Hydrodynamic stability and mixing at high Reynolds number: MSRI Graduate Summer school 2015

Jacob Bedrossian

August 24, 2015

Contents

1 Introduction

1.1 Hydrodynamic stability: a bit of history and context

Hydrodynamic stability is one of the oldest fields of fluid mechanics, dating back to the 1800s and attracting the attention of many greats such as Reynolds, Stokes, Lord Kelvin, Lord Rayleigh, and many others. This field is concerned with understanding the stability of *laminar* flow configurations and, specifically, describing when and how these flows become unstable; modern applications naturally include techniques for inhibiting, triggering, or otherwise controlling these instabilities. Aside from these questions being theoretically natural, they are also of practical relevance and, indeed, hydrodynamic stability is one of the main pillars of applied fluid mechanics. As we will see, it will also provide us an interesting weakly nonlinear regime to study some fundamental processes in fluid mechanics which are difficult to get a handle on otherwise.

The term *Reynolds number* is due to Reynolds' hydrodynamic stability experiments flow through a pipe. For a cylindrical flow configuration $(r, \theta, z) \in [0, R] \times [0, 2\pi) \times \mathbb{R}$, one can verify that

$$u_z = \frac{AR^2}{2\nu}\left(1 - \left(\frac{r}{R}\right)^2\right), \quad u_\theta = 0, \quad u_r = 0, \quad p = -Az$$

is a solution to the 3D Navier-Stokes equations; this configuration represents pressure driven flow in a pipe. What Reynolds did in [Rey83] was force fluid through a pipe while varying the various parameters and observed that for small **Re** this laminar flow is stable but that for sufficiently large **Re** he observed instability and spontaneous transition to turbulence. As theory and experiment progressed it became clear that this transition to turbulence was occurring in systems which are *spectrally stable* (see below)– this kind of transition is called *subcritical transition* or *by-pass transition*. This raised many interesting questions, especially since linear stability and nonlinear stability were almost always assumed to go together in applied mathematics at the time.

Since the work of Reynolds, countless experiments and computer simulations have been done on hydrodynamic stability problems in both 2D and 3D and the subject is both vast and rich; see e.g. the texts [DR81, Yag12, SH01, Dra02] and the references therein. We will discuss some of the many facets of the theory as we go along.

1.2 Notions of stability

Let N be a given (possibly nonlinear) operator N and suppose we have the abstract evolution equation

$$\partial_t f = N[f], \tag{1.1a}$$
$$f(0) = f_{in}, \tag{1.1b}$$

with the equilibrium point $N[f_0] = 0$. We will not trouble ourselves with well-posedness issues here, so we can assume the abstract system (1.1) is well-posed in whatever spaces we care about.

We will not be discussing spectral theory much since in fluid mechanics, most of the linear operators are non-normal, which means $AA^* \neq A^*A$. The spectral theorem shows that it is reasonable to think of this as the correct generalization of "non-diagonalizable" from basic linear algebra; see e.g. [RS79]. In particular, for non-normal operators, the spectrum of A may not tell us enough information about the linear evolution $\partial_t f = Af$ for us to really "understand" the behavior. We will see some examples for finite dimensional linear ODEs below. Recall that the definition of spectrum for unbounded operators is the following; see e.g. [EN00, RS79].

Definition 1. Let H be a Hilbert space and $A : D(A) \to H$ be a closed operator with domain $D(A) \subset H$[1]. The *resolvent set* $\rho(A) \subset \mathbb{C}$ is given by

$$\rho(A) = \{\lambda \in \mathbb{C} : A - \lambda I \text{ is invertible on } D(A) \to H\}.$$

The *spectrum* is then defined as $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

Remark 1. In infinite dimensions, not everything in the spectrum corresponds to eigenvalues. We will be studying a very important linear operator $L = y\partial_x f$ on $\mathbb{T} \times \mathbb{R}$, which is skew-adjoint on $L^2(\mathbb{T} \times \mathbb{R})$.

Exercise 1.1. Verify that all eigenfunctions in L^2 of L are independent of x and correspond to the zero eigenvalue. That is, prove that if $f \in L^2$ and $y\partial_x f = \lambda f$, then $f(x, y) = \phi(y)$ for some ϕ (up to redefinition on a set of measure zero).

We see from Exercise 1.1 that $0 \in \sigma(L)$. However, let us verify that $i\lambda \in \sigma(L)$ for all $\lambda \in \mathbb{R}$, that is, the entire imaginary axis is also in the spectrum. For this, we need to ensure that the $i\lambda$ is not in the resolvent by proving that $L - i\lambda I$ is not invertible. We know already that we cannot do this by exhibiting an eigenfunction, however, we can do this by exhibiting a sequence of functions f_n such that

$$(L - i\lambda)f_n \to 0$$

but that $\|f_n\|_{L^2} \approx 1$.

Let ϕ be a smooth, compactly supported bump function. Fix a $k \in \mathbb{R}$ and consider now the sequence

$$f_n(x, y) = e^{ikx} n^{1/2} \phi\left(n(y - \frac{\lambda}{k})\right).$$

We can immediately check that $\|f_n\|_{L^2} = \sqrt{2\pi}\, \|\phi\|_{L^2}$, which is a fixed number. We can directly compute that

$$(L - i\lambda)f_n = (iky - i\lambda)f_n$$

However

$$\begin{aligned}
\|(iky - i\lambda)f_n\|_{L^2}^2 &= \int n\,|\lambda - ky|^2\, \phi^2\left(n(y - \lambda/k)\right) dxdy \\
&= \int n\,|kz|^2\, \phi^2\left(nz\right) dxdz \\
&= n^{-2} \int |kv|^2\, \phi^2(v) dxdv.
\end{aligned}$$

We see that this goes to zero as $n \to 0$ which shows that $L - i\lambda I$ is not invertible. Therefore, $i\lambda \in \sigma(L)$.

Regardless of all the potential issues with using the spectrum to assess stability, it is still a very useful notion, and one that is still the most common used by engineers and physicists.

[1] Recall that for unbounded operators, for example $A = -\Delta$ on $H = L^2$, we cannot make sense of the operator over the entire Hilbert space, hence we define the domain $D(A)$ as the subset for which it makes sense to consider $A : D(A) \to L^2$ (for example, if $A = -\Delta$ and $L^2(\mathbb{T}^2)$ then we should take $D(A) = H^2(\mathbb{T}^2)$).

Definition 2 (Spectral stability)**.** Let $Lf = DN[f_0]f$ be the linearization of N. The equilibrium f_0 is called *spectrally stable* in a Hilbert space X if $\sigma(L) \cap \{c \in \mathbb{C} : \mathrm{Re}\, c > 0\} = \emptyset$, where $\sigma(L)$ denotes the spectrum of L in X. The evolution is called *spectrally unstable* if $\sigma(L) \cap \{c \in \mathbb{C} : \mathrm{Re}\, c > 0\} \neq \emptyset$.

Remark 2. Often times, if system is spectrally stable but the spectrum intersects the imaginary axis, the terminology *neutrally stable* is used.

A different notion of stability, which is usually attributed to Lyapunov, is the following. When a mathematician says "stable", this is normally what he or she means.

Definition 3 (Lyapunov/nonlinear stability)**.** Given two Banach spaces X and Y (usually the same), the equilibrium f_0 is called *stable* (from X to Y) if for all $\epsilon > 0$, there exists a $\delta > 0$ such that if

$$\|f_{in} - f_0\|_X < \delta.$$

then for all $t > 0$, the solution to (1.1) satisfies

$$\|f(t) - f_0\|_Y < \epsilon.$$

We say f_0 is *unstable* if it is not stable.

Remark 3. In many applications, $X = Y$, however in some applications this is too much to ask and we instead take X to be a smaller space than Y (e.g. the restriction on the initial data is stronger than the norm in which stability is deduced).

Even for *finite dimensional* ***linear*** *systems*, these two definitions of stability are *not* equivalent.

Exercise 1.2. Consider a finite dimensional ODE $\partial_t X = AX$ for a given fixed matrix A (and the equilibrium is of course $f_0 = 0$). Prove that if A is diagonalizable then spectral stability implies Lyapunuv stability. However, prove that if A is not diagonalizable, then spectral stability does *not* necessarily imply Lyapunov stability.

Exercise 1.3. Consider a finite dimensional ODE $\partial_t X = AX$ for a given fixed matrix A. Prove that spectral instability always implies instability in the sense of Lyapunov.

Exercise 1.4. Consider a nonlinear finite dimensional ODE $\partial_t X = F(X)$ with equilibrium $F(X_0) = 0$. Show that if $\nabla F(X_0)$ is diagonalizable and $\sigma(\nabla F(X_0)) \subset \{c \in \mathbb{C} : \mathbf{Re}\, c < 0\}$, then the equilibrium X_0 is Lyapunov stable. Give an example that shows nonlinear stability can fail if we only assume $\sigma(\nabla F(X_0)) \subset \{c \in \mathbb{C} : \mathbf{Re}\, c \leq 0\}$ (hence, even for diagonalizable systems, neutral stability does not imply Lyapunov stability).

A good rule of thumb is that spectral stability is in some sense the *weakest* kind of stability but spectral *instability* tends to be a pretty strong kind of instability, and in many settings, spectral instability is enough to deduce nonlinear instability.

We would also like to mention that in infinite dimensions, stability depends a lot on the norms that you are measuring. For example, consider the (very relevant) PDE:

$$\partial_t f + y\partial_x f = 0 \tag{1.2a}$$
$$f(0) = f_{in}. \tag{1.2b}$$

Exercise 1.5. Prove that the the equilibrium $f \equiv 0$ for (1.2) is spectrally stable in L^2, Lyapunov stable in L^2, but Lyapunov unstable in H^1 (in fact any H^s with $s > 0$).

2 2D inviscid planar shear flows and Rayleigh's theorem

We begin our study of hydrodynamic stability of inviscid shear flows, the simplest of all flows:

$$u(x,y) = \begin{pmatrix} U(y) \\ 0 \end{pmatrix},$$

where we will assume that $U(y) \in C^\infty$. We want to investigate the stability of this configuration in the 2D Euler equations to start with. Later we will consider 3D and finite Reynolds number generalizations.

One of the first works on the theoretical side of hydrodynamic stability was that of Lord Rayleigh who in 1880 attempted to determine when a certain inviscid shear flow is stable or not in the 2D Euler equations [Ray80]. He was interested in *spectral* stability, that is, to determine when there does or does not exist unstable eigenvalues to the linearized problem. For the following, we will mostly follow the treatment in [Dra02]. Lord Rayleigh derived a *necessary* condition for spectral instability of 2D shear flows by what is sometimes referred to as the normal mode method. This is just a name for the method of looking for sets of orthogonal eigenfunctions and eigenvalues for the linearized problem. Rayleigh considered the linearization of 2D Euler about a given shear flow $(U(y), 0)$:

$$u_t + U(y)\partial_x u + \begin{pmatrix} u_2 U'(y) \\ 0 \end{pmatrix} = -\nabla p \tag{2.1a}$$

$$-\Delta p = 2U'\partial_x u_2 \tag{2.1b}$$

$$u \cdot n = 0. \tag{2.1c}$$

say for $y \in [-1,1]$, a bounded channel.

Exercise 2.1. Write (2.1) as $\partial_t u + Lu = 0$ for a linear operator L and verify that $L^*L \neq LL^*$ (and hence L is non-normal).

In vorticity form, this becomes

$$\omega_t + U(y)\partial_x\omega - U''\partial_x\psi = 0$$
$$\Delta\psi = \omega.$$

Notice that the zero-th Fourier mode in x $\omega_0(t,y) = \hat{\omega}(t,0,y) = \frac{1}{2\pi}\int \omega(t,x,y)dx$ is conserved by the evolution, so we may assume without loss of generality that $\int \omega_{in}(x,y)dx = 0$ and hence that remains true forward (and backward) in time (note that this conservation law is not true for the nonlinear problem). We re-write on the streamfunction:

$$\left(\partial_t + U(y)\partial_x\right)\Delta\psi - U''(y)\partial_x\psi = 0. \tag{2.2}$$

Note that if we are in a bounded channel, this comes with boundary conditions, in particular $\nabla\psi \cdot \tau = 0$ on the upper and lower boundaries.

The problem is still translation invariant with respect to x so we can use the Fourier transform in this direction. Therefore, we look for a solution of the form:

$$\psi(t,x,y) = \phi(y)e^{i\alpha(x-ct)}, \tag{2.3}$$

where $\alpha \in \mathbb{R}$ is non-zero and $c \in \mathbb{C}$. Notice that if (2.3) is going to satisfy the boundary condition $\nabla\psi \cdot \tau = 0$, we are going to need to impose $\phi|_{\pm 1} = 0$. Upon substitution of (2.3) into (2.2) we get *Rayleigh's problem*:

$$(U(y) - c)\left(\phi'' - \alpha^2\phi\right) - U''(y)\phi = 0. \tag{2.4}$$

We will have an unstable mode if we can find ϕ, c and α which solve (2.4) with $\mathrm{Im}c \neq 0$ (noting that if α, c, ϕ is a solution, then so is $-\alpha, c, \phi$). The horrible degeneracy at the point $U(y) = c$ is known as the "critical layer" in the applied literature; this is only possible when $c \in \mathbb{R}$ and so is not relevant when we are looking for unstablities (it *is* relevant when looking for neutrally stable modes). The degeneracy is connected with the continuous spectrum which lies on the imaginary axis. However, let us not be too concerned with this (yet).

Theorem 2.1 (Rayleigh [Ray80])**.** *Consider the 2D Euler equations linearized around the shear flow $(U(y), 0)$ with $y \in [-1, 1]$ (or $y \in \mathbb{R}$) given in* (2.1)*. If the linearized 2D Euler equations have an unstable eigenmode in H^1, then U'' must vanish at least at one point (hence, any flow without an inflection point is spectrally stable).*

Proof. The proof proceeds by proving that if there is an unstable eigenvalue, then necessarily there is an inflection point somewhere in the flow. This can be proved by an energy-type estimate on (2.4). If there is an unstable eigenvalue, then there is a solution to (2.4) with $\alpha \geq 0$, $\mathrm{Im}c > 0$, and ϕ non-trivial. Dividing (2.4) by $U - c$ (note that it is non-vanishing because $\mathrm{Im}c \neq 0$ and $U \in \mathbb{R}$) and multiplying by $\overline{\phi}$ an integrating by parts gives (note the boundary terms vanish due to $\phi|_{y=\pm 1} = 0$),

$$\int \left|\phi'\right|^2 + \alpha^2 \left|\phi\right|^2 dy + \int \frac{U''(y)}{U(y) - c} \left|\phi\right|^2 dy = 0.$$

Taking the imaginary part leaves us

$$\mathrm{Im}\, c \int \frac{U''(y)}{\left|U(y) - c\right|^2} \left|\phi\right|^2 dy = 0.$$

By assumption $\mathrm{Im}\, c > 0$, therefore we have,

$$\int \frac{U''(y)}{\left|U(y) - c\right|^2} \left|\phi\right|^2 dy = 0,$$

is a necessary condition for instability. This requires that $U''(y) = 0$ in at least one place, and hence the theorem follows. ☐

Remark 4. By elliptic regularity, if $\mathrm{Im}\, c \neq 0$, any H^1 solution to (2.4) will be C^∞.

There are sharper spectral stability conditions [Fjo50], however, to our knowledge, there is still no known sharp or nearly sharp condition for spectral stability. For example, the *Couette flow* $u = (y, 0)$ is linearly stable (as we will see) and so are flows that are nearby in a certain sense (see below).

As it turns out, the result of Lord Rayleigh unfortunately extends to 3D via a result known as *Squire's theorem* [Squ33]. This theorem shows that if the 3D planar shear flow has unstable eigenvalues, then so does the 2D problem, and hence if one is looking for eigenvalue instabilities, then studying 2D is sufficient in the sense that any planar shear flow which is spectrally stable in 2D is also spectrally stable in 3D. See e.g. [Dra02] for a proof (it is not too hard actually, you could do it as an exercise). I say "unfortunately" because this theorem is *horrifically* misleading as it suggests all interesting aspects of hydrodynamic stability can be found in 2D equations – this is *extremely false*. In fact, it can be shown[2] that pretty much *every* non-trivial shear flow is unstable in the sense of Lyapunov for the linearized 3D Euler equations in $(x, y, z) \in \mathbb{T} \times \mathbb{R}^2$!

Theorem 2.2 (Squire's theorem [Squ33]). *Consider the 3D linearized Euler equations near a planar shear flow $(U(y), 0, 0)$ between parallel plates $[-1, 1]$. If there is a 3D unstable mode, then there is a 2D unstable mode with a faster or equally as fast growth rate. As a consequence, any planar shear flow which is spectrally stable in the linear 2D Euler equations is spectrally stable also in the linear 3D Euler equations.*

The method of normal modes sheds light on many classical problems in hydrodynamic stability (and many other questions in hundreds of other fields), such as understanding the Rayleigh-Benard convection cells, the Kelvin-Helmholtz instability, the Rayleigh-Taylor instability, the instability of jets, and more; see e.g. [DR81, Dra02]. Unfortunately (its quite fortunate in another sense) there are two major shortcomings of this method for hydrodynamic stability at high Reynolds number: (A) the linearized inviscid problems usually do not have a purely discrete spectrum and so looking for eigenfunctions is not sufficient[3] and (B) the linearized problems are usually not normal operators and so knowing the spectrum is usually not sufficient. Recall a linear operator is normal if $AA^\star = A^\star A$ in a suitable sense (let us not dwell on the finer points of spectral and unbounded operator theory here). It is precisely this class of operators such that nice versions of the spectral mapping theorems hold. Recall, spectral mapping theorems roughly tell you that if you know the spectrum of A (and you have several technical conditions satisfied) then you not only know the spectrum of e^{tA} but you also know the norm of this semigroup (see e.g. [EN00] for more precise details). This is not true of non-normal operators: if A is not normal, than the norm of e^{tA} can vary wildly from what its spectrum suggests [TE05]. The simplest example is the ODE (which actually is pretty relevant to hydrodynamic stability as we will see):

$$\partial_t X = \begin{pmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{pmatrix}.$$

The eigenvalues of the matrix are $-\epsilon < 0$, hence if the operator were normal, the solutions would be decaying and the operator norm of e^{tA} would be uniformly bounded in ϵ. However, instead we have $\|X(t)\| \approx \|X_0\| \langle t \rangle e^{-\epsilon t} \lesssim \epsilon^{-1} \|X_0\|$, which shows that solutions can undergo a large transient growth before eventually decaying. It was Orr [Orr07] who, to our knowledge, first pointed out the potentially pivotal importance of non-normal transient growth in fluid mechanics. We will see several examples of linear (and nonlinear) problems in fluid mechanics which can undergo a large transient growth like the above.

3 Arnold's nonlinear stability theorem for shear flows in a channel

Due to the non-normal nature of the linearizations, one can worry that *nonlinear* stability at high or infinite Reynolds number can be hard to come by. This is because transient linear growth could carry small perturbations out of the linear regime and into the fully nonlinear regime, triggering a secondary instability, as first suggested by Orr [Orr07] (we will return to this idea later, which is by now classical in applied fluid mechanics, see e.g. [TTRD93, RSBH98, SH01, TE05, Yag12]).

[2] See future lectures, or do it as an exercise by taking x averages of the 3D analogue of (2.1) and considering the PDE that is left over.

[3] The viscous problem may have a discrete spectrum (depending on the boundary conditions or functions spaces) however in the inviscid limit it will develop a continuous spectrum.

However, there is a beautiful and classical result of Arnold [Arn65] which uses a variational method to provide a simple proof of Lyapunov stability (in the H^1 norm of u) for a class of 2D shear flows in the nonlinear Euler equations. The basic ideas also generalize to many other situations, see e.g. the review by [HMRW85] and the references therein.

To make life simple, let us explain the idea of Arnold in the channel $(x, y) \in \mathbb{T} \times [-1, 1]$ with no-penetration boundaries on the top and bottom edges; much more generality is possible [HMRW85]. Notice that this domain is *not simply connected* which means we have to be a bit careful about the vorticity-streamfunction formulation of the equations.

3.0.1 Conservation laws and vorticity-streamfunction formulation in the channel $\mathbb{T} \times [-1, 1]$

We consider smooth solutions to the 2D Euler equations

$$\partial_t u + u \cdot \nabla u = -\nabla p \tag{3.1a}$$

$$\nabla \cdot u = 0 \tag{3.1b}$$

$$u \cdot n|_{y=\pm 1} = u_2|_{y=\pm 1} = 0. \tag{3.1c}$$

The 2D incompressible Euler equations can be thought of formally as a Hamiltonian system for the kinetic energy if interpreted correctly (also due to Arnold originally [Arn66]). Recall that the energy is

$$E[u] = \frac{1}{2} \int |u|^2 \, dx.$$

Since we are in 2D, for any smooth function Φ, the associated *Casimir*,

$$C_\Phi[\omega] = \int \Phi(\omega) dx$$

is conserved, where $\omega = \partial_x u_2 - \partial_y u_1$ is the (scalar) vorticity. These Casimirs provide a useful infinite set of conservation laws, which is one of the reasons that the 2D Euler equations are very different than the 3D Euler equations.

Next, due to the Kelvin circulation theorem, the circulation around every connected component of the boundary is constant. That is,

$$\Gamma_i = \int_{\partial D_i} u(t) \cdot ds = \int_{\partial D_i} u(0) \cdot ds.$$

This is a general fact, in this case, it applies to the lines $y = \pm 1$. Due to the shape of the domain (in particular, the translation invariance in x), the total x momentum is still conserved. That is,

$$M_x = \int_{-1}^{1} \int u_1(t, x, y) dx dy = \int_{-1}^{1} \int u_1(0, x, y) dx dy.$$

Before continuing, let us briefly discuss the vorticity stream-function formulation in the domain $\mathbb{T} \times [-1, 1]$. Let $(0, 0, \omega) = \nabla \times u$ be the scalar vorticity. Next, consider looking for a streamfunction ψ which satisfies $\Delta \psi = \omega$ and $u = \nabla^\perp \psi$. In general, we know that since $u \cdot n|_{\partial D_i} = \nabla^\perp \psi \cdot n|_{\partial D_i} = \nabla \psi \cdot \tau|_{\partial D_i}$, ψ is constant on each connected component of the boundary. The streamfunction is determined only up to a constant, but since there are two disconnected pieces of the boundary, the difference between the constants associated with each boundary is not determined. In this case it is easy to see what the difference is. Consider taking x averages of u_1:

$$\begin{aligned} u_1(t,x,y) &= -\partial_y \psi(t,x,y) \\ \langle u_1 \rangle_x(t,y) &= -\partial_y \langle \psi \rangle_x(t,y). \end{aligned}$$

Integrating this in y gives

$$\frac{1}{2\pi} M_x = \frac{1}{2\pi} \int u_1(t,x,y)dxdy = \langle \psi \rangle_x(t,-1) - \langle \psi \rangle_x(t,1).$$

The LHS is a fixed number in time (the mean flow across the torus) and so the difference between the value of the streamfunction at the top and bottom is this constant. Without loss of generality we may as well take

$$\begin{aligned} \langle \psi \rangle_x(t,-1) &= 0 \\ \langle \psi \rangle_x(t,1) &= -\frac{1}{2\pi} M_x. \end{aligned}$$

Hence, to find the streamfunction given the vorticity, we can solve the Dirichlet problem

$$\begin{aligned} \Delta \psi &= \omega \\ \psi(x,-1) &= 0 \\ \psi(x,1) &= -\frac{1}{2\pi} M_x. \end{aligned}$$

Finally, let us note that when viewed in terms of the vorticity and streamfunction, (ω, ψ) the energy becomes

$$\begin{aligned} E[\omega] &= \frac{1}{2} \int \left| \nabla^{\perp} \psi \right|^2 dx \\ &= \frac{1}{2} \int |\nabla \psi|^2 dx \\ &= \frac{1}{2} \sum_i \int_{\partial D_i} \psi \nabla \psi \cdot n ds - \frac{1}{2} \int \psi \omega dx. \end{aligned}$$

Note that because ψ is constant along the boundaries, this becomes

$$E[\omega] = \frac{1}{2} \sum_i \psi|_{\partial D_i} \int_{\partial D_i} u \cdot ds - \frac{1}{2} \int \psi \omega dx.$$

Hence, $-\frac{1}{2} \int \psi \omega dx$ is conserved, since the energy, the circulations at the top and bottom, and the value of $\psi|_{\partial D_i}$ are all individually conserved.

3.0.2 Variational stability

The general scheme of Arnold is to find suitable Φ and a_i such that the equilibrium, ω_E, is a critical point of the conserved energy functional

$$H_C[\omega] = \frac{1}{2} \int |u|^2 dx + \int \Phi(\omega) dx + \sum_{\partial D_i} a_i \Gamma_i$$

and that H_C is locally convex and suitably coercive with respect to the L^2 norms of the velocity and vorticity in a neighborhood of the critical point. Recall, coercive with respect to a norm means that control on the energy functional controls the norm; see below.

First let us compute the first variation of H_C at ω_E due to perturbations ω, ψ. The perturbations are assumed to preserve the mean flow rate, and therefore $\psi(x,-1) = \psi(x,1) = 0$. Computing the first variation and integrating by parts (using $u_E = \nabla^\perp \psi_E$ and $u = \nabla^\perp \psi$) gives

$$\begin{aligned} DH_C[\omega_E]\omega &= \int u_E \cdot u dx + \int \Phi'(\omega_E)\omega dx + \sum_{\partial D_i} a_i \int_{\partial D_i} u \cdot ds \\ &= \int \nabla^\perp \psi_E \cdot \nabla^\perp \psi dx + \int \Phi'(\omega_E)\omega dx + \sum_{\partial D_i} a_i \int_{\partial D_i} u \cdot ds \\ &= \int \nabla \psi_E \cdot \nabla \psi dx + \int \Phi'(\omega_E)\omega dx + \sum_{\partial D_i} a_i \int_{\partial D_i} u \cdot ds \\ &= \int \psi_E \nabla \psi \cdot n ds - \int \psi_E \omega dx + \int \Phi'(\omega_E)\omega dx + \sum_{\partial D_i} a_i \int_{\partial D_i} u \cdot ds \\ &= -\int \psi_E \omega dx + \int \Phi'(\omega_E)\omega dx + \sum_{\partial D_i} (a_i - \psi_E|_{\partial D_i}) \int_{\partial D_i} u \cdot ds. \end{aligned}$$

In the last line we used that $\nabla\psi \cdot n = u \cdot \tau$ and that $\psi_E|_{\partial D_i}$ is constant along the boundaries. Hence, ω_E will be a critical point of H_C as soon as

$$\psi_E = \Phi'(\omega_E), \tag{3.2a}$$

$$a_i = \psi_E|_{\partial D_i}. \tag{3.2b}$$

Suppose that we have a functional relationship $\psi_E = \Psi(\omega_E)$. In this case, it suffices to have

$$\Psi(\omega_E) = \Phi'(\omega_E),$$

and hence we can then take, for some constant λ,

$$\Phi(t) = \int_0^t \Psi(\tau) d\tau + \lambda.$$

We see that the equilibrium determines Φ; now it suffices to see what kind of equillibria are such that H_C is convex near ω_E. Let $\omega_E(x) + \omega(t,x)$ solve the full nonlinear 2D Euler equations in the periodic strip. Consider now the conserved functional (the first term is conserved, last term is zero by construction, and the second to last term is a constant in time):

$$F[\omega] = H_C[\omega + \omega_E] - H_C[\omega_E] - DH_C[\omega_E]\omega.$$

Computing this out gives

$$F[\omega] = \int \frac{1}{2} |u|^2 dx + \int \Phi(\omega_E + \omega) - \Phi(\omega_E) - \Phi'(\omega_E)\omega dx.$$

If there exists a constant $\delta > 0$ such that $\Phi'' \geq \delta$, then Φ is uniformly convex and therefore

$$\int \Phi(\omega_E + \omega) - \Phi(\omega_E) - \Phi'(\omega_E)\omega dx \geq \frac{\delta}{2} \int |\omega|^2 dx.$$

In this case we have

$$F[\omega(0)] = F[\omega(t)] \geq \frac{1}{2} \left\| u(t) \right\|_{L^2}^2 + \frac{\delta}{2} \left\| \omega(t) \right\|_{L^2}^2 .$$

This means that the kinetic energy and the enstrophy of the perturbation are uniformly bounded. If there is a constant $C > 0$ such that $\Phi'' \leq C$ then we also get

$$F[\omega(0)] \leq \frac{1}{2} \left\| u(0) \right\|_{L^2}^2 + \frac{C}{2} \left\| \omega(0) \right\|_{L^2}^2 ,$$

which will imply the desired global nonlinear stability in L^2 of the velocity and vorticity (sometimes called the "energy norm" and "enstrophy norm"). Due to the divergence free condition, this is equivalent to the H^1 norm on the velocity.

The general procedure can be extended to some more general domains and other equillibria which satisfy some kind of functional relationship $\psi_E = \Psi(\omega_E)$ for some Ψ, however, let us continue to just think about what this means for shear flows. In the case of a shear flow $u_E = (U(y), 0)$,

$$\begin{aligned} \omega_E(y) &= -U'(y) \\ U(y) &= -\partial_y \psi(y). \end{aligned}$$

Therefore, taking y derivatives of (3.2a) gives

$$-U(y) = -\Phi''(\omega_E(y))U''(y),$$

or

$$\frac{U(y)}{U''(y)} = \Phi''(\omega_E(y)).$$

Putting everything together, we have proved the following beautiful theorem.

Theorem 3.1 (Arnold's nonlinear stability)**.** *Let $u_E = (U(y), 0)$ be a shear flow on $\mathbb{T} \times [-1, 1]$ such that U is smooth and such that there is some smooth function $\Psi : \mathbb{R} \to \mathbb{R}$ such that $\psi_E = \Psi(\omega_E)$ (where $\omega_E = -U'$ is the vorticity). Suppose (up to Galilean invariance) that there exists constants $\infty > C > \delta > 0$ which satisfy*

$$C > \frac{U(y)}{U''(y)} \geq \delta.$$

Then the equilibrium u_E is globally nonlinearly stable for the 2D incompressible Euler equations in the energy and enstrophy norms if we restrict to perturbations which conserve M_x the mean zero momentum. That is, if $u + u_E$ solves the 2D Euler equations in $\mathbb{T} \times [-1, 1]$ and $\omega = \partial_x u_2 - \partial_y u_1$ and satisfies $\int_{\mathbb{T}\times\mathbb{R}} u_1 dx = 0$, then there holds uniformly forward and backward in time (note that it is $u + u_E$ which solves the 2D Euler equations):

$$\left\| u(t) \right\|_{L^2}^2 + \left\| \omega(t) \right\|_{L^2}^2 \lesssim \frac{C}{\delta} \left(\left\| u(0) \right\|_{L^2}^2 + \left\| \omega(0) \right\|_{L^2}^2 \right) . \tag{3.3}$$

Remark 5. Due to the divergence free condition, (3.3) is equivalent to

$$\left\| u(t) \right\|_{H^1} \lesssim \frac{C}{\delta} \left\| u(0) \right\|_{H^1} .$$

Remark 6. The above theorem proves, for example, that the equilibrium $U(y) = 2 + y^2$ is nonlinearly stable in energy and enstrophy norms in the 2D Euler equations.

Remark 7. We will see later that it would be unreasonable to expect L^2 stability at the level of the velocity only, that is, we will NOT be seeing any inequalities of the kind:

$$\|u(t)\|_{L^2}^2 \lesssim \|u(0)\|_{L^2}^2 .$$

It will also be clear later that it is unreasonable to expect any kind of H^s stability for $s > 1$, that is, we will NOT be seeing any inequalities of the following kind for $s > 1$:

$$\|u(t)\|_{H^s}^2 \lesssim \|u(0)\|_{H^s}^2 .$$

In this sense, the choice of H^1 in Arnold's theorem is very natural and was dictated purely by the variational structure of the equations.

Just to recap: the basic idea is to take advantage of the large numbers of conserved quantities by finding one for which the equilibrium is a local minimizer and satisfies some sort of convexity property. In some cases this is easier said than done, but in other cases, like the above, its not so bad. Much more general results are possible, in particular, one can consider more general equillibria and also more general stability criteria even for shear flows; see [HMRW85] and the references therein. Moreover, the general idea and variations thereof applies to a very wide variety of applications throughout plasma physics, atmospheric dynamics, and galaxy dynamics, to name a few.

As a last comment, Theorem 3.1 already tells us a lot, but on the other hand, it doesn't tell us about the behavior of higher norms. In this sense it doesn't settle certain questions about the actual dynamics of the solution: do solutions oscillate around in periodic or quasi-periodic orbits? do solutions develop all kinds of crazy small scales rapidly and become increasingly turbulent at the small scales? Do solutions settle back to shear flows in one way or another? The question of the long-time dynamics is in general very poorly understood – and here we do not mean we just cannot prove things, it is poorly understood even in the way *physicists* mean "understand". Naturally, it is this question we will be focusing on for the remainder of the course.

4 Mixing and dissipation in passive scalar flows at high Péclet number

Previously, we were mainly concerned with deducing spectral stability or nonlinear stability for planar shear flows. However, we neglected entirely the question of what the actual dynamics look like, which are far more interesting than the previous discussion makes it sound. One of the main dynamics we neglected to discuss was *mixing*. Here we will begin the discussion of these kinds of dynamics by first focusing on passive scalar flows, rather than the linearized fluid equations.

The mixing and dissipation of passive scalars in a given incompressible velocity field is given by the linear equation

$$\partial_t f + u \cdot \nabla f = \kappa \Delta f \tag{4.1a}$$
$$f(0) = f_{in}. \tag{4.1b}$$

for a scalar f, a given velocity field $u(t,x)$ with $\nabla \cdot u = 0$ and a diffusivity $\kappa > 0$. In these lectures we will not be concerned with regularity and well-posedness issues (not in the traditional sense anyway) so it will suffice to assume $u \in C^\infty_{t,x}$ and $f_{in} \in C^\infty_{t,x}$.

The advection diffusion equation is a classical and very important problem of practical and theoretical interest and certainly deserves to be studied in its own right. The hope is then that understanding certain aspects of this problem will also tell us something about hydrodynamic stability.

To non-dimensionalize we can replace (since the equation is linear we don't really need to rescale f)

$$f^\star(t,x) = f\left(\frac{t}{LU^{-1}}, \frac{x}{L}\right)$$
$$u^\star(t,x) = \frac{1}{U} u\left(\frac{t}{LU^{-1}}, \frac{x}{L}\right)$$

and we have

$$\partial_t f^\star + u^\star \cdot \nabla f^\star = \frac{\kappa}{LU} \Delta f.$$

The dimensionless number in front of the Δ is called the (inverse) *Péclet number*,

$$Pe = \frac{UL}{\kappa}.$$

It is a ratio of the time-scale of advective transport to the diffusive transport.

4.1 Passive scalar in Couette flow

Let us begin with the simplest of all examples: the planar Couette flow (dropping the $\star$'s and using $\kappa = Pe^{-1}$ as the inverse Peclet number)

$$\partial_t f + y \partial_x f = \kappa \Delta f \tag{4.2a}$$
$$f(0) = f_{in}. \tag{4.2b}$$

We will take periodic boundary conditions in x and infinite in y, so we have our problem on a cylinder $(x,y) \in \mathbb{T} \times \mathbb{R}$ (we could also consider in 3D or higher but nothing is different for passive scalars). The problem (4.2) was first solved by Lord Kelvin in 1887 [Kel87]. First, let us consider the case $\kappa = 0$. In this case, the solution is just

$$f(t, x, y) = f_{in}(x - ty, y).$$

Taking the Fourier transform gives the following:

$$\begin{aligned}\hat{f}(t, k, \eta) &= \frac{1}{2\pi} \int e^{-ikx - i\eta y} f_{in}(x - ty, y) dx dy \\ &= \frac{1}{2\pi} \int e^{-ikx - ikty - i\eta y} f_{in}(x, y) dx dy \\ &= \hat{f}_{in}(k, \eta + kt).\end{aligned}$$

For each k this is a linear-in-time transfer of information to high frequencies. This implies the lack of compactness in L^2 and the weak convergence back to equilibrium:

Exercise 4.1. For $f_{in} \in L^2$, prove that if f solves (4.2) with $f(0) = f_{in}$, then $f(t) \rightharpoonup \langle f_{in} \rangle_x$ in L^2. Show that this convergence is only strong if $f(t) = \langle f_{in} \rangle_x$ for all t. Similarly, prove that if $f_{in} \neq \langle f_{in} \rangle_x$ and $f_{in} \in H^n$, then $\|f(t)\|_{H^n} \approx \langle t \rangle^n \|f_{in}\|_{H^n}$ (we are denoting $\langle t \rangle = (1 + |t|^2)^{1/2}$).

The phenomenon of weak convergence despite the fact that we are on a compact set (sort of) is usually called *mixing*. Draw a picture or two of the Couette flow evolution to convince yourself that the linear evolution is not so dissimilar from some of the fundamental processes that take place when you stir milk into coffee.

Notice that the behavior in Exercise 4.1 is not possible in finite dimensional Hamiltonian systems. For the rest of the lectures, we will denote

$$\begin{aligned}|k, \eta| &= |k| + |\eta| \\ \langle k, \eta \rangle &= (1 + |k, \eta|^2)^{1/2}.\end{aligned}$$

We get the decay in negative Sobolev norms, at a price. For all $s \geq 0$,

$$\begin{aligned}\|f - \langle f \rangle_x\|_{H^{-s}} &\lesssim \sum_{k \neq 0} \int \frac{1}{\langle k, \eta \rangle^{2s}} \left| \hat{f}_{in}(k, \eta + kt) \right|^2 d\eta \\ &\lesssim \sum_{k \neq 0} \int \frac{1}{\langle k, \eta \rangle^{2s} \langle \eta + kt \rangle^{2s}} \left| \langle \eta + kt \rangle^{2s} \hat{f}_{in}(k, \eta + kt) \right|^2 d\eta \\ &\lesssim \frac{1}{\langle t \rangle^s} \|f_{in} - \langle f_{in} \rangle_x\|_{H^s}.\end{aligned}$$

One can view this as a more quantitative estimate on the weak convergence. We will re-visit the loss of regularity in this formula at length later. Notice that if we take the Fourier transform of $\partial_t f + y \partial_x f = 0$ we get

$$\partial_t \hat{f} - k \partial_\eta \hat{f} = 0,$$

which is still a shear flow. This is an important point: mixing in Couette flow is transport to infinity *in frequency.*

Consider now the diffusive case, for $\kappa > 0$

$$\partial_t f + y \partial_x f = \kappa \Delta f.$$

In this case, Lord Kelvin, in [Kel87], defined the variables $X = x - ty$ and $g(t, X, y) = f(t, X + ty, y)$, which then solves

$$\partial_t g = \kappa \Delta_L g$$
$$\Delta_L = \partial_{XX} + (\partial_y - t\partial_X)^2.$$

The 'L' stands for 'linear' for reasons which will make more sense later. Taking the Fourier transform and then integrating gives

$$\partial_t \hat{g} = -\kappa \left(k^2 + |\eta - kt|^2 \right) \hat{g},$$

and

$$\hat{g}(t, k, \eta) = \hat{f}_{in}(k, \eta) \exp \left[-\kappa \int_0^t \left(k^2 + |\eta - k\tau|^2 \right) d\tau \right]. \tag{4.3}$$

Notice that we have the following bound

$$\int_0^t |\eta - k\tau|^2 \, d\tau \gtrsim \min(|\eta|^2 t, k^2 t^3).$$

To see this, consider separately contributions to the integral from $\tau \leq \frac{\eta}{2k}$ and $\tau \geq 2\frac{\eta}{k}$. From (4.3), this shows that we get the following *enhanced dissipation* estimate for some c (which happens to be $< 1/3$),

$$\|g_{\neq}\|_{L^2} \lesssim \|g_{in}\|_{L^2} e^{-c\nu t^3}$$

The key point to the decay is the relationship between ν and t. The characteristic time-scale when the dissipation begins to dominate is $\tau_{ED} \sim \nu^{-1/3}$, which is significantly faster than the ν^{-1} time-scale associated with the heat equation. The Couette flow is sending information to high frequencies linearly in time, and this is where the 3 comes from (order of the Laplacian in the damping plus one from the integral). One can imagine this relaxation mechanism like the way a cup of coffee relaxes after you stir it up into a vortex. First, the angular dependence is eliminated as the fluid stirring itself mixes information to high frequencies where it is rapidly dissipated, like the relaxation of $k \neq 0$ modes. Over a longer time scale the (approximately) radially symmetric mean vortex relaxes in place as a laminar flow.

4.2 More general shear flows

The Couette flow is so easy we might get the impression that more general problems will continue to be super easy. This is not correct, and surprisingly little is known in mathematical rigor about more general flows (which is not to say that nothing is known or that there exists no good work on this – that is far from the truth, see e.g. [CKRZ, BW13, GGN09, Den13, BCZGH15, Zil14a, Zil14b] to name a small subset of related works, however our knowledge is still quite limited relative to what might be desired).

Following how we approached the Couette flow, the first goal is to consider the inviscid problem and study the transfer of information to high frequencies. Getting enhanced dissipation rates is in general harder, though see [CKRZ, BW13] for some information on this. Let us first try with a humble goal of considering more general shear flows than Couette; in this case results are not difficult.

We will prove the following two basic theorems, the proof is in a style similar to some found in [Zil14a] combined with some ideas from the method of stationary phase for classical oscillatory integrals [Ste93]. I will state a version on $\mathbb{T} \times \mathbb{R}$ and a version on $\mathbb{T}^2$, but it admits suitable generalizations to more general tori, channels bounded by walls with no flux boundaries, and higher dimensions.

Theorem 4.1 (Mixing by shear flows in $\mathbb{T} \times \mathbb{R}$)**.** *Let $U(y) \in C^\infty$ and let f solve the PDE*

$$\partial_t f + U(y)\partial_x f = 0.$$

Then

(i) If there is some $\delta > 0$ such that $|U'(y)| \geq \delta$ for all y, then

$$\|f(t) - \langle f \rangle_x\|_{H^{-1}} \lesssim \langle t \rangle^{-1} \|f(0)\|_{L^2_x H^1_y}. \tag{4.4}$$

(ii) If there is some $\delta > 0$ such that $|U''(y)| \geq \delta$ for all y, then

$$\|f(t) - \langle f \rangle_x\|_{H^{-1}} \lesssim \langle t \rangle^{-1/2} \|f(0)\|_{L^2_x H^1_y}. \tag{4.5}$$

(iii) More generally, suppose there is some R and δ such that $|U'(y)| \geq \delta$ for all $|y| \geq R$ and further that there are finitely many points y_i, $1 \leq i \leq K$, such that $U'(y_i) = 0$ and finitely many inflection points where $U''(\tilde{y}_i) = 0$. Further suppose that U' degenerates only to finite order: that is, there is a finite $n \in \mathbb{N}$ which is the minimal integer such that $U^{(n)}(y_i) \neq 0$ for all y_i (the critical points of the flow). Then,

$$\|f(t) - \langle f \rangle_x\|_{H^{-1}} \lesssim \langle t \rangle^{-1/n} \|f(0)\|_{L^2_x H^1_y}. \tag{4.6}$$

Remark 8. Items (i) and (ii) are special cases of item (iii), however, the statements are less technical and the proofs can be made much more direct in the cases of (i) and (ii).

This theorem has the following analogue on $\mathbb{T}^2$, which we only state in the general case now.

Theorem 4.2 (Mixing by shear flows in $\mathbb{T}^2$)**.** *Let $U(y)$ be C^∞ with finitely many points y_i, $1 \leq i \leq K$, such that $U'(y_i) = 0$ and finitely many inflection points where $U''(\tilde{y}_i) = 0$. Suppose that U' degenerates only to finite order: that is, there is a finite $n \in \mathbb{N}$ which is the minimal integer such that $U^{(n)}(y_i) \neq 0$ for all y_i (the critical points of the flow); necessarily $n \geq 2$. Let f solve the PDE*

$$\partial_t f + U(y)\partial_x f = 0.$$

Then, we have the decay rate:

$$\|f(t) - \langle f \rangle_x\|_{H^{-1}} \lesssim \langle t \rangle^{-1/n} \|f(0)\|_{L^2_x H^1_y}. \tag{4.7}$$

Remark 9. Notice the loss of regularity in (4.7). From the example of the Couette flow, we can surmise that this is necessary to deduce the pointwise-in-time decay estimate.

Remark 10. The proof will show that one can be a tiny bit more precise about exactly the norm that appears on the RHS of (4.7).

Proof. We may assume without loss of generality $\langle f\rangle_x = 0$ and $t \geq 1$. First observe that

$$\|f\|_{H^{-1}} = \sup_{\phi\in H^1:\|\phi\|_{H^1}=1;\phi} \int f\phi dA.$$

Let ϕ be such an arbitrary test function. Fourier transform in x only, denoting this as $\hat{f}_k(t,y)$, we get

$$\begin{aligned}\left|\int f\phi dA\right| &= \left|\sum_k \int \hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| \\ &\leq \sum_k \left|\int \hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right|.\end{aligned}$$

The result will follow by the method of stationary phase (see e.g. [Ste93]). Indeed, let $\chi_\epsilon(y)$ be a smooth cutoff function supported in ϵ intervals around the critical points y_i. Then consider the two contributions separately:

$$\begin{aligned}\left|\int \hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| &\leq \left|\int (1-\chi_\epsilon(y))\,\hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| + \left|\int \chi_\epsilon(y)\hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| \\ &= T1 + T2.\end{aligned}$$

On $T2$, the phase $ikU(y)$ is stationary so we cannot use any integration by parts. However, we can employ (together with the $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$ embedding):

$$T2 \lesssim \|\chi_\epsilon\|_{L^1} \sum_k \|f_k\|_{L^\infty_y} \|\phi_{-k}\|_{L^\infty_y} \lesssim \epsilon \|f\|_{L^2_x H^1_y}.$$

For $T1$ we have to be a little more precise. On $T1$ we may employ $\frac{1}{ikU'(y)t}\frac{d}{dy}e^{ikU(y)t} = e^{ikU(y)t}$ and integrate by parts to deduce

$$\begin{aligned}T1 &\lesssim \frac{1}{\langle t\rangle}\sup_y \frac{\mathbf{1}_{|y-y_i|>\epsilon}}{|U'(y)|}\sum_k \frac{1}{|k|}\left(\left\|\hat{f}_k\right\|_{L^2_y}\|\nabla\phi_{-k}\|_{L^2_y} + \left\|\nabla\hat{f}_k\right\|_{L^2_y}\|\phi_{-k}\|_{L^2_y}\right) \\ &\quad + \sum_{k\neq 0}\frac{1}{t|k|}\left|\int e^{ikU(y)t}\frac{1}{U'(y)}\partial_y\chi_\epsilon(y)\hat{f}(k,y)\overline{\hat{\phi}}(k,y)dy\right| \\ &\quad + \frac{1}{t|k|}\left|\int e^{ikU(y)t}\left(\frac{d}{dy}\frac{1}{U'(y)}\right)(1-\chi_\epsilon(y))\hat{f}(k,y)\overline{\hat{\phi}}(k,y)dy\right| \\ &= T1_0 + T1_1 + T1_2.\end{aligned}$$

The treatment of $T1_1$ is straightforward

$$\begin{aligned}T1_1 &\lesssim \frac{1}{\langle t\rangle}\sup_y \frac{\mathbf{1}_{|y-y_i|>\epsilon}}{|U'(y)|}\sum_{k\neq 0}\|\partial_y\chi_\epsilon\|_{L^1}\left\|\hat{f}_k\right\|_{H^1}\left\|\hat{\phi}_k\right\|_{H^1}. \\ &\lesssim \frac{1}{\langle t\rangle}\sup_y \frac{\mathbf{1}_{|y-y_i|>\epsilon}}{|U'(y)|}\left\|\hat{f}_k\right\|_{L^2H^1}\left\|\hat{\phi}_k\right\|_{L^2H^1}.\end{aligned}$$

By non-degeneracy (and Taylor's theorem) we get

$$\sup_y \frac{|1-\chi_\epsilon(y)|}{|U'(y)|} \approx \sup_y \frac{\mathbf{1}_{|y-y_i|>\epsilon}}{|U'(y)|} \lesssim \epsilon^{-n+1}.$$

For $T1_2$ we have to be a bit more precise due to the derivative of the nearly singular $(U')^{-1}$. Let $\tilde{\chi}_\epsilon(y)$ be a smooth cutoff supported in ϵ^{n-1} intervals around the *inflection* points. Then we further divide

$$\begin{aligned}
T1_2 &\leq \frac{1}{\langle t\rangle}\left|\int (1-\chi_\epsilon(y))\,(1-\tilde{\chi}_\epsilon(y))\left(\frac{d}{dy}\frac{1}{U'(y)}\right)\hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| \\
&\quad + \frac{1}{\langle t\rangle}\left|\int (1-\chi_\epsilon(y))\,\tilde{\chi}_\epsilon(y)\left(\frac{d}{dy}\frac{1}{U'(y)}\right)\hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| \\
&= T1_{21} + T1_{22}.
\end{aligned}$$

For $T1_{21}$ we note that the U' is monotone on each disjoint interval where the integral is supported (call these intervals J_i, $1 \leq i \leq N$) and U' is bounded below by $|U'| \gtrsim \epsilon^{1-n}$ on the support of the integrand. Therefore (using again the embedding $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$),

$$\begin{aligned}
T1_{21} &\lesssim \frac{1}{\langle t\rangle}\|f\|_{L^2L^\infty}\|\phi\|_{L^2L^\infty}\sum_{i=1}^N\int_{J_i}\left|\frac{d}{dy}\frac{1}{U'(y)}\right|dy \\
&= \frac{1}{\langle t\rangle}\|f\|_{L^2H^1}\|\phi\|_{L^2H^1}\sum_{i=1}^N\left|\int_{J_i}\frac{d}{dy}\frac{1}{U'(y)}dy\right| \\
&\leq \frac{2N}{\langle t\rangle}\|f\|_{L^2H^1}\|\phi\|_{L^2H^1}\sup_{y\in J_i}\frac{1}{|U'(y)|} \\
&\lesssim \frac{1}{\langle t\rangle\epsilon^{n-1}}\|f\|_{L^2H^1}\|\phi\|_{L^2H^1}.
\end{aligned}$$

For $T1_{22}$ we use instead that $|\tilde{\chi}_\epsilon U''| \lesssim \epsilon^{n-1}$ because U is smooth. From this we have

$$\begin{aligned}
T1_{22} &\lesssim \frac{1}{\langle t\rangle}\left|\int (1-\chi_\epsilon(y))\,\tilde{\chi}_\epsilon(y)\left(\frac{U''(y)}{(U'(y))^2}\right)\hat{f}_k(0,y)e^{ikU(y)t}\overline{\hat{\phi}}_k(y)dy\right| \\
&\lesssim \frac{1}{\langle t\rangle}\left\|\tilde{\chi}_\epsilon U''\right\|_{L^\infty_y}\epsilon^{2-2n}\|f\|_{L^2H^1}\|\phi\|_{L^2H^1} \\
&\lesssim \frac{1}{\langle t\rangle}\epsilon^{1-n}\|f\|_{L^2H^1}\|\phi\|_{L^2H^1}.
\end{aligned}$$

Putting everything together we get

$$\|f(t)\|_{H^{-1}} \lesssim T1 + T2 \lesssim \left(\epsilon + \frac{1}{\langle t\rangle\epsilon^{n-1}}\right)\|f(0)\|_{L^2H^1}.$$

Hence, we can choose the optimal ϵ,

$$\epsilon \approx \langle t\rangle^{-\frac{1}{n}}.$$

and deduce

$$\|f(t)\|_{H^{-1}} \lesssim \langle t\rangle^{-\frac{1}{n}}\|f(0)\|_{L^2_xH^1_y},$$

which completes the proof. □

5 Linearized 2D Euler and Navier-Stokes equations revisited

There is one case in which we can understand the entire nonlinear dynamics of Navier-Stokes (in 2D and 3D) and Euler (in 2D), at least for very smooth solutions, near a shear flow. This shear flow is the Couette flow $(y, 0)$ (or $(y, 0, 0)$ in 3D of course) with $(x, y) \in \mathbb{T} \times \mathbb{R}$ or $(x, y, z) \in \mathbb{T} \times \mathbb{R} \times \mathbb{T}$. The first work was 2D Euler [BM13], then 2D Navier-Stokes [BMV14] then 3D Navier-Stokes [BGM15a, BGM15b] (in both chronological and complexity ordering). The flow is quite nice and canonical but is also by far the easiest at the linear level and it is apparent that in order to understand the nonlinear problem, we will need an extremely precise understanding of the linear problem.

Due to the fact that the nonlinearity does not leave the x averages of the vorticity invariant, in order to understand nonlinear perturbations of Couette, we will need to be able to understand nearby flows (in some sense). Consider first the linearized 2D Navier-Stokes equations in vorticity form near an arbitrary shear flow:

$$\begin{aligned}\partial_t \omega + U(y)\partial_x \omega - U''\partial_x \psi &= \nu \Delta \omega \\ \Delta \psi &= \omega.\end{aligned}$$

For the Couette flow, $U(y) = y$, $U'' = 0$ and so the linearized Navier-Stokes equation reduces back to the passive scalar equation. However, one thing we did not consider is what happens to ψ (and hence the velocity). From previous lectures we saw that if we wrote $f(t, z, y) = \omega(t, z + ty, y)$ and $\phi(t, z, y) = \psi(t, z + ty, y)$, we derive

$$\begin{aligned}\partial_t f &= \nu \Delta_L f \\ \Delta_L \phi &= f \\ \Delta_L &= \partial_{zz} + (\partial_y - t\partial_z)^2.\end{aligned}$$

By taking Fourier transforms and integrating, we derived (actually Kelvin derived [Kel87]),

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

which implies the enhanced dissipation estimate (denoting $f_{\neq} = f - f_0 = f - \frac{1}{2\pi}\int f(t, x, y)dx$):

$$\|f_{\neq}\|_{H^\sigma} \lesssim \|\omega_{in}\|_{H^\sigma} e^{-c\nu t^3},$$

for some $0 < c < 1/3$. From the equation for ϕ we further derive

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

Now we observe something that Orr (basically) observed in 1907 [Orr07]: by using $\langle \eta \rangle \langle \eta - kt \rangle \gtrsim \langle kt \rangle$, we get the following for any $\beta \in [0, 2]$,

$$\|\phi_{k\neq 0}(t)\|_{H^\sigma} \lesssim \left(\sum_{k\neq 0} \int \langle k, \eta \rangle^{2\sigma} \frac{\left|\hat{f}(t, k, \eta)\right|^2}{\left(k^2 + |\eta - kt|^2\right)^2} d\eta\right)^{1/2} \lesssim \frac{1}{\langle t \rangle^\beta} \|f_{k\neq 0}(t)\|_{H^{\sigma+\beta}}. \tag{5.1}$$

This provides a decay of the streamfunction in the new variables (z, y). The crucial point here is that the decay is *independent of Reynolds number*. This effect is called *inviscid damping*[4]. Now using that

$$u_1(t,x,y) = -\partial_y \psi(t,x,y) = -\left((\partial_y - t\partial_x)\phi\right)(t, x-ty, y)$$
$$u_2(t,x,y) = \partial_x \psi(t,x,y) = -\left(\partial_x \phi\right)(t, x-ty, y),$$

we derive

$$\left\|u^1_{k\neq 0}(t)\right\|_{L^2} + \langle t\rangle \left\|u^2_{k\neq 0}(t)\right\|_{L^2} \lesssim \frac{1}{\langle t\rangle} \left\|f_{in}\right\|_{H^4}.$$

For the linear equations, we have $\widehat{u_1}(t,0,\eta) = \widehat{u_{1\,in}}(0,\eta)$ and $u^2_0(t,x,y) = 0$ (the latter due to the incompressibility) and hence

$$\left\| u(t) - \begin{pmatrix} u_0(0,y) \\ 0 \end{pmatrix} \right\|_{L^2} \lesssim \langle t\rangle^{-1} \left\|f_{in}\right\|_{H^4},$$

which shows convergence back to a shear flow – *back to equilibrium*. This is a purely inviscid stabilization mechanism wherein the shear flow opposes any non-shear flow motion and so returns back to equillibrium (at least at the linear level).

Although Orr predated Sobolev spaces, Orr understood where the regularity loss in (5.1) comes from. He pointed out that if one considers modes with $\eta/k \geq 0$, then the streamfunction associated with these modes first grows, reaches a maximum in amplitude at $t = \frac{\eta}{k}$ and the decays again. Orr referred to the time $kt = \eta$ as the *critical time*, which is the terminology we will use here, In particular, the *kinetic energy* associated with the mode (k, η) is amplified by a factor of $\approx |\eta/k|$. This mechanism of transient growth then decay is known as the *Orr mechanism* and has a number of interesting implications also in atmospheric dynamics [Boy83, Lin88]. Using this general idea we can prove that the linearized 2D Euler equations in velocity form are Lyapunov unstable with respect to the L^2 norm (despite being spectrally stable).

Theorem 5.1. *Consider the linearized 2D Euler equations around Couette flow:*

$$\partial_t u + y\partial_x u + \begin{pmatrix} u_2 \\ 0 \end{pmatrix} = -\nabla p \tag{5.2a}$$
$$\nabla \cdot u = 0 \tag{5.2b}$$
$$u(0) = u_{in}. \tag{5.2c}$$

These equations are Lyapunov unstable in the L^2 norm in the sense that for all $\epsilon > 0$, there exists smooth initial data with $\|u_{in}\|_{L^2} < \epsilon$ such that the resulting solution $u(t)$ to (5.2) satisfies

$$\sup_{t\in[0,\infty)} \|u(t)\|_{L^2} = 1.$$

[4] The name "inviscid damping" seems to have been first adopted by plasma physicists due to its similarity with a related mechanism in the Vlasov equations of kinetic theory known as *Landau damping*, discovered later by Landau [Lan46] (see e.g. [MV11, BMM13] and the references therein). In fact, many people simply refer to both Landau and inviscid damping as Landau damping, however, we prefer to use a different terminology to emphasize that the two mechanisms are not totally isomorphic, instead they are different examples of a unifying concept known as "phase mixing". See [BM13] and the references therein for more information.

Proof. Consider the initial *vorticity* $\tilde{\omega}(x,y) = c_0 \sin x \phi(y)$ for a smooth, non-trivial, compactly supported cut-off ϕ and a constant c_0 chosen such that $\left\|\nabla^\perp \Delta^{-1}\tilde{\omega}\right\|_{L^2} = 1$. Define for $\lambda > 0$ to be chosen

$$\tilde{\omega}_\lambda(x,y) = c_0 \sin(x + \lambda y)\phi(y).$$

By the inviscid damping estimate (5.1) and the arguments that follow, we get

$$\left\|\nabla^\perp \Delta^{-1}\tilde{\omega}_\lambda\right\|_{L^2} \lesssim \lambda^{-1}.$$

Next consider initial data $\omega(0,x,y) = \tilde{\omega}_\lambda(x,y)$ and $\partial_t \omega + y\partial_x \omega = 0$, then

$$\omega(t,x,y) = \tilde{\omega}_{\lambda - t}(x,y).$$

Moreover, note that $u(t,x,y)$ solves the linearized 2D Euler equations. Putting everything together, we have a solution to the Euler equations with

$$\begin{aligned} \|u(0,x,y)\|_{L^2} &\lesssim \lambda^{-1} \\ \|u(\lambda,x,y)\|_{L^2} &= 1. \end{aligned}$$

□

5.1 More general shear flows

Naturally, one wants to investigate inviscid damping (and enhanced dissipation) for more general shear flows, beginning of course at the linear level. Recall from previous sections that the vorticity equation is *not* a passive scalar equation for flows that are not Couette flow:

$$\partial_t \omega + U(y)\partial_x \omega - U''\partial_x \psi = 0 \tag{5.3a}$$

$$\Delta \psi = \omega. \tag{5.3b}$$

If the $U''\partial_x\psi$ term were absent, we would have a passive scalar equation and we know from Theorem 4.2 that inviscid damping would occur (at least at some rate). The extra term in the linearized equations may be lower order, but its presence is still problematic since we cannot so easily deduce the streamfunction is decaying in order to treat it as in the passive scalar case. However, we can still do something in a neighborhood of flows which look like the Couette flow.

Let me show a *relatively* simple argument for getting inviscid damping estimates on (5.3). I will make no attempt to be as accurate or as general as possible; for better and more general results, see Zillinger [Zil14a] on linear inviscid damping of monotone shear flows which satisfy a smallness condition on U'' (this can be interpreted as a condition requiring the shear be locally close to the Couette flow). The proof is also a warm-up for the more technically challenging issues in the nonlinear problem and, in particular, the proof is stylistically closer to some techniques used in [BM13, BMV14, BGM15a, BGM15b].

Theorem 5.2 (Zillinger [Zil14a])**.** *There exists a universal $\epsilon_0 > 0$ such that if $\|U' - 1\|_{H^6} < \epsilon_0$ then the solution to the linearized 2D Euler equations around the shear flow $u_E = (U,0)$ in $\mathbb{T}\times\mathbb{R}$ experiences the following inviscid damping for $\omega_{in} \in H^3$:*

$$\begin{aligned} \left\|u^1_{k\neq 0}\right\|_{L^2} &\lesssim \frac{1}{\langle t\rangle}\|\omega_{in}\|_{H^2} \\ \left\|u^2_{k\neq 0}\right\|_{L^2} &\lesssim \frac{1}{\langle t\rangle^2}\|\omega_{in}\|_{H^3}. \end{aligned}$$

Proof. **Change of variables**

First, it suffices to assume that $\frac{1}{2\pi}\int_{\mathbb{T}}\omega_{in}(x,y)dx = 0$ as the functions which are independent of x are in the nullspace of the linear operator. Next, we want to reformulate the problem a bit. First, we want to rewind by the characteristics but we also want to re-scale the y coordinates to make the shear look more like Couette flow (this will be pivotal later for slightly subtle reasons). That is, we define

$$X = x - tU(y) \tag{5.4a}$$
$$Y = U(y). \tag{5.4b}$$

By choosing ϵ_0 small enough, $\|U' - 1\|_{L^\infty} \lesssim \|U' - 1\|_{H^1} < \epsilon_0$ implies U is strictly monotone and hence by the implicit function theorem, this coordinate transform is always invertible and we can always solve for x and y in terms of (X,Y): $x = x(t,X,Y)$, $y = y(Y)$. Now define the new variables

$$f(t,X,Y) = \omega(t,x(t,X,Y),y(Y))$$
$$\phi(t,X,Y) = \psi(t,x(t,X,Y),y(Y)).$$

We can check now that f and ϕ satisfy the following, where $g(Y) = U'(U^{-1}(Y))$ and $b = U''(U^{-1}(Y))$,

$$\partial_t f = b\partial_X\phi \tag{5.5a}$$
$$\Delta_t\phi := \partial_X^2\phi + g(\partial_Y - t\partial_X)^2\phi + b(\partial_Y - t\partial_X)\phi = f. \tag{5.5b}$$

Our goal now is to first get a uniform in time H^3 estimate on f, then to (hopefully) be able to prove this implies inviscid damping on ϕ, and then transfer this information to inviscid damping on u.

Constant coefficient toy model

Suppose we just wanted to get a simple L^2 estimate on f. To do so, we're not going to use the L^2 norm, instead, we will design a norm which is well-adapted to the Orr mechanism lurking on the RHS of the f equation. One can also draw parallels with the ghost weight energy method of Alinhac [Ali01]. Imagine that we had the following equation now for $b \in \mathbb{R}$:

$$\partial_t f = b\partial_x\Delta_L^{-1}f.$$

First, if $b \in \mathbb{R}$ one can solve this equation explicitly in Fourier, moreover, there is an energy structure here that ensures the L^2 norm and all H^s norms are conserved exactly (because $|\Delta_L|^{-1/2}$ is self-adjoint, which is not going to generally work in the the variable coefficient problem). Let us suppose we are too dumb to notice this. On the Fourier side, suppose we take absolute values, then we get

$$\begin{aligned}\partial_t\left|\hat{f}\right|(t,k,\eta) &\leq \frac{|b|\,|k|}{|k|^2 + |\eta - kt|^2}\left|\hat{f}\right|(t,k,\eta)\\ &\leq \frac{|b|}{1 + \left|\frac{\eta}{k} - t\right|^2}\left|\hat{f}\right|(t,k,\eta)\end{aligned} \tag{5.6}$$

By the comparison for ODEs, if $|b| \leq 1$ (its clear this is not a requirement we really "need" here but it will be important for the general problem below) we can deduce the following: if $w(t,k,\eta)$ such that $w(t,0,\eta) = 1$ and for $k \neq 0$ we have

$$\partial_t w(t,k,\eta) = \frac{1}{1 + \left|\frac{\eta}{k} - t\right|^2}w(t,k,\eta)$$

then

$$\frac{\left|\hat{f}\right|(t,k,\eta)}{w(t,k,\eta)} \leq \left|\hat{f}\right|(0,k,\eta).$$

Note that $1 \lesssim w^{-1} \leq 1$. Now the idea is to use $w(t,k,\eta)^{-1}$ as a Fourier multiplier, and we immediately deduce

$$\|f(t)\|_{H^s} \lesssim \left\|\frac{1}{w(t,\nabla)}f(t)\right\|_{H^s} \leq \|f_{in}\|_{H^s}.$$

Variable coefficient estimates

Now we turn to the full PDE and begin an H^s energy estimate using w^{-1} as a Fourier multiplier:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left\|w^{-1}(t,\nabla)f\right\|_{H^s}^2 &= -\sum\int \frac{\partial_t w}{w^3}\langle k,\eta\rangle^{2s}\left|\hat{f}\right|^2 d\eta + \left(w^{-1}f, w^{-1}(b\partial_x\phi)\right)_{H^s} \\ &= CK_w + L. \end{aligned} \tag{5.7}$$

The CK_w stands for "Cauchy-Kovalevskaya" because this trick of weakening the norm to introduce negative terms goes all the way back there (though in analytic regularity). Note that the CK_w term is written as (we are using that w and $\partial_t w$ are both strictly positive)

$$CK_w = -\left\|\sqrt{\frac{\partial_t w}{w}}\frac{1}{w}f\right\|_{H^s}^2.$$

For the term L, our goal is to eventually deduce:

$$L \lesssim \|b\|_{H^\sigma}\left\|\sqrt{\frac{\partial_t w}{w}}f\right\|_{H^s}^2,$$

from which the result will follow if b is small since we can absorb the entire contribution with the CK_w term. The approach hinges on being able to approximate $\Delta_t\phi$ with $\Delta_L\phi$ in a rather specific manner. We will use the following three basic inequalities

$$\langle k,\eta\rangle^s \lesssim_s \langle\eta-\xi\rangle^s + \langle k,\xi\rangle^s \tag{5.8a}$$

$$w(t,k,\eta) \approx w(t,k,\xi) \tag{5.8b}$$

$$\frac{\partial_t w}{w}(t,k,\eta) = \frac{1}{1+\left|\frac{\eta}{k}-t\right|^2} \lesssim \frac{\langle\eta-\xi\rangle^2}{1+\left|\frac{\xi}{k}-t\right|^2} \approx \langle\eta-\xi\rangle^2\frac{\partial_t w}{w}(t,k,\xi). \tag{5.8c}$$

Armed with these observations, we note

$$\begin{aligned} L &\lesssim \sum\int_{\eta,\xi}\frac{\langle k,\eta\rangle^s}{w(t,k,\eta)}\left|\hat{f}(t,k,\eta)\right|\frac{\langle k,\eta\rangle^s|k|}{w(t,k,\eta)}\left|\hat{b}(\eta-\xi)\right|\left|\hat{\phi}(t,k,\xi)\right|d\xi d\eta \\ &\lesssim \sum\int_{\eta,\xi}\frac{\langle k,\eta\rangle^s}{w(t,k,\eta)}\left|\hat{f}(t,k,\eta)\right|\left|\langle\eta-\xi\rangle^s\hat{b}(\eta-\xi)\right|\frac{|k|}{|k|^2+|\xi-tk|^2}\frac{\langle k,\xi\rangle^s}{w(t,k,\xi)}\left|\widehat{\Delta_L\phi}(t,k,\xi)\right|d\xi d\eta. \end{aligned}$$

Note that what we did here in the second line was to multiply and divide by Δ_L. Now, we use a few more of the above inequalities:

$$\begin{aligned}
L &\lesssim \sum \int_{\eta,\xi} \frac{\langle k,\eta\rangle^s}{w(t,k,\eta)} \left|\hat{f}(t,k,\eta)\right| \left|\langle \eta-\xi\rangle^s \hat{b}(\eta-\xi)\right| \frac{1}{1+\left|\frac{\xi}{k}-t\right|^2} \frac{\langle k,\xi\rangle^s}{w(t,k,\xi)} \left|\widehat{\Delta_L\phi}(t,k,\xi)\right| d\xi d\eta \\
&\lesssim \sum \int_{\eta,\xi} \sqrt{\frac{\partial_t w}{w}} \frac{\langle k,\eta\rangle^s}{w(t,k,\eta)} \left|\hat{f}(t,k,\eta)\right| \left|\langle \eta-\xi\rangle^{s+1} \hat{b}(\eta-\xi)\right| \sqrt{\frac{\partial_t w}{w}} \frac{\langle k,\xi\rangle^s}{w(t,k,\xi)} \left|\widehat{\Delta_L\phi}(t,k,\xi)\right| d\xi d\eta \\
&\lesssim \left\|\langle \eta\rangle^{s+1}\hat{b}(\eta)\right\|_{L^1_\eta} \left\|\sqrt{\frac{\partial_t w}{w}} w^{-1} f\right\|_{H^s} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s} \\
&\lesssim \|b\|_{H^{s+3}} \left\|\sqrt{\frac{\partial_t w}{w}} w^{-1} f\right\|_{H^s} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s}. \qquad (5.9)
\end{aligned}$$

where in the second to last line we used Cauchy-Schwarz and then Young's convolution inequality in frequency. At this point we would be done if $\Delta_L = \Delta_t$ and b is sufficiently small. However, this is not true. Instead, we need another lemma:

Lemma 1. *For $\|1-g\|_{H^{s+3}} + \|b\|_{H^{s+3}}$ sufficiently small, there holds*

$$\left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s} \lesssim \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w} f\right\|_{H^s}.$$

Proof. The main key insight is the first step:

$$\Delta_L\phi = f + (1-g)(\partial_Y - t\partial_X)^2\phi - b(\partial_Y - t\partial_X)\phi.$$

Next, we will mostly just use (5.8) in a very similar manner to our estimates on L above. Indeed,

$$\begin{aligned}
\left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s} &\lesssim \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w} f\right\|_{H^s} + \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\left((1-g)(\partial_Y - t\partial_X)^2\phi\right)\right\|_{H^s} \\
&\quad + \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\left(b(\partial_Y - t\partial_X)\phi\right)\right\|_{H^s} \\
&:= \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w} f\right\|_{H^s} + E1 + E2.
\end{aligned}$$

Let us just consider $E1$, $E2$ is similar. Using (5.8) again we deduce,

$$\begin{aligned}
E1 &\approx \left\|\int_\xi \sqrt{\frac{\partial_t w}{w}}(k,\eta) \frac{\langle k,\eta\rangle^s}{w(t,k,\eta)} \Big((1-\hat{g})(\eta-\xi)(\xi-tk)^2\hat{\phi}(k,\xi)\Big) d\xi\right\|_{L^2_{k,\eta}} \\
&\lesssim \left\|\int_\xi \langle \eta-\xi\rangle^{s+1}(1-\hat{g})(\eta-\xi)(\xi-tk)^2 \sqrt{\frac{\partial_t w}{w}} \frac{\langle k,\xi\rangle^s}{w(t,k,\xi)}\hat{\phi}(k,\xi) d\xi\right\|_{L^2_{k,\eta}}.
\end{aligned}$$

Using Young's convolution inequality we then get

$$\begin{aligned}
E1 &\lesssim \left\|\langle \eta\rangle^{s+1}(1-\hat{g})(\eta)\right\|_{L^1_\eta} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s} \\
&\lesssim \|1-g\|_{H^{s+3}} \left\|\sqrt{\frac{\partial_t w}{w}} \frac{1}{w}\Delta_L\phi\right\|_{H^s}.
\end{aligned}$$

For $E2$ we get a similar estimate. Therefore,

$$\left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} \Delta_L \phi \right\|_{H^s} \lesssim \left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} f \right\|_{H^s} + \|1 - g\|_{H^{s+3}} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} \Delta_L \phi \right\|_{H^s} + \|b\|_{H^{s+3}} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} \Delta_L \phi \right\|_{H^s}.$$

Since the linear equations are globally well-posed in H^3 on the vorticity, we know that the norms in question are all a priori finite. This allows us to move terms to the LHS and divide, giving us (for some $C > 0$):

$$\left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} \Delta_L \phi \right\|_{H^s} \leq \frac{1}{1 - C \|1 - g\|_{H^{s+3}} - C \|b\|_{H^{s+3}}} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{1}{w} f \right\|_{H^s},$$

from which the result follows. □

Connecting now Lemma 1 with (5.9) and (5.7) we deduce that for $\|b\|_{H^{s+3}}$ sufficiently small,

$$\|f(t)\|_{H^s} \approx \left\|w(t, \nabla)^{-1} f(t)\right\|_{H^s} \leq \|f_{in}\|_{H^s}.$$

It remains now to see why this uniform H^s bound implies Theorem 5.2. First, we note the following lemma.

Lemma 2. *For $\|1 - g\|_{H^{s+3}} + \|b\|_{H^{s+3}}$ sufficiently small, there holds*

$$\|\phi(t)\|_{H^{s-2}} \lesssim \frac{1}{\langle t \rangle^2} \|f(t)\|_{H^s}.$$

Proof. Arguing as in (5.1) we have

$$\|\phi(t)\|_{H^{s-2}} \lesssim \frac{1}{\langle t \rangle^2} \|\Delta_L \phi(t)\|_{H^s}.$$

From here the lemma follows in a manner very similar to Lemma 1. □

Now we use

$$\begin{aligned} u_1(t, x, y) &= -U'(y) \left((\partial_v - t\partial_z)\phi \right)(t, x - tU(y), U(y)) \\ u_2(t, x, y) &= (\partial_z \phi)(t, x - tU(y), U(y)), \end{aligned}$$

to deduce (we are also using that composition under $U(y)$ is bounded on L^2, which comes from the uniform monotonicity of U),

$$\begin{aligned} \|u_1(t)\|_{L^2} &\lesssim \langle t \rangle \|\nabla \phi(t)\|_{H^1} \lesssim \langle t \rangle^{-1} \|f_{in}\|_{H^3} \\ \|u_2(t)\|_{L^2} &\lesssim \|\nabla \phi(t)\|_{H^1} \lesssim \langle t \rangle^{-2} \|f_{in}\|_{H^3}. \end{aligned}$$

Theorem 5.2 now follows. □

6 Nonlinear 2D Euler equations near the Couette flow

Now we come to the main question: does inviscid damping and enhanced dissipation hold for the *nonlinear dynamics* near the Couette flow? This is a question with an answer which is far from obvious as there is no clear, convincing, reason a priori to expect the linear behavior to dominate for all time. Numerics, physical experiments, and formal asymptotics were not conclusive either for a few subtle reasons (see e.g. [Shn12] and the references therein). Hence, this is a nice example of a place where mathematical analysis can answer questions that even physicists consider open.

Consider first 2D Euler in vorticity form near the Couette flow with $(x, y) \in \mathbb{T} \times \mathbb{R}$.

$$\begin{aligned}\partial_t \omega + y\partial_x \omega + \nabla^\perp \psi \cdot \nabla \omega = 0 \\ \Delta \psi = \omega.\end{aligned}$$

The results of [BM13] essentially show that, up to a correction due to the quasilinearity, the nonlinear dynamics are qualitatively similar to the linear dynamics provided the solutions are small[5] and smooth The regularity we need to work in is called Gevrey-$\frac{1}{s}$ with $s \in (0, 1]$ [Gev18] and is quantified with a scale of norms for $\lambda > 0$

$$\|f\|_{\mathcal{G}^{\lambda;s}} = \left\| e^{\lambda |\nabla|^s} f \right\|_2 .$$

Notice that $s = 1$ corresponds to real analytic whereas $s < 1$ does not (and there are compactly supported functions in $\mathcal{G}^{\lambda;s}$ for $s < 1$). This is a scale of very nice spaces that connects the two (somewhat pathological) extremes of C^∞ and real analytic. However, this is not why we are working in this regularity class – we will see that it's "necessity" here is predicted by formal weakly nonlinear analysis. In [BM13] we prove (slightly paraphrased and updated):

Theorem 6.1 (Nonlinear inviscid damping [BM13])**.** *Let $s \in (1/2, 1]$ and let $\lambda > \lambda' > 0$ and $\delta > 0$. There exists an $\epsilon_0 = \epsilon_0(\lambda, \lambda', s, \delta)$ such that if*

$$\|u_{in}\|_{L^2} + \|\omega_{in}\|_{\mathcal{G}^{\lambda;s}} = \epsilon < \epsilon_0,$$

then there exists some $\omega_\infty \in \mathcal{G}^{\lambda';s}(\mathbb{R}^2)$ and $u_\infty, \phi(t) \in \mathcal{G}^{\lambda';s}$ such that the following holds with all implicit constants independent of ϵ and t:

$$u_\infty(y) = -\partial_y \partial_{yy}^{-1} \langle \omega_\infty \rangle_x \tag{6.1}$$

$$\|\omega(t, x + ty + t\phi(t, y), y) - \omega_\infty\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon^2}{\langle t \rangle} \tag{6.2}$$

$$\left\|u^1_{\neq}(t)\right\|_{L^2} + \langle t \rangle \left\|u^2(t)\right\|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle} \tag{6.3}$$

$$\left\|\langle u^1(t)\rangle_x - u_\infty\right\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle t \rangle^2} \tag{6.4}$$

$$\left\|\phi(t) - \langle u^1(t)\rangle_x\right\|_{\mathcal{G}^{\lambda';s}} \lesssim \epsilon \langle t \rangle^{\delta - 1}. \tag{6.5}$$

Remark 11. Theorem 6.1 implies $\omega(t) \rightharpoonup \omega_\infty$ (a good exercise).

[5] You might ask "small relative to what in what units?" but remember we have scaled the velocities and lengths so that we have a problem on a torus of side-length 2π and a background shear flow with slope 1. In units, the smallness is relative to these parameters, in particular, for a torus of side-length of $2\pi L$ and shear of slope β, we should get a smallness condition which is linear in β but the dependence on L is a bit more opaque due to the Gevrey-$1/s$ norm, as rescaling the variables changes the λ and the ϵ_0 depends in a very complicated manner on λ (although by dimensional analysis, we expect it to be roughly $\approx \exp(-CL^p)$ for some C and $p = p(s)$).

The proof of Theorem 6.1 can also be used to prove the following, more or less relying on the fact that $\omega_\infty - \omega_{in} = O(\epsilon^2)$:

Corollary 1. *There is an open set in $\mathcal{G}^{\lambda;s}$ of initial data which result in solutions $\omega(t)$ satisfying the following for all $\sigma > 0$ and $t \in \mathbb{R}$:*

$$\begin{aligned} \|\omega_\infty\|_2 &< \|\omega(t)\|_2 \\ \epsilon\langle t\rangle^\sigma &\approx \|\omega(t)\|_{H^\sigma}. \end{aligned}$$

As far as Sobolev norm explosion results go, this is not so impressive given that faster rates are known in other settings; see e.g. [Den09, KŠ14] (although this gives very precise controls on all finite regularity norms which is lacking in other norm explosion results). However, the two more interesting points are (A) the orbits are not pre-compact in L^2 and (B) there is an open set of solutions whose norm does *NOT* grow exponentially. This setting is arguably somewhat contrived, however, neither of these have been verified in any setting of 2D Euler before so they are still of note.

We certainly do not have time to cover the entire proof, however, there should be time to discuss the most important points. The general schematic follows the technique we used to prove Theorem 5.2 above.

6.1 Nonlinear coordinate transform

In the linearized case, we used the coordinate transform (5.4). In the nonlinear case, we have the added complications that our shear flow is not stationary in time, not known a priori, and cannot be assumed to have significantly more regularity than the solution itself. The first step is to find the analogue of (5.4). Let us suppose that we are unsure how to guess it (for 2D Euler it is not so hard to guess, but for NSE it is less obvious, especially in 3D), and hence it makes sense to search for a coordinate system of the general form

$$\begin{aligned} X &= x - ty - th(t,y) \\ Y &= y + h(t,y). \end{aligned}$$

The first thing to note is that this is going to require h' small, so that we may invert the coordinate transform to recover information in (x,y) from information in (X,Y). This precise form is dictated so that the y derivatives transform in a specific way. Namely, if we denote $f(t,X(t,x,y),Y(t,y)) = \omega(t,x,y)$ and $\phi(t,X(t,x,y),Y(t,y)) = \psi(t,x,y)$, then

$$\begin{aligned} \partial_x\omega(t,x,y) &= (\partial_X f)(t,X(t,x,y),Y(t,y)) \\ \partial_y\omega &= (1+\partial_y h(t,y))\left((\partial_Y - t\partial_X)f\right)(t,X(t,x,y),Y(t,y)), \end{aligned}$$

and similar for ϕ. Similar to (5.5), we derive (note the new term involving $\partial_t h$),

$$\begin{aligned} \partial_t f + \left[\begin{pmatrix} -\frac{d}{dt}(th)(t,y) - \bar{u}_1(t,y) \\ \partial_t h \end{pmatrix} + \begin{pmatrix} -(1+\partial_y h)(\partial_Y - t\partial_X)\phi_{k\neq 0} \\ \partial_X\phi \end{pmatrix}\right] \cdot \begin{pmatrix} \partial_X f \\ -(1+\partial_y h)(\partial_Y - t\partial_X)f \end{pmatrix} &= 0 \\ \Delta_t\phi &= f \\ \partial_{XX} + (1+\partial_y h)^2(\partial_Y - t\partial_X)^2 + \partial_{yy}h(\partial_Y - t\partial_X) &:= \Delta_t. \end{aligned}$$

We have kept for a moment $\bar{u}_1(t,y)$ written in the y coordinates; note that $\bar{u}_1(t,y) = -(1 + \partial_y h(t,y))(\partial_Y\bar{\phi})(t,Y(t,y))$. The idea is then to choose h to eliminate the shear flow contribution, which implies

$$\frac{d}{dt}(th) = \bar{u}_1.$$

Then, notice the very crucial cancellation that ensures that the governing equations become

$$\begin{aligned} \partial_t f + \partial_t h \partial_Y f + (1 + \partial_y h)\nabla^{\perp}_{X,Y}\phi_{k\neq 0} \cdot \nabla_{X,Y} f = 0 \\ \Delta_t \phi = f. \end{aligned}$$

However, we are not done here. We need to be able to get good estimates on h and its derivatives, and we need to get them in the (X, Y) variables. It turns out there are a variety of good reasons to derive PDEs for h and its derivatives in (X, Y), but the most obvious reason is because compositions, like those involved in moving back and forth between variables, is poorly behaved in infinite regularity classes (see e.g. [MV11, BM13] and the references therein). If we want to avoid measuring high regularity norms of compositions, then we will have to get our highest norm estimates on everything in the (X, Y) variables. First, using $C(t, X, Y) = h(t, x, y)$, $U_i(t, X, Y) = u_i(t, x, y)$, $g = (\partial_t h)$, $v' = 1 + (\partial_y h)$ (in the same sense as above) and so forth, we get

$$g = \frac{\bar{U}_1 - C}{t},$$

which by

$$\partial_t C + g\partial_Y C = \frac{1}{t}\left[\bar{u}_1 - C\right]$$

and (by taking x-averages of the momentum equation),

$$\partial_t \bar{U}_1 + g\partial_Y \bar{U}_1 + v'\overline{(U_{\neq} \cdot \nabla(U_1)_{\neq})} = 0,$$

gives (by taking a time derivative)

$$\partial_t g + g\partial_Y g = -\frac{2}{t}g - \frac{v'}{t}\overline{(U_{\neq} \cdot \nabla(U_1)_{\neq})}.$$

As a side note, we used the x average structure to eliminate the $-v'\overline{U_1 t\partial_X U_1}$ term, which could have been problematic (this cancellation is much more crucial in 3D). By taking a y derivative of the definition of h and changing coordinates, we also get

$$\partial_t(v' - 1) + g\partial_Y(v' - 1) = \frac{1}{t}\left(-\bar{f} - (v' - 1)\right).$$

In the proof of [BM13, BMV14] it is actually useful to derive yet another PDE for the quantity on the RHS. We will not get so deep into the estimates here; as an exercise you could derive the PDE and think about why it might be useful.

6.2 Weakly nonlinear heuristics and the toy model

We know already that an H^2 bound on $\omega(t, x - ty, y)$ in the linear case implies inviscid damping of the (linearized) velocity field. With this in mind, it hence makes sense that an H^2 bound on f will give us inviscid damping for the nonlinear problem. Indeed, if we deduce sufficient regularity estimates on the coefficients, then the proof of Lemma 2 adapts and we do indeed get inviscid damping, as long as we have sufficient Sobolev regularity on g, v', and f. Now return to the vorticity equation in the new variables and see about deducing such an H^s bound:

$$\begin{aligned}\partial_t f + g\partial_Y f + v'\nabla^{\perp}_{X,Y}\phi_{k\neq 0}\cdot\nabla_{X,Y} f &= 0\\ \Delta_t\phi &= f.\end{aligned}$$

Let us think for a moment about the following simpler system, for which we will already see the problem:

$$\partial_t f + \nabla^{\perp}_{X,Y}\phi_{k\neq 0}\cdot\nabla_{X,Y} f = 0 \tag{6.6a}$$

$$\Delta_L\phi = 0. \tag{6.6b}$$

Let us see what happens when we attempt an H^s estimate:

$$\frac{1}{2}\frac{d}{dt}\|f(t)\|^2_{H^s} = -\langle f, \nabla^{\perp}\phi\cdot\nabla f\rangle_{H^s}.$$

First assume that $s > 1$ so that $H^s(\mathbb{T}\times\mathbb{R}) \hookrightarrow L^\infty(\mathbb{T}\times\mathbb{R})$. The two leading terms are when all the derivatives land on $\nabla^{\perp}\phi$ or when all the derivatives land on ∇f. The latter term cancels and so we basically have the following (just using that $3/2 > 1 = d/2$):

$$-\langle f, \nabla^{\perp}\phi\cdot\nabla f\rangle_{H^s} \lesssim \left\|\nabla^{\perp}\phi\right\|_{H^{3/2}}\|f\|^2_{H^s} + \left\|\nabla^{\perp}\phi\right\|_{H^s}\|f\|_{H^s}\|\nabla f\|_{H^{3/2}}. \tag{6.7}$$

For the first term, we can pay regularity on f for decay using (5.1), and deduce

$$\left\|\nabla^{\perp}\phi\right\|_{H^{3/2}} \lesssim \langle t\rangle^{-2}\|f\|_{H^{9/2}}$$

Hence, if we take $s \geq 9/2$, the first term in the energy estimate becomes

$$\left\|\nabla^{\perp}\phi\right\|_{H^{3/2}}\|f\|^2_{H^s} \lesssim \langle t\rangle^{-2}\|f\|^3_{H^s}.$$

If this were the only term, this would be sufficient for stability, since we could get a global in time estimate for small enough initial data. This can be done via a continuity/bootstrap argument:

Exercise 6.1. Use a continuity argument to prove that if $X(t)$ is non-negative, continuous in time, and satisfies

$$\dot{X}(t) \leq \frac{C}{\langle t\rangle^2}(X(t))^2,$$

for some $C > 0$ then there exists an ϵ_0 depending only on C such that if $X(0) < \epsilon_0$ initially then $X(t) < 2\epsilon_0$ for all time.

Recall that a "continuity argument" amounts to assuming that $X(t) \leq 4\epsilon_0$ on a maximal time interval containing zero $[0, T^\star]$ and deducing that for ϵ_0 sufficiently small, one can deduce the stronger estimate $X(t) < 2\epsilon_0$. Continuity is used in two places: first, to prove that $T^\star > 0$ and second, to prove that the latter estimate implies $T^\star = +\infty$. The argument is basically a continuous form of induction; see [Tao06] for more discussion.

Now turn to the second term in (6.7), which is the difficult term. If we attempt to pay regularity again, we would run into trouble if we used (5.1), since in order to get something integrable we would need to use

$$\left\|\nabla^{\perp}\phi(t)\right\|_{H^s} \lesssim \frac{1}{\langle t\rangle^{1+\delta}}\|f(t)\|_{H^{s+2+\delta}}.$$

However, this is a massive loss of derivatives, worse than even the backwards heat equation, and we definitely cannot close a nonlinear estimate with this. Hence, if we don't think of something more sophisticated we have no chance of going forward.

Let us briefly consider the physical mechanism behind this loss of regularity. For a single Fourier mode, the linear evolution is given by $e^{ik(x-ty)+i\eta y} = e^{ikx+i(\eta-kt)y}$ and we are writing our vorticity as a linear combination of these waves (in the linear case). For sufficiently small data, we hope to prove that the linear term dominates (forgetting about the time-dependence in the shear flow), in which case maybe the nonlinear problem is in some sense written also as a superposition of these waves except now with amplitudes which are coupled together through the nonlinear term. We saw above that the critical time, when $t = \eta/k$, is when each mode in the vorticity has the largest nonlinear effect, not coincidentally, the regularity loss in (5.1) is required exactly to control the contribution of this time. From here one imagines a kind of nonlinear effect wherein a mode near its critical time has a strong nonlinear reaction and moves enstrophy to a mode which is still yet to reach its critical time, making another strong nonlinear response later in the future and potentially beginning a self-sustaining cycle. One could hope that this is disallowed by some special structure of the equations, however, it is not. Precisely this kind of behavior has been isolated in 2D Euler *experimentally* in 2005 [YDO05, YD02], following the much earlier famous experiments producing a similar kind of phenomenon in plasmas in 1968 [MWGO68] (in fact, both experiments were performed by the same laboratory). The effect is called an *echo* or when there is a sequence of echoes feeding into one another, an *echo cascade*. See also [VMW98, Van02]. Physically, we are worried about an echo cascade triggering a larger instability and transition away from Couette flow. In the proof of Landau damping in the Vlasov equations, analysis of these plasma echoes plays a major role; see [MV11, BMM13] (understanding the plasma echoes are central to both). See [FR14] for an example showing that in the absence of plasma echoes, a relatively easy proof is available for the Vlasov equations (they study a version of the Vlasov equations which does not permit echo cascades due to the particular structure of the equations).

Let us return to mathematics (sort of) and consider this loss of regularity more carefully. First, we saw in the above argument that the regularity loss occurs when the streamfunction has all the derivatives landing on it. So it makes sense to reduce our model (6.6) further to a toy model with a given smooth function q,

$$\partial_t f + \nabla^\perp \Delta_L^{-1} f_{k\neq 0} \cdot \nabla q = 0,$$

which is now a linear model. Next, we note that because $\widehat{\nabla^\perp \Delta_L^{-1} f} = i(\eta, -k)\left(|k|^2 + |\eta - kt|^2\right)^{-1} \hat{f}(k,\eta)$, the η term is the most dangerous since there is a $|k|$ in the denominator, so we reduce our toy model further:

$$\partial_t f - \partial_Y \Delta_L^{-1} f_{k\neq 0} \partial_z q = 0.$$

On the Fourier side this becomes the convolution:

$$\partial_t \hat{f}(t,k,\eta) = \sum_{\ell\neq 0} \int \frac{\xi(k-\ell)\hat{f}(t,\ell,\xi)}{|\ell|^2 + |\xi - t\ell|^2} \hat{q}(k-\ell, \eta-\xi) d\xi.$$

We are going to have to take absolute values since we won't have much information about the phase of q relative to f. Hence,

$$\partial_t \left|\hat{f}(t,k,\eta)\right| \lesssim \sum_{\ell\neq 0} \int \frac{\left|\xi(k-\ell)\hat{f}(t,\ell,\xi)\right|}{|\ell|^2 + |\xi - t\ell|^2} \left|\hat{q}(k-\ell,\eta-\xi)\right| d\xi.$$

We want to reduce this further by eliminating the convolution and approximating q as very well localized in Fourier transform (after all, q is supposed to be very smooth). This gives us the further reduced system:

$$\partial_t \left|\hat{f}(t,k,\eta)\right| \approx \sum_{\ell\neq 0;\ell\neq k} \frac{\left|\eta \hat{f}(t,\ell,\eta)\right|}{|\ell|^2 + |\eta - t\ell|^2}.$$

Now we have done something interesting, since for each t there's at most one specific ℓ_c, if t is close to η/ℓ_c, which is "critical", and all others are "non-critical". In fact, it will be sufficient to consider just two modes, ℓ_c and $\ell_c - 1$ (let us suppose ℓ is non-negative):

$$\begin{aligned}
\partial_t \left|\hat{f}(t,\ell_c,\eta)\right| &\approx \frac{\left|\eta \hat{f}(t,\ell_c-1,\eta)\right|}{|\ell_c-1|^2 + |\eta - t(\ell_c-1)|^2} \\
\partial_t \left|\hat{f}(t,\ell_c-1,\eta)\right| &\approx \frac{\left|\eta \hat{f}(t,\ell_c,\eta)\right|}{|\ell_c|^2 + |\eta - t\ell_c|^2}.
\end{aligned}$$

Next, we note that the most dangerous case is $|\eta| \gtrsim |\ell_c|^2$, so we will further reduce ourselves to under that condition. Finally, we note that there is a an interval of length $\approx \eta/\ell_c$ around the critical time over which $|\eta - t(\ell_c - t)| \gtrsim \eta/\ell_c$. Therefore, we further reduce this to

$$\begin{aligned}
\partial_t \left|\hat{f}(t,\ell_c,\eta)\right| &\approx \frac{\left|\ell_c^2 \hat{f}(t,\ell_c-1,\eta)\right|}{|\eta|} \\
\partial_t \left|\hat{f}(t,\ell_c-1,\eta)\right| &\approx \frac{\left|\eta \hat{f}(t,\ell_c,\eta)\right|}{|\ell_c|^2 + |\eta - t\ell_c|^2}.
\end{aligned}$$

This will be our toy model, the analogue of (5.6) from the linear case. As we did above, we want to find a Fourier multiplier which will encode an upper bound on the growth of these modes to use as our norm. This suggests we use a weight like the following:

$$\partial_t w_C(t,k,\eta) \approx \frac{|k|^2 \, w_{NC}(t,k,\eta)}{|\eta|} \tag{6.8a}$$

$$\partial_t w_{NC}(t,k,\eta) \approx \frac{|\eta| \, w_C(t,k,\eta)}{|k|^2 + |\eta - tk|^2}, \tag{6.8b}$$

where the idea is that if $t \approx \eta/k$ we will use w_C as the weight for the mode (k,η) and w_{NC} as the weight for modes (j,η) with $j \neq k$. What we have to do next is to determine the w_C and w_{NC} actually are. This ODE can actually be approximately solved; see [BM13]. However, we do not need to solve it, we only need to find an approximate super-solution, after all, we are using it to approximately upper bound the dynamics mode-by-mode. We can guess that the functions look something like the following, over one critical interval, for a constant $\kappa > 1$ which depends on the constants preferred in (6.8) (for the 2D works [BM13, BMV14] the constant is not important but in the 3D works [BGM15a, BGM15b] it is chosen with respect to some universal constants):

$$\begin{aligned}
w_{NC}(t,\eta) &\approx \left(\frac{k^2}{|\eta|}\left(1 + \left|\frac{\eta}{k} - t\right|\right)\right)^{-1-\kappa}, & \frac{\eta}{k} - \frac{\eta}{k(k+1)} &\lesssim t \leq \frac{\eta}{k} \\
w_{NC}(t,\eta) &\approx w_{NC}\left(\frac{\eta}{k},\eta\right)\left(1 + \left|\frac{\eta}{k} - t\right|\right)^{\kappa}, & \frac{\eta}{k} &\lesssim t \leq \frac{\eta}{k} + \frac{\eta}{k(k-1)}.
\end{aligned}$$

Of course we actually need to append this behavior together for each critical time. For technical reasons which are a bit obscure, this is done counting "backwards" from the last echo forward. Precisely, if we define $t_{k,\eta} = \frac{\eta}{k} - \frac{|\eta|}{|k(k-1)|}$ and $t_{0,\eta} = 2\eta$ we make the definition

$$w_{NC}(2|\eta|,\eta) = 1 \tag{6.9a}$$

$$w_{NC}(t,\eta) = \Big(\frac{k^2}{\eta}\Big[1 + b_{k,\eta}|t - \frac{\eta}{k}|\Big]\Big)^{\kappa} w_{NC}(t_{k-1,\eta},\eta), \qquad \forall t \in I^R_{k,\eta} = \Big[\frac{\eta}{k}, t_{k-1,\eta}\Big], \tag{6.9b}$$

$$w_{NC}(t,\eta) = \Big(1 + a_{k,\eta}|t - \frac{\eta}{k}|\Big)^{-1-\kappa} w_{NC}\Big(\frac{\eta}{k},\eta\Big), \qquad \forall t \in I^L_{k,\eta} = \Big[t_{k,\eta}, \frac{\eta}{k}\Big], \tag{6.9c}$$

for all $k^2 \le |\eta|$, where the constants $a_{k,\eta}, b_{k,\eta}$ are chosen such that $\Big(\frac{k^2}{\eta}\big[1 + b_{k,\eta}|t_{k-1,\eta} - \frac{\eta}{k}|\big]\Big) = 1$ and $\Big(1 + a_{k,\eta}|t_{k,\eta} - \frac{\eta}{k}|\Big) = (\eta/k^2)^{1+2\kappa} w(t_{k-1,\eta})$. These choices ensure that the multiplier is continuous in time (yes, the constants do go to zero as $|k|$ gets close to $\sqrt{|\eta|}$, which will be a slight but largely irrelevant technical difficulty). To get the critical regularity, we set

$$w_C(t,\eta) = \Big(\frac{k^2}{\eta}\Big[1 + b_{k,\eta}|t - \frac{\eta}{k}|\Big]\Big) w_{NC}(t,\eta), \qquad \forall t \in I^R_{k,\eta} = \Big[\frac{\eta}{k}, t_{k-1,\eta}\Big],$$

$$w_C(t,\eta) = \frac{k^2}{\eta}\Big(1 + a_{k,\eta}|t - \frac{\eta}{k}|\Big) w_{NC}(t,\eta), \qquad \forall t \in I^L_{k,\eta} = \Big[t_{k,\eta}, \frac{\eta}{k}\Big],$$

and we note that $w_C(t_{k,\eta},\eta) = w_{NC}(t_{k,\eta},\eta)$ and $w_C(\eta/k,\eta) = \frac{k^2}{\eta} w_{NC}(\eta/k,\eta)$ so that at the critical times, the regularity is separated by η/k^2. Now define $I_{k,\eta} = I^L_{k,\eta} \cup I^R_{k,\eta}$. To define the weight we will use in the energy estimates, we then get (denothing $E(r)$ to be the largest integer with $E(r) \le r$),

$$w(t,k,\eta) = \begin{cases} w(t_{E(\sqrt{|\eta|}),\eta}, k, \eta) & t < t_{E(\sqrt{|\eta|}),\eta} \\ w_{NC}(t,\eta) & t \in [t_{E(\sqrt{|\eta|}),\eta}, 2|\eta|] \setminus I_{k,\eta} \\ w_C(t,\eta) & t \in I_{k,\eta} \\ 1 & t \ge 2|\eta|. \end{cases} \tag{6.10}$$

By design, $w(t,k,\eta)$ is Lipschitz continuous in time.

6.2.1 Analysis of w

Analyzing the properties of w defined in (6.10) is not so easy as the definition is a bit complicated. The first question is...how big is it? First, by design $w(t,k,\eta) \le 1$, so actually we will care about a lower bound. By design we note

$$\frac{w(t_{k,\eta},k,\eta)}{w(t_{k+1,\eta},k,\eta)} = \Big(\frac{|\eta|}{k^2}\Big)^{1+2\kappa}.$$

Hence, the total growth is something like the following, denoting $N = E(\sqrt{|\eta|})$,

$$\frac{w(2|\eta|,k,\eta)}{w(1,k,\eta)} = \Big(\frac{|\eta|}{N^2}\Big)^{1+2\kappa}\Big(\frac{|\eta|}{(N-1)^2}\Big)^{1+2\kappa} \cdots \Big(\frac{|\eta|}{2^2}\Big)^{1+2\kappa}\Big(\frac{|\eta|}{1^2}\Big)^{1+2\kappa}.$$

From Stirling's formula we can derive:

Lemma 3 (Lemma 3.1 [BM13]). *Let $|\eta| \ge 1$. Then, there is some constant $\mu = \mu(\kappa) > 0$ such that*

$$\frac{1}{w(1,k,\eta)} \approx \frac{1}{|\eta|^{\mu/8}} e^{\frac{\mu}{2}\sqrt{|\eta|}}.$$

Another thing we should take a close look at is how critical and non-critical modes are related and what the time derivatives are. The most basic property is the following, which we can see from the definition in (6.10),

Lemma 4. *Let* $2\sqrt{|\eta|} \le t$, $t \in I_{k,\eta}$ *and* $k \neq \ell$, *then*

$$\frac{w(t,\ell,\eta)}{w(t,k,\eta)} \approx \frac{|\eta|}{k^2\left(1 + \left|t - \frac{\eta}{k}\right|\right)}.$$

Moreover,

$$\frac{\partial_t w(t,k,\eta)}{w(t,k,\eta)} \approx \frac{\partial_t w(t,\ell,\eta)}{w(t,\ell,\eta)} \approx \frac{1}{1 + \left|t - \frac{\eta}{k}\right|}.$$

We will also need various analogues of (5.8), however, clearly these will be much more difficult here than they were in (5.8). We will not get into the very technical details here, so let me just explain the most basic facts about them, noting that the full proof requires a bit more details.

Lemma 5 (From [BM13]). *There is a* $C > 0$ *such that*

$$\frac{w(t,k,\eta)}{w(t,k,\xi)} \lesssim e^{C|\eta-\xi|^{1/2}}. \tag{6.11}$$

For $2\max(\sqrt{|\xi|}, \sqrt{|\eta|}) \le t \le 2\min(|\eta|, |\xi|)$,

$$\frac{\partial_t w(t,k,\eta)}{w(t,k,\eta)} \lesssim \frac{\partial_t w(t,\ell,\eta)}{w(t,\ell,\eta)} \langle k-\ell, \eta-\xi\rangle^2.$$

The key point above is that relating w at different frequencies requires losing Gevrey regularity.

6.3 The set up

Next, the idea is to define a norm which is capable of measuring the solution and the coefficients. For important, but difficult to see, reasons we need to make our norm even more complicated than you might be thinking. Pick $s \in (1/2, 1)$ and $\sigma \gg 1$ fixed, a time-dependent regularity index $\lambda(t)$, and define the Fourier multiplier

$$A(t,k,\eta) = e^{\lambda(t)|k,\eta|^s} \langle k,\eta\rangle^\sigma \left[\frac{e^{\mu|\eta|^{1/2}}}{w(t,k,\eta)} + e^{\mu|k|^{1/2}}\right]. \tag{6.12}$$

The norm we will use to measure the vorticity f is then

$$\left\| A(t,\nabla) f(t) \right\|_{L^2}.$$

The index $\lambda(t)$ is strictly decreasing and bounded below by $\lambda(t) > (\lambda + \lambda')/2$ (from the statement of Theorem 6.1), so in the end we retain uniform control on f in Gevrey-$1/s$ regularity. At very high frequencies, the multiplier is more or less the same as Gevrey-$1/s$ since the corrections involving w are still much smaller (as $s > 1/2$). There is an important reason that $s < 1$ which will be more clear once one attempts to carry out energy estimates in Gevrey class – it has to do with controlling the Gevrey errors in (6.11) (hence, if we want to keep track of analytic regularity in the proof, the norm needs to be altered a little further; see [BM13]). The extra $e^{\mu|k|^{1/2}}$ has to do with difficulties that arise when dealing with commutators that arise when treating transport in Gevrey regularity;

see [BM13] for more on this important but subtle technicality. Along with the controls on the vorticity, we will need a similar kind of control on the coefficients g and v'. A difference is that we of course want to deduce decay on g; we eventually have a decay like $\|g\| \lesssim t^{C\epsilon-2}$ in suitable norms for some constant C (so for ϵ sufficiently small, this implies decay close to t^{-2}). We will not get into such technical difficulties here, however, the norms used to control g and v' are similar to that used to control f (a bit simpler since $k = 0$ for the coefficients) and so denote the combination of these norms as $\mathcal{E}_v(t)$; see [BM13] for the full definition.

Once we have our norms, the idea is to use a large bootstrap argument of the form (essentially): Let $[0, T^\star]$ be the largest interval such that

$$\|A(t,\nabla)f(t)\|_2^2 + \mathcal{E}_v(t) \leq 4\epsilon, \tag{6.13}$$

then for $\epsilon \leq \epsilon_0$ chosen sufficiently small, on $[0, T^\star]$ there holds

$$\|A(t,\nabla)f(t)\|_2^2 + \mathcal{E}_v(t) < 2\epsilon, \tag{6.14}$$

and therefore the vorticity f and the coefficients are uniformly controlled in Gevrey-$1/s$ regularity in the coordinate system (X, Y) for all $t \in [0, \infty)$. With a bit of effort to undo the coordinate system and transform (6.14) into information in (x, y) variables (not a trivial step, this requires some estimates on Gevrey regularity under compositions and a Gevrey class inverse function theorem – see [BM13]) the uniform bound is enough to prove Theorem 6.1). The hard part is proving that (6.13) implies (6.14).

6.4 Basic energy estimates

Let us consider how we would use (6.12) to prove that (6.13) implies (6.14) in the case of the simpler model

$$\begin{aligned} \partial_t f + \nabla^\perp \phi \cdot \nabla f &= 0 \\ \Delta_L \phi &= 0. \end{aligned}$$

Of course we start with:

$$\frac{1}{2}\frac{d}{dt}\|Af\|_{L^2}^2 = (Af, \partial_t Af) + \left(Af, A(\nabla^\perp \phi \cdot \nabla f)\right).$$

The idea is to break the second term down into terms which can be absorbed with the first term and errors which are small and integrable in time. As when we were deducing H^s bounds on Euler, we want to eliminate the term where "all the derivatives" land on ∇f the second inner product. The analogue of this is, using the divergence free condition:

$$\frac{1}{2}\frac{d}{dt}\|Af\|_{L^2}^2 = (Af, \partial_t Af) + \left(Af, A(\nabla^\perp \phi \cdot \nabla f)\right) - \left(Af, \nabla^\perp \phi \cdot A\nabla f\right).$$

Next, we want to make more rigorous what we mean by distributing a Fourier multiplier such as A like one would distribute derivatives. For this we use what is called a "paraproduct", which as far as we are concerned with here, just provides a clever technique for breaking up contributions in frequency using Littlewood-Paley decompositions (introduced by Bony [Bon81]),

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left\|Af\right\|_{L^2}^2 &= (Af, \partial_t Af) + \sum_{N\in 2^{\mathbb{Z}}} \Big(Af, A(\nabla^{\perp}\phi_N \cdot \nabla f_{<N/8})\Big) - \Big(Af, \nabla^{\perp}\phi_N \cdot A\nabla f_{<N/8}\Big) \\ &\quad + \sum_{N\in 2^{\mathbb{Z}}} \Big(Af, A(\nabla^{\perp}\phi_{<N/8} \cdot \nabla f_N)\Big) - \Big(Af, \nabla^{\perp}\phi_{<N/8} \cdot A\nabla f_N\Big) \\ &\quad + \sum_{N\in 2^{\mathbb{Z}}} \sum_{N/8\le N'\le 8N} \Big(Af, A(\nabla^{\perp}\phi_{N'} \cdot \nabla f_N)\Big) - \Big(Af, \nabla^{\perp}\phi_{N'} \cdot A\nabla f_N\Big) \\ &= CK_A + R + T + \mathcal{R}.\end{aligned}$$

The R stands for "reaction" due to the fact that it is more or less analogous to the "reaction" terms in Mouhot and Villani's original proof of Landau damping [MV11]. The T stands for "transport" and $\mathcal{R}$ stands for "remainder". We can think of the reaction term as the term where "all" of A "lands" on ϕ and the transport term where "all" of A "lands" on ∇f. The remainder term is the analogue of the all the cross terms in Liebniz's rule. The commutator we introduced by using the divergence free condition is actually only for dealing with the transport term, as is the extra $e^{\mu|k|^{1/2}}$ in (6.12).

The reaction term is the one which we used to derive the toy model in §6.2 (in fact, just one term of it). To see exactly how to use w in this context we will need to learn how to make Littlewood-Paley projections and Gevrey regularity play together in a nice way, which requires some concavity estimates on x^s for $s < 1$ which will allow us to say that in the reaction term "most of A" "lands" on ϕ and vice-versa in the transport terms. That is, we need some precise equivalent of (5.8a) for Gevrey regularity. What is used is the following: for $|x - y| < |x|/2$, if $s \in (0, 1)$, there is a constant $c = c(s) \in (0, 1)$ such that

$$e^{\lambda|x|^s} \le e^{\lambda|y|^s} e^{c\lambda|x-y|^s}. \tag{6.15}$$

If one considers one of the reaction terms:

$$\sum_{N\in 2^{\mathbb{Z}}} \Big(Af, A(\nabla^{\perp}\phi_N \cdot \nabla f_{<N/8})\Big) = \sum_{N\in 2^{\mathbb{Z}}} \Big(\widehat{Af}, A(\widehat{\nabla^{\perp}\phi_N} * \widehat{\nabla f_{<N/8}})\Big)$$

then its observed that the term involving $\nabla^{\perp}\phi_N \cdot \nabla f_{<N/8}$ only concerns frequencies of approximately N and so that contribution to the sum also only involves frequencies of Af which are comparable to N as well. Hence (6.15) implies that "most" of the Gevrey regularity in the multiplier A becomes associated with $\nabla^{\perp}\phi_N$, rather than $\nabla f_{<N/8}$. Just as (5.8) allowed us to treat the linear problem in Theorem 5.2 as a perturbation of the linear toy model in (5.6), the inequalities and properties we have enumerated on A will allow us to treat the reaction term as a perturbation of the nonlinear toy model we derived above in (6.8). See [BM13] for more details on exactly how to carry out this procedure.

7 2D Navier-Stokes near Couette flow

Theorem 6.1 confirms that the inviscid damping predicted by Orr on the linear level persists in the 2D Euler equations. The next natural question is whether or not the enhanced dissipation predicted by Lord Kelvin persists in the Navier-Stokes equations. The answer is indeed yes, as we verfied in [BMV14]. In order to see this, what we need to do is to take the Navier-Stokes equations with inverse Reynolds number $\nu > 0$ and then study the vanishing viscosity limit $\nu \to 0$ in a way which gives us information that is uniform in t. This is an important point: if one considers a finite time interval $[0, T]$, then we know that Navier-Stokes simply converges to Euler as $\nu \to 0$ (because there are no boundaries!) whereas if we fix ν and then send $T \to \infty$ we just get that the Navier-Stokes equations behaves asymptotically like the heat equation and everything dissipates[6]. What we need to do is a sort of "long-time" inviscid limit, which allows us to retain uniform control on the dynamics at all times simultaneously as we send $\nu \to 0$, which is harder than just studying $\nu > 0$ fixed or setting $\nu = 0$ as in theorem 6.1. More precisely, the theorem we prove is the following (abbreviated and simplified for presentation), which is necessarily harder than Theorem 6.1 (basically the main estimates in the proof of Theorem 6.1 become lemmas in the proof of Theorem 7.1):

Theorem 7.1 (Long-time inviscid limit for the 2D Navier-Stokes equations [BMV14]). *Let $s \in (1/2, 1]$, let $\lambda > \lambda' > 0$, and $\alpha \geq 1$. There exists an $\epsilon_0 = \epsilon_0(\lambda, \lambda', s, \alpha)$ such that if*

$$\|u_{in}\|_{L^2} + \|\omega_{in}\|_{\mathcal{G}^{\lambda;s}} = \epsilon < \epsilon_0,$$

then the following holds with implicit constants that are independent of ν, ϵ, and t:

$$\|\omega_{\neq}(t, x + ty + t\phi(t, y), y)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t^3 \rangle^{\alpha}} \tag{7.1}$$

$$\|\overline{\omega}(t, y)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t \rangle^{1/4}} \tag{7.2}$$

$$\|u_{1,\neq}(t)\|_{L^2} + \langle t \rangle \|u_2(t)\|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle \langle \nu t^3 \rangle^{\alpha}}. \tag{7.3}$$

Moreover, for $t \ll \nu^{-1/3}$ one can deduce qualitative behavior similar to Theorem 6.1; see [BMV14] for more details.

Remark 12. The regularity requirement can be relaxed a tiny bit using the instant regularization due to the viscosity. That is, we don't really need our initial data to be *literally* Gevrey-$\frac{1}{s}$, it just needs to be small in Gevrey-$\frac{1}{s}$ by time $t = 1$ say. In [BMV14], we show that one can take initial data which can be split into smooth and rough part: $\omega_{in} = \omega_S + \omega_R$ with $\|\omega_S\|_{\mathcal{G}^{\lambda;s}} < \epsilon/2$ and $e^{K\nu^{-p}} \|\omega_R\|_{L^2} < \epsilon/2$ for suitable $K = K(s, \lambda) > 0$ and $p = p(s) > 0$. Hence, one can have data which is technically only L^2 but which is *very* close to Gevrey-$\frac{1}{s}$ as $\nu \to 0$. We remark that this class of initial data is very natural for inviscid limits.

Remark 13. It is worth emphasizing again that even in Theorem 7.1, the enhanced dissipation is not really the mechanism which is providing the stability – it is still the inviscid damping that is really stabilizing the flow. This is because we are allowing ϵ, t, and ν to be totally unrelated. In the 3D works [BGM15a, BGM15b], we will see that it is different: the enhanced dissipation plays the role of the main stabilizing effect and ϵ and ν must be related.

Remark 14. We will return to this point later, but note that because we have a uniform basin of stability as $\nu \to 0$, we can say that there is *no* "subcritical transition", as described in §1.1 in the 2D Couette flow for sufficiently smooth perturbations.

[6] simply using the divergence free condition, one can easily prove that the solution dissipates with the same rate as the heat equation, but we did not get so much precise information about the asymptotics to say it really behaves exactly like the heat equation (though it does for times $t \gg \nu^{-1}$)

7.1 New nonlinear coordinate system

The proof of Theorem 7.1 is an amplification of the proof of Theorem 6.1 in several ways. First, we need to find a coordinate system. As in §6.1, we look for a coordinate transform of the form

$$\begin{aligned} X &= x - ty - th(t,y) \\ Y &= y + h(t,y), \end{aligned}$$

so that we still have the important property of how derivatives transform: if $f(t, X(t,x,y), Y(t,y)) = \omega(t,x,y)$ and $\phi(t, X(t,x,y), Y(t,y)) = \psi(t,x,y)$, then

$$\begin{aligned} \partial_x \omega(t,x,y) &= (\partial_X f)(t, X(t,x,y), Y(t,y)) \\ \partial_y \omega &= (1 + \partial_y h(t,y)) \left((\partial_Y - t\partial_X) f \right)(t, X(t,x,y), Y(t,y)), \end{aligned}$$

and similar for ϕ. If we derive as we did in §6.1, then we arrive at basically the same system we had in the 2D Euler case except now with the dissipation on the RHS:

$$\begin{aligned} \partial_t f + \begin{pmatrix} -\frac{d}{dt}(th) + \bar{u}_1 \\ \partial_t h \end{pmatrix} \cdot \begin{pmatrix} \partial_X f \\ \partial_Y f \end{pmatrix} + (1 + \partial_y h) \nabla^{\perp}_{X,Y} \phi_{k \neq 0} \cdot \nabla_{X,Y} f &= \nu \Delta_t f \\ \Delta_t \phi &= f. \end{aligned}$$

It is very natural to think that we should again pick

$$\frac{d}{dt}(th) = \bar{u}_1,$$

however, $\bar{u}_1$ is not the only contribution to the shear flow anymore! Indeed, recall

$$\partial_{XX} + (1 + \partial_y h)^2 (\partial_Y - t\partial_X)^2 + \partial_{yy} h (\partial_Y - t\partial_X) := \Delta_t,$$

and hence the system is more naturally written as

$$\begin{aligned} \partial_t f + \begin{pmatrix} -\frac{d}{dt}(th) + \bar{u}_1 + t\partial_{yy} h \\ \partial_t h - \partial_{yy} h \end{pmatrix} \cdot \begin{pmatrix} \partial_X f \\ \partial_Y f \end{pmatrix} + (1 + \partial_y h) \nabla^{\perp}_{X,Y} \phi_{k \neq 0} \cdot \nabla_{X,Y} f &= \nu \tilde{\Delta}_t f \\ \Delta_t \phi &= f, \end{aligned}$$

where the operator on the RHS is given by:

$$\partial_{XX} + (1 + \partial_y h)^2 (\partial_Y - t\partial_X)^2 := \tilde{\Delta}_t.$$

Now, it becomes more clear that we should *actually* pick h to solve the parabolic equation:

$$\frac{d}{dt}(th) = \bar{u}_1 + \nu t \partial_{yy} h, \tag{7.4}$$

as this eliminates the shear flow in the first coordinate of the velocity field. Moreover, if we again denote $C(t, Y(t,y)) = h(t,y)$ and $\overline{U_1}(t, Y(t,y)) = \overline{u_1}(t,y)$, and

$$g = \frac{\overline{U_1} - C}{t},$$

then we derive

$$\partial_t f + g \partial_Y f + (1 + \partial_y h) \nabla^{\perp}_{X,Y} \phi_{k \neq 0} \cdot \nabla_{X,Y} f = \nu \tilde{\Delta}_t f \tag{7.5}$$

$$\Delta_t \phi = f. \tag{7.6}$$

Although in some sense we have define g in the same way as we did in the Euler case, note that its relationship to h is a bit different (it is no longer related to $\partial_t h$ but instead $\partial_t h - \partial_{yy} h$). If one works it out (good exercise) one derives for $v' = 1 + \partial_y h$ and g:

$$\partial_t g + g\partial_Y g = -\frac{2}{t}g - \frac{v'}{t}\overline{(U_{\neq} \cdot \nabla(U_1)_{\neq})} + \nu\tilde{\Delta}_t g$$
$$\partial_t(v'-1) + g\partial_Y(v'-1) = \frac{1}{t}\left(-\bar{f} - (v'-1)\right) + \nu\tilde{\Delta}_t(v'-1).$$

One can also check that if we did not include dissipation in the coordinate system, we would not have nice dissipation terms in these equations (actually, this is how we first derived the coordinate system – it was pointed out to us by Pierre Germain during the writing of [BGM15a, BGM15b] that there was an easier but equivalent, and more systematic, way of viewing the same coordinate system).

7.2 Enhanced dissipation estimates

Next, our main goal is to get energy estimates on f which are uniform in ν, ϵ, and t and also to quantify the enhanced dissipation. Since we want estimates which basically match those made on Euler, it makes sense to use the same A and $\mathcal{E}_v$ that we employed therein. To quantify the enhanced dissipation, we use a simpler multiplier:

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

with β fixed such that $\sigma \gg \beta + 3\alpha$ and D defined via

$$D(t,\eta) = \frac{1}{3\alpha}\min\left(|\eta|^3, \frac{1}{8}t^3\right).$$

The purpose of this multiplier is to make the norm *stronger* as time increases for $t > 2|\eta|$. This time is well past all of the critical times because $k \in \mathbb{Z}$, hence this multiplier is adapted to the Orr mechanism. Indeed, transient un-mixing will slow down the enhanced dissipation back to regular (slow) dissipation near $t \sim \eta/k$. We could try to use a much more precise D which keeps track of this in a way which is more dependent on k, however, it turns out the following is sufficient. In particular, this definition of D is sufficient to quantify the enhanced dissipation in the sense that:

$$\left\|\langle\nabla\rangle^\beta f_{\neq}\right\|_{\mathcal{G}^{\lambda;s}} \lesssim \langle \nu t^3\rangle^{-\alpha}\left\|A^\nu(t,\nabla)f\right\|_{L^2},$$

so that control of $\left\|A^\nu(t,\nabla)f\right\|_{L^2}$ implies the enhanced dissipation of f.

Exercise 7.1. Verify that

$$(A^\nu f, \partial_t A^\nu f) + \nu(A^\nu f, A^\nu \Delta_L f) \leq \dot{\lambda}\left\||\nabla|^{s/2}A^\nu f\right\|_{L^2}^2 - \frac{\nu}{8}\left\|\sqrt{-\Delta_L}A^\nu f\right\|_{L^2}^2.$$

Hence, we have an alternative (rather silly and inefficient) proof of the enhanced dissipation of the PDE

$$\partial_t f = \nu\Delta_L f,$$

(which of course we have already solved exactly). Note that getting enhanced dissipation estimates using $A^\nu(t,\nabla)$ is very precise, but it does cost regularity, and in some settings the would be considered quite high (in this particular setting of course, losing a couple hundred Sobolev derivatives is totally irrelevant compared to the Gevrey regularity we are already condemned to lose). There are other methods of quantifying enhanced dissipation which are less costly but less precise, however, they can get the job done as well [CKRZ, BW13, BGM15c]...

Also note that $\sigma \gg \beta+3\alpha$ (as well as the term involving w) ensures that $A(t,\nabla)$ is much stronger at high frequencies than A^{ν}. This will make the energy estimate to control $A^{\nu}(t,\nabla)f$ *much* easier than if we had tried to make one "master" multiplier that involved $A(t,\nabla)$ from Euler as well as D, because now when we are estimating A^{ν}, the frequencies $t \leq 2|\eta|$ are not very relevant, as $A^{\nu}\mathbf{1}_{t<2|\eta|} \lesssim \langle t\rangle^{-\gamma} A\mathbf{1}_{t<2|\eta|}$ for any $\gamma > 0$, so any nonlinear effects occuring in that range of frequencies is easily controlled by an estimate on A. This trick of coupling together a high norm estimate for a uniform bound on regularity and a low norm estimate for decay estimates is a classical, tried and true method in PDEs, and is very, very useful. Hence, the basic idea of the proof of Theorem 7.1 is basically the bootstrap: Let $[0,T^{\star}]$ be the largest interval such that, for some suitable C,

$$\|A(t,\nabla)f(t)\|_{L^2}^2 + \|A^{\nu}(t,\nabla)f(t)\|_{L^2}^2 + \mathcal{E}_v(t) \leq 4C\epsilon, \tag{7.7}$$

then for $\epsilon \leq \epsilon_0$ chosen sufficiently small, on $[0,T^{\star}]$ there holds

$$\|A(t,\nabla)f(t)\|_{L^2}^2 + \|A^{\nu}(t,\nabla)f(t)\|_{L^2}^2 + \mathcal{E}_v(t) < 2C\epsilon. \tag{7.8}$$

A natural thing to expect is that the estimates on $\|Af\|_{L^2}$ and $\mathcal{E}_v(t)$ are exactly the same as in the proof of Theorem 6.1, after all, we only added dissipation to (7.5), so how could we make the energy estimate worse? Well, actually we did: the point is that, unlike if we just had Δ or Δ_L, neither $(f,\tilde{\Delta}_t f)$ nor $\left(Af, A\left(\tilde{\Delta}_t f\right)\right)$ are obviously negative definite because of the coefficients coming from v'. After some integration by parts this might look easier if v' is small in a suitble sense, but this loses too many derivatives on v', which remember is an unknown we need to solve for and v' isn't really much smoother than f (it actually turns out to be a tiny bit smoother than f for subtle reasons; see [BM13] for more information on this). In fact, dealing with this requires some genuine work, and in particular, one actually has to use the enhanced dissipation estimates in order to deal with the error terms in the $\|Af\|_{L^2}$ estimate. In this sense, the estimate even at the higher norm $\|Af\|_2$ is not "free". The full details however, are a bit too technical for now, see [BMV14] for the full story...

8 Dynamics near the subcritical transition of the 3D Couette flow

Squire's theorem is misleading because it suggests that 3D dynamics is somehow very similar to 2D dynamics. However, this is far from correct. In this section, we will discuss how 2D and 3D differ in the case of the Couette flow.

Recall that at the start of the course, we used the example of subcritical transition to partially motivate the entire course. However, Theorems 6.1 and 7.1 show that for smooth enough disturbances, *subcritical transition does not occur in 2D*. This is in glaring contrast to experiments, which suggest that basically every 3D laminar flow is practically unstable at high Reynolds number (it is very hard to do experiments on 2D laminar flows). One could suggest that it is just a matter of regularity and that 2D flows are the same as 3D flows. While it is almost surely true that 2D flows will experience subcritical transition at low enough regularities, this is a very unsatisfactory answer that would completely fail to explain the 3D experiments [SH01] which clearly show that the instabilities are fundamentally 3D and would (if it were true, which it is not) arguably emphasize fundamental limitations of using mathematical analysis to understand subtle physical phenomena. Fortunately for us mathematicians, this is not the case – indeed, subtle phenomena, where experiments, numerics, and formal asymptotics cannot quite tread, are *exactly* the place where mathematical analysis *should* be *most* useful. Instead, in this section, we will discuss a set of theorems which show that, even for smooth, Gevery class perturbations, 3D flows are subject to 3D specific instabilities which result in subcritical transition (further emphasizing the limitations of Squire's theorem and clearly displaying the difference between 2D and 3D hydrodynamic stability).

We consider the problem in the domain $(x, z) \in \mathbb{T}^2$ and $y \in \mathbb{R}$: if $u + (y, 0, 0)^t$ solves the Navier-Stokes equations,

$$\partial_t u + y\partial_x u + u \cdot \nabla u + \nabla p^{NL} = \begin{pmatrix} -u_2 \\ 0 \\ 0 \end{pmatrix} - \nabla p^L + \nu \Delta u \tag{8.1a}$$

$$\Delta p^{NL} = -\partial_i u_j \partial_j u_i \tag{8.1b}$$

$$\Delta p^L = -2\partial_x u_2 \tag{8.1c}$$

$$\nabla \cdot u = 0, \tag{8.1d}$$

where $\nu = \mathbf{Re}^{-1}$ denotes the inverse Reynolds number, p^{NL} is the nonlinear contribution to the pressure due to the disturbance and p^L is the linear contribution to the pressure due to the interaction between the disturbance and the Couette flow.

It was suggested by Lord Kelvin [Kel87] (and basically Reynolds too [Rey83]) that subcritical transition is due to the fact that while the flow may be technically stable at all finite Reynolds number, it is progressively more unstable as $\nu \to 0$. It becomes natural to, for each norm N, try to find a $\gamma = \gamma(N)$ such that

$$\begin{aligned} \|u_{in}\|_N \lesssim \nu^\gamma \quad &\Rightarrow \quad \text{stability} \\ \|u_{in}\|_N \gg \nu^\gamma \quad &\Rightarrow \quad \text{possible instability.} \end{aligned}$$

Hence, the goal is not just to determine that there *is* a basin of stability, but to obtain nearly sharp estimates on the size of the basin as $\nu \to 0$ (the latter generally being much harder for nonlinear problems). This γ (if it exists) is called the *transition threshold*. A lot of work has been done attempting to estimate it for various laminar flows; see e.g. [SH01, Yag12] and the references therein (and further references below).

8.1 Streaks

In 2D, we saw that the x-dependence of the solutions rapidly decayed via enhanced dissipation and inviscid damping. In 3D, we have a much larger class of x-independent solutions, which are usually referred to as *streaks*, due the streak-like appearance of the relatively fast fluid in experiments and computations [TTRD93, SH01, TE05, BDDM98]. We will adopt this terminology here as well. In particular, we have the following large class of global solutions to 3D Navier-Stokes/Euler:

Proposition 1 (Streak solutions). *Let $\nu \in [0, \infty)$, $u_{in} \in H^{5/2+}$ be divergence free and independent of x, that is, $u_{in}(x, y, z) = u_{in}(y, z)$, and denote by $u(t)$ the corresponding unique strong solution to* (8.1) *with initial data u_{in}. Then $u(t)$ is global in time and for all $T > 0$, $u(t) \in L^\infty((0, T); H^{5/2+}(\mathbb{R}^3))$. Moreover, the pair $(u_2(t), u_3(t))$ solves the 2D Navier-Stokes/Euler equations on $(y, z) \in \mathbb{R} \times \mathbb{T}$:*

$$\partial_t u_i + (u_2, u_3) \cdot \nabla u_i = -\partial_i p + \nu \Delta u_i \tag{8.2a}$$

$$\partial_y u_2 + \partial_z u_3 = 0, \tag{8.2b}$$

and u^1 solves the (linear) forced advection-diffusion equation

$$\partial_t u_1 + (u_2, u_3) \cdot \nabla u_1 = -u_2 + \nu \Delta u_1. \tag{8.3}$$

From (8.3) we see something interesting that is not present in 2D: the x-average of $-u_2$ will disappear in 2D due to the divergence free condition, however, this term does not vanish in 3D and will create a large forcing on the x-average of u_1. This effect is called the *lift-up effect*, and was first noticed by Ellingsen and Palm in [EP75] and studied later in [Lan80] in the non-periodic setting. One can verify that every non-trivial shear flow will have a similar effect, an instability which is completely missed by the notion of spectral stability since it is essentially of the form (to leading order; see below for a more careful derivation):

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \nu\Delta & -1 \\ 0 & \nu\Delta \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

which is basically one of the canonical non-diagonalizable systems of 2x2 linear ODEs.

8.2 Linearized 3D Euler and Navier-Stokes

We will see that the streaks in Proposition 1 rapidly attract all sufficiently smooth solutions which are not too large and hence dominate all of the dynamics. The 2D work suggests that this could be possible via the enhanced dissipation effect, however, we will need a much more comprehensive and clear picture of the linear dynamics.

First, let us study the linearized Navier-Stokes equations:

$$\partial_t u + y\partial_x u = \begin{pmatrix} -u_2 \\ 0 \\ 0 \end{pmatrix} - \nabla p^L + \nu\Delta u \tag{8.4a}$$

$$\Delta p^L = -2\partial_x u_2 \tag{8.4b}$$

$$\nabla \cdot u = 0. \tag{8.4c}$$

It turns out that (8.4) can be solved analytically. The trick goes back to Lord Kelvin [Kel87]: define a new variable $q_2 = \Delta u_2$, and a computation shows that it becomes

$$\partial_t q + y\partial_x q = \nu\Delta q.$$

The same passive scalar as the vorticity solves in the 2D case! However, note carefully that q is NOT a direct analogue of the vorticity – it is two derivatives instead of one derivative. This will have a number of very important implications. Therefore, if $Q(t, X, y, z) = q_2(t, X + ty, y, z)$ there holds

$$\hat{Q}(t, k, \eta, l) = \hat{Q}(0, t, k, l) \exp\left[-\nu \int_0^t |k|^2 + |\eta - k\tau|^2 + |l|^2 \, d\tau\right].$$

As in the 2D vorticity equations, this implies the enhanced dissipation:

$$\|Q_{\neq}(t)\|_{H^\sigma} \lesssim \|Q(0)\|_{H^\sigma} e^{-c\nu t^3}.$$

If we define $U_2(t, X, y, z) = u(t, X + ty, y, z)$ then $Q = \Delta_L U_2$ and hence

$$\widehat{U_2}(t, k, \eta, l) = -\frac{\widehat{Q}(t, k, \eta, l)}{|k|^2 + |\eta - kt|^2 + |l|^2}.$$

Hence we get the analogue of (5.1):

$$\|U_{2,\neq}(t)\|_{H^\sigma} \lesssim \langle t\rangle^{-2} \|Q_{\neq}(t)\|_{H^{\sigma+2}}.$$

Hence $U_{2,\neq}$ still experiences inviscid damping as in the 2D case. Now consider the equations that $U_{1,3}$ satisfy

$$\begin{aligned}\partial_t U_1 &= -U_2 - \partial_X P^L + \nu \Delta_L U_1 \\ \partial_t U_3 &= -\partial_z P^L + \nu \Delta_L U_3 \\ \Delta_L P^L &= -2\partial_X U_2.\end{aligned}$$

From argument similar to (5.1) we get the very rapid decay of the pressure:

$$\left\| P^L_{\neq} \right\|_{H^\sigma} \lesssim \langle t \rangle^{-2} \left\| U_{2,\neq}(t) \right\|_{H^{\sigma+3}} \lesssim \langle t \rangle^{-4} \left\| Q_{\neq}(t) \right\|_{H^{\sigma+5}}.$$

Hence U_1 and U_3 do not really change much between the times $1 \lesssim t \lesssim \nu^{-1/3}$. This shows two things: first, there is in general *no* inviscid damping on U_1 and U_3 (although by the divergence free condition, there is on $\partial_X U_1 + \partial_z U_3$); and second, for those intermediate times, we get something like $u_{1,3}(t,x,y,z) \approx f_{1,3}(t, x-ty, y, z)$ and so these components are being mixed similar to a passive scalar until the enhanced dissipation kicks in. Hence, we see a cascade of *kinetic energy* to high frequencies in 3D, whereas in 2D we see only a cascade of *enstrophy* to high frequencies. Although inviscid damping will not occur, we still get enhanced dissipation:

$$\left\| U_{1,\neq} \right\|_{H^\sigma} + \left\| U_{3,\neq} \right\|_{H^\sigma} \lesssim \left\| U(0) \right\|_{H^{\sigma+5}} e^{-c\nu t^3}.$$

If one takes x averages of the equations for u^i we derive the non-normal system

$$\partial_t \begin{pmatrix} \overline{u_1} \\ \overline{u_2} \\ \overline{u_3} \end{pmatrix} = \begin{pmatrix} \nu\Delta & -1 & 0 \\ 0 & \nu\Delta & 0 \\ 0 & 0 & \nu\Delta \end{pmatrix} \begin{pmatrix} \overline{u_1} \\ \overline{u_2} \\ \overline{u_3} \end{pmatrix}.$$

In fact we can solve it and see that $\overline{u_1}(t) = e^{\nu t \Delta}\left(\overline{u_1}(0) - t\overline{u_2}(0)\right)$, which clearly shows the lift-up effect.

8.3 Subcritical transition

On the linear level, the x dependence of the solution gets wiped out by the dissipation after times $t \gtrsim \nu^{-1/3}$. For data smaller than $O(\nu^{1/3})$ in size, this is sooner than the lift-up effect will dominate. Hence, as long as the perturbations aren't that large, we expect the streaks to attract all dynamics. It is important to notice that, unlike in 2D, in 3D the enhanced dissipation is the primary stability mechanism: the inviscid damping only controls one component of the solution and definitely would not force solutions back to streaks the way inviscid damping causes convergence back to shear flows in 2D. This is going to change the nature of the proofs a lot, as most nonlinear effects are absorbed by the dissipation rather than dealt with using inviscid damping. Instead, the inviscid damping in 3D provides a kind of "null structure" that damps the top most dangerous leading order terms (c.f. the null condition in quasilinear wave equations [Kla86, Chr86]).

That the dynamics of all solutions near the transition threshold (which is determined to be $\gamma = 1$) are determined by the streaks to which they converge is the content of our main results in [BGM15a, BGM15b] (for initial data which is sufficiently smooth). In [BGM15a] we consider solutions below the $\gamma = 1$ threshold and prove the following. Note that the Gevrey regularity class is the same as in 2D. This is not exactly a coincidence, but it is also by no means clear a priori. The 2D nonlinear effects are far too weak to be relevant in the proof of Theorem 8.1 (more precisely, they are easily overpowered by the enhanced dissipation) and the 3D analogue of (6.8) is rather different. However, in the end, both predict the same regularity class (although the 2D and 3D toy models do not predict the same norm, though they are similar).

Theorem 8.1 (Below threshold case from [BGM15a]). *For all $s \in (1/2, 1)$, all $\lambda_0 > \lambda' > 0$, all integers $\alpha \geq 10$, all $\delta_1 > 0$, and all $\nu \in (0, 1]$, there exists constants $c_{00} = c_{00}(s, \lambda_0, \lambda', \alpha, \delta_1)$ and $K_0 = K_0(s, \lambda_0, \lambda')$ (both independent of ν), such that for all $c_0 \leq c_{00}$ and $\epsilon < c_0\nu$, if $u_{in} \in L^2$ is a divergence-free vector field that can be written $u_{in} = u_S + u_R$ (both also divergence-free) with*

$$\|u_S\|_{\mathcal{G}^{\lambda;s}} + e^{K_0 \nu^{-\frac{3s}{2(1-s)}}} \|u_R\|_{H^3} < \epsilon, \tag{8.5}$$

then the unique, classical solution $u(t)$ to (8.1) with initial data u_{in} is global in time and the following estimates hold with all implicit constants independent of ν, ϵ, t and c_0:

(i) transient growth of the streak: if $t < \frac{1}{\nu}$,

$$\left\|\overline{u_1}(t) - \left(e^{\nu t\Delta}\left(\overline{u_1}(0) - t\overline{u_2}(0)\right)\right)\right\|_{\mathcal{G}^{\lambda';s}} \lesssim c_0^2 \tag{8.6a}$$

$$\left\|\overline{u_2}(t) - e^{\nu t\Delta}\overline{u_2}(0)\right\|_{\mathcal{G}^{\lambda';s}} + \left\|\overline{u_3}(t) - e^{\nu t\Delta}\overline{u_3}(0)\right\|_{\mathcal{G}^{\lambda';s}} \lesssim c_0\epsilon \tag{8.6b}$$

(ii) uniform bounds and decay of the background streak

$$\|\overline{u_1}(t)\|_{\mathcal{G}^{\lambda';s}} \lesssim \min\left(\epsilon\langle t\rangle, c_0\right) \tag{8.7a}$$

$$\|\overline{u_2}(t)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t\rangle^{\alpha}} \tag{8.7b}$$

$$\|\overline{u_3}(t)\|_{\mathcal{G}^{\lambda';s}} \lesssim \epsilon \tag{8.7c}$$

$$\|\overline{u_1}(t)\|_4 \lesssim \frac{c_0}{\langle \nu t\rangle^{1/4}} \tag{8.7d}$$

$$\|\overline{u_3}(t)\|_4 \lesssim \frac{\epsilon}{\langle \nu t\rangle^{1/4}}; \tag{8.7e}$$

(iii) the rapid convergence to a streak

$$\|u_{1,\neq}(t, x+ty+t\psi(t,y,z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon\langle t\rangle^{\delta_1}}{\langle \nu t^3\rangle^{\alpha}} \tag{8.8a}$$

$$\|u_{2,\neq}(t, x+ty+t\psi(t,y,z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle t\rangle^{2-\delta_1}\langle \nu t^3\rangle^{\alpha}}, \tag{8.8b}$$

$$\|u_{3,\neq}(t, x+ty+t\psi(t,y,z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t^3\rangle^{\alpha}}. \tag{8.8c}$$

Here $\psi(t,y,z)$ is an $O(c_0)$ correction to the mixing which is approximately $\overline{u_1}(t)$.

Remark 15. Do not be misled: the solutions in question are $O(1)$ solutions to the Navier-Stokes equations, it is only that they are $O(\nu)$ away from the Couette flow. This is an important distinction. We saw in the previous course that if solutions to Navier-Stokes are initially $O(\nu)$ (in the appropriate norms) then they are automatically global. However, if one tries to use a similar argument here without a very intricate analysis, it will be hard to get results which are much better than logarithmic or a much worse power of ν! Also note that easier methods do still get you *some* nonlinear stability, but remember, the main interest of Theorem 8.1 is finding the largest size that still ensures stability (that is, we want to get an estimate of the basin of stability which is at least close to optimal in terms of the norm of the initial data).

Remark 16. If $\overline{u_2}(0)$ is such that $\|\overline{u_2}(0)\|_{\mathcal{G}^{\lambda';s}} \geq \frac{1}{4}\epsilon = \frac{1}{16}c_0\nu$ then (8.10) shows that for c_0 small (but independent of ϵ and ν), the streak $\overline{u_1}(t)$ reaches the maximal amplitude of $\|\overline{u_1}(t)\|_2 \gtrsim c_0$. In this sense, Theorem 8.1 includes perturbations which undergo dramatic growth of kinetic energy before decaying, so much in fact, that the solutions go from $O(\epsilon)$ to $O(c_0)$ before eventually decaying. Since c_0 is independent of ϵ, these solutions are necessarily on the edge of the weakly nonlinear regime.

Theorem 8.1 covers only solutions which do not undergo any kind of transition, despite the fact that they grow very large relative to ν even in L^2 (in general, the growth in higher Sobolev norms is much larger; see [BGM15a]). In the end however, they do not get quite large enough to trigger a secondary instability. To see solutions where this kind of behavior might occur we need to look for solutions which are larger than $O(\nu)$. This is done in [BGM15b] wherein we prove the following theorem. We cannot quite follow our solutions all the way through a secondary instability, however, we at least follow the solutions until they enter a fully nonlinear regime. The $\nu^{2/3}$ threshold is discussed after the theorem.

Theorem 8.2 (Above threshold dynamics)**.** *For all $s \in (1/2, 1)$, all $\lambda_0 > \lambda' > 0$, all integers $\alpha \geq 10$ and all $\delta > 0$, there exists a constant $c_{00} = c_{00}(s, \lambda_0, \lambda', \alpha, \delta)$, a constant $K_0 = K_0(s, \lambda_0, \lambda')$, and a constant $\nu_0 = \nu_0(s, \lambda_0, \lambda', \alpha, \delta)$ such that for all $\delta_1 > 0$ sufficiently small relative to δ, all $\nu \leq \nu_0$, $c_0 \leq c_{00}$, and $\epsilon < \nu^{2/3+\delta}$, if $u_{in} \in L^2$ is a divergence-free vector field that can be written $u_{in} = u_S + u_R$ (both also divergence-free) with*

$$\|u_S\|_{\mathcal{G}^{\lambda_0;s}} + e^{K_0 \nu^{-\frac{3s}{2(1-s)}}} \|u_R\|_{H^3} \leq \epsilon, \tag{8.9}$$

then the unique, classical solution to (8.1) *with initial data u_{in} exists at least until time $T_F = c_0 \epsilon^{-1}$ and the following estimates hold with all implicit constants independent of ν, ϵ, c_0 and t:*

(i) Transient growth of the streak for $t < T_F$:

$$\left\| \overline{u_1}(t) - e^{\nu t \Delta} \left(\overline{u_1}(0) - t \overline{u_2}(0) \right) \right\|_{\mathcal{G}^{\lambda';s}} \lesssim c_0^2 \tag{8.10}$$

$$\left\| \overline{u_2}(t) - e^{\nu t \Delta} \overline{u_2}(0) \right\|_{\mathcal{G}^{\lambda';s}} + \left\| \overline{u_3}(t) - e^{\nu t \Delta} \overline{u_3}(0) \right\|_{\mathcal{G}^{\lambda';s}} \lesssim c_0 \epsilon; \tag{8.11}$$

(ii) uniform control of the background streak for $t < T_F$:

$$\|\overline{u_1}(t)\|_{\mathcal{G}^{\lambda';s}} \lesssim \epsilon \langle t \rangle \tag{8.12a}$$

$$\|\overline{u_2}(t)\|_{\mathcal{G}^{\lambda';s}} + \|\overline{u_3}(t)\|_{\mathcal{G}^{\lambda';s}} \lesssim \epsilon; \tag{8.12b}$$

(iii) the rapid convergence to a streak by the mixing-enhanced dissipation and inviscid damping of x-dependent modes:

$$\|u_{1,\neq}(t, x + ty + t\psi(t, y, z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon t^{\delta_1}}{\langle \nu t^3 \rangle^{\alpha}} \tag{8.13a}$$

$$\|u_{3,\neq}(t, x + ty + t\psi(t, y, z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle \nu t^3 \rangle^{\alpha}} \tag{8.13b}$$

$$\|u_{2,\neq}(t, x + ty + t\psi(t, y, z), y, z)\|_{\mathcal{G}^{\lambda';s}} \lesssim \frac{\epsilon}{\langle t \rangle \langle \nu t^3 \rangle^{\alpha}}, \tag{8.13c}$$

where $\psi(t, y, z)$ is an $O(\epsilon t)$ correction to the mixing.

The fact that we prove results for initial data as large as $\nu^{2/3+\delta}$ shows that the streak growth scenario for transition is generic even for initial data which is far larger than the $O(\nu)$ threshold, at least for data which is sufficiently regular. We are not aware of the $2/3$ exponent appearing anywhere in the applied mathematics or physics literature previously despite being a threshold of natural interest. This is likely because it is very hard to estimate convincingly. In [BGM15b], the 3D toy model which is analogous to (6.8) is used to estimate this exponent, and to design the norm necessary to prove it (note that these heuristics would not be very convincing without Theorem 8.2 and its proof to back it up). Basically, the toy model suggests that the natural time-scale before a general x-dependent solution could potentially become fully nonlinear, τ_{NL}, is *at least* $\tau_{NL} \gtrsim \epsilon^{-1/2}$. On the other hand, the enhanced dissipation occurs on time-scales like $\tau_{ED} \sim \nu^{-1/3}$. Hence, if the enhanced dissipation is to dominate the 3D effects and relax the solution to a streak, then we need the latter time scale to be shorter than the former:

$$\tau_{ED} \sim \nu^{-1/3} \ll \epsilon^{-1/2} \lesssim \tau_{NL}. \tag{8.14}$$

This is the origin of the requirement $\epsilon \lesssim \nu^{2/3+\delta}$ (the $\delta > 0$ is to provide a little technical room to work with in the estimates). After $t \gg \tau_{ED}$ the solution is very close to a streak and, due to the lift-up effect, in general $\overline{u_1}(t)$ is growing like $\epsilon\langle t\rangle$ until times $t \sim \epsilon^{-1}$, at which point the streak will become fully nonlinear (see [BGM15a, BGM15b, RSBH98, Cha02] and the references therein).

8.4 Brief summary of the proof

The proofs of Theorems 8.1 and 8.2 are quite intricate and depend on a lot of subtle cancellation structures hidden in the equations. However, we can briefly outline the main ideas here.

8.4.1 New dependent variables

Although quite unusual relative to many linear and formal weakly nonlinear studies (see e.g. [Cha02, SH01]) we found it natural to define the full set of auxillary unknowns $q^i = \Delta u_i$ for $i = 1, 2, 3$. A computation shows that (q^i) solves

$$\begin{cases} \partial_t q^1 + y\partial_x q^1 + 2\partial_{xy} u_1 + u \cdot \nabla q^1 = -q^2 + 2\partial_{xx} u_2 - q^j \partial_j u_1 + \partial_x (\partial_i u_j \partial_j u_i) - 2\partial_i u_j \partial_{ij} u_1 + \nu\Delta q^1 \\ \partial_t q^2 + y\partial_x q^2 + u \cdot \nabla q^2 = -q^j \partial_j u_2 + \partial_y (\partial_i u_j \partial_j u_i) - 2\partial_i u_j \partial_{ij} u_2 + \nu\Delta q^2 \\ \partial_t q^3 + y\partial_x q^3 + 2\partial_{xy} u_3 + u \cdot \nabla q^3 = 2\partial_{zx} u_2 - q^j \partial_j u_3 + \partial_z (\partial_i u_j \partial_j u_i) - 2\partial_i u_j \partial_{ij} u_3 + \nu\Delta q^3. \end{cases} \tag{8.15}$$

The linear terms have disappeared from the q^2 equation, leaving only the nonlinear terms on the RHS. Note that the equations on q^1 and q^3 are far less favorable in that they contain linear terms which are associated with the vortex stretching. From the linear level, we expect $q^{1,3}$ to grow quadratically (as otherwise, $u_{1,3}$ would experience inviscid damping!).

8.4.2 New independent variables

As in the 2D works, we need to find a good coordinate system to work in to correct for the mixing due to the Couette flow as well as $\overline{u_1}$, which will be very large relative to ν. Moreover, we have a much less clear view of what the change should look like, however, as in 2D where we want the derivatives to transform well, we can start with the following ansatz:

$$\begin{cases} X = x - ty - t\psi(t, y, z) \\ Y = y + \psi(t, y, z) \\ Z = z, \end{cases}$$

Consider the simple convection diffusion equation on a passive scalar $f(t, x, y, z)$

$$\partial_t f + y\partial_x f + u \cdot \nabla f = \nu\Delta f.$$

Denoting $F(t,X,Y,Z) = f(t,x,y,z)$ and $U(t,X,Y,Z) = u(t,x,y,z)$, and Δ_t and ∇^t for the expressions for Δ and ∇ in the new coordinates, this simple equation becomes

$$\partial_t F + \begin{pmatrix} u_1 - t(1+\partial_y\psi)u_2 - t\partial_z\psi u_3 - \frac{d}{dt}(t\psi) + \nu t\Delta\psi \\ (1+\partial_y\psi)u_2 + \partial_z\psi u_3 + \partial_t\psi - \nu\Delta\psi \\ u_3 \end{pmatrix} \cdot \nabla_{X,Y,Z}F = \nu\tilde{\Delta}_t F, \tag{8.16}$$

where $\tilde{\Delta}_t$ is a variant of Δ_t without lower order terms:

$$\tilde{\Delta}_t F = \partial_{XX}F + (1+\partial_y\psi)^2(\partial_Y - t\partial_X)^2 F + 2\partial_z\psi(\partial_Y - t\partial_X)\partial_Z F.$$

The shear flow in the first component of the velocity field is the largest and slowest decaying contribution. Hence, just like in the 2D works, it makes sense to choose the coordinates to cancel it:

$$\frac{d}{dt}(t\psi) + \overline{u_2}t\partial_y\psi + \overline{u_3}t\partial_z\psi = \overline{u_1} - t\overline{u_2} + \nu t\Delta\psi.$$

Unlike perhaps in [BM13], this change would be rather difficult to guess just based on intuition. As in 2D, we want to express this PDE in terms of entirely (Y,Z) variables rather than (y,z). The way this is done in 3D is analogous to, but not quite the same, as it is done in 2D. Write $U^i(t,X,Y,Z) = u_i(t,x,y,z)$ (as usual, consider X,Y,Z as functions of (x,y,z) or vice versa). We recast this equation on ψ in terms of $C(t,Y,Z) = \psi(t,y,z)$ and the auxillary unknown $g = \frac{1}{t}(U_0^1 - C)$ (as in 2D, this roughly measures the time-oscillations of C). A variety of cancellations which take advantage of the precise structures give

$$\begin{cases} \partial_t C + \tilde{U}_0 \cdot \nabla_{Y,Z}C = g - U_0^2 + \nu\tilde{\Delta}_t C, \\ \partial_t g + \tilde{U}_0 \cdot \nabla_{Y,Z}g = -\frac{2}{t}g - \frac{1}{t}\overline{\left(U_{\neq} \cdot \nabla^t U^1_{\neq}\right)} + \nu\tilde{\Delta}_t g, \end{cases} \tag{8.17}$$

where $\tilde{U} = \begin{pmatrix} U^1_{\neq} - t(1+\psi_y)U^2_{\neq} - t\psi_z U^3_{\neq} \\ (1+\psi_y)U^2_{\neq} + \psi_z U^3_{\neq} + g \\ U^3 \end{pmatrix}$. Unlike in 2D, we do estimates on the PDE for C and then use suitable nonlinear combinations of derivatives of C to recover the coefficients $\partial_i\psi$ that appear in the PDE (see [BGM15a] or re-derive on your own for more information). This procedure turns out to be better suited in 3D; note that in 3D the regularity gap between q^i and ψ is different than between ω and v in 2D (a small but important distinction when dealing with the coordinate transforms).

Coming back to (8.15), we further derive in the new coordinates ($Q(t,X,Y,Z) = q(t,x,y,z)$).

$$\begin{cases} Q^1_t + \tilde{U}\cdot\nabla_{X,Y,Z}Q^1 = -Q^2 - 2\partial^t_{XY}U^1 + 2\partial_{XX}U^2 - Q^j\partial^t_j U^1 - 2\partial^t_i U^j\partial^t_{ij}U^1 + \partial_X(\partial^t_i U^j\partial^t_j U^i) + \nu\tilde{\Delta}_t Q^1 \\ Q^2_t + \tilde{U}\cdot\nabla_{X,Y,Z}Q^2 = -Q^j\partial^t_j U^2 - 2\partial^t_i U^j\partial^t_{ij}U^2 + \partial^t_Y(\partial^t_i U^j\partial^t_j U^i) + \nu\tilde{\Delta}_t Q^2 \\ Q^3_t + \tilde{U}\cdot\nabla_{X,Y,Z}Q^3 = -2\partial^t_{XY}U^3 + 2\partial^t_{XZ}U^2 - Q^j\partial^t_j U^3 - 2\partial^t_i U^j\partial^t_{ij}U^3 + \partial^t_Z(\partial^t_i U^j\partial^t_j U^i) + \nu\tilde{\Delta}_t Q^3, \end{cases} \tag{8.18}$$

We perform most of our estimates on the coupled systems (8.18) and (8.17), recovering U^i from Q^i.

8.4.3 Choice of the norms

The proofs of Theorems 8.1 and 8.2 rely on a bootstrap argument as the 2D works did. The norms are quite complicated; we will only discuss those used in the proof of Theorem 8.1 as the norms used in Theorem 8.2 are even more complicated and will be too technical to motivate or explain (although both are primarily motivated by the same 3D toy model). Each Q^i is measured with a slightly different norm, of the form $\left\|A^i(t,\nabla)Q^i(t)\right\|_2$ where $A^i(t,\nabla)$ are special Fourier multipliers. Let us just describe the norm used to measure Q^3, the rest are similar (but not quite the same):

$$A^3_k(t,\eta,\ell) = e^{\lambda(t)|k,\eta,l|^s}\langle k,\eta,l\rangle^\sigma \frac{e^{\mu|\eta|^{1/2}}}{w(t,\eta)w_L(t,k,\eta,l)}\left(\mathbf{1}_{k\neq 0}\min\left(1,\frac{\langle \eta, l\rangle^2}{t^2}\right) + \mathbf{1}_{k=0}\right).$$

We now comment on the different components: $e^{\lambda(t)|k,\eta,\ell|^s}$ corresponds to a Gevrey-$\frac{1}{s}$ norm, with decreasing radius, while $\langle k,\eta,\sigma\rangle^\sigma$ gives a Sobolev correction (mainly for technical convenience). The next factors correspond to important physical effects. The factor w comes from the 3D toy model, described in [BGM15a]. Notice that, unlike in 2D, *it is independent of* k. Hence, unlike the 2D works, there is no strange unbalancing between critical and non-critical modes (there is, unfortunately, in the proof of Theorem 8.2 in [BGM15b]). The factor w is otherwise similar to the w in the 2D works (despite the fact that the 3D toy model is different). Roughly speaking, it is taken to satisfy the following for $|k|^2 \lesssim |\eta|$ (hence $\sqrt{|\eta|} \lesssim t \lesssim |\eta|$),

$$\frac{\partial_t w(t,\eta)}{w(t,\eta)} \sim \frac{1}{1+|t-\frac{\eta}{k}|}, \qquad \text{when } \left|t-\tfrac{\eta}{k}\right| \lesssim \tfrac{\eta}{k^2} \quad \text{and} \quad w(1,\eta)=1.$$

As in 2D, w is increasing and from Stirling's formula this this gives a growth like $\frac{w(2|\eta|,\eta)}{w(1,\eta)} \approx e^{\frac{\mu}{2}|\eta|^{1/2}}$ (up to a small polynomial correction), hence the choice of the numerator in $\frac{e^{\mu|\eta|^{1/2}}}{w(t,\eta)}$, and the Gevrey-2 regularity requirement. The multiplier w_L is a uniformly bounded multiplier that corrects for the anisotropy of the bounded growth experienced due to linear pressure effects (the L stands for 'linear'); this multiplier is not quite the same as, but very similar to, the multiplier we used in the proof of Theorem 5.2 above regarding the linearized 2D problem.

The last factor in the norm corresponds to a growth occuring for times large compared to the frequency due to the linear vortex stretching. That Q^1 and Q^3 ultimately grow at least quadratically is evident on the linear level: no inviscid damping occurs in general on $u^{1,3}$. That the growth can (and should) be taken only for frequencies small relative to time is far from clear and is predicted by the 3D toy model; see [BGM15a] for more discussion on this.

While the norm which was just sketched corresponds to the highest regularity estimate, estimates at lower regularity are also needed, in particular to quantify the enhanced dissipation. For this, we use an approach similar to that employed for 2D NSE in [BMV14], and define a set of semi-norms of the type $\left\|A^{\nu;i}(t,\nabla)Q^i\right\|_2$ for specially designed Fourier multipliers $A^{\nu;i}(t,\nabla)$. For example, for Q^3:

$$A^{\nu;3} = e^{\lambda(t)|k,\eta,\ell|^s}\langle k,\eta,\sigma\rangle^\beta \langle D(t,\eta)\rangle^\alpha \frac{1}{w_L(t,k,\eta,l)}\min\left(1,\frac{\langle\eta,l\rangle^2}{t^2}\right)\mathbf{1}_{k\neq 0},$$

where $D(t,\eta) \gtrsim \nu t^3$, is the same as in 2D.

8.4.4 Basic weakly nonlinear heuristics

There are several nonlinear mechanisms which have the potential to cause instabilities other than the lift-up effect and many have been proposed as important in the applied mathematics and physics literature for understanding transition, see e.g. [Cra71, TTRD93, RSBH98, SH01] and the references therein. We are particularly worried mechanisms similar to the 2D echo efffect [TTRD93, TE05, Wal95, BM13]: nonlinear interactions that repeatedly excite growing linear modes.

We will classify the main effects by the x frequency of the interacting functions: denote for instance $0 \cdot \neq \to \neq$ for the interaction of a zero mode (in x) and a non-zero mode (in x) giving a non-zero mode (in x), and similarly, with obvious notations, $0 \cdot 0 \to 0$, $\neq \cdot \neq \to \neq$, and $\neq \cdot \neq \to 0$.

(2.5NS) ($0 \cdot 0 \to 0$) For *2.5D Navier-Stokes*, these terms arise from the 2D nonlinear behavior of the streaks. Since we have global regularity for 2D Navier-Stokes, we are generally not so worried about these terms (although subtleties arise in [BGM15b]).

(SI) ($0 \cdot \neq \to \neq$) For *secondary instability*, this effect is the transfer of momentum from the large zero-modes in x to non-zero modes in x. These involve a zero frequency and non-zero frequency k interacting to amplify the same mode k, or the k mode of a different component, e.g. u_0^1 and u_k^3 together force u_k^2. These interactions are those that arise when linearizing an x-dependent perturbation of a streak and so are what ultimately give rise to the secondary instabilities observed in larger streaks (hence the terminology) [RSBH98, Cha02].

(3DE) ($\neq \cdot \neq \to \neq$) For *three dimensional echoes*, these effects are 3D variants of the 2D hydrodynamic echo phenomenon: nonlinear interactions of x-dependent modes forcing unmixing modes [Mor98, Van02, BM13]. In 3D, the range of possible interactions is much wider (see e.g. [Cra71, SH01, BGM15a, BGM15b]). This involves two non-zero frequencies k_1, k_2 interacting to force mode $k_1 + k_2$ with $k_1, k_2, k_1 + k_2 \neq 0$. Since these involve the interaction of only non-zero frequencies, they should only be problematic for short times: for $t \gtrsim \nu^{-1/3}$, this effect should be wiped out by the enhanced dissipation.

(F) ($\neq \cdot \neq \to 0$) For *nonlinear forcing*, this is the effect of the forcing from x-dependent modes back into x-independent modes. This involves two non-zero frequencies k and $-k$ interacting to force a zero frequency (in general this could involve a variety of the components). Similar to **(3DE)**, this effect is over-powered by the enhanced dissipation after $t \gtrsim \nu^{-1/3}$.

All of these are coupled to one another, and one can imagine bootstrap mechanisms involving several of them. It is the need to consider exactly these kinds of nonlinear bootstraps, the ones which involve all effects simultaneously, that eventually defines the toy model and precipitates the Gevrey-2 regularity requirement. The toy model also helps to determine the design of the norms we are using. Unfortunately, the toy model in 3D is significantly more complicated than in 2D (it is a 6x6 ODE, though not fully coupled), and we do not have time to derive and discuss it here. In a general sense, the derivation is analogous to how we derived (6.8) in §6.2, however, identifying which of the plethora of nonlinear terms are really respresentative of the “worst” and which can really be ignored is a lot trickier and requires a bit of intuition and patience. Moreover, the resulting system cannot really be solved by hand in any approximate sense at all, and figuring out approximate super solutions also requires a bit of intuition; in [BM13], we orginally approximately solved the 2D toy model (6.8) first and then found the simplified super-solutions. See [BGM15a] for information on this and all of the many details regarding the proof...

References

[Ali01] S. Alinhac. The null condition for quasilinear wave equations in two space dimensions I. *Invent. Math.*, 145(3):597–618, 2001.

[Arn65] VIi Arnold. Conditions for non-linear stability of plane steady curvilinear flows of an ideal fluid. *Doklady Akademii Nauk SSSR*, 162(5):773–777, 1965.

[Arn66] Vladimir Arnold. Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In *Annales de l'institut Fourier*, volume 16, pages 319–361, 1966.

[BCZGH15] J. Bedrossian, M. Coti Zelati, and N. Glatt-Holtz. Invariant measures for passive scalars in the small noise inviscid limit. *arXiv:1505.07356*, 2015.

[BDDM98] S. Bottin, O. Dauchot, F. Daviaud, and P. Manneville. Experimental evidence of streamwise vortices as finite amplitude solutions in transitional plane Couette flow. *Physics of Fluids*, 10:2597, 1998.

[BGM15a] J. Bedrossian, P. Germain, and N. Masmoudi. Dynamics near the subcritical transition of the 3D Couette flow I: Below threshold. *arXiv preprint*, 2015.

[BGM15b] J. Bedrossian, P. Germain, and N. Masmoudi. Dynamics near the subcritical transition of the 3D Couette flow II: Above threshold. *arXiv preprint*, 2015.

[BGM15c] J. Bedrossian, P. Germain, and N. Masmoudi. Working title: On the subcritical transition threshold for the 3D Couette flow in Sobolev regularity. *In preparation*, 2015.

[BM13] 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*, pages 1–106, 2013.

[BMM13] J. Bedrossian, N. Masmoudi, and C. Mouhot. Landau damping: paraproducts and Gevrey regularity. *arXiv:1311.2870*, 2013.

[BMV14] J. Bedrossian, N. Masmoudi, and V. Vicol. Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the 2D Couette flow. *To appear in Arch. Rat. Mech. Anal.*, 2014.

[Bon81] J.-M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup. (4)*, 14(2):209–246, 1981.

[Boy83] J.P. Boyd. The continuous spectrum of linear Couette flow with the beta effect. *Journal of the atmospheric sciences*, 40(9):2304–2308, 1983.

[BW13] M Beck and C E Wayne. Metastability and rapid convergence to quasi-stationary bar states for the two-dimensional Navier–Stokes equations. *Proc. Royal Soc. of Edinburgh: Sec. A Mathematics*, 143(05):905–927, 2013.

[Cha02] S.J. Chapman. Subcritical transition in channel flows. *Journal of Fluid Mechanics*, 451:35–98, 2002.

[Chr86] D. Christodoulou. Global solutions of nonlinear hyperbolic equations for small initial data. *Comm. Pure Appl. Math.*, 39(2):267–282, 1986.

[CKRZ] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168:643–674.

[Cra71] Alex DD Craik. Non-linear resonant instability in boundary layers. *Journal of Fluid Mechanics*, 50(02):393–413, 1971.

[Den09] Sergey A. Denisov. Infinite superlinear growth of the gradient for the two-dimensional Euler equation. *Discrete Contin. Dyn. Syst.*, 23(3):755–764, 2009.

[Den13] Wen Deng. Pseudospectrum for Oseen vortices operators. *Inter. Math. Res. Notices*, 2013(9):1935–1999, 2013.

[DR81] P.G. Drazin and W.H. Reid. *Hydrodynamic stability*. Cambridge U. Press, Cambridge, 1981.

[Dra02] PG Drazin. Introduction to hydrodynamic instability. *Cambridge Texts in Applied Mathematics*, 2002.

[EN00] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*. Graduate Texts in Mathematics. Springer-Verlag, 2000.

[EP75] T Ellingsen and E Palm. Stability of linear flow. *Physics of Fluids*, 18:487, 1975.

[Fjo50] Ragnar Fjortoft. Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex. *Geofys. Publ*, 17(6):1–52, 1950.

[FR14] E. Faou and F. Rousset. Landau damping in Sobolev spaces for the Vlasov-HMF model. *arXiv:1403.1668*, 2014.

[Gev18] M. Gevrey. Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire. *Ann. Sci. École Norm. Sup. (3)*, 35:129–190, 1918.

[GGN09] Isabelle Gallagher, Thierry Gallay, and Francis Nier. Spectral asymptotics for large skew-symmetric perturbations of the harmonic oscillator. *International Mathematics Research Notices*, page rnp013, 2009.

[HMRW85] Darryl D Holm, Jerrold E Marsden, Tudor Ratiu, and Alan Weinstein. Nonlinear stability of fluid and plasma equilibria. *Physics reports*, 123(1):1–116, 1985.

[Kel87] Lord Kelvin. Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates. *Phil. Mag.*, (24):188, 1887.

[Kla86] S. Klainerman. The null condition and global existence to nonlinear wave equations. *Lect. Appl. Math*, 23:293–326, 1986.

[KŠ14] A. Kiselev and V. Šverák. Small scale creation for solutions of the incompressible two-dimensional Euler equation. *Annals of Mathematics*, 180(3):1205–1220, 2014.

[Lan46] L. Landau. On the vibration of the electronic plasma. *J. Phys. USSR*, 10(25), 1946.

[Lan80] MT Landahl. A note on an algebraic instability of inviscid parallel shear flows. *Journal of Fluid Mechanics*, 98(02):243–251, 1980.

[Lin88] R. Lindzen. Instability of plane parallel shear flow (toward a mechanistic picture of how it works). *PAGEOPH*, 126(1), 1988.

[Mor98] P. J. Morrison. Hamiltonian description of the ideal fluid. *Rev. Modern Phys.*, 70(2):467–521, 1998.

[MV11] Clément Mouhot and Cédric Villani. On Landau damping. *Acta Math.*, 207(1):29–201, 2011.

[MWGO68] J. Malmberg, C. Wharton, C. Gould, and T. O'Neil. Plasma wave echo. *Phys. Rev. Lett.*, 20(3):95–97, 1968.

[Orr07] W. Orr. The stability or instability of steady motions of a perfect liquid and of a viscous liquid, Part I: a perfect liquid. *Proc. Royal Irish Acad. Sec. A: Math. Phys. Sci.*, 27:9–68, 1907.

[Ray80] Lord Rayleigh. On the Stability, or Instability, of certain Fluid Motions. *Proc. London Math. Soc.*, S1-11(1):57, 1880.

[Rey83] O Reynolds. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Proc. R. Soc. Lond.*, (35):84, 1883.

[RS79] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I.* Academic Press, 1979.

[RSBH98] S.C. Reddy, P.J. Schmid, J.S. Baggett, and D.S. Henningson. On stability of streamwise streaks and transition thresholds in plane channel flows. *J. of Fluid Mechanics*, 365:269–303, 1998.

[SH01] Peter J. Schmid and Dan S. Henningson. *Stability and transition in shear flows*, volume 142 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2001.

[Shn12] A. Shnirelman. On the analyticity of particle trajectories in the ideal incompressible fluid. *arXiv preprint arXiv:1205.5837*, 2012.

[Squ33] HB Squire. On the stability for three-dimensional disturbances of viscous fluid flow between parallel walls. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, pages 621–628, 1933.

[Ste93] E. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

[Tao06] T. Tao. *Nonlinear dispersive equations*, volume 106. American Mathematical Society, Providence, RI, 2006.

[TE05] Lloyd Nicholas Trefethen and Mark Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, 2005.

[TTRD93] L.N. Trefethen, A.E. Trefethen, S.C. Reddy, and T.A. Driscoll. Hydrodynamic stability without eigenvalues. *Science*, 261(5121):578–584, 1993.

[Van02] J. Vanneste. Nonlinear dynamics of anisotropic disturbances in plane Couette flow. *SIAM J. Appl. Math.*, 62(3):924–944 (electronic), 2002.

[VMW98] J. Vanneste, P.J. Morrison, and T. Warn. Strong echo effect and nonlinear transient growth in shear flows. *Physics of Fluids*, 10:1398, 1998.

[Wal95] Fabian Waleffe. Transition in shear flows. nonlinear normality versus non-normal linearity. *Physics of Fluids (1994-present)*, 7(12):3060–3066, 1995.

[Yag12] A.M. Yaglom. *Hydrodynamic Instability and Transition to Turbulence*, volume 100. Springer, 2012.

[YD02] J.H. Yu and C.F. Driscoll. Diocotron wave echoes in a pure electron plasma. *IEEE Trans. Plasma Sci.*, 30(1), 2002.

[YDO05] J.H. Yu, C.F. Driscoll, and T.M. O'Neil. Phase mixing and echoes in a pure electron plasma. *Phys. of Plasmas*, 12(055701), 2005.

[Zil14a] C. Zillinger. Linear inviscid damping for monotone shear flows. *arXiv:1410.7341*, 2014.

[Zil14b] C. Zillinger. Linear inviscid damping for monotone shear flows in a finite periodic channel, boundary effects, blow-up and critical Sobolev regularity. *arXiv:1506.04010*, 2014.