencies, which could lead to potential instability by arresting the convergence of the background shear flow (which we are ruling out). First divide into leading order and higher order contributions:

$$\begin{aligned} F &= \sum_{k\neq 0} 2t^{-1}\langle t\rangle^{2+2s} \int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left(\nabla^\perp\phi_k\cdot\nabla f_{-k}\right)dv \\ &\quad + \sum_{k\neq 0} 2t^{-1}\langle t\rangle^{2+2s} \int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left(h\nabla^\perp\phi_k\cdot\nabla f_{-k}\right)dv \\ &= F^0 + F^\epsilon. \end{aligned}$$

As suggested by Section 6.2.1, F^ϵ is not significantly harder than F^0, in fact the primary complications that arise in the treatment of $R_N^{\epsilon,1}$ do not arise in the treatment of F^ϵ. Hence we focus only on F^0; the control of F^ϵ results in, at worst, similar contributions with an additional power of ϵ. We begin the treatment of F^0 with a paraproduct in v only:

$$\begin{aligned} F^0 &= 2\sum_{M\geq 8}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left(\left(\nabla^\perp\phi_k\right)_{<M/8}\cdot(\nabla f_{-k})_M\right)dv \\ &\quad + 2\sum_{M\geq 8}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left(\left(\nabla^\perp\phi_k\right)_{M}\cdot(\nabla f_{-k})_{<M/8}\right)dv \\ &\quad + 2\sum_{M\in\mathbf{D}}\sum_{M'\approx M}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int \frac{A}{\langle\partial_v\rangle^s}\bar{h}\frac{A}{\langle\partial_v\rangle^s}\left(\left(\nabla^\perp\phi_k\right)_{M'}\cdot(\nabla f_{-k})_{M}\right)dv \\ &= F^0_{LH} + F^0_{HL} + F^0_{\mathcal{R}}. \end{aligned}$$

The term F^0_{HL} looks something like the reaction term as studied in Section 6; dealing with it in a way that allows us to deduce (2.30b) requires the use of the regularity gap between $\bar{h}$ and ϕ. Consider a single dyadic shell and subdivide based on whether ϕ has resonant frequency or not:

$$\begin{aligned} F^0_{HL;M} &= \frac{1}{\pi}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int_{\eta,\xi}\left(\chi^R+\chi^{NR}\right)\frac{A_0(\eta)}{\langle\eta\rangle^s}\bar{\hat{h}}(\eta)\frac{A_0(\eta)}{\langle\eta\rangle^s} \\ &\quad \times\left(\widehat{\nabla^\perp\phi}_k(\xi)_M\cdot\widehat{\nabla f}_{-k}(\eta-\xi)_{<M/8}\right)d\eta d\xi \\ &= F^{0;R}_{HL;M} + F^{0;NR}_{HL;M}, \end{aligned}$$

where $\chi^R = \mathbf{1}_{t\in\mathbf{I}_{k,\xi}}$ and $\chi^{NR} = 1-\chi^R$.

Consider first the NR contribution. Subdivide further based on the relationship between time and frequency:

(**8.12**)
$$\begin{aligned} \mathrm{F}^{0;\mathrm{NR}}_{\mathrm{HL;M}} &= \frac{1}{\pi}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int_{10|\eta|\geq t}\chi^{\mathrm{NR}}\frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s}\bar{\hat{h}}(\eta)\frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s} \\ &\quad\times\left(\widehat{\nabla^{\perp}\phi}_k(\xi)_{\mathrm{M}}\cdot\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right)d\eta d\xi \\ &\quad+\frac{1}{\pi}\sum_{k\neq 0} t^{-1}\langle t\rangle^{2+2s}\int_{10|\eta|<t}\chi^{\mathrm{NR}}\frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s}\bar{\hat{h}}(\eta)\frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s} \\ &\quad\times\left(\widehat{\nabla^{\perp}\phi}_k(\xi)_{\mathrm{M}}\cdot\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right)d\eta d\xi \\ &= \mathrm{F}^{0;\mathrm{NR,S}}_{\mathrm{HL;M}}+\mathrm{F}^{0;\mathrm{NR,L}}_{\mathrm{HL;M}}, \end{aligned}$$

where the ‘S’ is for ‘short time’ and the ‘L’ is for ‘long time’ (relative to frequency). Consider first the ‘S’ contribution, for which we take advantage of the $\langle\partial_v\rangle^{-s}$ to reduce the power of t (also using $|\eta|\approx|\xi|\gtrsim 1$),

$$\begin{aligned} \mathrm{F}^{0;\mathrm{NR,S}}_{\mathrm{HL;M}} &\lesssim \sum_{k\neq 0} t^{1+s}\int_{10|\eta|\geq t}\chi^{\mathrm{NR}} \\ &\quad\times\left|\frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s}\hat{h}(\eta)\mathrm{A}_0(\eta)\left(\widehat{\nabla^{\perp}\phi}_k(\xi)_{\mathrm{M}}\cdot\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right)\right|d\eta d\xi \\ &\lesssim \sum_{k\neq 0} t^{1+s}\int_{10|\eta|\geq t}\chi^{\mathrm{NR}}\left|\frac{|\eta|^{s/2}\mathrm{A}_0(\eta)}{\langle\eta\rangle^s}\hat{h}(\eta)\mathrm{A}_0(\eta)|k,\xi|^{1-s/2}\hat{\phi}_k(\xi)_{\mathrm{M}}|k|^{s/2}\right. \\ &\quad\times\left.\left|\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right|\right|d\eta d\xi. \end{aligned}$$

Hence, by (3.31), (A.7) (using the frequency localizations as usual), (A.11), (A.12), (A.4) and the bootstrap hypotheses (along with (A.1)),

(**8.13**)
$$\begin{aligned} \sum_{\mathrm{M}\geq 8}\mathrm{F}^{0;\mathrm{NR,S}}_{\mathrm{HL;M}} &\lesssim \sum_{\mathrm{M}\geq 8}\epsilon t^{1+s}\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}\bar{h}_{\sim\mathrm{M}}\right\|_2\sum_{k\neq 0}\langle k\rangle^{-2}\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}(\mathrm{A}\phi_k)_{\mathrm{M}}\right\|_2 \\ &\lesssim \epsilon t^2\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2+\epsilon\langle t\rangle^{2s}\left\||\nabla|^{1-s/2}\chi^{\mathrm{NR}}\mathrm{A}\phi\right\|_2^2. \end{aligned}$$

The first term can be absorbed by $\mathrm{CK}^{v,2}_{\lambda}$ whereas the latter requires (6.11) and then Proposition 2.4 with the bootstrap controls on the CCK integrals.

Turn next to the ‘long’ contribution in (8.12). By (3.31), (A.7) (using also that $||\eta|-|\xi||\leq|\eta-\xi|<3|\eta|/16$ on the support of the integrand) and (A.11) we have for some

$c \in (0,1)$,

$$F^{0;NR,L}_{HL;M} \lesssim \sum_{k\neq 0} t^{1+2s} \int_{10|\eta|<t} \chi^{NR} \left| \frac{A_0(\eta)}{\langle \eta \rangle^s} \widehat{\bar{h}}(\eta) \right| |k,\xi|^{1-s} \left| A\phi_k(\xi)_M \right| e^{c\lambda |k,\eta-\xi|^s} \times |k|^s \left| \widehat{\nabla f}_k(\eta-\xi)_{<M/8} \right| d\eta d\xi.$$

Then we use the assumption $s \geq 1/2$ in order to deduce $|k,\xi|^{1-s} \leq |k,\xi|^s$ together with $|\xi - kt| \gtrsim t$, which holds on the support of the integrand since $10|\eta| < t$ and $1 \lesssim 13|\eta|/16 \leq |\xi| \leq 19|\eta|/16$. Hence,

$$F^{0;NR,L}_{HL;M} \lesssim \sum_{k\neq 0} t^{2s-1} \int_{10|\eta|<t} \chi^{NR} |\eta|^{s/2} \left| \frac{A_0(\eta)}{\langle \eta \rangle^s} \widehat{\bar{h}}(\eta) \right| |k,\xi|^{s/2} \left(k^2 + |\xi - kt|^2\right) \times \left| A\phi_k(\xi)_M \right| e^{c\lambda |k,\eta-\xi|^s} |k|^2 \left| \widehat{\nabla f}_k(\eta-\xi)_{<M/8} \right| d\eta d\xi.$$

Therefore (also multiplying by $1 \approx \langle \frac{\xi}{kt} \rangle^{-1}$),

$$F^{0;NR,L}_{HL;M} \lesssim \sum_{k\neq 0} t^{3s-1} \int_{10|\eta|<t} \chi^{NR} |\eta|^{s/2} \left| \frac{A_0(\eta)}{\langle \eta \rangle^s} \widehat{\bar{h}}(\eta) \right| \left\langle \frac{\xi}{tk} \right\rangle^{-1} \left(k^2 + |\xi - kt|^2\right) \times \frac{|k,\xi|^{s/2}}{\langle t \rangle^s} \left| A\phi_k(\xi)_M \right| e^{c\lambda |k,\eta-\xi|^s} \langle k \rangle^2 \left| \widehat{\nabla f}_{-k}(\eta-\xi)_{<M/8} \right| d\eta d\xi.$$

Summing in k and M (using (A.1)) and applying (A.4) and (A.12) we have,

$$\sum_{M\geq 8} F^{0;NR,L}_{HL;M} \lesssim \epsilon \sum_{M \geq 8} t^{3s-1} \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \bar{h}_{\sim M} \right\|_2 \times \sum_{k \neq 0} k^{-2} \left\| \left\langle \frac{\partial_v}{t\partial_z} \right\rangle^{-1} \Delta_L \frac{|\nabla|^{s/2}}{\langle t \rangle^s} (A\phi_k)_M \right\|_2$$
$$\lesssim \epsilon t^{6s-2} \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \epsilon \left\| \left\langle \frac{\partial_v}{t\partial_z} \right\rangle^{-1} \Delta_L \frac{|\nabla|^{s/2}}{\langle t \rangle^s} A\phi \right\|_2^2. \quad (\mathbf{8.14})$$

One can verify that we may always choose $\tilde{q} > 1/2$ such that $6s - 2 < 2 + 2s - 2\tilde{q}$ provided that we take the restriction that $s < 3/4$; this is an artifact due to the fact that we took s in Proposition 2.4 rather than $\tilde{q}$. However, as discussed in Section 2.3, we can without loss of generality take s close to $1/2$. Therefore, the first term is absorbed by $CK^{v,2}_\lambda$ and the second term is controlled by Proposition 2.4 and the bootstrap controls on the CCK terms, completing the treatment of $F^{0;NR,L}_{HL}$.

Next consider $F^{0;R}_{HL;M}$ where, similar to the $R^{NR,R}_N$ term in Section 6, we have to use the gain from passing A onto ϕ. By (6.6), (3.32), (A.7), (A.12) and (A.11) we deduce for some $c \in (0,1)$,

$$\left|\mathrm{F}^{0;\mathrm{R}}_{\mathrm{HL;M}}\right| \lesssim \sum_{k\neq 0} t^{1+2s} \int \chi^{\mathrm{R}} \left| \left(\sqrt{\frac{\partial_t w_0(\eta)}{w_0(\eta)}} + \frac{|\eta|^{s/2}}{\langle t\rangle^s} \right) \frac{\mathrm{A}_0(\eta)}{\langle \eta\rangle^s} \widehat{\bar{h}}(\eta) \right| \times \sqrt{\frac{w_k(\xi)}{\partial_t w_k(\xi)}} \frac{|k,\xi|}{\langle\eta\rangle^s} \frac{w_{\mathrm{R}}(\xi)}{w_{\mathrm{NR}}(\xi)} \left|\mathrm{A}\widehat{\phi}_k(\xi)_{\mathrm{M}}\right| \left|\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right| \times e^{c\lambda|k,\eta-\xi|^s} d\eta d\xi$$

$$\lesssim \sum_{k\neq 0} t^{1+s} \int \chi^{\mathrm{R}} \left| \left(\sqrt{\frac{\partial_t w_0(\eta)}{w_0(\eta)}} + \frac{|\eta|^{s/2}}{\langle t\rangle^s} \right) \frac{\mathrm{A}_0(\eta)}{\langle \eta\rangle^s} \widehat{\bar{h}}(\eta) \right| \times \sqrt{\frac{w_k(\xi)}{\partial_t w_k(\xi)}} |k,\xi| \frac{w_{\mathrm{R}}(\xi)}{w_{\mathrm{NR}}(\xi)} \left|\mathrm{A}\widehat{\phi}_k(\xi)_{\mathrm{M}}\right| \left|\widehat{\nabla f}_{-k}(\eta-\xi)_{<\mathrm{M}/8}\right| \times e^{c\lambda|k,\eta-\xi|^s} d\eta d\xi,$$

where the last line followed since $\eta \approx kt$. Hence, (A.4) and the bootstrap hypotheses imply (also using (A.12) and the low frequency factor to sum in k as well as $k^2 \lesssim |\eta|$ to replace A with $\tilde{\mathrm{A}}$):

$$\textbf{(8.15)} \qquad \sum_{\mathrm{M}\geq 8} \left|\mathrm{F}^{0;\mathrm{R}}_{\mathrm{HL;M}}\right| \lesssim \epsilon t^2 \left\| |\partial_v|^{s/2} \frac{\mathrm{A}}{\langle\partial_v\rangle^s} \bar{h} \right\|_2^2 + \epsilon t^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{\mathrm{A}}{\langle\partial_v\rangle^s} \bar{h} \right\|_2^2 + \epsilon \left\| \sqrt{\frac{w}{\partial_t w}} |\nabla| \frac{w_{\mathrm{R}}}{w_{\mathrm{NR}}} \chi^{\mathrm{R}} \tilde{\mathrm{A}}\phi \right\|_2^2.$$

The first two terms are absorbed by the $\mathrm{CK}^{v,2}$ terms and hence this suffices for $\mathrm{F}^{0;\mathrm{R}}_{\mathrm{HL}}$ after (6.11), Proposition 2.4 and the bootstrap controls on the CCK terms.

Next, we deal with $\mathrm{F}^0_{\mathrm{LH}}$:

$$\mathrm{F}^0_{\mathrm{LH;M}} \lesssim \sum_{k\neq 0} t^{1+2s} \int \left| \frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s} \hat{\bar{h}}(\eta) \right| \frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s} \left| \widehat{\nabla^\perp \phi}_{-k}(\eta-\xi)_{<\mathrm{M}/8} \right| \times \left| \widehat{\nabla f}_k(\xi)_{\mathrm{M}} \right| d\eta d\xi.$$

Here there is a serious loss of regularity from the ∇f factor, and we will make fundamental use of $s \geq 1/2$. Indeed, by the frequency localization $|\eta - \xi| \leq 3|\eta|/16$, (3.31) and (A.11) we get the following, by absorbing s derivatives using the regularity gap and $\mathrm{M} \geq 8$,

$$\left|\mathrm{F}^0_{\mathrm{LH;M}}\right| \lesssim \sum_{k\neq 0} t^{1+2s} \int \left| \frac{\mathrm{A}_0(\eta)}{\langle\eta\rangle^s} \hat{\bar{h}}(\eta) \right| |\xi|^{1-s} e^{\lambda|k,\eta-\xi|^s} \langle k\rangle \left| \widehat{\nabla^\perp \phi}_{-k}(\eta-\xi)_{<\mathrm{M}/8} \right|$$

$$\times \left|A\widehat{f_k}(\xi)_M\right| d\eta d\xi.$$

Then crucially we use $1-s \leq s$ and $|\eta| \approx |\xi|$ to deduce

$$\left|F^0_{LH;M}\right| \lesssim \sum_{k\neq 0} t^{1+2s} \int |\eta|^{s/2} \left|\frac{A_0(\eta)}{\langle\eta\rangle^s}\hat{\bar{h}}(\eta)\right| e^{\lambda|k,\eta-\xi|^s}\langle k\rangle \left|\widehat{\nabla^\perp\phi}_{-k}(\eta-\xi)_{<M/8}\right|$$
$$\times \left||\xi|^{s/2}A\widehat{f_k}(\xi)_M\right| d\eta d\xi.$$

Therefore, we apply (A.4) and Lemma 4.1 to deduce (also gaining additional powers in k to sum):

$$\left|F^0_{LH;M}\right| \lesssim \epsilon \sum_{k\neq 0} k^{-2} t^{2s-1} \left\| |\partial_v|^{s/2}\frac{A}{\langle\partial_v\rangle^s}\bar{h}_{\sim M}\right\|_2 \left\| |\partial_v|^{s/2}(Af_k)_M\right\|_2.$$

Therefore, summing in k and M,

$$\sum_{M\geq 8}\left|F^0_{LH;M}\right| \lesssim \epsilon t^{2+2s-2\tilde{q}} \left\| |\partial_v|^{s/2}\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2 + \epsilon t^{2(s+\tilde{q})-4}\left\| |\partial_v|^{s/2}Af\right\|_2^2. \tag{8.16}$$

We have $2(s+\tilde{q})-4 < -2\tilde{q}$ if we choose $1/2 < \tilde{q} < 1-s/2$, which can always be done for s close to $1/2$. Therefore for ϵ sufficiently small, the first term (8.16) is absorbed by $CK^{v,2}_\lambda$ and the latter term is controlled by ϵCK_λ.

The remainder term $F^0_{\mathcal{R}}$ is easy to handle, as in Section 7. We omit the treatment and conclude from the bootstrap hypotheses

$$\left|F^0_{\mathcal{R}}\right| \lesssim \langle t\rangle^{2s-1}\epsilon^2 \left\|\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2 \lesssim \epsilon t^{1+2s}\left\|\frac{A}{\langle\partial_v\rangle^s}\bar{h}\right\|_2^2 + \epsilon^3\langle t\rangle^{2s-3}. \tag{8.17}$$

The first term is absorbed by $CK^{v,2}_L$ and the latter is time integrable since $s<1$.

Combining (8.10), (8.11), (8.13), (8.14), (8.15), (8.16), (8.17) with the bootstrap hypotheses and Proposition 2.4 completes the proof of (2.30b).

Proof of (2.30a): From (8.7) we have

$$\frac{1}{2}\frac{d}{dt}\left\|A^R h\right\|_2^2 = -CK^h_\lambda - CK^h_w - \int A^R h A^R\big([\partial_t v]\partial_v h\big)dv + \int A^R h A^R \bar{h} dv, \tag{8.18}$$

where

$$CK^h_w(\tau) = \left\|\sqrt{\frac{\partial_t w}{w}}A^R h(\tau)\right\|_2^2 \tag{8.19a}$$

$$CK^h_\lambda(\tau) = \big(-\dot{\lambda}(\tau)\big)\left\||\partial_v|^{s/2}A^R h(\tau)\right\|_2^2. \tag{8.19b}$$

Using the integration by parts trick as in Section 2.3:

$$
\begin{aligned}
&-\int \mathrm{A}^{\mathrm{R}} h \mathrm{A}^{\mathrm{R}}\left([\partial_t v]\partial_v h\right) dv \\
&= \frac{1}{2}\int \partial_v[\partial_t v]\left|\mathrm{A}^{\mathrm{R}} h\right|^2 dv + \int \mathrm{A}^{\mathrm{R}} h\left([\partial_t v]\mathrm{A}^{\mathrm{R}}\partial_v h - \mathrm{A}^{\mathrm{R}}\left([\partial_t v]\partial_v h\right)\right) dv \\
&= \mathcal{E} + \mathcal{M}.
\end{aligned}
$$

By Sobolev embedding and the bootstrap control on $[\partial_t v]$,

$$
\mathcal{E} \lesssim \frac{\epsilon}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}\left\|\mathrm{A}^{\mathrm{R}} h\right\|_2^2.
$$

As in Section 2.3, the commutator $\mathcal{M}$ is decomposed with a paraproduct,

$$
\begin{aligned}
\mathcal{M} &= \sum_{\mathrm{M}\geq 8}\int \mathrm{A}^{\mathrm{R}} h\left([\partial_t v]_{<\mathrm{M}/8}\mathrm{A}^{\mathrm{R}}\partial_v h_{\mathrm{M}} - \mathrm{A}^{\mathrm{R}}\left([\partial_t v]_{<\mathrm{M}/8}\partial_v h_{\mathrm{M}}\right)\right) dv \\
&\quad + \sum_{\mathrm{M}\geq 8}\int \mathrm{A}^{\mathrm{R}} h\left([\partial_t v]_{\mathrm{M}}\mathrm{A}^{\mathrm{R}}\partial_v h_{<\mathrm{M}/8} - \mathrm{A}^{\mathrm{R}}\left([\partial_t v]_{\mathrm{M}}\partial_v h_{<\mathrm{M}/8}\right)\right) dv \\
&\quad + \sum_{\mathrm{M}\in\mathbf{D}}\sum_{\mathrm{M}'\approx\mathrm{M}}\int \mathrm{A}^{\mathrm{R}} h\left([\partial_t v]_{\mathrm{M}}\mathrm{A}^{\mathrm{R}}\partial_v h_{\mathrm{M}'} - \mathrm{A}^{\mathrm{R}}\left([\partial_t v]_{\mathrm{M}}\partial_v h_{\mathrm{M}'}\right)\right) dv \\
&= \mathrm{T}^h + \mathrm{R}^h + \mathcal{R}^h.
\end{aligned}
$$

The T^h and $\mathcal{R}^h$ terms can be treated as in the methods used in Sections 5 and 7 with u replaced just with $[\partial_t v]$ (also there are no R vs NR losses from Lemma 3.6 since the high frequency factors all have the same 'resonant' regularity). Hence, we omit the treatment and simply conclude

$$
\mathrm{T}^h + \mathcal{R}^h \lesssim \epsilon \mathrm{CK}^h_\lambda + \epsilon^3\langle t\rangle^{-2+\mathrm{K_D}\epsilon/2}. \tag{8.20}
$$

In the 'reaction' term, R^h, we have an issue: a loss of regularity as $[\partial_t v]$ has only 'non-resonant' regularity. Consider one dyadic shell on the frequency side as in Section 6,

$$
\begin{aligned}
\mathrm{R}^h_{\mathrm{M}} &= -\frac{1}{2\pi}\int_{\eta,\xi} \mathrm{A}^{\mathrm{R}}\bar{\hat{h}}(\eta)_{\sim\mathrm{M}}\left[\mathrm{A}^{\mathrm{R}}(\eta) - \mathrm{A}^{\mathrm{R}}(\eta-\xi)\right]\widehat{[\partial_t v]}(\xi)_{\mathrm{M}} \\
&\quad \times \widehat{\partial_v h}(\eta-\xi)_{<\mathrm{M}/8}\, d\eta\, d\xi \\
&= \mathrm{R}^{h;1}_{\mathrm{M}} + \mathrm{R}^{h;2}_{\mathrm{M}}.
\end{aligned}
$$

The treatment of $\mathrm{R}^{h;2}_{\mathrm{M}}$ is essentially the same as the analogous $\mathrm{R}^3_{\mathrm{N}}$ in Section 6 and is hence omitted:

$$
\sum_{\mathrm{M}\geq 8}\left|\mathrm{R}^{h;2}_{\mathrm{M}}\right| \lesssim \frac{\epsilon^3}{\langle t\rangle^{2-\mathrm{K_D}\epsilon/2}}. \tag{8.21}
$$

The more interesting $R_M^{h,1}$ is treated in a manner similar to Section 6.3 above. We first divide the integral into the contributions where $A^R(t,\eta)$ disagrees noticeably with $A_0(t,\xi)$ and where it does not:

$$\begin{aligned} R_M^{h,1} &= -\frac{1}{2\pi}\int_{\eta,\xi}\left[\sum_{k\neq 0}\mathbf{1}_{t\in\mathbf{I}_{k,\eta}}\mathbf{1}_{t\in\mathbf{I}_{k,\xi}}+\chi^*\right] \\ &\qquad \times A^R\bar{\hat{h}}(\eta)_{\sim M}A^R(\eta)\widehat{[\partial_t v]}(\xi)_M\widehat{\partial_v h}(\eta-\xi)_{<M/8}d\eta d\xi \\ &= \left(\sum_{k\neq 0}R_{M,k}^{h,1;D}\right)+R_M^{h,1;*}, \end{aligned}$$

where $\chi^* = 1-\sum_{k\neq 0}\mathbf{1}_{t\in\mathbf{I}_{k,\eta}}\mathbf{1}_{t\in\mathbf{I}_{k,\xi}}$. We first treat $R_M^{h,1;*}$. Due to the presence of χ^*, we do not lose much by replacing $J^R(t,\eta)$ with $J_0(t,\xi)$ (recall the definition of A^R and J^R in (3.10)). Indeed, by the proof of Lemma 3.6 we have on the support of the integrand:

$$\chi^*\frac{J^R(t,\eta)}{J_0(t,\xi)} \lesssim e^{10\mu|\eta-\xi|^{1/2}}\chi^*.$$

Therefore, since $||\eta|-|\xi|| \leq |\eta-\xi| < 3|\eta|/16$ on the support of the integrand, (A.7), (A.11) and (A.12) together imply that for some $c\in(0,1)$ we have,

$$\begin{aligned} \left|R_M^{h,1;*}\right| &\lesssim \int_{\eta,\xi}\chi^*\left|A^R\hat{h}(\eta)_{\sim M}\right|\left|A_0(\xi)\widehat{[\partial_t v]}(\xi)_M e^{c\lambda|\eta-\xi|^s}\right. \\ &\qquad \times \left.\widehat{\partial_v h}(\eta-\xi)_{<M/8}\right|d\eta d\xi. \end{aligned}$$

Using that $|\eta|\approx|\xi|\gtrsim 1$ on the support of the integrand and applying (A.4) (with the bootstrap hypotheses) we deduce,

$$\begin{aligned} \left|R_M^{h,1;*}\right| &\lesssim \int_{\eta,\xi}\chi^*\left|A^R\hat{h}(\eta)_{\sim M}\right|\frac{|\eta|^{s/2}|\xi|^{s/2}}{\langle\xi\rangle^s}\left|A_0(\xi)\widehat{[\partial_t v]}(\xi)_M\right. \\ &\qquad \times \left. e^{c\lambda|\eta-\xi|^s}\widehat{\partial_v h}(\eta-\xi)_{<M/8}\right|d\eta d\xi \\ &\lesssim \epsilon\left\||\partial_v|^{s/2}A^R h_{\sim M}\right\|_2\left\||\partial_v|^{s/2}\frac{A}{\langle\partial_v\rangle^s}[\partial_t v]_M\right\|_2 \\ &\lesssim \epsilon\langle t\rangle^{-2s}\left\||\partial_v|^{s/2}A^R h_{\sim M}\right\|_2^2+\epsilon\langle t\rangle^{2s}\left\||\partial_v|^{s/2}\frac{A}{\langle\partial_v\rangle^s}[\partial_t v]_M\right\|_2^2. \end{aligned} \tag{8.22}$$

For ϵ sufficiently small, the first term is absorbed by CK_λ^h and the second term is controlled by $CK_\lambda^{v,1}$ since $2s < 2+2s-2\tilde{q}$.

Next turn to $R_{M,k}^{h,1;D}$. The following version of (3.34) applies on the support of the integrand with an analogous proof (recall (3.10)):

$$\frac{J^R(\eta)}{J_0(\xi)} \lesssim \frac{|\eta|}{k^2}\sqrt{\frac{\partial_t w_0(t,\eta)}{w_0(t,\eta)}}\sqrt{\frac{\partial_t w_0(t,\xi)}{w_0(t,\xi)}}e^{20\mu|\eta-\xi|^{1/2}}. \tag{8.23}$$

Therefore, by (A.7) and (A.11) there exists $c \in (0,1)$ such that

$$\begin{aligned}\left|R_{M,k}^{h,1;D}\right| \lesssim & \int_{\eta,\xi} \mathbf{1}_{t\in \mathbf{I}_{k,\eta}}\mathbf{1}_{t\in \mathbf{I}_{k,\xi}}\left|\sqrt{\frac{\partial_t w}{w}}A^R\hat{h}(\eta)_{\sim M}\right|\frac{|\eta|}{k^2} \\ & \times \left|\sqrt{\frac{\partial_t w}{w}}A\widehat{[\partial_t v]}(\xi)_M e^{c\lambda|\eta-\xi|^s}\widehat{\partial_v h}(\eta-\xi)_{<M/8}\right|d\eta d\xi.\end{aligned}$$

On the support of the integrand $kt \approx \eta \approx \xi$, so by (A.4) and the bootstrap hypotheses,

$$\begin{aligned}\left|R_{M,k}^{h,1;D}\right| \lesssim & \, t^{1+s}\int_{\eta,\xi} \mathbf{1}_{t\in \mathbf{I}_{k,\eta}}\mathbf{1}_{t\in \mathbf{I}_{k,\xi}}\left|\sqrt{\frac{\partial_t w}{w}}A^R\hat{h}(\eta)_{\sim M}\right|\frac{1}{\langle\xi\rangle^s} \\ & \times \left|\sqrt{\frac{\partial_t w}{w}}A\widehat{[\partial_t v]}(\xi)_M e^{c\lambda|\eta-\xi|^s}\widehat{\partial_v h}(\eta-\xi)_{<M/8}\right|d\eta d\xi \\ \lesssim & \, \epsilon t^{1+s}\left\|\sqrt{\frac{\partial_t w}{w}}A^R\mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}}h_{\sim M}\right\|_2\left\|\sqrt{\frac{\partial_t w}{w}}\frac{A}{\langle\partial_v\rangle^s}\mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}}[\partial_t v]_M\right\|_2,\end{aligned}$$

where we are denoting the Fourier multiplier $(\widehat{\mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}}f})(t,\eta) = \mathbf{1}_{t\in \mathbf{I}_{k,\eta}}\hat{f}(t,\eta)$. For t fixed, the supports of these multipliers for different k are disjoint in frequency, and hence we can sum:

$$\begin{aligned}\sum_{k\neq 0}\left|R_{M,k}^{h,1;D}\right| \lesssim & \sum_{k\neq 0}\epsilon\left\|\sqrt{\frac{\partial_t w}{w}}A^R\mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}}h_{\sim M}\right\|_2^2 \\ & + \epsilon t^{2+2s}\left\|\sqrt{\frac{\partial_t w}{w}}\frac{A}{\langle\partial_v\rangle^s}\mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}}[\partial_t v]_M\right\|_2^2 \\ \approx & \, \epsilon\left\|\sqrt{\frac{\partial_t w}{w}}A^R h_{\sim M}\right\|_2^2 + \epsilon t^{2+2s}\left\|\sqrt{\frac{\partial_t w}{w}}\frac{A}{\langle\partial_v\rangle^s}[\partial_t v]_M\right\|_2^2,\end{aligned} \tag{8.24}$$

which are respectively absorbed by CK_w^h and bounded by $CK_w^{v,1}$. This concludes the treatment of the first (non CK) term in (8.18).

The treatment of the term involving $\bar{h}$ in (8.18) is similar to $R_M^{h,1}$. That the dependence is linear is the reason for the presence of the large constant K_v in (2.30a). As in the treatment of $R_M^{h,1}$ we divide based on frequency:

$$\int A^R h A^R \bar{h} dv = \int_\eta \Big[\sum_{k\neq 0} \mathbf{1}_{t\in \mathbf{I}_{k,\eta}} + \chi^*\Big] A^R \hat{h}(\eta) A^R(\eta) \widehat{\bar{h}}(\eta) d\eta$$
$$= \Big(\sum_{k\neq 0} H_k^D\Big) + H^*,$$

where $\chi^* = 1 - \sum_{k\neq 0} \mathbf{1}_{t\in \mathbf{I}_{k,\eta}}$.

First turn to H_k^D. Here, (8.23) with $\xi = \eta$ holds on the support of the integrand, as does $|\eta| \gtrsim 1$, and hence

$$\left|H_k^D\right| \lesssim \int_\eta \mathbf{1}_{t\in \mathbf{I}_{k,\eta}} \left|A^R \hat{h}(\eta)\right| \frac{\partial_t w(\eta)}{w(\eta)} \frac{|\eta|^{1+s}}{k^2} \left|\frac{A_0(\eta)}{\langle \eta\rangle^s} \widehat{\bar{h}}(\eta)\right| d\eta.$$

As the Fourier restrictions have disjoint support and since $kt \approx \eta$ on the support of the integrand, we get, by Cauchy-Schwarz, the following for some fixed constant $C > 0$,

$$\sum_{k\neq 0} \left|H_k^D\right| \le \sum_{k\neq 0} \frac{1}{4} \left\|\sqrt{\frac{\partial_t w}{w}} A^R \mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}} h\right\|_2^2 + Ct^{2+2s} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v\rangle^s} \mathbf{1}_{t\in \mathbf{I}_{k,\partial_v}} \bar{h}\right\|_2^2$$
$$\le \frac{1}{4} \left\|\sqrt{\frac{\partial_t w}{w}} A^R h\right\|_2^2 + Ct^{2+2s} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v\rangle^s} \bar{h}\right\|_2^2. \tag{8.25}$$

The first term is absorbed by CK_w^h and the latter controlled by the bootstrap hypothesis on $CK_w^{v,2}$, provided we choose $K_v \gg C$.

Next we turn to H^*. Due to the presence of χ^*, $A^R(t,\eta) \approx A_0(t,\eta)$ on the support of the integrand, by the same proof as (3.31). Hence, we do not need to concern ourselves with the distinction. However, we still need to recover the gap of s derivatives. We treat high and low frequencies separately: for some $C > 0$, we have by Cauchy-Schwarz,

$$\left|H^*\right| \lesssim \int_{|\eta|\ge 1} \chi^* \left|A^R h(\eta)\right| |\eta|^s \left|\frac{A}{\langle \eta\rangle^s} \bar{h}(\eta)\right| d\eta$$
$$+ \int_{|\eta|<1} \chi^* \left|A^R h(\eta)\right| \left|\frac{A}{\langle \eta\rangle^s} \bar{h}(\eta)\right| d\eta$$
$$\lesssim \left\||\partial_v|^{s/2} A^R h\right\|_2 \left\||\partial_v|^{s/2} \frac{A}{\langle \partial_v\rangle^s} \bar{h}\right\|_2 + \left\|A^R h\right\|_2 \left\|\frac{A}{\langle \partial_v\rangle^s} \bar{h}\right\|_2$$
$$\le \frac{\delta_\lambda}{10 t^{2\tilde{q}}} \left\||\partial_v|^{s/2} A^R h\right\|_2^2 + \frac{C}{\delta_\lambda} t^{2\tilde{q}} \left\||\partial_v|^{s/2} \frac{A}{\langle \partial_v\rangle^s} \bar{h}\right\|_2^2 + C\left\|A^R h\right\|_2 \left\|\frac{A}{\langle \partial_v\rangle^s} \bar{h}\right\|_2. \tag{8.26}$$

The first term is absorbed by CK_λ^h and the latter terms are controlled by the bootstrap hypotheses on $\mathrm{CK}_\lambda^{v,2}$ and $\bar{h}$ provided we choose $\mathrm{K}_v \gg \mathrm{C}\delta_\lambda^{-2}$ and $2+2s > 4\tilde{q}$.

Finally, summing together (8.20), (8.21), (8.22), (8.24), (8.25), (8.26) (using almost orthogonality) with (8.18) and the bootstrap hypotheses implies that for ϵ chosen sufficiently small,

$$\left\|\mathrm{A}^{\mathrm{R}} h(t)\right\|_2^2 + \frac{1}{2}\int_1^t \mathrm{CK}_w^h(\tau)d\tau + \frac{1}{2}\int_1^t \mathrm{CK}_\lambda^h(\tau)d\tau \lesssim \mathrm{K}_v\epsilon^2, \tag{8.27}$$

which almost proves (2.30a). To complete the proof we need to apply the product rules in Lemma 3.8. Indeed, writing $(v')^2 - 1 = (v'-1)^2 + 2(v'-1)$, and applying Lemma 3.8 and (8.27) gives,

$$\mathrm{CCK}_w^1 + \mathrm{CCK}_\lambda^1 \lesssim \mathrm{CK}_w^h + \mathrm{CK}_\lambda^h + \epsilon^2\left(\mathrm{CK}_w^h + \mathrm{CK}_\lambda^h\right). \tag{8.28}$$

Similarly, for CCK_λ^2 and CCK_w^2 terms, write $v'\partial_v v' = \partial_v(v'-1) + (v'-1)\partial_v(v'-1)$ and apply Lemma 3.8:

$$\begin{aligned}\mathrm{CCK}_\lambda^2 + \mathrm{CCK}_w^2 &\lesssim \mathrm{CK}_w^h + \mathrm{CK}_\lambda^h + \left\|\sqrt{\frac{\partial_t w}{w}}\frac{\mathrm{A}^{\mathrm{R}}}{\langle\partial_v\rangle}\left(\left(1-v'\right)\partial_v v'\right)\right\|_2^2 \\ &\quad + \left\||\partial_v|^{s/2}\frac{\mathrm{A}^{\mathrm{R}}}{\langle\partial_v\rangle}\left(\left(1-v'\right)\partial_v v'\right)\right\|_2^2 \\ &\lesssim \mathrm{CK}_w^h + \mathrm{CK}_\lambda^h + \epsilon^2\left(\mathrm{CK}_w^h + \mathrm{CK}_\lambda^h\right).\end{aligned} \tag{8.29}$$

Hence, by possibly adjusting K_v and choosing ϵ even smaller, (8.27), (8.28) and (8.29) imply (2.30a).

Proof of (2.30d)*:* Both of the terms in (2.30d) are controlled in essentially the same way. To see the control on the first term, start with dividing into high and low frequency and use the bootstrap control on $[\partial_t v]$:

$$\begin{aligned}&\langle\tau\rangle^{2+2s}\left|\dot{\lambda}(\tau)\right|\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}[\partial_t v](\tau)\right\|_2^2 \\ &\quad\le \langle\tau\rangle^{2+2s}\left|\dot{\lambda}(\tau)\right|\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}[\partial_t v](\tau)_{\le 1}\right\|_2^2 \\ &\qquad + \langle\tau\rangle^{2+2s}\left|\dot{\lambda}(\tau)\right|\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}[\partial_t v](\tau)_{>1}\right\|_2^2 \\ &\quad\lesssim \langle\tau\rangle^{2s-2-2\tilde{q}+\mathrm{K}_{\mathrm{D}}\epsilon}\epsilon^2 + \langle\tau\rangle^{2+2s}\left|\dot{\lambda}(\tau)\right|\left\||\partial_v|^{s/2}\frac{\mathrm{A}}{\langle\partial_v\rangle^s}\partial_v[\partial_t v](\tau)_{>1}\right\|_2^2.\end{aligned} \tag{8.30}$$

The first term is integrable for ϵ small and the latter term is bounded by

$$\left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \partial_v [\partial_t v](\tau) \right\|_2^2 \lesssim \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \big(v' \partial_v [\partial_t v](\tau) \big) \right\|_2^2 + \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \left(\frac{v'-1}{v'} v' \partial_v [\partial_t v](\tau) \right) \right\|_2^2. \tag{8.31}$$

The first term in (8.31) is already controlled by $CK_\lambda^{v,2}$. By (3.39b), the second term is bounded by the following for some $c \in (0,1)$:

$$\begin{aligned} &\left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \left(\frac{v'-1}{v'} v' \partial_v [\partial_t v](\tau) \right) \right\|_2^2 \\ &\quad \lesssim \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \left(\frac{v'-1}{v'} \right) \right\|_2^2 \left\| v' \partial_v [\partial_t v](\tau) \right\|_{\mathcal{G}^{c\lambda,\sigma}}^2 \\ &\qquad + \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} v' \partial_v [\partial_t v](\tau) \right\|_2^2 \left\| \frac{v'-1}{v'} \right\|_{\mathcal{G}^{c\lambda,\sigma}}^2. \end{aligned} \tag{8.32}$$

By the bootstrap hypotheses, we may write $(v')^{-1}$ as the uniformly convergent geometric series

$$\frac{1}{v'(t,v)} = 1 + \sum_{n=1}^{\infty} \big(v'(t,v) - 1 \big)^n.$$

Therefore, by (A.10) and the bootstrap control on $v'-1$, for some $C>0$ we have for ϵ small:

$$\left\| \frac{v'-1}{v'} \right\|_{\mathcal{G}^{c\lambda,\sigma}} = \left\| \sum_{n=1}^{\infty} \big(v'-1 \big)^n \right\|_{\mathcal{G}^{c\lambda,\sigma}} \leq \sum_{n=1}^{\infty} (C\epsilon)^n \lesssim \epsilon. \tag{8.33}$$

Together with the bootstrap control on $CK_\lambda^{v,2}$, this suffices to treat the second term in (8.32). To control the first term, we use a repeated application of (3.39a) and (3.40), choosing ϵ sufficiently small to sum the resulting series (denoting constants in these inequalities as C_p and C_a respectively),

$$\begin{aligned} &\left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \left(\frac{v'-1}{v'} \right) \right\|_2^2 \\ &\quad = \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \sum_{n=1}^{\infty} \big(v'-1 \big)^n \right\|_2^2 \end{aligned}$$

$$\leq \left(\sum_{n=1}^{\infty} \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} (v'-1)^n \right\|_2 \right)^2$$
$$\leq \left(\left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} (v'-1) \right\|_2 \sum_{n=1}^{\infty} \epsilon^{n-1} 4^{n-1} K_v^{n-1} \sum_{j=1}^{n} C_p^{n-j} C_a^{j-1} \right)^2$$
$$\leq \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} (v'-1) \right\|_2^2 \left(\sum_{n=1}^{\infty} \epsilon^{n-1} n 4^{n-1} C_p^{n-1} K_v^{n-1} C_a^{n-1} \right)^2$$
$$\lesssim \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} (v'-1) \right\|_2^2.$$

Therefore, putting this together with (8.30), (8.31), (8.32), (8.33) and the bootstrap control on $v'\partial_v[\partial_t v]$ implies,

$$\langle \tau \rangle^{2+2s} \left|\dot{\lambda}(\tau)\right| \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \partial_v[\partial_t v](\tau) \right\|_2^2$$
$$\lesssim CK_\lambda^{v,2} + \epsilon^2 \left|\dot{\lambda}(\tau)\right| \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} (v'-1) \right\|_2^2.$$

The first term is integrable by the bootstrap hypotheses and the latter term is integrable by the CK_λ^h bound in (8.27). From (8.30), this completes the treatment of the first term in (2.30d).

To see control on the second term in (2.30d), first divide into high and low frequency and use that $\partial_t w$ is only supported in frequencies larger than $1/2$ (see (3.9)):

$$\langle \tau \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} [\partial_t v](\tau) \right\|_2^2 \lesssim \langle \tau \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} \partial_v[\partial_t v](\tau)_{>1/2} \right\|_2^2.$$

From here, it is very similar to the treatment of the first term in (2.30d), except now we apply (3.39b) as opposed to (3.39a). We omit this for the sake of brevity.

Proof of (2.30c)*:* This estimate is relatively straightforward to prove since, due to the lower regularity, we can utilize Lemma 4.1 as opposed to Proposition 2.4. It is possible to only use a gap of s derivatives and apply Proposition 2.4, however, one will be repeating a more intricate version of the arguments used to deduce (2.30b).

Computing the evolution of $[\partial_t v]$ gives

$$\frac{d}{dt}\left(\langle t \rangle^{4-K_D\epsilon} \left\| [\partial_t v] \right\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 \right) = (4-K_D\epsilon) t \langle t \rangle^{2-K_D\epsilon} \left\| [\partial_t v] \right\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 + \langle t \rangle^{4-K_D\epsilon} \frac{d}{dt} \left\| \frac{A}{\langle \partial_v \rangle^s} [\partial_t v] \right\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2. \tag{8.34}$$

Denoting the multiplier $A^S(t,\partial_v) = e^{\lambda(t)|\partial_v|^s}\langle\partial_v\rangle^{\sigma-6}$ ('S' for 'simple'), the latter term gives

$$\langle t\rangle^{4-K_D\epsilon}\frac{d}{dt}\left\|[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda(t),\sigma-6}} = 2\langle t\rangle^{4-K_D\epsilon}\dot\lambda(t)\left\||\partial_v|^{s/2}[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda(t),\sigma-6}} + 2\langle t\rangle^{4-K_D\epsilon}\int A^S[\partial_t v]A^S\partial_t[\partial_t v]dv. \tag{8.35}$$

From (2.13b),

$$\begin{aligned} 2\langle t\rangle^{4-K_D\epsilon}\int A^S[\partial_t v]A^S\partial_t[\partial_t v]dv &= -\frac{4\langle t\rangle^{4-K_D\epsilon}}{t}\left\|[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda(t),\sigma-6}} \\ &\quad - 2\langle t\rangle^{4-K_D\epsilon}\int A^S[\partial_t v]A^S\big([\partial_t v]\partial_v[\partial_t v]\big)dv \\ &\quad - \frac{2\langle t\rangle^{4-K_D\epsilon}}{t}\int A^S[\partial_t v]A^S\big(v'\langle\nabla^\perp P_{\neq 0}\phi\cdot\nabla\tilde u\rangle\big)dv \\ &= V_1 + V_2 + V_3. \end{aligned} \tag{8.36}$$

By (A.10), an argument analogous to (8.33) and the bootstrap hypotheses we have,

$$\begin{aligned} V_2 &\lesssim \langle t\rangle^{4-K_D\epsilon}\left\|[\partial_t v]\right\|_{\mathcal{G}^{\lambda,\sigma-6}}\left\|[\partial_t v]\left(1+\frac{1-v'}{v'}\right)v'\partial_v[\partial_t v]\right\|_{\mathcal{G}^{\lambda,\sigma-6}} \\ &\lesssim \langle t\rangle^{4-K_D\epsilon}\left\|[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda,\sigma-6}}\left\|v'\partial_v[\partial_t v]\right\|_{\mathcal{G}^{\lambda,\sigma-6}}\left(1+\left\|\frac{v'-1}{v'}\right\|_{\mathcal{G}^{\lambda,\sigma-6}}\right) \\ &\leq \frac{K_D\epsilon}{2}\langle t\rangle^{3-K_D\epsilon-s}\left\|[\partial_t v]\right\|^2_{\mathcal{G}^{\lambda,\sigma-6}}, \end{aligned} \tag{8.37}$$

where we define K_D to be the maximum of the constant appearing in this term and one other below.

Treating V_3 is not hard due to the regularity gap of 6 derivatives. Note that

$$\nabla\tilde u = -\begin{pmatrix} v'(\partial_v - t\partial_z)\partial_z\phi \\ \partial_v v'(\partial_v - t\partial_z)\phi + v'(\partial_v - t\partial_z)\partial_v\phi \end{pmatrix}, \tag{8.38}$$

and therefore by (A.10), (8.38), Lemma 4.1 and the bootstrap hypotheses,

$$\left\|\nabla P_{\neq 0}\tilde u(t)\right\|_{\mathcal{G}^{\lambda(t),\sigma-5}} \lesssim \frac{\epsilon}{\langle t\rangle}. \tag{8.39}$$

Projecting to individual frequencies in z and noting that $k \neq 0$ (by the z average and the projection on ϕ), by (A.10) and Cauchy-Schwarz we have

$$V_3 = \sum_{k\neq 0}\frac{2\langle t\rangle^{4-K_D\epsilon}}{t}\int A^S[\partial_t v]A^S\big[v'\nabla^\perp\phi_k\cdot\nabla\tilde u_{-k}\big]dv$$

$$\lesssim \sum_{k\neq 0} \frac{\langle t\rangle^{4-\mathrm{K_D}\epsilon}}{t} \left\| \mathrm{A^S}[\partial_t v] \right\|_2 \left\| \mathrm{A^S}\nabla^\perp \phi_k \right\|_2 \left\| \mathrm{A^S}\nabla \tilde{u}_{-k} \right\|_2$$
$$\times \left(1 + \left\| \mathrm{A^S}(v'-1) \right\|_2\right)$$
$$\lesssim \frac{\langle t\rangle^{4-\mathrm{K_D}\epsilon}}{t} \left\| [\partial_t v] \right\|_{\mathcal{G}^{\lambda,\sigma-6}} \left\| \nabla^\perp \mathrm{P}_{\neq 0}\phi \right\|_{\mathcal{G}^{\lambda,\sigma-6}} \| \nabla \mathrm{P}_{\neq 0}\tilde{u} \|_{\mathcal{G}^{\lambda,\sigma-6}}$$
$$\times \left(1 + \left\| v' - 1 \right\|_{\mathcal{G}^{\lambda,\sigma-6}}\right).$$

By (8.39), the bootstrap hypotheses on $v' - 1$, and Lemma 4.1 we then have for some $\mathrm{C} > 0$,

$$\mathrm{V}^3 \lesssim \langle t\rangle^{-\mathrm{K_D}\epsilon} \epsilon^2 \left\| [\partial_t v] \right\|_{\mathcal{G}^{\lambda,\sigma-6}} \leq \frac{\mathrm{K_D}\epsilon \langle t\rangle^{4-\mathrm{K_D}\epsilon}}{2t} \left\| [\partial_t v] \right\|^2_{\mathcal{G}^{\lambda,\sigma-6}} + \mathrm{C}\epsilon^3 t^{-3-\mathrm{K_D}\epsilon}. \tag{8.40}$$

Putting together (8.37),(8.40) with (8.36) and (8.35) and integrate, we derive (2.30c). This completes the proof of Proposition 2.5. Being the last Proposition remaining, the proof of Theorem 1 is complete.

9. Conclusion

Our proof is very different from that of Mouhot and Villani [70], however there are still some analogies and common themes that are worth pointing out. Let us mention the most important mathematical parallels:

- In [70], the role of the plasma echoes has similarities with that of the Orr critical times, being that both are responsible for the main nonlinear growth. The moment estimates in [70] controls this growth whereas here the "toy model" plays this role. Note that the exponential growth in time of [70] is replaced here by a controlled regularity loss.
- The use of para-differential calculus (combined with the well-established existence theory) allows to avoid the use of the Newton iteration in [70]. For example, the paraproduct decomposition permits us to separate the natural transport effects and the reaction effects.
- Our treatment of transport in the energy estimates and the well-chosen change of variables allow us to avoid the use of 'deflection maps' as in [70].

Although the Euler and Vlasov-Poisson systems have several fundamental differences, after completing this work we succeeded, with C. Mouhot, to extend the Landau damping result in [70] to all Gevrey class smaller than three (e.g. $s > 1/3$) using some of the ideas of this work [12].

An obvious question about Theorem 1 is whether or not $s > 1/2$ is optimal. The toy model (3.3), which estimates the worst possible growth and is behind the regularity requirement, took absolute values and hence does not take into account the potential for oscillations and cancellations. In fact, numerical simulations of the 3-wave model of [92] show dramatic oscillations. Even when there are infinitely many interacting modes, oscillations may often cancel and yield much weaker growth than the one given by (3.3). However, we think it would be extremely difficult to rule out the possibility that there exist rare configurations which are "resonant" in some sense and lead to a growth which matches our toy model (we expect that such configurations are highly non-generic).

There are many other related problems where the linear operator predicts a decay by mixing similar to (1.1). The most obvious extension is to include a more general class of shear flows where now we linearize around the flow $(V(y), 0)$. We believe that this requires the incorporation of several non-trivial enhancements, since it fundamentally changes the structure of the critical times which would manifest in our approach most clearly in the elliptic estimates. A related extension would be to study the problem in the presence of no-penetration boundaries in y. A third problem would be to remove the periodicity in x, altering the physical mechanism from mixing to *filamentation* (in fact the behavior may be different, as is possible in Landau damping, see [40, 41]). Here there are fundamental difficulties: our proof relies heavily on the 'well-separation' of critical times, which no longer holds in the unbounded case. Other examples in fluid mechanics pointed out in Section 1 include the β-plane model, stratified flows and the vortex axisymmetrization problem. Proving decay by mixing on the nonlinear level for any of these models seems to be a very interesting question. Of all of these, we expect the β-plane model to be the easiest.

Acknowledgements

The authors would like to warmly thank Clement Mouhot for sharing his many insights about this problem and for enlightening discussions on this topic and the topic of Landau damping. We would also like to thank the following people for references and suggestions: Antoine Cerfon, George Hagstrom, Susan Friedlander, Andrew Majda, Vladimir Sverak, Vlad Vicol, Cedric Villani and Chongchun Zeng as well as David Gérard-Varet for pointing out an error in the first draft. J. Bedrossian would like to thank David Uminsky for pointing him to Landau damping and its connection with 2D Euler.

Appendix A

A.1 Littlewood-Paley decomposition and paraproducts

In this section we fix conventions and notation regarding Fourier analysis, Littlewood-Paley and paraproduct decompositions. See e.g. [4, 14] for more details.

For $f(z, v)$ in the Schwartz space, we define the Fourier transform $\hat{f}_k(\eta)$ where $(k, \eta) \in \mathbf{Z} \times \mathbf{R}$,

$$\hat{f}_k(\eta) = \frac{1}{2\pi} \int_{\mathbf{T}\times\mathbf{R}} e^{-izk-iv\eta} f(z, v)\, dz dv.$$

Similarly we have the Fourier inversion formula,

$$f(z, v) = \frac{1}{2\pi} \sum_{k\in\mathbf{Z}} \int_{\mathbf{R}} e^{izk+iv\eta} \hat{f}_k(\eta)\, d\eta.$$

As usual the Fourier transform and its inverse are extended to L^2 via duality. We also need to apply the Fourier transform to functions of v alone, for which we use analogous conventions. With these conventions note,

$$\int f(z, v)\overline{g}(z, v)\, dz dv = \sum_k \int \hat{f}_k(\eta)\overline{\hat{g}}_k(\eta)\, d\eta$$

$$\widehat{fg} = \frac{1}{2\pi} \hat{f} * \hat{g}$$

$$(\widehat{\nabla f})_k(\eta) = (ik, i\eta)\hat{f}_k(\eta).$$

This work makes heavy use of the Littlewood-Paley dyadic decomposition. Here we fix conventions and review the basic properties of this classical theory, see e.g. [4] for more details. First we define the Littlewood-Paley decomposition only in the v variable. Let $\psi \in C_0^\infty(\mathbf{R}; \mathbf{R})$ be such that $\psi(\xi) = 1$ for $|\xi| \le 1/2$ and $\psi(\xi) = 0$ for $|\xi| \ge 3/4$ and define $\rho(\xi) = \psi(\xi/2) - \psi(\xi)$, supported in the range $\xi \in (1/2, 3/2)$. Then we have the partition of unity

$$1 = \psi(\xi) + \sum_{M\in 2^{\mathbf{N}}} \rho\left(M^{-1}\xi\right),$$

where we mean that the sum runs over the dyadic numbers $M = 1, 2, 4, 8, \ldots, 2^j, \ldots$ and we define the cut-off $\rho_M(\xi) = \rho(M^{-1}\xi)$, each supported in $M/2 \le |\xi| \le 3M/2$. For $f \in L^2(\mathbf{R})$ we define

$$f_M = \rho_M\left(|\partial_v|\right)f,$$

$$f_{\frac{1}{2}} = \psi\left(|\partial_v|\right)f,$$

$$f_{<M} = f_{\frac{1}{2}} + \sum_{K\in 2^{\mathbf{N}}: K<M} f_K,$$

which defines the decomposition

$$f = f_{\frac{1}{2}} + \sum_{M\in 2^{\mathbf{N}}} f_M.$$

Normally one would use f_0 rather than the slightly inconsistent $f_{\frac{1}{2}}$, however f_0 is reserved for the much more commonly used projection onto the zero mode only in z (or x). Recall the definition of $\mathbf{D}$ from Section 1.2. There holds the almost orthogonality and the approximate projection property

$$\|f\|_2^2 \approx \sum_{M\in\mathbf{D}} \|f_M\|_2^2 \tag{A.1a}$$

$$\|f_M\|_2 \approx \left\|(f_M)_M\right\|_2. \tag{A.1b}$$

The following is also clear:

$$\left\| |\partial_v| f_M \right\|_2 \approx M\|f_M\|_2.$$

We make use of the notation

$$f_{\sim M} = \sum_{K\in\mathbf{D}:\frac{1}{C}M\leq K\leq CM} f_K,$$

for some constant C which is independent of M. Generally the exact value of C which is being used is not important; what is important is that it is finite and independent of M. Similar to (A.1) but more generally, if $f = \sum_k D_k$ for any D_k with $\frac{1}{C}2^k \subset \operatorname{supp} D_k \subset C2^k$ it follows that

$$\|f\|_2^2 \approx_C \sum_k \|D_k\|_2^2. \tag{A.2}$$

During much of the proof we are also working with Littlewood-Paley decompositions defined in the (z, v) variables, with the notation conventions being analogous. Our convention is to use N to denote Littlewood-Paley projections in (z, v) and M to denote projections only in the v direction.

We have opted to use the compact notation above, rather than the commonly used alternatives

$$\Delta_j f = f_{2^j}, \qquad S_j f = f_{<2^j},$$

in order to reduce the number of characters in long formulas. The last unusual notation we use is

$$P_{\neq 0} f = f - \langle f\rangle,$$

which denotes projection onto the non-zero modes in z.

Another key Fourier analysis tool employed in this work is the paraproduct decomposition, introduced by Bony [14] (see also [4]). Given suitable functions f, g we may define the paraproduct decomposition (in either (z, v) or just v),

$$fg = T_f g + T_g f + \mathcal{R}(f, g)$$

$$= \sum_{N\geq 8} f_{<N/8} g_N + \sum_{N\geq 8} g_{<N/8} f_N + \sum_{N\in \mathbf{D}} \sum_{N/8\leq N'\leq 8N} g_{N'} f_N,$$

where all the sums are understood to run over $\mathbf{D}$. In our work we do not employ the notation in the first line since at most steps in the proof we are forced to explicitly write the sums and treat them term-by-term anyway. This is due to the fact that we are working in non-standard regularity spaces and, more crucially, are usually applying multipliers which do not satisfy any version of $AT_f g \approx T_f Ag$. Hence, we have to prove almost everything 'from scratch' and can only rely on standard para-differential calculus as a guide.

A.2 Elementary inequalities and Gevrey spaces

In the sequel we show some basic inequalities which are extremely useful for working in this scale of spaces. The first three are versions of Young's inequality (applied on the frequency-side here).

Lemma **A.1.** — *Let* $f(\xi), g(\xi) \in L^2_\xi(\mathbf{R}^d)$, $\langle\xi\rangle^\sigma h(\xi) \in L^2_\xi(\mathbf{R}^d)$ *and* $\langle\xi\rangle^\sigma b(\xi) \in L^2_\xi(\mathbf{R}^d)$ *for* $\sigma > d/2$, *Then we have*

$$\|f * h\|_2 \lesssim_{\sigma,d} \|f\|_2 \left\| \langle\cdot\rangle^\sigma h \right\|_2, \tag{A.3}$$

$$\int \left| f(\xi)(g*h)(\xi) \right| d\xi \lesssim_{\sigma,d} \|f\|_2 \|g\|_2 \left\| \langle\cdot\rangle^\sigma h \right\|_2 \tag{A.4}$$

$$\int \left| f(\xi)(g*h*b)(\xi) \right| d\xi \lesssim_{\sigma,d} \|f\|_2 \|g\|_2 \left\| \langle\cdot\rangle^\sigma h \right\|_2 \left\| \langle\cdot\rangle^\sigma b \right\|_2. \tag{A.5}$$

Proof. — Inequality (A.3) follows from the $L^2 \times L^1 \to L^2$ Young's inequality and Cauchy-Schwarz:

$$\int \left| h(\xi) \right| d\xi \leq \left(\int \frac{1}{\langle\xi\rangle^{2\sigma}} d\xi \right)^{1/2} \left\| \langle\cdot\rangle^\sigma h \right\|_2 \lesssim \left\| \langle\cdot\rangle^\sigma h \right\|_2.$$

Inequality (A.4) follows from Cauchy-Schwarz and (A.3). For (A.5), apply Young's inequality twice:

$$\begin{aligned} \int \left| f(\xi)(g*h*b)(\xi) \right| d\xi &\lesssim \|f\|_2 \|g*h*b\|_2 \lesssim \|f\|_2 \|g\|_2 \|h*b\|_1 \\ &\lesssim \|f\|_2 \|g\|_2 \|h\|_1 \|b\|_1, \end{aligned}$$

and proceed as above. □

The next set of inequalities show that one can often gain on the index of regularity when comparing frequencies which are not too far apart (provided $0 < s < 1$).

Lemma **A.2.** — *Let* $0 < s < 1$ *and* $x \geq y \geq 0$ *(without loss of generality).*

(i) *If* $x + y > 0$,

$$\left|x^s - y^s\right| \lesssim_s \frac{1}{x^{1-s} + y^{1-s}} |x - y|. \tag{A.6}$$

(ii) *If* $|x - y| \leq x/K$ *for some* $K > 1$ *then*

$$\left|x^s - y^s\right| \leq \frac{s}{(K-1)^{1-s}} |x - y|^s. \tag{A.7}$$

Note $\frac{s}{(K-1)^{1-s}} < 1$ *as soon as* $s^{\frac{1}{1-s}} + 1 < K$.

(iii) *In general,*

$$|x + y|^s \leq \left(\frac{x}{x+y}\right)^{1-s} \left(x^s + y^s\right).$$

In particular, if $y \leq x \leq Ky$ *for some* $K < \infty$ *then*

$$|x + y|^s \leq \left(\frac{K}{1+K}\right)^{1-s} \left(x^s + y^s\right). \tag{A.8}$$

Proof. — Inequality (A.6) follows easily from considering separately $x \geq 2y$ and $x < 2y$.

To prove (A.7) we use that in this case $y^{-1} \leq K/(K-1)x^{-1}$ and hence,

$$x^s \leq y^s + \frac{s}{y^{1-s}}(x - y) \leq y^s + \frac{s}{(K-1)^{1-s}} |x - y|^s.$$

To see (A.8),

$$|x + y|^s = \left(\frac{x}{x+y}\right)|x + y|^s + \left(\frac{y}{x+y}\right)|x + y|^s \leq \left(\frac{x}{x+y}\right)^{1-s} \left(x^s + y^s\right). \qquad \square$$

Using (A.4), (A.7) and (A.8) together with a paraproduct expansion, the following product lemma is relatively straightforward. For contrast, the lemma holds when $s = 1$ only for $c = 1$.

Lemma **A.3** *(Product lemma).* — *For all* $0 < s < 1$, $\sigma \geq 0$ *and* $\sigma_0 > 1$, *there exists* $c = c(s) \in (0, 1)$ *such that the following holds for all* $f, g \in \mathcal{G}^{\lambda,\sigma;s}$*:*

$$\|fg\|_{\mathcal{G}^{\lambda,\sigma;s}} \lesssim \|f\|_{\mathcal{G}^{c\lambda,\sigma_0;s}} \|g\|_{\mathcal{G}^{\lambda,\sigma;s}} + \|g\|_{\mathcal{G}^{c\lambda,\sigma_0;s}} \|f\|_{\mathcal{G}^{\lambda,\sigma;s}}, \tag{A.9a}$$

in particular, $\mathcal{G}^{\lambda,\sigma;s}$ *has the algebra property:*

$$\|fg\|_{\mathcal{G}^{\lambda,\sigma}} \lesssim \|f\|_{\mathcal{G}^{\lambda,\sigma}} \|g\|_{\mathcal{G}^{\lambda,\sigma}}. \tag{A.10}$$

As $\sigma_0 > 1 = d/2$ *(where* d *is the dimension), the implicit constants are independent of* λ.

Gevrey and Sobolev regularities can be related with the following two inequalities.

(i) For all $x \geq 0$, $\alpha > \beta \geq 0$, $C, \delta > 0$,

$$e^{Cx^\beta} \leq e^{C(\frac{C}{\delta})^{\frac{\beta}{\alpha-\beta}}} e^{\delta x^\alpha}; \tag{A.11}$$

(ii) For all $x \geq 0$, $\alpha, \sigma, \delta > 0$,

$$e^{-\delta x^\alpha} \lesssim \frac{1}{\delta^{\frac{\sigma}{\alpha}} \langle x \rangle^\sigma}. \tag{A.12}$$

Together these inequalities show that for $\alpha > \beta \geq 0$, $\|f\|_{\mathcal{G}^{C,\sigma;\beta}} \lesssim_{\alpha,\beta,C,\delta,\sigma} \|f\|_{\mathcal{G}^{\delta,0;\alpha}}$.

A.3 Coordinate transformations in Gevrey spaces

The proof of Theorem 1 requires moving from (x, y) to (z, v) coordinates at the beginning (in Lemma 2.1) and then back again in Section 2.4. It is crucial to notice the regularity losses incurred in this section, as discussed in more depth in [70] where related inequalities play an important role.

It is well-known (see e.g. [54]) that the $\mathcal{G}^{\lambda;s}$ norms have an equivalent 'physical-side' representation which will be convenient here:

$$\|f\|_{\mathcal{G}^\lambda} \approx \left[\sum_{n=0}^\infty \left(\frac{\lambda^n}{(n!)^{\frac{1}{s}}} \left\|D^n f\right\|_2\right)^2\right]^{1/2}. \tag{A.13}$$

In this section it will also be useful to have a slightly more general scale of norms:

$$\|f\|_{\ell^q L^p;\lambda} = \left[\sum_{n=0}^\infty \left(\frac{\lambda^n}{(n!)^{\frac{1}{s}}} \left\|D^n f\right\|_p\right)^q\right]^{1/q}. \tag{A.14}$$

By Hölder's inequality and Sobolev embedding (also (A.12)): for $\lambda > \lambda'$ and $p, q \in [1, \infty]$,

$$\|f\|_{\ell^p L^q;\lambda'} \leq \|f\|_{\ell^1 L^q;\lambda'} \lesssim_{\lambda-\lambda'} \|f\|_{\ell^p L^q;\lambda}, \tag{A.15a}$$

$$\|f\|_{\ell^2 L^\infty;\lambda'} \lesssim_{\lambda-\lambda'} \|f\|_{\ell^2 L^2;\lambda}. \tag{A.15b}$$

One of the norms in the scale (A.14) satisfies an algebra property:

$$\|fg\|_{\ell^1 L^\infty;\lambda} \leq \|f\|_{\ell^1 L^\infty;\lambda} \|g\|_{\ell^1 L^\infty;\lambda}, \tag{A.16}$$

which follows from Leibniz's rule and Young's inequality in a manner similar to several proofs in [70]; we omit the details. A more sophisticated inequality is the following, which estimates the effect of composition on Gevrey regularity.

Lemma **A.4** *(Composition inequality). — For all* $s \in (0, 1]$, $p \in [1, \infty]$ *and* $\lambda > 0$,

$$\left\| F \circ (\mathrm{Id} + G) \right\|_{l^1 L^p;\lambda} \leq \left\| \det (\mathrm{Id} + \nabla G)^{-1} \right\|_\infty^{1/p} \| F \|_{l^1 L^p;\lambda+\nu}, \tag{A.17}$$

where

$$\nu = \| G \|_{l^1 L^\infty;\lambda}.$$

Proof. — For simplicity we restrict the following proof to one dimension, however it holds also in higher dimensions with a similar proof. Denote $H = \mathrm{I}d + G$. Proceeding as in [70], by the Faà di Bruno formula we have

$$\| F \circ H \|_{l^1 L^p;\lambda} \leq \sum_{k=0}^\infty \left\| \left(D^k F \right) \circ H \right\|_p \sum_{\sum_{j=1}^n j m_j = n, \sum_{j=1}^n m_j = k} \frac{\lambda^n}{(n!)^{\frac{1}{s}-1} m_1! \ldots m_n!} \times \prod_{j=1}^n (j!)^{(\frac{1}{s}-1)m_j} \left\| \frac{D^j H}{(j!)^{\frac{1}{s}}} \right\|_\infty^{m_j},$$

where the second summation runs over all possible combinations which satisfy the conditions indicated. Under these conditions, we claim by induction on k that the following always holds:

$$k! \prod_{j=1}^n (j!)^{m_j} \leq n!. \tag{A.18}$$

Since $k \leq n$ it is trivial for $k = n = 1$. To see the inductive step, suppose it is true for a given combination of m_j with $\sum_{j=1}^n m_j = k$ and $\sum_{j=1}^n j m_j = n$ for some choices of k and n. Now consider replacing $m_{j_0} \mapsto m_{j_0} + 1$ for some $1 \leq j_0 \leq n + 1$ (where we consider $m_{n+1} = 0$). This increments k by one and n by j_0 and hence by the inductive hypotheses we need only check

$$(k+1) j_0! \leq (n+1) \ldots (n+j_0),$$

which is clear since $n \geq k$, from which (A.18) follows. Hence by (A.18),

$$\begin{aligned} \| F \circ H \|_{l^1 L^p;\lambda} &\leq \sum_{k=0}^\infty \left\| \left(D^k F \right) \circ H \right\|_p \frac{1}{(k!)^{\frac{1}{s}-1}} \\ &\quad \times \sum_{\sum_{j=1}^n j m_j = n, \sum_{j=1}^n m_j = k} \frac{\lambda^n}{m_1! \ldots m_n!} \prod_{j=1}^n \left\| \frac{D^j H}{(j!)^{\frac{1}{s}}} \right\|_\infty^{m_j} \\ &= \sum_{k=0}^\infty \left\| \left(D^k F \right) \circ H \right\|_p \frac{1}{(k!)^{\frac{1}{s}}} \left[\sum_{j=1}^\infty \lambda^j \left\| \frac{D^j H}{(j!)^{\frac{1}{s}}} \right\|_\infty \right]^k, \end{aligned}$$

$$= \sum_{k=0}^{\infty} \left\| \left(\mathrm{D}^k \mathrm{F}\right) \circ \mathrm{H} \right\|_p \frac{1}{(k!)^{\frac{1}{s}}} \left[\lambda + \sum_{j=1}^{\infty} \lambda^j \left\| \frac{\mathrm{D}^j \mathrm{G}}{(j!)^{\frac{1}{s}}} \right\|_{\infty} \right]^k,$$

where the second to last line followed from the multinomial formula. The proof is completed by changing variables in the Lebesgue norm. □

The next tool is the following inverse function theorem in Gevrey regularity.

Lemma **A.5** *(Inverse function theorem). — Let* $\alpha(x) : \mathbf{T} \times \mathbf{R} \to \mathbf{T} \times \mathbf{R}$ *be a given smooth function and consider the equation*

$$\beta(x) = \alpha\big(x + \beta(x)\big). \tag{A.19}$$

For all $\lambda > \lambda' > 0$, *there exists an* $\epsilon_0 = \epsilon_0(\lambda, \lambda') > 0$ *such that if* $\|\alpha\|_{l^1 \mathrm{L}^\infty;\lambda} < \epsilon_0$ *then* (A.19) *has a smooth solution* β *which satisfies*

$$\|\beta\|_{l^1 \mathrm{L}^\infty;\lambda'} \lesssim \|\alpha\|_{l^1 \mathrm{L}^\infty;\lambda}.$$

By (A.19) *and Lemma A.4 we have the following for* $p \in [1, \infty]$ *(the RHS is not necessarily finite)*

$$\|\beta\|_{l^1 \mathrm{L}^p;\lambda'} \lesssim \|\alpha\|_{l^1 \mathrm{L}^p;\lambda' + \|\beta\|_{l^1 \mathrm{L}^\infty;\lambda'}}. \tag{A.20}$$

Proof. — The proof follows from a Picard iteration. It is important that we chose a norm which has an algebra property (A.16) and for which composition does not lose too much regularity; otherwise one would have to use a Newton iteration. Let $\beta_0(x) = \alpha(x)$ and inductively define

$$\beta_{k+1}(x) = \alpha\big(x + \beta_k(x)\big).$$

Hence,

$$\beta_{k+1}(x) - \beta_k(x) = \int_0^1 \mathrm{D}\alpha\big(x + s\beta_{k-1}(x) + (1-s)\beta_k(x)\big) \times \big(\beta_k(x) - \beta_{k-1}(x)\big) ds.$$

Therefore from (A.16) and Lemma A.4.

$$\|\beta_{k+1} - \beta_k\|_{l^1 \mathrm{L}^\infty;\lambda'} \leq \left(\sup_{s \in (0,1)} \|\mathrm{D}\alpha\|_{l^1 \mathrm{L}^\infty;\lambda' + s\|\beta_{k-1}\|_{l^1 \mathrm{L}^\infty;\lambda'} + (1-s)\|\beta_k\|_{l^1 \mathrm{L}^\infty;\lambda'}} \right) \times \|\beta_k - \beta_{k-1}\|_{l^1 \mathrm{L}^\infty;\lambda'}.$$

By induction, for ϵ_0 chosen sufficiently small the iteration converges and the lemma follows. □

Using Lemma A.5 we may prove Lemma 2.1 stated in Section 2.3.

Lemma 2.1. — From the results of [9, 33, 54], it follows that any initial data $\omega_{in} \in \mathcal{G}^{\mu,\sigma;\beta}$ for $\beta \in (0,1]$ which is mean zero and satisfies $\int |y\omega_{in}(x,y)| dxdy < \infty$ (to ensure finite kinetic energy as in (2.33)) will result in a unique global solution $\omega(t) \in \mathcal{G}^{\mu(t),\sigma;\beta}$ for some $\mu(t) > 0$ and all $t \in \mathbf{R}$ which remains mean zero and satisfies $\int |y\omega(t,x,y)| dxdy < \infty$ for all $t \in \mathbf{R}$. In particular, this holds for $\beta = s$ and $\mu = \mu(0) = \lambda_0$ as well as for analytic data, which shows that if we regularize ω_{in} to be analytic at time zero, it remains so for all time. Using the characteristics, we may also assert the spatial localization for any fixed $\bar{\epsilon} < \epsilon$ (possibly after reducing ϵ'),

$$\max_{t\in[0,1]} \int \left| y\omega(t,x,y) \right| dxdy \leq \bar{\epsilon},$$

as well as the a priori estimate on the kinetic energy of the zero frequency velocity field (recall this is not a conserved quantity):

$$\max_{t\in[0,1]} \left\| U_0^x(t) \right\|_2 \leq \bar{\epsilon}.$$

We next want to show a short-time estimate on $\omega(t)$ which has the mixing due to the Couette flow removed and then transfer this information to (z,v) coordinates. We first use the linear change of variables of (2.2) (although we will change the notation to not clash with the proof itself). Define $q(t,\bar{z},y) = \omega(t,\bar{z}+ty,y)$ and $\tilde{\phi}(t,\bar{z},y) = \psi(t,\bar{z}+ty,y)$, which satisfies (2.7):

$$\partial_t q + \nabla^{\perp}_{\bar{z},y}\tilde{\phi} \cdot \nabla_{\bar{z},y} q = 0 \tag{A.21a}$$

$$\Delta_L \tilde{\phi} = q. \tag{A.21b}$$

From (2.4), it is easy to verify for any $\mu(t)$ that

$$\left\| \nabla^{\perp} P_{\neq 0}\tilde{\phi} \right\|_{\mathcal{G}^{\mu,\sigma}} \lesssim t\|q\|_{\mathcal{G}^{\mu,\sigma}} \tag{A.22a}$$

$$\left\| \nabla^{\perp} P_{\neq 0}\tilde{\phi} \right\|_{\mathcal{G}^{\mu,\sigma-3}} \lesssim \langle t\rangle^{-2}\|q\|_{\mathcal{G}^{\mu,\sigma}}. \tag{A.22b}$$

Moreover, we have the same spatial localization

$$\max_{t\in[0,1]} \int \left| yq(t,\bar{z},y) \right| d\bar{z}dy \leq \bar{\epsilon}$$

and the associated a priori estimate on the kinetic energy in the shear flow

$$\max_{t\in[0,1]} \left\| \partial_y \tilde{\phi}_0(t) \right\|_2 = \max_{t\in[0,1]} \left\| U_0^x(t) \right\|_2 \leq \bar{\epsilon}. \tag{A.23}$$

Let $\mu(t)$ be chosen to satisfy

$$\dot{\mu}(t) = -\bar{\epsilon}^{1/2}\mu(t),$$

$$\mu(0) = 9\lambda_0/10 + \lambda'/10,$$

where we will choose $\bar{\epsilon}$ such that $\mu(1) > 8\lambda_0/9 + \lambda'/9$. Using the methodology of Section 5, in particular, Section 5.1 combined with the integration by parts trick discussed in (2.23) (see also e.g. [49, 54]), it follows from (A.22) and (A.23), for $\bar{\epsilon}$ chosen sufficiently small, that

$$\max_{t\in[0,1]} \left\| q(t) \right\|_{\mathcal{G}^{\mu(t),\sigma}} < 2\bar{\epsilon}.$$

Via the Biot-Savart law and an argument like that used in (2.33),

$$\max_{t\in[0,1]} \left\| U_0^x(t) \right\|_{\mathcal{G}^{\mu(t),\sigma}} \lesssim \bar{\epsilon}. \tag{A.24}$$

Next we now need to convert estimates on q into estimates on f and the associated nonlinear coordinate system. From (A.24) and (2.9), we have Gevrey control on $v = v(t, y)$ and $z = z(t, x, y)$. Notice that we can also write z in terms of $\bar{z}$ and y via

$$z(t, \bar{z}, y) = \bar{z} - t\big(v(t, y) - y\big), \tag{A.25}$$

and $f(t, z, v) = q(t, \bar{z}(t, z, v), y(t, v))$. Therefore, in order to control the Gevrey norm of f we need to solve for $\bar{z}, y$ in terms of z, v. From (2.9) we have (on $t \leq 1$),

$$\left\| v(t, y) - y \right\|_{\mathcal{G}^{\mu,\sigma}} \lesssim \bar{\epsilon}.$$

Hence, writing $\alpha(y) = y - v(t, y)$, $\beta(v) = y(t, v) - v$ and $\beta(v) = \alpha(v + \beta(v))$ we may apply (A.15) and Lemma A.5 (adjusting $\bar{\epsilon}$ if necessary) to solve for $y(t, v) - v$ with $\|y(t, v) - v\|_{\mathcal{G}^{\mu'}} \lesssim \bar{\epsilon}$, with $\mu' < \mu(1)$ and $3\lambda_0/4 + \lambda'/4 < \mu'$. In turn, from (2.9), this allows us to write $\|\bar{z}(t, z, v) - z\|_{\mathcal{G}^{\mu'}} \lesssim \bar{\epsilon}$. Then by (A.15) and Lemma A.4 we can deduce (for $\bar{\epsilon}$ sufficiently small)

$$\sup_{t\in(0,1)} \left\| f(t) \right\|_{3\lambda_0/4+\lambda'/4} < \epsilon$$

$$E(1) < \epsilon^2$$

$$\left\| 1 - v' \right\|_\infty < 6/10.$$

This completes the proof of Lemma 2.1. □

A.4 Rapid convergence of background flow

The proof of (1.7a) found in Section 2.4 follows from writing the x average of (1.2) in the (z, y) variables to (2.39) and integrating using that the priori estimates (2.32) provide decay estimates in the (z, y) variables. In fact, the derivation of (2.39) is capturing a subtle cancellation between the vorticity and U^y that originates from the structure of the linear problem. Consider (1.2) and take x averages of the first equation. Then one derives the following:

$$\partial_t \langle U^x \rangle + \langle U^y \partial_y U^x \rangle = 0.$$

We will consider the nonlinear term with U^y and U^x replaced by solutions to the linearized Euler equations. Since y derivatives are growing linearly, from a rough order of magnitude approximation one would expect that the nonlinear term decays like $O(t^{-2})$. In fact, we have $\langle U^y \partial_y U^x \rangle = -\langle U^y \omega \rangle$, so anything faster than a $O(t^{-2})$ decay indicates that something interesting is happening. First note from the Biot-Savart law:

$$\langle U^y \partial_y U^x \rangle = -\frac{1}{2\pi} \int \psi_x \psi_{yy} dx.$$

Writing ψ_{yy} on the Fourier side and using the (x, y) analogue of (1.4):

$$\widehat{\psi_{yy}}(t,k,\xi) = \frac{\xi^2 \widehat{\omega_{in}}(k, \xi + kt)}{k^2+\xi^2} = \left(|kt|^2 - 2kt(\xi+kt) + |\xi+kt|^2\right) \times \frac{\widehat{\omega_{in}}(k,\xi+kt)}{k^2+\xi^2}. \tag{A.26}$$

Therefore, while $P_{\neq}\phi_{yy}$ is not decaying or strongly converging to anything, we have the remarkable property that $\psi_{yy} = t^2 \psi_{xx} + O(t)$. Indeed, from (A.26) we have the following, since the leading order cancels due to the x average:

$$\begin{aligned}
&\left| \left(\widehat{\frac{1}{2\pi}\int \psi_x \psi_{yy} dx} \right)(t,\eta) \right| \\
&\quad = \left| \frac{i}{2\pi} \sum_{k\neq 0} \int_\xi k\left(|\xi+kt|^2 - 2kt(\xi+kt)\right) \right. \\
&\qquad \left. \times \frac{\omega_{in}(-k,\eta-\xi-kt)\omega_{in}(k,\xi+kt)}{(k^2+\xi^2)(k^2+|\eta-\xi|^2)} d\xi \right| \\
&\quad \lesssim t \sum_{k\neq 0} \int_\xi \left| \frac{\omega_{in}(-k,\eta-\xi-kt)\langle k,\xi+kt\rangle^3 \omega_{in}(k,\xi+kt)}{(k^2+\xi^2)(k^2+|\eta-\xi|^2)} \right| d\xi \\
&\quad \lesssim t \sum_{k\neq 0} \int_\xi \left| \frac{\omega_{in}(-k,\eta-\xi-kt)\langle k,\eta-\xi-kt\rangle^2 \langle k,\xi+kt\rangle^5 \omega_{in}(k,\xi+kt)}{(k^2+\xi^2)(k^2+|\eta-\xi|^2)\langle k,\eta-\xi-kt\rangle^2 \langle k,\xi+kt\rangle^2} \right| d\xi \\
&\quad \lesssim \frac{t}{\langle t\rangle^4} \sum_{k\neq 0} \int_\xi \langle -k,\eta-\xi-kt\rangle^2 \\
&\qquad \times \left| \omega_{in}(-k,\eta-\xi-kt)\langle k,\xi+kt\rangle^5 \omega_{in}(k,\xi+kt) \right| d\xi.
\end{aligned}$$

Then from (A.3) we have (without making an attempt to be optimal),

$$\left\| \langle U^y \partial_y U^x \rangle \right\|_2 = \left\| \frac{1}{2\pi} \int \psi_x \psi_{yy} dx \right\|_2 \lesssim \frac{1}{\langle t \rangle^3} \|\omega_{in}\|_{H^5}^2.$$

Since the nonlinear behavior matches the linear behavior to leading order, this indicates that indeed, (1.7a) should be expected on the nonlinear level.

REFERENCES

1. V. I. Arnold and B. A. Khesin, *Topological Methods in Hydrodynamics*, Applied Mathematical Sciences, vol. 125, Springer, New York, 1998.
2. J. Baggett, T. Driscoll and L. Trefethen, A mostly linear model of transition of turbulence, *Phys. Fluids*, **7** (1995), 833–838.
3. H. Bahouri and J.-Y. Chemin, Équations de transport relatives á des champs de vecteurs non-lipschitziens et mécanique des fluides, *Arch. Ration. Mech. Anal.*, **127** (1994), 159–181.
4. H. Bahouri, J.-Y. Chemin and R. Danchin, *Fourier Analysis and Nonlinear Partial Differential Equations*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 343, Springer, Heidelberg, 2011.
5. N. Balmforth and P. Morrison, Normal modes and continuous spectra, *Ann. N.Y. Acad. Sci.*, **773** (1995), 80–94.
6. N. Balmforth and P. Morrison, Singular eigenfunctions for shearing fluids I, *Institute for Fusion Studies Report, University of Texas-Austin*, **692** (1995), 1–80.
7. N. Balmforth, P. Morrison and J.-L. Thiffeault, *Pattern formation in Hamiltonian systems with continuous spectra; a normal-form single-wave model*, preprint (2013).
8. N. J. Balmforth and P. J. Morrison, Hamiltonian description of shear flow, in *Large-Scale Atmosphere-Ocean Dynamics*, vol. II, pp. 117–142, Cambridge Univ. Press, Cambridge, 2002.
9. C. Bardos and S. Benachour, Domaine d'analycité des solutions de l'équation d'Euler dans un ouvert de R^n, *Ann. Sc. Norm. Super. Pisa, Cl. Sci.*, **4** (1977), 647–687.
10. C. Bardos, Y. Guo and W. Strauss, Stable and unstable ideal plane flows, *Chin. Ann. Math., Ser. B*, **23** (2002), 149–164. Dedicated to the memory of Jacques-Louis Lions.
11. A. Bassom and A. Gilbert, The spiral wind-up of vorticity in an inviscid planar vortex, *J. Fluid Mech.*, **371** (1998), 109–140.
12. J. Bedrossian, N. Masmoudi and C. Mouhot, Landau damping: paraproducts and Gevrey regularity, arXiv:1311.2870, 2013.
13. J. Bedrossian, N. Masmoudi and V. Vicol, Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the 2D Couette flow, arXiv:1408.4754, 2014.
14. J. Bony, Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non lináires, *Ann. Sci. Éc. Norm. Super.*, **14** (1981), 209–246.
15. S. Bottin, O. Dauchot, F. Daviaud and P. Manneville, Experimental evidence of streamwise vortices as finite amplitude solutions in transitional plane Couette flow, *Phys. Fluids*, **10** (1998), 2597.
16. F. Bouchet and H. Morita, Large time behavior and asymptotic stability of the 2D Euler and linearized Euler equations, *Physica D*, **239** (2010), 948–966.
17. J. Boyd, The continuous spectrum of linear Couette flow with the beta effect, *J. Atmos. Sci.*, **40** (1983), 2304–2308.
18. R. Briggs, J. Daugherty and R. Levy, Role of Landau damping in crossed-field electron beams and inviscid shear flow, *Phys. Fluids*, **13** (1970).
19. E. Caglioti and C. Maffei, Time asymptotics for solutions of Vlasov-Poisson equation in a circle, *J. Stat. Phys.*, **92** (1998).
20. R. Camassa and C. Viotti, Transient dynamics by continuous-spectrum perturbations in stratified shear flows, *J. Fluid Mech.*, **717** (2013).
21. K. M. Case, Plasma oscillations, *Ann. Phys.*, **7** (1959), 349–364.
22. K. M. Case, Stability of inviscid plane Couette flow, *Phys. Fluids*, **3** (1960), 143–148.
23. A. Cerfon, J. Freidberg, F. Parra and T. Antaya, Analytic fluid theory of beam spiraling in high-intensity cyclotrons, *Phys. Rev. ST Accel. Beams*, **16** (2013), 024202.

24. J.-Y. Chemin, Le système de Navier-Stokes incompressible soixante dix ans après Jean Leray, in *Actes des Journées Mathématiques à la Mémoire de Jean Leray*, Sémin. Congr, vol. 9, pp. 99–123, Soc. Math. France, Paris, 2004.
25. J.-Y. Chemin, I. Gallagher and M. Paicu, Global regularity for some classes of large solutions to the Navier-Stokes equations, *Ann. Math. (2)*, **173** (2011), 983–1012.
26. J.-Y. Chemin and N. Masmoudi, About lifespan of regular solutions of equations related to viscoelastic fluids, *SIAM J. Math. Anal.*, **33** (2001), 84–112 (electronic).
27. P. Constantin, A. Kiselev, L. Ryzhik and A. Zlatoš, Diffusion and mixing in fluid flow, *Ann. Math. (2)*, **168** (2008), 643–674.
28. P. Degond, Spectral theory of the linearized Vlasov-Poisson equation, *Trans. Am. Math. Soc.*, **294** (1986), 435–453.
29. S. A. Denisov, Infinite superlinear growth of the gradient for the two-dimensional Euler equation, *Discrete Contin. Dyn. Syst.*, **23** (2009), 755–764.
30. P. G. Drazin and W. H. Reid, *Hydrodynamic Stability*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge University Press, Cambridge, 1981.
31. T. Ellingsen and E. Palm, Stability of linear flow, *Phys. Fluids*, **18** (1975), 487.
32. A. B. Ferrari and E. S. Titi, Gevrey regularity for nonlinear analytic parabolic equations, *Commun. Partial Differ. Equ.*, **23** (1998), 1–16.
33. C. Foias and R. Temam, Gevrey class regularity for solutions of the Navier-Stokes equations, *J. Funct. Anal.*, **87** (1989), 359–369.
34. S. Friedlander, W. Strauss and M. Vishik, Nonlinear instability in an ideal fluid, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, **14** (1997), 187–209.
35. D. Gérard-Varet and N. Masmoudi, *Well-posedness for the Prandtl system without analyticity or monotonicity*, preprint (2013).
36. P. Germain, N. Masmoudi and J. Shatah, Global solutions for the gravity water waves equation in dimension 3, *Ann. Math. (2)*, **175** (2012), 691–754.
37. M. Gevrey, Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire, *Ann. Sci. Éc. Norm. Super.*, **3** (1918), 129–190.
38. A. Gilbert, Spiral structures and spectra in two-dimensional turbulence, *J. Fluid Mech.*, **193** (1988), 475–497.
39. J. Ginibre and G. Velo, Long range scattering and modified wave operators for some Hartree type equations. I, *Rev. Math. Phys.*, **12** (2000), 361–429.
40. R. Glassey and J. Schaeffer, Time decay for solutions to the linearized Vlasov equation, *Transp. Theory Stat. Phys.*, **23** (1994), 411–453.
41. R. Glassey and J. Schaeffer, On time decay rates in landau damping, *Commun. Partial Differ. Equ.*, **20** (1995), 647–676.
42. N. Glatt-Holtz, V. Sverak and V. Vicol, On inviscid limits for the stochastic Navier-Stokes equations and related models, arXiv:1302.0542, 2013.
43. E. Grenier, On the nonlinear instability of Euler and Prandtl equations, *Commun. Pure Appl. Math.*, **53** (2000), 1067–1091.
44. Y. Guo and G. Rein, Isotropic steady states in galactic dynamics, *Commun. Math. Phys.*, **219** (2001), 607–629.
45. G. Hagstrom and P. Morrison, Caldeira-Leggett model, Landau damping and the Vlasov-Poisson system, *Physica D*, **240** (2011), 1652–1660.
46. H. Hwang and J. Velażquez, On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem, *Indiana Univ. Math. J.*, (2009), 2623–2660.
47. L. Kelvin, Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates, *Philos. Mag.*, **24** (1887), 188.
48. R. Kraichnan, Inertial ranges in two-dimensional turbulence, *Phys. Fluids*, **10** (1967).
49. I. Kukavica and V. Vicol, On the radius of analyticity of solutions to the three-dimensional Euler equations, *Proc. Am. Math. Soc.*, **137** (2009), 669–677.
50. S. B. Kuksin and A. Shirikyan, *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, Cambridge, 2012.
51. L. Landau, On the vibration of the electronic plasma, *J. Phys. USSR*, **10** (1946).
52. P. D. Lax and R. S. Phillips, *Scattering Theory*, vol. 26, Academic Press, San Diego, 1990.
53. M. Lemou, F. Méhats and P. Raphaël, Orbital stability of spherical galactic models, *Invent. Math.*, **187** (2012), 145–194.
54. D. Levermore and M. Oliver, Analyticity of solutions for a generalized Euler equation, *J. Differ. Equ.*, **133** (1997), 321–339.
55. Y. Li and Z. Lin, A resolution of the Sommerfeld paradox, *SIAM J. Math. Anal.*, **43** (2011), 1923–1954.

56. C. C. Lin, *The Theory of Hydrodynamic Stability*, Cambridge University Press, Cambridge, 1955.
57. Z. Lin, Nonlinear instability of ideal plane flows, *Int. Math. Res. Not.*, **41** (2004), 2147–2178.
58. Z. Lin and C. Zeng, Inviscid dynamic structures near Couette flow, *Arch. Ration. Mech. Anal.*, **200** (2011), 1075–1097.
59. Z. Lin and C. Zeng, Small BGK waves and nonlinear Landau damping, *Commun. Math. Phys.*, **306** (2011), 291–331.
60. H. Lindblad and I. Rodnianski, Global existence for the Einstein vacuum equations in wave coordinates, *Commun. Math. Phys.*, **256** (2005), 43–110.
61. R. Lindzen, Instability of plane parallel shear flow (toward a mechanistic picture of how it works), *PAGEOPH*, **126** (1988).
62. A. Lundbladh and A. V. Johansson, Direct simulation of turbulent spots in plane Couette flow, *J. Fluid Mech.*, **229** (1991), 499–516.
63. A. Majda, *Introduction to PDEs and Waves for the Atmosphere and Ocean*, Courant Lecture Notes in Mathematics, vol. 9, New York University Courant Institute of Mathematical Sciences, New York, 2003.
64. J. Malmberg and C. Wharton, Collisionless damping of electrostatic plasma waves, *Phys. Rev. Lett.*, **13** (1964), 184–186.
65. J. Malmberg, C. Wharton, C. Gould and T. O'Neil, Plasma wave echo, *Phys. Rev. Lett.*, **20** (1968), 95–97.
66. P. S. Marcus and W. H. Press, On Green's functions for small disturbances of plane Couette flow, *J. Fluid Mech.*, **79** (1977), 525–534.
67. N. Masmoudi and K. Nakanishi, Energy convergence for singular limits of Zakharov type systems, *Invent. Math.*, **172** (2008), 535–583.
68. P. J. Morrison, Hamiltonian description of the ideal fluid, *Rev. Mod. Phys.*, **70** (1998), 467–521.
69. P. J. Morrison, Hamiltonian description of Vlasov dynamics: action-angle variables for the continuous spectrum, *Transp. Theory Stat. Phys.*, **29** (2000), 397–414.
70. C. Mouhot and C. Villani, On Landau damping, *Acta Math.*, **207** (2011), 29–201.
71. K. Nakanishi, Modified wave operators for the Hartree equation with data, image and convergence in the same space, *Commun. Pure Appl. Anal.*, **1** (2002), 237–252.
72. L. Nirenberg, An abstract form of the nonlinear Cauchy-Kowalewski theorem, *J. Differ. Geom.*, **6** (1972), 561–576.
73. T. Nishida, A note on a theorem of Nirenberg, *J. Differ. Geom.*, **12** (1977), 629–633.
74. W. Orr, The stability or instability of steady motions of a perfect liquid and of a viscous liquid, Part I: a perfect liquid, *Proc. R. Ir. Acad., A Math. Phys. Sci.*, **27** (1907), 9–68.
75. S. A. Orszag and L. C. Kells, Transition to turbulence in plane Poiseuille and plane Couette flow, *J. Fluid Mech.*, **96** (1980), 159–205.
76. L. Rayleigh, On the stability, or instability, of certain fluid motions, *Proc. Lond. Math. Soc.*, **S1-11** (1880), 57.
77. S. C. Reddy, P. J. Schmid and D. S. Henningson, Pseudospectra of the Orr-Sommerfeld operator, *SIAM J. Appl. Math.*, **53** (1993), 15–47.
78. O. Reynolds, An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels, *Proc. R. Soc. Lond.*, **35** (1883), 84.
79. D. Ryutov, Landau damping: half a century with the great discovery, *Plasma Phys. Control. Fusion*, **41** (1999), A1.
80. D. Schecter, D. Dubin, A. Cass, C. Driscoll and I.L. et al., Inviscid damping of asymmetries on a two-dimensional vortex, *Phys. Fluids*, **12** (2000).
81. P. J. Schmid and D. S. Henningson, *Stability and Transition in Shear Flows*, Applied Mathematical Sciences., vol. 142, Springer, New York, 2001.
82. J. Shatah and C. Zeng, Geometry and a priori estimates for free boundary problems of the Euler equation, *Commun. Pure Appl. Math.*, **61** (2008), 698–744.
83. A. Shnirelman, *On the long time behavior of fluid flows*, preprint (2012).
84. S. Strogatz, R. Mirollow and P. Matthews, Coupled nonlinear oscillators below the synchronization threshold: relaxation by generalized Landau damping, *Phys. Rev. Lett.*, **68** (1992), 2730–2733.
85. J. Tataronis and W. Grossmann, Decay of MHD waves by phase mixing, *Z. Phys.*, **261** (1973), 203–216.
86. N. Tillmark and P. Alfredsson, Experiments on transition in plane Couette flow, *J. Fluid Mech.*, **235** (1992), 89–102.
87. L. Trefethen, A. Trefethen, S. Reddy and T. Driscoll, Hydrodynamic stability without eigenvalues, *Science*, **261** (1993), 578–584.
88. L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, 2005.
89. K. Tung, Initial-value problems for Rossby waves in a shear flow with critical level, *J. Fluid Mech.*, **133** (1983), 443–469.

90. N. van Kampen, On the theory of stationary waves in plasmas, *Physica*, **21** (1955), 949–963.
91. J. Vanneste, Nonlinear dynamics of anisotropic disturbances in plane Couette flow, *SIAM J. Appl. Math.*, **62** (2002), 924–944 (electronic).
92. J. Vanneste, P. Morrison and T. Warn, Strong echo effect and nonlinear transient growth in shear flows, *Phys. Fluids*, **10** (1998), 1398.
93. A. Yaglom, *Hydrodynamic Instability and Transition to Turbulence*, vol. 100, Springer, Berlin, 2012.
94. J. Yu and C. Driscoll, Diocotron wave echoes in a pure electron plasma, *IEEE Trans. Plasma Sci.*, **30** (2002).
95. J. Yu, C. Driscoll and T. O'Neil, Phase mixing and echoes in a pure electron plasma, *Phys. Plasmas*, **12** (2005), 055701.

J. B.
Courant Institute of Mathematical Sciences,
New York University,
New York,
USA
jacob@cims.nyu.edu

N. M.
Courant Institute of Mathematical Sciences,
New York University,
New York,
USA
masmoudi@cims.nyu.edu

Manuscrit reçu le 26 février 2014
Manuscrit accepté le 9 février 2015
publié en ligne le 25 février 2015.