PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2155–2184)

THE ORR MECHANISM: STABILITY/INSTABILITY OF THE COUETTE FLOW FOR THE 2D EULER DYNAMIC

JACOB BEDROSSIAN, YU DENG AND NADER MASMOUDI

Abstract

We review our works on the nonlinear asymptotic stability and instability of the Couette flow for the 2D incompressible Euler dynamic. In the fits part of the work we prove that perturbations to the Couette flow which are small in Gevrey spaces G^s of class $1/s$ with $s > 1/2$ converge strongly in L^2 to a shear flow which is close to the Couette flow. Moreover in a well chosen coordinate system, the solution converges in the same Gevrey space to some limit profile. In a later work, we proved the existence of small perturbations in G^s with $s < 1/2$ such that the solution becomes large in Sobolev regularity and hence yields instability. In this note we discuss the most important physical and mathematical aspects of these two results and the key ideas of the proofs.

Contents

MSC2010: primary 35Q31; secondary 35B35.

1 Introduction and physical background

The theory of hydrodynamic stability at high Reynolds number started already in the 19th century, with the likes of Stokes, Reynolds, Kelvin, Orr and others. Some of the first early theoretical works were done by Rayleigh [1879/80, 1887/88, 1895/96], including for example, the inflection point theorem of the spectral instability on inviscid planar shear flows, and the exact solutions for Coutte flow in the absence of boundaries constructed by Kelvin [1887] which showed linear stability independent of Reynolds number. These solutions were later followed up by Orr [1907], Dikiĭ [1961], and Case [1960] to show linearized stability of the inviscid Couette flow also in a channel. The early experiments of Reynolds [1883] clearly showed instability at all high Reynolds number for flow in a pipe, which although that linearized problem still to this day has not been conclusively solved, seemed slightly in contradiction with most of the theoretical results of the time. Kelvin proposed the solution to this 'paradox' (sometimes called the 'Sommerfeld paradox' Li and Lin [2011]): although the fluid might be stable at all Reynolds numbers, as the Reynolds number increases, the fluid becomes increasing sensitive to small perturbations. This phenomena now often called *subcritical transition* and it is ubiquitous in 3D fluid mechanics (as well as plasma physics). It has also been observed in cases with spectral instability at high Reynolds number in the sense that the flows often go unstable at lower Reynolds number than that predicted by linearized theory or in a way completely unrelated to the unstable eigenvalues; see e.g. Drazin and Reid [1982], Schmid and Henningson [2001], and Yaglom [2012] and the references therein for discussions on this effect in the physics literature. In the case of linearized stability at all Reynolds number, we could then phrase the following question: given a norm $\|\cdot\|_X$, and an initial perturbation $\tilde{u}$, we could look for a $\gamma = \gamma(X)$ such that:

(1a) $\quad \|\tilde{u}\|_X \lesssim \mathbf{Re}^{-\gamma} \Rightarrow$ stability and linear-dominated behavior $\forall t$,

(1b) $\quad \|\tilde{u}\|_X \gg \mathbf{Re}^{-\gamma} \Rightarrow$ (potential) nonlinear-dominated behavior and instability.

The exponent γ is sometimes called the *transition threshold*. More recent insights show that there are two interesting aspects of this effect:

- a/ The linearized problem may contain transient growth, and these could be triggering nonlinear instabilities, especially if the growth becomes larger as Reynolds number goes to infinity.
- b/ The dependence on on the stability threshold may depend strongly on the topology in which one measures perturbations. It is this aspect that we will be reviewing here.

Modern research shows several fundamental differences between 2D and 3D hydrodynamic stability, both at the linear and nonlinear level. This is due to the fact that (A) the 3D linearized equations have a wider range of more extreme transient growth mechanisms than 2D and (B) 3D has a much more complicated structure of 'resonances', that is, the weakly nonlinear structure is much more complicated (see below for more discussion). Unlike what might be suggested by Squire's theorem Squire [1933], as a result, 2D stability studies do not seem to give significant physical insight into 3D fluids in the sense that theoretical or numerical results on 2D hydrodynamic stability give no specific, direct information on 3D flows. However, many important physical applications are well-approximated to leading order by 2D fluids, such as many atmospheric and oceanic phenomena, so it is still important to give careful consideration to 2D fluids. Moreover, when it comes to hydrodynamic stability questions, the dynamics of 2D fluids are significantly simpler than 3D fluids in many ways, and hence it is reasonable to begin mathematical studies with the former rather than the latter: a theorem about 2D flow in a channel does not give much specific physical insight into 3D flow in a pipe, but the mathematics developed therein hopefully will. Indeed, this was clear in the works Bedrossian and Masmoudi [2013] and Bedrossian, Masmoudi, and Vicol [2016] vs Bedrossian, Germain, and Masmoudi [2015a,b, 2017a]: the dynamics and nonlinear structures might be very different in 2D and 3D, but nevertheless, the subsequent proofs all used certain mathematical tools originally designed for the stability of 2D Euler in Bedrossian and Masmoudi [2013], or at least used ideas heavily influenced by the insight obtained therein.

With the above discussions in mind, we will focus in this paper on the 2D case, and even mostly on the simpler case of infinite Reynolds number, e.g. the 2D incompressible Euler equations (we will see that the inviscid problem is a reasonable place to start, even though (1) is phrased in terms of Reynolds number). See Bedrossian, Germain, and Masmoudi [2017b] for a review of the related 3D stability problems. Moreover, we are interested in nonlinear questions, and progress has mostly only been made on one shear flow: the Couette flow $u = (y, 0)$ on $(x, y) \in \mathbb{T} \times \mathbb{R}$. In this case, the 2D Euler system in the vorticity formulation with the background shear flow becomes:

$$\begin{cases} \omega_t + y\partial_x\omega + U \cdot \nabla\omega = 0, \\ U = \nabla^\perp(\Delta)^{-1}\omega, \qquad \omega(t=0) = \omega_{in}. \end{cases} \tag{2}$$

Here, $(x, y) \in \mathbb{T} \times \mathbb{R}$, $\nabla^\perp = (-\partial_y, \partial_x)$ and (U, ω) are periodic in the x variable with period normalized to 2π. The physical velocity is $(y, 0) + U$ where $U = (U^x, U^y)$ denotes the velocity perturbation and the total vorticity is $-1 + \omega$. In what follows, we define the streamfunction $\psi = \Delta^{-1}\omega$.

We first state the result of Bedrossian and Masmoudi [2013] and then attempt to elucidate some of the interesting physical and mathematical concepts which are involved in the proof. We then state the instability result of Deng and Masmoudi [2018] which shows the criticality of the $G^{1/2}$ space, namely the Gevrey space of class 2. It is worth noting that this space appears here due to a nonlinear mechanism, not a linear one. The relationship with Landau damping in the Vlasov equations of plasma physics and the recent work of Mouhot and Villani [2011] will also be discussed.

2 Linear dynamics

Linearizing the 2D Euler equations as written in (2) just means dropping the quadratic term:

$$\begin{cases} \partial_t \omega + y\partial_x \omega = 0 \\ \Delta\phi = \omega. \end{cases} \tag{3}$$

This is readily solved:

$$\begin{aligned} \omega(t,x,y) &= \omega_{in}(x-ty,y) \\ \widehat{\omega}(t,k,\eta) &= \widehat{\omega_{in}}(k,\eta+kt). \end{aligned} \tag{4}$$

From (4) we can see a linear-in-time transfer of enstrophy to high frequencies. Since Δ^{-1} gains two derivatives, we should intuitively guess that $P_{k\neq 0}\phi$ decays like $\langle t\rangle^{-2}$ in L^2. This might seem circuitous for solving the linear problem, but let us follow Kelvin and Orr and introduce the following change of variables:

$$z = x - ty \tag{5a}$$

$$f(t,z,y) = \omega(t,z+ty,y) = \omega(t,x,y) \tag{5b}$$

$$\phi(t,z,y) = \psi(t,z+ty,y) = \psi(t,x,y). \tag{5c}$$

This is nothing more than rewinding by the linear propagator associated to the Couette flow. From (3) we have

$$\partial_t f = 0 \tag{6a}$$

$$\partial_{zz}\phi + (\partial_y - t\partial_x)^2\phi = f, \tag{6b}$$

which gives:

$$\hat{\phi}(t,k,\eta) = -\frac{\hat{f}(t,k,\eta)}{k^2+|\eta-kt|^2} = -\frac{\hat{\omega}_{in}(k,\eta)}{k^2+|\eta-kt|^2}. \tag{7}$$

From (7) we derive the fundamental decay-by-mixing estimate: for any $\sigma \in [0,\infty)$ and $\beta \in [0,2]$,

$$\|P_{\neq 0}\phi\|_{H^\sigma} \lesssim \frac{1}{\langle t\rangle^\beta}\|f\|_{H^{\sigma+\beta}} = \frac{1}{\langle t\rangle^\beta}\|\omega_{in}\|_{H^{\sigma+\beta}}, \tag{8}$$

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

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

we get the inviscid damping predicted by Orr [1907]:

$$\|P_{\neq 0}U^x\|_{L^2} \lesssim \langle t\rangle \|\nabla\phi\|_{L^2} \lesssim \langle t\rangle^{-1}\|\omega_{in}\|_{H^3} \tag{9a}$$
$$\|U^y\|_{L^2} \lesssim \|\partial_x\phi\|_{L^2} \lesssim \langle t\rangle^{-2}\|\omega_{in}\|_{H^3}. \tag{9b}$$

This shows that on the linear level, we have the convergence $(y+U^x,U^y) \to (y+ < U^x_{in} > (y), 0)$ in L^2 as time goes to infinity. Hence, the velocity field converges strongly back to a shear flow but not back to the Couette flow. As discussed in the introduction, Orr had a second observation from (7), which is that modes with $\eta k > 0$ undergo first a transient growth in ϕ before decaying. Indeed, the denominator (7) is minimal the time $t = \frac{\eta}{k}$, which Orr called the *critical time*. Therefore, if $|\eta| >> |k|$, the velocity field is *amplified* by a large factor between $t = 0$ and the critical time $t = \frac{\eta}{k}$. In physical terms, this transient growth is due to the fact that the mode of the vorticity in question is initially well-mixed, and then proceeds to *unmix* under the Couette flow evolution. See figure (1) for how this mixing/un-mixing effect appears on each Fourier mode of the vorticity. The relevance of the Orr mechanism to hydrodynamic stability has been debated over the years; see e.g. Orr [1907], Boyd [1983], and Lindzen [1988] and Yaglom [2012] for a detailed account of how the literature on the topic developed. However, our work verifies the crucial importance of the Orr mechanism for nonlinear stability problems at high Reynolds numbers in 2D fluid mechanics, or at least for the Couette flow.

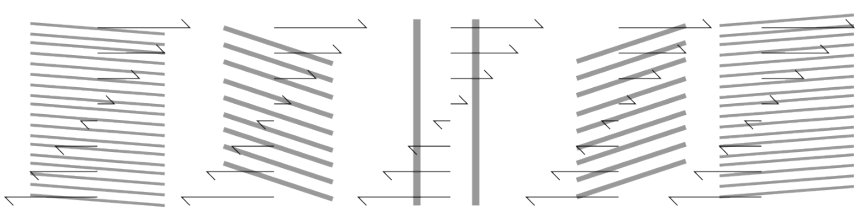

Figure 1: A mode-by-mode visualization of the Orr mechanism: the arrows represent the background flow, while the stripes are the level sets of the function $e^{ik(x-ty)+i\eta y}$ with $\eta/k \gg 1$. Time increases from left to right, and the center image is the critical time $t = \eta/k$. The full linearized solution is simply a superposition of these tilting waves.

3 Overview of the mathematical results

In this section we will summarize the results of Bedrossian and Masmoudi [2013] and Deng and Masmoudi [2018] in the 2D Euler equations. Each has analogies in the Vlasov-Poisson equations of kinetic theory. The original work of Mouhot and Villani proved the nonlinear Landau damping in $\mathbb{T}^d \times \mathbb{R}^d$, as predicted by the linearized Vlasov (see also Bedrossian, Masmoudi, and Mouhot [2016b]). These results are the analogue of the positive stability results of Bedrossian and Masmoudi [2013] (Theorem 1 below), though broadly speaking, Theorem 1 seems to require a much more subtle proof for several reasons. For the instability results, something related to Theorem 2 below was proved previously for Vlasov-Poisson in Bedrossian [2016] in Sobolev spaces; both the proof and the nature of the nonlinear behavior demonstrated by the solutions are different, but closely related.

3.1 Stability result.

The main result of Bedrossian and Masmoudi [2013] is the following, which shows that if one is willing to take Gevrey-1/2 regularity on the initial data, the *nonlinear* 2D Euler equations display an inviscid damping essentially the same as that predicted by Orr.

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

(10) $$\int_{\mathbb{T}\times\mathbb{R}} (1+|y|)\cdot|\omega_{in}(x,y)|\,\mathrm{d}x\mathrm{d}y \leq \varepsilon, \quad \int_{\mathbb{T}\times\mathbb{R}} \omega_{in}(x,y)\,\mathrm{d}x\mathrm{d}y = 0 \quad \textit{and}$$

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

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

(11) $$\|\omega(t, x+ty+\Phi(t,y), y) - f_\infty(x,y)\|_{\lambda'} \lesssim \frac{\epsilon^2}{\langle t\rangle},$$

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

(12) $$\Phi(t,y) = \frac{1}{2\pi}\int_0^t \int_{\mathbb{T}} U^x(s,x,y)dxds = u_\infty(y)t + \theta(t,y),$$

with $u_\infty = \partial_y \partial_{yy}^{-1} \frac{1}{2\pi} \int_{\mathbb{T}} f_\infty(x,y)dx$ *and* $|\theta(t,y)| \lesssim \epsilon^2$*. Moreover, the velocity field* U *decays as*

$$\left\| \int U^x(t,x,\cdot)dx - u_\infty \right\|_{\lambda'} \lesssim \frac{\epsilon^2}{\langle t \rangle}, \tag{13a}$$

$$\left\| U^x(t) - \int U^x(t,x,\cdot)dx \right\|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle}, \tag{13b}$$

$$\| U^y(t) \|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle^2}. \tag{13c}$$

The above result was extended to a uniform-in-Reynolds number statement about the 2D Navier-Stokes equations in Bedrossian, Masmoudi, and Vicol [2016] (note, that such a statement is always strictly harder than an infinite Reynolds number result). This work also shows that the mixing due to the Couette flow enhances the effect of the viscosity, an effect which plays an important role in 3D as well Bedrossian, Germain, and Masmoudi [2015a,b, 2017a]. These papers show that for 2D Couette flow, in Gevrey-2 regularity, there is *no subcritical transition*. Note that in 3D, there *is* subcritical transition, even in Gevrey-2 Bedrossian, Germain, and Masmoudi [2015b].

3.2 Instability result. It is known that the dynamics of Theorem 1 may not happen in low regularities. In Lin and Zeng [2011a], time periodic solutions to Equation (2) are constructed in Sobolev spaces H^s where $s < 3/2$; for Vlasov-Poisson equations, the same result was proved in Lin and Zeng [2011b] and, as mentioned before, Bedrossian [2016] has proved instability in any Sobolev space H^s.

The main result of Deng and Masmoudi [2018] fills the gap between these stability and instability results, by proving instability of the $2D$ Couette flow in any Gevrey-s regularity for $s < 1/2$. In fact something slightly stronger is proved, where Gevrey-s is replaced by a log-corrected version of Gevrey-$1/2$.

Theorem 2. *Let* $N_0 = 9000$*,* $N_1 = 60000$*, and denote* $\log^+(x) = \log(2 + |x|)$*. For a function* $f : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$*, define the Gevrey-type norm* $\mathcal{G}^*$ *by*

$$\|F\|_{\mathcal{G}^*}^2 = \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}} e^{2\kappa(k,\xi)} |\widehat{F}(k,\xi)|^2 \, \mathrm{d}\xi, \quad \kappa(k,\xi) = \frac{(|k|+|\xi|)^{1/2}}{(\log^+(|k|+|\xi|))^{N_1}}. \tag{14}$$

Then, for any sufficiently small $\varepsilon > 0$*, there exists a solution* $\omega = \omega(t,x,y)$ *to* (2)*, such that:*

1. *The initial data* $\omega(0)$ *satisfies the assumptions (10), and that*

$$\|\omega(0)\|_{\mathcal{G}^*} \le \varepsilon; \tag{15}$$

2. *At some later time T, the solution ω satisfies that*

$$\|\langle\partial_x\rangle^{N_0}\omega(T,x,y)\|_{L^2(\mathbb{T}\times\mathbb{R})} \geq \frac{1}{\varepsilon}. \tag{16}$$

4 The nonlinear dynamics: The toy model

In both of the theorems above, we are interested in understanding the *weakly nonlinear* effects. It is a classical idea that transient growth in a linear problem can interact badly with the nonlinearity to trigger instability, for instance see the discussion in L. N. Trefethen, A. E. Trefethen, Reddy, and T. A. Driscoll [1993]. The basic mechanism is as follows. Heuristically, in the weakly nonlinear regime we can imagine the solution as an interacting superposition of waves undergoing linear shear. Through the nonlinear term, each mode has a strong effect at its critical time during which it strongly forces the others, potentially putting information into modes which have not yet reached their critical time and are hence still growing. At a later time, these modes have a large effect and continue to excite other growing modes and so forth, perpetuating a so-called self-sustaining 'nonlinear bootstrap' (see L. N. Trefethen, A. E. Trefethen, Reddy, and T. A. Driscoll [1993], Baggett, T. A. Driscoll, and L. N. Trefethen [1995], Vanneste, Morrison, and Warn [1998], and Vanneste [2001/02] and the references therein for discussions in the fluid mechanics context). Since the measurable effect of a nonlinear interaction can occur long after the event, this mechanism permits *nonlinear echoes*, in which the electric field of the plasma, or kinetic energy of the fluid disturbance, is highly concentrated at specific times. These spectacular displays of reversibility were captured experimentally for Vlasov, there known as *plasma echoes*, in the work of Malmberg, Wharton, Gould, and O'Neil [1968]. The analogous 'Euler echoes' were recently studied and observed both numerically Vanneste, Morrison, and Warn [1998] and Vanneste [2001/02] and experimentally Yu and C. Driscoll [2002] and Yu, C. Driscoll, and O'Neil [2005].

The careful analysis of plasma echoes in the Vlasov equations is crucial in the proof of Mouhot and Villani [2011] (and also in Bedrossian, Masmoudi, and Mouhot [2016b,a]), as these are the dominant weakly nonlinear effect that could lead to instability. Later, the work of Bedrossian [2016] confirmed this viewpoint by constructing arbitrarily small perturbations in H^s (for any finite s) which give rise to arbitrarily many distinct nonlinear oscillations in the electric field (similar to the experiments of Malmberg, Wharton, Gould, and O'Neil [1968] but with arbitrarily long chains of echoes). This resonance is also used to prove Theorem 2.

Let us try to begin the analysis in the natural way, by first making the change $f(t,z,y) = \omega(t, z+ty, y)$. The nonlinear Euler equations then become (note the gratuitious cancellation)

$$\partial_t f + \nabla^\perp \phi \cdot \nabla f = 0 \tag{17a}$$

$$\partial_{zz}\phi + (\partial_y - t\partial_x)^2 \phi = f. \tag{17b}$$

There are actually two immediate problems. First of all, the contribution of the velocity field $\frac{1}{2\pi}\int_0^{2\pi} \partial_y \phi(t,z,y) dz$ will not decay, indeed, the linear problem leaves these modes invariant. Hence, we will unavoidably have $\|\partial_y^j f\| \approx \epsilon \langle t \rangle^j$ in general. This clearly shows that for long-time estimates, we are working in the wrong coordinate system.

Even if we ignore this problem, we have another problem. In order to obtain inviscid damping, the goal is simply to obtain uniform-in-time H^s estimates on f. Imagine we forget about the non-decaying modes and write (up to commutators that are not important here):

$$\begin{aligned} \frac{1}{2}\partial_t \|\langle \nabla \rangle^s f\|_2^2 &= \langle \langle \nabla \rangle^s f, \langle \nabla \rangle^s \left(P_{k\neq 0} \nabla^\perp \phi \cdot \nabla f \right) \rangle \approx \\ &\approx \langle \langle \nabla \rangle^s f, \langle \nabla \rangle^s P_{k\neq 0} \nabla^\perp \phi \cdot \nabla f \rangle + \langle \langle \nabla \rangle^s f, P_{k\neq 0} |\nabla| \nabla^\perp \phi \cdot \nabla \langle \nabla \rangle^{s-1} f \rangle + \cdots \end{aligned} \tag{18}$$

The second term we can pay regularity to get decay of the velocity field provided s is large enough, however, it is far from clear how to obtain any kind of decay on the first term. Indeed, getting decay from inviscid damping seems to always require more regularity than we have; and it seems to require so much that even if we are willing to work in analytic regularity via some kind of Cauchy-Kovalevskaya argument, it would still not be enough to close any estimates.

Let us now look closer. Since we must pay regularity to deduce decay on the velocity u, it is natural to consider the frequency interactions in the product $u \cdot \nabla f$ with the frequencies of u much larger than f. This leads us to study a simpler model

$$\partial_t f = -u \cdot \nabla f_{lo}, \tag{19}$$

where f_{lo} is a given function that we think of as much smoother than f. Let us just focus on what should be the worst:

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

This problem is linear on the Fourier side:

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

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

$$\partial_t \hat{f}(t,k,\eta) = \frac{1}{2\pi} \sum_{l \neq 0} \frac{\eta(k-l)}{l^2 + |\eta - lt|^2} \hat{f}(l,\eta) f_{lo}(t,k-l,0). \tag{20}$$

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

$$\partial_t f_R = \kappa \frac{k^2}{|\eta|} f_{NR}, \tag{21a}$$

$$\partial_t f_{NR} = \kappa \frac{|\eta|}{k^2 + |\eta - kt|^2} f_R, \tag{21b}$$

where we think of f_R as being the evolution of the k mode and f_{NR} being the evolution of a nearby mode l with $l \neq k$. The factor $k^2/\,|\eta|$ in the ODE for f_R is an upper bound on the strongest interaction a non-resonant mode, for example the $k-1$ mode, can have with the resonant mode. It is important to note that if at the beginning of the interval I_k, we have $f_R = f_{NR}$, then over the interval I_k, both f_R and f_{NR} are at most amplified by roughly the same factor $C\,(\frac{\eta}{k^2})^{1+2C\kappa}$ (though they crucially are not amplified by the same amount on the left and right parts of the interval). Taking the product of these amplifications for $k = E(\sqrt{\eta}), E(\sqrt{\eta}) - 1, ..., 1$ yields a total amplification which is $O(e^{C\sqrt{\eta}})$. This indicates that unless there is some special structure or cancellation not taken into account, the growth of high frequencies will cause a loss of Gevrey-2 regularity of the solution as $t \to \infty$. Therefore, in order to maintain control, the initial data must have at least this much regularity to lose, and this is the origin of the requirement $s > 1/2$ (or at least $s \geq 1/2$).

5 Proof of the stability result

Hence, we have two main challenges to overcome. The first is to choose a coordinate system that is properly adapted to the shear flow which is mixing the solution. Note that this shear flow is changing in time and cannot be determined directly from the initial data. We carry this out in Section 5.1 below. The next step is to get global-in-time, uniform regularity estimates on the resulting f. To do this we will design a special norm with which to measure the solution that accounts for the nonlinear Orr mechanism described above.

5.1 Coordinate transform. The original equations in vorticity form are (2), and we are trying essentially to prove that

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

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

$$z(t,x,y) = x - tv \tag{22a}$$

$$v(t,y) = y + \frac{1}{t}\int_0^t < U^x > (\tau, y)d\tau, \tag{22b}$$

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

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

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

where

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

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

$$\partial_t f + [\partial_t v]\partial_v f + \partial_t z \partial_z f = -y\partial_z f + v'\left[\partial_v \phi + \partial_z \phi \partial_v z - \partial_z \phi \partial_v z\right]\partial_z f - v'\partial_z \phi \partial_v f.$$

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

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

The Biot-Savart law also gets transformed into:

$$f = \partial_{zz}\phi + (v')^2\left(\partial_v - t\partial_z\right)^2 \phi + v''\left(\partial_v - t\partial_z\right)\phi = \Delta_t \phi. \tag{23}$$

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

$$\begin{cases} \partial_t f + u \cdot \nabla_{z,v} f = 0, \\ u = (0, [\partial_t v]) + v'\nabla^{\perp}_{z,v} P_{\neq 0}\phi, \\ \phi = \Delta_t^{-1}[f]. \end{cases} \tag{24}$$

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

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

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

$$\partial_t \tilde{u}_0 + [\partial_t v]\partial_v \tilde{u}_0 + v' < \nabla^{\perp} P_{\neq 0}\phi \cdot \nabla \tilde{u} >= 0. \tag{25}$$

Finally, one can express v' and $[\partial_t v]$ as solutions to a system of PDE in the (t, v) variables coupled to (24) (see Bedrossian and Masmoudi [2013] for details):

$$\partial_t(t(v'-1)) + [\partial_t v]t\partial_v v' = -f_0 \tag{26a}$$

$$\partial_t[\partial_t v] + \frac{2}{t}[\partial_t v] + [\partial_t v]\partial_v[\partial_t v] = -\frac{v'}{t} < \nabla^{\perp} P_{\neq 0}\phi \cdot \nabla \tilde{u} > \tag{26b}$$

$$v''(t,v) = v'(t,v)\partial_v v'(t,v). \tag{26c}$$

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

$$\partial_v \tilde{u}_0(t,v) = \frac{1}{v'(t,v)}\partial_y U_0^x(t,y) = -\frac{1}{v'(t,v)}\omega_0(t,y) = -\frac{1}{v'(t,v)} f_0(t,v). \tag{27}$$

Given a priori estimates on the system (24), (26), we can recover estimates on the original system (2) by the inverse function theorem as long as $v'-1$ remains sufficiently small (see Bedrossian and Masmoudi [ibid.] for details).

5.2 Construction of the toy model norm. For simplicity of notation in this section we usually take $\eta, k > 0$ but the work applies equally well to $\eta, k < 0$. Note that modes where $\eta k < 0$ do not have resonances for positive times. Keeping with the intuition from the derivation of (21), in this section we will think of η as a fixed parameter and time varying. Accordingly, in this section we will use $I_{k,\eta} = [\frac{\eta}{k} - \frac{\eta}{2k(k+1)}, \frac{\eta}{k} + \frac{\eta}{2k(k-1)}]$ to denote any resonant interval with $\eta/k^2 \geq 1$ (with the modification $[\eta - \frac{\eta}{4}, 2\eta]$ if $k = 1$). A key feature of the methods in Bedrossian and Masmoudi [ibid.] is how the toy model is used to construct a norm which precisely matches the estimated worst-case behavior that the reaction terms create, done by choosing $w_k(t,\eta)$ to be an approximate solution to (21). First we have the following (easy to check) Proposition.

Proposition 1. *Let $\tau = t - \frac{\eta}{k}$ and consider the solution $(f_R(\tau), f_{NR}(\tau))$ to (21) with $f_R\left(-\frac{\eta}{k^2}\right) = f_{NR}\left(-\frac{\eta}{k^2}\right) = 1$. There exists a constant C such that for all $\kappa < 1/2$ and $\frac{\eta}{k^2} \geq 1$,*

$$f_R(\tau) \leq C\left(\frac{k^2}{\eta}(1+|\tau|)\right)^{-C\kappa} \qquad -\frac{\eta}{k^2} \leq \tau \leq 0,$$

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

$$f_R(\tau) \leq C\left(\frac{\eta}{k^2}\right)^{C\kappa}(1+|\tau|)^{1+C\kappa} \qquad 0 \leq \tau \leq \frac{\eta}{k^2},$$

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

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

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

On each interval $I_{k,\eta}$, growth of the resonant mode (k, η) will be modeled by w_R and the rest of the modes (which are non-resonant) will be modeled by w_{NR}. By Proposition 1, we will able to choose w such that the total growth of w_R and w_{NR} exactly agree.

The construction is done backward in time, starting with $k = 1$. For $t \in I_{k,\eta}$ and $\tau = t - \frac{\eta}{k}$, we will choose (w_{NR}, w_R) such that over the interval $I_{k,\eta}$ they approximately satisfy (21):

$$\partial_\tau w_R \approx \kappa \frac{k^2}{\eta} w_{NR}, \qquad (28)$$
$$\partial_\tau w_{NR} \approx \kappa \frac{\eta}{k^2(1+\tau^2)} w_R,$$

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

Let w_{NR} be a non-decreasing function of time with $w_{NR}(t, \eta) = 1$ for $t \geq 2\eta$. For $k \geq 1$, we assume that $w_{NR}(t_{k-1,\eta})$ was computed. To compute w_{NR} on the interval $I_{k,\eta}$, we use the growth predicted by Proposition 1: for $k = 1, 2, 3, ..., E(\sqrt{\eta})$, we define

(29a)
$$w_{NR}(t,\eta) = \left(1 + a_{k,\eta}|t - \frac{\eta}{k}|\right)^{-1-C\kappa} w_{NR}\left(\frac{\eta}{k}\right), \qquad \forall t \in I^L_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right],$$
(29b)
$$w_{NR}(t,\eta) = \left(\frac{k^2}{\eta}\left[1 + b_{k,\eta}|t - \frac{\eta}{k}|\right]\right)^{C\kappa} w_{NR}(t_{k-1,\eta}), \qquad \forall t \in I^R_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right].$$

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

$$b_{k,\eta} = \frac{2(k-1)}{k}\left(1 - \frac{k^2}{\eta}\right) \qquad (30)$$

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

$$a_{k,\eta} = \frac{2(k+1)}{k}\left(1 - \frac{k^2}{\eta}\right). \tag{31}$$

Hence, $w_{NR}(\frac{\eta}{k}) = w_{NR}(t_{k-1,\eta})\left(\frac{k^2}{\eta}\right)^{C\kappa}$ and $w_{NR}(t_{k,\eta}) = w_{NR}(t_{k-1,\eta})\left(\frac{k^2}{\eta}\right)^{1+2C\kappa}$. The choice of $a_{k,\eta}$ and $b_{k,\eta}$ was made to ensure that the ratio between $w_{NR}(t_{k,\eta})$ and $w_{NR}(t_{k-1,\eta})$ is exactly $\left(\frac{k^2}{\eta}\right)^{1+2C\kappa}$. 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 $t \in [0, t_{E(\sqrt{\eta}),\eta}]$. Note that we always have $0 \le b_{k,\eta} < 1$ and $0 \le a_{k,\eta} < 4$, but that $a_{k,\eta}$ and $b_{k,\eta}$ approach 0 when k approaches $E(\sqrt{\eta})$. This will present minor technical difficulties in the sequel since this implies that $\partial_t w$ vanishes near this time and hence a loss of the lower bounds in (28).

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

$$w_R(t,\eta) = \frac{k^2}{\eta}\left(1 + a_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{NR}(t,\eta), \qquad \forall t \in I^L_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right], \tag{32a}$$

$$w_R(t,\eta) = \frac{k^2}{\eta}\left(1 + b_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{NR}(t,\eta), \qquad \forall t \in I^R_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right]. \tag{32b}$$

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

$$w_k(t,\eta) = \begin{cases} w_k(t_{E(\sqrt{\eta}),\eta},\eta) & t < t_{E(\sqrt{\eta}),\eta} \\ w_{NR}(t,\eta) & t \in [t_{E(\sqrt{\eta}),\eta}, 2\eta] \setminus I_{k,\eta} \\ w_R(t,\eta) & t \in I_{k,\eta} \\ 1 & t \ge 2\eta. \end{cases} \tag{33}$$

Since w_R and w_{NR} agree at the end-points of $I_{k,\eta}$, $w_k(t,\eta)$ is Lipschitz continuous in time. This completes the construction of w which appears in the J defined above (36).

The following lemma shows that the toy model predicts a growth of high frequencies which amounts to a loss of Gevrey-2 regularity, which is the primary origin of the restriction $s > 1/2$ in Theorem 1, though naturally the main application of the toy model is to design the norm.

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

$$\frac{1}{w_k(0,\eta)} = \frac{1}{w_k(t_{E(\sqrt{\eta}),\eta},\eta)} \sim \frac{1}{\eta^{\mu/8}} e^{\frac{\mu}{2}\sqrt{\eta}}. \tag{34}$$

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

Proof. Counting the growth over each interval implied by (33) gives the exact formula:

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

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

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

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

5.3 Main energy estimate. Our goal is to control solutions to (24) uniformly in a suitable norm as $t \to \infty$. The key idea we use for this is the carefully designed time-dependent norm written as

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

The multiplier A has several components,

$$A_k(t,\eta) = e^{\lambda(t)|k,\eta|^s} \langle k,\eta\rangle^\sigma J_k(t,\eta).$$

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

$$\lambda(t) = \frac{3}{4}\lambda_0 + \frac{1}{4}\lambda', \quad t \le 1 \tag{35a}$$

$$\dot{\lambda}(t) = -\frac{\delta_\lambda}{\langle t\rangle^{2\tilde{q}}}(1+\lambda(t)), \quad t > 1 \tag{35b}$$

where $\delta_\lambda \approx \lambda_0 - \lambda'$ is a small parameter that ensures $\lambda(t) > \lambda_0/2 + \lambda'/2$ and $1/2 < \tilde{q} \le s/8 + 7/16$ is a parameter chosen by the proof. The reason for (35a) is to account for the behavior of the solution on the time-interval $[0,1]$; see Bedrossian and Masmoudi [2013] for this minor detail. The main multiplier for dealing with the Orr mechanism and the associated nonlinear growth is

$$J_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)} + e^{\mu|k|^{1/2}}, \tag{36}$$

where $w_k(t,\eta)$ was constructed above and describes the expected 'worst-case' growth due to nonlinear interactions at the critical times. What will be important is that J imposes more regularity on modes which satisfy $t \sim \frac{\eta}{k}$ (the 'resonant modes') than those that do not (the 'non-resonant modes'). by controlled loss of regularity and is reminiscent of the notion of losing regularity estimates used in e.g. Bahouri and Chemin [1994] and Chemin and Masmoudi [2001]. One of the main differences is that here we have to be more precise in the sense that the loss of regularity occurs for different frequencies during different time intervals.

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

$$E(t) = \frac{1}{2}\|A(t)f(t)\|_2^2 + E_v(t), \tag{37}$$

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

$$E_v(t) = \langle t\rangle^{2+2s}\|\frac{A}{\langle\partial_v\rangle^s}v'\partial_v[\partial_t v](t)\|_2^2 + \\ + \langle t\rangle^{4-K_D\epsilon}\|[\partial_t v](t)\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 + \frac{1}{K_v}\|A^R(v'-1)(t)\|_2^2. \tag{38}$$

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

By the well-posedness theory for 2D Euler in Gevrey spaces Bardos and Benachour [1977], Ferrari and Titi [1998], Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009] we may safely ignore the time interval (say) $[0, 1]$ by further restricting the size of the initial data. See Bedrossian and Masmoudi [2013] for a slightly more detailed discussion. The goal is next to prove by a continuity argument that the energy $E(t)$ (together with some related quantities) is uniformly bounded for all time if ϵ is sufficiently small. We define the following controls referred to in the sequel as the bootstrap hypotheses for $t \geq 1$,

(B1) $E(t) \leq 4\epsilon^2$;

(B2) $\|v' - 1\|_\infty \leq \frac{3}{4}$

(B3) 'CK' integral estimates (for 'Cauchy-Kovalevskaya'):

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

The CK terms above that appear without the K_v^{-1} prefactor arise from the time derivatives of $A(t)$ and are naturally controlled by the energy estimates we are making. The others are related quantities that are controlled separately in Proposition 6 below. These both will be defined below when discussing the energy estimates.

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

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

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

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

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

The remainder of the section is devoted to the proof of Proposition 2, the primary step being to show that on $[1, T^\star]$, we have

$$E(t) + \frac{1}{2}\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + \\ + K_v^{-1}\left(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\right) + K_v^{-1}\sum_{i=1}^{2}\left(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\right)\Big] d\tau \le E(1) + K\epsilon^3 \tag{39}$$

for some constant K which is independent of ϵ and $T^\star$. If ϵ is sufficiently small then (39) implies Proposition 2. Indeed, the control $\|1-v'\| < 5/8$ is an immediate consequence of (B1) by Sobolev embedding for ϵ sufficiently small.

To prove (39), it is natural to compute the time evolution of $E(t)$,

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

The first contribution is of the form

$$\frac{1}{2}\frac{d}{dt}\int |Af|^2\,dx = -CK_\lambda - CK_w - \int AfA(u\cdot\nabla f)dx, \tag{40}$$

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

$$CK_\lambda = -\dot{\lambda}(t)\|\,|\nabla|^{s/2}Af\|_2^2 \tag{41a}$$

$$CK_w = \sum_k \int \frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)} e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)} A_k(t,\eta)\left|\hat{f}_k(t,\eta)\right|^2 d\eta. \tag{41b}$$

In what follows we define

$$\tilde{J}_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)}, \tag{42a}$$

$$\tilde{A}_k(t,\eta) = e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \tilde{J}_k(t,\eta). \tag{42b}$$

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

Strictly speaking, equality (40) is not quite rigorous since it involves a derivative of Af, which is not a priori well-defined. To make this calculation rigorous, we have first to approximate the initial data of (2) by (for instance) analytic initial data and use that the global solutions of (2) stay analytic for all time (see Bardos and Benachour [1977], Foias and Temam [1989], and Ferrari and Titi [1998]). Hence, we can perform all calculations on these solutions with regularized initial data and then perform a passage to the limit to infer that (39) still holds.

To treat the main term in (40), begin by integrating by parts, as in the techniques Foias and Temam [1989], Levermore and Oliver [1997], Kukavica and Vicol [2009], and Gerard-Varet and Masmoudi [2015]

(43)

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

Notice that the relative velocity is not divergence free:

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

The first term is controlled by the bootstrap hypothesis (B1). For the second term we pay regularity and show that under the bootstrap hypotheses we have

$$\| P_{\neq 0} \phi(t) \|_{\mathcal{G}^{\lambda(t),\sigma-3}} \lesssim \frac{\epsilon}{\langle t \rangle^2}. \tag{44}$$

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

$$\left| \int \nabla \cdot u \, |Af|^2 \, dx \right| \leq \|\nabla u\|_\infty \|Af\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D \epsilon/2}} \|Af\|_2^2 \lesssim \frac{\epsilon^3}{\langle t \rangle^{2-K_D \epsilon/2}}. \tag{45}$$

To handle the commutator, $\int Af \, [A(u \cdot \nabla f) - u \cdot \nabla Af] \, dx$, we use a paraproduct decomposition (see e.g. Bony [1981] and Bahouri, Chemin, and Danchin [2011]). Precisely, we define three main contributions: *transport*, *reaction* and *remainder*:

$$\int Af \, [A(u \cdot \nabla f) - u \cdot \nabla Af] \, dx = \frac{1}{2\pi} \sum_{N \geq 8} T_N + \frac{1}{2\pi} \sum_{N \geq 8} R_N + \frac{1}{2\pi} \mathfrak{R}, \tag{46}$$

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

$$\begin{aligned} T_N &= 2\pi \int Af \left[A(u_{<N/8} \cdot \nabla f_N) - u_{<N/8} \cdot \nabla A f_N \right] dx \\ R_N &= 2\pi \int Af \left[A(u_N \cdot \nabla f_{<N/8}) - u_N \cdot \nabla A f_{<N/8} \right] dx \\ \mathfrak{R} &= 2\pi \sum_{N \in \mathbb{D}} \sum_{\frac{1}{8} N \leq N' \leq 8N} \int Af \, [A(u_N \cdot \nabla f_{N'}) - u_N \cdot \nabla A f_{N'}] \, dx. \end{aligned}$$

Here $N \in \mathbb{D} = \{\frac{1}{2}, 1, 2, 4, \ldots, 2^j, \ldots\}$ and g_N denotes the N-th Littlewood-Paley projection and $g_{<N}$ means the Littlewood-Paley projection onto frequencies less than N. Formally, the paraproduct decomposition (46) represents a kind of 'linearization' for the evolution of higher frequencies around the lower frequencies. The terminology 'reaction' is borrowed from Mouhot and Villani [2011].

To control the transport term, we use

Proposition 3 (Transport)**.** *Under the bootstrap hypotheses,*

$$\sum_{N \geq 8} |T_N| \lesssim \epsilon C K_\lambda + \epsilon C K_w + \frac{\epsilon^3}{\langle t \rangle^{2-K_D \epsilon/2}}.$$

The proof of Proposition 3 uses ideas from the works of Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009]. Since the velocity u is restricted to 'low frequency', we will have the available regularity required to apply (44). However, the methods of Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009] do not adapt immediately since $J_k(t,\eta)$ is imposing slightly different regularities to certain frequencies, which is problematic. Physically speaking, we need to ensure that resonant frequencies do not incur a very large growth due to nonlinear interactions with non-resonant frequencies (which are permitted to be slightly larger than the resonant frequencies). Controlling this imbalance is why CK_w appears in Proposition 3.

Controlling the reaction contribution in (46) is one of the main tasks. Here we cannot apply (44), as an estimate on this term requires u in the highest norm on which we have control, and hence we have no regularity to spare. Physically, here in the reaction term is where the dangerous nonlinear effects are expressed and a great deal of precision is required to control them.

Proposition 4 (Reaction). *Under the bootstrap hypotheses,*

$$\sum_{N\geq 8}|R_N| \lesssim \epsilon CK_\lambda + \epsilon CK_w + \frac{\epsilon^3}{\langle t\rangle^{2-K_D\epsilon/2}} + \epsilon CK_\lambda^{v,1} + \epsilon CK_w^{v,1}$$
$$+\epsilon\|\langle\frac{\partial_v}{t\partial_z}\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\|_2^2. \tag{47}$$

The $CK^{v,1}$ terms are defined below in (51). The first step to controlling the term in (47) involving ϕ is Proposition 5. This proposition treats Δ_t as a perturbation of $\partial_{zz}+(\partial_v - t\partial_z)^2$ and passes the multipliers in the last term of (47) onto f *and* the coefficients of Δ_t. Physically, these latter contributions are indicating the nonlinear interactions between the higher modes of f and the coefficients v', v'' (which involve time-averages of f_0 (26)). Analogous lemmas have continued to play important roles in the theory.

Proposition 5 (Precision elliptic control). *Under the bootstrap hypotheses,*

$$\|\langle\frac{\partial_v}{t\partial_z}\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\|_2^2$$
$$\lesssim CK_\lambda + CK_w + \epsilon^2\sum_{i=1}^{2} CCK_\lambda^i + CCK_w^i, \tag{48}$$

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

(49a) $$CCK_\lambda^1 = -\dot{\lambda}(t) \| \, |\partial_v|^{s/2} A^R \left(1-(v')^2\right) \|_2^2,$$

(49b) $$CCK_w^1 = \| \sqrt{\frac{\partial_t w}{w}} A^R \left(1-(v')^2\right) \|_2^2,$$

(49c) $$CCK_\lambda^2 = -\dot{\lambda}(t) \| \, |\partial_v|^{s/2} \frac{A^R}{\langle \partial_v \rangle} v'' \|_2^2,$$

(49d) $$CCK_w^2 = \| \sqrt{\frac{\partial_t w}{w}} \frac{A^R}{\langle \partial_v \rangle} v'' \|_2^2.$$

The next step in the bootstrap is to provide good estimates on the coordinate system and the associated CK and CCK terms. The following proposition provides controls on $v'-1$, the CCK terms arising in (49), the pair $[\partial_t v]$, $v'\partial_v[\partial_t v]$ and finally all of the $CK^{v,i}$ terms. The norm defined by $A^R(t)$ is stronger than that defined by $A(t)$, which we use to measure f. It turns out that we will be able to propagate this stronger regularity on $v'-1$ due to a time-averaging effect, derived via energy estimates on (26). By contrast, $[\partial_t v]$ is expected basically to have the regularity of $\tilde{u}_0$ and hence even (50b) has s fewer derivatives than expected. On the other hand, it has a significant amount of time decay, which near critical times can be converted into regularity.

Proposition 6 (Coordinate system controls). *Under the bootstrap hypotheses, for ϵ sufficiently small and K_v sufficiently large there is a $K>0$ such that*

(50a)
$$\| A^R(v'-1)(t)\|_2^2 + \frac{1}{2}\int_1^t \sum_{i=1}^2 CCK_w^i(\tau)d\tau + \frac{1}{2}\int_1^t \sum_{i=1}^2 CCK_\lambda^i(\tau)d\tau \le \frac{1}{2}K_v\epsilon^2$$

(50b) $$\langle t\rangle^{2+2s} \| \frac{A}{\langle \partial_v \rangle^s} v'\partial_v[\partial_t v]\|_2^2 + \frac{1}{2}\int_1^t CK_\lambda^{v,2}(\tau) + CK_w^{v,2}(\tau)d\tau \le \epsilon^2 + K\epsilon^3$$

(50c) $$\langle t\rangle^{4-K_D\epsilon} \| [\partial_t v]\|^2_{\mathcal{G}^{\lambda(t),\sigma-6}} \le \epsilon^2 + K\epsilon^3$$

(50d) $$\int_1^t CK_\lambda^{v,1}(\tau) + CK_w^{v,1}(\tau)d\tau \le \frac{1}{2}K_v\epsilon^2,$$

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

$$CK_w^{v,2}(\tau) = \langle\tau\rangle^{2+2s} \| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle\partial_v\rangle^s} v' \partial_v [\partial_t v](\tau)\|_2^2 \tag{51a}$$

$$CK_\lambda^{v,2}(\tau) = \langle\tau\rangle^{2+2s} (-\dot{\lambda}(\tau)) \| \, |\partial_v|^{s/2} \frac{A}{\langle\partial_v\rangle^s} v' \partial_v [\partial_t v](\tau)\|_2^2 \tag{51b}$$

$$CK_w^{v,1}(\tau) = \langle\tau\rangle^{2+2s} \| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle\partial_v\rangle^s} [\partial_t v](\tau)\|_2^2 \tag{51c}$$

$$CK_\lambda^{v,1}(\tau) = \langle\tau\rangle^{2+2s} (-\dot{\lambda}(\tau)) \| \, |\partial_v|^{s/2} \frac{A}{\langle\partial_v\rangle^s} [\partial_t v](\tau)\|_2^2. \tag{51d}$$

Note that neither (50b) nor (50c) controls the other: at higher frequencies the former is stronger than the latter and at lower frequencies the opposite is true. One of the advantages of this scheme is that $v'\partial_v[\partial_t v]$ satisfies an equation that is simpler than $[\partial_t v]$ and so is easier to get good estimates on. Both (50b) and (50c) are linked to the convergence of the background shear flow; in particular, they rule out that the background flow oscillates or wanders due to nonlinear effects.

Finally we need to control the remainder term in (46). This is straightforward and is detailed in Bedrossian and Masmoudi [2013].

Proposition 7 (Remainders)**.** *Under the bootstrap hypotheses,*

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

Collecting Propositions 3, 4, 5, 6, 7 with (46) and (45), we have finally (39) for ϵ sufficiently small with constants independent of both ϵ and $T^\star$; hence for ϵ sufficiently small we may propagate the bootstrap control and prove Proposition 2.

6 Proof of the instability result

6.1 Ideas of the proof. The proof starts by performing the same coordinate change $(x, y) \mapsto (z, v)$, as defined in Equation (22a) and Equation (22b). Following the calculations of Section 5.1, this then gets rid of the badly behaving zeroth mode, and reduces Equation (2) to a system for f, $v' - 1$, and $[\partial_t v]$. These are recorded in Equation (52) $\sim$ Equation (53) below, where for simplicity we denote $h := v' - 1$, $\theta := [\partial_t v]$ and $g := (f, h, \theta)$:

The goal is to find a solution (f, h, θ) to the system

$$\begin{cases} \partial_t f = -\theta \cdot \partial_v f - (h+1)\nabla^{\perp}\phi \cdot \nabla f, \\ \partial_t h = -\theta \partial_v h - \dfrac{\mathbb{P}_0 f + h}{t}, \\ \partial_t \theta = -\dfrac{2\theta}{t} - \theta \partial_v \theta + \dfrac{1}{t}\mathbb{P}_0(f \cdot \partial_z \phi), \end{cases} \tag{52}$$

where the relevant quantities are defined as

$$\begin{cases} \phi = \mathbb{P}_{\neq 0}\Delta_t^{-1} f, \\ \Delta_t = \partial_z^2 + (h+1)^2(\partial_v - t\partial_z)^2 + (h+1)(\partial_v h)(\partial_v - t\partial_z). \end{cases} \tag{53}$$

Moreover, for $t \in I$ we have

$$t(h+1)\partial_v \theta(t) + \mathbb{P}_0 f(t) + h(t) = 0; \quad \int_{\mathbb{R}} \frac{h(t,v)}{h(t,v)+1}\, \mathrm{d}v = 0. \tag{54}$$

6.1.1 The choice of data, and setup. We will construct (f, h, θ) that satisfies the required instability assumptions. This solution will be constructed as the superposition of a background solution $(\underline{f}, \underline{h}, \underline{\theta})$, and a perturbation (f^*, h^*, θ^*) (which is a second order perturbation of a much smaller size but lower regularity). It turns out that (h, θ) plays a relatively less important role in the proof, so for simplicity, here we will consider f only.

The background solution $\underline{f}$ is guaranteed to exist by Theorem 1; we will assume it has analytic regularity (i.e. $s = 1$ in Theorem 1), and has size $\varepsilon_0 \ll \varepsilon$. More precisley, we define the background solution $\underline{g} := (\underline{f}, \underline{h}, \underline{\theta})$, which solves (52)∼(53) with initial data

$$\underline{f}(1, z, v) = \varepsilon_0 \cos z \cdot \varphi_b(v), \quad \underline{h}(1, z, v) = \underline{\theta}(1, z, v) = 0, \tag{55}$$

where

$$\varphi_b(v) = e^{-(C_0^{-1} v)^{18}}.$$

By Theorem 1, we know that $\underline{g}$ exists on $[1, +\infty)$, and satisfies the following properties, where recall that all constants here depend on C_0:

1. $\underline{f}$ and $\underline{h}$ are real-valued and even, $\underline{\theta}$ is real-valued and odd, and

$$\|\underline{f}(t)\|_{\mathfrak{G}_{C_0}} + \|\underline{h}(t)\|_{\mathfrak{G}_{C_0}} + \|\underline{\theta}(t)\|_{\mathfrak{G}_{C_0}} \lesssim \varepsilon_0, \quad \|\underline{\theta}(t)\|_{\mathfrak{G}_{C_0 - 1}} \lesssim \frac{\varepsilon_0}{t^2}; \tag{56}$$

2. $\underline{f}$ and $\underline{h}$ converge as $t \to \infty$,

$$\|\underline{f}(t) - f_\infty\|_{\mathfrak{G}_{C_0 - 1}} \lesssim \frac{\varepsilon_0^2}{t}, \quad \|\underline{h}(t) + \mathbb{P}_0 f_\infty\|_{\mathfrak{G}_{C_0 - 1}} \lesssim \frac{\varepsilon_0}{t}; \tag{57}$$

3. The limit f_∞ is close to the specific profile we choose, namely

$$\left\| f_\infty - \varepsilon_0 \cos z \cdot \varphi_b(v) \right\|_{\mathcal{G}_{C_0-1}} \lesssim \varepsilon_0^2. \tag{58}$$

We also define the function $\underline{\phi}$ and the operator $\underline{\Delta_t}$, corresponding to $(\underline{f}, \underline{h}, \underline{\theta})$, as in (53). In practice we will think of $\underline{f}$ as only having low frequency components, as it is much more regular than the perturbation f^* we will construct.

The perturbation f^* will be fixed by assigning the data at some time $t = T_0$:

$$f^*(T_0) = \varepsilon_1 \cos(k_0 z + \eta_0 v)\varphi_p(k_0\sqrt{\sigma} v),$$

where φ_p is a suitable Schwartz function. The parameters $(\varepsilon_0, k_0, T_0, \eta_0, \sigma)$ are related by (where $N_2 = 30$, $N_3 = 30000$):

$$\begin{aligned} &\sigma = (\log k_0)^{-N_2}, \quad && \alpha = 1 + (\log k_0)^{-2N_2}, \quad \varepsilon_0 = (\log k_0)^{-N_3}; \\ &\eta_0 = \frac{2k_0^2\alpha}{\pi\varepsilon_0}, && T_0 = \frac{2\eta_0}{2k_0+1}. \end{aligned} \tag{59}$$

Note that $\alpha - 1 = \sigma^2$ and $\varepsilon_0 = \sigma^{1000}$. For now it suffices to note that $\varepsilon_1 \ll \varepsilon_0$, and $\widehat{f^*(T_0)}$ is concentrated near only two frequencies, (k_0, η_0) and $(-k_0, -\eta_0)$, where (k_0, η_0) is considered the high frequency mode compared to $\underline{g}$. We also fix two times $T_1 \in [T_0, 2T_0]$ and $T_2 \le T_0$, define by

$$\begin{aligned} &t_m = \frac{2\eta_0}{2m+1}, \quad && T_j = t_{k_j}, \ 1 \le j \le 2; \\ &k_1 = (1-\sigma)k_0, && k_2 = \varepsilon_0^{-1/40}\sqrt{\eta_0} \end{aligned} \tag{60}$$

Note that $k_2 > k_0 > k_1$ and $T_1 > T_0 > T_2$. For simplicity we will assume all the k_i's are integers (otherwise take their integer parts).

6.1.2 The linearized system. Since $\varepsilon_1 \ll \varepsilon_0$, it is natural to first study the linearization of (52)∼(53) at the background solution $\underline{f}$. This linear system has the form $\partial_t f' = \mathfrak{L} f'$, where $\mathfrak{L}$ is a linear operator and $f'(T_0) = f^*(T_0)$. Following the observation made in Bedrossian and Masmoudi [2013], we know $\mathfrak{L}$ consists of two parts: the first one is a "transport" part,

$$\mathfrak{L}^T f' = \underline{\Phi} \cdot \nabla f', \tag{61}$$

where $\underline{\Phi}$ is a combination of the background solution, which has much higher regularity than f' (and thus contributes low frequencies only), and moreover decays like t^{-2}.

The second one is a "reaction" term, which is responsible for the Orr growth mechanism,

$$\mathfrak{L}^R f' = \underline{F} \cdot \nabla \underline{\Delta_t}^{-1} f', \tag{62}$$

where $\underline{F}$ again comes from the background solution, but has no decay in time (one can think $\underline{\Phi} \sim t^{-2}\underline{F}$); the operator $\underline{\Delta_t}^{-1}$ is defined, up to some error terms, by

$$\widehat{\underline{\Delta_t}^{-1} F}(t,k,\xi) = \frac{1}{(\xi - kt)^2 + k^2} \widehat{F}(t,k,\xi);$$

Notice that, if one compares (61) and (62), say at a critical time $t = \xi/k$, and assume that $\underline{\Phi} \sim t^{-2}\underline{F}$, then $\mathfrak{L}^R$ dominates $\mathfrak{L}^T$ if $t \gg k$ or equivalently $t \gtrsim \sqrt{|\xi|}$, and $\mathfrak{L}^T$ dominates $\mathfrak{L}^R$ if $t \lesssim \sqrt{|\xi|}$.

Our strategy is to show that the size of $\widehat{f}'$, say near $(\pm k_0, \pm\eta_0)$, exhibits growth at critical times *between* T_0 *and* T_1 by the Orr mechanism, and in fact saturates the upper bound proved in Bedrossian and Masmoudi [2013]. Note that this also explains the seemingly strange choice of assigning data at $t = T_0$ instead of $t = 1$, since we only know how to saturate the optimal growth on $[T_0, T_1]$.

Moreover, we need to go *backwards* from T_0 and recover the control for f' at time $t = 1$. There are two regimes here: when t is small (namely $t \le T_2$; note that T_2 is almost $\sqrt{\eta_0}$, see(60)), the transport term dominates, and the growth of f' can be easily controlled by an energy-type inequality for transport equations. When t is large, namely $t \in [T_2, T_0]$, the reaction term dominates and the situation will be much similar to what happens on $[T_0, T_1]$, except that only an upper bound is needed.

Summing up, we need to obtain a lower bound for f' on $[T_0, T_1]$, and an upper bound for f' on $[T_2, T_0]$, of form

$$|\widehat{f}'(T_1, \pm k_0, \pm\eta_0)| \gtrsim e^{c\sqrt{\eta_0}}\varepsilon_1; \quad |\widehat{f}'(T_2, \pm k_0, \pm\eta_0)| \lesssim e^{c'\sqrt{\eta_0}}\varepsilon_1, \tag{63}$$

for some suitable $c > c' > 0$. This would then imply that $f'(T_1)$ is large in H^N, and that $f'(T_2)$ (and hence $f'(1)$) is small in $\mathcal{G}^*$, upon choosing ε_1 appropriately. In both cases it is crucial to obtain precise bounds on the size of $\widehat{f}'$ near frequency $(\pm k_0, \pm\eta_0)$, which is the next step of the proof.

6.1.3 Linear analysis, and a more precise toy model. We may now restrict the linearized system to time $t \in [T_2, T_1]$, where the transport term plays essentially no role, so we will focus on the reaction term only. Recall the expression in (62); for simplicity we assume that $\underline{F}$ is independent of time and has only $k = \pm 1$ modes, say $\widehat{\underline{F}}(t,k,\xi) = \varepsilon_0 \mathbf{1}_{k=\pm 1}\varphi(\xi)/2$ with a Schwartz function φ.

By (62), we then write down the equation

$$\begin{aligned}(64)\quad \partial_t \widehat{f}'(t,k,\xi) = &\int_{\mathbb{R}} \frac{\varepsilon_0\eta/2}{(\eta - t(k+1))^2 + (k+1)^2}\widehat{\varphi}(\xi-\eta)\widehat{f}'(t,k+1,\eta)\,\mathrm{d}\eta \\ &- \int_{\mathbb{R}} \frac{\varepsilon_0\eta/2}{(\eta - t(k-1))^2 + (k-1)^2}\widehat{\varphi}(\xi-\eta)\widehat{f}'(t,k-1,\eta)\,\mathrm{d}\eta\end{aligned}$$

for $\widehat{f}'$. In Bedrossian and Masmoudi [ibid.], the authors replaced the function φ on the right hand side of (64) by the δ function, obtaining and ODE *toy model* which is essentially an "envelope" of (64) and can be solved explicitly. This is perfect for obtaining an *upper bound* for solutions to (64), but in order to get a *lower bound* a more accurate approximation will be needed - which is precisely what we are able to obtain here, under the assumption $\eta \approx \eta_0$ and $t \in [T_2, T_1]$.

For simplicity, let us assume $t \sim T_0$; recall from (59) that $T_0 \sim \sqrt{\eta_0/\varepsilon_0}$. Since $\eta \approx \eta_0$ due to the definition of $f'(T_0)$, we know that (64) plays a significant role only near the critical times η_0/m, where $m \sim \sqrt{\varepsilon_0\eta_0}$. We thus cut the time interval into subintervals, each containing exactly one critical time. Define, see also (60),

$$t_m = \frac{2\eta_0}{2m+1}; \qquad \frac{\eta_0}{m} \in [t_m, t_{m-1}] := I_m;$$

then on each I_m, according to (64), only the modes $k = m \pm 1$ will be active (i.e. has significant increments), since

$$\frac{1}{(\eta_0 - kt)^2 + k^2} \lesssim \frac{1}{t^2} \ll \frac{1}{m^2}, \qquad \forall t \in I_m, k \neq m.$$

We can therefore solve (64) approximately and explicitly[1], obtaining an approximate recurrence relation (see (59) for definition of parameters):

$$(65)\quad \mathcal{F} f'(t_{m-1},k,v) = \mathcal{F} f'(t_m,k,v) + \mathfrak{R} + \begin{cases} 0, & k \neq m \pm 1; \\ \mp \dfrac{\alpha k_0^2}{m^2}\varphi(v)\cdot \mathcal{F} f'(t_m,m,v), & k = m \pm 1, \end{cases}$$

after taking inverse Fourier transform in ξ, where the error term $\mathfrak{R}$ is small in L^2.

The recurrence relation (65) then plays the role of the toy model in Bedrossian and Masmoudi [ibid.]. In fact, if we choose φ such that $\|\varphi\|_{L^\infty} = 1$, then this already suffices

[1] Note that this argument works precisely when $t \in [T_2, T_1]$: when t si too small transport terms will come in, and when t is too large the $(m \pm 1)$ modes $\widehat{f}'(t, m \pm 1, \xi)$ will grow too much and destroy the approximate decoupling.

to prove the upper bound on $[T_2, T_0]$, since (65) essentially implies that

$$\sup_k \|\mathcal{F} f'(t_m, k, \cdot)\|_{L^2} \lesssim \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \cdot \sup_k \|\mathcal{F} f'(t_{m-1}, k, \cdot)\|_{L^2},$$

and thus by iteration,

$$\sup_k \|\mathcal{F} f'(T_2, k, \cdot)\|_{L^2} \lesssim \varepsilon_1 \prod_{m=k_2}^{k_0} \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \sim e^{c'\sqrt{\eta_0}}\varepsilon_1. \tag{66}$$

We turn to the lower bound for f' on $[T_0, T_1]$. If φ were identically 1, then in view of the *smallness* of $\mathfrak{R}$, we can use the same argument to obtain that

$$\sup_k \|\mathcal{F} f'(t_{m-1}, k, \cdot)\|_{L^2} \gtrsim \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \cdot \sup_k \|\mathcal{F} f'(t_m, k, \cdot)\|_{L^2},$$

and hence

$$\sup_k \|\mathcal{F} f'(T_1, k, \cdot)\|_{L^2} \gtrsim \varepsilon_1 \prod_{m=k_0}^{k_1} \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \sim e^{c\sqrt{\eta_1}}\varepsilon_1. \tag{67}$$

Comparing (66) and (67) we obtain the desired inequality (63) by direct computations, due to our choice of parameters.

Nevertheless φ cannot be identically 1, and moreover the error term $\mathfrak{R}$ is not local. To recover (67), in view of the factor $\varphi(v)$ on the right hand side of (65), we thus need to localize v in the region where $\varphi(v)$ is equal or close to 1. This localization is achieved by going back to physical space and performing an energy-type estimate for an L^2 norm with exponential weight in physical space.

6.1.4 Nonlinear analysis, and the Taylor expansion. Up to now we have only considered f', which is the solution to the linearized system $\partial_t f' = \mathfrak{L} f'$. The full nonlinear system (52)~(53), in terms of f^*, can be written as

$$\partial_f f^* = \mathfrak{L} f^* + \mathfrak{N}(f^*, f^*), \tag{68}$$

if, say, we consider only quadratic nonlinearities. Note that f' can also be regarded as the first order term in a formal Taylor expansion of f^*; we may write out the higher order terms by $f^{(1)} = f'$ and

$$\partial_t f^{(n)} = \mathfrak{L} f^{(n)} + \sum_{q_1+q_2=n-1} \mathfrak{N}(f^{(q_1)}, f^{(q_2)}), \quad f^{(n)}(T_0) = 0;$$

Our next step is to prove that, in some sense, we have[2]

$$\text{(the size of } f^{(n)}) \lesssim (\text{the size of } f^{(1)})^n, \tag{69}$$

Since the size of $f^{(1)}$ is $O(\varepsilon_1)$, the bound (69) guarantees that the contribution of $f^{(n)}$ with $n \geq 2$ will be negligible, and thus Theorem 2 follows from the estimates for $f^{(1)}$ obtained above.

The proof of (69) follows from an inductive argument, where at each step we combine the multilinear estimates for the nonlinear term $\mathfrak{N}$ with the linear estimates for the inhomogeneous equation

$$\partial_t f = \mathfrak{L} f + \mathfrak{N},$$

which is proved in the same way as the linear homogeneous case. Here the main difficulty is that $f^{(n)}$, being essentially the n-th power of $f^{(1)}$, is supported in Fourier space at (say) the frequency $(nk_0, n\eta_0)$. We thus need to run the arguments above for this particular choice of frequency, instead of (k_0, η_0). Fortunately this just corresponds to changing of parameters in the Orr growth mechanism, and most of the arguments above still go through.

Finally, to avoid the divergence issue caused by doing the Taylor expansion directly, we will close the whole proof by fixing some very large n_0 and claiming that

$$f^{(1)} + f^{(2)} + \cdots + f^{(n_0)}$$

is an approximate solution to (68), with error term so small that an actual solution to (68) can be constructed by a perturbative argument on the interval $[1, T_1]$.

6.2 Further discussions.

We mention two possible further questions related to Theorem 2.

6.2.1 Asymptotic instability.

Given Theorem 2, an immediate question is whether asymtotic instability can also be proved for (2). We believe this can be done by repeatedly applying the arguments in this paper.

Roughly speaking, we fix the background solution $\underline{f}$ and construct the perturbation $f^* = f_1^*$ as in Theorem 2. Note that f_1^* and grows from some time T_0^1 to some later time T_1^1; We now take $\underline{f} + f_1^*$ as the new background and construct a further perturbation f_2^* which grows from time T_0^2 to T_1^2, and so on. We then pile up a sequence of perturbations and define

$$f := \underline{f} + f_1^* + f_2^* + \cdots,$$

[2] Note however that $f^{(n)}$ is supported at higher and higher frequencies, namely $(nk_0, n\eta_0)$; thus this fact cannot be captured by a bootstrap argument in a single Gevrey norm, and this formal Taylor expansion seems necessary.

which we expect to satisfy that

$$\|f(1)\|_{\mathcal{G}^*} \leq \varepsilon, \quad \lim_{t\to\infty} \|\langle\partial_x\rangle^{N_0} f(t)\|_{L^2} = +\infty.$$

The main difficulty here is to control the evolution of f_1^* after time T_1^1; we then have to extend our arguments, which now covers only critical times η_0/m with $m \gtrsim k_0$, to *all* critical times up to $m = 1$. We believe that a suitable combination of the techniques used in this paper and the weighted energy method used in Bedrossian and Masmoudi [2013] should be the key to solving this problem.

6.2.2 Genericity. Another natural question is whether the Orr growth mechanism is generic, i.e., whether the full upper bound of growth can be saturated for "most" solutions in a suitable sense. To study this problem, we have to consider solutions with general distribution in frequencies, instead of the f^* we choose here, which essentially has only two modes. In such cases we no longer have the simple decoupling as in Section 6.1.3, nor the recurrence relation (65); the main challenge is thus to find a substitute to (65) and to approximate (64), and it would be crucial to be able to separate the different components of the solution that evolve differently. It seems that some further physical-space based techniques will be needed.

Another challenge is the possible cancellations for the toy model (if we can find one) in the generic case. This also depends on how well are different frequencies and different physical space locations separated - if they are mixed together then we would have less control of the solution.

References

Jeffrey S. Baggett, Tobin A. Driscoll, and Lloyd N. Trefethen (1995). "A mostly linear model of transition to turbulence". *Phys. Fluids* 7.4, pp. 833–838. MR: 1324952 (cit. on p. 2162).

H. Bahouri and J.-Y. Chemin (1994). "Équations de transport relatives á des champs de vecteurs non-lipschitziens et mécanique des fluides". *Arch. Rational Mech. Anal.* 127.2, pp. 159–181. MR: 1288809 (cit. on p. 2171).

Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin (2011). *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, pp. xvi+523. MR: 2768550 (cit. on p. 2174).

C. Bardos and S. Benachour (1977). "Domaine d'analycité des solutions de l'équation d'Euler dans un ouvert de R^n". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 4.4, pp. 647–687. MR: 0454413 (cit. on pp. 2171, 2173).

Jacob Bedrossian (May 2016). "Nonlinear echoes and Landau damping with insufficient regularity". arXiv: 1605.06841 (cit. on pp. 2160–2162).

Jacob Bedrossian, Pierre Germain, and Nader Masmoudi (June 2015a). "Dynamics near the subcritical transition of the 3D Couette flow I: Below threshold case". To appear in Mem. Amer. Math. Soc. arXiv: 1506.03720 (cit. on pp. 2157, 2161).

– (June 2015b). "Dynamics near the subcritical transition of the 3D Couette flow II: Above threshold case". arXiv: 1506.03721 (cit. on pp. 2157, 2161).

– (2017a). "On the stability threshold for the 3D Couette flow in Sobolev regularity". *Ann. of Math. (2)* 185.2, pp. 541–608. MR: 3612004 (cit. on pp. 2157, 2161).

– (Dec. 2017b). "Stability of the Couette flow at high Reynolds number in 2D and 3D". arXiv: 1712.02855 (cit. on p. 2157).

Jacob Bedrossian and Nader Masmoudi (2013). "Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations". *Publ. math. de l'IHÉS* 1306, pp. 1–106 (cit. on pp. 2157, 2160, 2167, 2170, 2171, 2177, 2179–2181, 2184).

Jacob Bedrossian, Nader Masmoudi, and Clément Mouhot (2016a). "Landau damping: paraproducts and Gevrey regularity". *Ann. PDE* 2.1, Art. 4, 71. MR: 3489904 (cit. on p. 2162).

– (2016b). "Landau damping: paraproducts and Gevrey regularity". *Ann. PDE* 2.1, Art. 4, 71. MR: 3489904 (cit. on pp. 2160, 2162).

Jacob Bedrossian, Nader Masmoudi, and Vlad Vicol (2016). "Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the two dimensional Couette flow". *Arch. Ration. Mech. Anal.* 219.3, pp. 1087–1159. MR: 3448924 (cit. on pp. 2157, 2161).

Jean-Michel Bony (1981). "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires". *Ann. Sci. École Norm. Sup. (4)* 14.2, pp. 209–246. MR: 631751 (cit. on p. 2174).

John P Boyd (1983). "The continuous spectrum of linear Couette flow with the beta effect". *Journal of the atmospheric sciences* 40.9, pp. 2304–2308 (cit. on p. 2159).

K. M. Case (1960). "Stability of inviscid plane Couette flow". *Phys. Fluids* 3, pp. 143–148. MR: 0128230 (cit. on p. 2156).

Jean-Yves Chemin and Nader Masmoudi (2001). "About lifespan of regular solutions of equations related to viscoelastic fluids". *SIAM J. Math. Anal.* 33.1, pp. 84–112. MR: 1857990 (cit. on p. 2171).

Yu Deng and Nader Masmoudi (2018). *Long time instability of the couette flow in low gevrey spaces*. Preprint (cit. on pp. 2157, 2160, 2161).

L. A. Dikiĭ (1961). "The stability of plane-parallel flows of an ideal fluid". *Soviet Physics Dokl.* 5, pp. 1179–1182. MR: 0147072 (cit. on p. 2156).

P. G. Drazin and W. H. Reid (1982). *Hydrodynamic stability*. Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, Cambridge-New York, pp. xiv+527. MR: 684214 (cit. on p. 2156).

Andrew B. Ferrari and Edriss S. Titi (1998). "Gevrey regularity for nonlinear analytic parabolic equations". *Comm. Partial Differential Equations* 23.1-2, pp. 1–16. MR: 1608488 (cit. on pp. 2171, 2173).

C. Foias and R. Temam (1989). "Gevrey class regularity for the solutions of the Navier-Stokes equations". *J. Funct. Anal.* 87.2, pp. 359–369. MR: 1026858 (cit. on pp. 2171, 2173, 2175).

David Gerard-Varet and Nader Masmoudi (2015). "Well-posedness for the Prandtl system without analyticity or monotonicity". *Ann. Sci. Éc. Norm. Supér.* 48.6, pp. 1273–1325 (cit. on p. 2173).

Lord Kelvin (1887). "Stability of fluid motion: rectilinear motion of viscous fluid between two parallel plates". *Phil. Mag* 24.5, pp. 188–196 (cit. on p. 2156).

Igor Kukavica and Vlad Vicol (2009). "On the radius of analyticity of solutions to the three-dimensional Euler equations". *Proc. Amer. Math. Soc.* 137.2, pp. 669–677. MR: 2448589 (cit. on pp. 2171, 2173, 2175).

C. David Levermore and Marcel Oliver (1997). "Analyticity of solutions for a generalized Euler equation". *J. Differential Equations* 133.2, pp. 321–339. MR: 1427856 (cit. on pp. 2171, 2173, 2175).

Y. Charles Li and Zhiwu Lin (2011). "A resolution of the Sommerfeld paradox". *SIAM J. Math. Anal.* 43.4, pp. 1923–1954. MR: 2831254 (cit. on p. 2156).

Zhiwu Lin and Chongchun Zeng (2011a). "Inviscid dynamical structures near Couette flow". *Arch. Ration. Mech. Anal.* 200.3, pp. 1075–1097. MR: 2796139 (cit. on p. 2161).

– (2011b). "Small BGK waves and nonlinear Landau damping". *Comm. Math. Phys.* 306.2, pp. 291–331. MR: 2824473 (cit. on p. 2161).

Richard S Lindzen (1988). "Instability of plane parallel shear flow (toward a mechanistic picture of how it works)". *Pure and applied geophysics* 126.1, pp. 103–121 (cit. on p. 2159).

J. H. Malmberg, C. B. Wharton, R. W. Gould, and T. M. O'Neil (1968). "Plasma wave echo experiment". *Physical Review Letters* 20.3, pp. 95–97 (cit. on p. 2162).

Clément Mouhot and Cédric Villani (2011). "On Landau damping". *Acta Math.* 207.1, pp. 29–201. MR: 2863910 (cit. on pp. 2158, 2162, 2174).

William M'F Orr (1907). "The Stability or Instability of the Steady Motions of a Perfect Liquid and of a Viscous Liquid. Part I: A Perfect Liquid". In: *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*. Vol. 27. JSTOR, pp. 9–68 (cit. on pp. 2156, 2159).

Lord Rayleigh (1879/80). "On the Stability, or Instability, of certain Fluid Motions". *Proc. Lond. Math. Soc.* 11, pp. 57–70. MR: 1575266 (cit. on p. 2156).

– (1895/96). "On the Stability or Instability of certain Fluid Motions (III.)" *Proc. Lond. Math. Soc.* 27, pp. 5–12. MR: 1576484 (cit. on p. 2156).

– (1887/88). "On the Stability or Instability of certain Fluid Motions, II". *Proc. Lond. Math. Soc.* 19, pp. 67–74. MR: 1576885 (cit. on p. 2156).

Osborne Reynolds (1883). "An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall Be Direct or Sinuous, and of the Law of Resistance in Parallel Channels." *Proceedings of the Royal Society of London* 35.224-226, pp. 84–99 (cit. on p. 2156).

Peter J. Schmid and Dan S. Henningson (2001). *Stability and transition in shear flows*. Vol. 142. Applied Mathematical Sciences. Springer-Verlag, New York, pp. xiv+556. MR: 1801992 (cit. on p. 2156).

Herbert Brian Squire (1933). "On the stability for three-dimensional disturbances of viscous fluid flow between parallel walls". In: *Proc. R. Soc. Lond. A*. Vol. 142. 847. The Royal Society, pp. 621–628 (cit. on p. 2157).

Lloyd N. Trefethen, Anne E. Trefethen, Satish C. Reddy, and Tobin A. Driscoll (1993). "Hydrodynamic stability without eigenvalues". *Science* 261.5121, pp. 578–584. MR: 1229495 (cit. on p. 2162).

J. Vanneste (2001/02). "Nonlinear dynamics of anisotropic disturbances in plane Couette flow". *SIAM J. Appl. Math.* 62.3, pp. 924–944. MR: 1897729 (cit. on p. 2162).

J Vanneste, PJ Morrison, and T Warn (1998). "Strong echo effect and nonlinear transient growth in shear flows". *Physics of Fluids* 10.6, pp. 1398–1404 (cit. on p. 2162).

Akiva M. Yaglom (2012). *Hydrodynamic instability and transition to turbulence*. Vol. 100. Fluid Mechanics and its Applications. With a foreword by Uriel Frisch and a memorial note for Yaglom by Peter Bradshaw. Springer, Dordrecht, pp. xii+600. MR: 3185102 (cit. on pp. 2156, 2159).

JH Yu and CF Driscoll (2002). "Diocotron wave echoes in a pure electron plasma". *IEEE transactions on plasma science* 30.1, pp. 24–25 (cit. on p. 2162).

JH Yu, CF Driscoll, and TM O'Neil (2005). "Phase mixing and echoes in a pure electron plasma". *Physics of plasmas* 12.5, p. 055701 (cit. on p. 2162).

Received 2018-02-25.

JACOB BEDROSSIAN
jacob@cscamm.umd.edu
jacob@math.umd.edu

YU DENG
yudeng@cims.nyu.edu

NADER MASMOUDI
masmoudi@cims.nyu.edu