



PAPER • OPEN ACCESS

On variable viscosity and enhanced dissipation

To cite this article: Xian Liao and Christian Zillinger 2023 *Nonlinearity* **36** 6071

View the [article online](#) for updates and enhancements.

You may also like

- [Aubry–Mather theory on graphs](#)
Antonio Siconolfi and Alfonso Sorrentino
- [Isodrastic magnetic fields for suppressing transitions in guiding-centre motion](#)
J W Burby, R S MacKay and S Naik
- [Impact of temperature dependent diffusion coefficients on heat and mass transport in viscoelastic liquid using generalized Fourier theory](#)
M Nawaz, Uzma Arif and Imran Haider Qureshi

On variable viscosity and enhanced dissipation

Xian Liao and Christian Zillinger* 

Karlsruhe Institute of Technology, Englerstraße 2, 76131 Karlsruhe, Germany

E-mail: christian.zillinger@kit.edu

Received 23 November 2021; revised 27 August 2023

Accepted for publication 29 September 2023

Published 12 October 2023

Recommended by Professor Beatrice Pelloni



CrossMark

Abstract

In this article we consider the two-dimensional Navier–Stokes equations with variable viscosity depending on the vertical position. As our main result we establish linear enhanced dissipation near the *non-affine* stationary states replacing Couette flow. For instance, these shear flows may grow exponentially. Moreover it turns out that, in contrast to the constant viscosity case, *decreasing* viscosity leads to *stronger* enhanced dissipation and *increasing* viscosity leads to *weaker* dissipation.

Mathematics subject classification: 35Q30, 76D05, 76E05

Keywords: Navier–Stokes, variable viscosity, mixing, enhanced dissipation, stability

1. Introduction

In the present paper we are concerned with the two-dimensional incompressible Navier–Stokes equations in the presence of large (stratified) viscosity variations

$$\begin{cases} \partial_t v + v \cdot \nabla v - \operatorname{div}(\mu S v) + \nabla p = 0, \\ \operatorname{div} v = 0. \end{cases} \quad (1)$$

Here $t \in [0, \infty)$ and $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{T} \times \mathbb{R}$ denote the time and space variables respectively. The vector-valued function $v = v(t, x, y) : [0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and the scalar function $p = p(t, x, y) :$

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

$[0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}$ denote the unknown velocity vector field and the unknown pressure of the two-dimensional flow, respectively.

The symmetric part of the velocity gradient

$$\frac{1}{2}Sv := \frac{1}{2} \left(\nabla v + (\nabla v)^T \right)$$

denotes the symmetric deformation tensor. The viscosity coefficient μ is a given non-constant positive scalar function, and in the present paper we consider the case of stratified viscosity

$$\mu = \mu(y)$$

depending on the vertical direction only, and study its interplay with 2D shear flows.

Viscous stratification is a typical phenomenon not only in nature (e.g. in the atmosphere and ocean flows) but also in industrial application (e.g. in the chemical and food industry). The (in)stabilities in viscosity-stratified flows have attracted large interest in the physics community [Cra69, GS14, Hei85, Lin44, HB87, Yih67]. While additional dissipation at first sight suggests stabilization¹, in experiments viscosity exhibits dual roles [Dra02, chapter 8, p 160]:

(R1) A stabilizing role due to the dissipation of energy;

(R2) A more subtle destabilizing role.

Yih [Yih67] showed that the instability in a low Reynolds number flow can be caused by viscosity stratifications (see also Craik [Cra69] for the study of flows with continuous viscosity stratification). These results motivated decades of active research on the instability caused by viscosity interfaces, see [GS14] for a review paper on this topic.

In this paper we aim to show (R2) by investigating the stability (R1) for the case of two-dimensional shear flows. More precisely, we consider the model (1) of the fluids with equal density/temperature but different viscosities, which can for instance be used to describe the transport of the highly viscous oil and an immiscible low viscous lubricant (see e.g. [JRR84, PV91] for the relevant instability analysis). We then study the asymptotic behaviour of perturbations to the shear flow solutions

$$\mu = \mu(y), \quad v = \begin{pmatrix} U(y) \\ 0 \end{pmatrix}, \quad (2)$$

which satisfy the hydrostatic balance

$$\partial_y (\mu \partial_y U) = 0. \quad (3)$$

As a consequence of (3), one already observes that the variable viscosity coefficient changes the slope of the underlying velocity profile, such that the viscous stratification comes into play, even at high Reynolds numbers $\mu \ll 1$.²

In recent years there has been extensive research on the stability study of the shear flows (2) for the inviscid fluids with

¹ The Orr–Sommerfeld eigenvalue problem has only positive eigenvalues for Couette flows, which implies the stability of Couette flows for all Reynolds number, but experiments showed instability under small but finite perturbations.

² The viscosity variations increase the order of the Orr–Sommerfeld equation from two to four, which makes a difference in the dynamics even at high Reynolds numbers (contrary to the intuitive expectation of negligible viscous effect).

$$\mu = 0,$$

and for the viscous fluids with constant viscosity

$$\mu = \text{const.} > 0.$$

Since the literature is extensive, we here do not provide a complete overview but refer the interested readers to the following recent works for further discussion [Jia20, LX19, WZZ18, IJ20a, IJ20b, Wid18, YL18, BMV16, EW15, LZ11, WZZ20, BGM17, BM15, DWZ21]. We, in particular, recall that for linearized equations around Couette flow

$$\mu = \mu_1 > 0, U(y) = y,$$

it can be shown by explicit calculations that the interplay of shearing and dissipation leads to damping with a rate

$$\exp\left(-C(\mu_1)^{\frac{1}{3}} t\right), \quad (4)$$

and thus on a time scale $(\mu_1)^{-\frac{1}{3}}$ much smaller than the dissipation time scale $(\mu_1)^{-1}$, if $\mu_1 > 0$ is a small constant. This phenomenon is hence called *enhanced dissipation* (see [BVW18] for further discussion and the analysis of the nonlinear problem), which highlights the stabilizing role (R1) of the viscosity: the larger the viscosity is, the stronger dissipation the flow exhibits, and hence the more stable the shear flows are.

We remark that settings of large dissipation, $\mu \gg 1$, can be treated by energy estimates in a rather straightforward way. Hence, a main focus in this article is on the setting of small dissipation, $\mu < 1$. Our main questions in this article are:

- (R1): Does enhanced dissipation also hold in settings where μ can vary by many orders of magnitude and, if so, in which spaces?
- (R2): How does the enhanced dissipation rate depend on μ ? In particular, given $\mu(y)$ in some region, how much should μ increase or decrease to change the (local) enhanced dissipation rate by a given factor?

In view of the *non-local* structure of the Biot–Savart law, one cannot simply ‘localize’ estimates, making both questions a very challenging problem. As we will discuss in section 4, a natural (sufficient) compatibility condition to connect adapted estimates with the non-local Biot–Savart law is given by a control of the Lipschitz constant of

$$\log(\mu).$$

We remark that by the balance relation (3), this also implies a control of $\log(\partial_y U(y))$. However, we stress that we will *not* require that μ is close to constant or that $U(y)$ is close to Couette flow. Indeed, a prototypical example is given by

$$\mu(y) = \mu_0 e^{\delta y}, U(y) = U_0 e^{-\delta y}, \quad y \in \mathbb{R}, \quad (5)$$

where $U_0, \delta \in \mathbb{R}$, $\mu_0 > 0$ are constants and $|\delta|$ is small (see remark 1.1 for further discussion). This profile is *locally* bilipschitz and we may thus construct ‘localized’ Fourier weights (see section 3) which can then be ‘glued’ together (see section 4).

Recall that by the hydrostatic balance (3):

$$\mu \partial_y U = \text{const.} =: \sigma \neq 0,$$

and hence the above heuristic from the constant viscosity case suggests that the ‘local damping rate’ should be given by

$$\left(\mu (\partial_y U)^2\right)^{\frac{1}{3}} = \frac{\sigma^{\frac{2}{3}}}{\mu^{\frac{1}{3}}(y)}.$$

It can thus vary by *arbitrarily many orders of magnitude* and is proportional to an *enhanced, negative* power of $\mu(y)$. This dependence shows both the stabilizing role (R1) for $\mu > 0$, as well as ‘a subtle destabilizing role’ (R2) in the sense that the larger the viscosity is, the weaker dissipation the flow exhibits and hence the less stable the shear flows are. This also helps to explain wall heating or cooling techniques (corresponding to the liquid flows or gas flows respectively) in industrial application, which produce less viscous flow near the wall, and hence stabilize the flows [BG81].

Theorem 1.1. *Let $\mu = \mu(y) \in C^2(\mathbb{R})$ with $\mu > 0$ be a given stratified viscosity profile. Then a shear flow $v = v(x, y) = \begin{pmatrix} U(y) \\ 0 \end{pmatrix}$, $(x, y) \in \mathbb{T} \times \mathbb{R}$ such that*

$$\mu \partial_y U = \text{const} \tag{6}$$

is a stationary solution of the Navier–Stokes equation (1), and the linearized equations around this solution in vorticity formulation read

$$\begin{aligned} \partial_t \omega + U \partial_x \omega &= U'' v_2 + \text{div}(\mu \nabla \omega) - \text{div}(\mu' \nabla v_1) - \mu'' \partial_x v_2, \\ v &= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -\nabla^\perp (-\Delta)^{-1} \omega, \end{aligned} \tag{7}$$

where $U' = \partial_y U$, $U'' = \partial_y^2 U$, $\mu' = \partial_y \mu$, $\mu'' = \partial_y^2 \mu$ denote y derivatives, and $\nabla = \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}$, $\nabla^\perp = \begin{pmatrix} -\partial_y \\ \partial_x \end{pmatrix}$, $\text{div} = \nabla \cdot$.

Suppose also that $\mu > 0$ only varies gradually, in the sense that

$$\|(\ln \mu)'\|_{W^{1,\infty}} = \left\| \frac{\mu'}{\mu} \right\|_{L^\infty(\mathbb{R})} + \left\| \partial_y \frac{\mu'}{\mu} \right\|_{L^\infty(\mathbb{R})} < 0.0001, \tag{8}$$

and let $a(y) > 0$ be a weight function that also only varies gradually (e.g. $a = \mu$, $a = 1$ or $a = \frac{1}{\mu}$). Then the linearized equation (7) are stable in $L^2(a \, dx dy)$ and exhibit enhanced dissipation. More precisely, there exist a time-dependent family of operators $A(t)$ with

$$0.1 \|\omega(t)\|_{L^2(\mathbb{T} \times \mathbb{R}, a \, dx dy)}^2 \leq \|A(t)\omega(t)\|_{L^2(\mathbb{T} \times \mathbb{R}, a \, dx dy)}^2 \leq \|\omega(t)\|_{L^2(\mathbb{T} \times \mathbb{R}, a \, dx dy)}^2,$$

and a constant $C > 0$, such that, if the x -average of the initial vorticity vanishes: $\int_{\mathbb{T}} \omega_0 dx = 0$, then for all times $t > 0$ it holds that

$$\frac{d}{dt} \|A(t)\omega(t)\|_{L^2(\mathbb{T} \times \mathbb{R}, a \, dx dy)}^2 \leq -C \left\| \sqrt{\mu + (\mu(U')^2)^{1/3}} A(t)\omega \right\|_{L^2(\mathbb{T} \times \mathbb{R}, a \, dx dy)}^2. \tag{9}$$

Moreover, under further regularity assumptions these results also extend to stability of the ‘profile’ $W(t, x, y) := \omega(t, x + tU(y), y)$ in higher Sobolev norms H^N (see proposition 5.1 for a precise statement).

Remark 1.1. Let us comment on these results:

- While $\ln(\mu)$ is only allowed to vary gradually, μ may be arbitrarily large or small and grow or decrease at an up to exponential rate (with a small constant). A prototypical model case here is given by (5):

$$\mu(y) = \mu_0 e^{\delta y}, \quad U(y) = U_0 e^{-\delta y}.$$

In particular, μ is not close to constant and can vary by many orders of magnitude.

- A key challenge of the analysis lies in the fact that in view of the big changes in the size of μ we require suitably ‘localized’ estimates, which also hold for the case when μ is not close to constant. On the other hand the non-local structure of the Biot–Savart law implies that purely local estimates are not possible. We thus combine robust elliptic estimates (see section 3) and ‘gluing’ results (see section 4) to construct global energy functionals.
- That is, we note that for any given interval $I_j = (a_j, b_j)$ the restriction of U to I_j is bilipschitz and similarly the restriction of μ is bounded above and below. In section 3, as a model case we thus consider (smooth) affine/constant extensions

$$\partial_y U_j(y) = \begin{cases} \partial_y U(y) & \text{in } I_j, \\ \partial_y U(b_j) & \text{if } y \gg b_j, \\ \partial_y U(a_j) & \text{if } y \ll a_j, \end{cases}$$

$$\mu_j(y) = \frac{\sigma}{\partial_y U_j(y)}.$$

Unlike U and μ , these functions are globally bilipschitz or bounded above and below, respectively, and hence allow for an explicit construction of (localized) pseudo-differential operators A_j .

In a second step, in section 4, we partition and ‘glue’ the localized models of section 3. More precisely, given U and μ , we partition \mathbb{R} into intervals I_j and introduce an associated partition of unity $\sum_j \chi_j^2 = 1$. Using the results of section 3, we then define the operator

$$A = \sum_j \chi_j A_j \chi_j.$$

While this partition, of course, introduces several error and commutator terms, we show that these terms can be controlled (in suitably weighted L^2, H^1 and H^{-1} spaces) even if U is not globally bilipschitz and μ is not globally bounded away from 0 and ∞ . Instead we only require smallness of the logarithmic derivative

$$\frac{\partial_y \mu(y)}{\mu(y)} = -\frac{\partial_y^2 U(y)}{\partial_y U(y)},$$

which in the prototypical case (5) reduces to the constant function δ .

- For this prototypical case the local dissipation rate in (9) is given by

$$\mu_0 e^{\delta y} + (\mu_0 U_0^2)^{1/3} \delta^{2/3} e^{-\frac{1}{3}\delta y}. \quad (10)$$

In particular, this rate may vary by many orders of magnitude and the enhanced dissipation effect is visible in those regimes where both horizontal and vertical dissipation are much smaller than 1 and hence

$$(\mu_0 U_0^2)^{1/3} \delta^{2/3} e^{-\frac{1}{3}\delta y} \gg \mu_0 U_0^2 \delta^2 e^{-\delta y}.$$

We stress that this enhanced vertical rate (not just for this example) is proportional to $\mu^{-1/3}$. In particular, a *decrease* of μ by a factor 1000 corresponds to an *increase* of the ‘local’ dissipation rate by a factor 10. Conversely, *increasing* the viscosity corresponds to *weaker* dissipation.

- The small absolute constant $C > 0$ in (9) accounts for losses of factors due to error terms. That is, given a localized bound with constant C_1 for the ‘main dissipation term’ (see lemma 3.4 of section 3), we obtain a slightly worse constant at least of size $C_1/10$ for the full localized problem (see proposition 3.1) and finally a constant $C \geq C_1/100$ with a possible loss due to gluing errors (see proposition 4.1). In order to simplify notation we do not track these absolute constants throughout the article, and they may change from line to line.
- Unlike in the constant viscosity setting, for the shear flow considered in this article the second derivative of the shear U'' is non-trivial and does not approach zero under the (variable viscosity) heat flow.
- The nonlinear stability problem of the Navier–Stokes equations with constant viscosity has been studied in [BVW18]. This article extends these results in the linearized case to the stratified viscosity problem. In particular, we extend the by now common Cauchy–Kowalewskaya approach to the setting where $U'(y)$ and $\mu(y)$ may vary by many orders of magnitude (but may do so only gradually). We expect these methods to be of interest of their own for the wider community and applicable also to other related problems (e.g. the well-posedness issue of the variable viscosity Boussinesq equations).

Remark 1.2. Based on the local dissipation rate in theorem 1.1, at first sight one might also conjecture an estimate of the form

$$\|\exp(\nu t) \omega(t)\|_{L^2} \leq C \|\omega(0)\|_{L^2}.$$

However, such an estimate cannot be expected to hold in general, since the Biot–Savart law is non-local and not decaying fast enough. More precisely, if ω is highly localized in a region M , then the velocity field generated by ω exhibits decay away from M in terms of a power law of the distance $\text{dist}(y, M)$. In particular, if M' is a different region with much higher damping rate, then the decay of the Biot–Savart law in terms of $\text{dist}(M, M')$ is not sufficiently strong to compensate for the difference in dissipation rates.

The remainder of our article is structured as follows:

- In section 2 we introduce function spaces, changes of variables and notational conventions used throughout the article.
- As a first model setting in section 3 we establish linear L^2 stability for the case when μ varies only by a bounded factor: $\frac{\sup(\mu)}{\inf(\mu)} \leq 2$. This allows us to more transparently present the main tools of our proofs and discuss the necessity of assumptions.

- In section 4 we extend these L^2 stability results to the general setting by constructing local versions of several estimates. Here the non-local structure of the Biot–Savart law and the interaction of the localization and dissipation require careful analysis.
- Using the linear L^2 stability results as a building block, in section 5 we establish linear stability in H^N and thus prove theorem 1.1.

2. Stationary solutions and notation

In this section we establish that the shear flow $(U(y), 0)^T$ given in theorem 1.1 indeed is a stationary solution. Furthermore, we derive the linearized equation around this state in vorticity formulation.

In our analysis of the Navier–Stokes equations it is often convenient to work in Lagrangian coordinates moving with the underlying shear flow $(U(y), 0)^T$. Moreover, since we assume that U is strictly monotone there exists a change of coordinates $y \mapsto z = U(y)$ which straightens out the flow lines. However, as we discuss in section 4 this change of variables may be highly degenerate (not Lipschitz). For this reason, in the latter section we instead introduce families of bilipschitz changes of coordinates, which locally agree with z up to a factor.

Lemma 2.1 (Stationary solution). *Let $\mu = \mu(y) \in C^2(\mathbb{R})$ be a given function with $\mu > 0$. Let $\sigma \in \mathbb{R} \setminus \{0\}$, and let $U = U(y) \in C^3(\mathbb{R})$ satisfy*

$$\mu(y) U'(y) = \sigma.$$

Then $v(x, y) = (U(y), 0)^T \in C^3(\mathbb{T} \times \mathbb{R}; \mathbb{R}^2)$ is a stationary solution of the Navier–Stokes equation (1) with viscosity μ .

The linearized equations in vorticity formulation around this stationary solution are given by

$$\begin{aligned} \partial_t \omega + U \partial_x \omega - U'' v_2 &= \operatorname{div}(\mu \nabla \omega) - \operatorname{div}(\mu' \nabla v_1) - \mu'' \partial_x v_2, \\ v &= (v_1 \ v_2)^T = \nabla^\perp \Delta^{-1} \omega. \end{aligned} \tag{11}$$

Proof of lemma 2.1. Following theorem 1.1 we make the ansatz

$$\mu = \mu(y), \quad v = \begin{pmatrix} U(y) \\ 0 \end{pmatrix}. \tag{12}$$

The Navier–Stokes equation (1) then reduce to the following equations

$$\begin{pmatrix} -\partial_y(\mu \partial_y U) + \partial_x p \\ \partial_y p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The second equation $\partial_y p = 0$ implies $p = P(x)$ for some function P depending only on x , while $\partial_y(\mu \partial_y U)$ depends only on y . Hence, both functions need to equal a common constant, which yields the hydrostatic balance relation

$$\partial_y(\mu \partial_y U) = C_0 \tag{13}$$

and $p = P(x) = C_0 x + C_1$, where $C_0, C_1 \in \mathbb{R}$ are constants. In particular, specializing to the case $C_0 = 0$, we verify that our choice of U yields a stationary solution.

If we also allow for C_0 to be possibly non-trivial there are many solutions of potential interest:

- The Uniform flow: $U = \text{const.}$
- The Couette flow: $U = y$, with $\mu = \text{const.}$ or $\mu = C_0y + C_2$.
- The Poiseuille flow: $U = y(1 - y)$, with $\mu = \text{const.}$, $y \in [0, 1]$.
- The shear layer: $U = \tanh(y)$, with $\mu = \text{sech}^{-2}(y)$.
- The jet or wake: $U = \text{sech}^2(y)$, with $\mu = y \cosh^2(y) \coth(y)$.

In this article we restrict to the case $C_0 = 0$ since then for non-vanishing viscosity the (non-trivial) shear flow U has no critical points, which would pose an obstacle to damping estimates. Furthermore, in view of physical applications we are mainly interested in the case when the effective damping rate $\mu(\partial_y U)^2$ is not large (indeed extremely small).

In the following let U, μ be solutions of (13) which hence are solutions of the Navier–Stokes equations in velocity formulation. We may then obtain the equation for the vorticity

$$\omega = \nabla^\perp \cdot v, \quad \text{with } \nabla^\perp = \begin{pmatrix} -\partial_y \\ \partial_x \end{pmatrix},$$

by applying the operator $\nabla^\perp \cdot$ to the velocity equation (1). Notice that

$$\begin{aligned} \text{div}(\mu Sv) &= \begin{pmatrix} 2\partial_x \mu \partial_x v_1 & \partial_y \mu \partial_y v_1 + \partial_y \mu \partial_x v_2 \\ \partial_x \mu \partial_x v_1 + \partial_x \mu \partial_x v_2 & 2\partial_y \mu \partial_y v_2 \end{pmatrix}, \\ v &= \nabla^\perp \Delta^{-1} \omega = \begin{pmatrix} -\partial_y \Delta^{-1} \omega \\ \partial_x \Delta^{-1} \omega \end{pmatrix}. \end{aligned}$$

We may calculate (see also [HL20])

$$\nabla^\perp \cdot \text{div}(\mu Sv) = [(\partial_{yy} - \partial_{xx})\mu(\partial_{yy} - \partial_{xx}) + (2\partial_{xy})\mu(2\partial_{xy})] \Delta^{-1} \omega,$$

which can be equivalently expressed as

$$\Delta(\mu\omega) - 2\mu''\partial_x v_2 = \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2.$$

Thus we arrive at the vorticity formulation for the Navier–Stokes equations with viscosity μ :

$$\partial_t \omega + v \cdot \nabla \omega = \Delta(\mu\omega) - 2\mu''\partial_x v_2 \equiv \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2. \tag{14}$$

Finally, we linearize the vorticity equation (14) around this shear flow to arrive at the following linearized equation

$$\begin{aligned} \partial_t \omega + U(y)\partial_x \omega - U''(y)v_2 &= \Delta(\mu\omega) - 2\mu''\partial_x v_2 \\ &\equiv \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2. \end{aligned} \tag{15}$$

□

In the following we introduce some equivalent reformulations of linearized equation (11) in order to simplify our notation. We first observe that in the equation (11), all coefficient

functions do not depend on the x variable. Hence the evolution of the x -average of the vorticity which we denote by $\omega_{=}$ decouples as

$$\partial_t \omega_{=} = \partial_y (\mu \partial_y \omega_{=}) + \partial_y (\mu' \omega_{=}) = \partial_{yy} (\mu \omega_{=}).$$

The x average hence evolves as in a variable coefficient heat equation and does not influence the evolution of the orthogonal complement

$$\omega_{\neq} = \omega - \omega_{=}.$$

For this reason we in the following without loss of generality assume that initially

$$\omega_{=} = 0, \tag{16}$$

which then remains the case for all times.

As another consequence of the lack of x -dependence, the equations decouple after a Fourier transform in x , which we denote by

$$\hat{\omega}(t, k, y) = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ikx} \omega(t, x, y) dx.$$

Our equations read:

$$\partial_t \hat{\omega} + ikU(y) \hat{\omega} - U''(y) \frac{ik}{-k^2 + \partial_{yy}} \hat{\omega} = (-k^2 + \partial_{yy}) (\mu \hat{\omega}) + 2\mu'' \frac{k^2}{-k^2 + \partial_{yy}} \hat{\omega}.$$

We may further consider the vorticity moving with the underlying shear

$$W(t, x, y) = \omega(t, x + tU(y), y). \tag{17}$$

Expressed in Fourier variables it holds that

$$\mathcal{F}_x W(t, k, y) = e^{iktU(y)} \hat{\omega}(t, k, y),$$

and hence

$$\begin{aligned} \partial_t (\mathcal{F}_x W) - \frac{ikU''(y)}{-k^2 + (\partial_y - iktU'(y))^2} (\mathcal{F}_x W) &= \left(-k^2 + (\partial_y - iktU'(y))^2 \right) (\mu \mathcal{F}_x W) \\ &+ \frac{2\mu'' k^2}{-k^2 + (\partial_y - iktU'(y))^2} (\mathcal{F}_x W). \end{aligned} \tag{18}$$

Notice that after the Fourier transform with respect to the x variable, the equation (18) for $(\mathcal{F}_x W)(t, k, y)$ are decoupled with respect to k . Since

$$\|W(t, x, y)\|_{L^2} = \left\| \left\| (\mathcal{F}_x W)(t, k, y) \right\|_{L^2(\mathbb{R}, dy)} \right\|_{\ell_x^2(\mathbb{Z})},$$

it suffices to consider the evolution of the norm

$$\left\| (\mathcal{F}_x W)(t, k, y) \right\|_{L^2(\mathbb{R}, dy)}.$$

In the remainder of the article we will thus focus on the evolution of W which is localized at an arbitrary but fixed frequency $k \neq 0$.

In the following sections we establish asymptotic stability of W in Sobolev regularity. More precisely, we will first consider the special case where U is globally bilipschitz with comparable upper and lower Lipschitz constants in section 3. Building on these results, in section 4 we consider the general case, where we further introduce modified changes of coordinates adapted to the local behaviour of the coefficient functions. We remark already here that this construction requires further refinement for the general situation, but provides a good description if one additionally assumes that μ is globally comparable to a constant, which is the model setting of section 3. In section 4 we replace this global change of variables by a family of suitably localized coordinate changes, which accounts for the fact that μ and hence U' may change by many orders of magnitude. Finally, in section 5 we bootstrap the stability results in L^2 to establish stability in H^N .

3. A model case and L^2 estimates

In this section we impose the *additional assumption* that $\mu(y)$ is bounded above and below and require that

$$\frac{\sup(\mu)}{\inf(\mu)} \leq 2, \tag{19}$$

and we also impose such an assumption on the weight a . As we discuss in section 4 this can, of course, not be expected to hold in general. However, since $\mu(y)$ is assumed to be slowly varying, it does hold when restricting to suitable intervals (see remark 1.1 and lemma 4.2). A main challenge of the latter section is thus to ‘localize’ and ‘glue’ estimates in a way compatible with non-local interaction by the Biot–Savart law.

We remark that by the hydrostatic balance (6), the condition (19) also implies that U' is bounded above and below (and without loss of generality strictly positive) and

$$\frac{\sup(U')}{\inf(U')} \leq 2.$$

Therefore, with this additional assumption the change of variables

$$y \mapsto z = \frac{U(y)}{u}$$

with

$$u := \inf U', \tag{20}$$

is *globally bilipschitz* with constants bounded by $\frac{1}{2}$ and 2.

Given a solution ω of the linearized Navier–Stokes equation (11), we may thus equivalently consider

$$W(t, x, z) := \omega(t, x + tU(y), y) \tag{21}$$

and study stability in $L^2(dx dz)$ in place of $L^2(dx dy)$. Here, with slight abuse of notation compared to (17), we identify W depending on y and on z , respectively.

For this unknown the linearized Navier–Stokes equations read as in (18):

$$\partial_t W = U'' V_2 + \operatorname{div}_t (\mu \nabla_t W) - \operatorname{div}_t (\mu' \nabla_t V_1) - \mu'' \partial_x V_2, \tag{22}$$

where $\nabla_t := \left(\frac{\partial_x}{u'}(\partial_z - tu\partial_x) \right)$ and V denotes the velocity in these new coordinates. We here note that $\frac{U'}{u}$ is comparable to 1 and that $(\partial_z - tu\partial_x)$ is a Fourier multiplier.

The following proposition summarizes our main results for this section and employs a by now common Lyapunov functional/energy approach (see for instance [MSHZ20, BMV16, TW19, Lis20]), where a key challenge lies in constructing a suitable time, frequency and space-dependent operator A which captures possible growth in the evolution of solutions.

Proposition 3.1. *Let $\mu \in C^2(\mathbb{R}; (0, \infty))$, $U \in C^3(\mathbb{R}; \mathbb{R})$, $a \in C^2(\mathbb{R}; (0, \infty))$ satisfy the assumptions of theorem 1.1 and additionally suppose that μ and a are bounded above and below and satisfy (19). Then the results of theorem 1.1 hold.*

More precisely, let W be a solution of the linearized Navier–Stokes equation (22). Then there exists a time-dependent family of operators $A(t)$ such that for any initial data $\omega_0 \in L^2$ with $\int_{\mathbb{T}} \omega_0 dx = 0$, it holds that

$$c \|W(t)\|_{L^2(\mathbb{T} \times \mathbb{R}; a dx dz)} \leq \|A(t)W(t)\|_{L^2(\mathbb{T} \times \mathbb{R}; a dx dz)} \leq \|W(t)\|_{L^2(\mathbb{T} \times \mathbb{R}; a dx dz)},$$

for some positive constant $c \in [\frac{1}{2}, 1]$ (see lemma 3.3 below for details). Furthermore, there exists a constant $C > 0$ such that

$$\begin{aligned} \frac{d}{dt} \|AW\|_{L^2(\mathbb{T} \times \mathbb{R}; a dx dz)}^2 &\leq -C \|a\|_{L^\infty} \sum_k \int_{\mathbb{R}} \left(\inf(\mu)k^2 + (\nu k^2)^{\frac{1}{3}} + \nu \left(\frac{\xi}{u} - kt \right)^2 + \frac{u}{1 + (\xi - ktu)^2} \right) \\ &\quad \times |\mathcal{F}_{x,z}(AW)|^2 d\xi, \end{aligned} \tag{23}$$

where the effective damping rate ν is defined as

$$\nu := \inf \mu (U')^2, \tag{24}$$

and u denotes the infimum of U' as in (20).

Let us comment on these results:

- (1) In this model case all of a, μ, U' are additionally assumed to be bounded and comparable to their supremum and infimum. The general case, where these quantities may slowly vary, is established in section 4.
- (2) The operator $A(t)$ will be constructed in terms of a Fourier multiplier in definition 3.2 in section 3.1.
- (3) The decay rate $\inf(\mu)k^2 + (\nu k^2)^{\frac{1}{3}} + \nu \left(\frac{\xi}{u} - kt \right)^2$ quantifies the enhanced dissipation mechanism. More precisely, the first term corresponds to horizontal dissipation, which is not enhanced. For the vertical dissipation we distinguish between frequency regions. If ξ is far from resonant, that is, $|\frac{\xi}{uk} - t| \gtrsim (\nu k^2)^{-\frac{1}{3}}$, then the latter term dominates. If instead ξ is close to resonant, we still obtain a dependence on $(\nu k^2)^{\frac{1}{3}}$.
- (4) The last multiplier $\frac{u}{1 + (\xi - ktu)^2}$ allows us to control the error due to $U''V_2$. We remark that in regimes of large enhanced dissipation this error can be easily absorbed. It is hence only of relevance in regimes where νk^2 is very small.

In order to obtain the estimate of proposition 3.1, we need to control various error terms appearing in

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|AW\|_{L^2(\mathbb{T} \times \mathbb{R}; dx dz)}^2 &= \langle \dot{A}W, AW \rangle + \langle AW, AU''V_2 \rangle \\ &\quad + \langle AW, A \operatorname{div}_t(\mu \nabla_t W) \rangle \\ &\quad - \langle AW, A \operatorname{div}_t(\mu' \nabla_t V_1) \rangle \\ &\quad - \langle AW, A \mu'' \partial_x V_2 \rangle. \end{aligned} \tag{25}$$

In particular, these include various commutation errors involving the multiplier A and the variable coefficients in the differential operators. We thus split the proof of proposition 3.1 into the following subsections.

3.1. The Fourier multiplier m and operator A

Based on the heuristics of the constant viscosity case, a major stabilizing effect in (25) is expected to be given by the dissipation term

$$- \langle \nabla_t AW, \mu \nabla_t AW \rangle,$$

where we also need to control several commutator errors. In particular, in view of the time dependence of ∇_t , this dissipation is very strong when considering frequencies very far from resonant, $|\xi - ukt| \gg 1$. If instead $|\xi - ukt| \lesssim 1$, this term is possibly too small to control errors and we therefore need to rely on

$$\langle \dot{A}W, AW \rangle \leq 0$$

to absorb errors. Moreover, also in case where the (vertical) dissipation is small, we need control the velocity error term

$$\langle AW, AU''V_2 \rangle = \left\langle AW, A \frac{U''}{U'} U' V_2 \right\rangle$$

by the decay of A . This motivates the following definition.

Definition 3.2 (Decreasing multiplier and Fourier sets). Let μ, U' be given as in proposition 3.1, and let ν, u be the local dissipation rate and the local shear rate defined in (24) and (20) respectively.

We define the good set $G_t \subset \mathbb{Z} \times \mathbb{R}$ in the frequency space with respect to (x, z) by

$$G_t = \left\{ (k, \xi) \in \mathbb{Z} \times \mathbb{R} \mid k \neq 0, \quad \left| \frac{\xi}{ku} - t \right| \geq 0.1 (\nu k^2)^{-\frac{1}{3}} \right\},$$

and the bad set B_t as the complement (excluding $k = 0$)

$$B_t = \left\{ (k, \xi) \in \mathbb{Z} \times \mathbb{R} \mid k \neq 0, \quad \left| \frac{\xi}{ku} - t \right| < 0.1 (\nu k^2)^{-\frac{1}{3}} \right\}.$$

For any fixed k , if $B_t \cap \{k\} \times \mathbb{R}$ is non-empty, the set $G_t \cap \{k\} \times \mathbb{R}$ has two connected components, where we denote by $G_t^- \subset G_t$ such that

$$G_t^- \cap \{k\} \times \mathbb{R} = \{k\} \times (-\infty, ktu - 0.1|ku|(\nu k^2)^{-\frac{1}{3}}]$$

the half-line extending to $-\infty$, and by $G_t^+ \subset G_t$ such that

$$G_t^+ \cap \{k\} \times \mathbb{R} = \{k\} \times [ktu + 0.1|ku|(\nu k^2)^{-\frac{1}{3}}, +\infty)$$

the half-line extending to $+\infty$.

Associated with this partition we define a Fourier multiplier $m = m(t, k, \xi)$ by

$$\partial_t m(t, k, \xi) = \begin{cases} m(t, k, \xi) \left(-(\nu k^2)^{\frac{1}{3}} - 0.1 \frac{u}{1+u^2(\frac{\xi}{ku}-t)^2} \right), & \text{if } (k, \xi) \in B_t, \\ 0, & \text{else,} \end{cases} \quad (26)$$

and the asymptotic condition $\lim_{t \rightarrow -\infty} m(t, k, \xi) = 1$.

We denote the operator associated with the Fourier multiplier m by $A(t)$:

$$A\phi = \mathcal{F}^{-1}m\mathcal{F}\phi,$$

where \mathcal{F} denotes the Fourier transform with respect to $(x, z) \in \mathbb{T} \times \mathbb{R}$.

Remark 3.1. This multiplier combines features of the inviscid multiplier of [Zil17] and the constant viscosity multiplier of [Lis20, BVW18].

- The relative decay of m by $-(\nu k^2)^{1/3}$ compensates for the relatively weak dissipation in the *bad* Fourier region. Here the decay of A allows to establish damping of $\|AW\|_{L^2}^2$.
- The term $-\frac{u}{1+u^2(\frac{\xi}{ku}-t)^2}$ allows us to estimate contributions by $U''V_2$. As we discuss in Subsection 3.4, unlike in the constant viscosity setting here U'' might be very large. However, by our assumption that U' is slowly varying, we can exploit some smallness of $\frac{U''}{U'}$ in our estimates.

As we prove in the following subsection the multiplier m (and hence the operator A) satisfies several useful bounds and, in particular, serves to control various error terms when W is concentrated in the *bad* set.

Lemma 3.3. *Let m be as in definition 3.2. Then m satisfies the following estimates:*

(1) *There exists a constant $c \in [0.5, 1]$ independent of ξ and t such that*

$$c \leq m \leq 1.$$

(2) *The multiplier m is constant (independent of ξ and t , but might depend on k) for large positive or negative times. By the conventions of our definition one of these constants is chosen as 1 and the other as c :*

$$m(t, k, \xi) = c \text{ if } t > \frac{\xi}{ku} + 0.1(\nu k^2)^{-\frac{1}{3}},$$

$$m(t, k, \xi) = 1 \text{ if } t < \frac{\xi}{ku} - 0.1(\nu k^2)^{-\frac{1}{3}}.$$

(3) *The operator A is a continuous invertible operator from L^2 to L^2 and satisfies*

$$c\|\phi\|_{L^2} \leq \|A\phi\|_{L^2} \leq \|\phi\|_{L^2}$$

for all $\phi \in L^2(dx dz)$.

We begin by discussing the properties of the multiplier m , which may be computed explicitly in terms of integrals.

Proof of lemma 3.3. By definition it holds that $\partial_t m \leq 0$ and hence

$$m(t) = m(-\infty) + \int_{-\infty}^t \partial_t m \leq m(-\infty) = 1.$$

Furthermore, we may explicitly compute m as

$$m(t, k, \xi) = \exp \left(\int_{-\infty}^t \frac{\partial_t m}{m} 1_{B_\tau}(k, \xi) \, d\tau \right), \tag{27}$$

where we used that $m(-\infty, k, \xi) = 1$. Since $\frac{\partial_t m}{m} = -((\nu k^2)^{\frac{1}{3}} + 0.1 \frac{u}{1+u^2(\frac{\xi}{ku}-t)^2})$ inside B_t such that $-0.1(\nu k^2)^{-\frac{1}{3}} < t - \frac{\xi}{ku} < 0.1(\nu k^2)^{-\frac{1}{3}}$ while $\frac{\partial_t m}{m} = 0$ outside B_t , we define the constant c to be

$$c := \exp \left(- \int_{-0.1(\nu k^2)^{-\frac{1}{3}}}^{0.1(\nu k^2)^{-\frac{1}{3}}} \left((\nu k^2)^{\frac{1}{3}} + 0.1 \frac{u}{1+u^2 t^2} \right) dt \right),$$

such that (1) and (2) in lemma 3.3 hold. We now estimate the size of c : Since

$$\begin{aligned} & - \int_{-0.1(\nu k^2)^{-\frac{1}{3}}}^{0.1(\nu k^2)^{-\frac{1}{3}}} (\nu k^2)^{\frac{1}{3}} dt = -0.2, \\ & - \int_{-0.1(\nu k^2)^{-\frac{1}{3}}}^{0.1(\nu k^2)^{-\frac{1}{3}}} \frac{u}{1+u^2 t^2} dt = - \arctan(\tau) \Big|_{\tau=-0.1u(\nu k^2)^{-\frac{1}{3}}}^{0.1u(\nu k^2)^{-\frac{1}{3}}} \geq -\pi, \end{aligned}$$

we have

$$c \geq \exp(-0.2 - 0.1\pi) > 0.5.$$

Finally, by Parseval’s identity these bounds for the multiplier m are equivalent to L^2 bounds for the operator A . □

3.2. Proof of proposition 3.1

Given our multiplier m in definition 3.2, our main task in the following is to establish suitable estimates for (25):

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|AW\|_{L^2}^2 &= \langle AW, \dot{A}W \rangle \\ &+ \langle AW, A \operatorname{div}_t(\mu \nabla_t) W \rangle \\ &- \langle AW, A \operatorname{div}_t(\mu' \nabla_t) V_1 \rangle - \langle AW, A \mu'' \partial_x V_2 \rangle \\ &+ \langle AW, A U'' V_2 \rangle. \end{aligned} \tag{28}$$

More precisely, we need to show that the dissipation and the decay of $m(t)$ are strong enough to absorb possible growth and that hence $\|AW\|^2$ is non-increasing in time. Integrating these estimates we thus obtain a Lyapunov functional, which allows us to prove proposition 3.1.

The following lemmas summarize our estimates for the main dissipation term, the error due variable viscosity and the error due to the non-affine shear profile, which will allow us to prove proposition 3.1. The proof of each lemma is then given in the following subsections.

Lemma 3.4. *Let μ, U, W be as in proposition 3.1 and let A be given as in definition 3.2. Then for all times $t > 0$ the following dissipation estimate holds:*

$$0.2 \langle AW, \dot{AW} \rangle + \langle AW, A(\operatorname{div}_t(\mu \nabla_t) W) \rangle \leq -0.001 \sum_k \int_{\mathbb{R}} \left(\inf(\mu) k^2 + (\nu k^2)^{\frac{1}{3}} + \nu \left(\frac{\xi}{u} - kt \right)^2 + \frac{0.1u}{1 + u^2 \left(\frac{\xi}{ku} - t \right)^2} \right) |\mathcal{F}_{x,z}(AW)|^2 d\xi. \tag{DE}$$

Lemma 3.5. *Let μ, U, W be as in proposition 3.1 and let A be given as in definition 3.2. Then for all times $t > 0$ the following viscosity error estimate holds:*

$$0.2 \langle AW, \dot{AW} \rangle - \langle AW, A(\operatorname{div}_t(\mu' \nabla_t) V_1 + \mu'' \partial_x V_2) \rangle \leq 0.0005 \sum_k \int_{\mathbb{R}} \left(\inf(\mu) k^2 + (\nu k^2)^{\frac{1}{3}} + \nu \left(\frac{\xi}{u} - kt \right)^2 + \frac{0.1u}{1 + u^2 \left(\frac{\xi}{ku} - t \right)^2} \right) |\mathcal{F}_{x,z}(AW)|^2 d\xi. \tag{EE}$$

Lemma 3.6. *Let μ, U, W be as in proposition 3.1 and let A be given as in definition 3.2. Then for all times $t > 0$ the following velocity error estimate holds:*

$$0.2 \langle AW, \dot{AW} \rangle + \langle AW, AU''V_2 \rangle \leq 0. \tag{VE}$$

Before proceeding to the proof of these lemmas, let us discuss how they can be used to establish proposition 3.1.

Proof of proposition 3.1. Let W be a given solution, let A be as in definition 3.2 and consider the time derivative of the energy $\|AW\|_{L^2}^2$ as computed in (28). Then by the results of lemmas 3.4–3.6 it holds that

$$\frac{1}{2} \frac{d}{dt} \|AW\|_{L^2}^2 \leq -10^{-5} \sum_k \int_{\mathbb{R}} \left(\inf(\mu) k^2 + (\nu k^2)^{\frac{1}{3}} + \nu \left(\frac{\xi}{u} - kt \right)^2 + \frac{0.1u}{1 + u^2 \left(\frac{\xi}{ku} - t \right)^2} \right) |\mathcal{F}_{x,z}(AW)|^2 d\xi,$$

which is the desired estimate. Moreover, by the results of lemma 3.3 it holds that

$$c^2 \|W\|_{L^2}^2 \leq \|AW\|_{L^2}^2 \leq \|W\|_{L^2}^2.$$

□

3.3. Proof of the dissipation estimate, lemma 3.4

We recall that by definition 3.2, at any time $t > 0$ the frequency space can be decomposed into the three regions G_t^+, B_t, G_t^- and that the operator A is a multiple of the identity, when restricted to G_t^+ or G_t^- . We therefore also split

$$W = W_1 + W_2 + W_3 \tag{29}$$

according to these Fourier regions and study

$$\langle AW_i, \text{Adiv}_t(\mu \nabla_t) W_j \rangle, \quad i, j \in \{1, 2, 3\}.$$

As we discuss in the following steps, here the ‘diagonal terms’ $i = j$ are well behaved, while terms with $i \neq j$ require us to exploit some cancellation properties (for the case of constant viscosity and affine shear these terms identically vanish due to the disjoint Fourier support).

3.3.1. Estimates for $\langle AW_j, \text{Adiv}_t(\mu \nabla_t) W_j \rangle, j = 1$ or 3. Since $AW_1 = W_1$, we may explicitly compute that (with $a = 1$)

$$\begin{aligned} \langle AW_1, \text{Adiv}_t(\mu \nabla_t) W_1 \rangle &= \langle W_1, \text{div}_t(\mu \nabla_t) W_1 \rangle \\ &= \left\langle W_1, \mu \partial_x^2 W_1 + U' \left(\frac{1}{u} \partial_z - t \partial_x \right) \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \right\rangle \\ &= - \langle \partial_x W_1, \mu \partial_x W_1 \rangle - \left\langle U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \right\rangle \\ &\quad - \left\langle \frac{1}{u} \partial_z U' W_1, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \right\rangle \\ &= - \|\sqrt{\mu} \partial_x W_1\|_{L^2}^2 - \|\sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1\|_{L^2}^2 - \left\langle \frac{\partial_z U'}{u} \sqrt{\mu} W_1, \sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \right\rangle. \end{aligned}$$

We thus obtain both the desired horizontal and vertical dissipation terms, as well as one error term. For the error term, we use the fact that by the chain rule

$$\frac{1}{u} \partial_z U' = \frac{U''}{U'} \Big|_z$$

and that this factor is thus small by assumption. Furthermore, since $\int W dx = 0$, we may apply Poincaré’s inequality in x and thus estimate

$$\left| \left\langle \frac{\partial_z U'}{u} \sqrt{\mu} W_1, \sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \right\rangle \right| \leq \left\| \frac{1}{u} \partial_z U' \right\|_{L^\infty} \|\sqrt{\mu} \partial_x W_1\|_{L^2} \|\sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1\|_{L^2}$$

and use Young’s inequality to absorb this error into the decay. The estimate for $AW_3 = cW_3$ is analogous.

If the weight a is not constant, we additionally obtain contributions in the error term:

$$\int_{\mathbb{T} \times \mathbb{R}} W_1 \mu \frac{U'}{u} \frac{\partial_z a}{a} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1 \, a dx dz \leq 2 \left\| \frac{\partial_z a}{a} \right\|_{L^\infty} \|\sqrt{\mu} \partial_x W_1\|_{L^2} \|\sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_1\|_{L^2},$$

and

$$\left\langle AW_1, \frac{1}{a} [A, a] \text{div}_t(\mu \nabla_t) W_1 \right\rangle \leq \left\| \frac{1}{a} [A, a] \right\|_{L^\infty} \langle W_1, \text{div}_t(\mu \nabla_t) W_1 \rangle,$$

where A is a Fourier multiplier with a Lipschitz symbol m such that

$$\left\| \frac{1}{a} [A, a] \right\|_{L^\infty} \leq C_m \left\| \frac{\partial_z a}{a} \right\|_{L^\infty}.$$

Since $\left\| \frac{\partial_z a}{a} \right\|_{L^\infty} = \left\| \frac{u}{U'} \frac{\partial_z a}{a} \right\|_{L^\infty} \ll 1$ if a is gradually varying, this term can be easily absorbed into the dissipation terms with a very small loss of constant. For simplicity of presentation, in the following we thus focus on the case $a \equiv 1$.

3.3.2. Estimates for $\langle AW_i, A \operatorname{div}_t(\mu \nabla_t) W_j \rangle$ with $(i, j) = (1, 3)$ or $(i, j) = (3, 1)$. While W_1, W_3 are disjointly supported in Fourier space, since μ and U' are not constant, the L^2 product in general does not vanish. In the following we simply take $a = 1$.

We observe that

$$\langle AW_i, A \operatorname{div}_t(\mu \nabla_t) W_j \rangle = \langle A^2 W_i, \operatorname{div}_t(\mu \nabla_t) W_j \rangle$$

and that A^2 is a multiple of the identity. In the following we may thus for simplicity of notation instead consider

$$\begin{aligned} \langle W_i, \operatorname{div}_t(\mu \nabla_t) W_j \rangle &= -\langle \partial_x W_i, \mu \partial_x W_j \rangle - \left\langle U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_i, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle \\ &\quad + \left\langle \frac{1}{u} \partial_z U' W_i, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle. \end{aligned}$$

By the same argument as above, the quantity

$$\left| \left\langle \frac{1}{u} \partial_z U' W_i, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle \right| \leq \left\| \frac{1}{u} \partial_z U' \right\|_{L^\infty} \|\sqrt{\mu} \partial_x W_i\|_{L^2} \|\sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j\|_{L^2}$$

can be considered a negligible error term.

For the remaining terms, a simple estimate by Hölder's or Young's inequality is not sufficient and we hence instead need to exploit that the Fourier supports of W_1, W_3 are contained in

$$\begin{aligned} G_{k,t}^- &:= \left\{ \xi \in \mathbb{R} \mid \frac{\xi}{u} - kt \leq -0.1 \nu^{-\frac{1}{3}} |k|^{\frac{1}{3}} \right\}, \\ G_{k,t}^+ &:= \left\{ \xi \in \mathbb{R} \mid \frac{\xi}{u} - kt \geq 0.1 \nu^{-\frac{1}{3}} |k|^{\frac{1}{3}} \right\}, \end{aligned}$$

and are hence very well separated. In particular, by Plancherel's theorem, we may introduce a (smooth Littlewood–Payley) frequency projection operator P to frequencies larger than $0.1 \nu^{-1/3} u$ (a lower bound on the distance between $G_{k,t}^-$ and $G_{k,t}^+$). Then it holds that

$$\langle \partial_x W_i, \mu \partial_x W_j \rangle = \langle \partial_x W_i, (P\mu) \partial_x W_j \rangle \leq \|P\mu\|_{L^\infty} \|\partial_x W_i\|_{L^2} \|\partial_x W_j\|_{L^2}.$$

Recalling the fact that in this section we assume that μ is comparable to its supremum by (19) it thus suffices to show that

$$\|P\mu\|_{L^\infty} / \|\mu\|_{L^\infty} \tag{30}$$

is small, which then allows us to absorb this error term using Young's inequality.

Indeed, if χ denotes the smooth Fourier multiplier corresponding to P , then we may write

$$P\mu = \mathcal{F}^{-1} \left(\chi \frac{\xi}{\xi} \mathcal{F}\mu \right)$$

to bound this by a constant times

$$\nu^{1/3} \left\| \frac{1}{u} \partial_z \mu \right\|_{L^\infty} = \nu^{1/3} \left\| \frac{\mu}{U'} \frac{\partial_y \mu}{\mu} \right\|_{L^\infty} \leq 2 \left\| \left(\frac{\mu}{U'} \right)^{\frac{1}{3}} \right\|_{L^\infty} \left\| \mu \frac{\partial_y \mu}{\mu} \right\|_{L^\infty}.$$

As previously remarked, in regimes of large horizontal dissipation the proof greatly simplifies. Hence, without loss of generality, here and in the following we restrict to the regime, where

$$\mu \ll \left(\mu U'^2 \right)^{\frac{1}{3}} \Leftrightarrow \frac{\mu}{U'} \ll 1, \tag{31}$$

such that $\nu^{1/3} \left\| \frac{1}{u} \partial_z \mu \right\|_{L^\infty}$ is much smaller compared to μ , by our assumption that μ is slowly varying.

The estimates for

$$\left\langle U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_i, \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle = \mu U' \left\langle (PU') \left(\frac{1}{u} \partial_z - t \partial_x \right) W_i, \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle$$

are analogous.

3.3.3. Estimates for terms involving W_2 . It remains to discuss the influence of the Fourier-localized part W_2 in the *bad set*. We still take $a = 1$.

We first study the self-interaction term:

$$\begin{aligned} \langle AW_2, A \operatorname{div}_t (\mu \nabla_t) W_2 \rangle &= \langle AW_2, A \partial_x \mu \partial_x W_2 \rangle \\ &+ \left\langle AW_2, A \left(\frac{1}{u} \partial_z - t \partial_x \right) U' \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_2 \right\rangle \\ &- \left\langle AW_2, A \left(\frac{1}{u} \partial_z U' \right) \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_2 \right\rangle. \end{aligned}$$

Since none of A , μ and U' are constant, we cannot easily appeal to the negativity of the elliptic operator in this regime. Instead, we use that the dissipation is small and can hence be controlled by the decay of A .

More precisely, we note that the differential operator $\left(\frac{1}{u} \partial_z - t \partial_x \right)$ is bounded on the *bad set* with

$$\left\| \left(\frac{1}{u} \partial_z - t \partial_x \right) 1_{B_t} (D_{x,z}) \right\|_{L^2 \rightarrow L^2} \leq 0.1 (\nu k^2)^{-\frac{1}{3}} |k| = 0.1 \nu^{-\frac{1}{3}} |k|^{\frac{1}{3}}. \tag{32}$$

We thus may estimate

$$\begin{aligned} \left\langle AW_2, A \left(\frac{1}{u} \partial_z - t \partial_x \right) U' \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_2 \right\rangle_{L^2_z} &\leq c^{-1} \cdot \nu \cdot \left(0.1 \nu^{-\frac{1}{3}} |k|^{\frac{1}{3}} \right)^2 \|AW_2\|_{L^2_z}^2 \\ &= 0.02 (\nu k^2)^{\frac{1}{3}} \|AW_2\|_{L^2_z}^2. \end{aligned}$$

Recalling the decay of A and that

$$\langle AW_2, \dot{A}W_2 \rangle_{L^2_x} \leq -(\nu k^2)^{\frac{1}{3}} \|AW_2\|_{L^2_x}^2,$$

the above contribution can hence be absorbed.

Concerning the horizontal dissipation term, we note that if the horizontal dissipation is larger than the vertical dissipation, we may simply replace A by the identity in our construction. It hence suffices to consider the case where μk^2 (recall that we assume that μ is comparable to its supremum (19)) is much smaller than $(\nu k^2)^{1/3}$, similar as in (31). In this case, we simply estimate $\|A\|_{L^2 \rightarrow L^2}$ by a constant and use that μ is independent of x to obtain that

$$\langle AW_2, A\partial_x \mu \partial_x W_2 \rangle \leq \sup(\mu) c^{-1} \|\partial_x AW_2\|_{L^2}^2 < 0.01 (\nu k^2)^{\frac{1}{3}} \|AW_2\|_{L^2}^2$$

and similarly control

$$\begin{aligned} & \left| \left\langle AW_2, A \left(\frac{1}{u} \partial_z U' \right) \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_2 \right\rangle \right| \\ & \leq \left\| \frac{1}{u} \partial_z U' \right\|_{L^\infty} 2 \|\sqrt{\mu} U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_2\|_{L^2} \|\sqrt{\mu} AW_2\|_{L^2} \\ & \leq 0.001 (\nu k^2)^{\frac{1}{3}} \|AW_2\|_{L^2}^2. \end{aligned}$$

These terms can hence be absorbed into the decay due to \dot{A} .

Finally, it remains to discuss the cross terms

$$\begin{aligned} \langle AW_i, A \operatorname{div}_t (\mu \nabla_t) W_j \rangle &= \langle AW_i, A \partial_x \mu \partial_x W_j \rangle \\ &+ \left\langle AW_i, A \left(\frac{1}{u} \partial_z - t \partial_x \right) U' \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle \\ &+ \left\langle AW_i, A \left(\frac{1}{u} \partial_z U' \right) \mu U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W_j \right\rangle, \end{aligned}$$

where one i, j equals 2. By the preceding arguments we may again control the operator norm of A by a constant and use Hölder’s inequality to obtain estimates involving, for instance,

$$\|\sqrt{\mu} A \partial_x W_2\|_{L^2} \|\sqrt{\mu} A \partial_x W_1\|_{L^2}.$$

We thus use Young’s inequality with factors 10 and $\frac{1}{10}$, where the contribution by W_2 is absorbed into the decay of A and the contribution by W_1 is absorbed into the damping estimated established previously.

3.4. Proofs of lemmas 3.5 and 3.6

In the previous subsection we have established a dissipation estimate (DE) due the ‘main’ terms of the dissipation operator. This decay lies at the core of our damping mechanism. In the following we show that all other contributions to $\frac{d}{dt} \|AW\|^2$,

$$\begin{aligned} & - \langle AW, A \operatorname{div}_t (\mu' \nabla_t) V_1 \rangle - \langle AW, A \mu'' \partial_x V_2 \rangle \\ & + \langle AW, A U'' V_2 \rangle, \end{aligned}$$

can be considered as errors. In particular, all terms involving higher derivatives of μ can be considered as lower order.

We first recall that by definition of the vorticity

$$W = -U' \left(\frac{1}{u} \partial_z - t \partial_x \right) V_1 + \partial_x V_2,$$

and that $\frac{1}{u} \partial_z \mu' = \frac{\mu''}{U'}$. We may hence eliminate V_1 from this equation:

$$\begin{aligned} & \langle AW, A(\operatorname{div}_t(\mu' \nabla_t V_1) + \mu'' \partial_x V_2 - U'' V_2) \rangle \\ &= \left\langle A^2 W, \mu' \partial_{xx} V_1 + U' \left(\frac{1}{u} \partial_z - t \partial_x \right) \mu' U' \left(\frac{1}{u} \partial_z - t \partial_x \right) V_1 + \mu'' \partial_x V_2 - U'' V_2 \right\rangle \\ &= \left\langle A^2 W, \mu' \left(\partial_{xx} + \left(U' \left(\frac{1}{u} \partial_z - t \partial_x \right) \right)^2 \right) V_1 \right\rangle \\ &\quad + \left\langle A^2 W, U' \left(\frac{1}{u} \partial_z \mu' \right) U' \left(\frac{1}{u} \partial_z - t \partial_x \right) V_1 + \mu'' \partial_x V_2 - U'' V_2 \right\rangle \\ &= \left\langle A^2 W, -\mu' U' \left(\frac{1}{u} \partial_z - t \partial_x \right) W \right\rangle + \langle A^2 W, -\mu'' W \rangle + \langle A^2 W, 2\mu'' \partial_x V_2 - U'' V_2 \rangle. \end{aligned}$$

The first two terms can be absorbed into the horizontal and vertical dissipation by recalling that

$$\mu' = \frac{\mu'}{\mu} \mu \ll \mu$$

and that

$$\mu'' = \left(\partial_y \frac{\mu'}{\mu} + \left(\frac{\mu'}{\mu} \right)^2 \right) \mu \ll \mu.$$

Concerning the V_2 -terms, we argue similarly as in the inviscid case [CZZ19]. That is, the control of V corresponds to an elliptic estimate of the stream function

$$\left(\partial_x^2 + \left(U' \left(\frac{1}{u} \partial_z - t \partial_x \right) \right)^2 \right) \phi = W,$$

where we use that

$$V = \nabla_t^\perp \phi = \left(-U' \left(\frac{1}{u} \partial_z - t \partial_x \right), \partial_x \right)^T \phi.$$

Thus, if we define another simpler stream function and modified velocity by

$$\left(\partial_x^2 + u^2 \left(\frac{1}{u} \partial_z - t \partial_x \right)^2 \right) \psi = W,$$

with the constant $u = \inf U'$ and

$$\tilde{V} = \tilde{\nabla}_t^\perp \psi = \left(-(\partial_z - tu \partial_x), \partial_x \right)^T \psi,$$

then these are comparable quantities in the sense of bilinear forms. More precisely, testing these equations with either ψ or ϕ one obtains that the energies satisfy

$$\begin{aligned} \|\tilde{V}\|_{L^2}^2 &= \|(\partial_x, \partial_z - ut\partial_x)\psi\|_{L^2}^2 = \langle W, -\psi \rangle = \left\langle \left(\partial_x^2 + \left(U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \right)^2 \right) \phi, -\psi \right\rangle \\ &= \langle \partial_x \phi, \partial_x \psi \rangle + \left\langle U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \phi, U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \psi \right\rangle \\ &\quad + \left\langle U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \phi, \left(\frac{1}{u} \partial_z U' \right) \psi \right\rangle, \\ \|V\|_{L^2}^2 &= \left\| \left(\partial_x, U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \right) \phi \right\|_{L^2}^2 = \langle W, -\phi \rangle = \left\langle \left(\partial_x^2 + (\partial_z - ut\partial_x)^2 \right) \psi, -\phi \right\rangle \\ &= \langle \partial_x \psi, \partial_x \phi \rangle + \left\langle u \left(\frac{1}{u} \partial_z - t\partial_x \right) \psi, u \left(\frac{1}{u} \partial_z - t\partial_x \right) \phi \right\rangle. \end{aligned}$$

Hence, by Hölder’ inequality and using that $\sup U' \leq 2 \inf U' = 2u$, we deduce that

$$\|V\|_{L^2} = \left\| \left(\partial_x, U' \left(\frac{1}{u} \partial_z - t\partial_x \right) \right) \phi \right\|_{L^2} \approx \left\| \left(\partial_x, u \left(\frac{1}{u} \partial_z - t\partial_x \right) \right) \psi \right\|_{L^2} = \|\tilde{V}\|_{L^2} \tag{33}$$

are indeed comparable.

With this preparation we may express

$$\langle A^2 W, -U'' V_2 \rangle = \left\langle A^2 \left(\partial_x^2 + u^2 \left(\frac{1}{u} \partial_z - t\partial_x \right)^2 \right) \psi, -U'' V_2 \right\rangle,$$

where the constant coefficient operator now commutes with A and integrate by parts. We thus obtain a bound by

$$\|U''\|_{W^{1,\infty}} \|A \partial_x \tilde{V}\|_{L^2}^2 = \left\| \frac{U''}{u} \right\|_{W^{1,\infty}} \| \partial_x A \tilde{V} \|_{L^2}^2.$$

The estimate thus follows by the explicit characterization

$$u \|\partial_x A \tilde{V}\|_{L^2}^2 = \sum_k \int \frac{uk^2}{k^2 + u^2 \left(\frac{1}{u} \xi - kt \right)^2} |\mathcal{F}AW|^2 d\xi$$

and the fact that $\|U''/u\|_{W^{1,\infty}}$ is small by assumption.

For the last remaining term, we may simply estimate $\|\partial_x V_2\|_{L^2} \leq \|\partial_x W\|_{L^2}$ and use the smallness of μ'' to bound this contribution by the horizontal dissipation.

4. Localization and non-local Interactions

In this section we continue to consider the linearized equation (11) for the vorticity. Unlike in section 3 we here allow for μ (and hence also U') to vary by many orders of magnitude and also allow for a slowly varying weight a .

Since $U(y)$ is not Lipschitz and U' may vary by many orders of magnitude, this setting cannot be treated perturbatively and we cannot introduce a change of variables $z = \frac{U}{u}$ as in section 3.

Instead we need to:

- partition \mathbb{R} into suitable regions where μ is controlled. That is, μ can be arbitrarily large or small and need not be close to constant, but the infimum and supremum in a given region are comparable.
- extend each localized problem to a global problem and construct robust energy functionals. Here we also construct localized changes of variables z_j .
- show that all localized estimates can be ‘glued’ together in such a way that estimates close, even though the Biot–Savart law is non-local. Here is where we require that $\log(\mu)$ is slowly varying.

In order to prove these results we rely on robust comparison arguments for elliptic operators, which allow us to treat the case where μ is far from constant and on partition of unity methods for H^{-1} estimates. The latter method is shown to naturally involve a size constraint on $\frac{U''}{U'} = -\frac{\mu'}{\mu}$ and on $\frac{a'}{a}$, but does not require any smallness of U'' itself.

Our main result of this section are summarized in the following proposition.

Proposition 4.1. *Let μ, U and a satisfy the assumptions of theorem 1.1. Let $\omega_0 \in L^2$ with zero x -average be any initial data. Let $W(t, x, y) := \omega(t, x + tU(y), y)$ solves the linearized equation (7) with the initial data ω_0 in the new coordinates:*

$$\partial_t W = U'' V_2 + \operatorname{div}_t(\mu \nabla_t W) - \operatorname{div}_t(\mu' \nabla_t V_1) - \mu'' \partial_x V_2, \tag{34}$$

where $\nabla_t := \begin{pmatrix} \partial_x \\ \partial_y - tU' \partial_x \end{pmatrix}$ and V denotes the velocity in these new coordinates.

Then there exists a time-dependent family of operators $A(t)$ such that

$$c \|W(t)\|_{L^2} \leq \|A(t) W(t)\|_{L^2} \leq \|W(t)\|_{L^2}.$$

Furthermore, there exists an absolute constant $C > 0$ such that

$$\frac{d}{dt} \|AW\|_{L^2}^2 \leq -C \left(\|\sqrt{\mu} \partial_x W\|_{L^2}^2 + \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} W\|_{L^2}^2 + \|\sqrt{\mu} (\partial_y - tU' \partial_x) W\|_{L^2}^2 \right).$$

We recall the smallness assumptions (8) on $\frac{\mu'}{\mu} = -\frac{U''}{U'}$:

$$\left| \frac{\partial_y U'}{U'} \right| + \left| \partial_y \frac{\partial_y U'}{U'} \right| = \left| \frac{\partial_y \mu}{\mu} \right| + \left| \partial_y \frac{\partial_y \mu}{\mu} \right| < 0.0001. \tag{35}$$

This quantifies the requirement that μ may only change gradually, but since \mathbb{R} is unbounded μ may change by many orders of magnitude over all. However, this constraint on the relative rate of change then further implies that when restricted to any interval I of suitable size, it holds that

$$\frac{\max_I \mu}{\min_I \mu} \leq 2.$$

Thus, if we extend the restrictions $\mu|_I, U|_I$ by constants to functions μ_I, U_I on all of \mathbb{R} , then these extensions satisfy the assumptions (19): $\frac{\sup(\mu_I)}{\inf(\mu_I)} \leq 2$ of section 3. Thus we may ‘locally’ reduce to that model setting. However, these restrictions and extensions have to be related to the actual whole space problem (22) (see lemma 4.4) and have to be combined to control growth of the whole space problem (see lemmas 4.5–4.7).

Our main challenges in the following are to formalize this intuition and to control non-local errors. More precisely, since the velocity is non-local and so are several commutator terms, it is not possible to just restrict W and reduce estimates to the ones of section 3. Instead we will show that in the sum over all localized estimates still holds.

The following lemma establishes the existence of a partition of \mathbb{R} such that on each interval of the partition μ (and hence U') is comparable to a constant. Furthermore, the sizes of these intervals is bounded from below and hence cut-off functions and partitions of unity corresponding to this partition have controlled $W^{k,\infty}$ norms. Using these partitions we may also construct extensions of the restrictions of μ, U which satisfy the assumptions of the model setting studied in section 3.

Lemma 4.2 (Partitions). *Let $\mu = \mu(y) \in C^2(\mathbb{R}; \mathbb{R}^+)$, $U = U(y) \in C^3(\mathbb{R}; \mathbb{R})$ be as in theorem 1.1 and, in particular, assume that*

$$\|(\ln \mu)'\|_{W^{1,\infty}} < 0.0001. \tag{36}$$

Then there exists a partition $(I_j)_{j \in \mathbb{Z}}$, $I_j = [y_j, y_{j+1})$ of \mathbb{R} into intervals such that

$$\frac{\sup_{3I_j} \mu}{\inf_{3I_j} \mu} \leq 1.5 \tag{37}$$

for all j , where $3I_j$ denotes the rescaled intervals with the same center. Furthermore, the length of each interval I_j is bounded from below by 1000.

Associated with this partition there exists a family of non-negative functions $\chi_j \in C_c^\infty$ with $\text{supp}(\chi_j^2) \subset 3I_j$ such that χ_j^2 is a partition of unity and all the derivatives are uniformly small:

$$\|\partial_y \chi_j\|_{W^{1,\infty}} \leq 0.001, \tag{38}$$

and all higher derivatives are bounded uniformly in j .

Moreover, for each j there exist $\mu_j \in C^2(\mathbb{R})$, $U_j \in C^3(\mathbb{R})$ such that

$$\begin{aligned} \mu_j &= \mu, U_j = U \text{ in } I_j, \\ \mu_j U_j' &= \text{const. in } \mathbb{R} \end{aligned}$$

and so that μ_j and U_j' are constant outside $3I_j$ and

$$\frac{\max_{\mathbb{R}} \mu_j}{\min_{\mathbb{R}} \mu_j} \leq 2.$$

If furthermore $\mu = \mu(y) \in C^{N+2}$, $U = U(y) \in C^{N+3}$ for some $N \geq 1$ then $\mu_j \in C^{N+2}$, $U_j \in C^{N+3}$.

As an example, if $\mu(y) = \mu_0 e^{\delta y}$, then

$$I_j = \frac{1}{\delta} [j, j + 1)$$

is of size δ^{-1} and hence $|\partial_y \chi_j| = \mathcal{O}(\delta)$ is small.

Proof of lemma 4.2. In the case that μ is monotone, we may simply define $3I_j$ as the preimages of $(1.5^j, 1.5^{j+1})$ if $U'(y_j) \geq 1$. Since $\log(\mu)(y)$ is slowly varying by assumption (8), it follows that the size of I_j is bounded below by a large constant, say 1000, and hence our partition covers all of \mathbb{R} . And if $U'(y_j) < 1$, then we rescale the preimage by $(U'(y_j))^{-1}$.

More generally, we can define I_j greedily. That is, given a point y_j we define y_{j+1} and

$$I_j = [y_j, y_{j+1})$$

as the largest interval such that

$$\frac{1}{1.5} \mu(y_j) \leq \mu(y) \leq 1.5 \mu(y_j), \quad \forall y \in I_j.$$

In case this definition yields an unbounded interval, we may truncate it.

It is a classical result that given such a partition into intervals, there exists a partition of unity for which the square root of each function is still smooth and such that bounds on C^k norms are uniform in j (since they only depend on the size of each I_j , which is bounded from below).

Furthermore, given this partition of unity, we may construct μ_j as agreeing with μ on I_j and constant when a distance away from this support. The associated shear profile U_j is then constructed by integrating

$$\partial_y U_j := \frac{\text{const.}}{\mu_j}$$

with the constant of integration chosen such that $U_j(y) = U(y)$ in I_j .

This then directly implies the desired bounds, where we used that the derivatives of the partition of unity are bounded and hence the estimate (37) only possibly deteriorates by a small factor under this extension. \square

We remark that by construction for each j it holds that U'_j is comparable to a constant $U'_j(y_j)$. Therefore we may introduce the localized change of coordinates

$$y \mapsto z_j := \frac{U_j(y)}{u_j} \text{ where } u_j := \inf U'_j,$$

which is globally bilipschitz with constants between $\frac{1}{2}$ and 1. We note that the shear flow

$$U_j(y) = u_j z_j$$

is affine in these coordinates. Furthermore, for any $j \in \mathbb{Z}$ and any function $f \in C_c^\infty(\mathbb{T} \times \mathbb{R})$ it holds that

$$\frac{1}{2} \|f\|_{L^2(adxdz_j)} \leq \|f\|_{L^2(adxdy)} \leq 2 \|f\|_{L^2(adxdz_j)}. \tag{39}$$

Given these partitions we may naturally define operators acting on $\chi_j W$ by using the results of section 3.

Definition 4.3 (Localized Fourier weights). Let χ_j^2 be the partition of unity of lemma 4.2 and let μ_j, U_j be the collection of viscosities and shear associated with these partitions.

We then define A_j to be the operator as given in definition 3.2 for μ, U replaced by μ_j, U_j . Furthermore, we define

$$W_j(t, x, z_j) := \chi_j(y) W(t, x, y)$$

and the energy functional

$$E(t) = \sum_j \langle A_j W_j, A_j W_j \rangle_{L^2(dz_j)}. \tag{40}$$

In case of a non-constant weight a as in theorem 1.1 we instead consider

$$E(t) = \sum_j a_j \langle A_j W_j, A_j W_j \rangle_{L^2(dz_j)}.$$

with $a_j = \inf_{I_j} a \geq \frac{1}{2} \sup_{I_j} a$.

We remark that here for each interval I_j we consider the L^2 inner product with respect to z_j so that A_j indeed is a Fourier multiplier. However, by (39) each L^2 norm is comparable to the one for $L^2(dy)$. Hence, we may transparently switch between these spaces in our estimates and for simplicity of notation do not explicitly note the j dependence of our inner products.

Since χ_j^2 is a partition of unity, the norms of W and the sum of the norms of W_j are comparable.

Lemma 4.4 (Norm estimates). *Let χ_j and W_j be as in definition 4.3.*

Then there exist constants $0 < c_1 < c_2 < \infty$ such that the L^2 norms satisfy

$$c_1 \|W\|^2 \leq \sum_j \|W_j\|^2 \leq c_2 \|W\|^2.$$

Moreover, for any $N \in \mathbb{N}$ there exist constants d_0, \dots, d_N with $d_N = 1$ and c_1, c_2 such that

$$c_1 \|W\|_{H^N}^2 \leq \sum_{l=0}^N \sum_j \sum_{|\alpha|=l} d_l \|\partial_y^\alpha W_j\|^2 \leq c_2 \|W\|_{H^N}^2.$$

We remark that here $\|W\|_{H^N}$ consists of norms $\|\partial_y^k W\|_{L^2(ad_y)}$, where a is a slowly varying weight. Moreover, using our assumption that μ and hence U' is slowly varying, in the above estimates we may freely replace ∂_y by ∂_{z_j} for any j .

Proof of lemma 4.4. Since χ_j^2 is a partition of unity,

$$W^2 = W \sum_j \chi_j^2 W = \sum_j W_j^2$$

and hence the estimate for $N = 0$ trivially holds true with $c_1 = c_2 = 1$ and equality; also for arbitrary weighted L^2 norms.

For the case $N \geq 1$ we instead need to exploit the size of the derivative of χ_j and of $\ln(a)$. For instance, we may express

$$\begin{aligned} \int a |\partial_y W|^2 &= \sum_j \int a \partial_y W \partial_y (\chi_j W_j) = \sum_j \int a \partial_y W \chi_j \partial_y W_j + a \partial_y W \chi_j' W_j \\ &= \sum_j \int a |\partial_y W_j|^2 - a \chi_j' W \partial_y W_j + a \partial_y W \chi_j' W_j. \end{aligned}$$

Since $\|\chi_j'\|_{W^{1,\infty}}$ is bounded (and even small) we may use Young's inequality to absorb the last two terms as an error, provided d_0 is large enough.

For $N > 1$ we argue by induction. More precisely, for any given index α we may expand

$$\partial_y^\alpha W_j = \partial_y^\alpha (\chi_j W) = \chi_j \partial_y^\alpha W + \sum_{\beta+\gamma=\alpha} (\partial_y^\beta \chi_j) \partial_y^\gamma W.$$

By the same argument as in the L^2 case it holds that

$$\sum_j \|\chi_j \partial_y^\alpha W\|_{L^2}^2 = \|\partial_y^\alpha W\|_{L^2}^2.$$

For all other terms we note that by (38) and the smallness condition (36)

$$|\partial_y^\beta \chi_j| < \infty$$

is bounded uniformly in j and is supported in the same bounded region as χ_j . Since the supports of the functions χ_j cover \mathbb{R} at most twice, we thus may use Hölder’s inequality to control

$$\sum_{\beta+\gamma=\alpha} \|(\partial_y^\beta \chi_j) \partial_y^\gamma W\|_{L^2}^2 \leq C_\alpha \sum_{m < N} \|W\|_{H^m}^2,$$

which can be controlled in terms of

$$\sum_{l=0}^{N-1} \sum_j \sum_{|\alpha|=l} d_l \|\partial^\alpha W_j\|^2$$

by the induction assumption. □

Given this definition of an energy (40), we next need to verify that it indeed is a Lyapunov functional and thus study (according to the vorticity equation (22))

$$\begin{aligned} \frac{d}{dt} E/2 &= \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_j \langle A_j W_j, A_j \chi_j \partial_t W \rangle \\ &= \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_{jj'} \langle A_j W_j, A_j \chi_j (\operatorname{div}_t (\mu \nabla_t \chi_{j'} W_{j'})) \rangle \\ &\quad - \sum_j \langle A_j W_j, A_j \chi_j (\operatorname{div}_t (\mu' \nabla_t V_1) + \mu'' \partial_x V_2 - U'' V_2) \rangle. \end{aligned} \tag{41}$$

Compared to the results of section 3 we here encounter several additional challenges:

- The Biot–Savart law is *non-local*. Therefore $\chi_j V$ depends on all $(W_{j'})_{j'}$, not just W_j . We thus need to compare various localizations of the Biot–Savart law, while at the same time also localizing in frequency.
- The evolution of W_j hence also depends on all $(W_{j'})_{j'}$.
- In the dissipation term we have a double sum with respect to j and j' . Here we observe that for $|j - j'| \geq 2$ the support of χ_j and $\chi_{j'}$ are disjoint and hence we only need to consider $j' \in \{j - 1, j, j + 1\}$ (only neighbours instead of full non-local interaction as for the velocity). However, the coupling introduced by this interaction implies that we cannot hope to control $\langle A_j W_j, A_j W_j \rangle$ in terms of itself, but rather have to control sums over all j .

The following lemma generalizes lemma 3.4 to the present setting.

Lemma 4.5 (Localized dissipation estimates). *Let $W \in \mathcal{S}$, then it holds that*

$$\begin{aligned}
 & 0.2 \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_j \langle A_j W_j, A_j \chi_j (\operatorname{div}_t (\mu \nabla_t W)) \rangle \\
 & \leq -0.001 \sum_j \left(\|\sqrt{\mu_j} U_j' \left(\frac{1}{u_j} \partial_{z_j} - t \partial_x \right) W_j\|^2 \right. \\
 & \quad \left. + \|\left(\mu_j (U_j')^2 \right)^{1/6} (-\partial_x^2)^{1/6} W_j\|^2 + \|\sqrt{\mu_j} \partial_x W_j\|^2 \right). \tag{42}
 \end{aligned}$$

Proof of lemma 4.5. We note that in (42) the dissipation involves W and not just W_j and we thus have to control the interaction with other intervals. However, by construction only neighbouring functions $\chi_j, \chi_{j'}$ with $j' \in \{j-1, j, j+1\}$ have intersections of their supports.

We thus expand

$$\chi_j (\operatorname{div}_t (\mu \nabla_t W)) = \operatorname{div}_t (\mu \nabla_t W_j) + [\operatorname{div}_t \mu \nabla_t, \chi_j] \sum_{j' \in \{j-1, j, j+1\}} \chi_{j'} W_{j'}.$$

Here the ‘diagonal term’

$$\langle A_j W_j, A_j \operatorname{div}_t (\mu \nabla_t W_j) \rangle$$

can be controlled by using lemma 3.4 of section 3.

For the other terms we note that

$$\begin{aligned}
 [\operatorname{div}_t \mu \nabla_t, \chi_j] &= [(\partial_y - tU' \partial_x) \mu (\partial_y - tU' \partial_x), \chi_j] \\
 &= (\mu \chi_j')' + (\mu \chi_j' + (\mu \chi_j)') (\partial_y - tU' \partial_x)
 \end{aligned}$$

where $(\partial_y - tU' \partial_x) = U' \left(\frac{U_j'}{U'} \frac{1}{u_j} \partial_{z_j} - t \partial_x \right)$ is a first order differential operator and U_j' is equal to U' on the support of χ_j . We can use Young’s inequality to absorb these terms into the dissipation, by the slow variations (36) and (38) in μ and χ_j . \square

Similarly, lemma 3.6 is generalized as follows.

Lemma 4.6 (Non-local velocity estimates). *Let $W \in \mathcal{S}$, then it holds that*

$$\sum_j 0.2 \langle A_j W_j, \dot{A}_j W_j \rangle + \sum_j \langle A_j W_j, A_j \chi_j (U'' V_2) \rangle \leq 0.$$

Proof of lemma 4.6. We observe that unlike in lemma 4.5, here $\chi_j V_2$ depends on $W_{j'}$ for all j' and not just $j' \in \{j-1, j, j+1\}$.

Instead of estimating in terms of j' as in lemma 4.5, we generalize the elliptic estimates of [CZZ19] to the present setting.

More precisely, let ϕ_j be the stream function generated by W_j :

$$\Delta_j \phi_j = W_j = \chi_j W, \text{ with } \Delta_j = \partial_x^2 + (\partial_{z_j} - t u_j \partial_x)^2,$$

and let ϕ denote the stream function generated by W :

$$\Delta_t \phi = W = \sum_j \chi_j W_j, \text{ with } \Delta_t = \partial_x^2 + (\partial_y - tU' \partial_x)^2.$$

Then by testing the above equations with $-\phi_j$ and $-\phi$, respectively, we observe that

$$\begin{aligned} \|\nabla_j \phi_j\|^2 &\leq \langle \nabla_t \phi, \nabla_j(\chi_j \phi_j) \rangle \\ &\leq \|\nabla_t \phi\|_{L^2(\text{supp}(\chi_j))} (\|\nabla_j \phi_j\| + \|\chi_j'\|_{L^\infty} \|\phi_j\|) \end{aligned}$$

and

$$\begin{aligned} \|\nabla_t \phi\|^2 &\leq \sum_j \langle \nabla_t(\chi_j \phi), \nabla_j \phi_j \rangle \\ &\leq \sqrt{\sum_j \|\nabla_j \phi_j\|^2} \sqrt{\sum_j \|\nabla_t(\chi_j \phi)\|^2}, \end{aligned}$$

where we used that ∇_t and ∇_j agree on the support of χ_j . Using the fact that derivatives of χ_j are bounded, it thus follows that (recalling Poincaré’s inequality for functions with zero x -average)

$$\|\nabla_t \phi\|^2 \approx \sum_j \|\nabla_j \phi_j\|^2.$$

Thus errors in velocity can be controlled in terms of sums of $\nabla_j \phi_j$ (see also lemma 4.4). In order to conclude, we note that by the definition of U_j, μ_j and W_j each such contribution can be controlled in terms of the decay of the multiplier A_j and the dissipation. Hence the velocity errors can be absorbed.

We remark that this proof also immediately extends to the case of $L^2(adxdy)$ for a slowly varying. That is, testing the equation

$$\Delta_t \phi = W$$

with $-a\phi$ instead, we obtain

$$\int a |\nabla_t \phi|^2 + \frac{a'}{a} a \phi (\partial_y - tU' \partial_x) \phi.$$

Since $|\frac{a'}{a}| \ll 1$, we may use Young’s inequality and the Poincaré inequality to show that this energy remains positive definite and comparable to

$$\int a |\nabla_t \phi|^2$$

for any choice of a .

□

Lemma 4.7 (Viscosity errors). *Let $W \in \mathcal{S}$, then it holds that*

$$\begin{aligned} &0.2 \sum_j \langle A_j W_j, \dot{A}_j W_j \rangle - \sum_j \langle A_j W_j, A_j \chi_j (\text{div}_t(\mu' \nabla_t) V_1 + \mu'' \partial_x V_2) \rangle \\ &\leq 0.0005 \sum_j \left(\|\mu_j (U_j')^2\|^{1/6} (-\partial_x^2)^{1/6} W_j \right)^2 \\ &\quad + \|\sqrt{\mu_j} \partial_x W_j\|^2 + \|\sqrt{\mu_j} U_j' \left(\frac{1}{u_j} \partial_{z_j} - t \partial_x \right) W_j\|^2. \end{aligned}$$

Proof of lemma 4.7. In order to prove these estimates we employ a combination of the methods used in the proofs of lemmas 3.5, 4.5 and 4.6.

More precisely, we first use the structure of the Biot–Savart law to express

$$(\operatorname{div}_t(\mu' \nabla_t) V_1 + \mu'' \partial_x V_2)$$

in terms of W and lower order terms. For the terms involving W we can then argue analogously as in lemma 3.5, using the decoupling of χ_j and $\chi_{j'}$ if j and j' are far apart as in lemma 4.5.

Finally, for the remaining terms involving the velocity, we argue as in lemma 4.6 and thus reduce to estimating $\nabla_j \phi_j$ in place of V . Summing over the ‘diagonal’ estimates as established in lemma 3.5 then concludes the proof. \square

Having established these estimates, we are now ready to prove proposition 4.1 and thus also prove part of theorem 1.1. An extension of these results to higher Sobolev norms H^N is given in section 5, which then completes the proof of theorem 1.1.

Proof of proposition 4.1. Let $\omega_0 \in L^2(\text{ady})$ be a given initial datum with zero x -average, let μ, U, a satisfy the assumptions of theorem 1.1 and let W denote the solution of (34) with this initial data.

Then by lemma 4.2 there exists a partition of \mathbb{R} into intervals I_j and an associated partition of unity χ_j^2 . We then define A_j and $W_j := \chi_j W$ as in definition 4.3, and study the evolution of the energy $E(t) := \sum_j a_j \langle A_j W_j, A_j W_j \rangle$.

Inserting the estimates derived in lemmas 4.4–4.7 to the time derivative of E in (41), we deduce that

$$\begin{aligned} \frac{d}{dt} E(t) \leq & -10^{-5} \sum_j a_j \left(\left\| \left(\mu_j (U_j')^2 \right)^{1/6} (-\partial_x^2)^{1/6} W_j \right\|^2 + \left\| \sqrt{\mu_j} \partial_x W_j \right\|^2 \right. \\ & \left. + \left\| \sqrt{\mu_j} U_j' \left(\frac{1}{u_j} \partial_{z_j} - t \partial_x \right) W_j \right\|^2 \right). \end{aligned}$$

Finally, by lemma 4.4 the energy $E(t)$ is comparable to $\|W(t)\|_{L^2(\text{ady})}^2$. This hence concludes the proof of proposition 4.1 where the symmetric operator A is defined such that

$$\|A(t) W(t)\|^2 := E(t).$$

\square

5. Stability in H^N

As the last step of our proof of theorem 1.1, in this section we extend the stability and damping estimates in L^2 established in section 4 to estimates in H^N . Here we follow an inductive approach introduced in [Zil21] in the inviscid setting. We consider the linearized equation (34)

$$\begin{aligned} \partial_t W &= U'' V_2 + \operatorname{div}_t(\mu \nabla_t W) - \operatorname{div}_t(\mu' \nabla_t V_1) - \mu'' \partial_x V_2 =: LW, \\ V_1 &= \frac{-(\partial_y - tU' \partial_x)}{\partial_x^2 + (\partial_y - tU' \partial_x)^2} W, \quad V_2 = \frac{\partial_x}{\partial_x^2 + (\partial_y - tU' \partial_x)^2} W, \end{aligned} \tag{43}$$

where we introduced the time-dependent linear operator L for brevity of notation. We remark that derivatives with respect to x can be identified with multiplication by ik , since the linearized equations decouple with respect to k . Hence higher derivatives in x can be estimated using the

L^2 energy. In the following we hence only consider derivatives with respect to y . Applying N derivatives to (43) we obtain that

$$\partial_t \partial_y^N W = L \partial_y^N W + [L, \partial_y^N] W. \tag{44}$$

In the following lemma we will show that the commutator term can be considered as an error term involving fewer than N derivatives, while $L \partial_y^N W$ can be treated in the same way as in the L^2 estimate. In this sense the L^2 estimate in section 4 forms the core of our argument.

Proposition 5.1. *Let μ, U, a satisfy the assumptions of theorem 1.1. In particular, let $N \in \mathbb{N}$ and suppose that $\partial_y \ln(\mu) \in W^{N+1, \infty}$. Let A be as in proposition 4.1, then there exist constants $c_0, c_1, \dots, c_N > 0$ depending only on the $W^{N+1, \infty}$ norms of $\partial_y \ln(\mu)$ such that*

$$E_N(t) = \sum_{l \leq N} c_l \langle A \partial_y^l W, A \partial_y^l W \rangle$$

is a Lyapunov functional and satisfies for some positive constant $C > 0$

$$\frac{d}{dt} E_N(t) \leq -C \sum_{l \leq N} \left\| \left(\sqrt{\mu} \partial_x \partial_y^l W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^l W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} \partial_y^l W \right) \right\|_{L^2}^2.$$

We remark that here we only require that the $W^{N+1, \infty}$ norm of $\partial_y \ln \mu$ is finite. Only the $W^{1, \infty}$ norm needs to be small in order to establish the L^2 stability estimate.

Proof of proposition 5.1. The case $N = 0$ has been established in proposition 4.1 with $c_0 = 1$.

We aim to proceed by induction. Hence, suppose that the estimates have been established for the case $N - 1$ and consider

$$E_N(t) = c_N \langle A \partial_y^N W, A \partial_y^N W \rangle + E_{N-1}(t)$$

with c_N to be determined later.

Then by the induction assumption it holds that

$$\begin{aligned} \frac{d}{dt} E_{N-1}(t) &\leq -C \sum_{l \leq N-1} \left\| \left(\sqrt{\mu} \partial_x \partial_y^l W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^l W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} \partial_y^l W \right) \right\|_{L^2}^2. \end{aligned} \tag{45}$$

In particular, all derivatives of W up to order $N - 1$ can be controlled by the induction assumption. We thus turn to the control of the ‘leading order’ term involving $\partial_y^N W$. Here, by the L^2 estimates of proposition 4.1 it holds that

$$\begin{aligned} \frac{d}{dt} c_N \langle A \partial_y^N W, A \partial_y^N W \rangle &= 2c_N \langle \dot{A} \partial_y^N W, A \partial_y^N W \rangle + 2c_N \langle A \partial_y^N W, A L \partial_y^N W \rangle + 2c_N \langle A \partial_y^N W, A [L, \partial_y^N] W \rangle \\ &\leq -10^{-5} c_N \left\| \left(\sqrt{\mu} \partial_x \partial_y^N W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^N W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} \partial_y^N W \right) \right\|_{L^2}^2 \\ &\quad + 2c_N \langle A \partial_y^N W, A [L, \partial_y^N] W \rangle. \end{aligned} \tag{46}$$

Combining the estimates (46) and (45) it thus suffices to show that for a suitable choice of c_N we may absorb the commutation error

$$2c_N \langle A \partial_y^N W, A [L, \partial_y^N] W \rangle.$$

into the decay in (45) and (46).

Let us first discuss the main dissipation term of L . Here we may iteratively expand

$$\begin{aligned} [\operatorname{div}_t(\mu \nabla_t), \partial_y^N] W &= [\operatorname{div}_t(\mu \nabla_t), \partial_y] \partial_y^{N-1} W \\ &\quad + [\operatorname{div}_t(\mu \nabla_t), \partial_y^{N-1}] \partial_y W - [[\operatorname{div}_t(\mu \nabla_t), \partial_y^{N-1}], \partial_y] W \\ &= \sum_{l < N} B_l \partial_y^l W, \end{aligned}$$

where the operators B_l are second order elliptic operators whose coefficient functions may be explicitly computed in terms of derivatives of U' and μ up to order $N - l$. In order to estimate

$$\langle A \partial_y^N W, A [\operatorname{div}_t(\mu \nabla_t), \partial_y^N] W \rangle = \sum_{l < N} \langle A \partial_y^N W, A B_l \partial_y^l W \rangle$$

we may thus argue as in the proof of lemma 4.5 and control

$$c_N \sum_{l < N} \langle A \partial_y^N W, A B_l \partial_y^l W \rangle \leq c_N \|\sqrt{\mu} \nabla_t \partial_y^N W\|_{L^2} \sum_{l < N} d_l \|\sqrt{\mu} \nabla_t \partial_y^l W\|_{L^2}. \tag{47}$$

Similarly we may iterative expand the equation satisfied by derivatives of the stream function

$$\Delta_t \partial_y^N \phi = \partial_y^N W + [\Delta_t, \partial_y^N] \phi$$

and thus obtain that

$$\partial_y^N \phi = \Delta_t^{-1} \partial_y^N W + \Delta_t^{-1} \sum_{l < N} \tilde{B}_l \Delta_t^{-1} \partial_y^l W,$$

where the second order operators \tilde{B}_l may again be explicitly computed. Thus, we may argue as in the proofs of lemmas 4.6 and 4.7 and again use Hölder's and Young's inequality to control

$$\begin{aligned} &\langle A \partial_y^N W, A [\operatorname{div}_t(\mu' \nabla_t) (\partial_y - tU' \partial_x) \Delta_t^{-1} + \mu'' \partial_x \Delta_t^{-1} - U'' \partial_x \Delta_t^{-1}, \partial_y^N] W \rangle \\ &\leq \left\| \left(\sqrt{\mu} \partial_x \partial_y^N W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^N W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} \partial_y^N W \right) \right\|_{L^2} \\ &\quad \times \sum_{l < N} \left\| \left(\sqrt{\mu} \partial_x \partial_y^l W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^l W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{1/6} \partial_y^l W \right) \right\|_{L^2}. \tag{48} \end{aligned}$$

We may thus conclude our estimate by using Young's inequality. More precisely, we first apply Young's inequality to the estimates (47) and (48) so that the contributions due to $\partial_y^N W$ can be bounded by

$$10^{-6} c_N \left\| \left(\sqrt{\mu} \partial_x \partial_y^N W, \sqrt{\mu} (\partial_y - tU' \partial_x) \partial_y^N W, \left(\mu (U')^2 \right)^{\frac{1}{6}} (-\partial_x^2)^{\frac{1}{6}} \partial_y^N W \right) \right\|_{L^2}$$

and can thus be absorbed into the decay in estimate (46). Then, choosing c_N sufficiently small the remaining terms obtained in the application of Young's inequality can be absorbed into the decay by (45). This concludes the proof. \square

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 258734477-SFB 1173.

ORCID iD

Christian Zillinger  <https://orcid.org/0000-0002-0480-2719>

References

- [BG81] Barker S J and Gile D 1981 Experiments on heat-stabilized laminar boundary layers in water *J. Fluid Mech.* **104** 139–58
- [BGM17] Bedrossian J, Germain P and Masmoudi N 2017 On the stability threshold for the 3d Couette flow in Sobolev regularity *Ann. Math.* **185** 541–608
- [BM15] Bedrossian J and Masmoudi N 2015 Inviscid damping and the asymptotic stability of planar shear flows in the 2d Euler equations *Publ. Math. l'IHÉS* **122** 195–300
- [BMV16] Bedrossian J, Masmoudi N and Vicol V 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** 1087–159
- [BVW18] Bedrossian J, Vicol V and Wang F 2018 The Sobolev stability threshold for 2D shear flows near Couette *J. Nonlinear Sci.* **28** 2051–75
- [Cra69] Craik A D D 1969 The stability of plane Couette flow with viscosity stratification *J. Fluid Mech.* **36** 685–93
- [CZZ19] Coti Zelati M and Zillinger C 2019 On degenerate circular and shear flows: the point vortex and power law circular flows *Commun. PDE* **44** 110–55
- [Dra02] Drazin P G 2002 *Introduction to Hydrodynamic Stability* vol 32 (Cambridge University Press)
- [DWZ21] Deng W, Wu J and Zhang P 2021 Stability of Couette flow for 2D Boussinesq system with vertical dissipation *J. Funct. Anal.* **281** 109255
- [EW15] Elgindi T M and Widmayer K 2015 Sharp decay estimates for an anisotropic linear semigroup and applications to the surface quasi-geostrophic and inviscid Boussinesq systems *SIAM J. Math. Anal.* **47** 4672–84
- [GS14] Govindarajan R and Chandra Sahu K 2014 Instabilities in viscosity-stratified flow *Annu. Rev. Fluid Mech.* **46** 331–53
- [HB87] Hooper A P and Boyd W G C 1987 Shear-flow instability due to a wall and a viscosity discontinuity at the interface *J. Fluid Mech.* **179** 201–25
- [Hei85] Heisenberg W 1985 Über Stabilität und Turbulenz von Flüssigkeitsströmen *Original Scientific Papers Wissenschaftliche Originalarbeiten* (Springer) pp 31–81
- [HL20] He Z and Liao X 2020 Solvability of the two-dimensional stationary incompressible inhomogeneous Navier–Stokes equations with variable viscosity coefficient (arXiv:2005.13277)
- [IJ20a] Ionescu A D and Jia H 2020 Inviscid damping near the Couette flow in a channel *Commun. Math. Phys.* **374** 2015–96

- [IJ20b] Ionescu A D and Jia H 2020 Nonlinear inviscid damping near monotonic shear flows (arXiv:2001.03087)
- [Jia20] Jia H 2020 Linear inviscid damping in Gevrey spaces *Arch. Ration. Mech. Anal.* **235** 1327–55
- [JRR84] Joseph D D, Renardy M and Renardy Y 1984 Instability of the flow of two immiscible liquids with different viscosities in a pipe *J. Fluid Mech.* **141** 309–17
- [Lin44] Lin C C 1944 On the stability of two-dimensional parallel flows *Proc. Natl Acad. Sci. USA* **30** 316
- [Lis20] Liss K 2020 On the Sobolev stability threshold of 3D Couette flow in a uniform magnetic field *Commun. Math. Phys.* **377** 1–50
- [LX19] Lin Z and Ming X 2019 Metastability of Kolmogorov flows and inviscid damping of shear flows *Arch. Ration. Mech. Anal.* **231** 1811–52
- [LZ11] Lin Z and Zeng C 2011 Inviscid dynamical structures near Couette flow *Arch. Ration. Mech. Anal.* **200** 1075–97
- [MSHZ20] Masmoudi N, Said-Houari B and Zhao W 2020 Stability of Couette flow for 2D Boussinesq system without thermal diffusivity (arXiv:2010.01612)
- [PV91] Power H and Villegas M 1991 Weakly nonlinear instability of the flow of two immiscible liquids with different viscosities in a pipe *Fluid Dyn. Res.* **7** 215
- [TW19] Tao L and Jiahong W 2019 The 2D Boussinesq equations with vertical dissipation and linear stability of shear flows *J. Differ. Equ.* **267** 1731–47
- [Wid18] Widmayer K 2018 Convergence to stratified flow for an inviscid 3D Boussinesq system *Commun. Math. Sci.* **16** 1713–28
- [WZZ18] Wei D, Zhang Z and Zhao W 2018 Linear inviscid damping for a class of monotone shear flow in Sobolev spaces *Commun. Pure Appl. Math.* **71** 617–87
- [WZZ20] Wei D, Zhang Z and Zhao W 2020 Linear inviscid damping and enhanced dissipation for the Kolmogorov flow *Adv. Math.* **362** 106963
- [Yih67] Yih C-S 1967 Instability due to viscosity stratification *J. Fluid Mech.* **27** 337–52
- [YL18] Yang J and Lin Z 2018 Linear inviscid damping for Couette flow in stratified fluid *J. Math. Fluid Mech.* **20** 445–72
- [Zil17] Zillinger C 2017 Linear inviscid damping for monotone shear flows *Trans. Am. Math. Soc.* **369** 8799–855
- [Zil21] Zillinger C 2021 Linear inviscid damping in Sobolev and Gevrey spaces *Nonlinear Anal.* **213** 112492