On variable viscosity and enhanced dissipation
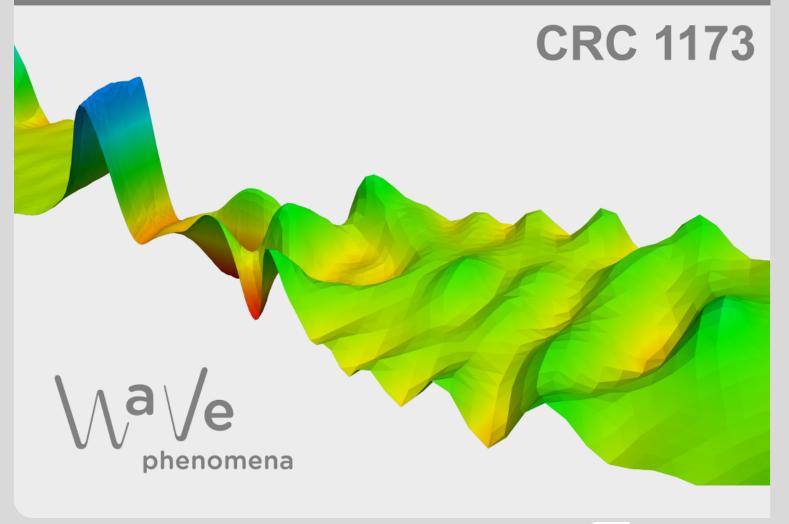
Xian Liao, Christian Zillinger

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173

Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

TU
WIEN

Funded by

DFG

ON VARIABLE VISCOSITY AND ENHANCED DISSIPATION

XIAN LIAO AND CHRISTIAN ZILLINGER

ABSTRACT. In this article we consider the 2D 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. Moreover it turns out that the shear flow overcompensates for weakening viscosity: decreasing viscosity leads to stronger enhanced dissipation and increasing viscosity leads to weaker dissipation than in the constant viscosity case.

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

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

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

The symmetric part of the velocity gradient

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

denotes the symmetric deformation tensor. The viscosity coefficient

$$\mu = \mu(x, y)$$

is a given non-constant positive scalar function. More precisely, we consider the case of stratified viscosity $\mu(y)$ 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 constant interests of physicists [Cra69, GS14, Hei85, Lin44, HB87, Yih67]. While additional dissipation at first sight suggest stabilization [1], in experiments viscosity exhibits dual roles [Dra02, Chapter 8, pp. 160]: a stabilizing role due to the dissipation of energy and a more subtle destabilizing role. Yih [Yih67] showed that the instability in a low Reynolds number flow can be caused by viscosity stratifications

---

Date: October 22, 2021.

[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.

(see also Craik [Cra69] for the study of flows with continuous viscosity stratification). These results motivated decades of active researches on the instability caused by viscosity interfaces, see [GS14] for a review paper on this topic.

In this paper 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 behavior of perturbations to the shear flow solutions

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

which satisfies the hydrostatic balance

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

This condition implies that as μ decreases $\partial_y U$ increases and vice versa.

As an additional assumption, while we allow μ and hence $\partial_y U$ to change by several orders of magnitude, we require that locally there is not too much oscillation:

$$(4) \qquad\qquad \left\| \partial_y \ln(\mu(y)) \right\|_{W^{1,\infty}} \leq 0.001.$$

Thus, for instance μ may decrease exponentially but only with a small exponent c.

As a consequence of the balance relation (11) one 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 [2]. More precisely, as we discuss following Theorem 1.1 as the viscosity *decreases* towards zero, the effective dissipation $\mu(\partial_y U)^2$ rate becomes *larger*. This also helps to explains 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].

In recent years there has been extensive research on the stability study of the shear flows (10) for the inviscid fluids with

$$\mu = 0,$$

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 reader to the following recent works for further discussion [Jia19, LX19, WZZ18, IJ18, Wid18, YL18, BMV16, EW15, LZ11, WZZ17, BGM17, BM15, DWZ20, LZ11]. We, in particular, recall that for linearized equations around Couette flow

$$\mu = \text{const.} > 0, U(y) = y,$$

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

$$\exp(-C \sqrt[3]{\mu} t),$$

---

[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).

and thus on a time scale $\mu^{-\frac{1}{3}}$ much smaller than the dissipation time scale μ^{-1}. This phenomenon is hence called *enhanced dissipation* (see [BVW18] for further discussion and the analysis of the nonlinear problem).

Given a particular value of the viscosity at a given point, $\mu(y_0)$, in this paper we are interested in the change of the (local, effective) dissipation rates if μ varies as $y > y_0$ increases. For instance, how much of an increase or decrease of μ is required to change the dissipation rate by a factor of 10? As our main results we establish stability of the linearized equations and prove damping with a local rate $\mu(U')^2$, which is *inversely* proportional to $\sqrt[3]{\mu}$.

Theorem 1.1. *Let $\mu \in C^2(\mathbb{R})$ with $\mu > 0$ be a given stratified viscosity profile. Then a stationary solution of the Navier-Stokes equations (1) on $\mathbb{T} \times \mathbb{R}$ in the presence of the variable viscosity μ is given by a shear flow $v = (U(y), 0)^T$ such that*

$$(5) \qquad \mu \partial_y U = const.$$

and the linearized equations around this solution in vorticity formulation read

$$(6) \qquad \begin{aligned} \partial_t \omega + U \partial_x \omega &= U'' v_2 + \mathrm{div}(\mu \nabla \omega) - \mathrm{div}(\mu' \nabla v_1) - \mu'' v_2, \\ v &= \nabla^\perp (-\Delta)^{-1} \omega, \end{aligned}$$

where $U' = \partial_y U$ and $U'' = \partial_y^2 U$ denote y derivatives.

Additionally suppose that μ only varies gradually, in the sense that

$$(7) \qquad \begin{aligned} \|\frac{\mu'}{\mu}\|_{L^\infty} + \|U' \partial_y \frac{\mu'}{\mu}\|_{L^\infty} &< 0.001, \\ \frac{1}{\sup \mu} \|\mathcal{F}(\mu')\|_{L^1(\mathbb{R} \setminus (-\sup(\mu)^{-1/3}, \sup(\mu)^{-1/3}))} &< 0.001, \end{aligned}$$

where $\mathcal{F}(\mu)$ denotes the (tempered) Fourier transform and that μ is bounded above and below and sufficiently small so that

$$(8) \qquad \frac{(\sup(\mu))^2}{\inf(\mu)} < 0.1.$$

For instance, μ may grow at a (small) exponential rate from a value ν^2 to 0.1ν with $\nu < 0.1$. For simplicity of presentation also assume that $U' \geq 1$ (which can be assumed without loss of generality after rescaling). Then the linearized equations (6) are stable and exhibit enhanced dissipation. More precisely, there exists a time-dependent family of operators $A(t)$ with

$$0.1 \|\omega(t)\|_{L^2(U' dy)} \leq \|A(t)\omega(t)\|_{L^2(U' dy)} \leq \|\omega(t)\|_{L^2(U' dy)}.$$

Furthermore, if $\int \omega_0 dx = 0$ then for all times $t > 0$ it holds that

$$\frac{d}{dt} \|A(t)\omega(t)\|_{L^2(U' dy)}^2 \leq -0.001 (\|(\mu(U')^2))^{1/6} \omega\|^2 + \|\sqrt{U'} v\|^2).$$

Moreover, under further regularity assumptions these results also extend to stability of the "profile" $W(t, x, U(y)) := \omega(t, x - tU(y), y)$ in higher Sobolev norms H^N (see Proposition 5.1 for a precise statement).

Let us comment on these results:

- We remark that due to the balance relation (5) it holds that
$$\mu U' = \text{const.} =: \sigma.$$
Hence, one observes that the local dissipation rate satisfies
$$\sqrt[3]{\mu(U')^2} = \sqrt[3]{\mu(U')^2 \mu \frac{1}{\mu}} = \sqrt[3]{\sigma^2} \frac{1}{\sqrt[3]{\mu(y)}}$$
and thus is proportional to $\mu(y)^{-1/3}$. Thus a *decrease* of μ by a factor 1000 corresponds to an *increase* of the dissipation rate by a factor 10. Conversely, *increasing* the viscosity compared to $\mu(y_0)$ corresponds to *weaker* dissipation.
- Our assumptions on μ ensure that the effective dissipation rate $\mu(U')^2$ is always smaller than one. The dependence in terms of a third root hence reflects the enhancement of the mixing rate due to shear.
- The nonlinear constant viscosity Navier-Stokes problem 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 variable viscosity Boussinesq equations).
- 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. Hence, as in the inviscid setting [Zil17] we require a smallness condition to control the correction term $U''v_2$ in the linearized equation (this condition further allows us to control derivatives of the viscosity). This motivates our assumption (7) (see Section 3.1 for further discussion).
- We remark that in view of (5) the shear flow U is strictly monotone and hence invertible (but U' might be very large). In our analysis it will prove advantageous to thus equivalently consider the variable $z = U(y)$ and the profile W moving with the shear. Moreover, stability is most naturally phrased in terms of spaces $L^2(dz) \simeq L^2(U'dy)$.
- The first condition in (7) allows μ to grow exponentially but only with a small exponent. In particular, this implies that level sets of the form $\{y : 10^j < \mu(y) < 10^{j+1}\}$ for $j \in \mathbb{Z}$ are bounded below in size, which we exploit in a partitioning construction in Section 4. The second condition in (7) should be understood as a regularity assumption on the relative change $\frac{\mu}{\sup(\mu)}$, which should decay at very large frequencies (larger than $\sup(\mu)^{-1/3}$).

As we discuss in the following corollary the weighted dissipation estimate implies exponential decay, if ω remains suitably localized.

Corollary 1.2. *Let $W(t, x, U(y)) := \omega(t, x - tU(y), y)$ be as in Theorem 1.1, let $M \subset \mathbb{R}$ and suppose that on a given time interval $(t_1, t_2) \subset (0, \infty)$ a fraction at least $\theta \in (0, 1)$ of the L^2 energy is localized in M. That is,*
$$\|W(t)\|_{L^2(M)}^2 \geq \theta \|W(t)\|_{L^2}^2.$$

Then for all $t \in (t_1, t_2)$ it holds that

$$\|W(t)\|_{L^2}^2 \leq \exp(-0.001\theta \min_M \sqrt[3]{\mu(U')^2}(t - t_1))\|W(t_1)\|_{L^2}^2$$

Proof. By Theorem 1.1 it holds that

$$\partial_t \|W(t)\|_{L^2}^2 \leq -0.01\|cW\|_{L^2}^2$$

with $c := (\mu(U')^2)^{1/6}$. By assumption the left-hand-side can be bounded from above by

$$-0.001 \min_M c \ \theta \|AW(t)\|^2,$$

which yields the result. $\qquad\square$

We stress that the time interval considered in this corollary might be very small if M is a region with a very fast effective decay rate, since then the L^2 energy (or enstrophy) in that region can be expected to decay much faster than in other regions and hence will correspond to a much smaller fraction than θ after a time.

Based on the local dissipation rate at first sight one might also conjecture an estimate of the form

$$\| \exp(\sqrt[3]{\mu(U')^2}t)W(t)\|_{L^2} \leq C\|W(0)\|_{L^2}.$$

to hold. However, such an estimate cannot be expected to hold in general, since the Biot-Savart law is non-local and not decaying quickly enough. More precisely, if W is highly localized in a region M, then the velocity field generated by W exhibits decay away from M in terms of a power law of the distance $\text{dist}(y, M)$. Hence, supposing for the moment that W remains localized and that M is a region with small decay rate, we expect W to decay with slower rate. 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. This allows us to more clearly 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 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 $(U(y), 0)$. 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. For easier reference the equivalent formulations of the equations with respect to these coordinates are also collected in this section. Moreover, we define Sobolev spaces and multipliers with respect to z.

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 $\partial_y U$ may change by many orders of magnitude.

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

$$U(y) = \frac{C}{\mu(y)}.$$

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

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

$$
\begin{align}
(9) \qquad \partial_t \omega + U(y)\partial_x \omega - U'' v_2 &= \operatorname{div}(\mu \nabla \omega) - \operatorname{div}(\mu' \nabla v_1) - \mu'' \partial_x v_2, \\
v &= \nabla^\perp \Delta^{-1} \omega
\end{align}
$$

Proof of Lemma 2.1. Following Theorem 1.1 we make the ansatz

$$(10) \qquad \mu = \mu(y), \quad v = \begin{pmatrix} U(y) \\ 0 \end{pmatrix}.$$

The Navier-Stokes equations (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

$$(11) \qquad \partial_y(\mu \partial_y U) = C_0$$

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 $\mu = C_0 y + C_2$.
- 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 additionally assume that the effective damping rate $\mu(\partial_y U)^2$ is not large.

In the following let U, μ be solutions of (11) 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

$$\text{div}(\mu S v) = \begin{pmatrix} 2\partial_x \mu \partial_x v_1 & \partial_y \mu \partial_y v_1 + \partial_y \mu \partial_x v_2 \\ \partial_x \mu \partial_x v_1 + \partial_x \mu \partial_x v_2 & 2\partial_y \mu \partial_y v_2 \end{pmatrix},$$

$$v = \nabla^\perp \Delta^{-1} \omega = \begin{pmatrix} -\partial_y \Delta^{-1} \omega \\ \partial_x \Delta^{-1} \omega \end{pmatrix}.$$

We may calculate(see also [HL20])

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

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:

$$(12) \qquad \partial_t\omega + v \cdot \nabla\omega = \Delta(\mu\omega) - 2\mu''\partial_x v_2 \equiv \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2.$$

Finally, we linearize the vorticity equation (12) around this shear flow to arrive at the following linearized equation

$$(13)$$
$$\partial_t\omega + U(y)\partial_x\omega - U''(y)v_2 = \Delta(\mu\omega) - 2\mu''\partial_x v_2 \equiv \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2.$$

$\square$

In the following we introduce some equivalent reformulations of linearized equations (9) in order to simplify our notation. We first observe that in the equations (9)

$$\partial_t\omega + U(y)\partial_x\omega - U''v_2 = \text{div}(\mu\nabla\omega) - \text{div}(\mu'\nabla v_1) - \mu''\partial_x v_2,$$
$$v = \nabla^\perp \Delta^{-1}\omega$$

all coefficient functions do not depend on x explicitly.

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

$$\omega_= = 0,$$

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

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

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

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

Expressed in Fourier variables it holds that

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

and hence

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

As a final step we observe that by assumption $U(y)$ is strictly monotone and hence there exists a change of variables $y \mapsto z = U(y)$, which serves as our main formulation in the following.

Lemma 2.2 (Vorticity formulation)**.** *Let U, μ be as in Theorem 1.1 and consider the linearized equations in vorticity formulation. Further denote the change of coordinates $y \mapsto z = U(y)$ and define*

$$W(t, x, z) = \omega(t, x + tU(y), y).$$

Then ω solves (9) *if and only if for every $k \in \mathbb{Z}$ the Fourier transform of W with respect to x solves:*

$$
\begin{aligned}
(15) \quad & \partial_t(\mathcal{F}_x W) - \frac{ikU''}{-k^2 + (U'(\partial_z - ikt))^2}(\mathcal{F}_x W) \\
& = (-k^2 + (U'(\partial_z - ikt))^2)(\mu\mathcal{F}_x W) + \frac{2\mu''k^2}{-k^2 + (U'(\partial_z - ikt))^2}(\mathcal{F}_x W),
\end{aligned}
$$

where with slight abuse of notation coefficient functions are evaluated in y such that $z = U(y)$, e.g. $U'' = \partial_y^2 U|_{y=U^{-1}(z)}$. In other words, by introducing

$$(16) \qquad \nabla_t := \begin{pmatrix} \partial_x \\ U'(\partial_z - t\partial_x) \end{pmatrix},$$

we may write the above equation for $\mathcal{F}_x W$ as

$$(17) \qquad \partial_t W - U''V_2 = \mathrm{div}_t(\mu\nabla_t W) - \mathrm{div}_t(\mu'\nabla_t V_1) - \mu''\partial_x V_2.$$

where

$$V_1 = \frac{-U'(\partial_z - t\partial_x)}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2}W,$$

$$V_2 = \frac{\partial_x}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2}W.$$

Proof of Lemma 2.2. Since by assumption

$$\mu \partial_y U$$

equals a non-trivial constant and $\mu > 0$ does not vanish, it follows that U is strictly monotone and hence invertible. Thus, the above claimed change of variables exists. Furthermore, it holds that

$$\partial_y - tU'\partial_x = U'\partial_z - tU'\partial_x = U'(\partial_z - t\partial_x),$$

which together with (14) concludes the proof. $\qquad\qquad\square$

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 behavior of the coefficient functions. Finally, in Section 5 we bootstrap the stability results in L^2 to establish stability in H^N.

Unless noted otherwise we here always work in coordinates with respect to z and may without loss of generality assume that W is localized at frequency $k \neq 0$, arbitrary but fixed, with respect to x. We thus briefly write

$$L^2 := L^2(\mathbb{R}, dz)$$

and use

$$\langle \cdot, \cdot \rangle$$

to refer to the inner product on that space.

3. A Model Case and L^2 Estimates

In this section we consider a special case of the linearized Navier-Stokes equations (9) in vorticity formulation

$$\partial_t \omega + U(y)\partial_x \omega - U'' v_2 = \mathrm{div}(\mu\nabla\omega) - \mathrm{div}(\mu\nabla v_1) - \mu''\partial_x v_2,$$
$$v = \nabla^\perp \Delta^{-1}\omega,$$

or rather

$$\partial_t W - U'' V_2 = \mathrm{div}_t(\mu\nabla_t W) - \mathrm{div}_t(\mu'\nabla_t V_1) - \mu''\partial_x V_2.$$
$$W(t, x, U(y)) = \omega(t, x - tU(y), y),$$

in which we additionally assume that μ is comparable to a constant globally instead of just locally. More precisely, in this section we additionally require that

$$\frac{\sup(\mu)}{\inf(\mu)} \leq 100.$$

We note that this further implies that U is bilipschitz and hence allows for a global change of variables to $z = U(y)$ (see Section 2 and Lemma 2.2). This simplification therefore allows us to more clearly present commutator estimates and introduce techniques of proof.

In a second step in Section 4 we use that such bounds are true *locally*, that is when μ and U are restricted to bounded intervals of a given length. Extending these restrictions to functions on the whole space which are bounded above and below we will hence be able to use this section's results for the "localized" problems. A

key challenge then lies in controlling non-linear interaction due to the Biot-Savart law and in "gluing" the various estimates in a way that preserves dissipation and decay estimates.

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 to (9).

Proposition 3.1. *Let μ, U satisfy the assumptions of Theorem 1.1 and additionally suppose that*

$$\frac{\sup(\mu)}{\inf(\mu)} \leq 100.$$

Then there exists a time-dependent family of operators $A(t)$ such that for any initial data $\omega_0 \in L^2$ it holds that

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

Furthermore, define $u = \inf U'$ and $\nu := \inf \mu(U')^2$ (note that $\nu = u\mu(0)U'(0)$, since $\mu U'$ is constant) then it holds that

$$\frac{d}{dt}\|AW\|_{L^2}^2 \leq -0.001\|\mathcal{F}(\nu^{1/3} + \nu(\xi - kt)^2 + \frac{1}{1 + u^2(\xi - kt)^2})\mathcal{F}^{-1}AW\|_{L^2}^2 \leq 0$$

In particular, the linear stability estimates in L^2 of Theorem 1.1 hold in this setting.

- The operator $A(t)$ is defined in terms of a Fourier multiplier in Definition 3.2.
- We remark that in the present setting by assumption μ and hence U' may only vary by a factor 100. Therefore u is also comparable to an average value of U'.
- The decay by $\nu^{1/3} + \nu(\xi - kt)^2$ quantifies the enhanced dissipation mechanism. More precisely, if $|\xi - kt| \geq \nu^{-1/3}$ the latter term dominates, but for frequencies with $|\xi - kt|$ smaller than this the enhanced rate $\nu^{1/3}$ still persists (see Definition 3.2 and Lemma 3.4 for further discussion).
- The last multiplier $\frac{u}{1 + u^2(\xi - kt)^2}$ corresponds to control of the velocity perturbation. Since u is bounded below this contribution is dominated by $\nu^{1/3}$ for large frequencies. The statement of the theorem thus is that even for frequencies where $|\xi - kt|$ is small this control persists. In particular, this control remains valid as ν tends to zero, which is crucial for the control of $U''v_2$ in the evolution equation (see Lemma 3.6).
- We note that here the choice of G involves a small factor 0.1, as do the dissipation rates in Theorem 1.1. These small factors allow use some flexibility in the control of interaction terms (see Section 3.1).

In the interest of clear presentation we formulate the main steps of the proof as a series of lemmas. We then show how to use these to establish Proposition 3.1 before proving the lemmas at the end of this section.

Definition 3.2 (Decreasing Multiplier and Fourier sets)**.** *Let μ, U be a given stationary solution as in Theorem 1.1 and define the* local dissipation rate ν *as*

$$\nu := \inf \mu(\partial_y U)^2.$$

and the local shear rate as

$$u := \inf |U'|.$$

Let further

$$G := 0.1\nu^{-1/3}.$$

We then define the good *set* $G_t \subset \mathbb{Z} \times \mathbb{R}$ *by*

$$G_t := \{(k, \xi) : k \neq 0, |\frac{\xi}{k} - t| \geq G\},$$

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

$$B_t = \{(k, \xi) \in \mathbb{Z} \times \mathbb{R} : k \neq 0, |\frac{\xi}{k} - t| < G\}.$$

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^- *the half-line extending to* $-\infty$ *and by* G_t^+ *the half-line extending to* $+\infty$*.*

Associated with this partition we define a Fourier multiplier m *by*

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

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

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 *and* $z = U(y)$*.*

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

- The relative decay of μ by $-\nu^{1/3}$ compensates for the relatively weak dissipation in the *bad* Fourier region. Here the decay of A allows to establish damping of AW.
- The second multiplier models the growth of v_2 as given by the Biot-Savart law or rather of $U'v_2$. As we discuss in the proof of Lemma 3.6 we then use that by assumption $U'' = \frac{U''}{U'}U'$ is small compared to U' and hence the linearization error $U''v_2$ can be controlled by this multiplier when we are in the *bad* region.

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 *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:*

$$m(t, k, \xi) = c \ \textit{if } t > \frac{\xi}{k} + G,$$

$$m(t, k, \xi) = 1 \ \textit{if } t < \frac{\xi}{k} - G.$$

(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 $\phi \in L^2(dz)$.

Given this definition of our multiplier our main task in the following is to establish suitable estimates for

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

where we used the equation (17) to rewrite $A\partial_t W$. More precisely, we intend to show that the dissipation and the decay of $m(t)$ are strong enough to absorb possible growth and that hence $\|AW\|^2$ is decreasing in time. Integrating these estimates we thus obtain a Lyapunov functions, which allows us to prove Proposition 3.1.

The following lemma quantifies the combined strength of the dissipation mechanism and the decay of the multiplier.

Lemma 3.4 (The dissipation term)**.** *Let $t \geq 0$, let A, G and m be given by Definition 3.2 and let $W \in L^2$ be a given function. Then it holds that*

$$\begin{aligned}
&0.2\langle AW, \dot{A}W \rangle + \langle AW, A\operatorname{div}_t(\mu\nabla_t)W \rangle \\
&\leq -0.1\left\|\mathcal{F}^{-1}\left(\nu^{1/3} + u^2(\xi - kt)^2 + \frac{u}{1 + u^2(\frac{\xi}{k} - t)^2}\right)\mathcal{F}AW\right\|^2,
\end{aligned}$$

where u and ν are defined as in Definition 3.2.

This decay lies at the core of our damping mechanism. In the following lemmas we show that all other contributions to $\frac{d}{dt}\|AW\|^2$ can be considered as errors.

We begin by the errors in the dissipation term due to the fact μ is non-constant. Here we recall that by the assumptions imposed in Theorem 1.1 the relative change of μ is required to be small:

$$|\frac{\mu'}{\mu}| \leq 0.001$$

as well as

$$|\partial_z \frac{\mu'}{\mu}| \leq 0.001.$$

This smallness together with the fact that commutator terms involve integro-differential operators of lower order than 2 allows us to control these errors.

Lemma 3.5 (Viscosity errors). *Let $t \geq 0$, let A, C, ν, u and m be given by Definition 3.2 and let $W \in L^2$ be a given function. Then it holds that*

$$0.2\langle AW, \dot{A}W \rangle - \langle AW, A(\operatorname{div}_t(\mu'\nabla_t)v_1 + \mu''\partial_x v_2) \rangle$$

$$\leq 0.01 \|\mathcal{F}^{-1}\left(\nu^{1/3} + u^2(\xi - kt)^2 + \frac{u}{1 + u^2(\frac{\xi}{k} - t)^2}\right)\mathcal{F}AW\|^2.$$

Lemma 3.6 (Velocity errors). *Let $t \geq 0$, let A, C, ν, u and m be given by Definition 3.2 and let $W \in L^2$ be a given function. Then it holds that*

$$0.2\langle AW, \dot{A}W \rangle - \langle AW, A(U''v_2) \rangle$$

$$\leq 0.01 \|\mathcal{F}^{-1}\left(\nu^{1/3} + u^2(\xi - kt)^2 + \frac{u}{1 + u^2(\frac{\xi}{k} - t)^2}\right)\mathcal{F}AW\|^2.$$

The proofs of Lemmas 3.3 to 3.6 are given in the following Section 3.1. We briefly discuss how to combine the estimates of the the lemmas to establish Proposition 3.1.

Proof of Proposition 3.1. Let A, m be given as in Definition 3.2, let W denote the solution of the linearized Navier-Stokes equations and consider the energy

$$E(t) = \|A(t)W(t)\|_{L^2}^2.$$

Then by the results of Lemmas 3.4, 3.5 and 3.6 it follows that

$$(18) \qquad \partial_t E \leq -0.05 \|\mathcal{F}^{-1}\nu^{1/3} + u^2(\xi - kt)^2 + \frac{u}{1 + u^2(\frac{\xi}{k} - t)^2}\mathcal{F}AW\|^2.$$

In order to conclude we recall that by Lemma 3.3 the Fourier multiplier m corresponding to A is bounded between c and 1 and we may hence relate E with $\|W\|_{L^2}^2$. $\qquad \square$

In Section 4 we will extend these estimates to the case where μ (and hence U') is allowed to vary by many orders of magnitude. In particular, the values of ν and μ then are only locally defined. A key challenge there is to show that non-local effects and interactions between different regions in space can be controlled in a sufficiently good way. Finally, in Section 5 we show that the damping estimates in L^2 can be bootstrapped to yield stability in arbitrary Sobolev regularity, following an argument of [Zil21].

3.1. **Proof of Lemmas.** 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

$$(19) \qquad m(t, k, \xi) = \exp\left(\int_{-\infty}^{t} \frac{\partial_t m}{m} 1_{B_\tau}(k, \xi)\, d\tau\right),$$

where we used that $m(-\infty, k, \xi) = 1$. It then holds that

$$\int_{-G}^{G} \nu^{1/3} dt = 2G\nu^{1/3} \leq 0.2$$

is bounded by a uniform constant (and further improves for k large). Furthermore, also

$$\int_{-G}^{G} \frac{u}{1 + u^2 t^2} dt = \arctan(\tau)|_{\tau = -uG}^{uG} \leq \pi$$

is uniformly controlled. Therefore, we may estimate

$$m(t, k, \xi) \geq \exp(-0.2 - \pi).$$

We further observe that for $t < \frac{\xi}{k} - G$ or $t > \frac{\xi}{k} - G$ it holds that

$$\frac{\partial_t m}{m} = 0$$

and m is thus constant on these intervals. On the left interval m equals $m(-\infty, k, \xi) = 1$, while on the right it equals

$$\exp(-\int_{-G}^{G} \nu^{1/3} + \frac{u}{1 + u^2 t^2} dt) := c.$$

Finally, by Parseval's identity these bounds for the multiplier m are equivalent to L^2 bounds for the operator A. $\qquad\square$

Having established these properties of the operator A we next turn to estimating the dissipation. Here for *good* frequencies in the sense of Definition 3.2 the dissipation is rather strong. If one instead considers *bad* (that is, close to resonant) frequencies we rely on the fact that m decreases in time t to provide sufficient decay.

Proof of Lemma 3.4. We recall that by Lemma 3.3 the multiplier m is multiple of the identity on the connected components of the *good* Fourier set G_t. In particular, when restricted to these sets A commutes with all other operators. In our proof we hence expand

$$W = \mathcal{F}^{-1} 1_{G_t^-} \mathcal{F} W + \mathcal{F}^{-1} 1_{B_t} \mathcal{F} W + \mathcal{F}^{-1} 1_{G_t^+} \mathcal{F} W =: W_1 + W_2 + W_3$$

according to this Fourier decomposition. That is, we study

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

with $i, j \in \{1, 2, 3\}$.

We begin by discussing the diagonal cases $i = j$ in the *good* regime.

Estimates for $\langle AW_j, A \operatorname{div}_t(\mu \nabla_t) W_j \rangle$, $j = 1$ or 3:

We recall that $A = \operatorname{Id}$ or $c \operatorname{Id}$ when applied to W_1 or W_3, respectively. Hence, we may explicitly compute that

$\langle AW_1, A \operatorname{div}_t(\mu \nabla_t) W_1 \rangle = \langle W_1, \operatorname{div}_t(\mu \nabla_t) W_1 \rangle$

$= \langle W_1, \mu \partial_x^2 + U'(\partial_z - t\partial_x)\mu U'(\partial_z - t\partial_x) W_1 \rangle$

$= -\langle \partial_x W_1, \mu \partial_x W_1 \rangle - \langle U'(\partial_z - t\partial_x) W_1, \mu U'(\partial_z - t\partial_x) W_1 \rangle - \langle (\partial_z U' W_1), \mu U'(\partial_z - t\partial_x) W_1 \rangle$

$\leq -\nu \|(\partial_z - t\partial_x) W_1\|^2 - \langle (\partial_z U') W_1, \mu U'(\partial_z - t\partial_x) W_1 \rangle,$

where we estimated the first summand by 0 from above and used that $\mu(U')^2 \geq \nu$ by definition. For the last term we recall that μ and hence U' is slowly varying and we may therefore control

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

$$\leq \|\frac{\partial_z U'}{U'}\|_{L^\infty}100G^{-1}\|u(\partial_z - t\partial_x)W_1\|\|\mu U'(\partial_z - t\partial_x)W_1\|$$

$$= \|\frac{\partial_z U'}{U'}\|_{L^\infty}100G^{-1}\nu\|(\partial_z - t\partial_x)W_1\|_{L^2}^2,$$

where we used that $(\partial_z - t\partial_x)$ is invertible with operator norm of the inverse map bounded by G^{-1} on the *good* set and controlled $|U'| \leq 100u$. Since G^{-1} is very small, as is $\|\frac{\partial_z U'}{U'}\|_{L^\infty}$, it holds that

$$\text{(Cond.1)} \qquad\qquad 100G^{-1}\|\frac{\partial_z U'}{U'}\|_{L^\infty} \leq 0.001$$

and hence this error can be absorbed into the decay. We thus note that

$$\langle AW_1, A\operatorname{div}_t(\mu\nabla_t)W_1\rangle \leq -\|\sqrt{\mu}\partial_x AW_1\|_{L^2}^2 - 0.9\nu\|(\partial_z - t\partial_x)AW_1\|^2$$

and by the same argument

$$\langle AW_3, A\operatorname{div}_t(\mu\nabla_t)W_3\rangle \leq -\|\sqrt{\mu}\partial_x AW_3\|_{L^2}^2 - 0.9\nu\|(\partial_z - t\partial_x)AW_3\|^2.$$

Since μ and U' are not constant there further is some non-trivial interaction between different *good* contributions.

Estimates for $\langle AW_i, A\operatorname{div}_t(\mu\nabla_t)W_j\rangle$ with $(i,j) = (1,3)$ or $(i,j) = (3,1)$:
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

$$\langle W_i, \operatorname{div}_t(\mu\nabla_t)W_j\rangle = -\langle\partial_x W_i, \mu\partial_x W_j\rangle - \langle U'(\partial_z - t\partial_x)W_i, \mu U'(\partial_z - t\partial_x)W_j\rangle$$

$$+ \langle U'W_i, \frac{\partial_z U'}{U'}\mu U'(\partial_z - t\partial_x)W_j\rangle.$$

By the same argument as above

$$|\langle U'W_i, \frac{\partial_z U'}{U'}\mu U'(\partial_z - t\partial_x)W_j\rangle| \leq \frac{1}{10}\nu\|(\partial_z - t\partial_x)W_1\|\|(\partial_z - t\partial_x)W_3\|$$

can be considered a negligible error term.

For the remaining terms we instead exploit that W_i and W_j have disjoint support in Fourier space and that these supports have distance at least $2G$. In particular, we may estimate

$$\langle\partial_x W_i, \mu\partial_x W_j\rangle = \int \mathcal{F}(\mu)(\xi_1)\mathcal{F}(\partial_x W_i)(\xi_2)\overline{\mathcal{F}(\partial_x W_i)}(\xi_1 + \xi_2)d\xi_1 d\xi_2$$

$$\leq \|\frac{1}{\xi_1}\mathcal{F}(\partial_z\mu)(\xi_1)\|_{L^1(|\xi_1| \geq 2G)}\|\partial_x W_1\|_{L^2}\|\partial_x W_3\|_{L^2}$$

$$\leq G^{-1}\|\mathcal{F}(\partial_z\mu)(\xi_1)\|_{L^1(|\xi_1| \geq G)}\frac{1}{\inf(\mu)}\|\sqrt{\mu}\partial_x W_1\|_{L^2}\|\sqrt{\mu}\partial_x W_3\|_{L^2}.$$

This interaction term can hence absorbed by the decay provided

$$(20) \qquad G^{-1}\|\mathcal{F}(\partial_z\mu)(\xi_1)\|_{L^1(|\xi_1|\geq G)}\frac{1}{\inf(\mu)} \leq 0.1,$$

which is part of our assumption (7) that the relative size of μ is slowly varying. We remark that here $\mathcal{F}(\mu)$ refers to the distributional Fourier transform, since μ is bounded but not in L^1.

By the same argument and using that

$$\frac{\partial_z U'}{U'} = -\frac{\partial_z\mu}{\mu}$$

we observe that

$$|\langle U'(\partial_z - t\partial_x)W_i, \mu U'(\partial_z - t\partial_x)W_j\rangle|$$
$$\leq \frac{\sup_{|\xi_1|\geq 2G}\mathcal{F}(U')(\xi_1)}{u}\nu\|(\partial_z - t\partial_x)W_1\|\|(\partial_z - t\partial_x)W_3\|,$$

which can be absorbed provided

$$(21) \qquad \frac{\sup_{|\xi_1|\geq 2G}|\mathcal{F}(U')(\xi_1)|}{\inf(U')} \leq 0.5,$$

which holds by assumption.

Estimates for terms involving W_2:
It remains to discuss the influence of the part W_2 Fourier-localized in the *bad set*.

We first study the self-interaction term:

$$\langle AW_2, A\operatorname{div}_t(\mu\nabla_t)W_2\rangle = \langle AW_2, A\partial_x\mu\partial_x W_2\rangle$$
$$+ \langle AW_2, A(\partial_z - t\partial_x)U'\mu U'(\partial_z - t\partial_x)W_2\rangle$$
$$- \langle AW_2, A(\partial_z U')\mu U'(\partial_z - t\partial_x)W_2\rangle.$$

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

$$\langle AW_2, A(\partial_z - t\partial_x)U'\mu U'(\partial_z - t\partial_x)W_2\rangle \leq C\nu G^2\|AW_2\|_{L^2}^2$$

since W_2 is localized in the Fourier set where $|\xi - kt|$ is not large (yet). We recall that here G was chosen in such a way that $\nu G^2 < 0.01\nu^{1/3}$ and thus the dissipation in this *bad* region is weaker than desired.

Similarly, we may estimate

$$(22) \qquad \langle AW_2, A\partial_x\mu\partial_x W_2\rangle \leq C\|\sqrt{\mu}\partial_x AW\|_{L^2}^2 \leq C\nu k^2\|AW\|_{L^2}^2,$$

where we used that $(U')^2 \geq 1$ and hence $\inf(\mu) \leq \nu$. As remarked following Definition 3.2 we may without loss of generality only consider those k for which

$$\nu k^2 < 0.001\nu^{1/3}$$

is much smaller than the enhanced dissipation rate, since otherwise this horizontal dissipation already achieves the desired decay. With this understanding we restrict to this case for the reminder of the article. It then also holds that

$$\langle AW_2, A\partial_x\mu\partial_x W_2\rangle \leq C\nu^{1/3}\|AW_2\|_{L^2}^2$$

with a small constant C. Furthermore, we may control

$$|\langle AW_2, A(\partial_z U')\mu U'(\partial_z - t\partial_x)W_2\rangle| \leq C\|\frac{\partial_z U'}{U'}\|_{L^\infty}\|\mu U'(\partial_z - t\partial_x)W_2\|_{L^2}\|A\|^2$$
$$(\|\partial_x W_2\|_{L^2} + \|u(\partial_z - t\partial_x)W_2\|_{L^2})$$
$$\leq C\nu^{1/3}\|AW_2\|_{L^2}^2$$

with a small absolute constant C, by our choice of cut-off G.

Since the decay of A yields that

$$\langle AW_2, \dot{A}W_2\rangle \leq -\nu^{1/3}\|AW_2\|_{L^2}^2$$

these contributions can be absorbed.

Finally, it remains to discuss the cross terms

$$\langle AW_i, \text{div}_t(A\mu\nabla_t)W_j\rangle = \langle AW_i, A\partial_x\mu\partial_x W_j\rangle$$
$$+ \langle AW_i, A(\partial_z - t\partial_x)U'\mu U'(\partial_z - t\partial_x)W_j\rangle$$
$$+ \langle AW_i, A(\frac{\partial_z U'}{U'})U'\mu U'(\partial_z - t\partial_x)W_j\rangle,$$

where one i, j equals 2. For the first term we estimate by

$$C\|\sqrt{\mu}\partial_x AW_i\|_{L^2}\|\sqrt{\mu}\partial_x AW_j\|_{L^2},$$

which can be absorbed as above. Similarly, the second term can be controlled by

$$C\|u(\partial_z - t\partial_x)W_{1,3}\|_{L^2}\nu^{1/3}\|W_2\|$$

and the last term by

$$CG^{-1}\|\frac{\partial_z U'}{U'}\|_{L^\infty}\|u(\partial_z - t\partial_x)W_{1,3}\|_{L^2}\nu^{1/3}\|W_2\|.$$

We may thus use use Young's inequality and the previously obtained damping estimates to absorb these error terms, which concludes the proof. □

In the previous lemma we have established dissipation due the "main" term of the dissipation operator. The following lemma shows that this decomposition is indeed and all terms involving higher derivatives of μ can be considered lower order. We recall here that by assumption μ is slowly varying in the sense that

$$\|\frac{\mu'}{\mu}\|_{L^\infty}$$

and

$$\|\partial_z \frac{\mu'}{\mu}\|_{L^\infty}$$

are small.

Proof of Lemma 3.5. Given the decay established in Lemma 3.4 we may use integration by parts, Hölder's inequality and Young's inequality to reduce our proof to establishing bounds on norms of

$$\mu'\nabla_t v_1$$

and of

$$\mu''\partial_x v_2.$$

In analogy to the proof of Lemma 3.4 we here again distinguish between the parts of v_1, v_2 generated by the vorticity in the *good* region W_1, W_3 and the one localized in the *bad* region W_2.

More precisely, we note that as in [CZZ19] in these estimates we may replace $v = \nabla_t \phi$, defined in terms of the usual stream function, by simpler potential using the averaged value u of U'. For this purpose we define ψ to satisfy

$$(\partial_x^2 + u^2(\partial_z - t\partial_x)^2)\psi = W = (\partial_x^2 + (U'(\partial_z - t\partial_x))^2)\phi$$

with suitable integrability assumptions at infinity. Then testing these equations with either ψ or ϕ one obtains that the energies

$$\|v\|_{H^N}^2 \approx \|\partial_x \psi\|^2 + \|u(\partial_y - t\partial_x)\psi\|^2$$

are comparable (in the sense of bilinear forms acting on W).

In the following we may thus discuss

$$V = (-u(\partial_z - t$$
$$p_x), \partial_x)\psi$$

in place of v.

We then observe that in the *good* Fourier region

$$\mathcal{F}(\partial_x V_2) = \frac{k^2}{k^2 + u^2(\xi - kt)^2}\mathcal{F}(W)$$

is controlled by $c^{-2}\mathcal{F}(W)$.

Similarly, in the *good* Fourier region

$$\nabla_t V_1$$

can be controlled by the dissipation, since

$$\frac{1}{1 + u^2(\frac{\xi}{k} - t)^2} \leq \frac{1}{1 + G^2} \ll \nu^{1/3}.$$

It thus only remains to discuss the *bad* Fourier region. However, there the weight A is chosen in just such a way that

$$\|(\partial_x, u(\partial_z - t\partial_x))\psi_A\|^2$$

can be absorbed using

$$\langle AW, \dot{A}W \rangle.$$

More precisely, we may estimate

$$u\|(\partial_x, u(\partial_z - t\partial_x))\psi_A\|^2 = \int \frac{u}{k^2 + u^2(\xi - kt)^2}|AW|^2$$
$$\leq \int (1_{B_t}(k, \xi)\frac{u}{k^2 + u^2(\xi - kt)^2} + 0.1\nu^{1/3}1_{G_t})|AW|^2,$$

since

$$\frac{u}{k^2 + u^2(\xi - kt)^2} < 0.1\nu^{1/3}$$

in the *good* region. The right-hand-side then is controlled by the decay of A and by the dissipation, which concludes the proof.

$\square$

Finally we turn to the control of the error in due to the convective term:

$$U''v_2 = \frac{U''}{U'}U'v_2.$$

Proof of Lemma 3.6. We remark that in the case of very small effective viscosity one cannot expect to control

$$\langle AW, AU''v_2 \rangle \tag{23}$$

in terms of the dissipation. Hence, we fall back to the estimates developed for the inviscid case in [Zil17]. More precisely, let again $u = \min U'$ and define the constant coefficient stream function ψ_A and ψ as the solution of the equations

$$(\partial_x^2 + u^2(\partial_z - t\partial_x)^2)\psi_A = AW,$$

and

$$(\partial_x^2 + u^2(\partial_z - t\partial_x)^2)\psi = W,$$

respectively. Note that both differential operators involve constant coefficients and hence both ψ and ψ_A can be explicitly computed in terms of Fourier multipliers.

Furthermore, integrating by parts and using that $v_2 = \partial_x \phi$ we control (23) by

$$\|(\partial_x, u(\partial_z - t\partial_x))\psi_A\|_{L^2}\|A\|\|U''\|_{W^{1,\infty}}\|(\partial_x, U'(\partial_z - t\partial_x))\partial_x\phi\|_{L^2},$$

where $\|A\|$ denotes the L^2 operator norm of A.

We next claim that it holds that

$$\begin{aligned}\|(\partial_x, U'(\partial_z - t\partial_x))\partial_x\phi\|_{L^2} &\leq \|(\partial_x, u(\partial_z - t\partial_x))\partial_x\psi\|_{L^2} \\ &\leq \|A^{-1}\|\|(\partial_x, u(\partial_z - t\partial_x))\partial_x\psi_A\|_{L^2}.\end{aligned} \tag{24}$$

Assuming this claim for the moment, we may compute

$$\|(\partial_x, u(\partial_z - t\partial_x))\psi_A\|_{L^2}^2 = \int \frac{1}{k^2 + u^2(\xi - kt)^2}|\mathcal{F}(AW)(t,k,\xi)|^2$$

and hence observe that the velocity error (23) can be estimated by

$$\int \|A\|\|A^{-1}\|\left\|\frac{U''}{u}\right\|_{W^{1,\infty}} \frac{|k|u}{k^2 + u^2(\xi - kt)^2}|\mathcal{F}(AW)(t,k,\xi)|^2$$

Since by assumption

$$\|A\|\|A^{-1}\|\|\frac{U''}{u}\|_{W^{1,\infty}}$$

is small this error can thus indeed be absorbed using the dissipation and the decay of $A(t)$.

It remains to prove the claim (24) for which argue as in [Zil17]. That is, we test the stream function equation

$$(\partial_x^2 + (U'(\partial_z - t\partial - x))^2)\phi = W = (\partial_x^2 + u^2(\partial_z - t\partial_x)^2)\psi$$

with ϕ (or rather $\frac{U'}{u}\phi$) to obtain that

$$\|(\partial_x, U'(\partial_z - t\partial_x))\phi\|_{L^2}^2 \leq \|(\partial_x, U'(\partial_z - t\partial_x))\phi\|_{L^2}\|(\partial_x, u(\partial_z - t\partial_x))\psi\|_{L^2},$$

which yields the first estimate of (24). The second estimate (24) immediately follows from the explicit characterization of ψ and ψ_A in terms of Fourier multipliers (which only differ by multiplication with the Fourier weight of A).

$$\square$$

We remark that unlike the estimates of Lemmas 3.4 and 3.5 the above estimate does not explicitly involve the viscosity and has been obtained in the inviscid case. This lemma hence imposes the strongest restrictions on the profile U (and hence equivalently on μ). As discussed following the statement of Theorem 1.1 we do not expect the smallness condition to be optimal, but rather a non-resonance/spectral condition as in [WZZ18]. However, the present stronger assumption allows for an approach in terms of a Lyapunov functional.

4. Localization and Non-local Interactions

In this section we consider the linearized equations (9) in vorticity formulation

$$\partial_t \omega + U(y)\partial_x \omega - U''v_2 = \operatorname{div}(\mu\nabla\omega) - \operatorname{div}(\mu\nabla v_1) - \mu''\partial_x v_2,$$

$$v = \nabla^\perp \Delta^{-1}\omega.$$

Unlike in Section 3 we here allow for μ (and hence also U') to vary by many orders of magnitude.

Our main result of this section, Proposition 4.1 then establishes the stability and damping results of Theorem 1.1 in L^2, for which a special case had been treated in Proposition 3.1.

Proposition 4.1. *Let μ, U satisfy the assumptions of Theorem 1.1. Then there exists a time-dependent family of operators $A(t)$ such that for any initial data $\omega_0 \in L^2$ the solution $W(t)$ with that initial data satisfies*

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

Furthermore, it holds that

$$\frac{d}{dt}\|AW\|_{L^2}^2 \le -0.001\left(\|(\mu(U')^2)^{1/6}W\|_{L^2}^2 + \|\sqrt{\mu}U'(\partial_z - t\partial_x)W\|_{L^2}^2 + \|\sqrt{U'}v\|_{L^2}^2\right).$$

We recall that as part of the assumptions of Theorem 1.1 we require that (7) holds:

$$\|\frac{\mu'}{\mu}\|_{W^{1,\infty}} < 0.1.$$

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

$$\frac{\max_I \mu}{\min_I \mu} \le 100.$$

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 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 (9) (see Lemma 4.4) and have to be combined to control growth of the whole space problem (see Lemmas 4.5, 4.6 and 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.

4.1. **Partitions and Non-local Interaction.** 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 below and hence cut-off functions and partitions of unity corresponding to this partition have controlled $W^{k,\infty}$ norms. Using these partitons 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 $N \in \mathbb{N}$ and $\mu \in C^{N+2}$ be as in Theorem 1.1. Then there exits a partition $(I_j)_{j\in\mathbb{Z}}$ of $\mathbb{R}$ into intervals such that*

$$(25) \qquad\qquad \frac{\sup_{3I_j} \mu}{\inf_{3I_j} \mu} \leq 50$$

for all j, where $3I_j$ denotes the rescaled intervals with the same center. Furthermore, the length of each interval I_j is bounded below by 1.

Associated with this partition there exists a family of non-negative functions $\chi_j \in C_c^\infty$ with $\operatorname{supp}(\chi_j^2) \subset 3I_j$ such that χ_j^2 is a partition of unity.

For each j there exist $\mu_j, U_j \in C^{N+2}(\mathbb{R})$ such that

$$\mu_j = \mu, U_j = U \text{ in } I_j,$$
$$\mu_j U_j \equiv const. \text{ in } \mathbb{R}$$

and so that μ_j and $\partial_y U_j$ are constant outside $3I_j$ and

$$\frac{\max_{\mathbb{R}} \mu_j}{\min_{\mathbb{R}} \mu_j} \leq 100.$$

Proof of Lemma 4.2. We recall that by assumption on μ the relative rate of change $\frac{\mu'}{\mu}$ is bounded. Hence, given any two points y_1, y_2 we observe that

$$\frac{\mu(y_2)}{\mu(y_1)} = \exp(\ln(\mu(y_2)) - \ln(\mu(y_1))) = \exp\left(\int_{y_1}^{y_2} \frac{\mu'}{\mu} dy\right)$$

is bounded in terms of $|y_2 - y_1|$, also when exchanging y_1 and y_2.

In order to construct the intervals I_j we thus pick an initial point $y_1 = 0$ and then choose $y_2 > 0$ (or $y_2 < 0$) maximally such that (25) holds with $I_j = (y_1, y_2)$. By the above calculation the size of I_j is bounded below. Therefore, iterating this procedure with y_1 chosen as a boundary point of a previously generated interval, we obtain the desired partition $(I_j)_j$ of $\mathbb{R}$ with the size of each I_j bounded below.

We remark that in this greedy procedure it is possible that for up to two choices of j the interval I_j is unbounded (in which case the above procedure only generates finitely many j). In this case we may instead impose that y_2 should be maximized under the additional constraint that $|y_2 - y_1| \leq 1000$.

It is a classical result that given such a partition of 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 the size of each I_j is bounded below).

Furthermore, given this partition unity we construct an extension of μ by

$$\mu_j = \mu(y_1) \sum_{l<j} \chi_l^2 + \mu\chi_j^2 + \mu(y_2) \sum_{l>j} \chi_l^2.$$

The associated shear profile U_j is then constructed by integrating

$$\partial_y U_j := \frac{C}{\mu_j}$$

with C and the the constant of integration chosen such that $U_j(y_1) = U(y_1)$ and $\partial_y U_j(y_1) = \partial_y U(y_1)$. This then directly implies the desired bounds, where we used that the derivatives of the partition of unity are bounded and hence the estimate (25) only possibly deteriorates by a small factor under this extension. $\qquad\square$

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

$$W_j := \chi_j W$$

and the energy functional

$$E(t) = \sum_j \langle A_j W_j, A_j W_j \rangle.$$

We remark that here for each interval I_j we consider the L^2 inner product on $L^2(dz_j)$. However, since χ_j is compactly supported in $3I_j$ we observe that

$$\langle A_j W_j, A_j \chi_j \partial_t W \rangle_{L^2(dz_j)} = \langle A_j^2 W_j, \chi_j \partial_t W \rangle \leq \|A_j^2 W_j\|_{L^2(dz_j)} \|\chi_j \partial_t W\|_{L^2(dz)},$$

where in the last step we used that z and z_j agree on this support. In view of this compatbility with $L^2(dz)$ we may hence transparently switch between the spaces $L^2(dz_j), j \in \mathbb{Z}$ in several estimates and thus suppress this formal j dependence in our notation.

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 and suppose that $\chi_j \in C_b^0$.*

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

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

Let next $N \in \mathbb{N}$ and suppose that $\chi_j \in C_b^N$. Then there exist constants $d_0, \ldots, d_N$ with $d_N = 1$ and c_1, c_2 such that

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

Proof of Lemma 4.4. Since χ_j^2 is a partition of unity, this estimate is actually trivially true with $c_1 = c_2 = 1$ and equality.

Moreover, if $\|A_j W\| \approx \|W\|$ with constants uniform in j, this also implies that

$$\sum_j \|A_j W_j\|^2 \approx \sum_j \|W_j\|^2 = \|W\|^2.$$

Here and in the following the notation $a \approx b$ states that there exist constants $0 < c_1 < c_2 < \infty$ such that $c_1 a \leq b \leq c_2 a$.

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

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

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

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

For all other terms we note that

$$|\partial^\beta \chi_j| \leq \|\chi_j\|_{C_b^{|\beta|}} 1_{\mathrm{supp}(\chi_j)}$$

and that the supports of the functions χ_j at most cover $\mathbb{R}$ twice. Hence, we may control

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

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.

We further remark that these comparisons remain true if W_j is replaced by $A_j W_j$. $\qquad\square$

Given this definition of an energy, we next need to verify that it indeed is a Lyapunov functional and thus study

$$\begin{aligned}
\frac{d}{dt} E &= \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_j \langle A_j W_j, A_j \chi_j \partial_t W \rangle \\
&= \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_{j,j'} \langle A_j W_j, A_j \chi_j (\mathrm{div}_t(\mu \nabla_t \chi_{j'} W_{j'})) \rangle \\
&\quad - \sum_j \langle A_j W_j, A_j \chi_j (\mathrm{div}_t(\mu' \nabla_t v_1) - \mu'' \partial_x v_2) \rangle \\
&\quad + \sum_j \langle A_j W_j, A_j \chi_j U'' v_2 \rangle.
\end{aligned}$$

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 neighbors

instead of full non-local interaction as for the velocity). However, the coupling introduced by this interaction implies that we cannot hope to control $\langle A_j W_j, A_j W \rangle$ 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*

$$(26) \quad \begin{aligned} &0.01 \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle + \sum_j \langle A_j W_j, A_j \chi_j (\operatorname{div}_t(\mu \nabla_t W)) \rangle \\ &\leq -0.01 \sum_j \| \sqrt{\mu} U'(\partial_z - t\partial_x) A_j W_j \|^2 + \|(\mu(U')^2)^{1/6} A_j W_j\|^2. \end{aligned}$$

Proof of Lemma 4.5. We note that in (26) the dissipation involves W and not just W_j and we thus have to control the interaction with other intervals. However, by construction only neighboring functions $\chi_j, \chi_{j'}$ with $j' \in \{j-1, h, j+1\}$ have intersections of their support.

We thus expand

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

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

$$[\operatorname{div}_t \mu \nabla_t, \chi_j]$$

is a first order differential operator. In the *good* region it can thus easily be controlled by the dissipation by the same argument as in the proof of Lemma 3.4.

In the *bad* region we have to require that derivatives of χ_j are not too large. As discussed in Lemma 4.2 this control of the derivatives is a consequence of our assumption that μ only varies gradually and that hence the sizes of the intervals I_j is bounded below by a (large) constant. This then implies that we can use Young's inequality to absorb these terms into the dissipation.

This smallness is a consequence of our assumptions on μ, which imply that that each χ_j is supported on intervals of size at least L and hence an n-th order derivative is controlled in terms of L^{-n}, which is much smaller than 1. $\qquad\square$

Lemma 4.6 (Non-local velocity estimates)**.** *Let $t \geq 0$, let A, C and m be given by Definition 3.2 and let $W \in H^N$ be a given function. Then it holds that*

$$\begin{aligned} &\sum_j 0.2 \langle A_j W_j, \dot{A}_j W_j \rangle - \sum_j \langle A_j W_j, A_j \chi_j(U'' v_2) \rangle \\ &\leq 0.1 \nu^{1/3} \sum_j \|A_j W_j\|^2 + 0.1 \sum_j \|\sqrt{\mu} \nabla_t A_j W_j\|^2 \end{aligned}$$

Proof of Lemma 4.6. We again observe that here the right-hand-side depends on all of W. However, 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:

$$\Delta_j \phi_j = W_j = \chi_j W,$$

and let ϕ denote the stream function generated by W:

$$\Delta \phi = W = \sum_j \chi_j W_j.$$

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

$$\|\nabla_j \phi_j\|^2 \leq \|\nabla_j \phi_j\| \|\nabla(\chi_j \phi)\|$$

and

$$\|\nabla \phi\|^2 \leq \sum_j \|\nabla(\chi_j \phi)\| \|\nabla_j \phi_j\|.$$

Using the fact that derivatives of χ_j are bounded, it thus follows that

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

Thus errors in velocity can be controlled in terms of sums of $\nabla_j \phi_j$ (see also Lemma 4.4). Moreover, the above argument extends to considering weighted spaces.

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. $\qquad\square$

Lemma 4.7 (Viscosity errors)**.** *Let $t \geq 0$, let A, C and m be given by Definition 3.2 and let $W \in H^N$ be a given function. Then it holds that*

$$0.2 \sum_j \langle A_j W_j, \dot{A}_j W_j \rangle - \sum_j \langle A_j W_j, A_j \chi_j (\mathrm{div}_t(\mu' \nabla_t) v_1 + \mu'' \partial_x v_2) \rangle$$
$$\leq 0.1 \nu^{1/3} \sum_j \|A_j W_j\|^2 + 0.1 \sum_j \|\sqrt{\mu} \nabla_t A_j W_j\|^2$$

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

$$(\mathrm{div}_t(\mu' \nabla_t) v_1 + \mu'' \partial_x v_2)$$

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 yields the result. $\qquad\square$

Having establised 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 $\omega_0 \in L^2(dz)$ be a given initial datum, let μ, U satsify the asssumptions of Theorem 1.1 and let W denote the solution of (17)

$$\partial_t W - U'' V_2 = \text{div}_t(\mu \nabla_t W) - \text{div}_t(\mu' \nabla_t v_1) - \mu'' \partial_x v_2.$$

with this initial data, where

$$(27) \qquad\qquad v_1 = \frac{-U'(\partial_z - t\partial_x)}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2} W,$$

$$(28) \qquad\qquad v_2 = \frac{\partial_x}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2} W.$$

Then by Lemma 4.2 there exists a parition 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

$$E(t) := \sum_j \langle A_j W_j, A_j W_j \rangle.$$

Inserting the evolution equation (17) we then have to estimate

$$\begin{aligned}
\frac{d}{dt} E(t) = {}& 2 \sum_j \langle \dot{A}_j W_j, A_j W_j \rangle \\
& + 2 \sum_j \langle A_j W_j, A_j \chi_j (\text{div}_t(\mu \nabla_t W)) \rangle \\
& - 2 \sum_j \langle A_j W_j, \chi_j \text{div}_t(\mu' \nabla_t v_1) \rangle \\
& - 2 \sum_j \langle A_j W_j, A_j \chi_j \mu'' \partial_x v_2 \rangle.
\end{aligned}$$

Combining the estimates of each summand, derived in Lemmas 4.4 to 4.7 we deduce that

$$\frac{d}{dt} E(t) \leq -0.01 \sum_j \left(\nu_j^{1/3} \|W_j\|^2 + \|\sqrt{\mu_j} U_j'(\partial_{z_j} - t\partial_x) W_j\|^2 + u_j \|(\partial_x, u_j(\partial_{z_j} - t\partial_x)) \psi_{A_j}\|^2 \right).$$

Finally, we recall that μ_j, U_j', z_j agree with μ, U', z on each interval I_j and that by Lemma 4.4 the energy $E(t)$ is comparable to $\|W(t)\|_{L^2(dz)}^2$. 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 4.1 to estimates in H^N. Here

we follow an inductive approach introduced in [Zil21] in the inviscid setting. We consider the linearized equations (17)

$$\partial_t W = U'' v_2 + \operatorname{div}_t(\mu \nabla_t W) - \operatorname{div}_t(\mu' \nabla_t v_1) - \mu'' \partial_x v_2 =: LW,$$
$$v_1 = \frac{-U'(\partial_z - t\partial_x)}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2} W,$$
$$v_2 = \frac{\partial_x}{\partial_x^2 + (U'(\partial_z - t\partial_x))^2} W,$$

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 z. Applying N derivatives to (17) we obtain that

$$(29) \qquad \partial_t \partial_z^N W = L \partial_z^N W + [L, \partial_z^N] W.$$

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

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

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

is a Lyapunov functional and satisfies

$$\frac{d}{dt} E(t) \leq -0.01 (\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^N W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}} \partial_z^N W\|_{L^2}^2).$$

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

Proof of Prosposition 5.1. The case $N = 0$ has been established in Proposition 4.1 with $c_0 = 1$. We hence aim to proceed by induction. Hence, suppose that the estimates have been established for the case $N - 1$ and consider

$$E_N(t) = c_N \langle A \partial_z^N W, A \partial_z^N W \rangle + E_{N-1}(t) + E_{N-2}(t) + \cdots + E_0(t)$$

with c_N to be determined later.
    Then by the induction assumption it holds that

$$(30) \qquad \frac{d}{dt} E_l(t) \leq -0.01 (\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^l W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}} \partial_z^l W\|_{L^2}^2)$$

for all $0 \leq l \leq N - 1$ 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 $\partial_z^N W$. Here, by the L^2 estimates of Proposition 4.1 it holds that

(31)
$$\frac{d}{dt} c_N \langle A\partial_z^N W, A\partial_z^N W\rangle = 2c_N \langle \dot{A}\partial_z^N W, A\partial_z^N W\rangle + 2c_n\langle A\partial_z^N W, AL\partial_z^N W\rangle$$
$$+ 2c_N\langle A\partial_z^N W, A[L, \partial_z^N]W\rangle$$
$$\leq -0.01 c_N(\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^{N-1}W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^{N-1}W\|_{L^2}^2)$$
$$+ 2c_N\langle A\partial_z^N W, A[L, \partial_z^N]W\rangle.$$

Combining the estimates (31) and (30) it thus suffices to show that for a suitable choice of c_N we may absorb the commutation error
$$2c_N\langle A\partial_z^N W, A[L, \partial_z^N]W\rangle.$$

into the decay
$$-0.01 c_N(\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^{N-1}W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^{N-1}W\|_{L^2}^2)$$
$$-0.01 \sum_{l<N}(\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^l W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^l W\|_{L^2}^2)$$

Let us first discuss the main dissipation term of L. Here we may iteratively expand
$$[\text{div}_t(\mu\nabla_t), \partial_z^N]W = [\text{div}_t(\mu\nabla_t), \partial_z]\partial_z^{N-1}W$$
$$+ [\text{div}_t(\mu\nabla_t), \partial_z^{N-1}]\partial_z^1 W + [[\text{div}_t(\mu\nabla_t), \partial_z^{N-1}], \partial_z]W$$
$$= \sum_{l<N} E_l\partial_z^l W,$$

where the operators E_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
$$\langle A\partial_z^N W, A[\text{div}_t(\mu\nabla_t), \partial_z^N]W\rangle = \sum_{l<N}\langle A\partial_z^N, AE_l\partial_z^l W\rangle$$

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

(32) $$c_N \sum_{l<N}\langle A\partial_z^N, AE_l\partial_z^l W\rangle \leq c_N\|\sqrt{\mu}\nabla_t A\partial_z^N W\|_{L^2}\sum_l d_l\|\sqrt{\mu}\nabla_t\partial_z^k lW\|_{L^2}.$$

Similarly we may iterative expand the equation satisfied by derivatives of the stream function
$$\Delta_t\partial_z^N\phi = \partial_z^N W + [\Delta_t, \partial_z^N]\phi$$

and thus obtain that
$$\partial_z^N\phi = \Delta_t^{-1}\partial_z^N W + \Delta_t^{-1}\sum_{l<N}\tilde{E}_l\Delta_t^{-1}\partial_z^l W,$$

where the second order operators $\tilde{E}_l$ 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

$$\langle A\partial_z^N W, A[\mathrm{div}_t(\mu'\nabla_t)U'(\partial_z - t\partial_x)\Delta_t^{-1} - \mu''\partial_x\Delta_t^{-1} + U''\partial_x\Delta_t^{-1}, \partial_z^N]W\rangle$$

$$(33) \quad \leq (\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^N W\|_{L^2} + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^N W\|_{L^2})$$

$$\sum_l (\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^l W\|_{L^2} + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^l W\|_{L^2}).$$

We may thus conclude our estimate by using Young's inequality. More precisely, we first apply Young's inequality to the estimates (32) and (33) so that the contributions due to $\partial_z^N W$ can be bounded by

$$0.00001c_N(\|\sqrt{\mu}(\partial_z - t\partial_x)\partial_z^N W\|_{L^2}^2 + \|(\mu(U')^2)^{\frac{1}{6}}\partial_z^N W\|_{L^2}^2)$$

and can thus be absorbed into the decay in estimate (31). Then, choosing c_N sufficiently small the remaining terms obtained in the application of Young's inequality can be absorbed into the decay by (30). This concludes the proof. $\qquad\square$

Acknowledgments

References

[BG81]   Steven J Barker and Douglas Gile. Experiments on heat-stabilized laminar boundary layers in water. *Journal of Fluid Mechanics*, 104:139–158, 1981.

[BGM17]   Jacob Bedrossian, Pierre Germain, and Nader Masmoudi. On the stability threshold for the 3d Couette flow in Sobolev regularity. *Annals of Mathematics*, pages 541–608, 2017.

[BM15]   Jacob Bedrossian and Nader Masmoudi. Inviscid damping and the asymptotic stability of planar shear flows in the 2d Euler equations. *Publications mathématiques de l'IHÉS*, 122(1):195–300, 2015.

[BMV16]   Jacob Bedrossian, Nader Masmoudi, and Vlad Vicol. Enhanced dissipation and inviscid damping in the inviscid limit of the Navier–Stokes equations near the two dimensional Couette flow. *Archive for Rational Mechanics and Analysis*, 219(3):1087–1159, 2016.

[BVW18]   Jacob Bedrossian, Vlad Vicol, and Fei Wang. The Sobolev stability threshold for 2D shear flows near Couette. *Journal of Nonlinear Science*, 28(6):2051–2075, 2018.

[Cra69]   Alex DD Craik. The stability of plane couette flow with viscosity stratification. *Journal of Fluid Mechanics*, 36(4):685–693, 1969.

[CZZ19]   Michele Coti Zelati and Christian Zillinger. On degenerate circular and shear flows: the point vortex and power law circular flows. *Communications in Partial Differential Equations*, 44(2):110–155, 2019.

[Dra02]   Philip G Drazin. *Introduction to hydrodynamic stability*, volume 32. Cambridge university press, 2002.

[DWZ20]   Wen Deng, Jiahong Wu, and Ping Zhang. Stability of couette flow for 2d Boussinesq system with vertical dissipation. *arXiv preprint arXiv:2004.09292*, 2020.

[EW15]   Tarek M Elgindi and Klaus Widmayer. Sharp decay estimates for an anisotropic linear semigroup and applications to the surface quasi-geostrophic and inviscid Boussinesq systems. *SIAM Journal on Mathematical Analysis*, 47(6):4672–4684, 2015.

[GS14]   Rama Govindarajan and Kirti Chandra Sahu. Instabilities in viscosity-stratified flow. *Annual review of fluid mechanics*, 46:331–353, 2014.

[HB87]   AP Hooper and WGC Boyd. Shear-flow instability due to a wall and a viscosity discontinuity at the interface. *Journal of Fluid Mechanics*, 179:201–225, 1987.

[Hei85]   Werner Heisenberg. Über stabilität und turbulenz von flüssigkeitsströmen. In *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pages 31–81. Springer, 1985.

[HL20]   Zihui He and Xian Liao. Solvability of the two-dimensional stationary incompressible in-
          homogeneous navier-stokes equations with variable viscosity coefficient. *arXiv preprint
          arXiv:2005.13277*, 2020.
[IJ18]   Alexandru Ionescu and Hao Jia. Inviscid damping near shear flows in a channel. *arXiv
          preprint arXiv:1808.04026*, 2018.
[Jia19]   Hao Jia. Linear inviscid damping in Gevrey spaces. *arXiv preprint arXiv:1904.01188*,
          2019.
[JRR84]   Daniel D Joseph, Michael Renardy, and Yuriko Renardy. Instability of the flow of
          two immiscible liquids with different viscosities in a pipe. *Journal of Fluid Mechanics*,
          141:309–317, 1984.
[Lin44]   CC Lin. On the stability of two-dimensional parallel flows. *Proceedings of the National
          Academy of Sciences of the United States of America*, 30(10):316, 1944.
[Lis20]   Kyle Liss. On the Sobolev stability threshold of 3d Couette flow in a uniform magnetic
          field. *Communications in Mathematical Physics*, pages 1–50, 2020.
[LX19]   Zhiwu Lin and Ming Xu. Metastability of kolmogorov flows and inviscid damping of
          shear flows. *Archive for Rational Mechanics and Analysis*, 231(3):1811–1852, 2019.
[LZ11]   Zhiwu Lin and Chongchun Zeng. Inviscid dynamical structures near Couette Flow.
          Archive for rational mechanics and analysis, 200(3):1075–1097, 2011.
[MSHZ20] Nader Masmoudi, Belkacem Said-Houari, and Weiren Zhao. Stability of Couette flow
          for 2d Boussinesq system without thermal diffusivity. *arXiv preprint arXiv:2010.01612*,
          2020.
[PV91]   Henry Power and Miguel Villegas. Weakly nonlinear instability of the flow of two
          immiscible liquids with different viscosities in a pipe. *Fluid dynamics research*, 7(5-
          6):215, 1991.
[TW19]   Lizheng Tao and Jiahong Wu. The 2d Boussinesq equations with vertical dissipation
          and linear stability of shear flows. *Journal of Differential Equations*, 267(3):1731–1747,
          2019.
[Wid18]   Klaus Widmayer. Convergence to stratified flow for an inviscid 3d Boussinesq system.
          Communications in Mathematical Sciences, 16(6):1713–1728, 2018.
[WZZ17]   D. Wei, Z. Zhang, and W. Zhao. Linear inviscid damping and enhanced dissipation for
          the Kolmogorov flow. *ArXiv e-prints*, November 2017.
[WZZ18]   Dongyi Wei, Zhifei Zhang, and Weiren Zhao. Linear inviscid damping for a class of
          monotone shear flow in Sobolev spaces. *Communications on Pure and Applied Math-
          ematics*, 71(4):617–687, 2018.
[Yih67]   Chia-Shun Yih. Instability due to viscosity stratification. *Journal of Fluid Mechanics*,
          27(2):337–352, 1967.
[YL18]   Jincheng Yang and Zhiwu Lin. Linear inviscid damping for Couette flow in stratified
          fluid. *Journal of Mathematical Fluid Mechanics*, 20(2):445–472, 2018.
[Zil17]   Christian Zillinger. Linear inviscid damping for monotone shear flows. *Trans. Amer.
          Math. Soc.*, 369(12):8799–8855, 2017.
[Zil21]   Christian Zillinger. Linear inviscid damping in sobolev and gevrey spaces. *Nonlinear
          Analysis*, 213:112492, 2021.

Karlsruhe Institute of Technology, Englerstrasse 2, 76131 Karlsruhe, Germany
Email address: `xian.liao@kit.edu`

Karlsruhe Institute of Technology, Englerstrasse 2, 76131 Karlsruhe, Germany
Email address: `christian.zillinger@kit.edu`