

# Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the 2D Couette flow

Jacob Bedrossian\* and Nader Masmoudi† and Vlad Vicol‡

August 20, 2014

## Abstract

In this work we study the long time, inviscid limit of the 2D Navier-Stokes equations near the periodic Couette flow, and in particular, we confirm at the nonlinear level the qualitative behavior predicted by Kelvin's 1887 linear analysis. At high Reynolds number  $\mathbf{Re}$ , we prove that the solution behaves qualitatively like 2D Euler for times  $t \lesssim \mathbf{Re}^{1/3}$ , and in particular exhibits inviscid damping (e.g. the vorticity weakly approaches a shear flow). For times  $t \gtrsim \mathbf{Re}^{1/3}$ , which is sooner than the natural dissipative time scale  $O(\mathbf{Re})$ , the viscosity becomes dominant and the streamwise dependence of the vorticity is rapidly eliminated by an enhanced dissipation effect. Afterward, the remaining shear flow decays on very long time scales  $t \gtrsim \mathbf{Re}$  back to the Couette flow. When properly defined, the dissipative length-scale in this setting is  $\ell_D \sim \mathbf{Re}^{-1/3}$ , larger than the scale  $\ell_D \sim \mathbf{Re}^{-1/2}$  predicted in classical Batchelor-Kraichnan 2D turbulence theory. The class of initial data we study is the sum of a sufficiently smooth function and a small (with respect to  $\mathbf{Re}^{-1}$ )  $L^2$  function.

## Contents

1	Introduction	2
2	Proof of Theorem 1	9
3	Properties of the norms	23
4	Elliptic estimates	29
5	High norm vorticity estimate (2.66a)	30
6	Enhanced dissipation estimate (2.67)	39
7	Coordinate system higher regularity controls	44
8	Low frequency estimates	50
A	Appendix	54

---

\**jacob@cscamm.umd.edu*, University of Maryland, College Park

†*masmoudi@cims.nyu.edu*, Courant Institute of Mathematical Sciences

‡*vvicol@math.princeton.edu*, Princeton University

# 1 Introduction

Consider the 2D incompressible Navier-Stokes equations in the vicinity of the Couette flow:

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

Here,  $t \in \mathbb{R}_+$ ,  $(x, y) \in \mathbb{T} \times \mathbb{R}$ ,  $\nabla^\perp = (-\partial_y, \partial_x)$ ,  $U = (U^x, U^y)$  is the velocity, and  $\omega$  is the vorticity of the *perturbation* to the Couette flow. The physical velocity is  $(y, 0) + U$  and the total vorticity of the flow is  $-1 + \omega$ . We denote the streamfunction of the perturbation by  $\psi = \Delta^{-1}\omega$ . We have already non-dimensionalized (1.1) by normalizing the Couette flow to  $(y, 0)$  and the side-length of the torus to  $2\pi$ , which together with the kinematic viscosity of the fluid, sets the Reynolds number  $\mathbf{Re}$ . We use the notation  $\nu = \mathbf{Re}^{-1}$ . If  $\nu = 0$ , equivalently  $\mathbf{Re} = \infty$ , the fluid is called *inviscid* and (1.1) corresponds to the incompressible 2D Euler equations. The velocity satisfies the momentum equation:

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

where  $P$  denotes the pressure. As we are interested in the behavior of solutions to (1.1) for high Reynolds number, without loss of generality we assume throughout the paper that  $\nu \in (0, 1]$ .

The study of (1.1) for small perturbations is an old problem in hydrodynamic stability, considered by both Rayleigh [47] and Kelvin [35], as well as by many modern authors with new perspectives (see e.g. the classical texts [24, 55] and the references therein). Rayleigh and Kelvin both studied the linearization of (1.1), which is simply

$$\begin{cases} \partial_t\omega + y\partial_x\omega = \nu\Delta\omega, \\ \Delta\psi = \omega, \\ \omega(0) = \omega_{in}. \end{cases} \quad (1.3)$$

Rayleigh first proved spectral stability (the absence of unstable eigenmodes) and Kelvin went further to explicitly solve (1.3). Indeed, if we denote by  $\hat{\omega}(t, k, \eta)$  the Fourier transform of  $\omega(t, x, y)$  (note  $k \in \mathbb{Z}$ ,  $\eta \in \mathbb{R}$ ), then it is straightforward to verify Kelvin's solution

$$\hat{\omega}(t, k, \eta) = \hat{\omega}_{in}(k, \eta + kt) \exp \left[ -\nu \int_0^t |k|^2 + |\eta - k\tau + kt|^2 d\tau \right], \quad (1.4a)$$

$$\hat{\psi}(t, k, \eta) = -\frac{\hat{\omega}_{in}(k, \eta + kt)}{k^2 + \eta^2} \exp \left[ -\nu \int_0^t |k|^2 + |\eta - k\tau + kt|^2 d\tau \right]. \quad (1.4b)$$

From (1.4) we see directly that if  $k \neq 0$  (that is, the mode depends on  $x$ ), the mode is strongly damped by the viscosity by the time  $t \gtrsim \nu^{-1/3}$ , whereas for  $k = 0$  the evolution is that of the 1D heat equation with a slower  $t \gtrsim \nu^{-1}$  dissipative time-scale; hence the  $k \neq 0$  modes experience *enhanced dissipation*. For shorter times, the behavior of (1.4) essentially matches that of the inviscid problem, studied in detail by Orr in [46]. One of Orr's key observations was the *inviscid damping* predicted by (1.4b), that is, even for  $\nu = 0$ , (1.4b) shows that if the vorticity is sufficiently regular, then the velocity field  $U = \nabla^\perp\psi$  will return to a shear flow at an algebraic rate. Both inviscid damping and enhanced dissipation are manifestations of the same *vorticity mixing* induced by the background shear flow, which gives rise to a linear-in-time (for each  $k \neq 0$ ) transfer of enstrophy to high frequencies (as seen in (1.4)). Together, the effects can be summarized by the following.

**Proposition 1.1** (Linearized behavior (Kelvin [35] and Orr [46]) ). *Let  $\omega(t)$  be a solution to (1.3) with initial data  $\omega_{in} \in H^\sigma$  for any  $\sigma \geq 3$  such that  $\int \omega_{in} dx dy = 0$  and  $\int |y \omega_{in}| dy dx < \infty$ . Then from (1.4) we have for some  $c > 0$  (writing also  $(U^x, U^y) = \nabla^\perp \psi$ )*

$$\|P_{\neq 0} \omega(t, x + ty, y)\|_{H^\sigma} \lesssim \|\omega_{in}\|_{H^\sigma} e^{-c\nu t^3} \quad (1.5a)$$

$$\|P_0 \omega(t, x, y)\|_{H^\sigma} \lesssim \frac{\|\omega_{in}\|_{H^\sigma} + \|\omega_{in}\|_1}{\langle \nu t \rangle^{1/4}} \quad (1.5b)$$

$$\|P_{\neq 0} U^x(t)\|_2 + \langle t \rangle \|U^y(t)\|_2 \lesssim \frac{\|\omega_{in}\|_{H^\sigma} e^{-c\nu t^3}}{\langle t \rangle}. \quad (1.5c)$$

where all implicit constants are independent of  $\nu$  and  $t$  and we are defining the orthogonal projections to  $x$ -independent modes  $P_0 g = \frac{1}{2\pi} \int g(x, y) dx$  and  $x$ -dependent modes  $P_{\neq 0} g = g - \frac{1}{2\pi} \int g(x, y) dx$ .

There are several details of Proposition 1.1 that warrant attention. First, is the requirement of three derivatives, which is necessary to deduce (1.5c). Indeed, in order to get the inviscid algebraic decay from (1.4b), one requires localization of the Fourier transform of the initial data. Physically, this is to control the amount of information which is being *un-mixed* by the Couette flow (see below and [9] for further discussion). Secondly, notice that the regularity estimates in (1.5a) are only after pulling back by the Couette flow characteristics, another evidence of the mixing and a common theme in [45, 9, 10]. Note further that this implies that in general, if one does not pull back by the shear flow we have  $\|\omega(t)\|_{H^\sigma} \approx \langle t \rangle^\sigma$  for  $\nu t^3 \ll 1$ . Thirdly, notice that the  $x$ -dependent contribution to the vorticity decays rapidly in (1.5a) after  $\nu t^3 \gg 1$  whereas the  $x$ -independent contribution takes much longer to decay. Finally, the reason that  $c < 1/3$  in the above theorem is due to the fact that the enhanced dissipation is anisotropic in time and frequency, as can be seen in the degeneracy of the exponentials in (1.4). Like the loss of regularity, this is also due to a transient un-mixing effect.

In our work we study the long time, inviscid limit of ‘small’ solutions to (1.1) and seek to find the appropriate nonlinear analogue of Proposition 1.1. The primary motivation of our work is to study the enhanced dissipation by mixing exhibited in (1.4). A second, deeply related, motivation for our work is to affirm the physical relevance of the inviscid ( $\nu = 0$ ) work completed by two of the authors in [9], where inviscid damping is exhibited in the Euler equations.

The enhanced dissipation effect is sometimes referred to by modern authors as the ‘shear-diffusion mechanism’ and has been explicitly or implicitly pointed out by numerous authors studying shear flows [48, 40, 20, 8], vortex axisymmetrization [12, 28, 29, 2] and 3D strained vortex filaments [43]. The work of [21] considers in a more general framework the connection between the spectral properties of the background flow and the enhanced dissipation effect in the context of passive tracers using the RAGE theorem (see also [11] for the elliptic case). It is pointed out in [8] that the effect can be linked to a form of ‘hypocoercivity’ [54]. Moreover, it is related to Taylor dispersion [51], as discussed in e.g. [48], which is not present in our work due to the periodic boundary conditions. Certain aspects of anomalous diffusions due to underlying drift have been studied in the context of SDEs; see for example [34, 56, 31, 19].

At high Reynolds number, the linear solution (1.4) exhibits a constant flux of enstrophy to high frequencies, independent of the Reynolds number, for times  $t \lesssim \mathbf{Re}^{1/3}$ . In [9] it is shown that this enstrophy flux persists in the Euler equations. Here, we show that this enstrophy flux is stable in the inviscid limit, and moreover gives rise to the enhanced dissipation. It is suggested in [43, 28, 29, 6], that the inviscid limit of this enhanced dissipation effect is important for understanding the fine scale features of turbulent flows predicted in Kolmogorov [36] and Batchelor/Kraichnan [7, 37] theories. However, notice that the constant enstrophy flux in (1.4) is *more efficient* than a *generic turbulent flow* at moving enstrophy to small scales (where it is annihilated by the diffusion). Indeed, one

can see this by considering the associated dissipative length scale  $\ell_D$ . Since the enstrophy transfer is linear in time, we see that in fact  $\ell_D \sim \mathbf{Re}^{-1/3}$  which is a much larger scale than the  $\mathbf{Re}^{-1/2}$  predicted by the Kraichnan statistical theory. Partly, this is due to the fact that the enstrophy transfer considered here is highly anisotropic, and so should not necessarily be in the same scaling regime as homogeneous, isotropic turbulence.

It is well known that studying the inviscid limit  $\nu \rightarrow 0$  and setting  $\nu = 0$  in (1.1) is not the same problem. Theorem 1 below (and [9]) implies that the 2D Euler equations qualitatively predict the correct behavior of (1.1) for times  $t \lesssim \mathbf{Re}^{1/3}$ , at which point they cease to do so. This time interval is much longer than what follows from brute force energy estimates, but still much shorter than for completely ‘laminar’ flows, which would behave qualitatively like 2D Euler until  $t \sim \mathbf{Re}$ . We emphasize that, in general, results which hold for the Euler evolution need not necessarily hold for the Navier-Stokes equations at high Reynolds number. For instance, in the presence of boundaries it was shown in [30] that a linearly spectrally stable Euler flow becomes unstable upon the addition of dissipation at large Reynolds number (see also [24]). Also, the recent constructions of non-unique energy dissipative solutions of the 3D Euler equations [22, 23, 33, 17] are not known to have an analogue in the presence of dissipation [5], with or without boundaries.

One of the consequences of our work is that (1.1) will exhibit inviscid damping for  $t \lesssim \mathbf{Re}^{1/3}$ . See [9] and the references therein for an in-depth discussion of the phenomenon. The term ‘inviscid damping’ appeared after it was noticed to be the hydrodynamic analogue of *Landau damping* [39, 49, 18, 32, 45, 10] in the kinetic description of plasmas which are sufficiently collisionless (the analogue of inviscid) to be well-described by the Vlasov equations. See e.g. [14, 50, 16, 3, 45, 4, 9] and the references therein for some discussion about the connections between inviscid damping and Landau damping and the more general concept of phase mixing.

Orr himself pointed out the main subtlety inherent with inviscid damping and Landau damping: in (1.4), vorticity can just as easily unmix to large scales and create growth in the velocity field as it can mix to small scales and create decay; together the two effects are known as the *Orr mechanism*. Indeed, if one takes initial data concentrated near frequencies  $\eta, k > 0$  with  $\eta$  large relative to  $k$ , one can observe from (1.4) that the stream-function reaches a maximum amplitude near the *critical time*  $t \sim \eta/k$ . Even on the linear level, this implies that one needs to *pay regularity* to deduce decay of the velocity, in order to control the amount of information that is being unmixed. See [15, 42, 9] for more discussion of the Orr mechanism. For similar reasons, the enhanced dissipation slows down near the critical times as the enstrophy passes through the  $O(1)$  length scales. Hence, it is a serious over-simplification to imagine the enhanced dissipation to be equivalent to replacing the  $\nu$  in front of the  $\partial_{xx}$  on the RHS of (1.1) with  $\nu^{1/3}$ ; it is acting differently for each frequency at each time. Our methods will deal with this effect by designing a semi-norm to measure the solution which respects the frequency-by-frequency anisotropy and allows to gradually pay regularity for enhanced dissipation of the vorticity; see §2.

The Orr mechanism is known to interact poorly with the nonlinear term, creating a weakly nonlinear effect referred to as an *echo*. In nearly collisionless plasmas, echoes were captured experimentally in [44] and in 2D Euler much later in [57, 58] (see also [53, 52]). The echoes can potentially chain into a cascade and greatly amplify the regularity loss already present in the linear theory (see e.g. [45, 9]), which is the reason that the nonlinear results in [18, 32, 45, 9, 10] all require at least Gevrey class regularity [27] (the work of [25] is the exception, as the model studied therein does not support infinite echo cascades). For (1.1), we will not need the data to be Gevrey class in a qualitative sense (which would be somewhat non-physical for a model involving a dissipative effect), however, the high regularity requirement will come as a quantitative requirement that the data be  $L^2$  close to Gevrey class (see Theorem 1 below). Specifically, Gevrey- $\frac{1}{s}$  with  $s \in (1/2, 1)$  data with a distance in  $L^2$  that diminishes as a function of Reynolds number, eventually collapsing

to a neighborhood in Gevrey- $\frac{1}{s}$  in the inviscid limit (consistent with [9]). In view of the regularity requirement needed to obtain inviscid damping in 2D Euler, the type of initial data we consider is natural as the  $L^2$  piece must be small enough to have negligible influence above the dissipative length scale (although it is not clear whether or not the diameter of the  $L^2$  neighborhood in (1.6) is optimal). The Gevrey- $\frac{1}{s}$  norm is given by

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

where  $\hat{f}_k(\eta)$  is the Fourier transform of  $f$  at frequency  $(k, \eta)$ .

## 1.1 Statement of Theorem

We first state the main result, and then follow the statement by remarks, a discussion and a brief summary of the proof.

**Theorem 1.** *For all  $s \in (1/2, 1)$ ,  $\lambda > \lambda' > 0$ ,  $\delta > 0$  and all integers  $\alpha \geq 1$  there exists  $\epsilon_0 = \epsilon_0(\alpha, s, \lambda, \lambda', \delta)$  and  $K_0 = K_0(s, \lambda, \lambda', \delta)$  such that if  $\omega_{in}^\nu$  is mean-zero and satisfies  $\omega_{in}^\nu = \omega_{S,in}^\nu + \omega_{R,in}^\nu$  with*

$$\| |\nabla|^{-1} \omega_{in}^\nu \|_2 + \|\omega_{S,in}^\nu\|_{\mathcal{G}^{\lambda;s}} + e^{K_0 \nu^{-\frac{3(1+\delta)s}{2(1-s)}}} \|\omega_{R,in}^\nu\|_2 = \epsilon \leq \epsilon_0, \quad (1.6)$$

*then for all  $\nu$  sufficiently small (independent of  $\epsilon$  and  $\epsilon_0$ ) the solution  $\omega^\nu(t)$  with initial data  $\omega_{in}^\nu$  satisfies the following properties (with all implicit constants independent of  $\nu$ ,  $t$  and  $\epsilon$ ),*

(i) *the uniform (in  $t$  and  $\nu$ ) estimate*

$$\|\omega^\nu(t, x + ty + \Phi(t, y), y)\|_{\mathcal{G}^{\lambda';s}} \lesssim \epsilon \quad (1.7)$$

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

$$\Phi(t, y) = \int_0^t e^{\nu(t-\tau)\partial_{yy}} \left( \frac{1}{2\pi} \int_{\mathbb{T}} U^x(\tau, x, y) dx \right) d\tau; \quad (1.8)$$

(ii) *inviscid damping on inviscid time-scales: for  $t \lesssim \nu^{-1/3}$ ,*

$$\left\| \frac{d}{dt} [\omega^\nu(t, x + ty + \Phi(t, y), y)] \right\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon^2}{\langle t \rangle^2} + \epsilon \nu \langle t \rangle^2, \quad (1.9a)$$

$$\|P_{\neq 0} U^x(t)\|_2 + \langle t \rangle \|U^y(t)\|_2 \lesssim \frac{\epsilon}{\langle t \rangle}, \quad (1.9b)$$

$$\left\| \frac{d}{dt} P_0 U^x(t) \right\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon^2}{\langle t \rangle^3} + \epsilon \nu; \quad (1.9c)$$

(iii) *enhanced dissipation for  $x$ -dependent modes on the fast viscous time-scale: for  $t \gtrsim \nu^{-1/3}$ ,*

$$\|P_{\neq 0} \omega^\nu(t, x + ty + \Phi(t, y), y)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t^3 \rangle^\alpha}, \quad (1.10a)$$

$$\|P_{\neq 0} U^x(t)\|_2 + \langle t \rangle \|U^y(t)\|_2 \lesssim \frac{\epsilon}{\langle t \rangle \langle \nu t^3 \rangle^\alpha}; \quad (1.10b)$$

(iv) and standard viscous decay for  $x$ -independent modes on the slow viscous time-scale: for  $t \gtrsim \nu^{-1}$ ,

$$\|P_0\omega^\nu(t)\|_2 \lesssim \frac{\epsilon}{\langle \nu t \rangle^{1/4}}. \quad (1.11)$$

**Remark 1.** It is useful to compare Theorem 1 with the linear behavior in Proposition 1.1. The qualitative behavior predicted by the nonlinear theorem is essentially the same except for two details: the algebraic rate of the enhanced dissipation and the correction to the shear flow characteristics  $\Phi(t, y)$ . The correction  $\Phi$  is to account for the fact that the perturbation induces a small shear,  $\langle U^x(t) \rangle$ , which varies in time and decays at a non-integrable rate on the very long  $O(\mathbf{Re})$  time scale (in particular, in the  $\nu = 0$  case the flow never returns to Couette flow). Therefore, gradients would grow without bound in (1.7) if we did not account for this. The algebraic rate of decay arises since, due to un-mixing, we pay regularity in return for enhanced dissipation. The current proof pays  $3\alpha$  derivatives; with some additional technical effort it seems we could trade Gevrey-2 regularity (or Gevrey- $\frac{1}{\beta}$  for any  $\beta < s$ ) and likely get a decay like  $\exp[-(\nu t^3)^{1/6}]$ . However, getting the decay seen on the linear level seems very difficult with the current method (we are unsure about whether or not we can expect the same linear decay rate from the nonlinear solution).

**Remark 2.** Inequality (1.11) is also true with the  $L^2$  norm replaced by  $\mathcal{G}^{\lambda'; s}$  with a very similar proof.

**Remark 3.** For  $\nu t^3 \ll 1$ , statement (ii) essentially reduces to the main results of [9].

**Remark 4.** We note that if  $\omega_{R, in}^\nu$  were in  $H^\sigma$  with  $\sigma > 5/2$ , instead of merely  $L^2$ , then one could take  $\delta = 0$  in the hypothesis (1.6), which is consistent with the instant regularization available from the heat equation.

Schematics of the qualitative behavior predicted by Theorem 1 are given in Figures 1 and 2. As Theorem 1 predicts essentially the same behavior as the main Theorem in [9] before the enhanced dissipation is activated, it is natural that the proof of Theorem 1 also contains the entire proof of the main theorem in [9]. To avoid being redundant, we simply quote without proof estimates that are proved in [9] and *concentrate only on the steps that are specific to Navier-Stokes*. The proof of Theorem 1 is outlined in §2, however let us quickly highlight the main ideas and the main challenges.

- The coordinate system used in [9] needs to be adapted to account for the viscous dissipation in the background shear flow (see (2.36) below). In the Lagrangian view, we are accounting for the stochastic wandering of particle trajectories in the  $y$  variable. In the Eulerian view, we are accounting for the momentum transport induced by viscosity. Note that this is physically relevant even for short times, as the viscous dissipation always dominates for high enough frequencies. Like the choice of coordinates in [9], the choice used here is determined essentially uniquely by the analytic properties we require.
- As can be seen from (1.4), the enhanced dissipation effect is highly anisotropic in time and frequency: it is only active for  $x$ -dependent frequencies and slows down near the critical time  $t \sim \eta/k$ . To correctly capture the enhanced dissipation we use a special semi-norm that is

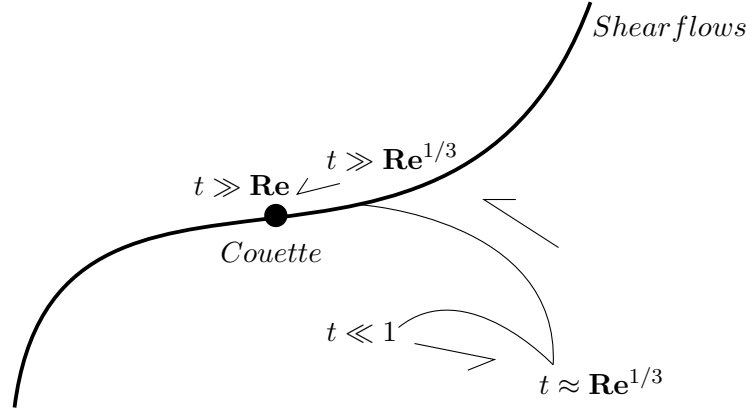


Figure 1: This is a schematic of the qualitative behavior of  $\omega^\nu(t)$  in a strong topology (e.g.  $H^\sigma$  for  $\sigma > 0$ ) as time evolves. The solution rapidly regularizes due to the viscosity, together with the assumption (1.6), and near time zero lies in a neighborhood of the Couette flow. For  $t \ll \mathbf{Re}^{1/3}$ , inviscid dynamics dominate and there is a transient growth of strong norms as the vorticity mixes (e.g. the  $H^\sigma$  norm will generally grow as  $\langle t \rangle^\sigma$ ). By times  $t \gg \mathbf{Re}^{1/3}$ , the dissipation dominates and the solution rapidly converges to a shear flow (but not necessarily the Couette flow). Finally, for times  $t \gg \mathbf{Re}$ , the remaining shear flow relaxes to the Couette flow.

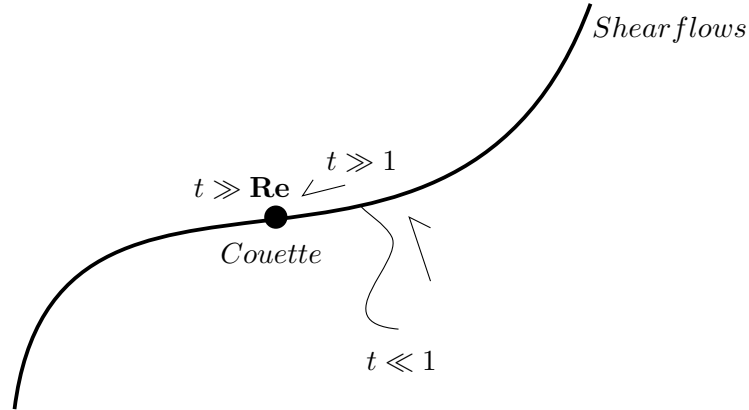


Figure 2: This is a schematic of the qualitative behavior of  $\omega^\nu(t)$  in the weak topology as time evolves. The solution begins in a neighborhood of the Couette flow and rapidly approaches a shear flow in the weak topology via inviscid damping on time scales independent of  $\mathbf{Re}$ . There is relatively little change until times  $t \gg \mathbf{Re}$  when finally viscous dissipation relaxes the shear flow back to the Couette flow.

well adapted to the anisotropy (defined in (2.55b) and analyzed in §3.2). The anisotropy also creates a huge imbalance between  $x$ -dependent modes and  $x$ -independent modes, and an important part of the analysis in §6 is to prove that very little enstrophy from the  $x$ -independent modes can be converted to  $x$ -dependent enstrophy at long times.

- The evolving background shear flow determines the rate at which dissipation occurs locally in  $y$ , a fully nonlinear effect as the shear flow is determined by the solution. Controlling this effect will be the main challenge to proving Theorem 1 and requires a careful analysis of the interplay between the regularity and decay of the background shear flow with the regularity and decay of the  $x$ -dependent modes (indeed, the structure of the proof is essentially dictated by this balance; see §2).

## 1.2 Notation and conventions

See §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  $\eta$  and  $\xi$  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 dyadic integers by  $M, N \in \mathbb{D}$  or  $2^{\mathbb{Z}}$  where

$$2^{\mathbb{Z}} = \left\{ \dots, 2^{-j}, \dots, \frac{1}{4}, \frac{1}{2}, 1, 2, \dots, 2^j, \dots \right\},$$

$$\mathbb{D} = \left\{ \frac{1}{2}, 1, 2, \dots, 2^j, \dots \right\} = 2^{\mathbb{D}} \cup \frac{1}{2}.$$

This will be useful when defining Littlewood-Paley projections and paraproduct decompositions, see §A.1. Given a function  $m \in L^\infty$ , we define the Fourier multiplier  $m(\nabla)f$  by

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

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 define  $f \gtrsim g$ ). 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  $\alpha$ . 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  $\mathbb{R}^n$  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  $\|f\|_p$ . We make common use of the Gevrey- $\frac{1}{s}$  norm with Sobolev correction defined by

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

Since most of the paper we are taking  $s$  as a fixed constant, it is normally omitted. Also, if  $\sigma = 0$ , it is omitted. We refer to this norm as the  $\mathcal{G}^{\lambda, \sigma; s}$  norm and occasionally refer to the space of functions

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



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

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

## 2 Proof of Theorem 1

### 2.1 Instant regularization and $O(1)$ times

A physical aspect of our work is to verify that the instant regularization due to the viscosity replaces the qualitative regularity requirements in [9] with the more physically natural quantitative regularity requirement (1.6). Indeed, we have the following result:

**Proposition 2.1** (Instant regularization). *For all  $\epsilon > 0$ ,  $\lambda > \lambda' > 0$ ,  $\delta > 0$ , and  $s \in (1/2, 1)$ , there exists an  $\epsilon' = \epsilon'(s, \lambda, \lambda', \delta)$ , a  $K_0 = K_0(s, \lambda, \lambda', \delta)$ , and a  $\nu_0 = \nu_0(s, \lambda, \lambda', \delta)$  such that if  $\omega_{in}^\nu$  is mean-zero and can be written as the sum of a smooth part and a rough part  $\omega_{in}^\nu = \omega_{S,in}^\nu + \omega_{R,in}^\nu$  with*

$$\| |\nabla|^{-1} \omega_{in}^\nu \|_2 + \|\omega_{S,in}^\nu\|_{\mathcal{G}^{\lambda,s}} + e^{K_0 \nu^{-\frac{(3+\delta)s}{2(1-s)}}} \|\omega_{R,in}^\nu\|_2 \leq \epsilon', \quad (2.1)$$

*then for all  $\nu \leq \nu_0$  the solution  $\omega^\nu$  to 2D NSE near Couette flow (1.1), with initial datum  $\omega_{in}^\nu$  satisfies*

$$\|u^\nu(T_{\lambda,\lambda'})\|_2 + \|\omega^\nu(T_{\lambda,\lambda'})\|_{\mathcal{G}^{\frac{3\lambda+\lambda'}{4},s}} \leq \epsilon \quad (2.2)$$

where

$$T_{\lambda,\lambda'} = \min \left\{ \frac{\lambda - \lambda'}{C_0}, 1 \right\}, \quad (2.3)$$

for some sufficiently large universal constant  $C_0 > 0$ .

*Proof.* The proof consists in four steps:

- (a) We propagate the low-frequency control of the solution, which is encoded in the  $L^2$  energy estimate on  $u^\nu$ .
- (b) We solve the Navier-Stokes equations perturbed about the Couette flow (1.1) with initial datum  $\omega_{S,in}^\nu$  and establish that the solution  $\omega_S^\nu(t)$  lies in  $\mathcal{G}^{\frac{4\lambda+\lambda'}{5},s} \cap \mathcal{G}^{\sqrt{\nu}t,1}$  for all  $t \leq T_2 \lesssim \nu^{1+\delta}$ , with a norm less than  $2\epsilon'$ .
- (c) We solve the perturbation of (1.1) around the solution  $\omega_S^\nu$ , with initial datum  $\omega_{R,in}^\nu$  and show that the corresponding solution lies in  $\mathcal{G}^{\sqrt{\nu}t,1}$  for all  $t \leq T_2$ , with a suitable control on the norm. In view of the smallness of the initial datum  $\omega_{R,in}^\nu$  in  $L^2$  it follows that the  $\mathcal{G}^{\frac{4\lambda+\lambda'}{5},s}$  norm of  $\omega_R^\nu$  is also less than  $2\epsilon'$ .

- (d) We propagate the Gevrey-class regularity of the solution to (1.1) with initial datum  $\omega_S^\nu(T_2) + \omega_R^\nu(T_2) \in \mathcal{G}^{\frac{4\lambda+\lambda'}{5};s}$  of size  $\leq 4\epsilon'$  and show that the corresponding solution  $\omega^\nu$  at time  $T_{\lambda,\lambda'}$  obeys the bound (2.2).

For the remainder of the proof, for ease of notation we drop the  $\nu$  upper index for  $\omega$  and  $u$ . Without loss of generality we may assume the “rough” part of the initial datum  $\omega_{R,in}$  is supported on frequencies away from the set  $\{k = 0, |\eta| \leq 1\}$ . Further, we define  $u_S(t) = \nabla^\perp \Delta^{-1} \omega_S(t)$  and  $u_R(t) = \nabla^\perp \Delta^{-1} \omega_R(t)$ , the velocities associated with the smooth and rough solutions (both in  $L^2$  initially due to the assumptions on the initial data and the frequency support of  $\omega_{R,in}$ ).

*Proof of (a).* The standard  $L^2$  energy estimate for  $u_S$  obtained from (1.2) yields

$$\frac{1}{2} \frac{d}{dt} \|u_S\|_2^2 + \nu \|\omega_S\|_2^2 \leq \left| \int u_S^x u_S^y \right| \leq \|P_{\neq 0} u_S\|_2^2$$

and therefore

$$\|u_S(t)\|_2 \leq 3 \|u_{S,in}\|_2 \leq 3(\|\nabla|^{-1} \omega_{in}\|_2 + \|\omega_{S,in}\|_2) \leq 6\epsilon' \quad (2.4)$$

for all  $t \leq 1$ . Note that the same bound for the full solution  $u$  holds on  $[0, 1]$ .

*Proof of (b).* Consider the analytic Fourier multiplier

$$M_A(t, k, \eta) = \langle k, \eta \rangle^\sigma e^{\sqrt{\nu} t |k, \eta|} \quad (2.5)$$

for some  $\sigma > 5/2$ . As in [26], the solution  $\omega_S$  of (1.1) with initial datum  $\omega_{S,in}$  obeys the a priori estimate

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|M_A \omega_S\|_2^2 + \nu \||\nabla| M_A \omega_S\|_2^2 \\ &= \sqrt{\nu} \||\nabla|^{1/2} M_A \omega_S\|_2^2 + \sum_k \int_\eta M_A(k, \eta)^2 \partial_\eta \hat{\omega}_S(k, \eta) k \overline{\hat{\omega}_S(k, \eta)} d\eta \\ &+ \sum_{k, l} \int_\eta \int_\xi M_A(k, \eta) \overline{\hat{\omega}_S(k, \eta)} M_A(l, \xi) \hat{\omega}_S(l, \xi) M_A(k-l, \eta-\xi) \hat{\omega}_S(k-l, \eta-\xi) \\ &\quad \times \frac{(l, \xi)^\perp \cdot (k, \eta)}{|l, \xi|^2} \frac{M_A(k, \eta)}{M_A(l, \xi) M_A(k-l, \eta-\xi)} d\xi d\eta \\ &= \sqrt{\nu} \||\nabla|^{1/2} M_A \omega_S\|_2^2 + T_A + N_A \end{aligned} \quad (2.6)$$

where  $T_A$  is the shear transport term and  $N_A$  is the nonlinear term. Upon integrating by parts in  $\eta$  and bounding  $|\partial_\eta M_A(k, \eta)|$  we immediately obtain

$$|T_A| \lesssim \sqrt{\nu} t \||\nabla|^{1/2} M_A \omega_S^\nu\|_2^2. \quad (2.7)$$

In order to bound the nonlinear term  $N_A$  we use the triangle inequality

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

followed by (A.5), which implies

$$\begin{aligned}
|N_A| &\leq \sum_{k,l} \int_{\eta} \int_{\xi} |k, \eta| M_A(k, \eta) |\hat{\omega}_S(k, \eta)| M_A(k-l, \eta-\xi) |\hat{\omega}_S(k-l, \eta-\xi)| \\
&\quad \times |l, \xi|^{\frac{\delta}{(2+\delta)}} M_A(l, \xi) |\hat{\omega}_S(l, \xi)| \frac{\mathbf{1}_{|(l, \xi)| \geq 1}}{|l, \xi|^{1+\frac{\delta}{(2+\delta)}}} d\xi d\eta \\
&\quad + \left| \sum_k \int_{\eta} \int_{\xi} k M_A(k, \eta) \overline{\hat{\omega}_S(k, \eta)} M_A(k, \eta-\xi) \hat{\omega}_S(k, \eta-\xi) M_A(0, \xi) \hat{u}_S(0, \xi) \mathbf{1}_{|\xi| \leq 1} d\xi d\eta \right| \\
&\lesssim_{\delta} \|\nabla |M_A \omega_S|\|_2 \|M_A \omega_S\|_2 \left\| |\nabla|^{\frac{\delta}{2+\delta}} M_A \omega_S \right\|_2 + \|\nabla |M_A \omega_S^{\nu}|\|_2 \|M_A \omega_S^{\nu}\|_2 \|\hat{u}_S(0, \xi) \mathbf{1}_{|\xi| \leq 1}\|_2 \\
&\lesssim \|\nabla |M_A \omega_S|\|_2^{\frac{2+2\delta}{2+\delta}} \|M_A \omega_S\|_2^{\frac{4+\delta}{2+\delta}} + \|\nabla |M_A \omega_S|\|_2 \|M_A \omega_S\|_2 \|u_S\|_2, \tag{2.8}
\end{aligned}$$

for any  $\delta > 0$ . In the last inequality we have used the Gagliardo-Nirenberg interpolation for the first term. In view of the bound (2.4), estimate (2.8) yields

$$\begin{aligned}
|N_A| &\leq \frac{C}{\nu} \|M_A \omega_S\|_2^2 \left( \frac{1}{\nu^{\delta}} \|M_A \omega_S\|_2^{2+\delta} + \|u_S\|_2^2 \right) + \frac{\nu}{2} \|\nabla |M_A \omega_S|\|_2^2 \\
&\leq \frac{C}{\nu^{1+\delta}} \|M_A \omega_S\|_2^2 (1 + \|M_A \omega_S\|_2)^{2+\delta} + \frac{\nu}{2} \|\nabla |M_A \omega_S|\|_2^2 \tag{2.9}
\end{aligned}$$

for some positive constant  $C \geq 1$ , that is independent of  $\nu, \epsilon' \leq 1$ , but depends on the choice of  $\delta > 0$ . Combining (2.6), (2.7), and (2.9) we arrive at

$$\begin{aligned}
\frac{d}{dt} \|M_A \omega_S\|_2^2 + \nu \|\nabla |M_A \omega_S|\|_2^2 &\leq 2\sqrt{\nu}(1+t) \left\| |\nabla|^{1/2} M_A \omega_S \right\|_2^2 + \frac{C}{\nu^{1+\delta}} \|M_A \omega_S\|_2^2 (1 + \|M_A \omega_S\|_2)^{2+\delta} \\
&\leq 2\sqrt{\nu}(1+t) \|\nabla |M_A \omega_S|\|_2 \|M_A \omega_S\|_2 + \frac{C}{\nu^{1+\delta}} \|M_A \omega_S\|_2^2 (1 + \|M_A \omega_S\|_2)^{2+\delta} \\
&\leq \frac{\nu}{2} \|\nabla |M_A \omega_S|\|_2^2 + 4(1+t)^2 \|M_A \omega_S\|_2^2 + \frac{C}{\nu^{1+\delta}} \|M_A \omega_S\|_2^2 (1 + \|M_A \omega_S\|_2)^{2+\delta},
\end{aligned}$$

for any  $\delta > 0$  and some constant  $C \geq 1$ . Therefore, for  $t \leq 1$

$$\frac{d}{dt} \|M_A \omega_S\|_2^2 \leq \frac{C_1}{\nu^{1+\delta}} \|M_A \omega_S\|_2^2 (1 + \|M_A \omega_S\|_2^2)^{1+\delta/2} \tag{2.10}$$

for some constant  $C_1 \geq 1$ . Letting

$$T_1 = \min \left\{ \frac{\nu^{1+\delta}}{2C_1}, T_{\lambda, \lambda'} \right\} = \frac{\nu^{1+\delta}}{2C_1} \quad (\text{without loss of generality, as } \nu \leq \nu_0), \tag{2.11}$$

we obtain from (2.10) and the assumption (2.1) on the  $H^{\sigma}$  norm of  $\omega_{S, in}$  that for  $\epsilon' \leq 1$

$$\|M_A(t) \omega_S(t)\|_2 \leq 2\epsilon' \quad \text{for all } t \leq T_1. \tag{2.12}$$

The Gevrey class estimates differ slightly from the real analytic bounds as we need to employ an additional commutator estimate. Consider the Gevrey- $\frac{1}{s}$  Fourier-multiplier

$$M_G(t, k, \eta) = \langle k, \eta \rangle^{\sigma} e^{\lambda(t)|k, \eta|^s} \tag{2.13}$$

where the decreasing function  $\lambda(t)$  shall be chosen precisely later, and  $\sigma > 5/2$  is a fixed Sobolev correction. Similarly to (2.6) we have

$$\frac{1}{2} \frac{d}{dt} \|M_G \omega_S\|_2^2 + \nu \|\nabla M_G \omega_S\|_2^2 = \dot{\lambda}(t) \left\| |\nabla|^{s/2} M_G \omega_S \right\|_2^2 + T_G + N_G \quad (2.14)$$

where  $T_G$  and  $N_G$  are defined precisely as  $T_A$  and  $N_A$  upon replacing all  $M_A$ 's by  $M_G$ 's. It is not hard to see that

$$|T_G| \lesssim \|M_G \omega_S\|_2^2 + \lambda(t) \left\| |\nabla|^{s/2} M_G \omega_S \right\|_2^2. \quad (2.15)$$

Bounding the nonlinear term  $N_G$  requires a commutator estimate which arises since  $u_S$  is divergence free, in the spirit of [41, 38]. More precisely, we have,

$$\begin{aligned} N_G &= \sum_{k,l} \int_{\eta} \int_{\xi} M_G(k, \eta) \overline{\hat{\omega}_S(k, \eta)} \hat{\omega}_S(l, \xi) \hat{\omega}_S(k-l, \eta-\xi) \\ &\quad \times \frac{(l, \xi)^\perp \cdot (k, \eta)}{|l, \xi|^2} (M_G(k, \eta) - M_G(k-l, \eta-\xi)) d\xi d\eta. \end{aligned}$$

By the mean-value theorem applied to  $M_G$ , we obtain that

$$\begin{aligned} |N_G| &\leq \lambda(t) \sum_{k,l} \int_{\eta} \int_{\xi} M_G(k, \eta) |\hat{\omega}_S(k, \eta)| M_G(k-l, \eta-\xi) |\hat{\omega}_S(k-l, \eta-\xi)| |\hat{\omega}_S(l, \xi)| \\ &\quad \times \left| \frac{(l, \xi)^\perp \cdot (k, \eta)}{|l, \xi|^2} \right| \frac{|l, \xi| e^{\lambda(t)|l, \xi|^s}}{|k, \eta|^{1-s} + |k-l, \eta-\xi|^{1-s}} d\xi d\eta \\ &\quad + \sum_{k,l} \int_{\eta} \int_{\xi} M_G(k, \eta) |\hat{\omega}_S(k, \eta)| M_G(k-l, \eta-\xi) |\hat{\omega}_S(k-l, \eta-\xi)| M_G(l, \xi) |\hat{\omega}_S(l, \xi)| \\ &\quad \times \left| \frac{(l, \xi)^\perp \cdot (k-l, \eta-\xi)}{|l, \xi|^2} \right| \left( \langle k-l, \eta-\xi \rangle^{\sigma-1} + \langle k, \eta \rangle^{\sigma-1} \right) \frac{C\sigma |l, \xi|}{\langle k-l, \eta-\xi \rangle^\sigma \langle l, \xi \rangle^\sigma} d\xi d\eta \\ &\leq \lambda(t) \sum_{k,l} \int_{\eta} \int_{\xi} M_G(k, \eta) |\hat{\omega}_S(k, \eta)| M_G(k-l, \eta-\xi) |\hat{\omega}_S(k-l, \eta-\xi)| M_G(l, \xi) |\hat{\omega}_S(l, \xi)| \\ &\quad \times \frac{|k, \eta|^{s/2} |k-l, \eta-\xi|^{s/2}}{\langle l, \xi \rangle^\sigma} d\xi d\eta \\ &\quad + C \sum_{k,l} \int_{\eta} \int_{\xi} M_G(k, \eta) |\hat{\omega}_S(k, \eta)| M_G(k-l, \eta-\xi) |\hat{\omega}_S(k-l, \eta-\xi)| M_G(l, \xi) |\hat{\omega}_S(l, \xi)| \\ &\quad \times \left( \frac{1}{\langle l, \xi \rangle^\sigma} + \frac{1}{\langle l, \xi \rangle \langle k-l, \eta-\xi \rangle^{\sigma-1}} \right) d\xi d\eta \\ &\leq C\lambda(t) \|M_G \omega_S\|_2 \left\| |\nabla|^{s/2} M_G \omega_S \right\|_2^2 + C \|M_G \omega_S\|_2^3 \end{aligned} \quad (2.16)$$

for some constant  $C \geq 1$  that may depend only on  $\sigma$ . Combining (2.14)–(2.16) we arrive at

$$\frac{1}{2} \frac{d}{dt} \|M_G \omega_S\|_2^2 \leq \left( \dot{\lambda}(t) + C_2 \lambda(t) (1 + \|M_G \omega_S\|_2) \right) \left\| |\nabla|^{s/2} M_G \omega_S \right\|_2^2 + C_2 \|M_G \omega_S\|_2^2 (1 + \|M_G \omega_S\|_2) \quad (2.17)$$

for some constant  $C_2 \geq 1$ . The initial datum by the assumption (2.1) obeys

$$\|M_G(0) \omega_{S,in}\|_2 \leq \epsilon'$$

whenever  $\lambda(0) \leq \lambda$ . Letting

$$\lambda(t) = \frac{9\lambda + \lambda'}{10} e^{-3C_2 t}$$

we obtain that (recall  $T_1$  is defined in (2.11)), for  $\epsilon'$  sufficiently small, depending on  $\sigma, \lambda, \lambda'$ , for all

$$t \leq T_2 = \min \left\{ T_1, \frac{1}{3C_2} \left( 1 + \log \frac{9\lambda + \lambda'}{8\lambda + 2\lambda'} \right) \right\} = \frac{\nu^{1+\delta}}{2C_1} \quad (\text{without loss of generality, as } \nu \leq \nu_0), \quad (2.18)$$

the following holds:

$$\|M_G(t)\omega_S(t)\|_2 \leq 2\epsilon' \quad \text{and} \quad \lambda(t) \geq \frac{4\lambda + \lambda'}{5}. \quad (2.19)$$

We have now completed the proof of step (b), and shown that  $\omega_S(t) \in \mathcal{G}^{\sqrt{\nu}t;1} \cap G^{\frac{4\lambda+\lambda'}{5};s}$ , with norm less than  $2\epsilon'$ , but only for times  $t \leq T_2 \lesssim \nu^{1+\delta}$ .

*Proof of (c).* We now wish to show that the “rough” perturbation  $\omega_R = \omega - \omega_S$ , which in view of (1.1), obeys

$$\begin{cases} \partial_t \omega_R + y \partial_x \omega_R + u_R \cdot \nabla \omega_R + u_S \cdot \nabla \omega_R + u_R \cdot \nabla \omega_S = \nu \Delta \omega_R, \\ u_S = \nabla^\perp (\Delta)^{-1} \omega_S, u_R = \nabla^\perp (\Delta)^{-1} \omega_R \\ \omega_R(0) = \omega_{R,in}, \end{cases} \quad (2.20)$$

is sufficiently small in  $\mathcal{G}^{\sqrt{\nu}t;1}$  at time  $t = T_2$ . The proof is very much similar to the first part of step (b), and we thus omit the redundant details. Before estimating the analytic norm of  $\omega_R$  we estimate the  $L^2$  norm of  $u_R$ , which solves (1.2) perturbed about  $u_S$ . From the energy estimate we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u_R\|_2^2 + \nu \|\omega_R\|_2^2 &= - \int u_R^y u_R^x - \int u_R \cdot \nabla u_S u_R \\ &\leq \|u_R\|_2^2 (1 + \|\nabla u_S\|_{L^\infty}) \leq 3 \|u_R\|_2^2 \end{aligned} \quad (2.21)$$

for all  $t \leq T_2$ , by the Sobolev embedding, which holds because  $\sigma > 5/2$ , and we have used the bound (2.19). In particular, in view of the assumption (2.1) on the initial datum  $u_{R,in}$  which is supported away from frequencies less than 1, we obtain that

$$\|u_R(t)\|_2 \leq 21 \|u_{R,in}\|_2 \leq 21 \|\omega_{R,in}\|_2 \leq 21 e^{-K_0 \nu^{-\frac{(3+2\delta)s}{2(1-s)}}} \epsilon', \quad (2.22)$$

for all  $t \leq T_2$ .

Since  $\omega_{R,in} \in L^2$ , and not necessarily in  $H^\sigma$ , we need to consider a new analytic Fourier-multiplier

$$M_R(t, k, \eta) = e^{\sqrt{\nu} t |k, \eta|}. \quad (2.23)$$

Then the solution  $\omega_R$  of (2.20) obeys the a priori estimate

$$\begin{aligned}
& \frac{1}{2} \frac{d}{dt} \|M_R \omega_R\|_2^2 + \nu \|\nabla M_R \omega_R\|_2^2 \\
&= \sqrt{\nu} \left\| |\nabla|^{1/2} M_R \omega_R \right\|_2^2 + \sum_k \int_{\eta} M_R(k, \eta)^2 k \overline{\hat{\omega}_R(k, \eta)} \partial_{\eta} \hat{\omega}_R(k, \eta) d\eta \\
&+ \sum_{k, l} \int_{\eta} \int_{\xi} M_R(k, \eta) \overline{\hat{\omega}_R(k, \eta)} M_R(l, \xi) \hat{\omega}_R(l, \xi) M_R(k-l, \eta-\xi) \hat{\omega}_R(k-l, \eta-\xi) \\
&\quad \times \frac{(l, \xi)^{\perp} \cdot (k, \eta)}{|l, \xi|^2} \frac{M_R(k, \eta)}{M_R(l, \xi) M_R(k-l, \eta-\xi)} d\eta d\xi \\
&+ \sum_{k, l} \int_{\eta} \int_{\xi} M_R(k, \eta) \overline{\hat{\omega}_R(k, \eta)} M_R(l, \xi) \hat{\omega}_S(l, \xi) M_R(k-l, \eta-\xi) \hat{\omega}_R(k-l, \eta-\xi) \\
&\quad \times \frac{(l, \xi)^{\perp} \cdot (k, \eta)}{|l, \xi|^2} \frac{M_R(k, \eta)}{M_R(l, \xi) M_R(k-l, \eta-\xi)} d\xi d\eta \\
&+ \sum_{k, l} \int_{\eta} \int_{\xi} M_R(k, \eta) \overline{\hat{\omega}_R(k, \eta)} M_R(l, \xi) \hat{\omega}_R(l, \xi) M_R(k-l, \eta-\xi) \hat{\omega}_S(k-l, \eta-\xi) \\
&\quad \times \frac{(l, \xi)^{\perp} \cdot (k, \eta)}{|l, \xi|^2} \frac{M_R(k, \eta)}{M_R(l, \xi) M_R(k-l, \eta-\xi)} d\xi d\eta \\
&= \sqrt{\nu} \left\| |\nabla|^{1/2} M_R \omega_S \right\|_2^2 + T_R + N_R + L_{S,R} + L_{R,S}.
\end{aligned} \tag{2.24}$$

As before, similarly to (2.7) we have

$$|T_R| \lesssim \sqrt{\nu} t \left\| |\nabla|^{1/2} M_R \omega_S \right\|_2^2 \tag{2.25}$$

and similarly to (2.9) we have

$$|N_R| \leq \frac{C}{\nu^{1+\delta}} \|M_R \omega_R\|_2^2 (1 + \|M_R \omega_R\|_2)^{2+\delta} + \frac{\nu}{2} \|\nabla M_R \omega_R\|_2^2 \tag{2.26}$$

for some constant  $C \geq 1$  which depends on  $\delta > 0$ . Bounding the first linear term  $L_{S,R}$  is done exactly as in (2.9) by splitting according to the relative size of  $|(l, \xi)|$  and 1, and using the already established bound (2.12). We obtain

$$\begin{aligned}
|L_{S,R}| &\lesssim \|\nabla M_R \omega_S\|_2 \|M_R \omega_R\|_2 \|\nabla M_R \omega_R\|_2 + \|u_S\|_2 \|M_R \omega_R\|_2 \|\nabla M_R \omega_R\|_2 \\
&\leq \frac{C(\epsilon')^2}{\nu} \|M_R \omega_R\|_2^2 + \frac{\nu}{4} \|\nabla M_R \omega_R\|_2^2
\end{aligned} \tag{2.27}$$

for all  $t \leq T_2$ . For the second linear term  $L_{R,S}$  we recall that for the analytic norm of  $\omega_S$  we have a Sobolev  $H^\sigma$  correction which allows us to deal with the derivative loss in  $u_R \cdot \nabla \omega_S$ , and in addition we have a good control on the low frequencies in view of (2.22). We then obtain

$$\begin{aligned}
|L_{R,S}| &\lesssim \|\nabla M_R \omega_R\|_2 \|M_R \omega_R\|_2 \|\nabla M_R \omega_S\|_2 + \|u_R\|_2 \|M_R \omega_R\|_2 \|\nabla M_R \omega_S\|_2 \\
&\leq \frac{C(\epsilon')^2}{\nu} \|M_R \omega_R\|_2^2 + \frac{\nu}{4} \|\nabla M_R \omega_R\|_2^2,
\end{aligned} \tag{2.28}$$

for some constant  $C \geq 1$ . Combining (2.22)–(2.28), and assuming without loss of generality that  $\epsilon'$  is sufficiently small (in terms of  $C_1, C_2$ ), we arrive at

$$\frac{d}{dt} \|M_R \omega_R\|_2^2 \leq \frac{3C_1}{2\nu^{1+\delta}} \|M_R \omega_R\|_2^2 (1 + \|M_R \omega_R\|_2^2)^{1+\delta/2} + C_1 \epsilon' \|\omega_{R,in}\|_2^2 \quad (2.29)$$

for all  $t \leq T_2$ , and  $\nu \leq 1$ , where  $C_1$  is the same constant as in (2.10). Since by definition (cf. (2.11)) we have  $T_2 \leq T_1 \leq \nu^{1+\delta}/(2C_1)$ , and  $M_R(0) = 1$ , we obtain from (2.29) that

$$\|M_R(t) \omega_R(t)\|_2^2 \leq 9 \|\omega_{R,in}\|_2^2.$$

for all  $t \in [0, T_2]$ . It thus follows from the assumption (2.1) that

$$\|M_R(T_2) \omega_R(T_2)\|_2 \leq 3\epsilon' e^{-K_0 \nu^{-\frac{(3+2\delta)s}{2(1-s)}}}. \quad (2.30)$$

We need the above estimate as it yields a Gevrey- $\frac{1}{s}$  bound on  $\omega_R^\nu$  at time  $T_2$ . Indeed, by (A.14) and (A.13),

$$\begin{aligned} \langle k, \xi \rangle^\sigma e^{\frac{4\lambda+\lambda'}{5}|k,\xi|^s} &\leq \langle k, \xi \rangle^\sigma e^{(\frac{4\lambda+\lambda'}{5})^{\frac{1}{1-s}} (2\sqrt{\nu}T_2)^{-\frac{s}{1-s}}} e^{\frac{\sqrt{\nu}}{2}T_2|k,\xi|} \\ &= \langle k, \xi \rangle^\sigma e^{-\frac{\nu^{\frac{3+2\delta}{2}}}{4C_1}|k,\xi|} e^{(\frac{4\lambda+\lambda'}{5})^{\frac{1}{1-s}} (4C_1)^{\frac{s}{1-s}} \nu^{-\frac{(3+2\delta)s}{2(1-s)}}} e^{\sqrt{\nu}T_2|k,\xi|} \\ &\leq e^{K_0 \nu^{-\frac{(3+2\delta)s}{2(1-s)}}} M_R(T_2) \end{aligned} \quad (2.31)$$

by letting  $K_0$  be sufficiently large, *independently* of  $\nu \leq 1$ . Here we have recalled the definition of  $T_2$  in (2.18). Combining (2.30)–(2.31), we thus have shown that for  $\epsilon'$  sufficiently small, depending only on  $s, \sigma, \lambda, \lambda'$ , and a sufficiently large  $K_0$ , depending on the same parameters, that

$$\|M_G(T_2) \omega_R(T_2)\|_2 \leq 3\epsilon', \quad \text{and} \quad \lambda(T_2) = \frac{4\lambda + \lambda'}{5} \quad (2.32)$$

which concludes the proof of step (c).

*Proof of (d).* To conclude the proof of the proposition, we now *mutatis-mutandi* run the Gevrey-class regularity propagation on the time interval  $[T_2, T_{\lambda,\lambda'}]$ , on the equation (1.1), with initial datum  $\omega_S(T_2) + \omega_R(T_2)$ . The estimate (2.17) holds on this interval, and thus, if we let

$$\lambda(t) = \frac{4\lambda + \lambda'}{5} e^{-3C_2(t-T_2)} \quad (2.33)$$

we see that

$$\|M_G(t) \omega(t)\|_2 \leq \epsilon \quad (2.34)$$

for all  $t \leq T_{\lambda,\lambda'}$ . Moreover, by letting the constant  $C_0$  in the definition of  $T_{\lambda,\lambda'}$  be sufficiently large, we can ensure that

$$\lambda(t) \geq \lambda(T_{\lambda,\lambda'}) \geq \frac{3\lambda + \lambda'}{4} \quad (2.35)$$

as claimed in (2.2). To conclude the proof of Proposition 2.1, we note that the bound  $\|u(T_{\lambda,\lambda'})\|_2 \leq \epsilon$  follows from a bound similar to (2.4), which holds for an  $O(1)$  time interval.  $\square$

## 2.2 Coordinate transform

One of the main insights in [9] was the use of a very specific change of coordinates that allows to simultaneously ‘mod out’ by the evolution of the time-dependent background shear flow and treat the mixing of this background shear as a perturbation of the Couette flow (in particular, to understand the nonlinear effect of the Orr mechanism). In this work we will adjust the change of coordinates used in [9] to account for the viscous dissipation in the following manner:

$$z(t, x, y) = x - tv(t, y), \quad (2.36a)$$

$$\frac{d}{dt}(t(v(t, y) - y)) = U_0^x(t, y) + \nu t \partial_{yy} v(t, y), \quad (2.36b)$$

supplemented with the initial condition

$$\lim_{t \rightarrow 0} t(v(t, y) - y) = 0.$$

Note that when  $\nu = 0$ , (2.36) reduces to the change of coordinates used in [9] for 2D Euler. As briefly discussed in §1.1, the coordinate transform is adapting to the momentum transport due to the viscosity in the background shear flow. It is important to note that the shear flow driving the mixing, and hence the enhanced dissipation, is the unknown, evolving flow  $(y + U_0^x(t, y), 0)$ .

Define the following quantities

$$C(t, v(t, y)) = v(t, y) - y, \quad (2.37a)$$

$$v'(t, v(t, y)) = (\partial_y v)(t, y), \quad (2.37b)$$

$$v''(t, v(t, y)) = (\partial_{yy} v)(t, y), \quad (2.37c)$$

$$[\partial_t v](t, v(t, y)) = (\partial_t v)(t, y), \quad (2.37d)$$

$$f(t, z(t, x, y), v(t, y)) = \omega(t, x, y), \quad (2.37e)$$

$$\phi(t, z(t, x, y), v(t, y)) = \psi(t, x, y), \quad (2.37f)$$

$$\tilde{u}(t, z(t, x, y), v(t, y)) = U^x(t, x, y). \quad (2.37g)$$

Denoting  $\Delta_t$  as the Laplacian in the coordinates (2.36), we derive the new Biot-Savart law:

$$\Delta_t \phi = \partial_{zz} \phi + (v')^2 (\partial_v - t \partial_z)^2 \phi + v'' (\partial_v - t \partial_z) \phi = f. \quad (2.38)$$

For future notational convenience denote the ‘linear’ Laplacian  $\Delta_L$  (the Laplacian associated with the pure Couette flow that arises in the coordinates  $z = x - tv, v = y$ ):

$$\Delta_L \phi = \partial_{zz} \phi + (\partial_v - t \partial_z)^2 \phi. \quad (2.39)$$

The vorticity equation (1.1) in the new variables becomes

$$\partial_t f + [\partial_t v] \partial_v f - \nu v'' t \partial_z f + v' \nabla_{z,v}^\perp P_{\neq 0} \phi \cdot \nabla_{z,v} f = \nu \Delta_t f. \quad (2.40)$$

Note the derivatives are with respect to  $z$  and  $v$ ; we will henceforth drop the subscripts and always assume this is the case unless otherwise indicated. Taking the  $x$  average of the first equation in (1.2) and transforming to the coordinates defined by (2.36) (using the notations (2.37)) gives

$$\partial_t \tilde{u}_0 + [\partial_t v] \partial_v \tilde{u}_0 + \langle v' \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} \rangle = \nu \Delta_t \tilde{u}_0. \quad (2.41)$$



Next, we derive expressions for the coordinate system that are amenable to energy estimates and derive a useful cancellation in (2.40) and (2.41). The parabolic aspect to the coordinate system in (2.36b) is crucial for both of these. Define the auxiliary function

$$g(t, v) = \frac{1}{t} (\tilde{u}_0(t, v) - C(t, v)), \quad (2.42)$$

which roughly measures how the background shear flow is converging in time. In 2D Euler [9], this quantity is exactly  $[\partial_t v]$ , however due to the presence of the dissipation in (2.36b) here we instead have the relation

$$[\partial_t v] = g + \nu v''. \quad (2.43)$$

We also derive from the chain rule:

$$v' \partial_v C(t, v) = v'(t, v) - 1, \quad (2.44a)$$

$$\partial_t C + [\partial_t v] \partial_v C = [\partial_t v], \quad (2.44b)$$

$$v' \partial_v v' = v'', \quad (2.44c)$$

$$v'' = \Delta_t C. \quad (2.44d)$$

Computing the time derivative of  $g$  from (2.42) using (2.44b), (2.43), (2.41) and (2.44d), yields

$$\begin{aligned} \partial_t g &= -\frac{g}{t} + \frac{1}{t} \left( -[\partial_t v] \partial_v \tilde{u}_0 - v' \langle \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} \rangle + \nu \Delta_t \tilde{u}_0 + [\partial_t v] \partial_v C - g - \nu v'' \right) \\ &= -\frac{2g}{t} - [\partial_t v] \partial_v g - \frac{v'}{t} \langle \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} \rangle + \nu \Delta_t g. \end{aligned} \quad (2.45)$$

The form of the dissipation in (2.36b) was made precisely so that a natural dissipation arises on the LHS of (2.45), as opposed to a dangerous forcing term such as  $\Delta_t \tilde{u}_0$ . The coordinate system is properly adapting itself to the effect of the dissipation that  $\tilde{u}_0$  is experiencing. This will also manifest in another way. Indeed, by (2.43) and the definition of  $\Delta_t$  (2.38), we reduce (2.45) to

$$\partial_t g + \frac{2g}{t} + g \partial_v g = -\frac{v'}{t} \langle \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} \rangle + \nu (v')^2 \partial_{vv} g. \quad (2.46)$$

The cancellation will turn out to be quite useful, as while  $[\partial_t v]$  will decay slowly for long times,  $g$  will decay essentially like  $O(t^{-2})$  uniformly in  $\nu$  (the same rate as  $[\partial_t v]$  in Euler [9]). This cancellation in (2.46) occurs throughout several equations so define:

$$\tilde{\Delta}_t = \partial_{zz} + (v')^2 (\partial_v - t \partial_z)^2. \quad (2.47)$$

From (2.40) and (2.43), we derive a similar cancellation on the vorticity evolution:

$$\partial_t f + u \cdot \nabla f = \nu \tilde{\Delta}_t f, \quad (2.48)$$

where we are defining the relative velocity field

$$u(t, z, v) = \begin{pmatrix} 0 \\ g \end{pmatrix} + v' \nabla^\perp P_{\neq 0} \phi. \quad (2.49)$$

Similarly, if we denote  $h = v' - 1$ , we derive from (2.36b) that

$$\partial_t h + g \partial_v h = \frac{-f_0 - h}{t} + \nu \tilde{\Delta}_t h. \quad (2.50)$$

Next, we define the unknown  $\bar{h} = \frac{-f_0 - h}{t}$  (which quantifies the convergence of the background vorticity) and derive from (2.48) and (2.50)

$$\partial_t \bar{h} + g \partial_v \bar{h} = -\frac{2}{t} \bar{h} + \frac{v'}{t} < \nabla^\perp P_{\neq 0} \phi \cdot \nabla f > + \nu \tilde{\Delta}_t \bar{h}. \quad (2.51)$$

Finally, note from the Biot-Savart law and the definitions of  $\bar{h}$  and  $g$  we have

$$v' \partial_v g = \bar{h}. \quad (2.52)$$

The cancellations in (2.46), (2.48), (2.50) and (2.51) imply that we have effectively modded out the slow decay of  $[\partial_t v]$  – or more precisely, we are moving the coordinate system in such a way so that the slow part of  $[\partial_t v]$  is balancing the lower order drift term coming from the  $\Delta_t$  in the dissipation. The choice of (2.36) appears to be the unique coordinate transformation that simultaneously satisfies this property as well as the properties used to obtain inviscid damping in [9].

Note that from (2.36), the conservation laws of (1.1) and Jensen's inequality, we have

$$\| < f > \|_p \lesssim \| f \|_p \lesssim \| (v')^{-1} \|_\infty^{1/p} \| \omega \|_p \lesssim \epsilon \| (v')^{-1} \|_\infty^{1/p}, \quad (2.53)$$

for all  $1 \leq p \leq \infty$ . Similarly, from (2.36) it follows that

$$\| v'(t) - 1 \|_1 \lesssim \| (v')^{-1} \|_\infty \| v_y(t) - 1 \|_1 \lesssim \| (v')^{-1} \|_\infty \left( \sup_{\tau \leq t} \| \omega_0(\tau) \|_1 \right) \lesssim \epsilon \| (v')^{-1} \|_\infty. \quad (2.54)$$

### 2.3 Energy Estimates

In this section we describe the bootstrap argument that propagates estimates from  $O(1)$  times to infinite times. The energy estimates will involve two main multipliers:

$$A(t, k, \eta) = e^{\lambda(t)|k, \eta|^s} \langle k, \eta \rangle^\sigma J_k(t, \eta), \quad (2.55a)$$

$$A^\nu(t, k, \eta) = e^{\lambda(t)|k, \eta|^s} \langle k, \eta \rangle^\beta \langle D(t, \eta) \rangle^\alpha \mathbf{1}_{k \neq 0}. \quad (2.55b)$$

The properties and detailed definitions of both multipliers can be found in §3 below. The index  $\lambda(t)$  is the main Gevrey  $-\frac{1}{s}$  regularity and is chosen to satisfy (recall definition (2.3)),

$$\lambda(t) = \frac{3}{4} \lambda + \frac{1}{4} \lambda', \quad t \leq T_{\lambda, \lambda'} \quad (2.56a)$$

$$\dot{\lambda}(t) = -\frac{\delta_\lambda}{\langle t \rangle^{2\tilde{q}}} (1 + \lambda(t)), \quad t > T_{\lambda, \lambda'} \quad (2.56b)$$

where  $\delta_\lambda \approx \lambda - \lambda'$  is a small parameter that ensures  $\lambda(t) > \lambda/2 + \lambda'/2$  and  $\tilde{q} \in (1/2, s/8 + 7/16)$  is a parameter chosen by the proof in [9]. Moreover, as in [9], we will assume without loss of generality that  $s < 2/3$ ; for larger  $s$ , an additional Gevrey regularity correction  $1/2 < \gamma < 2/3$  can be added to replace the role of  $s$  in the ensuing proof.

The high norm defined by  $A$  is the same used in the work on the Euler equations [9] (and in particular is independent of  $\nu$ ). The corrector  $J$  is defined in §3 of [9] and recalled here along with the relevant properties below in §3.1. We will also need the stronger variants  $A^R \geq A$ ,  $J^R \geq J$  defined in (3.7) below. The second multiplier,  $A^\nu$ , quantifies the enhanced dissipation effect through the inclusion of  $D$ , defined by

$$D(t, \eta) = \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu \left( t^3 - 8 |\eta|^3 \right)_+. \quad (2.57)$$

We will choose  $\beta + 3\alpha + 8 < \sigma$ , so that a sizable regularity gap is maintained between  $A$  and  $A^\nu$  and  $\beta > 3\alpha + 2$  (for less crucial technical convenience). Hence, for sufficiently high frequencies (depending on  $t$ ),  $A$  will define a stronger norm, whereas for lower frequencies  $A^\nu$  is stronger (for  $k \neq 0$ ). The point of introducing  $D$  is so that a uniform bound expresses a powerful enhanced dissipation effect:

$$\|P_{\neq 0}f(t)\|_{\mathcal{G}^{\lambda(t),\beta}} \lesssim_\alpha \langle \nu t^3 \rangle^{-\alpha} \|A^\nu f(t)\|_2. \quad (2.58)$$

The role this estimate plays in the proof is something like a parabolic analogue of the lossy elliptic Lemma 4.1.

The proof is based on a bootstrap argument. Let  $[T_{\lambda,\lambda'}, T^*]$  be the largest connected, closed interval with left endpoint  $T_{\lambda,\lambda'}$  such that the following *bootstrap hypotheses* hold for some fixed, positive constants  $K_v$  and  $K_D$  depending only on  $\lambda, \lambda', \alpha$  and  $s$  determined by the proof:

- The ‘high frequency’ controls (the first on the vorticity and the latter four on the coordinate system)

$$\|Af(t)\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda,\lambda'}}^t \left\| \sqrt{-\Delta_L} Af(\tau) \right\|_2^2 d\tau + \int_{T_{\lambda,\lambda'}}^t CK_\lambda + CK_w d\tau \leq 4\epsilon^2, \quad (2.59a)$$

$$t^{2+2s} \left\| \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda,\lambda'}}^t \tau^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 d\tau + \int_{T_{\lambda,\lambda'}}^t CK_\lambda^{v,2} + CK_w^{v,2} d\tau \leq 4\epsilon^2, \quad (2.59b)$$

$$t^{4-K_D\epsilon} \|g\|_{\mathcal{G}^{\lambda,\sigma-6}}^2 + \nu \int_{T_{\lambda,\lambda'}}^t \tau^{4-K_D\epsilon} \|\partial_v g\|_{\mathcal{G}^{\lambda(\tau),\sigma-6}}^2 d\tau \leq 4\epsilon^2, \quad (2.59c)$$

$$\int_{T_{\lambda,\lambda'}}^t CK_\lambda^{v,1} + CK_w^{v,1} d\tau \leq 4K_v\epsilon^2, \quad (2.59d)$$

$$\|A^R(v' - 1)\|_2^2 + \int_{T_{\lambda,\lambda'}}^t \sum_{i=1}^2 CCK_\lambda^i + CCK_w^i d\tau \leq 4K_v\epsilon^2; \quad (2.59e)$$

- the enhanced dissipation estimate

$$\|A^\nu f\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda,\lambda'}}^t \left\| \sqrt{-\Delta_L} A^\nu f(\tau) \right\|_2^2 d\tau + \int_{T_{\lambda,\lambda'}}^t CK_\lambda^\nu d\tau \leq 4\epsilon^2; \quad (2.60)$$

- the low frequency controls

$$\|v' - 1\|_\infty < \frac{3}{4}, \quad (2.61a)$$

$$\|v' - 1\|_2 + \langle \nu t \rangle^{1/2} \|\partial_v (v' - 1)\|_2 \leq 4K_v\epsilon \langle \nu t \rangle^{-1/4}, \quad (2.61b)$$

$$\|\tilde{u}_0(t)\|_2^2 + \frac{\nu}{10} \int_0^t \|v' \partial_v \tilde{u}_0(\tau)\|_2^2 d\tau \leq 4K_v\epsilon^2. \quad (2.61c)$$

The numerous ‘CK’ terms (for *Cauchy-Kovalevskaya* since the most important of these terms arise

naturally from the weakening of the norms in time) are given by

$$CK_\lambda(t) = -\dot{\lambda}(t) \left\| |\nabla|^{s/2} Af \right\|_2^2, \quad (2.62a)$$

$$CK_w(t) = \sum_k \int \frac{\partial_t w_k(t, \eta)}{w_k(t, \eta)} e^{\lambda(t)|k, \eta|^s} \langle k, \eta \rangle^\sigma \frac{e^{\mu|\eta|^{1/2}}}{w_k(t, \eta)} A_k(t, \eta) \left| \hat{f}_k(t, \eta) \right|^2 d\eta, \quad (2.62b)$$

$$CK_\lambda^\nu(t) = -\dot{\lambda}(t) \left\| |\nabla|^{s/2} A^\nu f(t) \right\|_2^2, \quad (2.62c)$$

$$CK_w^{v,2}(t) = \langle t \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} \bar{h}(t) \right\|_2^2, \quad (2.62d)$$

$$CK_\lambda^{v,2}(t) = \langle t \rangle^{2+2s} (-\dot{\lambda}(t)) \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} \bar{h}(t) \right\|_2^2, \quad (2.62e)$$

$$CK_w^{v,1}(t) = \langle t \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} g(t) \right\|_2^2, \quad (2.62f)$$

$$CK_\lambda^{v,1}(t) = \langle t \rangle^{2+2s} (-\dot{\lambda}(t)) \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} g(t) \right\|_2^2, \quad (2.62g)$$

along with the ‘coefficient Cauchy-Kovalevskaya’ terms

$$CCK_\lambda^1(t) = -\dot{\lambda}(t) \left\| |\partial_v|^{s/2} A^R (1 - (v')^2) (t) \right\|_2^2, \quad (2.63a)$$

$$CCK_w^1(t) = \left\| \sqrt{\frac{\partial_t w}{w}} A^R (1 - (v')^2) (t) \right\|_2^2, \quad (2.63b)$$

$$CCK_\lambda^2(t) = -\dot{\lambda}(t) \left\| |\partial_v|^{s/2} \frac{A^R}{\langle \partial_v \rangle} v''(t) \right\|_2^2, \quad (2.63c)$$

$$CCK_w^2(t) = \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A^R}{\langle \partial_v \rangle} v''(t) \right\|_2^2. \quad (2.63d)$$

For future convenience (see (2.62b) above) we also define

$$\tilde{J}_k(t, \eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t, \eta)}, \quad (2.64a)$$

$$\tilde{A}_k(t, \eta) = e^{\lambda(t)|k, \eta|^s} \langle k, \eta \rangle^\sigma \tilde{J}_k(t, \eta). \quad (2.64b)$$

Note that  $\tilde{A} \leq A$  and it turns out that if  $|k| \leq \frac{1}{4}|\eta|$  then  $A \lesssim \tilde{A}$  (see Lemma 3.1 below).

It is easy to show that the quantities on the left hand sides of (2.59), (2.60) and (2.61) take values continuously in time for  $t > 0$  due to the analyticity of solutions to (1.1) and similarly, using Proposition 2.1, that  $T^* > T_{\lambda, \lambda'}$ . Note by (2.61b), (2.61a) and (2.44c) we have

$$\|v''\|_2^2 = \|v' \partial_v v'\|_2^2 \lesssim \|v'\|_\infty^2 \|\partial_v v'\|_2^2 \lesssim \frac{\epsilon^2}{\langle \nu t \rangle^{3/2}}. \quad (2.65)$$

The primary step to the proof of Theorem 1 is the following, which shows that the bootstrap controls can be propagated indefinitely.

**Proposition 2.2.** *For  $\epsilon \leq \epsilon_0$  chosen sufficiently small depending only on  $s, \lambda, \lambda'$  and  $\alpha$ , the following estimates hold uniformly on  $t \in [T_{\lambda, \lambda'}, T^*]$ :*

- The ‘high frequency’ controls

$$\|Af(t)\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda, \lambda'}}^t \left\| \sqrt{-\Delta_L} Af(\tau) \right\|_2^2 d\tau + \int_{T_{\lambda, \lambda'}}^t CK_\lambda + CK_w d\tau \leq 2\epsilon^2 \quad (2.66a)$$

$$t^{2+2s} \left\| \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda, \lambda'}}^t \tau^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 d\tau + \int_{T_{\lambda, \lambda'}}^t CK_\lambda^{v,2} + CK_w^{v,2} d\tau \leq 2\epsilon^2 \quad (2.66b)$$

$$t^{4-K_D\epsilon} \|g\|_{\mathcal{G}^{\lambda, \sigma-6}}^2 + \frac{\nu}{10} \int_{T_{\lambda, \lambda'}}^t \tau^{4-K_D\epsilon} \|\partial_v g\|_{\mathcal{G}^{\lambda(\tau), \sigma-6}}^2 d\tau \leq 2\epsilon^2 \quad (2.66c)$$

$$\int_{T_{\lambda, \lambda'}}^t CK_\lambda^{v,1} + CK_w^{v,1} d\tau \leq 2K_v \epsilon^2 \quad (2.66d)$$

$$\|A^R(v' - 1)\|_2^2 + \int_{T_{\lambda, \lambda'}}^t \sum_{i=1}^2 CCK_\lambda^i + CCK_w^i d\tau \leq 2K_v \epsilon^2; \quad (2.66e)$$

- the enhanced dissipation estimate

$$\|A^\nu f\|_2^2 + \frac{\nu}{10} \int_{T_{\lambda, \lambda'}}^t \left\| \sqrt{-\Delta_L} A^\nu f(\tau) \right\|_2^2 d\tau + \int_{T_{\lambda, \lambda'}}^t CK_\lambda^\nu d\tau \leq 2\epsilon^2; \quad (2.67)$$

- the low frequency controls

$$\|v' - 1\|_\infty < \frac{5}{8} \quad (2.68a)$$

$$\|v' - 1\|_2 + \langle \nu t \rangle^{1/2} \|\partial_v (v' - 1)\|_2 \leq 2K_v \epsilon \langle \nu t \rangle^{-1/4}, \quad (2.68b)$$

$$\|\tilde{u}_0(t)\|_2^2 + \frac{\nu}{10} \int_0^t \|v' \partial_v \tilde{u}_0(\tau)\|_2^2 d\tau \leq 2\epsilon^2, \quad (2.68c)$$

$$\|f_0\|_2 \leq \frac{2K_v \epsilon}{\langle \nu t \rangle^{1/4}}. \quad (2.68d)$$

It therefore follows that  $T^* = \infty$ .

The proof of Proposition 2.2 constitutes the majority of our work; let us briefly outline it here. Except for the presence of the dissipation integrals, the estimates (2.66) are essentially the same as those propagated in the bootstrap argument for 2D Euler in [9] (along with (2.68a)). Since we want to be able to send  $\nu \rightarrow 0$  independently of  $t$  and  $\epsilon$ , the dissipation cannot help to get these estimates, and so the proof of (2.66) must also contain all the work done on the inviscid problem in [9] (we will try to repeat as little as possible). The main new challenge to deducing (2.66) for the Navier-Stokes equations (1.1) is ensuring that the presence of the variable coefficients in  $\tilde{\Delta}_t$  does not create growth or arrest the decay estimates. In (2.66a), one could worry that having rapid variations in the  $v$ -dependent enhanced dissipation rate could amplify gradients of  $f$ . Moreover, the fact that the enhanced dissipation slows down near the critical times will cause additional problems here. We will use (2.60) to provide the crucial decay necessary to handle the effect of

high frequencies in the coefficients of  $\tilde{\Delta}_t$  strongly forcing non-zero modes in  $z$ . This is detailed in §5 below. The NSE-specific challenge to deducing (2.66b), (2.66c) and (2.66e) is that having a  $v$ -dependent diffusivity in  $\tilde{\Delta}_t$  can induce effective bulk transport in  $v$ , a phenomenon that is mostly associated with low frequencies in the derivative of the coefficients. The effect is nonlinear, as the coefficients are coupled to  $f$ , but is controlled by the low frequency decay estimate (2.68b). This is detailed in §7.

The proof of the enhanced dissipation estimate (2.67) can be found in §6. The main issue confronted in §6 is that there are terms in (2.48) which are linear in the  $k \neq 0$  frequencies. For these terms, we cannot use the enhanced dissipation and so we must still carefully use some of the structural properties of the Euler nonlinearity (e.g. the transport structure), as if  $A^\nu f$  were the high norm. As a result, the chosen  $A^\nu$  still needs to satisfy certain amenable properties. Moreover, these linear terms also express the strong forcing of non-zero frequencies in  $z$  by the vastly larger zero frequencies in  $z$ . The regularity gap between  $A$  and  $A^\nu$  was chosen so that we could use (2.59a) to control this effect.

The low frequency controls (2.68) are relatively straightforward consequences of the slow viscous dissipation of  $z$ -independent (equivalently  $x$ -independent) quantities and are proved in §8 using the a priori high norm controls from (2.59a) and (2.60) and standard Moser iteration techniques.

## 2.4 Conclusion of the proof

To conclude the proof of Theorem 1 from Proposition 2.2 we need to translate the information back into information on the original unknowns,  $\omega$  and  $\psi$ . This is not trivial, but we may follow closely the corresponding steps in [9] with some minor alterations. By Proposition 2.2, we have the global uniform bounds

$$\|f(t)\|_{\mathcal{G}^{\lambda(t),\sigma}}^2 + \langle \nu t^3 \rangle^{2\alpha} \|P_{\neq 0} f(t)\|_{\mathcal{G}^{\lambda(t),\beta}}^2 + \langle \nu t \rangle^{1/2} \|f_0(t)\|_2^2 \lesssim \epsilon^2 \quad (2.69a)$$

$$\|\tilde{u}_0\|_{\mathcal{G}^{\lambda(t),\sigma}}^2 + \langle \nu t^3 \rangle^{2\alpha} \langle t \rangle^4 \|P_{\neq 0} \phi\|_{\mathcal{G}^{\lambda(t),\beta-3}}^2 + \|v' - 1\|_{\mathcal{G}^{\lambda(t),\sigma}}^2 \lesssim \epsilon^2. \quad (2.69b)$$

Define  $\lambda_\infty = \lim_{t \rightarrow \infty} \lambda(t)$ . In order to complete the proof of Theorem 1, we undo the change of coordinates in  $v$ , switching to the coordinates  $(z, y)$ . Writing  $\tilde{\omega}(t, z, y) = f(t, z, v) = \omega(t, x, y)$  and  $\tilde{\psi}(t, z, y) = \phi(t, z, v) = \psi(t, x, y)$ , one derives from (1.1), as in §2.2, that

$$\partial_t \tilde{\omega} + \nabla_{z,y}^\perp P_{\neq 0} \tilde{\psi} \cdot \nabla_{z,y} \tilde{\omega} = \nu \partial_{zz} \tilde{\omega} + \nu (\partial_y - t \partial_y U_0^x \partial_z)^2 \tilde{\omega}. \quad (2.70)$$

Similarly, denoting  $\tilde{U}(t, z, y) = \tilde{u}(t, z, v) = U^x(t, x, y)$ , we deduce from (1.2)

$$\partial_t U_0^x + \langle \nabla_{z,y}^\perp P_{\neq 0} \tilde{\psi} \cdot \nabla_{z,y} \tilde{U} \rangle = \nu \partial_{yy} U_0^x. \quad (2.71)$$

We may follow the argument in §2.3 of [9], apply the appropriate Gevrey inverse function theorem [9, Lemma A.5] and the Gevrey composition inequality [9, Lemma A.4] to deduce from (2.69) and (7.6) below that the estimates on  $f, \phi$  and  $\tilde{u}$  imply estimates on  $\tilde{\omega}, \tilde{\psi}$  and  $\tilde{U}$ . The main issue is inverting the coordinate transform  $y = y(t, v)$  and deducing good Gevrey regularity estimates on  $v(t, v) - y$  and  $y(t, v) - v$ . The only difference from [9] is in establishing the  $L^2$  control on  $v(t, y) - y$ , which here requires a straightforward estimate on the forced heat equation (2.36b). We omit the details for brevity and conclude that, for  $\epsilon$  sufficiently small, there exists some  $\lambda_\infty''' \in (\lambda', \lambda_\infty)$  such that

$$\|\tilde{\omega}(t)\|_{\mathcal{G}^{\lambda_\infty'''} }^2 + \langle \nu t^3 \rangle^{2\alpha} \|P_{\neq 0} \tilde{\omega}(t)\|_{\mathcal{G}^{\lambda_\infty'''} }^2 + \langle \nu t^3 \rangle^{2\alpha} \langle t \rangle^4 \|P_{\neq 0} \tilde{\psi}(t)\|_{\mathcal{G}^{\lambda_\infty'''} }^2 + \langle \nu t \rangle^{1/2} \|\tilde{\omega}(t)\|_2^2 \lesssim \epsilon^2 \quad (2.72a)$$

$$\|\tilde{U}_0\|_{\mathcal{G}^{\lambda_\infty'''} }^2 + \langle t \rangle^2 \langle \nu t^3 \rangle^{2\alpha} \|P_{\neq 0} \nabla_{z,y} \tilde{U}(t)\|_{\mathcal{G}^{\lambda_\infty'''} }^2 \lesssim \epsilon^2. \quad (2.72b)$$

In view of the definitions (2.36) and (1.8), estimate (2.72) above implies the claimed bounds on the vorticity (1.7), (1.10a) and (1.11). The  $L^2$  estimates (1.9b) and (1.10b) follow from shifting in the  $z$  variable and using the decay estimates on  $\psi$  in (2.72). From (2.70) and (2.72) together, we further deduce (1.9a). Similarly, (2.71) and (2.72) imply (1.9c). This completes the proof of Theorem 1.

### 3 Properties of the norms

In this section we detail the properties of the two multipliers  $A$  and  $A^\nu$ .

#### 3.1 Inviscid multiplier $A$

In §3 of [9], the corrector  $J$  is designed to deal with a possible frequency cascade caused by hydrodynamic echoes, the weakly nonlinear manifestation of the Orr mechanism [15, 46]. We briefly recall the definition here. For a constant  $\mu > 0$  fixed below, we define

$$J_k(t, \eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t, \eta)} + e^{\mu|k|^{1/2}}, \quad (3.1)$$

for  $w_k(t, \eta)$  specified next as an estimated ‘worst possible growth’. First recall the notation for critical intervals and times in §1.2. For the remainder of §3.1 assume for notational simplicity that  $\eta > 0$ . Finally, define  $\kappa \in (0, 1/2)$  a fixed constant.

Let  $w_{NR}$  be a non-decreasing function of time with  $w_{NR}(t, \eta) = 1$  for  $t \geq 2\eta$ . Further, set  $w_{NR}(t, \eta) = w_{NR}(t, 10)$  for  $|\eta| < 10$ . For  $k \geq 1$ , we assume that  $w_{NR}(t_{k-1, \eta})$  was computed and for  $k = 1, 2, 3, \dots, E(\sqrt{\eta})$ , we define

$$w_{NR}(t, \eta) = \left( \frac{k^2}{\eta} \left[ 1 + b_{k, \eta} \left| t - \frac{\eta}{k} \right| \right] \right)^{C\kappa} w_{NR}(t_{k-1, \eta}), \quad \forall t \in I_{k, \eta}^R = \left[ \frac{\eta}{k}, t_{k-1, \eta} \right], \quad (3.2a)$$

$$w_{NR}(t, \eta) = \left( 1 + a_{k, \eta} \left| t - \frac{\eta}{k} \right| \right)^{-1-C\kappa} w_{NR} \left( \frac{\eta}{k} \right), \quad \forall t \in I_{k, \eta}^L = \left[ t_{k, \eta}, \frac{\eta}{k} \right]. \quad (3.2b)$$

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

$$b_{k, \eta} = \frac{2(k-1)}{k} \left( 1 - \frac{k^2}{\eta} \right) \quad (3.3)$$

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

$$a_{k, \eta} = \frac{2(k+1)}{k} \left( 1 - \frac{k^2}{\eta} \right). \quad (3.4)$$

Finally, we take  $w_{NR}$  to be constant on the interval  $[0, t_{E(\sqrt{\eta}), \eta}]$ , namely  $w_{NR}(t, \eta) = w(t_{E(\sqrt{\eta}), \eta}, \eta)$  for  $t \in [0, t_{E(\sqrt{\eta}), \eta}]$ .

On each interval  $I_{k, \eta}$ , we define  $w_R(t, \eta)$  by

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

$$w_R(t, \eta) = \frac{k^2}{\eta} \left( 1 + a_{k, \eta} \left| t - \frac{\eta}{k} \right| \right) w_{NR}(t, \eta), \quad \forall t \in I_{k, \eta}^L = \left[ t_{k, \eta}, \frac{\eta}{k} \right]. \quad (3.5b)$$

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

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

$$w_k(t, \eta) = \begin{cases} w_k(t_{E(\sqrt{\eta}), \eta}, \eta) & t < t_{E(\sqrt{\eta}), \eta} \\ w_{NR}(t, \eta) & t \in [t_{E(\sqrt{\eta}), \eta}, 2\eta] \setminus I_{k, \eta} \\ w_R(t, \eta) & t \in I_{k, \eta} \\ 1 & t \geq 2\eta. \end{cases} \quad (3.6)$$

Note that  $w_k(t, \eta)$  is Lipschitz continuous in time. This completes the construction of the  $w_k(t, \eta)$  which appears in the  $J$  defined above in (3.1).

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

$$J^R(t, \eta) = \begin{cases} e^{\mu|\eta|^{1/2}} w_R^{-1}(t_{E(\sqrt{\eta}), \eta}, \eta) + 1 & t < t_{E(\sqrt{\eta}), \eta} \\ e^{\mu|\eta|^{1/2}} w_R^{-1}(t, \eta) + 1 & t \in [t_{E(\sqrt{\eta}), \eta}, 2\eta] \\ e^{\mu|\eta|^{1/2}} + 1 & t \geq 2\eta, \end{cases}$$

$$A^R(t, \eta) = e^{\lambda(t)|\eta|^s} \langle \eta \rangle^\sigma J^R(t, \eta). \quad (3.7)$$

From (3.5) we get that  $A^R(t, \eta) \geq A_0(t, \eta)$  and since the zero frequency is always non-resonant from (3.6), near the critical times,  $A^R$  can be as much as a factor of  $|\eta|$  larger.

We next recall basic properties of  $w$ ,  $J$  and  $A$ . The proofs can be found in §3 of [9]. The first shows the origin of the requirement  $s > 1/2$ .

**Lemma 3.1** (From [9]). *There exists some constants  $\mu > 0$ ,  $C_0 > 0$  such that for all  $|\eta| \geq 1$  there holds*

$$\frac{1}{w_k(0, \eta)} = \frac{1}{w_k(t_{E(\sqrt{\eta}), \eta}, \eta)} = \frac{C_0}{\eta^{\mu/8}} e^{\frac{\mu}{2}\sqrt{\eta}} + o_{|\eta| \rightarrow \infty} \left( \frac{C_0}{\eta^{\mu/8}} e^{\frac{\mu}{2}\sqrt{\eta}} \right). \quad (3.8)$$

The next lemma emphasizes the technical advantage of having well-separated critical times.

**Lemma 3.2** (From [9]). *Let  $\xi, \eta$  be such that there exists some  $K \geq 1$  with  $\frac{1}{K} |\xi| \leq |\eta| \leq K |\xi|$  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 the following holds:*

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

The next two fundamental lemmas allow first to compare  $\partial_t w/w$  at different frequencies and then  $J$  at different frequencies in nonlinear terms.

**Lemma 3.3** (From [9]). (i) *For  $t \geq 1$ , and  $k, l, \eta, \xi$  such that  $\max(2\sqrt{|\xi|}, \sqrt{|\eta|}) < t < 2\min(|\xi|, |\eta|)$ ,*

$$\frac{\partial_t w_k(t, \eta)}{w_k(t, \eta)} \frac{w_l(t, \xi)}{\partial_t w_l(t, \xi)} \lesssim \langle \eta - \xi \rangle \quad (3.9)$$

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

$$\sqrt{\frac{\partial_t w_l(t, \xi)}{w_l(t, \xi)}} \lesssim_K \left[ \sqrt{\frac{\partial_t w_k(t, \eta)}{w_k(t, \eta)}} + \frac{|\eta|^{s/2}}{\langle t \rangle^s} \right] \langle \eta - \xi \rangle. \quad (3.10)$$



(iii) For  $t \in I_{k,\eta}$  and  $t > 2\sqrt{|\eta|}$ , we have the following with  $\tau = t - \frac{\eta}{k}$  and all  $l \in \mathbb{Z}$ ,

$$\frac{\partial_t w_l(t, \eta)}{w_l(t, \eta)} \approx \frac{1}{1 + |\tau|} \approx \frac{\partial_t w_k(t, \eta)}{w_k(t, \eta)}. \quad (3.11)$$

**Lemma 3.4** (From [9]). *In general we have*

$$\frac{J_k(t, \eta)}{J_l(t, \xi)} \lesssim \frac{|\eta|}{k^2 (1 + |t - \frac{\eta}{k}|)} e^{9\mu|k-l, \eta-\xi|^{1/2}}. \quad (3.12)$$

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

$$\frac{J_k(t, \eta)}{J_l(t, \xi)} \lesssim e^{10\mu|k-l, \eta-\xi|^{1/2}}. \quad (3.13)$$

Finally if  $t \in I_{l,\xi}$ ,  $t \notin I_{k,\eta}$  and  $\frac{1}{K}|\xi| \leq |\eta| \leq K|\xi|$  for some  $K \geq 1$  then

$$\frac{J_k(t, \eta)}{J_l(t, \xi)} \lesssim \frac{l^2 \left(1 + \left|t - \frac{\xi}{l}\right|\right)}{|\xi|} e^{11\mu|k-l, \eta-\xi|^{1/2}}. \quad (3.14)$$

**Remark 5.** If  $t \in \mathbf{I}_{k,\eta} \cap \mathbf{I}_{k,\xi}$ ,  $k \neq l$ , then

$$\frac{J_k(t, \eta)}{J_l(t, \xi)} \lesssim \frac{|\eta|}{k^2} \sqrt{\frac{\partial_t w_k(t, \eta)}{w_k(t, \eta)}} \sqrt{\frac{\partial_t w_l(t, \xi)}{w_l(t, \xi)}} e^{20\mu|k-l, \eta-\xi|^{1/2}}. \quad (3.15)$$

While  $A$  does not define an algebra (by design), it does when restricted to the zero frequency in  $z$ , as is proved in the next lemma from [9].

**Lemma 3.5** (Product lemma (from [9])). *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  $v$ ,*

$$\left\| |\partial_v|^\alpha \langle \partial_v \rangle^\beta A(fg) \right\|_2 \lesssim \|f\|_{\mathcal{G}^{c\lambda, \sigma}} \left\| |\partial_v|^\alpha \langle \partial_v \rangle^\beta Ag \right\|_2 + \|g\|_{\mathcal{G}^{c\lambda, \sigma}} \left\| |\partial_v|^\alpha \langle \partial_v \rangle^\beta Af \right\|_2 \quad (3.16a)$$

$$\begin{aligned} \left\| \sqrt{\frac{\partial_t w}{w}} \langle \partial_v \rangle^\beta A(fg) \right\|_2 &\lesssim \|g\|_{\mathcal{G}^{c\lambda, \sigma}} \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \langle \partial_v \rangle^\beta Af \right\|_2 \\ &\quad + \|f\|_{\mathcal{G}^{c\lambda, \sigma}} \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) \langle \partial_v \rangle^\beta Ag \right\|_2. \end{aligned} \quad (3.16b)$$

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

$$\left\| \langle \partial_v \rangle^\beta A(fg) \right\|_2 \lesssim \left\| \langle \partial_v \rangle^\beta Af \right\|_2 \left\| \langle \partial_v \rangle^\beta Ag \right\|_2. \quad (3.17)$$

Moreover, (3.16) and (3.17) both hold with  $A$  replaced by  $A^R$ .

**Remark 6.** Writing  $(v')^2 - 1 = (v' - 1)^2 + 2(v' - 1)$  and  $v'' = \partial_v(v' - 1) + (v' - 1)\partial_v(v' - 1)$ , by (2.59e) combined with (3.17) we have,

$$\|A^R(1 - (v')^2)\|_2 \lesssim \|A^R(1 - v')\|_2 + \|A^R(1 - v')\|_2^2 \lesssim \epsilon \quad (3.18a)$$

$$\left\| \frac{A^R}{\langle \partial_v \rangle} v'' \right\|_2 = \left\| \frac{A^R}{\langle \partial_v \rangle} (v' \partial_v v') \right\|_2 \lesssim \|A^R(1 - v')\|_2 + \|A^R(1 - v')\|_2^2 \lesssim \epsilon. \quad (3.18b)$$

### 3.2 Enhanced dissipation multiplier $A^\nu$

In this section we focus on the relevant properties of the enhanced dissipation multiplier and the associated semi-norm. The following lemma summarizes the properties of  $D(t, \eta)$  (defined in (2.57)). The point is that  $D(t, \eta)$  essentially trades regularity for time-decay when  $t \geq 2|\eta|$  but we need to ensure that it does so in a way that will define a norm suitable for use in nonlinear estimates.

**Lemma 3.6** (Properties of  $D(t, \eta)$ ). *Uniformly in  $\eta, \xi, t$  and  $\nu$  we have:*

(a) *the lower bound*

$$\max\left(\nu|\eta|^3, \nu t^3\right) \lesssim \alpha D(t, \eta); \quad (3.19)$$

(b) *the ratio estimate:*

$$\frac{\langle D(t, \eta) \rangle}{\langle D(t, \xi) \rangle} \lesssim \langle \eta - \xi \rangle^3; \quad (3.20)$$

(c) *the difference estimate: if there is some  $K \geq 1$ , such that  $\frac{1}{K}|\xi| \leq |\eta| \leq K|\xi|$ , then*

$$|\langle D(t, \eta) \rangle^\alpha - \langle D(t, \xi) \rangle^\alpha| \lesssim_{\alpha, K} \frac{\langle D(t, \xi) \rangle^\alpha}{\langle \xi \rangle} \langle \eta - \xi \rangle^{3\alpha}. \quad (3.21)$$

*Proof.* First, (3.19) is immediate from (2.57).

For (3.20), write

$$\begin{aligned} \frac{\langle D(t, \eta) \rangle}{\langle D(t, \xi) \rangle} &= \frac{\left\langle \frac{1}{3\alpha}\nu|\eta|^3 + \frac{1}{24\alpha}\nu\left(t^3 - 8|\eta|^3\right)_+ \right\rangle}{\left\langle \frac{1}{3\alpha}\nu|\xi|^3 + \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3\right)_+ \right\rangle} \\ &\lesssim \langle \eta - \xi \rangle^3 + \frac{\left\langle \frac{1}{24\alpha}\nu\left(t^3 - 8|\eta|^3\right)_+ \right\rangle}{\left\langle \frac{1}{3\alpha}\nu|\xi|^3 + \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3\right)_+ \right\rangle}. \end{aligned}$$

If  $t \leq 2|\eta|$  then we have shown (3.20), so assume otherwise. Then,

$$\begin{aligned} \frac{\langle D(t, \eta) \rangle}{\langle D(t, \xi) \rangle} &\lesssim \langle \eta - \xi \rangle^3 + \frac{\left\langle \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3 + 8|\xi|^3 - 8|\eta|^3\right) \right\rangle}{\left\langle \frac{1}{3\alpha}\nu|\xi|^3 + \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3\right)_+ \right\rangle} \\ &\lesssim \langle \eta - \xi \rangle^3 + \frac{\left\langle \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3\right) \right\rangle}{\left\langle \frac{1}{3\alpha}\nu|\xi|^3 + \frac{1}{24\alpha}\nu\left(t^3 - 8|\xi|^3\right)_+ \right\rangle}. \end{aligned}$$

If  $t \geq 2|\xi|$  then (3.20) follows, so assume otherwise. Hence, the only case left is  $2|\eta| \leq t \leq 2|\xi|$ . However in this case, we get that (3.20) follows by  $|t - 2|\xi|| \leq 2||\eta| - |\xi|| \leq 2|\eta - \xi|$ . This covers all cases.

Finally turn to (3.21). First, from the mean value theorem and (3.20),

$$\begin{aligned}
|\langle D(t, \eta) \rangle^\alpha - \langle D(t, \xi) \rangle^\alpha| &= \left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle^\alpha - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle^\alpha \right| \\
&\lesssim \alpha \sup_{\xi^* \in [\eta, \xi]} \langle D(t, \xi^*) \rangle^{\alpha-1} \left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle \right| \\
&\lesssim \alpha \langle \eta - \xi \rangle^{3\alpha-3} \langle D(t, \xi) \rangle^{\alpha-1} \left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle \right|.
\end{aligned}$$

If  $t \geq 2 \max(|\eta|, |\xi|)$  then we just have

$$\left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle \right| = 0.$$

If  $t \leq 2 \min(|\eta|, |\xi|)$  then the factor reduces to

$$\left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle \right| = \left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 \right\rangle \right|.$$

This is essentially a statement about Sobolev regularity, and it follows from (3.19) that

$$\begin{aligned}
\left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 \right\rangle \right| &\lesssim \left( \langle \nu |\eta|^2 \rangle + \langle \nu |\xi|^2 \rangle \right) |\eta - \xi| \\
&\lesssim \frac{\langle D(t, \xi) \rangle}{\langle \xi \rangle} \langle \eta - \xi \rangle^2 |\eta - \xi|.
\end{aligned}$$

If  $2|\xi| < t < 2|\eta|$  then by (3.19) and (3.20) we have

$$\begin{aligned}
\left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\eta|^3)_+ \right\rangle - \left\langle \frac{1}{3\alpha} \nu |\xi|^3 + \frac{1}{24\alpha} \nu (t^3 - 8|\xi|^3)_+ \right\rangle \right| \\
= \left| \left\langle \frac{1}{3\alpha} \nu |\eta|^3 \right\rangle - \left\langle \frac{1}{3\alpha} \nu t^3 \right\rangle \right| \\
\lesssim \frac{1}{\alpha} \left( \langle \nu |\eta|^2 \rangle + \langle \nu t^2 \rangle \right) |\eta - t| \\
\lesssim \frac{\langle D(t, \xi) \rangle}{\langle \xi \rangle} |\eta - \xi|.
\end{aligned}$$

Since (3.20) is symmetric in  $\xi$  and  $\eta$ , this completes the proof of (3.20).  $\square$

From Lemma 3.6, and the choice of  $\beta$  and  $\sigma$ , we may deduce the following important product inequality about  $A^\nu$ . The lemma allows us to deduce the correct time decay from products of functions, only one of which might be decaying. The important detail to notice is the loss of  $3\alpha$  derivatives on the first factor in (3.22). The latter inequality, (3.23), is not directly used in the proof of Theorem 1 but regardless, the fact that it holds is important to consider when examining the main difficulties in the proof of (2.67) in §6.

**Lemma 3.7** ( $A^\nu$  Product Lemma). *The following holds for all  $q^1$  and  $q^2$  such that  $P_{\neq 0} q^2 = q^2$ ,*

$$\|A^\nu(q^1 q^2)\|_2 \lesssim \|q^1\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \|A^\nu q^2\|_2 \quad (3.22)$$

*If in addition we have  $P_{\neq 0} q^1 = q^1$  then it follows that*

$$\|A^\nu(q^1 q^2)\|_2 \lesssim \frac{1}{\langle \nu t^3 \rangle^\alpha} \|A^\nu q^1\|_2 \|A^\nu q^2\|_2. \quad (3.23)$$

*Proof.* First we prove (3.22). Expand via an inhomogeneous paraproduct

$$\begin{aligned}
A^\nu(\widehat{q^1 q^2})(k, \eta) &= \frac{1}{2\pi} \sum_{N \geq 8} \sum_{l \in \mathbb{Z}} \int_{\xi} A_k^\nu(\eta) \widehat{q^1}_l(\xi)_{<N/8} \widehat{q^2}_{k-l}(\eta - \xi)_N d\xi \\
&\quad + \frac{1}{2\pi} \sum_{N \geq 8} \sum_{l \in \mathbb{Z}} \int_{\xi} A_k^\nu(\eta) \widehat{q^1}_l(\xi)_N \widehat{q^2}_{k-l}(\eta - \xi)_{<N/8} d\xi \\
&\quad + \frac{1}{2\pi} \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} \sum_{l \in \mathbb{Z}} \int_{\xi} A_k^\nu(\eta) \widehat{q^1}_l(\xi)_{N'} \widehat{q^2}_{k-l}(\eta - \xi)_N d\xi \\
&= T_{LH} + T_{HL} + T_{\mathcal{R}}.
\end{aligned}$$

Consider first  $T_{LH}$ . Since on the support of the integrand (see §A.1),

$$||k, \eta| - |k - l, \eta - \xi|| \leq |l, \xi| \leq \frac{3}{16} |k - l, \eta - \xi|, \quad (3.24a)$$

$$\frac{13}{16} |k - l, \eta - \xi| \leq |k, \eta| \leq \frac{19}{16} |k - l, \eta - \xi|, \quad (3.24b)$$

inequalities (A.8) and (3.20) imply that for some  $c \in (0, 1)$  (depending only on  $s$  and our Littlewood-Paley conventions),

$$|T_{LH}| \lesssim \sum_{l \in \mathbb{Z}: l \neq k} \sum_{N \geq 8} \int_{\xi} \langle l, \xi \rangle^{3\alpha} e^{c\lambda|l, \xi|^s} \left| \widehat{q^1}_l(\xi)_{<N/8} \right| \langle k - l, \eta - \xi \rangle^\beta e^{\lambda|k-l, \eta-\xi|^s} \langle D(\eta - \xi) \rangle^\alpha \left| \widehat{q^2}_{k-l}(\eta - \xi)_N \right| d\xi.$$

Hence, by (A.4),  $\beta > 1$  and the almost orthogonality (A.2), we have

$$\begin{aligned}
\|T_{LH}\|_2^2 &\lesssim \sum_{N \geq 8} \left\| q_{<N/8}^1 \right\|_{\mathcal{G}^{c\lambda, 3\alpha+\beta}}^2 \|A^\nu q_N^2\|_2^2 \\
&\lesssim \|q^1\|_{\mathcal{G}^{\lambda, 3\alpha+\beta}}^2 \|A^\nu q^2\|_2^2.
\end{aligned} \quad (3.25)$$

Turn next to  $T_{HL}$ . On the support of the integrand we have,

$$||k, \eta| - |l, \xi|| \leq |k - l, \eta - \xi| \leq \frac{3}{16} |l, \xi|, \quad (3.26a)$$

$$\frac{13}{16} |l, \xi| \leq |k, \eta| \leq \frac{19}{16} |l, \xi|, \quad (3.26b)$$

which implies that, by (A.8) and (3.20), for some  $c \in (0, 1)$ , there holds

$$|T_{HL}| \lesssim \sum_{l \in \mathbb{Z}} \sum_{N \geq 8} \int_{\xi} \left| \langle l, \xi \rangle^{\beta+3\alpha} e^{\lambda|l, \xi|^s} \widehat{q^1}_l(\xi)_N \langle D(\eta - \xi) \rangle^\alpha \widehat{q^2}_{k-l}(\eta - \xi)_{<N/8} e^{c\lambda|k-l, \eta-\xi|^s} \right| d\xi.$$

Therefore, by (A.2), (A.4) and  $\sigma \geq \beta + 3\alpha$ ,  $\beta > 1$ ,

$$\begin{aligned}
\|T_{HL}\|_2^2 &\lesssim \sum_{N \geq 8} \|q_N^1\|_{\mathcal{G}^{\lambda, 3\alpha+\beta}}^2 \left\| \langle D(t, \partial_v) \rangle^\alpha q_{<N/8}^2 \right\|_{\mathcal{G}^{c\lambda, \beta}}^2 \\
&\lesssim \|q^1\|_{\mathcal{G}^{\lambda, 3\alpha+\beta}}^2 \|A^\nu q^2\|_2^2.
\end{aligned} \quad (3.27)$$

This completes the treatment of  $T_{HL}$ .

Finally turn to the remainder term  $T_{\mathcal{R}}$ . On the support of the integrand, there holds  $|l, \xi| \approx |k - l, \eta - \xi|$  and therefore by (A.10) and (3.20), there is some  $c \in (0, 1)$  such that

$$\begin{aligned} |T_{\mathcal{R}}| &\lesssim \sum_{N \in \mathbb{D}} \sum_{N \approx N'} \sum_{l \in \mathbb{Z}} \int_{\xi} \langle k, \eta \rangle^{\beta} \langle \xi \rangle^{3\alpha} e^{c|l, \xi|^s} \left| \widehat{q}_l^1(\xi)_{N'} e^{c|k-l, \eta-\xi|^s} \langle D(\eta - \xi) \rangle^{\alpha} \widehat{q}_{k-l}^2(\eta - \xi)_N \right| d\xi \\ &\lesssim \sum_{N \in \mathbb{D}} \sum_{N \approx N'} \sum_{l \in \mathbb{Z}} \int_{\xi} N^{-1} \langle k - l, \eta - \xi \rangle \langle l, \xi \rangle^{3\alpha+\beta} e^{c|l, \xi|^s} \left| \widehat{q}_l^1(\xi)_{N'} e^{c|k-l, \eta-\xi|^s} \langle D(\eta - \xi) \rangle^{\alpha} \widehat{q}_{k-l}^2(\eta - \xi)_N \right| d\xi \end{aligned}$$

Therefore, by (A.4) ( $\beta > 2$ ),

$$\begin{aligned} \|T_{\mathcal{R}}\|_2 &\lesssim \sum_{N \in \mathbb{D}} N^{-1} \|q^1\|_{\mathcal{G}^{e\lambda, 3\alpha+\beta}} \|A^{\nu} q^2\|_2 \\ &\lesssim \|q^1\|_{\mathcal{G}^{\lambda, 3\alpha+\beta}} \|A^{\nu} q^2\|_2. \end{aligned} \quad (3.28)$$

Upon taking square roots in (3.25) and (3.27) and combining with (3.28), this completes the proof of (3.22).

The proof of (3.23) is a slight variant of the proof of (3.22) except now the multiplier  $A^{\nu}$  is always passed to the ‘high frequency’ factor in the paraproduct and (2.58) is used on the ‘low frequency’ factor to introduce the additional decay. As (3.23) is not actually used in the proof of Theorem 1, we omit the details for brevity.  $\square$

## 4 Elliptic estimates

The following easy, but fundamental, lemma from [9] shows that by paying regularity, one can still deduce the same decay from  $\Delta_t^{-1}$  as from  $\Delta_L^{-1}$ . The loss of three derivatives comes from the presence of  $v'' = v' \partial_v v'$  in the coefficients of  $\Delta_t$ .

**Lemma 4.1** (Lossy elliptic estimate (from [9])). *Under the bootstrap hypotheses, for  $\epsilon$  sufficiently small,*

$$\|P_{\neq 0} \phi(t)\|_{\mathcal{G}^{\lambda(t), \sigma-3}} \lesssim \frac{\|f(t)\|_{\mathcal{G}^{\lambda(t), \sigma-1}}}{\langle t \rangle^2}. \quad (4.1)$$

The following lemma quantifies the rapid decay of the  $z$ -dependent velocity field due to the enhanced dissipation, crucial to the proof of (2.67) in §6. The lemma is ‘lossy’ for two reasons. First, due to the presence of  $A$  on the RHS of (4.2), and second due to the hidden loss of  $3\alpha$  derivatives on the coefficients of  $\Delta_t$  induced by (3.22). From (2.57) we see that for high frequencies  $D$  simply appears as differentiation and Lemma 4.2 is reduced to Lemma 4.1. For a given frequency, by the time  $D$  begins to increase, it is already past the critical time and we do not need to pay regularity for decay in  $\Delta_t^{-1}$ .

**Lemma 4.2** (Lossy elliptic estimate for  $A^{\nu}$ ). *Under the bootstrap hypotheses for  $\epsilon$  sufficiently small, for  $\sigma \geq \beta + 3\alpha + 4$  there holds*

$$\left\| A^{\nu} (\nabla^{\perp} P_{\neq 0} \phi) \right\|_2 \lesssim \frac{1}{\langle t \rangle^2} (\|A^{\nu} f\|_2 + \|Af\|_2). \quad (4.2)$$

*Proof.* We define the Fourier multipliers

$$\widehat{(\mathbf{1}_{t \leq 2|\partial_v|} f)}_k(\eta) = \mathbf{1}_{t \leq 2|\eta|} \widehat{f}_k(\eta), \quad (4.3)$$

$$\widehat{(\mathbf{1}_{t > 2|\partial_v|} f)}_k(\eta) = \mathbf{1}_{t > 2|\eta|} \widehat{f}_k(\eta), \quad (4.4)$$

and write

$$\left\| A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2 \leq \left\| \mathbf{1}_{t \leq 2|\partial_v|} A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2 + \left\| \mathbf{1}_{t > 2|\partial_v|} A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2. \quad (4.5)$$

Notice that if  $t \leq 2|\eta|$  there holds for  $\sigma \geq \beta + 3\alpha + 4$ ,

$$A_k^\nu(t, \eta) \approx e^{\lambda|k, \eta|^s} \langle k, \eta \rangle^{\beta+3\alpha} \lesssim e^{\lambda|k, \eta|^s} \langle k, \eta \rangle^{\sigma-4},$$

and therefore, by Lemma 4.1,

$$\left\| \mathbf{1}_{t \leq 2|\partial_v|} A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2 \lesssim \left\| \nabla^\perp P_{\neq 0} \phi \right\|_{\mathcal{G}^{\lambda, \sigma-4}} \lesssim \langle t \rangle^{-2} \|Af\|_2. \quad (4.6)$$

Next, focus on lower frequencies:

$$\begin{aligned} \left\| \mathbf{1}_{t > 2|\partial_v|} A^\nu P_{\neq 0} \phi \right\|_2^2 &= \sum_{k \neq 0} \int_{\eta} \mathbf{1}_{t > 2|\eta|} e^{2\lambda|k, \eta|^s} \langle k, \eta \rangle^{2\beta} \langle D(\eta) \rangle^\alpha \left| \hat{\phi}_k(\eta) \right|^2 d\eta \\ &= \sum_{k \neq 0} \int_{\eta} \mathbf{1}_{t > 2|\eta|} e^{2\lambda|(k, \eta)|^s} \frac{\langle k, \eta \rangle^{2\beta} \langle D(\eta) \rangle^\alpha}{(k^2 + |\eta - kt|^2)^2} (k^2 + |\eta - kt|^2)^2 \left| \hat{\phi}_k(\eta) \right|^2 d\eta \\ &\lesssim \frac{1}{\langle t \rangle^4} \left\| \Delta_L P_{\neq 0} A^\nu \phi \right\|_2^2. \end{aligned} \quad (4.7)$$

As in the proof of Lemma 4.1 (see [9]), we write  $\Delta_t$  as a perturbation of  $\Delta_L$  via

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

Applying (3.22) implies

$$\left\| \Delta_L A^\nu P_{\neq 0} \phi \right\|_2 \lesssim \|A^\nu f\|_2 + \left\| 1 - (v')^2 \right\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \left\| \Delta_L P_{\neq 0} A^\nu \phi \right\|_2 + \|v''\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \|(\partial_y - t\partial_z) P_{\neq 0} A^\nu \phi\|_2.$$

Therefore, (3.18) and  $\sigma \geq \beta + 3\alpha + 1$  imply

$$\left\| \Delta_L A^\nu P_{\neq 0} \phi \right\|_2 \lesssim \|A^\nu f\|_2 + \epsilon \left\| \Delta_L P_{\neq 0} A^\nu \phi \right\|_2.$$

Together with (4.7), for  $\epsilon$  is sufficiently small we get

$$\left\| \mathbf{1}_{t > 2|\partial_v|} A^\nu P_{\neq 0} \phi \right\|_2 \lesssim \langle t \rangle^{-2} \|A^\nu f\|_2,$$

which, with (4.6) and (4.5), completes the proof of Lemma 4.2.  $\square$

## 5 High norm vorticity estimate (2.66a)

We are now ready to begin the proof of (2.66a). Computing the evolution of  $Af$  from (2.48) gives,

$$\frac{d}{dt} \frac{1}{2} \|Af\|_2^2 = -CK_\lambda - CK_w - \int Af A(u \cdot \nabla f) dv dz + \nu \int Af A(\tilde{\Delta}_t f) dv dz. \quad (5.1)$$

Treating the third term, the Euler nonlinearity, comprises the majority of the work in [9]. Here, we need only replace the role of  $[\partial_t v]$  in [9] with  $g$  here and proceed in the same manner as used in §5, §6 and §7 of [9] to deduce (under the bootstrap hypotheses for  $\epsilon$  sufficiently small) that

$$\begin{aligned} - \int Af A(u \cdot \nabla f) dv dz &\lesssim \epsilon CK_\lambda + \epsilon CK_w + \epsilon CK_\lambda^{v,1} + \epsilon CK_w^{v,1} \\ &\quad + \epsilon^3 \left( \sum_{i=1}^2 CCK_\lambda^i + CCK_w^i \right) + \frac{\epsilon^3}{\langle t \rangle^{1+s}}. \end{aligned} \quad (5.2)$$

The new difficulty in deducing (2.66a) from (5.1) is commuting  $A$  and  $\tilde{\Delta}_t$  in the last term. We follow an approach which is consistent with the proofs of the elliptic estimates in §4 and [9]. Following those arguments we write,

$$\begin{aligned}
\nu \int AfA(\tilde{\Delta}_t f) &= \nu \int AfA(\Delta_L f) - \nu \int AfA[(1 - (v')^2)(\partial_v - t\partial_z)^2 f] dv dz \\
&= -\nu \left\| \sqrt{-\Delta_L} Af \right\|_2^2 - \nu \sum_{k \neq 0} \int Af_k A[(1 - (v')^2)(\partial_v - t\partial_z)^2 f_k] dv dz \\
&\quad - \nu \int Af_0 A[(1 - (v')^2)\partial_v^2 f_0] dv dz \\
&= -\nu \left\| \sqrt{-\Delta_L} Af \right\|_2^2 + E^\neq + E^0.
\end{aligned} \tag{5.3}$$

Note that while the first term looks like a powerful dissipation term ( $\sqrt{-\Delta_L}$  contains powers of  $t$ ), since we are sending  $\nu \rightarrow 0$  it can only really be useful for controlling the error terms  $E^\neq$  and  $E^0$ . Especially  $E^\neq$  is dangerous as while the leading order dissipation degenerates near the critical times, we will see that  $E^\neq$  can still be large. Since the operator  $\tilde{\Delta}_t$  is very anisotropic between non-zero frequencies in  $z$  and the zero frequencies, the two cases,  $E^0$  and  $E^\neq$ , are treated separately by different methods in §5.1 and §5.2 respectively.

### 5.1 Dissipation error term: non-zero frequencies

For future notational convenience, use the short hand

$$G(t, v) = (1 - (v')^2)(t, v). \tag{5.4}$$

We begin the treatment of  $E^\neq$  by decomposing with an inhomogeneous paraproduct in  $v$  only:

$$\begin{aligned}
E^\neq &= \frac{\nu}{2\pi} \sum_{M \geq 8} \sum_{k \neq 0} \int A\bar{\hat{f}}_k(t, \eta) A_k(t, \eta) \hat{G}(\eta - \xi)_M |\xi - tk|^2 \hat{f}_k(\xi)_{<M/8} d\xi d\eta \\
&\quad + \frac{\nu}{2\pi} \sum_{M \geq 8} \sum_{k \neq 0} \int A\bar{\hat{f}}_k(t, \eta) A_k(t, \eta) \hat{G}(\eta - \xi)_{<M/8} |\xi - tk|^2 \hat{f}_k(\xi)_M d\xi d\eta \\
&\quad + \frac{\nu}{2\pi} \sum_{M \in \mathbb{D}} \sum_{M/8 \leq M' \leq 8M} \sum_{k \neq 0} \int A\bar{\hat{f}}_k(t, \eta) A_k(t, \eta) \hat{G}(\eta - \xi)_M |\xi - tk|^2 \hat{f}_k(\xi)_{M'} d\xi d\eta \\
&= E_{HL}^\neq + E_{LH}^\neq + E_{\mathcal{R}}^\neq.
\end{aligned}$$

The  $E_{HL}^\neq$  term is the easier one, so we will treat this one first. The goal is to pass the  $A_k(\eta)$  onto  $\hat{f}_k(\xi)$  and split the  $(\partial_v - t\partial_z)^2$  between the two factors of  $f$ . The latter will use the following triangle inequality, which holds specifically for  $k \neq 0$ ,

$$|\xi - kt| \leq |\eta - \xi| + |\eta - kt| \leq \langle \eta - \xi \rangle (|k| + |\eta - kt|). \tag{5.5}$$

This allows to split the derivatives by losing regularity on the coefficient, which is fine since it is in ‘low frequency’ (we are essentially integrating by parts). Note that this will not be possible in treating  $E_{HL}^\neq$ . An important point is that (3.13) (as opposed to (3.12)) applies to transfer  $A_k(\eta)$  to  $A_k(\xi)$  - resonant vs non-resonant losses can only occur when comparing different frequencies in  $z$ .

Since on the support of the integrand there holds

$$||k, \eta| - |k, \xi|| \leq |\xi - \eta| \leq \frac{3}{16} |\xi| \leq \frac{3}{16} |k, \xi|, \quad (5.6a)$$

$$\frac{13}{16} |k, \xi| \leq |k, \eta| \leq \frac{19}{16} |k, \xi|, \quad (5.6b)$$

which implies by (A.8) we have for some  $c \in (0, 1)$ ,

$$|E_{LH}^\neq| \lesssim \nu \sum_{M \geq 8} \sum_{k \neq 0} \int \left| A \hat{f}_k(t, \eta) e^{c\lambda|\eta-\xi|^s} \hat{G}(\eta - \xi)_{<M/8} \frac{J_k(\eta)}{J_k(\xi)} |\xi - tk|^2 A \hat{f}_k(\xi)_M \right| d\xi d\eta.$$

Applying (5.5) together with (3.13) followed by (A.13) implies

$$|E_{LH}^\neq| \lesssim \nu \sum_{M \geq 8} \sum_{k \neq 0} \int \left| \sqrt{-\Delta_L} A \hat{f}_k(t, \eta) e^{\lambda|\eta-\xi|^s} \langle \eta - \xi \rangle \hat{G}(\eta - \xi)_{<M/8} |\xi - tk| A \hat{f}_k(\xi)_M \right| d\xi d\eta,$$

where we are slightly abusing notation by using  $\sqrt{-\Delta_L} \hat{f}_k(\eta) = (|k| + |\eta - kt|) \hat{f}_k(\eta)$ . Therefore, by (A.5) ( $\sigma > 1$ ), almost orthogonality (A.2) and (3.18) (recall the shorthand (5.4)),

$$\begin{aligned} |E_{LH}^\neq| &\lesssim \nu \sum_{M \geq 8} \sum_{k \neq 0} \left\| (1 - (v')^2)_{<M/8} \right\|_{\mathcal{G}^{\lambda, \sigma}} \left\| \sqrt{-\Delta_L} (A f_k)_{\sim M} \right\|_2^2 \\ &\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} A f \right\|_2^2. \end{aligned} \quad (5.7)$$

This term is then absorbed by the leading order dissipation term in (5.3) for  $\epsilon$  sufficiently small.

Now, let us turn to the more delicate HL term. Since the paraproduct decomposition was with respect to  $v$  but the norm depends on both  $z$  and  $v$ , we will need to divide into separate contributions corresponding to when  $k$  is large compared to  $\eta$  and vice-versa. A similar issue arose, for example, in [9, Proposition 2.5]. Hence,

$$\begin{aligned} |E_{HL}^\neq| &\lesssim \nu \sum_{M \geq 8} \int \left[ \mathbf{1}_{|k| \geq \frac{1}{16}|\eta|} + \mathbf{1}_{|k| < \frac{1}{16}|\eta|} \right] \left| A \hat{f}_k(\eta)_{\sim M} A_k(t, \eta) \hat{G}(\eta - \xi)_M |\xi - tk|^2 \hat{f}_k(\xi)_{<M/8} \right| d\xi d\eta \\ &= E_{HL}^{\neq, z} + E_{HL}^{\neq, v}. \end{aligned}$$

In the ‘ $z$ ’ case, we can assume that the ‘derivatives’ are still landing on  $f_k(\xi)$ , and hence we may treat this term in a manner similar to  $E_{LH}^\neq$ . On the support of the integrand, we claim that there is some  $c \in (0, 1)$  such that,

$$|k, \eta|^s \leq |k, \xi|^s + c |\eta - \xi|^s. \quad (5.8)$$

To see (5.8), one can consider separately the cases  $\frac{1}{16} |\eta| \leq |k| \leq 16 |\eta|$  and  $|k| > 16 |\eta|$ , applying (A.10) and (A.8) respectively. Therefore, by (5.8) we have

$$|E_{HL}^{\neq, z}| \lesssim \nu \sum_{M \geq 8} \int \mathbf{1}_{|k| \geq \frac{1}{16}|\eta|} \left| A \hat{f}_k(\eta)_{\sim M} e^{c\lambda|\eta-\xi|^s} \hat{G}(\eta - \xi)_M J_k(\eta) |\xi - tk|^2 \langle k, \eta \rangle^\sigma e^{\lambda|k, \xi|^s} \hat{f}_k(\xi)_{<M/8} \right| d\xi d\eta.$$

Moreover, on the support of the integrand we have  $\langle k, \eta \rangle \lesssim \langle k \rangle \leq \langle k, \xi \rangle$ . Therefore, by (5.5) and (3.13) followed by (A.13) ( $c < 1$  and  $s > 1/2$ ) we get

$$|E_{HL}^{\neq, z}| \lesssim \nu \sum_{M \geq 8} \int \mathbf{1}_{|k| \geq \frac{1}{16}|\eta|} \left| \sqrt{-\Delta_L} A \hat{f}_k(\eta)_{\sim M} e^{\lambda|\eta-\xi|^s} \langle \eta - \xi \rangle \hat{G}(\eta - \xi)_M |\xi - tk| A \hat{f}_k(\xi)_{<M/8} \right| d\xi d\eta.$$



By (A.5) ( $\sigma > 2$ ), Cauchy-Schwarz in  $M$  and  $k$ , almost orthogonality (A.2) and (3.18) we get

$$\begin{aligned} |E_{HL}^{\neq, z}| &\lesssim \nu \sum_{M \geq 8} \sum_{k \neq 0} \|(1 - (v')^2)_M\|_{\mathcal{G}^{\lambda, \sigma}} \left\| \sqrt{-\Delta_L}(Af_k)_{\sim M} \right\|_2 \left\| \sqrt{-\Delta_L}(Af_k)_{< M/8} \right\|_2 \\ &\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} Af \right\|_2^2 \end{aligned} \quad (5.9)$$

This term is then absorbed by the leading order dissipation term in (5.3) for  $\epsilon$  sufficiently small.

Turn now to the more challenging  $v$  case,  $E_{HL}^{\neq, v}$ , which corresponds to  $|k, \eta| \approx |\eta - \xi| \gg |k, \xi|$ . The challenge here is that we cannot simply use (5.5), as this requires more regularity on the coefficients than we have to spend. Hence we will have to find another way of controlling this term and will find it most challenging near the critical times, as there the leading order dissipation could become weaker than the error term. Indeed, using that  $|\xi| \leq \frac{3}{16} |\eta - \xi|$  on the support of the integrand, analogous to (5.6) there holds on the support of the integrand:

$$|\eta - \xi| - |k, \eta| \leq |k, \xi| \leq \frac{1}{16} |\eta| + |\xi| \leq \frac{1}{16} |\eta - \xi| + \frac{17}{16} |\xi| \leq \frac{67}{256} |\eta - \xi|. \quad (5.10)$$

Hence we may apply (A.8) to show there exists some  $c \in (0, 1)$  such that

$$|E_{HL}^{\neq, v}| \lesssim \nu \sum_{M \geq 8} \int \mathbf{1}_{|k| < \frac{1}{16} |\eta|} \left| A \hat{f}_k(\eta)_{\sim M} A \hat{G}(\eta - \xi)_M \frac{J_k(\eta)}{J_0(\eta - \xi)} e^{c\lambda|k, \xi|^s} |\xi - tk|^2 \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.$$

Near the resonant times, the treatment will be different, so divide further into resonant and non-resonant contributions:

$$\begin{aligned} |E_{HL}^{\neq, v}| &\lesssim \nu \sum_{M \geq 8} \int [\chi^R + \chi^*] \left| A \hat{f}_k(\eta)_{\sim M} A \hat{G}(\eta - \xi)_M \frac{J_k(\eta)}{J_0(\eta - \xi)} e^{c\lambda|k, \xi|^s} |\xi - tk|^2 \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta \\ &= E_{HL}^{\neq, v, R} + E_{HL}^{\neq, v, *}, \end{aligned}$$

where  $\chi^R = \mathbf{1}_{t \in \mathbf{I}_{k, \eta - \xi} \cap \mathbf{I}_{k, \eta}} \mathbf{1}_{|k| < \frac{1}{16} |\eta|}$  and  $\chi^* = (1 - \mathbf{1}_{t \in \mathbf{I}_{k, \eta - \xi} \cap \mathbf{I}_{k, \eta}}) \mathbf{1}_{|k| < \frac{1}{16} |\eta|}$ .

Turn first to the ‘R’ contribution. Applying (3.12) followed by (A.13) implies (recall definitions (3.7) and (3.5)),

$$E_{HL}^{\neq, v, R} \lesssim \nu \sum_{k \neq 0} \sum_{M \geq 8} \int \chi^R \left| A \hat{f}_k(\eta)_{\sim M} A^R(\eta - \xi) \hat{G}(\eta - \xi)_M |\xi - tk|^2 e^{\lambda|k, \xi|^s} \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.$$

Since we are near the resonant time, the leading dissipation term in (5.3) is very weak, and in particular, it cannot directly recover the  $t^2$  present in the low frequency factor. In order to deal with this fundamental difficulty, we will use both the enhanced dissipation estimate (2.60) as well as the *CCK* control on  $G$  expressed in (2.59e). Using

$$|k| + |\xi - tk| \lesssim \langle t \rangle \langle k, \xi \rangle, \quad (5.11)$$

as well as (3.11), (3.10) and the definition of  $\chi^R$  we have

$$\begin{aligned}
E_{HL}^{\neq;v,R} &\lesssim \nu \sum_{k \neq 0} \sum_{M \geq 8} \int \chi^R (|k| + |\eta - kt|)^{1/2} \left| A \hat{f}_k(\eta)_{\sim M} \right| A^R(\eta - \xi) \left( \frac{1}{\sqrt{1 + \left| \frac{\eta}{k} - t \right|}} \right) \\
&\quad \times \widehat{G}(\eta - \xi)_M (|k| + |\xi - tk|)^{1/2} \langle t \rangle^{3/2} \langle k, \xi \rangle^{3/2} e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta \\
&\lesssim \nu \sum_{k \neq 0} \sum_{M \geq 8} \int \chi^R (|k| + |\eta - kt|)^{1/2} \left| A \hat{f}_k(\eta)_{\sim M} \left( \sqrt{\frac{\partial_t w(\eta - \xi)}{w(\eta - \xi)}} + \frac{|\eta - \xi|^{s/2}}{\langle t \rangle^s} \right) A^R(\eta - \xi) \right| \\
&\quad \times \widehat{G}(\eta - \xi)_M (|k| + |\xi - tk|)^{1/2} \langle t \rangle^{3/2} \langle k, \xi \rangle^{5/2} e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.
\end{aligned}$$

Therefore, by (A.14), (A.5) (using  $\beta > 4$ ) and Cauchy-Schwarz,

$$\begin{aligned}
E_{HL}^{\neq;v,R} &\lesssim \nu t^{3/2} \sum_{k \neq 0} \sum_{M \geq 8} \left\| (-\Delta_L)^{1/4} A(f_k)_{\sim M} \right\|_2 \left\| (-\Delta_L)^{1/4} f_k \right\|_{\mathcal{G}^{\lambda, \beta}} \\
&\quad \times \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2 \\
&\lesssim \nu t^{3/2} \sum_{M \geq 8} \left( \sum_{k \neq 0} \left\| (-\Delta_L)^{1/4} A(f_k)_{\sim M} \right\|_2^2 \right)^{1/2} \left( \sum_{k \neq 0} \left\| (-\Delta_L)^{1/4} f_k \right\|_{\mathcal{G}^{\lambda, \beta}}^2 \right)^{1/2} \\
&\quad \times \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2 \\
&\lesssim \nu t^{3/2} \sum_{M \geq 8} \|A f_{\sim M}\|_2^{1/2} \|f\|_{\mathcal{G}^{\lambda, \beta}}^{1/2} \left\| (-\Delta_L)^{1/2} A f_{\sim M} \right\|_2^{1/2} \left\| (-\Delta_L)^{1/2} A f \right\|_2^{1/2} \\
&\quad \times \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2.
\end{aligned}$$

Now we crucially apply the enhanced dissipation by using (2.58) and (2.60) to kill the extra powers of time,

$$\begin{aligned}
E_{HL}^{\neq;v,R} &\lesssim \frac{\nu t^{3/2}}{\langle \nu t^3 \rangle^{\alpha/2}} \sum_{M \geq 8} \|A f_{\sim M}\|_2^{1/2} \|A^\nu f\|_2^{1/2} \left\| (-\Delta_L)^{1/2} A f_{\sim M} \right\|_2^{1/2} \left\| (-\Delta_L)^{1/2} A f \right\|_2^{1/2} \\
&\quad \times \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2 \\
&\lesssim \sqrt{\nu \epsilon} \sum_{M \geq 8} \|A f_{\sim M}\|_2^{1/2} \left\| (-\Delta_L)^{1/2} A f_{\sim M} \right\|_2^{1/2} \left\| (-\Delta_L)^{1/2} A f \right\|_2^{1/2} \\
&\quad \times \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2.
\end{aligned}$$

Therefore, by (2.60), (2.59a), Cauchy-Schwarz in  $M$  and almost orthogonality (A.2),

$$\begin{aligned}
E_{HL}^{\neq;v,R} &\lesssim \nu \sum_{M \geq 8} \|Af_{\sim M}\| \left\| (-\Delta_L)^{1/2} Af_{\sim M} \right\|_2 \left\| (-\Delta_L)^{1/2} Af \right\|_2 \\
&\quad + \epsilon \sum_{M \geq 8} \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2)_M \right\|_2^2 \\
&\lesssim \epsilon \nu \left\| (-\Delta_L)^{1/2} Af \right\|_2^2 + \epsilon \left\| \left( \sqrt{\frac{\partial_t w}{w}} + \frac{|\partial_v|^{s/2}}{\langle t \rangle^s} \right) A^R (1 - (v')^2) \right\|_2^2. \tag{5.12}
\end{aligned}$$

The first term is absorbed by the dissipation term in (5.3) and the latter term is controlled by the bootstrap control on the  $CK^1$  terms in (2.59e).

The treatment of  $E_{HL}^{\neq;v,*}$  is easier as the leading order dissipation is much stronger at this set of frequencies (although we still cannot apply (5.5)). By (3.13) followed by (A.13) we have

$$E_{HL}^{\neq;v,*} \lesssim \nu \sum_{M \geq 8} \int A \left| \hat{f}_k(\eta)_{\sim M} \right| \chi^* A_0(\eta - \xi) \left| \hat{G}(\eta - \xi)_M \right| |\xi - tk|^2 e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.$$

We will introduce  $\sqrt{-\Delta_L}$  on the first factor with the goal of directly using the dissipation:

$$E_{HL}^{\neq;v,*} \lesssim \nu \sum_{M \geq 8} \int \sqrt{-\Delta_L} A \left| \hat{f}_k(\eta)_{\sim M} \right| \chi^* A_0(\eta - \xi) \left| \hat{G}(\eta - \xi)_M \right| \frac{|\xi - tk|^2}{\sqrt{k^2 + |\eta - kt|^2}} e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.$$

By the definition of  $\chi^*$  and Lemma 3.2 (using that  $\eta \approx \eta - \xi$ ), on the support of the integrand we have one of two possibilities: either  $|\eta - kt| \gtrsim t$  or  $|\xi| \gtrsim t$ . Therefore, using also (5.11),

$$E_{HL}^{\neq;v,*} \lesssim \nu \sum_{M \geq 8} \int \sqrt{-\Delta_L} A \left| \hat{f}_k(\eta)_{\sim M} \right| \chi^* A_0(\eta - \xi) \left| \hat{G}(\eta - \xi)_M \right| \langle t \rangle^{-1} |\xi - tk|^2 \langle \xi \rangle e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta.$$

Therefore, by (A.5) ( $\beta > 3$ ), Cauchy-Schwarz (in  $M$  and  $k$ ) and (3.18a) we have,

$$\begin{aligned}
E_{HL}^{\neq;v,*} &\lesssim \nu \sum_{k \neq 0} \sum_{M \geq 8} \int \left| \sqrt{-\Delta_L} A \hat{f}_k(\eta)_{\sim M} \chi^* A_0(\eta - \xi) \hat{G}(\eta - \xi)_M \right| \\
&\quad \times \langle k, \xi \rangle^2 |\xi - tk| e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{< M/8} \right| d\xi d\eta \\
&\lesssim \nu \sum_{M \geq 8} \|A(1 - (v')^2)_M\|_2 \left\| \sqrt{-\Delta_L} Af_{\sim M} \right\|_2 \left\| \sqrt{-\Delta_L} f_{< M/8} \right\|_{\mathcal{G}^{\lambda, \beta}} \\
&\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} Af \right\|_2^2. \tag{5.13}
\end{aligned}$$

Hence, for  $\epsilon$  chosen sufficiently small, this is absorbed by the dissipation term in (5.3).

Finally, turn to the remainder term  $E_{\mathcal{R}}^{\neq}$ . Dividing into two contributions depending on the relative size of the  $z$  vs  $v$  frequencies,

$$\begin{aligned}
E_{\mathcal{R}}^{\neq} &\lesssim \nu \sum_{M \in \mathbb{D}} \sum_{M/8 \leq M' \leq 8M} \sum_{k \neq 0} \int \left[ \mathbf{1}_{|k| > 100|\xi|} + \mathbf{1}_{|k| \leq 100|\xi|} \right] \left| A \hat{f}_k(t, \eta) A_k(t, \eta) \hat{G}(\eta - \xi)_M |\xi - tk|^2 \hat{f}_k(\xi)_{M'} \right| d\xi d\eta \\
&= E_{\mathcal{R}}^{\neq; z} + E_{\mathcal{R}}^{\neq; v}.
\end{aligned}$$

Consider  $E_{\mathcal{R}}^{\neq; z}$  first. On the support of the integrand,

$$||k, \eta| - |k, \xi|| \leq |\eta - \xi| \leq \frac{3M'}{2} \leq 12M \leq 24|\xi| \leq \frac{24}{100}|k, \xi|, \quad (5.14)$$

and hence inequality (A.8) implies there is some  $c \in (0, 1)$  such that

$$|E_{\mathcal{R}}^{\neq; z}| \lesssim \nu \sum_{M \in \mathbb{D}} \sum_{M \approx M'} \sum_{k \neq 0} \int \mathbf{1}_{|k| > 100|\xi|} \left| A \hat{f}_k(t, \eta) e^{c\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_M |\xi - tk|^2 \frac{J_k(\eta)}{J_k(\xi)} A \hat{f}_k(\xi)_{M'} \right| d\xi d\eta.$$

Therefore, by (5.5), (3.13), (A.13) and (A.14) (using  $c < 1$  and  $s > 1/2$ ) we have

$$|E_{\mathcal{R}}^{\neq; z}| \lesssim \nu \sum_{M \in \mathbb{D}} \sum_{M \approx M'} \sum_{k \neq 0} \int \mathbf{1}_{|k| > 100|\xi|} \left| \sqrt{-\Delta_L} A \hat{f}_k(t, \eta) e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_M |\xi - tk| A \hat{f}_k(\xi)_{M'} \right| d\xi d\eta.$$

Therefore, by (A.5) followed by Cauchy-Schwarz in  $k$  and  $M$ , almost orthogonality (A.2) and (3.18),

$$\begin{aligned} |E_{\mathcal{R}}^{\neq; z}| &\lesssim \nu \sum_{M \in \mathbb{D}} \sum_{k \neq 0} \left\| \sqrt{-\Delta_L} A f_k \right\|_2 \left\| (1 - (v')^2)_M \right\|_{\mathcal{G}^{\lambda, \sigma}} \left\| \sqrt{-\Delta_L} A(f_k)_{\sim M} \right\|_2 \\ &\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} A f \right\|_2^2, \end{aligned} \quad (5.15)$$

which can then be absorbed by the leading dissipation term in (5.3) for  $\epsilon$  sufficiently small.

Finally turn to  $E_{\mathcal{R}}^{\neq; v}$ . On the support of the integrand in this case,

$$\begin{aligned} |\eta - \xi| &\leq 24|\xi| \leq 24|k, \xi| \\ |k, \xi| &\leq 101|\xi| \leq 2424|\eta - \xi|. \end{aligned}$$

Therefore by (A.10) there exists a  $c \in (0, 1)$  such that

$$\begin{aligned} |E_{\mathcal{R}}^{\neq; v}| &\lesssim \nu \sum_{M \in \mathbb{D}} \sum_{M \approx M'} \sum_{k \neq 0} \int \mathbf{1}_{|k| \leq 100|\xi|} \left| A \hat{f}_k(\eta) e^{c\lambda|\eta - \xi|^s} \langle \xi - \eta \rangle^{\sigma/2} \widehat{G}(\eta - \xi)_M \right| \\ &\quad \times |\xi - tk|^2 J_k(\eta) e^{c\lambda|k, \xi|^s} \langle k, \xi \rangle^{\sigma/2} \left| \hat{f}_k(\xi)_{M'} \right| d\xi d\eta. \end{aligned}$$

By Lemma 3.1, (A.13), (5.5) and (A.14),

$$\begin{aligned} |E_{\mathcal{R}}^{\neq; v}| &\lesssim \nu \sum_{M \in \mathbb{D}} \sum_{M \approx M'} \sum_{k \neq 0} \int \mathbf{1}_{|k| \leq 100|\xi|} \left| \sqrt{-\Delta_L} A \hat{f}_k(t, \eta) e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_M \right| \\ &\quad \times |\xi - tk| e^{\lambda|k, \xi|^s} \left| \hat{f}_k(\xi)_{M'} \right| d\xi d\eta. \end{aligned}$$

Finally, by (A.5), Cauchy-Schwarz in  $M$  and  $k$ , almost orthogonality (A.2) and (3.18),

$$\begin{aligned} |E_{\mathcal{R}}^{\neq; v}| &\lesssim \nu \sum_{M \in \mathbb{D}} \sum_{k \neq 0} \left\| \sqrt{-\Delta_L} A f_k \right\|_2 \left\| (1 - (v')^2)_M \right\|_{\mathcal{G}^{\lambda, \sigma}} \left\| \sqrt{-\Delta_L} A(f_k)_{\sim M} \right\|_2 \\ &\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} A f \right\|_2^2, \end{aligned} \quad (5.16)$$

which can then be absorbed by the leading dissipation term in (5.3) for  $\epsilon$  sufficiently small. This completes the treatment of  $E_{\mathcal{R}}^{\neq; v}$  and of the entire non-zero frequency error term  $E^{\neq}$ .

## 5.2 Dissipation error term: zero frequencies

For treating  $E^0$  in (5.3), the main challenge is dealing with the contributions of low frequencies, where the leading dissipation term in (5.3) will not directly control  $E^0$ . Physically, there is an effective transport of enstrophy due to the motion of the coordinate system inducing a variable dissipation coefficient. This low frequency effect is controlled by (2.61b), which ensures that the coordinate system is relaxing sufficiently fast. Using again the shorthand (5.4), begin by decomposing  $E^0$  with a *homogeneous* paraproduct:

$$\begin{aligned} E^0 &= \nu \sum_{M \in 2^{\mathbb{Z}}} \int A f_0 A (G(t, v)_M \partial_{vv}(f_0)_{<M/8}) dv + \nu \sum_{M \in 2^{\mathbb{Z}}} \int A f_0 A (G(t, v)_{<M/8} \partial_{vv}(f_0)_M) dv \\ &\quad + \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int A f_0 A (G(t, v)_{M'} \partial_{vv}(f_0)_M) dv \\ &= E_{HL}^0 + E_{LH}^0 + E_{\mathcal{R}}^0. \end{aligned}$$

Consider

$$|E_{HL}^0| \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \int \left| A \hat{f}_0(\eta)_{\sim M} A_0(\eta) \widehat{G}(\eta - \xi)_M \xi^2 \hat{f}_0(\xi)_{<M/8} \right| d\eta.$$

On the support of the integrand,

$$||\eta| - |\eta - \xi|| \leq |\xi| \leq \frac{3}{16} |\eta - \xi|, \quad (5.17a)$$

$$\frac{13}{16} |\eta - \xi| \leq |\eta| \leq \frac{19}{16} |\eta - \xi|, \quad (5.17b)$$

which implies that by (A.8) and (3.13) followed by (A.13) we have,

$$|E_{HL}^0| \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \int \left| A \hat{f}_0(\eta)_{\sim M} A_0(\eta - \xi) \widehat{G}(\eta - \xi)_M \xi^2 \hat{f}_0(\xi)_{<M/8} e^{\lambda|\xi|^s} \right| d\eta.$$

Due to (5.17), on the support of the integrand there always holds  $|\xi| \lesssim |\eta|$  and therefore by (A.5), Cauchy-Schwarz in  $M$  and (3.18a) (also almost orthogonality (A.2) and  $\sigma > 1$ )

$$\begin{aligned} |E_{HL}^0| &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \int \left| \eta A \hat{f}_0(\eta)_{\sim M} A_0(\eta - \xi) \widehat{G}(\eta - \xi)_M \xi \hat{f}_0(\xi)_{<M/8} e^{\lambda|\xi|^s} \right| d\eta \\ &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \|\partial_v A(f_0)_{\sim M}\|_2 \|\partial_v(f_0)_{<M/8}\|_{\mathcal{G}^{\lambda, \sigma}} \|A(1 - (v')^2)_M\|_2 \\ &\lesssim \nu \epsilon \|\partial_v A f_0\|_2^2, \end{aligned} \quad (5.18)$$

which is then absorbed by the dissipation term in (5.3) for  $\epsilon$  sufficiently small.

The  $E_{LH}^0$  term is treated similarly. The analogue of (5.17) (with the role of  $\xi$  and  $\eta - \xi$  swapped) holds on the support of the integrand and hence from (A.8) and (3.13) followed by (A.13) ( $s > 1/2$ ), we have,

$$\begin{aligned} |E_{LH}^0| &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \int \left| A \hat{f}_0(\eta)_{\sim M} e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_{<M/8} \xi^2 A_0(\xi) \hat{f}_0(\xi)_M \right| d\eta \\ &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \int \left| \eta A \hat{f}_0(\eta)_{\sim M} e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_{<M/8} \xi A_0(\xi) \hat{f}_0(\xi)_M \right| d\eta. \end{aligned}$$

Then by (A.5), almost orthogonality (A.2), Cauchy-Schwarz in  $M$ ,  $\sigma > 1$  and (3.18a) we get,

$$\begin{aligned} |E_{LH}^0| &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \|\partial_v A(f_0)_{\sim M}\|_2 \|\partial_v A(f_0)_M\|_2 \left\| (1 - (v')^2)_{<M/8} \right\|_{\mathcal{G}^{\lambda, \sigma}} \\ &\lesssim \nu \epsilon \|\partial_v A f_0\|_2^2, \end{aligned} \quad (5.19)$$

which is again absorbed by the dissipation term in (5.3) for  $\epsilon$  sufficiently small.

Finally, consider the remainder term and divide into separate cases based on the output frequency:

$$\begin{aligned} E_{\mathcal{R}}^0 &= \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int (A f_0)_{\leq 1} A(G_M(\partial_{vv} f_0)_{M'}) dv \\ &\quad + \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int (A f_0)_{> 1} A(G_M(\partial_{vv} f_0)_{M'}) dv \\ &= E_{\mathcal{R}}^{0;L} + E_{\mathcal{R}}^{0;H}. \end{aligned}$$

The  $E_{\mathcal{R}}^{0;H}$  term is relatively straightforward. Indeed, on the support of the integrand we may apply (A.10) to deduce for some  $c \in (0, 1)$  that,

$$\begin{aligned} |E_{\mathcal{R}}^{0;H}| &\lesssim \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int_{\eta, \xi} \mathbf{1}_{|\eta| > 1} \left| A \widehat{f_0}(\eta) J_0(\eta) e^{\lambda|\eta|^s} \langle \eta - \xi \rangle^{\sigma/2} \widehat{G}(\eta - \xi)_M \langle \xi \rangle^{\sigma/2+1} \widehat{\partial_v f_0}(\xi)_{M'} \right| d\xi d\eta \\ &\lesssim \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int_{\eta, \xi} \mathbf{1}_{|\eta| > 1} \left| A \widehat{f_0}(\eta) J_0(\eta) e^{c\lambda|\eta - \xi|^s} \langle \eta - \xi \rangle^{\sigma/2} \widehat{G}(\eta - \xi)_M \right| \\ &\quad \times \left| \langle \xi \rangle^{\sigma/2+1} e^{c\lambda|\xi|^s} \widehat{\partial_v f_0}(\xi)_{M'} \right| d\xi d\eta. \end{aligned}$$

By Lemma 3.1 followed by (A.13) and (A.14) we have

$$\begin{aligned} |E_{\mathcal{R}}^{0;H}| &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int_{\eta, \xi} \mathbf{1}_{|\eta| > 1} \left| A \widehat{f_0}(\eta) e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_M e^{\lambda|\xi|^s} \widehat{\partial_v f_0}(\xi)_{M'} \right| d\xi d\eta \\ &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M/8 \leq M' \leq 8M} \int_{\eta, \xi} \mathbf{1}_{|\eta| > 1} \left| \eta A \widehat{f_0}(\eta) e^{\lambda|\eta - \xi|^s} \widehat{G}(\eta - \xi)_M e^{\lambda|\xi|^s} \widehat{\partial_v f_0}(\xi)_{M'} \right| d\xi d\eta, \end{aligned}$$

where in the last line we used  $|\eta| > 1$  on the support of the integrand. Therefore, by (A.5) followed by Cauchy-Schwarz in  $M$  and (3.18) (recall (5.4)),

$$\begin{aligned} |E_{\mathcal{R}}^{0;H}| &\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \|\partial_v A f_0\|_2 \left\| (1 - (v')^2)_M \right\|_{\mathcal{G}^{\lambda, 2}} \|\partial_v A(f_0)_{\sim M}\|_{\mathcal{G}^{\lambda, 0}} \\ &\lesssim \epsilon \nu \|\partial_v A f_0\|_2^2, \end{aligned} \quad (5.20)$$

which is then absorbed by the dissipation in (5.3) for  $\epsilon$  sufficiently small.

Next we treat  $E_{\mathcal{R}}^{0;L}$ , which requires more care than  $E_{\mathcal{R}}^{0;H}$ . By Cauchy-Schwarz and Bernstein's

inequalities,

$$\begin{aligned}
\left| E_{\mathcal{R}}^{0:L} \right| &\leq \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|P_{\leq 1} A f_0\|_2 \|P_{\leq 1} A (G_{M'} (\partial_{vv} f_0)_M)\|_2 \\
&\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|f_0\|_2 \|G_{M'} (\partial_{vv} f_0)_M\|_2 \\
&\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|f_0\|_2 \|(1 - (v')^2)_{M'}\|_{\infty} \|(\partial_{vv} f_0)_M\|_2 \\
&\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \|f_0\|_2 M^{3/2} \|(1 - (v')^2)_{\sim M}\|_2 \|(\partial_v f_0)_M\|_2 \\
&\lesssim \nu \sum_{M \in 2^{\mathbb{Z}}: M \leq 1} \|f_0\|_2 M^{1/2} \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|(\partial_v f_0)_M\|_2 \\
&\quad + \nu \sum_{M \in 2^{\mathbb{Z}}: M > 1} \|f_0\|_2 M^{1/2} \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|(\partial_v f_0)_M\|_2.
\end{aligned}$$

The first (low frequency) term is summed by Cauchy-Schwarz in  $M$  whereas the second term is summed by paying additional derivatives on the last factor to reduce the power of  $M$ . Therefore by (2.65) and (2.59a),

$$\begin{aligned}
\left| E_{\mathcal{R}}^{0:L} \right| &\lesssim \nu \|f_0\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v f_0\|_2 + \nu \|f_0\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v A f_0\|_2 \\
&\lesssim \nu \|f_0\|_2 \|\partial_v A f_0\|_2^2 + \nu \|f_0\|_2 \|\partial_v (1 - (v')^2)\|_2^2 \\
&\lesssim \epsilon \nu \|\partial_v A f_0\|_2^2 + \epsilon^3 \frac{\nu}{\langle \nu t \rangle^{3/2}}.
\end{aligned} \tag{5.21}$$

For  $\epsilon$  small, the first term is absorbed by the dissipation term in (5.3) while the latter is integrable in time uniformly in  $\nu$  and cubic in  $\epsilon$ .

Putting together all the contributions from (5.1): the Euler nonlinear bound (5.2), the dissipation terms (5.3), the non-zero frequency dissipation error terms (5.7), (5.9), (5.12), (5.13), (5.15) and (5.16) with the zero frequency dissipation error terms (5.19), (5.18), (5.20) and (5.21) and integrating in time gives the bound (2.66a) for  $\epsilon$  chosen sufficiently small.

## 6 Enhanced dissipation estimate (2.67)

Up to an adjustment of the constants in the bootstrap argument, it suffices to consider only  $t$  such that  $\nu t^3 \geq 1$  (say), as otherwise the decay estimate (2.67) follows trivially from the inviscid energy estimate (2.66a).

Computing the time evolution of  $\|A^\nu f\|_2$  from (2.48),

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|A^\nu f\|_2^2 &= -CK_\lambda^\nu + \alpha \int A^\nu f e^{\lambda(t)|\nabla|^s} \langle \nabla \rangle^\beta \langle D(t, \partial_v) \rangle^{\alpha-1} \frac{D(t, \partial_v)}{\langle D(t, \partial_v) \rangle} \partial_t D(t, \partial_v) P_{\neq 0} f dv dz \\
&\quad - \int A^\nu f A^\nu (u \cdot \nabla f) dv dz + \nu \int A^\nu f A^\nu (\tilde{\Delta}_t f) dv dz \\
&\leq -CK_\lambda^\nu + \frac{1}{8} \nu t^2 \|\mathbf{1}_{t \geq 2|\partial_v|} A^\nu P_{\neq 0} f\|_2^2 \\
&\quad - \int A^\nu f A^\nu (u \cdot \nabla f) dv dz + \nu \int A^\nu f A^\nu (\tilde{\Delta}_t f) dv dz.
\end{aligned} \tag{6.1}$$

As in (5.3) above, we write the dissipation term as a perturbation of  $\Delta_L$ ,

$$\begin{aligned} \nu \int A^\nu f A^\nu \Delta_t f dv dz &= -\nu \left\| \sqrt{-\Delta_L} A^\nu f \right\|_2^2 - \sum_{k \neq 0} \int A^\nu f_k A^\nu ((1 - (v')^2)(\partial_v - t \partial_z)^2 f_k) dv dz \\ &= -\nu \left\| \sqrt{-\Delta_L} A^\nu f \right\|_2^2 + E^\nu. \end{aligned} \quad (6.2)$$

First, we need to cancel the growing term in (6.1) that involves  $\nu t^2$  with part of the leading order dissipation in (6.2). Indeed, we have

$$\begin{aligned} \frac{1}{8} \nu t^2 \left\| \mathbf{1}_{t \geq 2|\partial_v|} A^\nu P_{\neq 0} f \right\|_2^2 - \nu \left\| \sqrt{-\Delta_L} A^\nu P_{\neq 0} f \right\|_2^2 \\ = \nu \sum_{k \neq 0} \int \left( \frac{1}{8} t^2 \mathbf{1}_{t \geq 2|\eta|} - |k|^2 - |\eta - kt|^2 \right) \left| A^\nu \hat{f}_k(\eta) \right|^2 d\eta \\ \leq -\frac{\nu}{8} \left\| \sqrt{-\Delta_L} A^\nu P_{\neq 0} f \right\|_2^2, \end{aligned}$$

which implies from (6.1),

$$\frac{1}{2} \frac{d}{dt} \|A^\nu f\|_2^2 \leq -CK_\lambda^\nu - \int A^\nu f A^\nu (u \cdot \nabla f) dv dz - \frac{\nu}{8} \left\| \sqrt{-\Delta_L} A^\nu f \right\|_2^2 + E^\nu. \quad (6.3)$$

There are two challenges here: the Euler nonlinearity (the second term) and the error from the dissipation (the final term).

## 6.1 Euler nonlinearity

We first divide into zero and non-zero frequency contributions, as they will be treated differently:

$$\begin{aligned} - \int A^\nu f A^\nu (u \cdot \nabla f) dv dz &= - \int A^\nu f A^\nu (g \partial_v f) dv dz - \int A^\nu f A^\nu (v' \nabla^\perp P_{\neq 0} \phi \cdot \nabla f) dv dz \\ &= \mathcal{E}_1 + \mathcal{E}_2. \end{aligned}$$

The reason for this is the large disparity that  $A^\nu$  imposes between the zero-in- $z$  mode and the non-zero-in- $z$  modes, which are mixed in  $\mathcal{E}_2$  due to the  $z$  dependence of the velocity field  $\nabla^\perp P_{\neq 0} \phi$ . Indeed, the commutator trick that is used to recover part of the derivative in the treatment of the Euler nonlinearity in [41, 9] normally requires that the norm not vary drastically in  $k$  (the discrete,  $z$  wavenumber). During the proof of (5.2) in §5 of [9], the disparity introduced by  $J$  is recovered using that it only occurs near the critical times, and hence decay of the velocity field can be transferred back to regularity. This will not work here; instead we will make use of the high norm estimate (2.66a) to recover the derivative and take advantage of the rapid decay of the velocity field given by (4.2).

First consider the estimation of  $\mathcal{E}_1$ , which as various techniques in e.g. [41, 38, 9], begins with the commutator trick (note that  $P_{=0}g = g$  and  $A^\nu P_{\neq 0} = A^\nu$ ):

$$- \int A^\nu f_k A^\nu (g \partial_v f) dv dz = \frac{1}{2} \int \partial_v g |A^\nu f|^2 dv dz + \int A^\nu f [g \partial_v A^\nu f - A^\nu (g \partial_v f)] dv dz. \quad (6.4)$$

The first term in (6.4) is controlled by Sobolev embedding and the decay from (2.59c),

$$\frac{1}{2} \int \partial_v g |A^\nu f|^2 dv dz \lesssim \|g\|_\infty \|A^\nu f\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D \epsilon/2}} \|A^\nu f\|_2^2. \quad (6.5)$$



The latter term in (6.4) is expanded with a paraproduct (in both  $z$  and  $v$ ):

$$\begin{aligned}
\int A^\nu f [g \partial_v A^\nu f - A^\nu (g \partial_v f)] dv dz &= \sum_{N \geq 8} \int A^\nu f [g_{<N/8} \partial_v A^\nu f_N - A^\nu (g_{<N/8} \partial_v f_N)] dv dz \\
&+ \sum_{N \geq 8} \int A^\nu f [g_N \partial_v A^\nu f_{<N/8} - A^\nu (g_N \partial_v f_{<N/8})] dv dz \\
&+ \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} \int A^\nu f [g_{N'} \partial_v A^\nu f_N - A^\nu (g_{N'} \partial_v f_N)] dv dz \\
&= \sum_{N \geq 8} T_N^0 + \sum_{N \geq 8} R_N^0 + \mathcal{R}^0.
\end{aligned} \tag{6.6}$$

On the Fourier side,

$$T_N^0 = -\frac{i}{2\pi} \sum_{k \neq 0} \int_{\eta, \xi} A^\nu \bar{\hat{f}}_k(\eta) [A_k^\nu(\eta) - A_k^\nu(\xi)] \hat{g}(\eta - \xi)_{<N/8} \xi \hat{f}_k(\xi)_N d\eta d\xi,$$

and on the support of the integrand there holds

$$||k, \eta| - |k, \xi|| \leq |\eta - \xi| \leq \frac{3}{16} |k, \xi|, \tag{6.7a}$$

$$\frac{13}{16} |k, \xi| \leq |k, \eta| \leq \frac{19}{16} |k, \xi|. \tag{6.7b}$$

For the commutator, we write,

$$\begin{aligned}
A_k^\nu(\eta) - A_k^\nu(\xi) &= \langle D(t, \eta) \rangle^\alpha \left[ e^{\lambda|k, \eta|^s} \langle k, \eta \rangle^\beta - e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta \right] \\
&+ e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta [\langle D(t, \eta) \rangle^\alpha - \langle D(t, \xi) \rangle^\alpha].
\end{aligned} \tag{6.8}$$

Then,

$$\begin{aligned}
T_N^0 &= -i \sum_{k \neq 0} \int_{\eta, \xi} A^\nu \bar{\hat{f}}_k(\eta) \langle D(\eta) \rangle^\alpha \left[ e^{\lambda|k, \eta|^s} \langle k, \eta \rangle^\beta - e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta \right] \hat{g}(\eta - \xi)_{<N/8} \xi \hat{f}_k(\xi)_N d\eta d\xi \\
&- i \sum_{k \neq 0} \int_{\eta, \xi} A^\nu \bar{\hat{f}}_k(\eta) e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta [\langle D(t, \eta) \rangle^\alpha - \langle D(t, \xi) \rangle^\alpha] \hat{g}(\eta - \xi)_{<N/8} \xi \hat{f}_k(\xi)_N d\eta d\xi \\
&= T_N^{0;1} + T_N^{0;2}.
\end{aligned}$$

Consider the more standard  $T_N^{0;1}$  first. We claim that by (6.7), on the support of the integrand there holds, for some  $c \in (0, 1)$  (depending only on  $s$  and our Littlewood-Paley conventions),

$$\left| e^{\lambda|k, \eta|^s} \langle k, \eta \rangle^\beta - e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta \right| \lesssim \frac{|\eta - \xi|}{\langle \xi \rangle^{1-s}} e^{c\lambda|\eta - \xi|^s} e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta; \tag{6.9}$$

see e.g. §5 of [9] or [41] combined with (A.8). Applying (6.9), (6.7) and (3.20) implies

$$\begin{aligned}
|T_N^{0;1}| &\lesssim \sum_{k \neq 0} \int_{\eta, \xi} \left| A^\nu \hat{f}_k(\eta) \langle \eta - \xi \rangle^{3\alpha+1} e^{c\lambda|\eta - \xi|^s} \hat{g}(\eta - \xi)_{<N/8} \frac{\xi}{\langle \xi \rangle^{1-s}} A^\nu \hat{f}_k(\xi)_N \right| d\eta d\xi \\
&\lesssim \sum_{k \neq 0} \int_{\eta, \xi} \left| |k, \eta|^{s/2} A^\nu \hat{f}_k(\eta) \langle \eta - \xi \rangle^{3\alpha+1} e^{c\lambda|\eta - \xi|^s} \hat{g}(\eta - \xi)_{<N/8} |k, \xi|^{s/2} A^\nu \hat{f}_k(\xi)_N \right| d\eta d\xi.
\end{aligned}$$

Therefore, by  $\sigma > 3\alpha + 8$  and (A.5) followed by (2.59c), we get

$$T_N^{0;1} \lesssim \|g_{<N/8}\|_{\mathcal{G}^{\lambda(t), \sigma-6}} \left\| |\nabla|^{s/2} A^\nu f_{\sim N} \right\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon}} \left\| |\nabla|^{s/2} A^\nu f_{\sim N} \right\|_2^2. \quad (6.10)$$

After summing in  $N$  and choosing  $\epsilon$  small, this term is absorbed by  $CK_\lambda^\nu$ .

For  $T_N^{0;2}$  we crucially use that (3.21) applies; this is due to the fact that  $g$  does not depend on  $z$  and will not work to treat  $\mathcal{E}_2$ . Indeed, (3.21) implies

$$T_N^{0;2} \lesssim \sum_{k \neq 0} \int_{\eta, \xi} \left| A^\nu \hat{f}_k(\eta) e^{\lambda|k, \xi|^s} \langle k, \xi \rangle^\beta \langle \eta - \xi \rangle^{3\alpha} \widehat{g}(\eta - \xi)_{<N/8} \langle D(t, \xi) \rangle^\alpha \hat{f}_k(\xi)_N \right| d\eta d\xi.$$

By  $\sigma > 3\alpha + 8$  and (A.5) followed by (2.59c) implies

$$T_N^{0;2} \lesssim \|g\|_{\mathcal{G}^{\lambda(t), \sigma-6}} \|A^\nu f_{\sim N}\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon}} \|A^\nu f_{\sim N}\|_2^2, \quad (6.11)$$

which is an integrable contribution.

The ‘reaction’ term  $R_N^0$  is dealt with easily using Lemma 3.7 and the bootstrap controls on the higher norms. First note that  $A^\nu(g\partial_v f) = A^\nu(g\partial_v P_{\neq 0} f)$ . Then, by Cauchy-Schwarz, (3.22),  $\sigma \geq \beta + 3\alpha + 7$ , Cauchy-Schwarz in  $N$  and (2.59c),

$$\begin{aligned} \left| \sum_{N \geq 8} \int A^\nu f A^\nu (g_N \partial_v P_{\neq 0} f_{<N/8}) dv dz \right| &\leq \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 \|A^\nu (g_N \partial_v P_{\neq 0} f_{<N/8})\|_2 \\ &\lesssim \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 \|g_N\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \|A^\nu \partial_v P_{\neq 0} f_{<N/8}\|_2 \\ &\lesssim \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 N \|g_N\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \|A^\nu P_{\neq 0} f_{<N/8}\|_2 \\ &\lesssim \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 \|g_N\|_{\mathcal{G}^{\lambda, \beta+3\alpha+1}} \|A^\nu P_{\neq 0} f_{<N/8}\|_2 \\ &\lesssim \|g\|_{\mathcal{G}^{\lambda, \sigma-6}} \|A^\nu f\|_2^2 \\ &\lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon/2}} \|A^\nu f\|_2^2, \end{aligned} \quad (6.12)$$

which is an integrable contribution. Similarly, by Bernstein’s inequalities, Cauchy-Schwarz in  $N$  and (2.59c) we have

$$\begin{aligned} \left| \sum_{N \geq 8} \int A^\nu f g_N \partial_v A^\nu P_{\neq 0} f_{<N/8} dv dz \right| &\leq \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 \|g_N\|_\infty \|\partial_v A^\nu P_{\neq 0} f_{<N/8}\|_2 \\ &\lesssim \sum_{N \geq 8} \|A^\nu f_{\sim N}\|_2 N^{3/2} \|g_N\|_2 \|A^\nu f_{<N/8}\|_2 \\ &\lesssim \|g\|_{H^{3/2}} \|A^\nu f\|_2^2 \\ &\lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon/2}} \|A^\nu f\|_2^2, \end{aligned} \quad (6.13)$$

which again is integrable. The remainder terms  $\mathcal{R}^0$  in (6.6) are treated very similar to the reaction terms  $R_N^0$  just completed; hence we omit the details and simply conclude

$$|\mathcal{R}^0| \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon/2}} \|A^\nu f\|_2^2. \quad (6.14)$$

Next turn to  $\mathcal{E}_2$ . If the zero mode in  $z$  interacts with non-zero modes, possible now due to the  $z$  dependence of the velocity field, then we can no longer apply (3.21) to gain regularity. Physically, the issue is that gradients in the large zero frequency can be converted to large gradients in non-zero frequencies by transport. By Cauchy-Schwarz and two applications of (3.22),

$$\begin{aligned}\mathcal{E}_2 &\leq \|A^\nu f\|_2 \left\| A^\nu \left( v' \nabla^\perp P_{\neq 0} \phi \right) \right\|_2 \|\nabla f\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \\ &\lesssim \|A^\nu f\|_2 \|f\|_{\mathcal{G}^{\lambda, \beta+3\alpha+1}} \left( \left\| A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2 + \left\| A^\nu \left( (v' - 1) \nabla^\perp P_{\neq 0} \phi \right) \right\|_2 \right) \\ &\lesssim \|A^\nu f\|_2 \|f\|_{\mathcal{G}^{\lambda, \beta+3\alpha+1}} \left( 1 + \|v' - 1\|_{\mathcal{G}^{\lambda, \beta+3\alpha}} \right) \left\| A^\nu \nabla^\perp P_{\neq 0} \phi \right\|_2.\end{aligned}$$

Then by  $\sigma \geq \beta + 3\alpha + 1$ , Lemma 4.2, (2.59a), (2.60) and (2.59e),

$$\mathcal{E}_2 \lesssim \|A^\nu f\|_2 \|Af\|_2 (1 + \|A(v' - 1)\|_2) \frac{\|Af\|_2 + \|A^\nu f\|_2}{\langle t \rangle^2} \lesssim \frac{\epsilon^3}{\langle t \rangle^2}, \quad (6.15)$$

an integrable, cubic in  $\epsilon$  contribution and hence this completes the treatment of  $\mathcal{E}_2$  and of the entire Euler nonlinearity.

## 6.2 Dissipation error term

Since at high frequencies  $A^\nu$  defines a weaker norm than  $A$ , it will be possible to apply (5.5) and the stronger control coming from (2.59e). Indeed, writing  $E^\nu$  on the Fourier side and applying (5.5) implies

$$\begin{aligned}|E^\nu| &\lesssim \nu \sum_{k \neq 0} \int_{\eta, \xi} \left| A^\nu \widehat{f}_k(\eta) A_k^\nu(\eta) \widehat{G}(\eta - \xi) |\xi - tk|^2 \widehat{f}_k(\xi) \right| d\eta d\xi \\ &\lesssim \nu \sum_{k \neq 0} \int_{\eta, \xi} \left| \sqrt{-\Delta_L} A^\nu \widehat{f}_k(\eta) A_k^\nu(\eta) \langle \eta - \xi \rangle \widehat{G}(\eta - \xi) |\xi - tk| \widehat{f}_k(\xi) \right| d\eta d\xi.\end{aligned}$$

By (A.9), (3.20) and  $\langle \eta \rangle^\beta \lesssim \langle \eta - \xi \rangle^\beta \langle \xi \rangle^\beta$  we have

$$|E^\nu| \lesssim \nu \sum_{k \neq 0} \int_{\eta, \xi} \left| \sqrt{-\Delta_L} A^\nu \widehat{f}_k(\eta) \langle \eta - \xi \rangle^{\beta+3\alpha+1} e^{\lambda|\eta-\xi|^s} \widehat{G}(\eta - \xi) |\xi - tk| A^\nu \widehat{f}_k(\xi) \right| d\eta d\xi.$$

Therefore, by (A.5),  $\sigma > \beta + 3\alpha + 2$  and  $k \neq 0$  followed by (3.18) we get

$$\begin{aligned}|E^\nu| &\lesssim \nu \sum_{k \neq 0} \|1 - (v')^2\|_{\mathcal{G}^{\lambda, \sigma}} \left\| \sqrt{-\Delta_L} A^\nu f_k \right\|_2^2 \\ &\lesssim \epsilon \nu \left\| \sqrt{-\Delta_L} A^\nu f \right\|_2^2,\end{aligned} \quad (6.16)$$

which is absorbed by the leading order dissipation in (6.3) for  $\epsilon$  small.

Putting together the estimates on (6.3) from the Euler nonlinearity ((6.5), (6.10), (6.11), (6.12), (6.13), (6.14) and (6.15)) and from the dissipation error terms (6.16), we deduce (2.67) after integrating in time and choosing  $\epsilon$  sufficiently small.

## 7 Coordinate system higher regularity controls

In §8 of [9], there are three main estimates on the coordinate system that need to be made, corresponding here to (2.66b), (2.66c) and (2.66e). That (2.66d) can be deduced from the proof of (2.66b), (2.66c) and (2.66e) is shown in §8 of [9] and is not repeated here as there is little difference.

The main new issue for the Navier-Stokes case is to confirm that the variable coefficients in  $\tilde{\Delta}_t$  do not slow down or otherwise impede the decay estimates. This could be possible, for example, if the diffusion coefficient oscillated too much relative to the strength of the damping terms, as gradients in the diffusion coefficient induce momentum transport. This effect is controlled by the gradient decay estimate in (2.68b).

### 7.1 Proof of (2.66c)

Computing from (2.46) implies

$$\begin{aligned} \frac{d}{dt} \left( \langle t \rangle^{4-K_D\epsilon} \|g\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2 \right) &= (4-K_D\epsilon) \langle t \rangle^{2-K_D\epsilon} \|g\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2 \\ &\quad + \langle t \rangle^{4-K_D\epsilon} \frac{d}{dt} \left\| \frac{A}{\langle \partial_v \rangle^s} g \right\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2. \end{aligned} \quad (7.1)$$

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

$$\langle t \rangle^{4-K_D\epsilon} \frac{d}{dt} \|g\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2 = 2 \langle t \rangle^{4-K_D\epsilon} \dot{\lambda}(t) \left\| |\partial_v|^{s/2} g \right\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2 + 2 \langle t \rangle^{4-K_D\epsilon} \int A^S g A^S \partial_t g dv, \quad (7.2)$$

From (2.46),

$$\begin{aligned} 2 \langle t \rangle^{4-K_D\epsilon} \int A^S g A^S \partial_t g dv &= -\frac{4 \langle t \rangle^{4-K_D\epsilon}}{t} \|g\|_{\mathcal{G}^{\lambda(t), \sigma-6}}^2 \\ &\quad - 2 \langle t \rangle^{4-K_D\epsilon} \int A^S g A^S (g \partial_v g) dv \\ &\quad - \frac{2 \langle t \rangle^{4-K_D\epsilon}}{t} \int A^S g A^S \left( v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} > \right) dv \\ &\quad + 2\nu \langle t \rangle^{4-K_D\epsilon} \int A^S g A^S \left( \tilde{\Delta}_t g \right) dv \\ &= V_1 + V_2 + V_3 + V_D. \end{aligned} \quad (7.3)$$

The first three terms are basically the same as in Euler, and are treated accordingly as in [9]. Indeed, in §8 of [9] it is shown that the bootstrap hypotheses together with (A.12) and (2.52) imply

$$V_2 \leq \frac{K_D\epsilon}{8} \langle t \rangle^{3-K_D\epsilon-s} \|g\|_{\mathcal{G}^{\lambda, \sigma-6}}^2, \quad (7.4)$$

where we define  $K_D$  to be the maximum of the constant appearing in this term and several other below.

Treating  $V_3$  is not hard due to the regularity gap of 6 derivatives. Note that

$$\nabla \tilde{u} = - \left( \begin{array}{c} v'(\partial_v - t\partial_z)\partial_z\phi \\ \partial_v v'(\partial_v - t\partial_z)\phi + v'(\partial_v - t\partial_z)\partial_v\phi \end{array} \right), \quad (7.5)$$

and therefore by (A.12), (7.5), Lemma 4.1 and the bootstrap hypotheses,

$$\|\nabla P_{\neq 0} \tilde{u}(t)\|_{\mathcal{G}^{\lambda(t), \sigma-5}} \lesssim \frac{\epsilon}{\langle t \rangle}. \quad (7.6)$$

It is shown in §8 of [9] that (7.6), together with (A.12), Lemma 4.1 and the bootstrap hypotheses, implies we have for some  $C > 0$ ,

$$V_3 \lesssim \langle t \rangle^{-K_D \epsilon} \epsilon^2 \|g\|_{\mathcal{G}^{\lambda, \sigma-6}} \leq \frac{K_D \epsilon \langle t \rangle^{4-K_D \epsilon}}{8t} \|g\|_{\mathcal{G}^{\lambda, \sigma-6}}^2 + C \epsilon^3 t^{-3-K_D \epsilon}. \quad (7.7)$$

Focus now on the term that is new for Navier-Stokes,  $V_D$ . As in many other estimates in this work, we write this as two contributions:

$$\begin{aligned} V_D &= -2\nu \langle t \rangle^{4-K_D \epsilon} \|\partial_v A^S g\|_2^2 - 2\nu \langle t \rangle^{4-K_D \epsilon} \int A^S g A^S (1 - (v')^2) \partial_{vv} g dv \\ &= -2\nu \langle t \rangle^{4-K_D \epsilon} \|\partial_v A^S g\|_2^2 - V_{DE}, \end{aligned} \quad (7.8)$$

a leading order dissipation term and an error term which remains to be controlled. We use again the shorthand (5.4) and decompose the error term with a homogeneous paraproduct

$$\begin{aligned} V_{DE} &= 2\nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D \epsilon} \int A^S g A^S (G_M \partial_{vv} g_{<M/8}) dv \\ &\quad + 2\nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D \epsilon} \int A^S g A^S (G_{<M/8} \partial_{vv} g_M) dv \\ &\quad + 2\nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D \epsilon} \sum_{M/8 \leq M' \leq 8M} \int A^S g A^S (G_{M'} \partial_{vv} g_M) dv \\ &= V_{DE,HL} + V_{DE,LH} + V_{DE,\mathcal{R}}. \end{aligned}$$

As in the zero frequency dissipation error terms in §5.2, the remainder term is the most delicate due to the contributions of low frequencies.

First we deal with  $V_{DE,LH}$ , which on the Fourier side is written as

$$V_{DE,LH} \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D \epsilon} \int_{\eta, \xi} \left| A^S \hat{g}(\eta) A^S(\eta) \widehat{G}(\eta - \xi)_{<M/8} |\xi|^2 \hat{g}(\xi)_M \right| d\eta d\xi.$$

On the support of the integrand, (5.17) holds with the role of  $\xi$  and  $\eta - \xi$  reversed. Hence, by (A.8), for some  $c \in (0, 1)$ :

$$V_{DE,LH} \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D \epsilon} \int_{\eta, \xi} \left| \eta A^S \hat{g}(\eta) e^{c\lambda|\eta-\xi|^s} \widehat{G}(\eta - \xi)_{<M/8} \xi A^S(\xi) \hat{g}(\xi)_M \right| d\eta d\xi.$$

Therefore by (A.5),  $\sigma > 8$  followed by (A.2) and (3.18a) we have

$$\begin{aligned} V_{DE,LH} &\lesssim \nu \langle t \rangle^{4-K_D \epsilon} \sum_{M \in 2^{\mathbb{Z}}} \|\partial_v A^S g_{\sim M}\|_2^2 \left\| A^S ((v')^2 - 1)_{<M/8} \right\|_2 \\ &\lesssim \epsilon \nu \langle t \rangle^{4-K_D \epsilon} \|\partial_v A^S g\|_2^2, \end{aligned} \quad (7.9)$$

which is absorbed by the leading order dissipation term in (7.8) for  $\epsilon$  small.

The treatment of  $V_{DE,HL}$  is similar to  $V_{DE,HL}$ . Indeed, we have

$$V_{DE,HL} \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D\epsilon} \int_{\eta, \xi} \left| A^S \hat{g}(\eta) A^S(\eta) \hat{G}(\eta - \xi)_M |\xi|^2 \hat{g}(\xi)_{<M/8} \right| d\eta d\xi,$$

and (5.17) holds on the support of the integrand. Therefore, on the support of the integrand there still holds  $|\xi| \lesssim |\eta|$ , and hence using (A.8) we have for some  $c \in (0, 1)$ ,

$$V_{DE,HL} \lesssim \nu \sum_{M \in 2^{\mathbb{Z}}} \langle t \rangle^{4-K_D\epsilon} \int_{\eta, \xi} \left| \eta A^S \hat{g}(\eta) A^S(\eta - \xi) \hat{G}(\eta - \xi)_M \xi e^{c\lambda|\xi|^s} \hat{g}(\xi)_{<M/8} \right| d\eta d\xi.$$

As in (7.9), it follows from (A.5),  $\sigma > 2$ , (A.1) and (3.18a) that

$$V_{DE,HL} \lesssim \epsilon \nu \langle t \rangle^{4-K_D\epsilon} \left\| \partial_v A^S g \right\|_2^2, \quad (7.10)$$

which is absorbed by the leading order dissipation in (7.8) for  $\epsilon$  small.

Finally turn to the remainder,  $V_{DE,\mathcal{R}}$ . As in §5.2, divide first into low and high frequencies

$$\begin{aligned} V_{DE,\mathcal{R}} &= \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \int (A^S g)_{\leq 1} A^S (G_{M'} \partial_{vv} g_M) dv \\ &\quad + \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \int (A^S g)_{> 1} A^S (G_{M'} \partial_{vv} g_M) dv \\ &= V_{DE,\mathcal{R}}^L + V_{DE,\mathcal{R}}^H. \end{aligned}$$

The high frequencies  $V_{DE,\mathcal{R}}^H$ , as in §5.2, can be treated by adding a derivative on the first factor and absorbing by the leading order dissipation in (7.8). We omit the details and simply conclude

$$V_{DE,\mathcal{R}}^H \lesssim \epsilon \nu \langle t \rangle^{4-K_D\epsilon} \left\| \partial_v A^S g \right\|_2^2, \quad (7.11)$$

which is absorbed by the leading order dissipation in (7.8) for  $\epsilon$  small.

To treat  $V_{DE,\mathcal{R}}^L$ , also similar to §5.2, we use Cauchy-Schwarz followed by Bernstein's inequalities,

$$\begin{aligned} V_{DE,\mathcal{R}}^L &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|P_{\leq 1} A^S g\|_2 \|P_{\leq 1} A^S (G_{M'} \partial_{vv} g_M)\|_2 \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|g_{<1}\|_2 \|G_{M'} \partial_{vv} g_M\|_2 \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \|g_{<1}\|_2 \|G_{M'}\|_{\infty} \|\partial_{vv} g_M\|_2 \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} M^{3/2} \|g\|_2 \|G_{M'}\|_2 \|\partial_v g_M\|_2 \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} M^{1/2} \|g\|_2 \|\partial_v G_{M'}\|_2 \|\partial_v g_M\|_2 \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}: M \leq 1} M^{1/2} \|g\|_2 \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|\partial_v g_M\|_2 \\ &\quad + \nu \langle t \rangle^{4-K_D\epsilon} \sum_{M \in 2^{\mathbb{Z}}: M > 1} M^{1/2} \|g\|_2 \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|\partial_v g_M\|_2. \end{aligned}$$

The first term is summed by Cauchy-Schwarz in  $M$  (and (A.2)) whereas the second term is summed by paying additional derivatives on the last factor to introduce a negative power of  $M$  and then applying Cauchy-Schwarz. Therefore, since  $\sigma > 7$ ,

$$\begin{aligned} V_{DE,\mathcal{R}}^L &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \|g\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v g\|_2 + \nu \langle t \rangle^{4-K_D\epsilon} \|g\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v g\|_{\mathcal{G}^{\lambda,\sigma-6}} \\ &\lesssim \nu \langle t \rangle^{4-K_D\epsilon} \|g\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v g\|_{\mathcal{G}^{\lambda,\sigma-6}} \\ &\lesssim \epsilon \nu \langle t \rangle^{4-K_D\epsilon} \|\partial_v g\|_{\mathcal{G}^{\lambda,\sigma-6}}^2 + \epsilon^{-1} \nu \langle t \rangle^{4-K_D\epsilon} \|g\|_2^2 \|v' \partial_v v'\|_2^2. \end{aligned}$$

Now we crucially use (2.65) (from (2.61b)) to deduce for some  $C > 0$

$$V_{DE,\mathcal{R}}^L \leq C \epsilon \nu \langle t \rangle^{4-K_D\epsilon} \|\partial_v g\|_{\mathcal{G}^{\lambda,\sigma-6}}^2 + \frac{K_D}{8} \epsilon \langle t \rangle^{3-K_D\epsilon} \|g\|_2^2. \quad (7.12)$$

The first term is absorbed by the leading order dissipation in (7.8) and the latter term is absorbed by the damping term  $V_1$  in (7.3) by choosing  $\epsilon$  sufficiently small.

By integrating (7.1) using (7.2), (7.3), (7.4), (7.7) (7.8), (7.9), (7.10), (7.11) and (7.12), the proof of (2.66c) is completed by choosing  $\epsilon$  sufficiently small.

## 7.2 Proof of (2.66b)

We extend the proof of the corresponding statement in [9]. From (2.51) we have

$$\begin{aligned} \frac{d}{dt} \left( \langle t \rangle^{2+2s} \left\| \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 \right) &= -(2-2s)t \langle t \rangle^{2s} \left\| \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 - CK_\lambda^{v,2} - CK_w^{v,2} \\ &\quad - 2 \langle t \rangle^{2+2s} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} (g \partial_v \bar{h}) dv \\ &\quad + 2t^{-1} \langle t \rangle^{2+2s} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} \langle v' \nabla^\perp P_{\neq 0} \phi \cdot \nabla f \rangle dv \\ &\quad + 2\nu \langle t \rangle^{2+2s} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} (\tilde{\Delta}_t \bar{h}) dv \\ &= -CK_L^{v,2} - CK_\lambda^{v,2} - CK_w^{v,2} + \mathcal{T}^h + F + \mathcal{D}^h. \end{aligned} \quad (7.13)$$

The main nonlinear terms,  $\mathcal{T}^h$  and  $F$ , are treated in §8.2 of [9]. Using the techniques therein implies:

$$|\mathcal{T}^h| \lesssim \epsilon CK_\lambda^{v,2} + \epsilon CK_\lambda^{v,1} + \epsilon \langle t \rangle^{2s+K_D\epsilon/2} \left\| \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2. \quad (7.14)$$

Controlling the ‘forcing’ term  $F$  is one of the key estimates made in [9] (found in §8.2):

$$\begin{aligned} |F| &\lesssim \epsilon CK_\lambda^{v,2} + \epsilon CK_w^{v,2} + \epsilon CK_\lambda + \epsilon CK_w \\ &\quad + \epsilon^3 \sum_{i=1}^2 CCK_\lambda^i + CCK_w^i + \epsilon CK_L^{v,2} + \epsilon^3 \langle t \rangle^{2s-3}. \end{aligned} \quad (7.15)$$

Hence we only need to focus on the dissipation error term  $\mathcal{D}^h$ , which we treat in a manner very similar to  $V_{DE}$ . Indeed, write

$$\begin{aligned} \mathcal{D}^h &= -2\nu \langle t \rangle^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 - 2\nu \langle t \rangle^{2+2s} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} ((1 - (v')^2) \partial_{vv} \bar{h}) dv \\ &= -2\nu \langle t \rangle^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \mathcal{E}^h \end{aligned} \quad (7.16)$$

As usual, we use the shorthand (5.4) and decompose the error term via homogeneous paraproduct:

$$\begin{aligned}
\mathcal{E}^h &= -2\nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} (G_M \partial_{vv} \bar{h}_{<M/8}) dv \\
&\quad - 2\nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} (G_{<M/8} \partial_{vv} \bar{h}_M) dv \\
&\quad - 2\nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \int \frac{A}{\langle \partial_v \rangle^s} \bar{h} \frac{A}{\langle \partial_v \rangle^s} (G_{M'} \partial_{vv} \bar{h}_M) dv \\
&= \mathcal{E}_{HL}^h + \mathcal{E}_{LH}^h + \mathcal{E}_{\mathcal{R}}^h.
\end{aligned}$$

The treatments of  $\mathcal{E}_{HL}^h$  and  $\mathcal{E}_{LH}^h$  mirror that of  $V_{DE,HL}$  and  $V_{DE,LH}$  except using also (3.13). The argument is essentially the same so it is omitted for brevity and we conclude that

$$\mathcal{E}_{HL}^h + \mathcal{E}_{LH}^h \lesssim \epsilon \nu \langle t \rangle^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2, \quad (7.17)$$

which is absorbed by the leading order dissipation in (7.16) after choosing  $\epsilon$  small.

The treatment of  $\mathcal{E}_{\mathcal{R}}^h$  is also very similar to the treatment of  $V_{DE,\mathcal{R}}$ . Indeed, we first divide into low and high frequencies:

$$\begin{aligned}
\mathcal{E}_{\mathcal{R}}^h &= -2\nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \int \left( \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right)_{\leq 1} \frac{A}{\langle \partial_v \rangle^s} (G_{M'} \partial_{vv} \bar{h}_M) dv \\
&\quad - 2\nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}} \sum_{M' \approx M} \int \left( \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right)_{> 1} \frac{A}{\langle \partial_v \rangle^s} (G_{M'} \partial_{vv} \bar{h}_M) dv \\
&= \mathcal{E}_{\mathcal{R}}^L + \mathcal{E}_{\mathcal{R}}^H.
\end{aligned}$$

As for  $V_{DE}$ , the high frequencies are treated without much effort as in §5.2 and absorbed by the leading order dissipation (7.16) after choosing  $\epsilon$  small. Hence this contribution is omitted. To treat the lower frequencies we use an argument similar to that used to treat  $V_{DE,\mathcal{R}}^L$ . Indeed by Cauchy-Schwarz and Bernstein's inequalities we have as above (skipping repetitive details):

$$\begin{aligned}
|\mathcal{E}_{\mathcal{R}}^L| &\lesssim \nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}: M \leq 1} M^{1/2} \|\bar{h}\|_2 \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|\partial_v \bar{h}_M\|_2 \\
&\quad + \nu \langle t \rangle^{2+2s} \sum_{M \in 2^{\mathbb{Z}}: M > 1} M^{1/2} \|\bar{h}\|_2 \|\partial_v (1 - (v')^2)_{\sim M}\|_2 \|\partial_v \bar{h}_M\|_2.
\end{aligned}$$

The first (low frequency) term is summed by Cauchy-Schwarz in  $M$  (and (A.2)) whereas the second term is summed by paying additional derivatives on the last factor to introduce a negative power of  $M$  and then applying Cauchy-Schwarz. As in §7.1, since  $\sigma > s + 1/2$ ,

$$\begin{aligned}
|\mathcal{E}_{\mathcal{R}}^L| &\lesssim \nu \langle t \rangle^{2+2s} \|\bar{h}\|_2 \|\partial_v (1 - (v')^2)\|_2 \|\partial_v \bar{h}\|_2 + \nu \langle t \rangle^{2+2s} \|\bar{h}\|_2 \|\partial_v (1 - (v')^2)\|_2 \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2 \\
&\lesssim \epsilon \nu \langle t \rangle^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \epsilon^{-1} \nu \langle t \rangle^{2+2s} \|\bar{h}\|_2^2 \|\partial_v (1 - (v')^2)\|_2^2.
\end{aligned}$$

By (2.65) we have

$$|\mathcal{E}_{\mathcal{R}}^L| \lesssim \epsilon \nu \langle t \rangle^{2+2s} \left\| \partial_v \frac{A}{\langle \partial_v \rangle^s} \bar{h} \right\|_2^2 + \epsilon \langle t \rangle^{1+2s} \|\bar{h}\|_2^2. \quad (7.18)$$



The first term is absorbed by the leading order dissipation in (7.16) and the latter term is absorbed by the damping term  $CK_L^{v,2}$  in (7.13) by choosing  $\epsilon$  sufficiently small.

Combining (7.13) with (7.14), (7.15), (7.17) and (7.18) and integrating in time implies (2.66b) after choosing  $\epsilon$  sufficiently small.

### 7.3 Proof of (2.66e)

From (2.50) we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|A^R h\|_2^2 &= -CK_\lambda^h - CK_w^h - \int A^R h A^R (g \partial_v h) dv \\ &\quad + \int A^R h A^R \bar{h} dv + \nu \int A^R h A^R (\tilde{\Delta}_t h) dv, \end{aligned} \quad (7.19)$$

where

$$CK_w^h(\tau) = \left\| \sqrt{\frac{\partial_t w}{w}} A^R h(\tau) \right\|_2^2 \quad (7.20a)$$

$$CK_\lambda^h(\tau) = (-\dot{\lambda}(\tau)) \left\| |\partial_v|^{s/2} A^R h(\tau) \right\|_2^2. \quad (7.20b)$$

The nonlinear transport term is controlled in §8 of [9]; here we simply recall the result:

$$\left| \int A^R h A^R (g \partial_v h) dv \right| \lesssim \epsilon CK_\lambda^h + \epsilon CK_w^h + \epsilon CK_\lambda^{v,1} + \epsilon CK_w^{v,1} + \epsilon^3 \langle t \rangle^{-2+K_D \epsilon/2}. \quad (7.21)$$

Similarly, the linear driving term from  $\bar{h}$  is treated in [9]; the result is for some  $C_h > 0$ :

$$\left| \int A^R h A^R \bar{h} dv \right| \leq \frac{1}{4} CK_\lambda^h + \frac{1}{4} CK_w^h + C_h \left( CK_\lambda^{v,2} + CK_w^{v,2} \right). \quad (7.22)$$

The presence of the constant  $C_h$  is the primary reason for the constant  $K_v$  in the main bootstrap.

The new term we need to treat in (7.19) is the dissipation term, which we treat in the same manner as the zero mode dissipation in the proof of (2.66a) in §5.2. As there, we write

$$\begin{aligned} \nu \int A^R h A^R (\tilde{\Delta}_t h) dv &= -\nu \|\partial_v A^R h\|_2^2 + \nu \int A^R h A^R ((1 - (v')^2) \partial_{vv} h) dv \\ &= -\nu \|\partial_v A^R h\|_2^2 + E^R. \end{aligned} \quad (7.23)$$

The treatment of  $E^R$  is essentially the same as  $E^0$  in §5.2 and is hence omitted for brevity. Indeed, the coefficients of  $\tilde{\Delta}_t$  have resonant regularity from (3.18) and hence there is little difference between  $A^R$  and  $A$  in the proof. Analogous to §5.2 we get for  $\epsilon$  sufficiently small,

$$E^R \lesssim \epsilon \nu \|\partial_v A^R h\|_2^2 + \frac{\nu \epsilon}{\langle \nu t \rangle^{3/2}} \|A^R h\|_2^2. \quad (7.24)$$

Putting together (7.19), (7.21), (7.22), (7.23) and (7.24) and integrating in time completes the proof of (2.66e) after choosing  $\epsilon$  small and  $K_v$  sufficiently large (independent of  $\epsilon$  of course).

## 8 Low frequency estimates

### 8.1 Kinetic energy and energy dissipation control: (2.68c)

Compute the time evolution from the momentum equation

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\tilde{u}_0\|_2^2 &= - \int \tilde{u}_0 (g \partial_v \tilde{u}_0) dv - \int \tilde{u}_0 v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} > dv \\ &\quad + \nu \int \tilde{u}_0 (v')^2 \partial_{vv} \tilde{u}_0 dv. \end{aligned} \quad (8.1)$$

For the first term we use integration by parts following by Sobolev embedding and (2.59c),

$$- \int \tilde{u}_0 (g \partial_v \tilde{u}_0) dv = \frac{1}{2} \int \partial_v g |\tilde{u}_0|^2 dv \lesssim \|\partial_v g\|_\infty \|\tilde{u}_0\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D \epsilon/2}} \|\tilde{u}_0\|_2^2. \quad (8.2)$$

For the forcing from the non-zero frequencies, we use Hölder's inequality followed by (2.61a), Gagliardo-Nirenberg-Sobolev and Cauchy-Schwarz in  $k$ :

$$\begin{aligned} \left| \int \tilde{u}_0 v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} > dv \right| &= \sum_{k \neq 0} \int \tilde{u}_0 v' \nabla^\perp \phi_k \cdot \nabla \tilde{u}_{-k} dv \\ &\leq \|\tilde{u}_0\|_2 (1 + \|v' - 1\|_\infty) \sum_{k \neq 0} \left\| \nabla^\perp \phi_k \right\|_4 \|\nabla \tilde{u}_{-k}\|_4 \\ &\lesssim \|\tilde{u}_0\|_2 \left\| \nabla^\perp P_{\neq 0} \phi \right\|_{H^2} \|\nabla P_{\neq 0} \tilde{u}\|_{H^2}. \end{aligned}$$

Therefore, by Lemma 4.1, (2.59a) and (7.6) we have

$$\left| \int \tilde{u}_0 v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla \tilde{u} > dv \right| \lesssim \frac{\epsilon^2}{\langle t \rangle^3} \|\tilde{u}_0\|_2 \lesssim \frac{\epsilon}{\langle t \rangle^3} \|\tilde{u}_0\|_2^2 + \frac{\epsilon^3}{\langle t \rangle^3}. \quad (8.3)$$

Finally we deal with the dissipation term (8.1):

$$\nu \int \tilde{u}_0 (v')^2 \partial_{vv} \tilde{u}_0 dv = -\nu \|v' \partial_v \tilde{u}_0\|_2^2 - 2\nu \int \tilde{u}_0 v' \partial_v v' \partial_v \tilde{u}_0 dv. \quad (8.4)$$

Note that by (2.61a),

$$\|\partial_v \tilde{u}_0\|_2^2 \approx \|v' \partial_v \tilde{u}_0\|_2^2.$$

To deal with the error term in (8.4), we employ the following Gagliardo-Nirenberg-Sobolev inequality for functions  $X = X(v)$ ,

$$\|X\|_\infty \lesssim \|X\|_2^{1/2} \|\partial_v X\|_2^{1/2}. \quad (8.5)$$

Applying (8.5) along with (2.68b) to the error term in (8.4) implies,

$$\begin{aligned} \left| 2\nu \int \tilde{u}_0 v' \partial_v v' \partial_v \tilde{u}_0 dv \right| &\leq 2\nu \|\tilde{u}_0\|_\infty \|\partial_v v'\|_2 \|v' \partial_v \tilde{u}_0\|_2 \\ &\lesssim \nu \|\tilde{u}_0\|_2^{1/2} \|\partial_v v'\|_2 \|v' \partial_v \tilde{u}_0\|_2^{3/2} \\ &\lesssim \nu \epsilon^{-3} \|\tilde{u}_0\|_2^2 \|\partial_v v'\|_2^4 + \epsilon \nu \|v' \partial_v \tilde{u}_0\|_2^2 \\ &\lesssim \frac{\nu \epsilon}{\langle \nu t \rangle^4} \|\tilde{u}_0\|_2^2 + \epsilon \nu \|v' \partial_v \tilde{u}_0\|_2^2. \end{aligned} \quad (8.6)$$

The second term is absorbed by the dissipation in (8.4) and the first term gives an integrable contribution. Hence, integrating (8.1) with (8.6), (8.4), (8.3) and (8.2) implies (2.68c).

## 8.2 Decay estimate (2.68b)

First, (2.68b) follows from (2.59a) for  $t \leq \nu^{-1}$ , so hence assume  $t > \nu^{-1}$  for the rest of the section. The  $L^2$  decay estimate is obtained via an energy estimate and the application of a suitable Gagliardo-Nirenberg-Sobolev inequality. The  $L^2$  gradient decay estimate is then obtained via iterating from the  $L^2$  estimate as in standard Moser iteration methods. The iteration procedure could be carried out for higher derivatives and higher  $L^p$  norms, but this is not necessary for our purposes.

Begin with the  $L^2$  norm by computing the time evolution from (2.50),

$$\frac{1}{2} \frac{d}{dt} \|h\|_2^2 = - \int h g \partial_v h dv + \int h \bar{h} dv + \nu \int h(v')^2 \partial_{vv} h dv. \quad (8.7)$$

The first term is controlled via integration by parts, Sobolev embedding and (2.59c),

$$- \int h g \partial_v h dv = \frac{1}{2} \int h^2 \partial_v g dv \leq \frac{1}{2} \|\partial_v g\|_\infty \|h\|_2^2 \lesssim \epsilon \langle t \rangle^{-2+K_D \epsilon/2} \|h\|_2^2. \quad (8.8)$$

By Cauchy-Schwarz, (2.52), (2.59c) and (2.61a) we have,

$$\begin{aligned} \int h \bar{h} dv &\leq \|h\|_2 \|v' \partial_v g\|_2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D \epsilon/2}} \|h\|_2 \\ &\lesssim \frac{1}{\langle t \rangle^{3/2-K_D \epsilon}} \|h\|_2^2 + \frac{\epsilon^2}{\langle t \rangle^{5/2}}; \end{aligned} \quad (8.9)$$

the choice of  $5/2$  is not sharp or significant.

Turn next to the dissipation term. By integration by parts,

$$\nu \int h(v')^2 \partial_{vv} h dv = -\nu \|v' \partial_v h\|_2^2 - 2\nu \int h v' \partial_v v' \partial_v h dv. \quad (8.10)$$

However, since  $\partial_v v' = \partial_v h$  we have by Sobolev embedding, (2.61a) and (2.59e),

$$\begin{aligned} \left| 2\nu \int h v' \partial_v v' \partial_v h dv \right| &\leq 2\nu \|h\|_\infty \|(v')^{-1}\|_\infty \|v' \partial_v h\|_2^2 \\ &\lesssim \nu \|h\|_{H^2} \|v' \partial_v h\|_2^2 \\ &\lesssim \epsilon \nu \|v' \partial_v h\|_2^2. \end{aligned} \quad (8.11)$$

Therefore, for  $\epsilon$  chosen sufficiently small, this term can be absorbed by the leading order dissipation in (8.10).

Combining (8.7), (8.8), (8.9), (8.10) and (8.11) with (2.61a) yields the differential inequality,

$$\frac{1}{2} \frac{d}{dt} \|h\|_2^2 \lesssim -\nu \|\partial_v h\|_2^2 + \frac{1}{\langle t \rangle^{3/2-K_D \epsilon}} \|h\|_2^2 + \frac{\epsilon^2}{\langle t \rangle^{5/2}}. \quad (8.12)$$

We next show that (8.12) implies  $\|h\|_2 \lesssim \epsilon \langle \nu t \rangle^{-1/4}$ . First, by the Gagliardo-Nirenberg-Sobolev inequality and the  $L^1$  control (2.54) we have

$$\|h\|_2 \lesssim \|h\|_1^{2/3} \|\partial_v h\|_2^{1/3} \lesssim \epsilon^{2/3} \|\partial_v h\|_2^{1/3}.$$

Hence, denoting  $X(t) = \|h(t)\|_2^2$ , this implies with (8.12) that for  $t \geq \nu^{-1}$  we have the following ((2.68b) is immediate from (2.66e) for earlier times),

$$\frac{d}{dt} \left( (\nu t)^{1/2} X(t) \right) \lesssim \frac{\sqrt{\nu}}{\sqrt{t}} X(t) - \nu (\nu t)^{1/2} \epsilon^{-4} X^3(t) + \frac{1}{\langle t \rangle^{3/2-K_D \epsilon}} (\nu t)^{1/2} X(t) + \frac{\epsilon^2 \nu^{1/2}}{\langle t \rangle^2}. \quad (8.13)$$

Next use the  $O(1)$  integrating factor

$$Y(t) = X(t) \exp \left[ - \int_{\nu^{-1}}^t \langle \tau \rangle^{K_D \epsilon - 3/2} d\tau \right];$$

note that  $X(t) \approx Y(t)$ . This collapses (8.13) to

$$\frac{d}{dt} \left( (\nu t)^{1/2} Y(t) \right) \lesssim \frac{\sqrt{\nu}}{\sqrt{t}} Y(t) - \nu (\nu t)^{1/2} \epsilon^{-4} Y^3(t) + \frac{\epsilon^2 \nu^{1/2}}{\langle t \rangle^2}. \quad (8.14)$$

For  $K'_h$  chosen sufficiently large (depending on the constants in (8.14)), the following is a supersolution of the differential equality corresponding to (8.14):

$$Y_S(t) = \frac{K'_h \epsilon^2}{\langle \nu t \rangle^{1/2}} + \frac{K'_h}{\langle \nu t \rangle^{1/2}} \int_{\nu^{-1}}^t \frac{\epsilon^2 \nu^{1/2}}{\langle \tau \rangle^2} d\tau.$$

Moreover, for  $K'_h$  large, we have  $Y(\nu^{-1}) \leq Y_S(\nu^{-1})$  and hence by comparison we have  $Y(t) \leq Y_S(t)$  for all  $t \geq \nu^{-1}$ , from which the  $L^2$  estimate in (2.68b) follows after possibly adjusting  $K_v$ .

Turn next to the  $\partial_v h$  bound stated in (2.68b), for which we will bootstrap off the  $L^2$  decay estimate. Taking a derivative of (2.50) gives

$$\partial_t \partial_v h + \partial_v g \partial_v h + g \partial_{vv} h = \partial_v \bar{h} + \nu \partial_v ((v')^2 \partial_{vv} h).$$

Hence calculating the evolution of the  $L^2$  norm and applying integration by parts:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\partial_v h\|_2^2 &= - \int \partial_v g |\partial_v h|^2 dv - \int \partial_v h g \partial_{vv} h dv \\ &\quad + \int \partial_v h \partial_v \bar{h} dv - \nu \|v' \partial_{vv} h\|_2^2 \\ &= -\frac{1}{2} \int |\partial_v h|^2 \partial_v g dv + \int \partial_v h \partial_v \bar{h} dv - \nu \|v' \partial_{vv} h\|_2^2. \end{aligned} \quad (8.15)$$

The first term is treated with Sobolev embedding and (2.59c),

$$-\frac{1}{2} \int |\partial_v h|^2 \partial_v g dv \lesssim \epsilon \langle t \rangle^{-2+K_D \epsilon/2} \|\partial_v h\|_2^2. \quad (8.16)$$

The second term in (8.15) is controlled by (2.59c) and (2.61a) via (2.52) and Sobolev embedding:

$$\begin{aligned} \left| \int \partial_v h \partial_v \bar{h} dv \right| &\leq \|\partial_v h\|_2 (\|v'\|_\infty \|\partial_{vv} h\|_2 + \|\partial_v h\|_2 \|\partial_v g\|_\infty) \\ &\lesssim \epsilon \langle t \rangle^{-2+K_D \epsilon/2} \|\partial_v h\|_2 + \epsilon \langle t \rangle^{-2+K_D \epsilon/2} \|\partial_v h\|_2^2 \\ &\lesssim \langle t \rangle^{-\frac{5}{4}+K_D \epsilon} \|\partial_v h\|_2^2 + \epsilon^2 \langle t \rangle^{-11/4}. \end{aligned} \quad (8.17)$$

We note that the  $11/4$  is not sharp or significant, however, it must be chosen larger than  $5/2$  to get the optimal decay (and cannot be as large as 3). Putting together (8.15), (8.16) and (8.17) with the bootstrap control on  $v'$ , we get

$$\frac{d}{dt} \|\partial_v h\|_2^2 \lesssim -\nu \|\partial_{vv} h\|_2^2 + \langle t \rangle^{-\frac{5}{4}+K_D \epsilon} \|\partial_v h\|_2^2 + \epsilon^2 \langle t \rangle^{-11/4}.$$

By interpolation and the  $L^2$  decay estimate in (2.61b) we have

$$\|\partial_v h\|_2 \leq \|h\|_2^{1/2} \|\partial_{vv} h\|_2^{1/2} \lesssim \langle \nu t \rangle^{-1/8} \epsilon^{1/2} \|\partial_{vv} h\|_2^{1/2},$$

which therefore implies,

$$\frac{d}{dt} \|\partial_v h\|_2^2 \lesssim -\nu \epsilon^{-2} \langle \nu t \rangle^{1/2} \|\partial_v h\|_2^4 + \langle t \rangle^{-5/4 + K_D \epsilon} \|\partial_v h\|_2^2 + \epsilon^2 \langle t \rangle^{-11/4}.$$

Following the proof of the  $L^2$  decay estimate, we define  $Y(t) = \|\partial_v h\|_2^2 \exp \left[ - \int_{\nu^{-1}}^t \langle \tau \rangle^{-5/4 + K_D \epsilon} d\tau \right]$  and for  $t \geq \nu^{-1}$  we have

$$\frac{d}{dt} \left( (\nu t)^{3/2} Y(t) \right) \lesssim \frac{(\nu t)^{3/2}}{t} Y(t) - \nu (\nu t)^2 \epsilon^{-2} Y^2(t) + \frac{\epsilon^2 \nu^{3/2}}{\langle t \rangle^{5/4}}.$$

By comparing against the supersolution (for  $K_h''$  sufficiently large),

$$Y_S(t) = \frac{K_h'' \epsilon^2}{(\nu t)^{3/2}} + \frac{K_h'' \epsilon^2}{(\nu t)^{3/2}} \int_{\nu^{-1}}^t \frac{\nu^{3/2}}{\langle \tau \rangle^{5/4}} d\tau,$$

and choosing  $K_h''$  large enough to ensure  $Y(\nu^{-1}) < Y_S(\nu^{-1})$  we deduce  $Y(t) \lesssim Y_S(t)$  for  $t \geq \nu^{-1}$ , which implies the  $\|\partial_v h\|_2$  decay estimate in (2.68b) (after possibly adjusting  $K_v$  further).

### 8.3 Zero frequency $L^2$ decay

Here we prove (2.68d), which is not necessary for the proof of the other statements in Proposition 2.2, but is included in the statement of Theorem 1 and gives the relaxation to the Couette flow. Begin by computing the evolution of the  $L^2$  norm of the zero frequency,

$$\frac{1}{2} \frac{d}{dt} \|f_0\|_2^2 = - \int f_0 g \partial_v f_0 dv - \int f_0 v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla f > dv + \nu \int f_0 (v')^2 \partial_{vv} f dv. \quad (8.18)$$

The first term is controlled via integration by parts, Sobolev embedding and (2.59c),

$$- \int f_0 g \partial_v f_0 dv = \frac{1}{2} \int |f_0|^2 \partial_v g dv \lesssim \epsilon \langle t \rangle^{-2 + K_D \epsilon/2} \|f_0\|_2^2. \quad (8.19)$$

Using the same proof as in (8.3), it follows from Lemma 4.1, (2.61a) and (2.59a) that

$$- \int f_0 v' < \nabla^\perp P_{\neq 0} \phi \cdot \nabla f > dv \lesssim \frac{\epsilon^2}{\langle t \rangle^2} \|f_0\|_2 \lesssim \frac{\epsilon}{\langle t \rangle^{3/2}} \|f_0\|_2^2 + \frac{\epsilon^3}{\langle t \rangle^{5/2}}. \quad (8.20)$$

Turn next to the dissipation term:

$$\nu \int f_0 (v')^2 \partial_{vv} f dv = -\nu \|v' \partial_v f_0\|_2^2 - 2\nu \int v' \partial_v v' f_0 \partial_v f dv. \quad (8.21)$$

By Hölder's inequality, the Gagliardo-Nirenberg-Sobolev inequality (8.5), and (2.65) we have

$$\begin{aligned} \nu \int f_0 v'' \partial_v f_0 dv &\leq \nu \|f_0\|_\infty \|v''\|_2 \|\partial_v f_0\|_2 \\ &\lesssim \nu \|f_0\|_2^{1/2} \|v''\|_2 \|\partial_v f_0\|_2^{3/2} \\ &\lesssim \epsilon^{-3} \nu \|v''\|_2^4 \|f_0\|_2^2 + \epsilon \nu \|\partial_v f_0\|_2^2 \\ &\lesssim \epsilon \nu \langle \nu t \rangle^{-3} \|f_0\|_2^2 + \epsilon \nu \|\partial_v f_0\|_2^2. \end{aligned} \quad (8.22)$$

Choosing  $\epsilon$  sufficiently small and putting (8.19), (8.20), (8.21), (8.22) together with (8.18) gives the differential inequality

$$\frac{1}{2} \frac{d}{dt} \|f_0\|_2^2 \lesssim -\nu \|\partial_v f_0\|_2^2 + \left( \frac{\epsilon}{\langle t \rangle^{3/2}} + \frac{\epsilon \nu}{\langle \nu t \rangle^3} \right) \|f_0\|_2^2 + \frac{\epsilon^3}{\langle t \rangle^{5/2}}, \quad (8.23)$$

(the  $5/2$  is not sharp or significant). By the Gagliardo-Nirenberg-Sobolev inequality together with (2.53) and (2.61a) we get,

$$\|f_0\|_2 \lesssim \|f_0\|_1^{2/3} \|\partial_v f_0\|_2^{1/3} \lesssim \epsilon^{2/3} \|\partial_v f_0\|_2^{1/3}. \quad (8.24)$$

As in the proof of (2.68b) in §8.2, we use the integrating factor

$$Y(t) = \|f_0(t)\|_2^2 \exp \left[ - \int_{\nu^{-1}}^t \left( \frac{\epsilon}{\langle \tau \rangle^{3/2}} + \frac{\epsilon \nu}{\langle \nu \tau \rangle^3} \right) d\tau \right],$$

together with (8.24), to reduce (8.23) to the differential inequality,

$$\frac{d}{dt} \left( (\nu t)^{1/2} Y(t) \right) \lesssim \frac{\sqrt{\nu}}{\sqrt{t}} Y(t) - \nu (\nu t)^{1/2} \epsilon^{-4} Y^3(t) + \frac{\epsilon^3 \nu^{1/2}}{\langle t \rangle^2}. \quad (8.25)$$

For  $K_f$  chosen sufficiently large (depending on the constants in (8.25)), the following is a supersolution of the differential equality corresponding to (8.25):

$$Y_S(t) = \frac{K_f \epsilon^2}{\langle \nu t \rangle^{1/2}} + \frac{K_f}{\langle \nu t \rangle^{1/2}} \int_{\nu^{-1}}^t \frac{\epsilon^3 \nu^{1/2}}{\langle \tau \rangle^2} d\tau.$$

For  $K_f$  large, we have  $Y(\nu^{-1}) \leq Y_S(\nu^{-1})$  and hence by comparison it follows that  $Y(t) \leq Y_S(t)$  for all  $t \geq \nu^{-1}$ , from which we deduce the  $L^2$  decay stated in (2.68d).

## Acknowledgments

The authors would like to thank the following people for references and suggestions: Margaret Beck, Steve Childress, Peter Constantin, Pierre Germain, Yan Guo and Gene Wayne. The work of JB was in part supported by NSF Postdoctoral Fellowship in Mathematical Sciences DMS-1103765 and NSF grant DMS-1413177, the work of NM was in part supported by the NSF grant DMS-1211806, while the work of VV was in part supported by the NSF grant DMS-1348193.

## A Appendix

### 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. [13, 1] for more details.

For  $f(z, v)$  in the Schwartz space, we define the Fourier transform  $\hat{f}_k(\eta)$  where  $(k, \eta) \in \mathbb{Z} \times \mathbb{R}$ ,

$$\hat{f}_k(\eta) = \frac{1}{2\pi} \int_{\mathbb{T} \times \mathbb{R}} e^{-izk - i\eta\eta} f(z, v) dz dv.$$

Similarly we have the Fourier inversion formula,

$$f(z, v) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} e^{izk + iv\eta} \hat{f}_k(\eta) d\eta.$$

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,

$$\begin{aligned} \int f(z, v) \bar{g}(z, v) dz dv &= \sum_k \int \hat{f}_k(\eta) \bar{g}_k(\eta) d\eta \\ \widehat{fg} &= \frac{1}{2\pi} \hat{f} * \hat{g} \\ (\widehat{\nabla f})_k(\eta) &= (ik, i\eta) \hat{f}_k(\eta). \end{aligned}$$

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. [1] for more details. First we define the Littlewood-Paley decomposition only in the  $v$  variable. Let  $\psi \in C_0^\infty(\mathbb{R}; \mathbb{R})$  be such that  $\psi(\xi) = 1$  for  $|\xi| \leq 1/2$  and  $\psi(\xi) = 0$  for  $|\xi| \geq 3/4$  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

$$1 = \psi(\xi) + \sum_{M \in 2^{\mathbb{N}}} \rho(M^{-1}\xi),$$

where we mean that the sum runs over the dyadic numbers  $M = 1, 2, 4, 8, \dots, 2^j, \dots$  and we define the cut-off  $\rho_M(\xi) = \rho(M^{-1}\xi)$ , each supported in  $M/2 \leq |\xi| \leq 3M/2$ . For  $f \in L^2(\mathbb{R})$  we define

$$\begin{aligned} f_M &= \rho_M(|\partial_v|)f, \\ f_{\frac{1}{2}} &= \psi(|\partial_v|)f, \\ f_{<M} &= f_{\frac{1}{2}} + \sum_{K \in 2^{\mathbb{N}}: K < M} f_K, \end{aligned}$$

which defines the decomposition

$$f = f_{\frac{1}{2}} + \sum_{M \in 2^{\mathbb{N}}} f_M.$$

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  $\mathbb{D}$  from §1.2. There holds the almost orthogonality and the approximate projection property

$$\|f\|_2^2 \approx \sum_{M \in \mathbb{D}} \|f_M\|_2^2 \tag{A.1a}$$

$$\|f_M\|_2 \approx \|(f_M)_M\|_2. \tag{A.1b}$$

The following is also clear:

$$\| |\partial_v| f_M \|_2 \approx M \|f_M\|_2.$$

We make use of the notation

$$f_{\sim M} = \sum_{K \in \mathbb{D}: \frac{1}{C}M \leq K \leq CM} f_K,$$

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  $\frac{1}{C}2^k \subset \text{supp } D_k \subset C2^k$  it follows that

$$\|f\|_2^2 \approx_C \sum_k \|D_k\|_2^2. \quad (\text{A.2})$$

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

$$\Delta_j f = f_{2^j}, \quad S_j f = f_{< 2^j},$$

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 [13] (see also [1]). Given suitable functions  $f, g$  we may define the *inhomogeneous* paraproduct decomposition (in either  $(z, v)$  or just  $v$ ),

$$\begin{aligned} fg &= T_f g + T_g f + \mathcal{R}(f, g) \\ &= \sum_{N \geq 8} f_{< N/8} g_N + \sum_{N \geq 8} g_{< N/8} f_N + \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} g_{N'} f_N, \end{aligned} \quad (\text{A.3})$$

where all the sums are understood to run over  $\mathbb{D}$ , or the *homogeneous* paraproduct

$$fg = \sum_{N \in 2^{\mathbb{Z}}} f_{< N/8} g_N + \sum_{N \in 2^{\mathbb{Z}}} g_{< N/8} f_N + \sum_{N \in 2^{\mathbb{Z}}} \sum_{N/8 \leq N' \leq 8N} g_{N'} f_N.$$

The choice of which one we employ depends on the role that low frequencies are playing in the proof. In our work we do not employ the notation in (A.3) since at most steps we are working in non-standard regularity spaces and are usually applying multipliers which do not satisfy any version of  $AT_f g \approx T_f A g$ . Hence, we normally have to prove most everything by hand and 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  $f(\xi), g(\xi) \in L_{\xi}^2(\mathbb{R}^d)$ ,  $\langle \xi \rangle^{\sigma} h(\xi) \in L_{\xi}^2(\mathbb{R}^d)$  and  $\langle \xi \rangle^{\sigma} b(\xi) \in L_{\xi}^2(\mathbb{R}^d)$  for  $\sigma > d/2$ . Then we have*

$$\|f * h\|_2 \lesssim_{\sigma, d} \|f\|_2 \|\langle \cdot \rangle^{\sigma} h\|_2, \quad (\text{A.4})$$

$$\int |f(\xi)(g * h)(\xi)| d\xi \lesssim_{\sigma, d} \|f\|_2 \|g\|_2 \|\langle \cdot \rangle^{\sigma} h\|_2 \quad (\text{A.5})$$

$$\int |f(\xi)(g * h * b)(\xi)| d\xi \lesssim_{\sigma, d} \|f\|_2 \|g\|_2 \|\langle \cdot \rangle^{\sigma} h\|_2 \|\langle \cdot \rangle^{\sigma} b\|_2. \quad (\text{A.6})$$



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  $0 < s < 1$ ). They are crucial for doing paradifferential calculus in Gevrey regularity.

**Lemma A.2.** *Let  $0 < s < 1$  and  $x \geq y \geq 0$  (without loss of generality).*

(i) *If  $x + y > 0$ ,*

$$|x^s - y^s| \lesssim_s \frac{1}{x^{1-s} + y^{1-s}} |x - y|. \quad (\text{A.7})$$

(ii) *If  $|x - y| \leq x/K$  for some  $K > 1$  then*

$$|x^s - y^s| \leq \frac{s}{(K-1)^{1-s}} |x - y|^s. \quad (\text{A.8})$$

*Note  $\frac{s}{(K-1)^{1-s}} < 1$  as soon as  $s^{\frac{1}{1-s}} + 1 < K$ .*

(iii) *In general,*

$$|x + y|^s \leq \left( \frac{x}{x + y} \right)^{1-s} (x^s + y^s). \quad (\text{A.9})$$

*In particular, if  $\frac{1}{K}y \leq x \leq Ky$  for some  $K < \infty$  then*

$$|x + y|^s \leq \left( \frac{K}{1 + K} \right)^{1-s} (x^s + y^s). \quad (\text{A.10})$$

Using (A.5), (A.8) and (A.10) 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  $c = c(s, \sigma, \sigma_0) \in (0, 1)$  such that the following holds for all  $f, g \in \mathcal{G}^{\lambda, \sigma; s}$ :*

$$\|fg\|_{\mathcal{G}^{\lambda, \sigma}} \lesssim \|f\|_{\mathcal{G}^{c\lambda, \sigma_0}} \|g\|_{\mathcal{G}^{\lambda, \sigma}} + \|g\|_{\mathcal{G}^{c\lambda, \sigma_0}} \|f\|_{\lambda, \sigma}, \quad (\text{A.11a})$$

*in particular,  $\mathcal{G}^{\lambda, \sigma; s}$  has the algebra property:*

$$\|fg\|_{\mathcal{G}^{\lambda, \sigma}} \lesssim \|f\|_{\mathcal{G}^{\lambda, \sigma}} \|g\|_{\mathcal{G}^{\lambda, \sigma}} \quad (\text{A.12})$$

Gevrey and Sobolev regularities can be related with the following two inequalities.

(i) For all  $x \geq 0$ ,  $\alpha > \beta \geq 0$ ,  $C, \delta > 0$ ,

$$e^{Cx^\beta} \leq e^{C(\frac{C}{\delta})^{\frac{\beta}{\alpha-\beta}}} e^{\delta x^\alpha}; \quad (\text{A.13})$$

(ii) For all  $x \geq 0$ ,  $\alpha, \sigma, \delta > 0$ ,

$$e^{-\delta x^\alpha} \lesssim \frac{1}{\delta^{\frac{\sigma}{\alpha}} \langle x \rangle^\sigma}. \quad (\text{A.14})$$

Together these inequalities show that for  $\alpha > \beta \geq 0$ ,  $\|f\|_{\mathcal{G}^{\beta; C, \sigma}} \lesssim_{\alpha, \beta, C, \delta, \sigma} \|f\|_{\mathcal{G}^{\alpha; \delta, 0}}$ .

## References

- [1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.
- [2] K. Bajer, A.P. Bassom, and A.D. Gilbert. Accelerated diffusion in the centre of a vortex. *Journal of Fluid Mechanics*, 437:395–411, 2001.
- [3] N.J. Balmforth and P.J. Morrison. Normal modes and continuous spectra. *Annals of the New York Academy of Sciences*, 773(1):80–94, 1995.
- [4] N.J. Balmforth, P.J. Morrison, and J.-L. Thiffeault. Pattern formation in Hamiltonian systems with continuous spectra; a normal-form single-wave model. *preprint*, 2013.
- [5] C. Bardos, E.S. Titi, and E. Wiedemann. The vanishing viscosity as a selection principle for the Euler equations: The case of 3D shear flow. *Comptes Rendus Mathematique*, 350(15):757–760, 2012.
- [6] A.P. Bassom and A.D. Gilbert. The spiral wind-up of vorticity in an inviscid planar vortex. *J. Fluid Mech.*, 371:109–140, 1998.
- [7] G.K. Batchelor. Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Physics of fluids*, 12(12):II–233, 1969.
- [8] M. Beck and C.E. Wayne. Metastability and rapid convergence to quasi-stationary bar states for the 2D Navier-Stokes equations. *arXiv:1108.3416*, 2011.
- [9] J. Bedrossian and N. Masmoudi. Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations. *arXiv:1306.5028*, 2013.
- [10] J. Bedrossian, N. Masmoudi, and C. Mouhot. Landau damping: paraproducts and Gevrey regularity. *arXiv:1311.2870*, 2013.
- [11] H. Berestycki, F. Hamel, and N. Nadirashvili. Elliptic eigenvalue problems with large drift and applications to nonlinear propagation phenomena. *Comm. Math. Phys.*, 253(2):451–480, 2005.
- [12] A.J. Bernoff and J.F. Lingeitch. Rapid relaxation of an axisymmetric vortex. *Phys. Fluids*, 6(3717), 1994.
- [13] J.M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann.Sc.E.N.S.*, 14:209–246, 1981.
- [14] F. Bouchet and H. Morita. Large time behavior and asymptotic stability of the 2D Euler and linearized Euler equations. *Physica D*, 239:948–966, 2010.
- [15] J.P. Boyd. The continuous spectrum of linear Couette flow with the beta effect. *Journal of the atmospheric sciences*, 40(9):2304–2308, 1983.
- [16] R.J. Briggs, J.D. Daugherty, and R.H. Levy. Role of Landau damping in crossed-field electron beams and inviscid shear flow. *Phys. Fl.*, 13(2), 1970.
- [17] T. Buckmaster, C. De Lellis, and L. Székelyhidi Jr. Dissipative euler flows with onsager-critical spatial regularity. *arXiv: 1404.6915*, 04 2014.

- [18] E. Caglioti and C. Maffei. Time asymptotics for solutions of Vlasov-Poisson equation in a circle. *J. Stat. Phys.*, 92(1/2), 1998.
- [19] O. Cardoso and P. Tabeling. Anomalous diffusion in a linear array of vortices. *EPL (Europhysics Letters)*, 7(3):225, 1988.
- [20] S.J. Chapman. Subcritical transition in channel flows. *Journal of Fluid Mechanics*, 451:35–98, 2002.
- [21] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168(2):643–674, 2008.
- [22] C. De Lellis and L. Székelyhidi, Jr. Dissipative Euler flows and Onsager’s conjecture. *arXiv preprint arXiv:1205.3626*, 2012.
- [23] C. De Lellis and L. Székelyhidi Jr. The h-principle and the equations of fluid dynamics. *Bulletin of the American Mathematical Society*, 49(3):347–375, 2012.
- [24] P.G. Drazin and W.H. Reid. *Hydrodynamic stability*. Cambridge U. Press, Cambridge, 1981.
- [25] E. Faou and F. Rousset. Landau damping in sobolev spaces for the Vlasov-HMF model. *arXiv:1403.1668*, 2014.
- [26] C. Foias and R. Temam. Gevrey class regularity for solutions of the Navier-Stokes equations. *J. Funct. Anal.*, 87:359–369, 1989.
- [27] M. Gevrey. Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire. *Ann. Sci. École Norm. Sup. (3)*, 35:129–190, 1918.
- [28] A.D. Gilbert. Spiral structures and spectra in two-dimensional turbulence. *J. Fluid Mech.*, 193:475–497, 1988.
- [29] A.D. Gilbert. A cascade interpretation of lundgren’s stretched spiral vortex model for turbulent fine structure. *Physics of Fluids A: Fluid Dynamics*, 5:2831, 1993.
- [30] E. Grenier, Y. Guo, and T. Nguyen. Spectral instability of characteristic boundary layer flows. *arXiv:1406.3862*, 06 2014.
- [31] P.H. Haynes and J. Vanneste. Dispersion in the large-deviation regime. Part 1: shear flows and periodic flows. *Journal of Fluid Mechanics*, 745:321–350, 2014.
- [32] H.J. Hwang and J.J.L. Velaázquez. On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem. *Indiana Univ. Math. J.*, pages 2623–2660, 2009.
- [33] P. Isett. Hölder continuous Euler flows in three dimensions with compact support in time. *arXiv preprint arXiv:1211.4065*, 2012.
- [34] G. Iyer and A. Novikov. Anomalous diffusion in fast cellular flows at intermediate time scales. *arXiv:1406.3881*, 06 2014.
- [35] Lord Kelvin. Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates. *Phil. Mag.*, (24):188, 1887.
- [36] A.N. Kolmogorov. Dissipation of energy in locally isotropic turbulence. In *Dokl. Akad. Nauk SSSR*, volume 32, pages 16–18, 1941.

- [37] R.H. Kraichnan. Inertial ranges in two-dimensional turbulence. *Phys. Fluids*, 10(7), 1967.
- [38] I. Kukavica and V. Vicol. On the radius of analyticity of solutions to the three-dimensional Euler equations. *Proc. Amer. Math. Soc.*, 137(2):669–677, 2009.
- [39] L. Landau. On the vibration of the electronic plasma. *J. Phys. USSR*, 10(25), 1946.
- [40] M. Latini and A.J. Bernoff. Transient anomalous diffusion in poiseuille flow. *Journal of Fluid Mechanics*, 441:399–411, 2001.
- [41] D. Levermore and M. Oliver. Analyticity of solutions for a generalized Euler equation. *J. Diff. Eqns.*, 133:321–339, 1997.
- [42] R. Lindzen. Instability of plane parallel shear flow (toward a mechanistic picture of how it works). *PAGEOPH*, 126(1), 1988.
- [43] T.S. Lundgren. Strained spiral vortex model for turbulent fine structure. *Phys. of Fl.*, 25:2193, 1982.
- [44] J. Malmberg, C. Wharton, C. Gould, and T. O’Neil. Plasma wave echo. *Phys. Rev. Lett.*, 20(3):95–97, 1968.
- [45] C. Mouhot and C. Villani. On Landau damping. *Acta Math.*, 207:29–201, 2011.
- [46] W. Orr. The stability or instability of steady motions of a perfect liquid and of a viscous liquid, Part I: a perfect liquid. *Proc. Royal Irish Acad. Sec. A: Math. Phys. Sci.*, 27:9–68, 1907.
- [47] Lord Rayleigh. On the Stability, or Instability, of certain Fluid Motions. *Proc. London Math. Soc.*, S1-11(1):57, 1880.
- [48] P.B. Rhines and W.R. Young. How rapidly is a passive scalar mixed within closed streamlines? *Journal of Fluid Mechanics*, 133:133–145, 1983.
- [49] D.D. Ryutov. Landau damping: half a century with the great discovery. *Plasma physics and controlled fusion*, 41(3A):A1, 1999.
- [50] D.A. Schecter, D. Dubin, A.C. Cass, C.F. Driscoll, and I.M. Lansky et. al. Inviscid damping of asymmetries on a two-dimensional vortex. *Phys. Fl.*, 12, 2000.
- [51] G. Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 219(1137):186–203, 1953.
- [52] J. Vanneste. Nonlinear dynamics of anisotropic disturbances in plane Couette flow. *SIAM J. Appl. Math.*, 62(3):924–944 (electronic), 2002.
- [53] J. Vanneste, P.J. Morrison, and T. Warn. Strong echo effect and nonlinear transient growth in shear flows. *Physics of Fluids*, 10:1398, 1998.
- [54] C. Villani. *Hypocoercivity*. American Mathematical Soc., 2009.
- [55] A.M. Yaglom. *Hydrodynamic Instability and Transition to Turbulence*, volume 100. Springer, 2012.

- [56] W. Young, A. Pumir, and Y. Pomeau. Anomalous diffusion of tracer in convection rolls. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 1(3):462–469, 1989.
- [57] J.H. Yu and C.F. Driscoll. Diocotron wave echoes in a pure electron plasma. *IEEE Trans. Plasma Sci.*, 30(1), 2002.
- [58] J.H. Yu, C.F. Driscoll, and T.M. O’Neil. Phase mixing and echoes in a pure electron plasma. *Phys. of Plasmas*, 12(055701), 2005.