# Shear Turbulence: Onset and Structure

Fabian Waleffe
notes by Martín Hoecker-Martínez and Chao Ma
revised and *massively* expanded by FW
WHOI GFD Lecture 1

June 20, 2011

## 1 Introduction

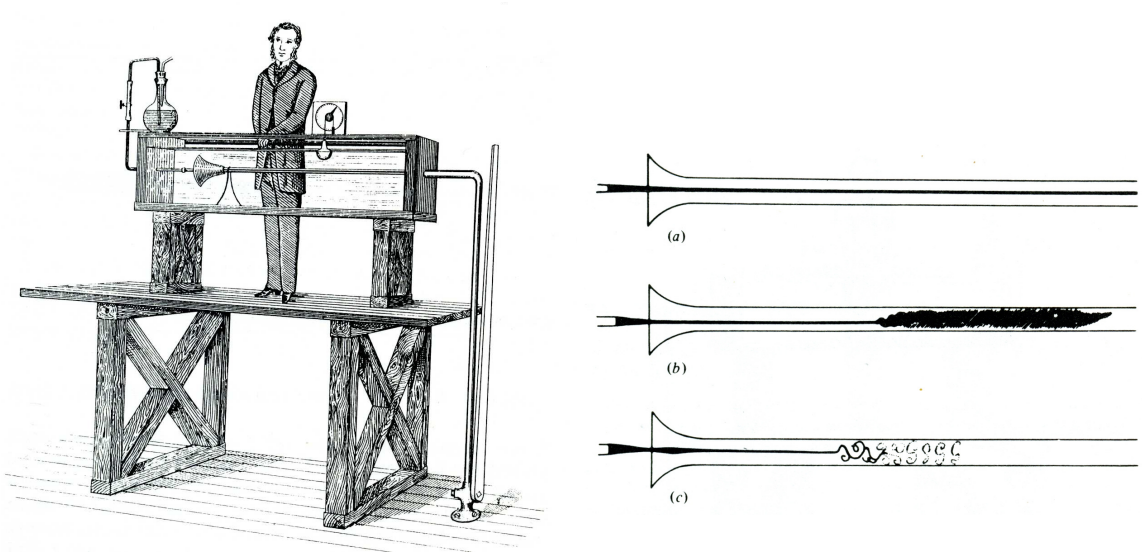


Figure 1: Reynolds' 1883 experiment, still operating at the University of Manchester, UK. Water flows down a long pipe and ink is introduced at the pipe centerline. On the right are the original drawings showing the dye filament in laminar flow (a), and the rapid mixing of the dye and water when the Reynolds number is larger than about 2000 in (b). (c) shows details of vortical structures in turbulent flow, visualized with a spark.

A classic series of educational fluid mechanics films is available online at

`http://web.mit.edu/hml/ncfmf.html`

The film on *Turbulence* provides an excellent visual introduction to turbulence in shear flows and reproduces the seminal experiments of Osborne Reynolds [24] on the transition from laminar to turbulent flow in a pipe, illustrated in Fig. 1. The 2011 *Annual Review*

*of Fluid Mechanics* article by Tom Mullin [21] provides many fascinating historical details and is also highly recommended reading.

Transition to turbulence results from the non-linearity of the governing Navier-Stokes equations and the relevant non-dimensional quantity is the Reynolds number

$$R \equiv \frac{UL}{\nu} \tag{1}$$

where $U$ and $L$ are characteristic velocity and length scales, respectively — such as the average (or 'bulk') flow speed and the diameter of the pipe — and $\nu$ is the kinematic viscosity of the fluid (*e.g.* $\nu \approx 10^{-5}$ m$^2$/s for air, and $\nu \approx 10^{-6}$ m$^2$/s for water). Pipe flow will typically be in a turbulent state when the Reynolds number is larger than about 2000 and laminar below that. For flow of water ($\nu \approx 10^{-6}$ m$^2$/s), in a pipe of diameter $D$=10cm = 0.1m, turbulent flow is observed for bulk velocities $U$ greater than about a mere 2 cm/s = 0.02 m/s.

### 1.1 Navier-Stokes equations

Although most fluids are compressible, the fundamental study of turbulence can be simplified if we consider an idealized incompressible flow. This is equivalent to assuming that the speed of sound in our medium, $c = \sqrt{(\partial P/\partial \rho)_S}$ ($\approx 340$ m/s in air and 1500 m/s in water), is much greater than any velocity in the fluid. The Navier-Stokes equations for constant density fluid flow (for a derivation see for example Acheson [2, Chap. 6] or Batchelor [4, Sect. 3.2]) are

$$\nabla \cdot \mathbf{v} = 0 \tag{2}$$

$$\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} + \nabla p = \nu \nabla^2 \mathbf{v} \tag{3}$$

where $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$ is the fluid velocity at position $\mathbf{r} \in \mathbb{R}^3$ at time $t$, $\nabla$ is the del operator, $p = p(\mathbf{r}, t)$ is the *kinematic* pressure (*i.e.* the pressure divided by the constant mass density $\rho$) and $\nu > 0$ is the *kinematic* viscosity and $\nabla^2 = \nabla \cdot \nabla$ is the Laplacian. The standard boundary condition for a viscous fluid is no-slip at the wall, that is $\mathbf{v}_{fluid} = \mathbf{v}_{wall}$.

For incompressible flow, the pressure $p$ is purely *mechanical*, not thermodynamic. The pressure is determined by the flow and the incompressibility constraint $\nabla \cdot \mathbf{v} = 0$, and there is no need for an equation of state. In fact, taking the divergence of (3) and using $\nabla \cdot \mathbf{v} = 0$, yields a Poisson equation for the pressure

$$\nabla^2 p = -\nabla \cdot (\mathbf{v} \cdot \nabla \mathbf{v}) \tag{4}$$

where $\nabla \cdot (\mathbf{v} \cdot \nabla \mathbf{v}) = \nabla \mathbf{v} : \nabla \mathbf{v} = (\partial_i v_j)(\partial_j v_i) = S_{ij} S_{ij} - \Omega_{ij} \Omega_{ij}$ in cartesian index notation, where $\partial_i = \partial/\partial x_i$ and $S_{ij} = (\partial_i v_j + \partial_j v_i)/2$, $\Omega_{ij} = (\partial_i v_j - \partial_j v_i)/2$ are the deformation and rotation rate tensors, respectively. Thus, the pressure gradient in (3) is in fact a *non-local, nonlinear term* for the Navier-Stokes equations. The nonlinearity of the Navier-Stokes equations is $\mathbf{v} \cdot \nabla \mathbf{v} + \nabla p$, with the pressure solving the elliptic Poisson equation (4).

However, there are difficulties, misunderstandings and controversies regarding the boundary conditions needed to solve the Poisson equation (4), as discussed in Rempfer [23]. The dream is that the pressure boundary conditions can be decoupled from the velocity – a Neumann boundary condition $\partial p/\partial n = 0$ where $n$ is normal to the wall, for instance – so

that one could update the velocity in time using (3) then update the pressure from (4) using the updated velocity. Such methods are usually referred to as *Pressure Poisson equation* (PPE) formulations. Such decoupling is not correct in general, and PPE formulations, as well as the related *fractional step methods*, may have particular difficulties with *steady*, *traveling wave* and *time-periodic* solutions of the Navier-Stokes equations. Such solutions are precisely the focus of these lectures. A general approach is to enforce the incompressibility constraint $\nabla \cdot \mathbf{v} = 0$ *ab initio* through a streamfunction-type (*i.e.* divergence-free) formulation for $\mathbf{v}$ and eliminate the pressure by taking the curl ($\nabla \times (\cdot)$) of the Navier-Stokes equations (3) or projecting those equations onto the space of divergence-free functions. The $v, \eta$ and 'poloidal-toroidal' representations fall under this general category of methods.

### 1.2 Laminar pipe flow

For flow in a nominally infinitely long pipe of radius $a$, driven by an imposed constant pressure gradient $dp_0/dx < 0$ in the axial streamwise direction $x$, GFD fellows easily verify that the Navier-Stokes equations (2), (3), have the parabolic solution

$$\mathbf{v} = \left(1 - \frac{r^2}{a^2}\right) U_c \, \hat{\mathbf{x}} \tag{5}$$

where $r = \sqrt{y^2 + z^2}$ is the distance to the pipe centerline, $U_c$ is the centerline velocity, $\hat{\mathbf{x}}$ is the unit vector in the pipe direction and

$$\frac{dp_0}{dx} = -\nu \frac{4U_c}{a^2}. \tag{6}$$

This is the *laminar* pipe flow solution, first found by Poiseuille [21] from his experimental data. Averaging (5) over the pipe cross section $A$, yields the laminar bulk velocity

$$U = \frac{1}{\pi a^2} \int_A \hat{\mathbf{x}} \cdot \mathbf{v} \, dA = \frac{2}{a^2} U_c \int_0^a r \left(1 - \frac{r^2}{a^2}\right) dr = \frac{U_c}{2} \tag{7}$$

which is simply half the centerline velocity for *laminar* pipe flow. This simple relationship is not true for *turbulent* pipe flow.

Defining the Reynolds number $R = UD/\nu$ in terms of the bulk velocity $U$ and the pipe diameter $D = 2a$, we can write a non-dimensional form of the relationship (6) between the pressure gradient and the laminar flow velocity, with $U_c = 2U$ from (7). This yields the the *friction factor*

$$f \equiv \left|\frac{dp_0}{dx}\right| \frac{D}{\frac{1}{2}U^2} = 64 \frac{\nu}{UD} = \frac{64}{R}. \tag{8}$$

Thus in laminar pipe flow, the friction factor $f$ is inversely proportional to the Reynolds number, $f = 64/R$. In experiments, this laminar flow regime is typically observed only for $Re \lesssim 2000$. For higher Reynolds number, the friction factor transitions to a much weaker dependence on Reynolds number, perhaps asymptoting to a non-zero constant as illustrated in the Moody diagram well-known to engineers (Fig. 2).

However, transition is more complex than suggested by the Moody diagram. Reynolds achieved laminar pipe flow up to $R \approx 13\,000$ and Pfenniger's world record is about $100\,000$

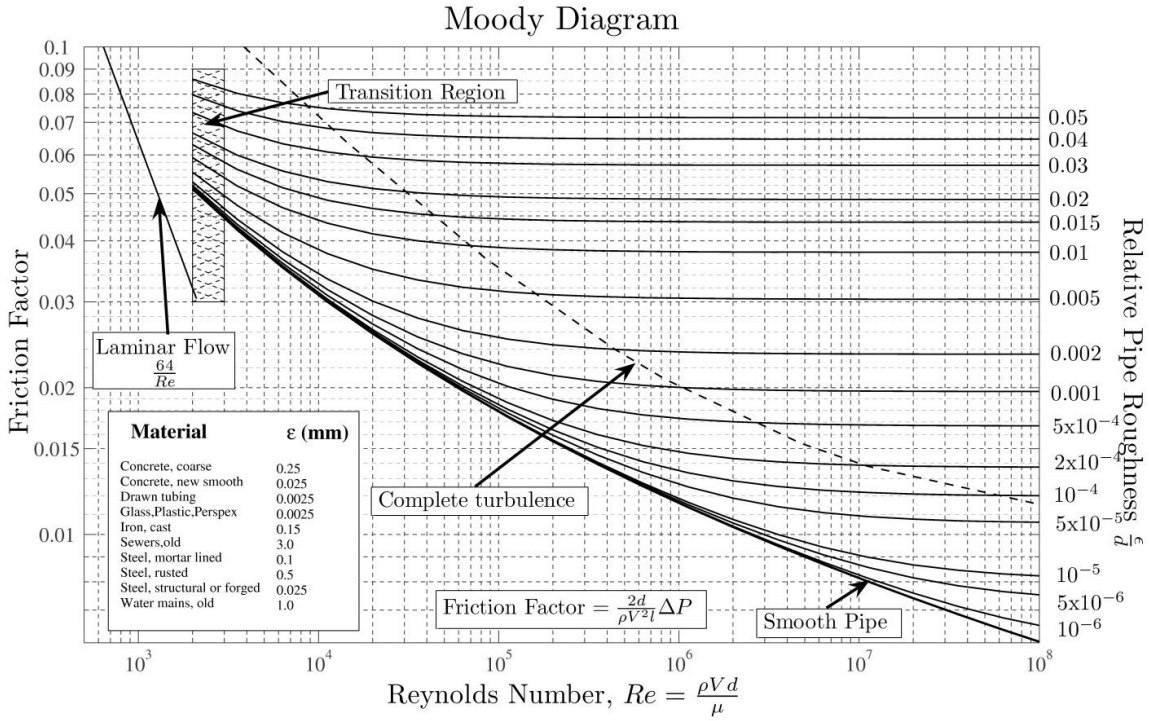


Figure 2: Moody diagram (from `http://en.wikipedia.org/wiki/Moody_chart`) showing the friction factor $f$ as a function of Reynolds number $R = UD/\nu$ (eqn. (8)). Each curve is associated with a surface roughness of the pipe. On this plot $V \equiv U$ is the bulk velocity, $d \equiv D$ is the pipe diameter, $\mu = \rho\nu$ is the dynamic viscosity, and $\Delta P/(\rho l) = |dP_0/dx|$. The laminar flow friction law $f = 64/R$ is experimentally observed only for $R \lesssim 2,000$, even for smooth pipes. For $R \gtrsim 2000$, the drag associated with turbulent flow is much larger than it would be for laminar flow at the same Reynolds number. 'Complete turbulence' is loosely defined as the region where the friction factor is independent of Reynolds number. At $R = 10^5$, $f \approx 0.018$ is about 30 times larger than the laminar value $64/R$.

[21]. Meseguer and Trefethen's [20] numerical calculations of the Navier-Stokes equations linearized about the laminar flow (5) show stability up to $R = 10^7$. Based on this experimental and numerical evidence, laminar pipe flow is believed to be *linearly stable for all Reynolds numbers.* However, no complete mathematical proof of linear stability is known to date.

### 1.3 Shear flows

Pipe flow is only one example of a *shear flow*, that is, a flow whose velocity varies in the direction perpendicular to the flow direction. Shear flows are a fundamental and ubiquitous class of fluid flow owing to the viscosity of real fluids and the no-slip boundary condition. Whenever a fluid flows by a wall, the no-slip boundary condition will lead to the generation of shear near the wall. This is the classic *Prandtl boundary layer* (Fig. 3) that can diffuse away from the wall temporally or spatially and even separate, shedding vortices (Fig. 4) (see *e.g.* [2], [4] for further information about this important and complex problem)

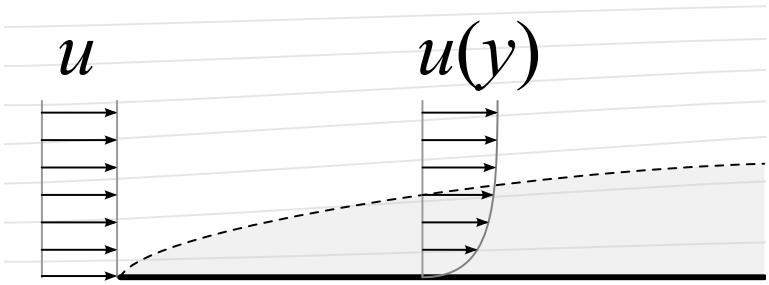


Figure 3: Shear flow developing spatially from viscosity and no-slip boundary condition as a fluid flows over a semi-infinite flat plate. The velocity is strongly dependent on the wall-normal $y$ direction in a narrow *boundary layer* near the wall but only weakly dependent on the streamwise direction $x$. In laminar flow, the boundary layer thickness $\delta$ scales as $\sqrt{\nu x/U_\infty}$. (From `http://en.wikipedia.org/wiki/Boundary_layer`)

This process of boundary layer development and possible vortex shedding occurs at the entrance to the pipe in Reynolds' experiment. The vortex shedding at the entrance has a strong effect on the onset of turbulence and it is to eliminate those perturbations that Reynolds used a funnel, as shown in Fig. 1. With careful control of the entrance flow and geometry, one obtains the fully developed, steady laminar pipe flow (5) that can be routinely observed up to $Re \approx 20\,000$.

Other canonical 'fully developed' shear flows are plane Poiseuille (a.k.a. channel) flow and plane Couette flow, sketched in Fig. 5. Plane Poiseuille flow is the flow driven by a pressure gradient in-between two infinitely long, parallel fixed planes and the laminar solution is $\mathbf{v} = (1 - y^2/h^2)\, U_c\, \hat{\mathbf{x}}$ where $y$ is the wall-normal direction and the walls are located at $y = \pm h$. Plane Couette flow is the flow between two infinite parallel walls and driven by the motion of those walls in opposite directions. The laminar plane Couette flow is $\mathbf{v} = U_w y/h\, \hat{\mathbf{x}}$ where the walls are at $y = \pm h$ and move at velocities $\pm U_w \hat{\mathbf{x}}$, respectively.

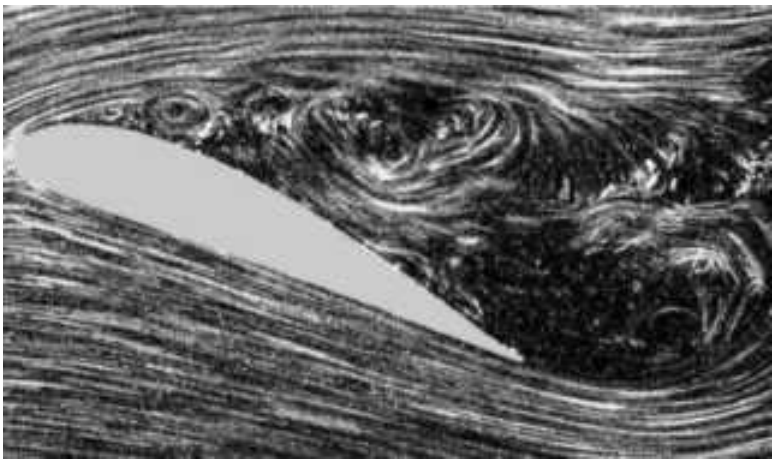

Figure 4: Boundary layer separation around an airfoil. This airfoil is *stalled*, the lift has collapsed, drag as increased, the flow is unsteady and turbulent behind the airfoil and vortices are continuously shed from the top front of the airfoil.
(from `http://en.wikipedia.org/wiki/Boundary_layer_separation`).

Plane Couette flow is linearly stable for *all* Reynolds numbers $R > 0$, as proved by Romanov [25], but experiments and simulations show transition to turbulence for $R = U_w h/\nu \gtrsim 350$. Contrary to its stable cylindrical cousin, plane Poiseuille flow has a linear instability for $R_c = U_c h/\nu > 5772$, where $U_c$ is the centerline velocity and $h$ the half-channel height. For plane Poiseuille flow, the bulk velocity $U = 2U_c/3$, so linear instability occurs for $R = UH/\nu > (4/3)\,5772 = 7696$, based on the bulk $U$ and the full channel height $H = 2h$. This linear instability is an intriguing instability that originates from viscosity *and* the no-slip boundary condition, as anticipated by Prandtl. This instability is governed by the Orr-Sommerfeld equation and was first revealed by the pioneering analyses of Heisenberg for channel flow in 1924 and Tollmien for the Blasius boundary layer flow in 1929 [9]. This instability is very weak and delicate, small changes in the flow or geometry can suppress it. In any case, transition to turbulence is observed for $R_c = U_c h/\nu \gtrsim 1500$ in channel flow, well below the linear instability threshold [19].

### 1.4 Transition threshold

If plane Couette and pipe flows are stable to infinitesimal perturbations but experiments show transition to turbulence, this transition must result from *finite amplitude effects*. This is also the case in plane Poiseuille and boundary layer flows, where transition is observed at significantly lower Reynolds number than predicted by linear analysis, and on much faster time scales than the slow viscous time scale of linear eigenmodes. In those cases, Morkovin coined the term *'bypass transition'* – transition that bypasses the linear instability process. Onset of turbulence is thus a Reynolds number *and* amplitude dependent phenomenon.

A natural question therefore is to ask whether there is a scaling relating the *threshold*

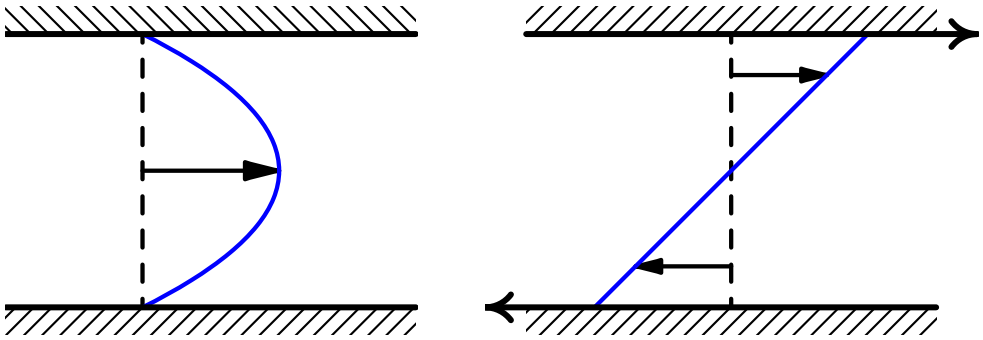

Figure 5: Canonical shear flows: pressure driven pipe and channel flows with a parabolic velocity profile $\mathbf{v} = (1-y^2)\,\hat{\mathbf{x}}$ (*left*) and wall driven plane Couette flow with laminar velocity profile $\mathbf{v} = y\,\hat{\mathbf{x}}$ (*right*), where $\hat{\mathbf{x}}$ is the flow direction, $y$ is the (signed) distance to the (midplane) centerline. The walls are at $y = \pm 1$, the centerline velocity is 1 for pipes and channels, and the wall velocities are $\pm 1$ for Couette.

transition amplitude, $\epsilon$ say, to the Reynolds number $R$, perhaps a power law scaling

$$\epsilon \sim R^a \tag{9}$$

where we expect $a < 0$. Surprisingly, this question does not seem to have been articulated before Trefethen *et al.* in 1993 [29, TTRD hereafter]. Closely related questions, and more specific nonlinear mechanisms, were being investigated before TTRD, but the focus was on inviscid processes, therefore those investigations focused on the relations between *time scale* and *amplitude*, rather than *Reynolds number* and *amplitude*. That the two questions are closely related can be understood as follows.

Non-dimensionalizing the velocity by a characteristic velocity $U$, length scales by a characteristic length $L$, time by $L/U$ and kinematic pressure by $U^2$, the Navier-Stokes equations (3) have the same form but with $\nu$ replaced by $1/R$

$$\partial_t \mathbf{v} - \frac{1}{R}\nabla^2 \mathbf{v} = -\mathbf{v}\cdot\nabla\mathbf{v} - \nabla p \tag{10}$$

where the right hand side is a quadratic nonlinearity since $\nabla p$ projects $\mathbf{v}\cdot\nabla\mathbf{v}$ onto the space of divergence-free fields (4). If $\epsilon$ is a measure of the amplitude of $\mathbf{v}$, we could estimate

$$\partial_t \mathbf{v} \sim \frac{\epsilon}{T}, \qquad \frac{1}{R}\nabla^2\mathbf{v} \sim \frac{\epsilon}{R}, \qquad \mathbf{v}\cdot\nabla\mathbf{v} + \nabla p \sim \epsilon^2 \tag{11}$$

where $T$ is a time scale to be determined. In the limit of large Reynolds number, $R \to \infty$, an inertial, inviscid scaling for (10) is then

$$\frac{\epsilon}{T} \sim \epsilon^2 \quad \Rightarrow \quad T \sim \frac{1}{\epsilon} \tag{12}$$

suggesting that the essential time scale for a non-trivial nonlinear process would be $T \sim 1/\epsilon$. This is much faster than the standard nonlinear time scale for nonlinear waves (*e.g.* Duffing

oscillator, weakly nonlinear pendulum, nonlinear Schrödinger, ...) where the nonlinear time scale is $T \sim 1/\epsilon^2 \gg 1/\epsilon$, for small $\epsilon$. The latter results from the fact that the quadratic interaction of a 'wave' or Fourier mode $\epsilon e^{ikx}$ of amplitude $\epsilon$ and wavenumber $k$, generates a harmonic $\epsilon^2 e^{i2kx}$ of amplitude $\epsilon^2$ and wavenumber $2k$. That harmonic then interacts with the complex conjugate of the fundamental wave $\epsilon e^{-ikx}$ to provide nonlinear feedback $\epsilon^3 e^{ikx}$ onto the fundamental wave, but only at order $\epsilon^3$, not $\epsilon^2$. The nonlinear wave balance in that case is not (12) but $\epsilon/T \sim \epsilon^3$ yielding the slow nonlinear time scale $T \sim 1/\epsilon^2$ (recall that $\epsilon$ is small and measures relative departure from an equilibrium).

Now if instead of $R \to \infty$ with large but finite $T$, we focus instead on *steady* as in steady state and traveling waves, or *statistically steady* as in periodic or turbulent solutions, that is $T \to \infty$ with $Re$ large but finite, the nonlinear balance and time scale argument (12) for (10) now yields

$$\frac{\epsilon}{R} \sim \epsilon^2 \quad \Rightarrow \quad \epsilon \sim \frac{1}{R} \tag{13}$$

and we interpret this as a measure of the minimum perturbation amplitude that could balance viscous damping – a *threshold amplitude*. Again this is a much smaller amplitude than the classical weakly nonlinear balance where feedback on the fundamental occurs at order $\epsilon^3$ and the balance is $\epsilon/R \sim \epsilon^3$ not $\epsilon^2$, yielding $\epsilon \sim 1/\sqrt{R} \gg 1/R$.

The simple balance (13) thus suggest that $a = -1$ in (9), stronger than the classic weakly nonlinear scaling that would have $a = -1/2$. However, while the scalings (12), (13) certainly apply to the *'subcritical logistic' equation*[1]

$$\frac{du}{dt} = -\frac{u}{R} + u^2, \tag{14}$$

that does have the threshold amplitude $u = 1/R$, it remains to be shown whether they do apply to the Navier-Stokes equations (10). Indeed, the brief discussion of classic weakly nonlinear feedback occurring at the weaker $\epsilon^3$ instead of $\epsilon^2$, suggests that the scalings (12) and (13) may not be allowed by the Navier-Stokes nonlinearity and it is necessary to investigate explicit nonlinear mechanisms for shear flows. The linear terms also require more investigation since they are trickier than the mere viscous damping in (14).

Decompose the full velocity $\mathbf{v}$ into a laminar flow $\mathbf{U}$ plus a perturbation $\mathbf{u}$. Substituting $\mathbf{v} = \mathbf{U} + \mathbf{u}$ into the Navier-Stokes equations (10), we obtain the equations for the solenoidal perturbation $\mathbf{u}$ in the schematic form

$$\frac{\partial \mathbf{u}}{\partial t} - \frac{1}{R}\nabla^2 \mathbf{u} = L(\mathbf{u}) + N(\mathbf{u}, \mathbf{u}) \tag{15}$$

where $N(\mathbf{u}, \mathbf{u})$ is a quadratic nonlinearity and $L(\mathbf{u}) = N(\mathbf{U}, \mathbf{u}) + N(\mathbf{u}, \mathbf{U})$ is linear in $\mathbf{u}$. Pressure is once again hidden in each of those operators to project them onto the space of solenoidal fields, so $N(\mathbf{u}, \mathbf{u}) = -\mathbf{u} \cdot \nabla \mathbf{u} - \nabla p$. The perturbation $\mathbf{u}$ is solenoidal $\nabla \cdot \mathbf{u} = 0$ and satisfies homogeneous boundary conditions, $\mathbf{u} = 0$ at no-slip walls since $\mathbf{U}$ satisfies $\mathbf{U}_{wall} = \mathbf{v}_{wall}$. Before looking into specific mechanisms and possible threshold exponent $a$ in (9), we first discuss some characteristics of the linear operator $L(\mathbf{u})$ for shear flows which was ignored in the 'fully nonlinear' scalings (11) and (13).

[1]This 'subcritical logistic' equation is sometimes used to suggest that the Navier-Stokes equations could develop a singularity for sufficiently large initial amplitudes.

# 2 Linear Theory: exponential and algebraic growth

## 2.1 Exponential growth

A linear stability analysis of the laminar flow $\mathbf{U}$ consists of studying the linearized equations

$$\frac{\partial \mathbf{u}'}{\partial t} - \frac{1}{R}\nabla^2 \mathbf{u}' = L(\mathbf{u}') \tag{16}$$

which is (15) for $\mathbf{u} = \epsilon \mathbf{u}'$ in the limit $\epsilon \to 0$, the 'infinitesimal amplitude' limit. If this linearized equation admits solutions that are *growing for all times*, then the laminar flow $\mathbf{U}$ is unstable. Growing perturbations typically have the form of *exponentially* growing eigenmodes of (16), solutions of the form $\mathbf{u}'(\mathbf{r}, t) = e^{\lambda t}\mathbf{u}(\mathbf{r})$ where $\lambda$ is the temporal eigenvalue, with $\Re(\lambda) > 0$ for instability, and $\mathbf{u}(\mathbf{r})$ is the *eigenmode*. For example, the simple 2-by-2 model

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -k_u^2/R & 1 \\ \sigma^2 & -k_v^2/R \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}, \tag{17}$$

where $k_u$, $k_v$ and $\sigma$ are real and positive, has two exponential eigenmodes, one of which is growing if $R > k_u k_v/\sigma$.

The linear analysis of exponentially growing modes for shear flows is exceedingly delicate and was a central problem in Applied Mathematics for many decades beginning with the reduction of the problem to the *Orr-Sommerfeld equation*, derived independently by Orr in 1907 and Sommerfeld in 1908, and followed by the intricate asymptotic analyses of that equation by Heisenberg (1924), Tollmien (1929, 1935, 1947), Schlichting (1933), C.C. Lin (1945, 1955, 1957, 1961) and the more rigorous and general analyses of Morawetz (1951) and Wasow (1953), among many others. Orszag (1971) provided accurate numerical solution of the Orr-Sommerfeld equation using Chebyshev polynomials, for moderate Reynolds number, and that work demonstrated the usefulness of *spectral methods* for fluid dynamics computations.

As mentioned in sect. 1.3, the result of those many decades of studies is that there is essentially *no* exponentially growing mode for viscous shear flows. It is proved that there are none in plane Couette flow [25]; there is strong evidence that there are none in pipe flow [20, 21]; there is a weak *viscous* growing mode in plane Poiseuille flow first identified by Heisenberg and a similar, slightly more significant, mode for Blasius boundary layer flow, first studied by Tollmien and Schlichting. Those weakly growing modes are called *Tollmien-Schlichting waves* [9], but Heisenberg already understood in 1924 that these weakly growing viscous linear modes did not explain the onset of turbulence in shear flows, ... and that may be partly why he quickly switched to developing the matrix formulation of quantum mechanics!

## 2.2 Algebraic growth, redistribution of base flow

Thus, as an analogy to shear flows, the model problem (17) should have $\sigma = 0$ and be stable for all $0 < R < \infty$. Degenerate *algebraic* instability can occur in (17) when $\sigma = 0$ and $R = \infty$,

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}, \tag{18}$$

the coupling matrix $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ is a *Jordan block*, it has the repeated eigenvalue $\lambda = 0$ but more importantly it has *only one* eigenvector, namely $(u, v) = (1, 0)$. Initial conditions $(u, v) = (0, 1)$ yield the algebraic pseudo-mode $(u, v) = (t, 1)$ since (18) is simply $dv/dt = 0$, $du/dt = v$. The general solution of (18) for initial conditions $(u, v) = (u_0, v_0)$ at $t = 0$ is

$$\begin{pmatrix} u \\ v \end{pmatrix} = u_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v_0 \begin{pmatrix} t \\ 1 \end{pmatrix}, \tag{19}$$

$v = v_0$ is constant but induces an algebraic growth $u = u_0 + v_0 t$.

Such algebraic growth is *generic* for the *inviscid linear* dynamics about shear flows. To show this, consider a general plane parallel shear flow $\mathbf{U} = U(y)\hat{\mathbf{x}}$, where $x$ is streamwise and $y$ shearwise. Substituting $\mathbf{v} = U(y)\hat{\mathbf{x}} + \mathbf{u}$ in (3) yields

$$\begin{aligned} \frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} + v\frac{dU}{dy} \quad &+ \mathbf{u}\cdot\nabla u = -\frac{\partial p}{\partial x} + \frac{1}{R}\nabla^2 u, \\ \frac{\partial v}{\partial t} + U\frac{\partial v}{\partial x} \qquad &+ \mathbf{u}\cdot\nabla v = -\frac{\partial p}{\partial y} + \frac{1}{R}\nabla^2 v, \\ \frac{\partial w}{\partial t} + U\frac{\partial w}{\partial x} \qquad &+ \mathbf{u}\cdot\nabla w = -\frac{\partial p}{\partial z} + \frac{1}{R}\nabla^2 w, \end{aligned} \tag{20}$$

where $U = U(y)$ is the laminar shear flow profile and $\mathbf{u} = (u, v, w)$ in cartesian coordinates $(x, y, z)$ with $\nabla\cdot\mathbf{u} = \partial u/\partial x + \partial v/\partial y + \partial w/\partial z = 0$. For streamwise independent perturbations, $\partial/\partial x = 0$, and linearizing in $\mathbf{u}$, (20) reduces to

$$\begin{aligned} \frac{\partial u}{\partial t} + v\frac{dU}{dy} \quad &= \qquad\quad \frac{1}{R}\nabla^2 u, \\ \frac{\partial v}{\partial t} \qquad &= -\frac{\partial p}{\partial y} \quad + \frac{1}{R}\nabla^2 v, \\ \frac{\partial w}{\partial t} \qquad &= -\frac{\partial p}{\partial z} \quad + \frac{1}{R}\nabla^2 w, \end{aligned} \tag{21}$$

and $\nabla\cdot\mathbf{u} = 0$ reduces to $\partial v/\partial y + \partial w/\partial z = 0$. Eliminating $p$ and $w$, we obtain

$$\begin{aligned} \frac{\partial u}{\partial t} + v\frac{dU}{dy} \quad &= \frac{1}{R}\nabla^2 u, \\ \frac{\partial}{\partial t}\nabla^2 v \qquad &= \frac{1}{R}\nabla^2\nabla^2 v, \end{aligned} \tag{22}$$

where $\nabla^2 = \partial^2/\partial y^2 + \partial^2/\partial z^2$ for streamwise independent perturbations. For planar geometries, the no-slip boundary conditions $\mathbf{u} = 0$ yields $u = v = \partial v/\partial y = 0$ at the walls, the latter following from $\partial v/\partial y + \partial w/\partial z = 0$ since $w = \partial w/\partial z = 0$ along the walls.

For inviscid flow, $R = \infty$, (22) is conceptually identical to the simple model (18) and a non-zero vertical velocity will remain constant, $v = v_0$, but induce an algebraic growth of the streamwise velocity perturbation $u$

$$\frac{\partial u}{\partial t} = -v_0\frac{dU}{dy} \quad \Rightarrow \quad u = u_0 - v_0\frac{dU}{dy}\,t. \tag{23}$$

This is the linearization of a simple *redistribution of streamwise velocity $U(y)$ by the shearwise velocity perturbation $v$*. If, for instance, $U' = dU/dy > 0$ and $v > 0$, then lower velocity

$U$ is lifted upward by the perturbation $v$ leading to a negative streamwise velocity perturbation $u = -v_0 U' t < 0$. This simple algebraic growth was perhaps first investigated by Benney and Lin in 1960 [7], rediscovered by Ellingsen and Palm in 1975 [10] and Gustavsson and Hultgren in 1980 and 1990, *e.g.* [12, 15, 11]. This redistribution was dubbed the *'lift-up mechanism'* by Marten Landahl, although it is actually *'pull-down'* when $v < 0$, and 'lift-up' usually comes with 'pull-down' because of conservation of volume. This is a trivial 'mechanism' yet an essential part of any instability or self-sustaining process that redistributes background momentum to release energy from the background shear.

But *linear* fluid dynamics is not *physical* fluid dynamics. Algebraic growth *for all times* cannot occur for bounded flows, even in the inviscid limit. Consider plane Couette flow $U(y) = y$ for $-1 \le y \le 1$. The largest perturbation that can be achieved is to 'lift-up' velocity $U = -1$ from the bottom wall all the way to the top wall and obtain $u = -2$ at $y = 1$, or likewise to 'pull-down' velocity $U = 1$ from the top wall all the way to the bottom wall, obtaining $u = 2$ at $y = -1$. Thus $|u| < 2$ in plane Couette and this is a strict upper bound since $v = 0$ at the walls, so it is actually not possible to lift-up or pull-down all the way from one wall to the other. In linear theory, *lift-up* can go on forever appearing mathematically as algebraic growth from the linearized advection term $v\,dU/dy$, but in real flows, lift-up does not go on forever. The term *redistribution* of streamwise velocity better captures the inherent limitation of the underlying simple advection of the background shear by the perturbation. If slow fluid is lifted up and fast fluid pulled down, then the mean shear has been reduced. Hence, there are obvious physical limits to the algebraic growth (19), (23), that would appear as nonlinear saturation effects in a more complete analysis. These and other issues are discussed in [31, 32] and in section 3.5, briefly, system (18) $\dot{u} = v$, $\dot{v} = 0$ with its algebraic growth $u = v_0 t$ is merely the linearization of

$$\frac{d}{dt}\begin{pmatrix} M \\ u \\ v \end{pmatrix} = \begin{pmatrix} -uv \\ Mv \\ 0 \end{pmatrix}, \tag{24}$$

about $M = 1$, $u = 0$, $v = 0$. System (24) does not have algebraic growth, it simply has continual *redistribution* of the mean shear $M$ into *'streaks'* $u$ then back to $M$, with a periodic general solution $M = M_0 \cos(v_0 t) - u_0 \sin(v_0 t)$, $u = u_0 \cos(v_0 t) + M_0 \sin(v_0 t)$, $v = v_0$.

**Viscous details: Stokes eigenmodes**

For finite $R$, it is clear that (22) has exponentially *decaying* eigenmodes with $u \neq 0$ but $v = 0$, and exponentially decaying $v$ eigenmodes that force a *transient growth* of $u$ with ultimate exponential decay. The former are simple eigenmodes of the heat equation for $u$, with even modes

$$u = \cos\left((2n-1)\frac{\pi}{2}y\right)\, e^{i\gamma z} \tag{25}$$

and odd modes

$$u = \sin(n\pi y)\, e^{i\gamma z} \tag{26}$$

for no-slip $u = 0$ at $y = \pm 1$, with $n = 1, 2, \ldots$ The latter are *Stokes eigenmodes* for $v$, *i.e.* the solutions of $\nabla^2\nabla^2 v = \lambda \nabla^2 v$. For channel geometries with no-slip walls at $y = \pm 1$, GFD

fellows easily derive or at least verify that the Stokes eigenmodes consist of even modes

$$v = \left( \frac{\cosh \gamma y}{\cosh \gamma} - \frac{\cos \beta y}{\cos \beta} \right) e^{i\gamma z} \qquad \text{with} \qquad \gamma \tanh \gamma + \beta \tan \beta = 0 \tag{27}$$

and odd modes

$$v = \left( \frac{\sinh \gamma y}{\sin \gamma} - \frac{\sin \beta y}{\sin \beta} \right) e^{i\gamma z} \qquad \text{with} \qquad \frac{\gamma}{\tanh \gamma} - \frac{\beta}{\tan \beta} = 0 \tag{28}$$

where $\gamma$ is a spanwise $z$ wavenumber. The hyperbolic terms satisfy $\nabla^2 v = 0$ and show up to enforce the clamped boundary conditions $v = \partial v/\partial y = 0$ at $y = \pm 1$. The trigonometric terms satisfy $\nabla^2 v = -(\beta^2 + \gamma^2)\ v$, so these Stokes eigenmodes decay like $e^{-(\beta^2+\gamma^2)t/Re}$ although $\beta$ is different for each mode, and there is a discrete infinity of $\beta$'s for each $\gamma$. A simple graphical analysis quickly shows that, for each $\gamma \neq 0$, the even modes (27) have $\beta = \beta_n$ with $(2n-1)\pi/2 < \beta_n < n\pi$ with $n = 1, 2, \ldots$ and $\beta_n \sim n\pi$ for $n$ large, or $\gamma$ small. Likewise, the odd modes have $n\pi < \beta_n < (2n+1)\pi/2$ and $\beta_n \sim (2n+1)\pi/2$ for large $n$ or small $\gamma$, with $n = 1, 2, \ldots$.

### 2.3 Transient growth, non-normal and Jordan matrices

Hence, the simple model (17) with $\sigma = 0$

$$\frac{d}{dt} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -k_u^2/R & 1 \\ 0 & -k_v^2/R \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \tag{29}$$

is almost an exact model of (22). For smallest viscous damping, $k_u^2 = \pi^2/4 + \gamma^2$ and $k_v^2 = \beta_1^2 + \gamma^2$ with $\pi/2 < \beta_1 < \pi$ the smallest solution of $\gamma \tanh \gamma + \beta \tanh \beta = 0$. These smallest $k_u^2$ and $k_v^2$ correspond to the first even modes for both (25) and (27) and $k_v^2 > k_u^2$.

System (29) does not have exponential growth and it does not have the algebraic growth for all times of the degenerate system (18), it has exponential decay and possible *transient algebraic growth* of $u$ with ultimate exponential decay. The general solution of (29) is

$$\begin{pmatrix} u \\ v \end{pmatrix} = u_0\, e^{-\nu k_u^2 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v_0\, e^{-\nu k_v^2 t} \begin{pmatrix} \frac{e^{\delta t}-1}{\delta} \\ 1 \end{pmatrix}, \tag{30}$$

where $\nu = 1/R \geq 0$ and $\delta = \nu(k_v^2 - k_u^2) \geq 0$, and the limit $R \to \infty$ is indeed (19). All components of (30) decay exponentially, except for the *forced response* $u_f$ of $u$ to $v$ which is such that

$$\frac{u_f}{v_0} = e^{-\nu k_v^2 t}\, \frac{e^{\delta t} - 1}{\delta} = \frac{e^{-\nu k_u^2 t} - e^{-\nu k_v^2 t}}{\delta} = e^{-\nu k_u^2 t}\, \frac{1 - e^{-\delta t}}{\delta}, \tag{31}$$

and this forced response is bounded from below and from above by

$$t\, e^{-\nu k_v^2 t} \leq \frac{u_f}{v_0} \leq t\, e^{-\nu k_u^2 t}. \tag{32}$$

These bounds follow easily from $1 + x \leq e^x$ and $1 - x \leq e^{-x}$ for all real $x$ (here $x \equiv \delta t$) and the equality occurs only at $t = 0$ if $\delta \neq 0$ or for all $t$ if $\delta = 0$. These bounds are interesting

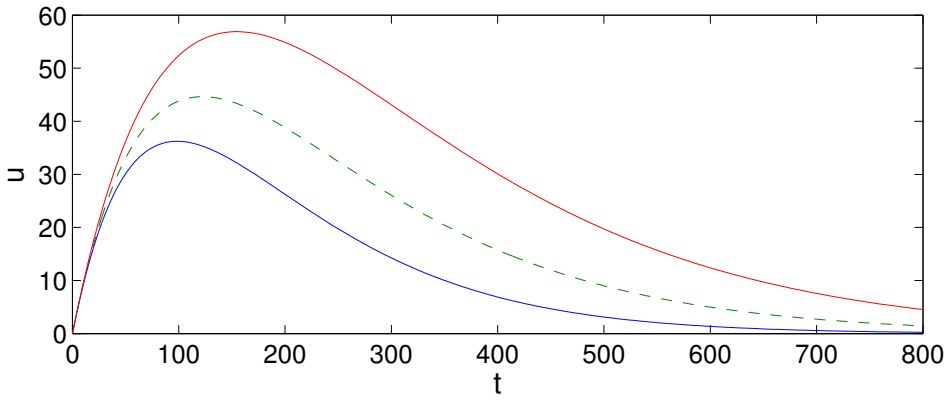


Figure 6: The forced response $u_f/v_0$ in (31), dashed, bracketed by the Jordan block bounds (32) for $R = 1000$ and $k_u^2 = \pi^2/4 + \gamma^2 \approx 6.4674$, $k_v^2 = \beta_1^2 + \gamma^2 \approx 10.1551$ for $\gamma = 2$ and $\beta_1 \approx 2.4809$ (27).

since they show that *the forced response is bounded by the forced responses of the degenerate forms of system (29),* that is (29) with $k_u \to k_v$ or $k_v \to k_u$ (Fig. 6). When both diagonal elements are $-\nu k_u^2$ or $-\nu k_v^2$, system (29) is again a Jordan block with only one eigenvector. The solution of those degenerate systems are (30) in the limit $\delta \to 0$ where $(e^{\delta t} - 1)/\delta \to t$ with $k_u$ or $k_v$ in both exponentials, depending on the case.

System (29) is not degenerate when $\delta = (k_v^2 - k_u^2)/R \neq 0$, there are two eigenvectors but those eigenvectors are almost parallel when $\delta \ll 1$ and become parallel as $\delta \to 0$. That system with $\delta \neq 0$ is *non-normal*[2] – its eigenvectors are not orthogonal except in the very viscous limit $R \to 0$. Non-normal systems have been emphasized and studied by Trefethen and co-workers and discussed at length by Schmid and Henningson in the context of shear flows [26], although that focus on general non-normal operators may be unnecessarily abstract for shear flows where there is a natural decoupling into exponentially decaying modes for $v$ with a *forced response* of the streamwise velocity perturbation $u$, as illustrated in (22) and the model (29). The bounds (32) also suggests that the *transient growth* associated with highly non-normal matrices – matrices with almost parallel eigenvectors – are tightly bounded by nearby degenerate systems (*i.e.* containing Jordan blocks) and that the initial conditions that lead to large transient growth are simply *orthogonal to the existing incomplete set of eigenvectors.* For the degenerate forms of (29), the only eigenvector is $(1, 0)$ and the initial condition giving most transient growth is $(0, 1)$.

It is now a simple Calculus exercise to show that the forced response (31) starts at $u_f = 0$ when $t = 0$, grows like $v_0 t$ for $(k_u^2 + k_v^2)t \ll R$ and ultimately decays like $(v_0/\delta)\, e^{-\nu k_u^2 t}$ for $t \gg 1/\delta$, while reaching a maximum of

$$\max(u_f) = v_0 \, \frac{R}{k_v^2} \, \exp\left(-k_u^2 \, \frac{\ln k_v^2 - \ln k_u^2}{k_v^2 - k_u^2}\right) \tag{33}$$

[2] *Normal* is in the sense of *orthogonal* here. An *orthogonal* matrix has columns orthogonal to each other. A *normal* matrix has *eigenvectors* orthogonal to each other.

at time

$$t_* = R\,\frac{\ln k_v^2 - \ln k_u^2}{k_v^2 - k_u^2} = R\,\frac{\ln(k_v^2/k_u^2)}{k_v^2 - k_u^2}. \tag{34}$$

Hence, there is a maximum growth of $O(v_0\,R)$ for $u_f$ occurring at a time of $O(R)$.

# 3 Transition threshold: mechanisms and scalings

## 3.1 Transient growth + nonlinear feedback and $\epsilon \sim R^{-3}$

Our discussions in the previous sections makes clear that the 'subcritical logistic' model $du/dt = -u/R + u^2$, eqn. (14), is too simple for shear flows, but a more plausible model might be

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\nu k_u^2 & 1 \\ 0 & -\nu k_v^2 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ u^2 \end{pmatrix}. \tag{35}$$

This is model (29) with $\nu = 1/R$ but with a quadratic nonlinear feedback from $u$ onto $v$. If $\epsilon$ is a measure of the initial $v_0$, that is simply $\epsilon = v_0$ for this simple model, the transient growth of $u$ will lead to a maximum $u \sim v_0 R = \epsilon R$, yielding a quadratic interaction $u^2 \sim \epsilon^2 (R)^2$ that must balance the viscous damping $v/R$ to reach the *transition threshold*. The transition scenario is

$$v_0 = \epsilon \;\longrightarrow\; u \sim \epsilon R \implies v \sim u^2 R \sim \epsilon^2 R^3 \sim \epsilon \tag{36}$$

and the threshold scaling would be $\epsilon \sim R^{-3}$ with $a = -3$ in (9). TTRD [29] and Baggett and Trefethen [3] discuss similar simple models to illustrate *'nonlinear recycling of [transiently amplified] outputs into inputs'*.

However there are at least two basic reasons why such simple models are not valid models for Navier-Stokes. First, direct nonlinear feedback as in (35) is not allowed in the Navier-Stokes equations, and many other classic nonlinear and weakly nonlinear systems, as we already briefly discussed in section 1.4. Second, there is in fact virtually *no* 'nonlinear recycling of outputs into inputs' and the primary nonlinear effect of transient growth is actually to wipe out the background shear and saturate the growth, not regenerate 'optimal' disturbances. This is exactly true for the strongly amplified $x$-independent disturbances, as we prove in section 3.5, although not exactly true for more weakly amplified *oblique* disturbances, and this has led to sustained confusion among people who have been seduced by linear non-normal operators but have shied away from detailed analysis of nonlinear interactions. These issues are illustrated using simple models in the following.

## 3.2 Benney-Gustavsson mechanism and $\epsilon \sim R^{-2}$

Benney and Gustavsson (1981) [6] knew that streamwise independent perturbations induce a strong redistribution of streamwise velocity but do not trigger transition inducing nonlinear effects. They searched then for a mechanism based on a *pair of oblique perturbations* and involving transient growth in the case of *direct resonance* – when $k_u^2 = k_v^2$ in (29) – and the limit of high $R$. Later studies of linear transient growth emphasized that direct resonances is not necessary for large growth, as discussed here in section 2.3 and figure 6. However,

oblique disturbances in shear flows travel in the streamwise $x$ direction, that is, the diagonal terms in (29) are complex in general and their imaginary parts do not scale like $1/R$, then large and distinct imaginary parts shut down the growth by phase shifting of the forcing $v$ and the forced $u$. Thus, although exact 'direct resonance' is not necessary for large transient growth, it still remains that the imaginary parts of the eigenvalues for the $v$ and $\eta$ modes of oblique disturbances must be sufficiently close to each other for growth to occur.

Let $v$ denote the amplitude of the vertical (*i.e.* shearwise) $y$ velocity for a pair of oblique perturbations, and $\eta$ represent the amplitude of the corresponding perturbation vertical vorticity ($\eta = \partial_z u - \partial_x w$). The latter $\eta$ substitutes for the transient amplified streamwise velocity $u$ in the $x$-independent dynamics. Benney and Gustavsson suggested the transition scenario

$$v = \epsilon \longrightarrow \eta \sim \epsilon t \Longrightarrow V, N \sim (\epsilon^2 t^2) \Longrightarrow v \sim (\epsilon^2 t^2)(\epsilon t)t = \epsilon^3 t^4, \tag{37}$$

in the limit of high $R$, where $v$ and $\eta$ are the amplitudes of $e^{i(\alpha x \pm \gamma z)}$ perturbations with linear eigenstructures in $y$, $V$ and $N$ are the amplitudes of $v$ and $\eta$ for the harmonics $e^{i2(\alpha x \pm \gamma z)}$ or $e^{i2\gamma z}$ and '$\longrightarrow$' is the linear transient growth, the first '$\Longrightarrow$' is the nonlinear interactions $\eta\eta$ and $\eta\eta^*$ and the second '$\Longrightarrow$' is the nonlinear interactions $V\eta$ and $N\eta$. This led them to suggest a nonlinear mechanism occurring on a time scale $t \sim 1/\sqrt{\epsilon}$, the time scale when the nonlinear feedback on $v$ is of the order of the original amplitude $\epsilon^3 t^4 \sim \epsilon$, as laid out in (37).

In essence, what Benney and Gustavsson said is what we already discussed in sect. 1.4, for Navier-Stokes, and many other physical nonlinear systems, if a mode $v$ linearly creates a mode $u$ it does not make sense to have the quadratic interaction of the latter, $u^2$, directly feedback onto $v$. For instance, if $v$ and $u$ correspond to $e^{i(\alpha x \pm \gamma z)}$ then $u^2$ would be associated with the harmonic $e^{i2(\alpha x \pm \gamma z)}$ and the latter is orthogonal to the original modes, so we cannot have $u^2$ forcing of $v$ as in (35). However, it is possible to have feedback at the next order of nonlinear interaction. In the context of our simple models, Benney and Gustavsson in effect said that (35) cannot be a suitable model for shear flows, but a possible model could be

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\nu k_u^2 & 1 \\ 0 & -\nu k_v^2 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ u^3 \end{pmatrix}, \tag{38}$$

with a *cubic* nonlinear feedback $u^3$ onto $v$, not quadratic feedback $u^2$. In the inviscid limit, $\nu = 1/R \to 0$, this model has the scalings $dv/dt = u^3 \sim \epsilon^3 t^3$, so $v \sim \epsilon^3 t^4$ as in (37), while the *threshold scaling* is $\nu k_v^2 v = u^3 = (v/(\nu k_u^2))^3$ with $v = \epsilon$, yielding

$$\frac{\epsilon}{R} \sim (\epsilon\, R)^3 \quad \Leftrightarrow \quad \epsilon \sim R^{-2} \tag{39}$$

with $a = -2$ in (9). This is weaker than $\epsilon \sim Re^{-3}$ in (36) but still a very strong mechanism, especially compared to the classic weakly nonlinear scaling of a system such as $du/dt = -u/R + u^3$ for which $\epsilon \sim R^{-1/2}$. However, there is doubt as to the relevnace of this model as discussed in the next section.

### 3.3 Waleffe-Kim-Hamilton mechanism and $\epsilon \sim R^{-3}$

Waleffe *et al.* (1991) [37] suggested that the Benney-Gustavsson mechanism might in fact be much stronger because of several additional transient amplifications that could occur.

They considered the scenario

$$v = \epsilon \longrightarrow \eta \sim \epsilon t \Longrightarrow V \sim (\epsilon^2 t^2)t \longrightarrow N \sim \epsilon^2 t^4 \Longrightarrow v \sim (\epsilon^2 t^4)(\epsilon t)t = \epsilon^3 t^6, \tag{40}$$

with feedback thus occurring when $\epsilon \sim \epsilon^3 t^6$ on the time scale $t \sim \epsilon^{-1/3}$ much faster than the Benney-Gustavsson $t \sim \epsilon^{-1/2}$. In essence, Waleffe *et al.* suggested that the dynamics of $V$ and $N$ could not be neglected and that a more powerful mechanism for shear flows might be modeled by

$$\frac{d}{dt}\begin{pmatrix} U \\ V \\ u \\ v \end{pmatrix} = \begin{pmatrix} -\nu k_U^2 & 1 & 0 & 0 \\ 0 & -\nu k_V^2 & 0 & 0 \\ 0 & 0 & -\nu k_u^2 & 1 \\ 0 & 0 & 0 & -\nu k_v^2 \end{pmatrix} \begin{pmatrix} U \\ V \\ u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ uu \\ 0 \\ uU \end{pmatrix}. \tag{41}$$

In this model, $u$ and $v$ represent the amplitudes of horizontal and vertical velocities, respectively, for *oblique* perturbations, $e^{i(\gamma z \pm \alpha x)}$ modes say, while $U$ and $V$ represent horizontal and vertical velocities for spanwise perturbations ($x$-independent), $e^{i2\gamma z}$ modes. Thus $u$ and $U$ in (41) represent the horizontal velocities associated with vertical *vorticity* modes $\eta$ and $N$ in (40) and that scenario corresponds to (41) for $\nu = 0$ and initial conditions $v = \epsilon$ with $u = U = V = 0$.

Both $v$ and $V$ lead to linear transient growth of $u$ and $U$ respectively, both from the 'lift-up' of the background laminar shear flow, showing up mathematically with the 1's on the off diagonal in (41). The quadratic interaction $u^2$ does not feedback on $v$ but it forces $V$ corresponding to an $e^{i2\gamma z}$ spanwise mode, which creates a large $U$ and finally the nonlinear interaction $uU$ feeds back on the original $v$. This yields the transition scenario

$$v = \epsilon \longrightarrow u \sim \epsilon R \Longrightarrow V \sim \epsilon^2 R^3 \longrightarrow U \sim \epsilon^2 R^4 \Longrightarrow v \sim \epsilon^3 R^6 \sim \epsilon, \tag{42}$$

and thus a threshold scaling $\epsilon \sim R^{-3}$.

So, is $\epsilon \sim Re^{-3}$, the same scaling as in the naive model (35), possible in shear flows?! Waleffe *et al.* [37] investigated these mechanisms using careful analysis of full Navier-Stokes simulations and showed that the nonlinear generation of $V$ came from $vv$ terms not $uu$, the latter being essentially zero! [37, Fig. 4], thus nonlinear interactions *completely bypassed the linear transient growth* of $u$ that was occurring simultaneously. The forcing of $v$ from $uU$ was not analyzed but is also believed to be insignificant at transitional Reynolds numbers. Thus, the nonlinear forcing of $v$ and $V$ could be, for instance,

$$\frac{d}{dt}\begin{pmatrix} U \\ V \\ u \\ v \end{pmatrix} = \begin{pmatrix} -\nu k_U^2 & 1 & 0 & 0 \\ 0 & -\nu k_V^2 & 0 & 0 \\ 0 & 0 & -\nu k_u^2 & 1 \\ 0 & 0 & 0 & -\nu k_v^2 \end{pmatrix} \begin{pmatrix} U \\ V \\ u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ vv \\ 0 \\ vV \end{pmatrix}. \tag{43}$$

with nonlinearities arising from $v$ and $V$, not from the transiently amplified $u$ and $U$. Transient growth of $u$ and $U$ occurs in (43) but is *not* involved in transition. System (43) has a threshold $\epsilon \sim R^{-1}$ from the scenario

$$v = \epsilon \Longrightarrow V \sim \epsilon^2 R \Longrightarrow v \sim \epsilon^3 Re^2 \sim \epsilon. \tag{44}$$

Keeping $uU$ instead of $vV$ for the $v$ forcing would yield a threshold $\epsilon \sim R^{-2}$, while a $Uv$ term would yield $\epsilon \sim R^{-3/2}$.

These observations also negate the Benney-Gustavsson scaling (39) since there would be no $u^3$ but instead perhaps only a $v^3$ in (38), leading back to the weakly nonlinear $\epsilon \sim R^{-1/2}$. Thus, careful analysis of numerical simulations must be made since many linear and nonlinear processes are occurring concurrently in full numerical simulations but they may not have cause and effect connections. *The nonlinear interactions leading to transition could completely bypass the linear transient growth.*

### 3.4 Chapman's viscous correction of the WKH mechanism

Chapman (2002) [8] — apparently unaware of earlier work[3] by Waleffe, Kim and Hamilton [37] on the Benney-Gustavsson mechanism and the self-sustaining process [13, 31, 32, 33, 34, 35]— considered transition scenarios essentially identical to that discussed in section 3.3. His toy model (2.6)–(2.9) which reads

$$\frac{d}{dt}\begin{pmatrix}\phi_1\\ \psi_1\\ \phi_2\\ \psi_2\end{pmatrix} = \begin{pmatrix}-\epsilon_c & 1 & 0 & 0\\ 0 & -2\epsilon_c & 0 & 0\\ 0 & 0 & -\delta & 1\\ 0 & 0 & 0 & -2\delta\end{pmatrix}\begin{pmatrix}\phi_1\\ \psi_1\\ \phi_2\\ \psi_2\end{pmatrix} + \begin{pmatrix}0\\ \phi_2^2\\ 0\\ \phi_1\phi_2\end{pmatrix}. \tag{45}$$

is essentially identical to model (41). The exact correspondence between Chapman's (2.6–9) and (41) is

$$(\phi_1, \psi_1, \phi_2, \psi_2) \equiv (U, V, u, v). \tag{46}$$

The order of Chapman's variables $(\phi, \psi)$ has been reversed from his (2.6)–(2.9) to match (41) and the double Jordan block-like structure.

Unfortunately, Chapman uses '$\epsilon$' for the viscous decay rate of the streaks $\phi_1 \equiv U$, which we call $\nu k_U^2$, while we have used $\epsilon$ for perturbation amplitude, so we use $\epsilon_c$ for his $\epsilon$ to distinguish from ours. He uses $\delta$ for our $\nu k_u^2$. He assumes that $k_V^2 = 2k_U^2$ and $k_v^2 = 2k_u^2$, but this is inconsequential for our questions of threshold scalings and mechanisms.

Chapman correctly includes the possibility that the viscous decay rates of streamwise-independent (his $\epsilon_c$) and oblique modes (his $\delta$, with $0 < \epsilon_c \ll \delta \ll 1$) scale differently with $R$, while we assumed[4] in section 3.3 that both scale like $R^{-1}$. The stronger decay rate represented by Chapman's $\delta \gg \epsilon_c$ arises from the critical layer structure of oblique eigenmodes of the linearized operators. That critical layer structure has a scale of $O(R^{-1/3})$ thus the effective wavenumbers $k_u$ and $k_v$ for oblique linear eigenmodes in (41) are both $O(R^{1/3})$ and $\delta = \nu k_u^2 \sim R^{-1}R^{2/3} = R^{-1/3}$.

The two time scales $\epsilon_c^{-1}$ and $\delta^{-1}$ in Chapman's model lead to two different scenarios with different scalings. His scenario (i) has $\psi_1 \equiv V$ transiently amplifying $\phi_1 \equiv U$ to an amplitude $1/\epsilon_c \equiv R/k_U^2$. That in itself will not trigger transition since it is obvious in (41), (45) that the nonlinear term vanishes identically if $\phi_2 \equiv u = 0$. The idea is that $\phi_1 \equiv U$ leads to an instability of the oblique modes $\phi_2 \equiv u, \psi_2 \equiv v$. To illustrate this and obtain

[3] An obvious failure of the JFM review process

[4] In the original Benney-Gustavsson and Waleffe *et al.* work, the focus was on nonlinear time scales in the $R \to \infty$ limit as discussed in sections 1.4, 3.2, 3.3.

the scaling we can rewrite Chapman's (2.8), (2.9), the $\phi_2$ and $\psi_2$ equations in (45), in the matrix form

$$\frac{d}{dt}\begin{pmatrix}\phi_2\\ \psi_2\end{pmatrix} = \begin{pmatrix}-\delta & 1\\ \phi_1 & -2\delta\end{pmatrix}\begin{pmatrix}\phi_2\\ \psi_2\end{pmatrix} \tag{47}$$

where $\delta > 0$ and this readily suggests that if $\phi_1 \gtrsim 2\delta^2$ there might be growth of $\phi_2$ and $\psi_2$ that might be interpreted as an instability of the 'streaks' $\phi_1 \equiv U$. A *streak instability* is a key part of the *self-sustaining process* discussed in sections 3.6, 3.7 below, and during a decade before Chapman's work [37, 13, 31, 32, 33, 34, 35], but what we have here is not quite a 'streak instability', it is an instability of the laminar shear (the 1 in the top right corner of the matrix in (47)) coupled with the streak amplitude $\phi_1 \equiv U$. A streak instability would have $\phi_1 \equiv U$ in *both* off-diagonal elements as in [33, eqns. (13)-(15)] and an inviscid ($\delta = 0$) growth rate of $U$, instead of $\sqrt{U} \equiv \sqrt{\phi_1}$ as in (47).

Since $\psi_1$ amplifies $\phi_1$ to $\phi_1 \sim \psi_1/\epsilon_c$, we can therefore conclude with Chapman that something similar to a 'streak instability' might occur if $\phi_1 \sim \psi_1/\epsilon_c \gtrsim \delta^2$. This growth would be on the $1/\delta$ time scale, much faster than the slower $1/\epsilon_c$ time scale. Scenario (i) therefore is

$$\psi_1 \longrightarrow \phi_1 \sim \frac{1}{\epsilon_c}\psi_1 \rightsquigarrow \psi_2, \phi_2 \Longrightarrow \psi_1 \quad \text{if} \quad \psi_1 \gtrsim \epsilon_c\delta^2. \tag{48}$$

where '$\longrightarrow$', '$\rightsquigarrow$' and '$\Longrightarrow$' denote linear amplification, instability and nonlinear feedback, respectively. If $\epsilon_c \sim R^{-1}$ and $\delta \sim R^{-1}$ this would yield a $\psi_1 \sim R^{-3}$ threshold as in (42), but if $\delta \sim R^{-1/3}$ as for oblique linear modes in shear flows, then the threshold would be $\psi_1 \sim R^{-5/3}$.

Chapman's scenario (ii) begins with the oblique rolls, $\psi_2 \equiv v$ that linearly amplify $\phi_2 \equiv u$ which quadratically forces $\psi_1 \equiv V$ that linearly amplifies $\phi_1 \equiv U$ and the quadratic interaction $\phi_1\phi_2 \equiv Uu$, at last, feeds back onto the original $\psi_2 \equiv v$. But there is a catch! now the initial transient growth of $\phi_2$ occurs (and peaks) on the faster time scale $1/\delta$ and $\psi_1$ only reaches $\sim \phi_2/\delta$ on that time scale. Diagrammatically,

$$\psi_2 \longrightarrow \phi_2 \sim \frac{1}{\delta}\psi_2 \Longrightarrow \psi_1 \sim \frac{1}{\delta}\phi_2^2 \longrightarrow \phi_1 \sim \frac{1}{\epsilon_c}\psi_1 \Longrightarrow \psi_2 \sim \frac{1}{\delta}\phi_1\phi_2 \tag{49}$$

putting it all together yields $\psi_2 \sim \psi_2^3/(\epsilon_c\delta^5)$ and the threshold scaling $\psi_2 \sim \epsilon_c^{1/2}\delta^{5/2}$ [8, Fig. 3]. For $\epsilon_c \sim R^{-1}$ and $\delta \sim R^{-1}$, the threshold would be $\psi_2 \sim R^{-3}$ as in (42), but for $\delta \sim R^{-1/3}$ the threshold would be $\psi_2 \sim R^{-4/3}$.

This is all very well, however a key problem is that it is built on the *assumption* that the quadratic interactions $uu$ and $Uu$ of the transiently amplified disturbances, $U \equiv \phi_1$ and $u \equiv \phi_2$ are the dominant nonlinear interactions, as in the naive model (35), the Benney-Gustavsson model (38) and the Waleffe, Kim & Hamilton model (41). Those nonlinear terms appear to be almost non-existent according to the numerical analysis of [37] as discussed in the previous section, so these transition scenarios and thresholds are likely to not be the effective scenarios for Navier-Stokes.

### 3.5 Nonlinear saturation of linear transient growth and '$\epsilon = \infty$'

Transient amplification by a factor of $R$ only occurs for streamwise $x$ independent perturbations, $\partial/\partial x = 0$, but nonlinear feedback does *not* occur for such disturbances. In fact,

$x$-independent perturbations form an *invariant manifold for shear flows for which the laminar flow is the global attractor*, that is, $x$-independent perturbations stay $x$-independent and the cross-stream velocities $v, w$ decouple from the transiently amplified streamwise velocity $u$, thus $v, w$ decay viscously and $u$ eventually returns to zero. This was proved by Joseph and Tao in 1963 [17], see also [16]. The proof is straightforward. Consider (20) with $\partial/\partial x = 0$, then $\mathbf{u} \cdot \nabla = v\partial_y + w\partial_z$ and $v$ and $w$ decouple from $u$, then multiply the $v$ equation by $v$, the $w$ equation by $w$, add the resulting equations and integrate over the cross-section. Integration by parts and the boundary conditions eliminate the advection and pressure terms and we are left with

$$\frac{d}{dt}\int_A \frac{v^2+w^2}{2}\,dA = -\frac{1}{R}\int_A \left(|\nabla v|^2 + |\nabla w|^2\right) dA \;\leq 0 \tag{50}$$

where $A$ is the flow cross-section and $dA$ is its area element. Thus $v$ and $w \to 0$ and in that limit the $u$ equation becomes a simple heat equation and $u \to 0$ also. This is a fully nonlinear result but only for $x$-independent perturbations. There is transient growth of $u$ but no nonlinear feedback on $v$.

The primary nonlinear effect, in fact, is to *reduce the transient growth of $u$ by reducing the background shear*. A model for streamwise independent perturbations in a shear flow, more physical than (29), is then [31]

$$\frac{d}{dt}\begin{pmatrix} S \\ U \\ V \end{pmatrix} = \begin{pmatrix} -\nu k_S^2 & 0 & 0 \\ 0 & -\nu k_U^2 & 1 \\ 0 & 0 & -\nu k_V^2 \end{pmatrix}\begin{pmatrix} S \\ U \\ V \end{pmatrix} + \begin{pmatrix} -UV \\ SV \\ 0 \end{pmatrix} \tag{51}$$

for which $S = U = V = 0$ is the global attractor since $V \to 0$, then $U$ and $S$ also $\to 0$.

In model (51), $V$ induces transient growth of $U$ that leads to $-UV < 0$ that creates $S < 0$ which reduces the forcing of $U$ from $V$ to $(1+S)V < V$, but there is no transition since there is no feedback on $V$. $S$ models the perturbation of the mean shear so that the total mean shear is $M = 1 + S$. The nonlinearity creates a 'Reynolds stress', $-UV$, that reduces the shear from 1 to $M = 1 + S < 1$ and the 'lift-up' term creating $U$ is then $(1+S)V = MV$ instead of $V$.

Note that the nonlinearity is energy conserving and

$$\frac{1}{2}\frac{d}{dt}\left(S^2 + U^2 + V^2\right) = UV - \nu\left(k_S^2 S^2 + k_U^2 U^2 + k_V^2 V^2\right) \tag{52}$$

which is entirely analogous to the perturbation energy equation in shear flows[5]

$$\frac{1}{2}\frac{d}{dt}\int_V |\mathbf{u}|^2 dV = \int_V -uv\frac{dU}{dy}dV - \nu\int_V |\nabla\mathbf{u}|^2 dV. \tag{53}$$

The latter is obtained by dotting the full perturbation equations (20) with $\mathbf{u}$ and integrating by parts over a volume $V$ with periodic or vanishing perturbations on its boundary.[9, §53]

[5]Note that there is a sign difference between the hydrodynamics (22) and our models since $uv < 0$ from 'lift-up' for $dU/dy > 0$ but our models have $UV > 0$ (these $U$'s are the forced response from $V$). This is only a difference in definition that can be removed by defining the flow so that $dU/dy < 0$ or considering that the $V$ in our models in fact corresponds to $-V$ and 'pull-down' in the hydrodynamics.

The $S$ equation in (51) is inspired by the mean flow equation, that is, let $\mathbf{v} = U(y)\hat{\mathbf{x}}+\mathbf{u} = U(y)\hat{\mathbf{x}} + \bar{u}(y,t)\hat{\mathbf{x}} + \tilde{\mathbf{u}}$ where $U(y)\hat{\mathbf{x}}$ is the laminar flow and $\bar{u}(y,t)\hat{\mathbf{x}}$ is the average of $\mathbf{u}$ over $x$ and $z$ with $\tilde{\mathbf{u}}$ the remaining fluctuating part. Averaging the $x$ component of the Navier-Stokes equations (3) yields the mean flow perturbation equation

$$\partial_t \bar{u} = -\partial_y \overline{\tilde{u}\tilde{v}} + \nu \partial_y^2 \bar{u}, \tag{54}$$

which leads to the $S$ equation in (51) if we assume a reasonable shape for $\overline{\tilde{u}\tilde{v}}$, say $\cos^2(\pi y/2)$ with $s$ being the amplitude of a $\pi^{-1}\sin\pi y$ shape for $\bar{u}(y,t)$, for instance.

This is only a justification for the model, not a derivation, but (51) does capture the 'fully nonlinear' physics of streamwise independent shear flows. The model is in fact linear since $V$ is decoupled from $S$ and $U$. Model (51) captures the full physics of *redistribution of streamwise velocity*, $(1+S)V$ in the $U$ equation together with $-UV$ in the $S$ equation, not simply the linearized *lift-up* which is merely the $V$ term in the $U$ equation. Substituting $S = M - 1$ in (51) yields

$$\frac{dM}{dt} = \nu k_S^2 (1-M) - UV, \qquad \frac{dU}{dt} = -\nu k_U^2 U + MV, \qquad \frac{dV}{dt} = -\nu k_V^2 V \tag{55}$$

which is the viscous version of system (24).

### 3.6 SSP and $\epsilon \sim R^{-3/2}$, $R^{-2}$

Waleffe [30] proposed a different mechanism — the Self-Sustaining Process (SSP)— that was later developed in a series of papers [37, 13, 31, 33], culminating in the construction of 'exact coherent states' for the full 3D Navier-Stokes equations [34, 35, 36]. The self-sustaining process was inspired by work of Benney (1984) [5] and experiments of Acarlar and Smith (1987) [1]. This process can be illustrated by a low order model that begins with the good model of $x$-independent dynamics (51), with $x$-independent modes $S$, $U$, $V$. We know from sect. 3.5 that such $x$-independent mean flows cannot be self-sustained, so we need an $x$-dependent fluctuation to obtain a self-sustaining process. Call that fluctuation $w$ for 'wave' and assume that it has a simple $e^{i\alpha x}$ form in $x$, with $w^*$ the amplitude of $e^{-i\alpha x}$ since the total flow must be real. The model is [31, 32]

$$\frac{d}{dt}\begin{pmatrix} S \\ U \\ V \\ w \end{pmatrix} = \begin{pmatrix} -\nu k_S^2 & 0 & 0 & 0 \\ 0 & -\nu k_U^2 & 1 & 0 \\ 0 & 0 & -\nu k_V^2 & 0 \\ 0 & 0 & 0 & -\nu k_w^2 \end{pmatrix} \begin{pmatrix} S \\ U \\ V \\ w \end{pmatrix} + \begin{pmatrix} -UV \\ SV - ww^* \\ ww^* \\ Uw - Vw \end{pmatrix}. \tag{56}$$

This is model (51) with an extra $w$ equation and a series of nonlinear interactions associated with $w$. The model reduces to (51) when $w = 0$, *i.e.* when the flow is $x$-independent.

This model contains not just the 'lift-up' of background laminar shear — the '1' in the row $U$, column $V$ of the matrix — but the complete *redistribution* of the mean shear $M = 1+S$ by $V$ that necessarily comes with the reduction of the mean shear $M = 1+S < 1$ by the Reynolds stress $-UV$. This redistribution leads to the formation of large $x$-independent streamwise velocity fluctuations $U$ whose nonlinear self-interactions $U^2$ are non-existent (unlike models (35, 38, 41)) but that may be unstable to an $x$-dependent $e^{i\alpha x}$ 'wave' of

amplitude $w$. The nonlinear interaction $ww^*$ of that mode with its complex conjugate $e^{-i\alpha x}$ yields a *negative feedback* on $U$, since the latter provides the energy source for $w$, but also a *positive feedback* on $V$. The latter comes with a negative feedback on $w$ and the nonlinear term is energy conserving so the energy equation for (56) is

$$\frac{1}{2}\frac{d}{dt}\left(S^2+U^2+W^2+w^2\right) = UV - \nu\left(k_S^2S^2+k_U^2U^2+k_V^2V^2+k_w^2w^2\right) \tag{57}$$

again entirely analogous to the hydrodynamic equivalent (53).

This mechanism is more complex and involves more physical steps. Models (35, 38, 41) are all essentially 2-steps: linear transient amplification of $\epsilon$ into $\epsilon R$ followed by quadratic interactions of the transiently amplified perturbation yielding *presumed* feedback onto $\epsilon$. Model (56) involves redistribution of mean shear and therefore 'nonlinear' reduction of mean shear, with exponential instability of transiently amplified fluctuations and various quadratic interactions of the latter growing mode leading to self-sustenance. We call it a *process* — the *self-sustaining process* — since it involves more steps.

Transition will not happen if $w \equiv 0$ in (56) since $w$ will stay zero and the model then reduces to (51) which will always decay to $S = U = V = 0$. The transition scenario based on this process is not as straightforward because the $w$ equation is linear in $w$ and does not directly determine its amplitude. The $V$, $U$ and $S$ equations yield, respectively,

$$V \sim w^2R, \qquad U \sim (1+S)w^2R^2 - w^2R, \qquad S \sim -Uw^2R^2. \tag{58}$$

This mean shear reduction $S < 0$ must not shut down production of $U$, so we need $(1+S)V \gtrsim w^2 + U/R$ and substituting for $S$ and $V$ from (58) this requires $w^2R^2 \gtrsim (1+w^4R^4)U$ to sustain $U$. We also need $U - V - \nu k_w^2 \gtrsim 0$ to sustain $w$ and since $V \sim w^2R$ and $\nu = 1/R$ that requires $U \gtrsim (1+w^2R^2)/R$. Sustenance of $U$ and $w$ thus requires

$$\frac{1+w^2R^2}{R} \lesssim U \lesssim \frac{w^2R^2}{1+w^4R^4} \tag{59}$$

which, in the limits $w\,R \ll 1$ and $w\,R \gg 1$, yields

$$R^{-3/2} \lesssim w \lesssim R^{-3/4}. \tag{60}$$

This would suggest a threshold exponent $a = -3/2$ in (9) with the transition scenario

$$w \sim R^{-3/2} \Longrightarrow V \sim R^{-2} \longrightarrow U \sim R^{-1},\, S \sim R^{-2} \;\rightsquigarrow w, \tag{61}$$

although we can also imagine the scenario

$$V \sim R^{-2} \longrightarrow U \sim R^{-1},\, S \sim R^{-2} \;\rightsquigarrow w \sim R^{-3/2} \Longrightarrow V \sim R^{-2} \tag{62}$$

where $w$ arises from an instability of $U$ and quickly grows to $w \sim R^{-3/2}$ to sustain the original $V$ perturbation. This closely related scenario would have a threshold exponent $a = -2$ in (9) as discussed in [3].

These (61), (62) are the scalings of *lower branch* steady states for (56) as given in [33, eqn. (24)]. The *upper branch* steady state for (56) pushes against the upper bound of (60) and has the scaling $w \sim R^{-3/4}$, $V \sim R^{-1/2}$, $U \sim R^{-1/2}$, $(1+S) \sim R^{-1}$ as given in [32, 33].

### 3.7 Derived SSP and $\epsilon \sim R^{-1}$

Model (56) was first presented at a Center for Turbulence Research seminar in 1990 but did not trigger much interest among engineers heavily involved into cutting-edge high resolution 3D numerical simulations of turbulent flows. The model did not appear in print until 1995 [31], prompted by the publication of models similar to (35) in TTRD [29]. Many simple models were proposed and analyzed in the mid 1990's [3], but few had any direct connection with the Navier-Stokes equations. A derivation of (56) from the Navier-Stokes equations was therefore attempted by Galerkin truncation as in the derivation of the well-known Lorenz-Saltzmann model of convection and chaos.

That derivation showed that a key interaction is missing in (56), there should be a $-(1+S)w \equiv -Mw$ term in the $w$ equation arising from the differential advection of that $x$-dependent mode by the mean shear $M = 1 + S$ (modes $A$, $B$, $C$, $D$, $E$ in [33, eqn. (10)] are coupled through $M$). For small $S$ this is a large extra damping for $w$ that has a direct impact on the transition threshold question. Thus the model should be

$$\frac{d}{dt}\begin{pmatrix} S \\ U \\ V \\ w \end{pmatrix} = \begin{pmatrix} -\nu k_S^2 & 0 & 0 & 0 \\ 0 & -\nu k_U^2 & 1 & 0 \\ 0 & 0 & -\nu k_V^2 & 0 \\ 0 & 0 & 0 & -\nu k_w^2 \end{pmatrix} \begin{pmatrix} S \\ U \\ V \\ w \end{pmatrix} + \begin{pmatrix} -UV + ww^* \\ SV - ww^* \\ ww^* \\ Uw - Vw - (1+S)w \end{pmatrix} \tag{63}$$

which is (56) with an extra $+ww^*$ in the $S$ equation and an extra $-(1+S)w$ in the $w$-equation. The energy equation is still (57) since we have merely added *redistribution* terms that exchange energy between $S$ and $w$ but do not change the total energy. That last term in the $w$ equation now includes a linear term, $-w$, and it is better therefore to rewrite the model in terms of the total mean shear $M = (1 + S)$, instead of the perturbation from laminar shear $S$, in which case (63) becomes

$$\frac{d}{dt}\begin{pmatrix} M \\ U \\ V \\ w \end{pmatrix} = \begin{pmatrix} -\nu k_S^2\, M \\ -\nu k_U^2\, U \\ -\nu k_V^2\, V \\ -\nu k_w^2\, w \end{pmatrix} + \begin{pmatrix} -UV + ww^* \\ MV - ww^* \\ ww^* \\ (U - V - M)w \end{pmatrix} + \begin{pmatrix} \nu k_S^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{64}$$

The '1' in row $U$, column $V$ of the linear coupling matrix in (63) has disappeared, it has been absorbed in the nonlinear term $MV$ in the $U$ equation in (64). The remaining matrix is now 'normal,' in fact it is diagonal and has been multiplied with the state vector. The full nonlinear term is energy conserving as in the Navier-Stokes equations, and a forcing has appeared to maintain the mean shear. Again, (64) is simply (63) with the change of variable $S = M - 1$. The laminar solution is $M = 1$, $U = V = w = 0$. The total energy equation reads

$$\frac{1}{2}\frac{d}{dt}\left(M^2 + U^2 + V^2 + w^2\right) = \nu k_S^2(1 - M) - \nu\left(k_U^2 U^2 + k_V^2 V^2 + k_w^2 w^2\right) \tag{65}$$

in lieu of (57). This energy equation shows that the total energy decays if $M > 1$, in other words $M < 1$ for transition.

As in model (56), transition will not happen if $w = 0$ and determining the amplitude of $w$ requires more delicate analysis. There are several cancellations and to keep track of them clearly it is better to label each interaction with distinct coefficients as in [33, eqn, (20)], but here we label only the $(M, w)$ interaction for simplicity. Let $\sigma_m > 0$ be the coefficient of interaction between $M$ and $w$, so $ww^* \to \sigma_m ww^*$ in the $M$ equation and $-Mw \to -\sigma_m Mw$ in the $w$-equation. The $V$, $U$ and $M$ equations yield the scalings

$$V \sim w^2 R, \qquad U \sim Mw^2R^2 - w^2R, \qquad M \sim 1 + \sigma_m w^2 R - Uw^2R^2 \tag{66}$$

instead of (58). Production of $U$ requires $MV \gtrsim k_u^2 U/R + ww^*$ and production of $w$ requires $U \gtrsim \sigma_m M + V + k_w^2/R$. Substituting for $M$ and $V$ from (66) into these inequalities gives, respectively,

$$(1 + \sigma_m w^2 R)w^2R^2 - w^2R \gtrsim (1 + w^4R^4)\, U, \tag{67}$$

and

$$(1 + \sigma_m w^2 R^2)\, U \gtrsim \sigma_m + \sigma_m^2 w^2 R + w^2 R + \frac{1}{R}. \tag{68}$$

Eliminating $U$ between those two inequalities gives

$$(1 + \sigma_m w^2 R^2)\left((1 + \sigma_m w^2 R)w^2R^2 - w^2R\right) \gtrsim (1 + w^4R^4)\left(\sigma_m + \sigma_m^2 w^2 R + w^2 R + \frac{1}{R}\right), \tag{69}$$

yielding

$$R\,X \;\gtrsim\; \sigma_m R + 1 + (\sigma_m^2 + 1)X + X^2 + X^3 \tag{70}$$

where $X \equiv w^2R^2$. This is related to the fixed point equation in [33, eqn. (21)] but all constants have been set to 1, except for $\sigma_m$ to help keep track of cancellations that occurred in deriving this inequality. The 1 has been kept on the right hand side of (70) so that the inequality reduces to (59) when $\sigma_m = 0$. The bounds on $w$ follow from investigating the two extremes $X \ll 1$ and $X \gg 1$. For $X \ll 1$ and $R \gg 1$, inequality (70) reduces to $\sigma_m \lesssim X$ so the smallest $X \sim 1$, not $X \ll 1$. For $X \gg 1$ the inequality reduces to $X^3 \lesssim R\,X$ and since $X \equiv w^2R^2$ these two limits provide the bounds on $w$

$$R^{-1} \lesssim w \lesssim R^{-3/4}, \tag{71}$$

suggesting that $a = -1$ in (9). The transition scenario would be

$$w \sim R^{-1} \Longrightarrow V \sim R^{-1} \longrightarrow U \sim M < 1 \;\rightsquigarrow w \tag{72}$$

or

$$V \sim R^{-1} \longrightarrow U \sim M < 1 \;\rightsquigarrow w \sim R^{-1} \Longrightarrow V \sim R^{-1} \tag{73}$$

with $w$ quickly growing to $w \sim R^{-1}$ from the instability of $U$ in the latter case. In either case, this gives $\epsilon \sim R^{-1}$ for the transition threshold.

Again these (72), (73) are the scalings of the lower branch steady solution [33, eqn. (23)] for model (64), while the upper limit of (71) yields the scaling of the upper branch steady solution [32, 33], that is $w \sim R^{-3/4}$, $V \sim R^{-1/2}$, $U \sim R^{-1/2}$, $M \sim R^{-1}$.

### 3.8 Summary of transition models and threshold scalings

The transient growth models (35), (38), (41), (43), (45) show transition thresholds ranging from $\epsilon \sim R^{-3}$ to $\epsilon \sim R^{-1}$, depending on which *nonlinear* interactions do or do not actually occur. The smallest thresholds and most negative exponents correspond to models that *assume* the 'nonlinear recycling' of transiently amplified disturbances into 'optimal' disturbances, but this has not been explicitly demonstrated in the Navier-Stokes equations. On the contrary, long ago, Waleffe, Kim & Hamilton [37] presented evidence suggesting that such nonlinear interactions are nil to negligible. Yet Chapman's later work [8] still assumes but does not demonstrate the predominance of those same nonlinear interactions of transiently amplified disturbances. His threshold scaling predictions are thus wanting.

Those transient growth models (35), (38), (41), (45) are all *weakly nonlinear* in the sense that the mean shear stays at its laminar value of 1 and the perturbations essentially consist of interacting eigenmodes of the linearized Navier-Stokes equations, *i.e.* eigenmodes of the Orr-Sommerfeld and Squire equations. Indeed, much of Chapman's analysis [8, §5] centers on estimating the *eigenvalues* of these linear operators in spite of the emphasis on non-normal algebraic growth. The models (41), (45) are superficially similar to the SSP model (64), they are 4th order with quadratic nonlinearity, and they *assume* an interaction $Uu \equiv \phi_1\phi_2$ in the $v$ equation that superficially appears to be a 'streak instability' when coupled with the $v$ term in the $u$ equation (sect. 3.4). However, the models (41), (45) are not actually models of transition since the mean shear remains at its laminar value 1 and the $Uu$ term would be the weak nonlinear interaction of Squire modes (eigenmodes of the vertical vorticity $\eta$ equation linearized about the laminar flow) instead of a real streak instability with its own eigenmodes.

So far, the only model that stands up to closer analysis of the Navier-Stokes nonlinearities is the SSP (sect. 3.7) with its $\epsilon \sim R^{-1}$ scaling. In the SSP model (64), the mean shear $M$ provides the energy source for $U$ through $V$ but tends to destroy $w$ that regenerates $V$. For transition in that model, the mean shear $M$ must be reduced from its laminar value of 1 and the streaks $U$ must be $O(1)$, in that sense, it is a fully nonlinear model as indicated by the $\epsilon \sim R^{-1}$ threshold scaling. In the SSP, $w$ corresponds to an eigenmode of a spanwise varying streaky flow $(U(y,z),0,0)$, not of the laminar flow [31, 33, 34, 36]. This strong nonlinearity is reflected in our model building that started from a linear system with a Jordan block-like structure in (18) but ended with a system (64) with multiple energy-conserving nonlinear redistribution terms where the Jordan-block structure is not anymore a relevant point of view, it is now all about nonlinear interactions, and determination of the threshold scaling is more involved because of the stronger nonlinearity. We have gone from a *non-normal linear* system to a *normal nonlinear* system [32].

Finally, a point often overlooked in the discussion of transition processes is that the very existence of a *threshold* is closely connected with the existence of *unstable nonlinear states*. This is already clear in the subcritical logistic equation (14) $\dot{u} = -u/R + u^2$ where the threshold $u = 1/R$ *is* an unstable steady state. In the SSP model (64) the threshold is not identical to the 'lower branch' steady state but the transition scenario (72) is closely connected to it. The existence of fully resolved 3D nonlinear lower branch steady states with the proper scaling in plane Couette flow (fig. 7) indicates that the threshold $\epsilon \sim R^{-1}$ is indeed relevant for transition in the full Navier-Stokes equations. Indeed, all aspects of the

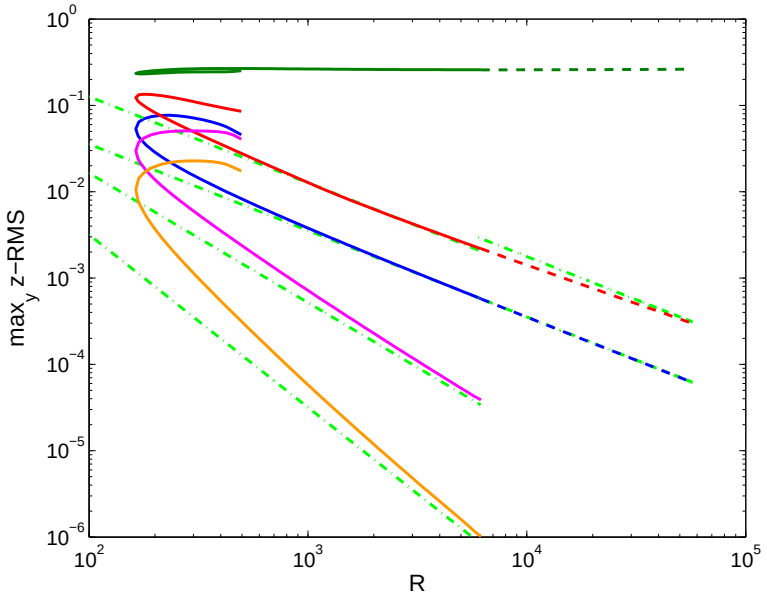


Figure 7: Scaling of 3D unstable lower branch steady solutions of the Navier-Stokes equations in plane Couette flow (see Wang, Gibson & Waleffe [38]). The streaks $U$ (top green curve) are $O(1)$, independent of $R$. The rolls $V$ scale like $R^{-1}$ (blue curve, 3rd from top), the fundamental wave $e^{i\alpha x}$ corresponding to $w$ in the SSP model (64) scales like $R^{-11/12}$ in this norm, slightly weaker than $R^{-1}$ because of a critical layer structure (red curve, 2nd from top). The bottom two curves (purple and orange) correspond to the 2nd and 3rd $x$ harmonics, $e^{i2\alpha x}$ and $e^{i3\alpha x}$, respectively. All harmonics are negligible for $R \gtrsim 6\,000$ and the solution is continued up to $R \approx 60\,000$ without them. The Reynolds number is based on the half channel height and the half wall velocity difference and this definition yields a Reynolds number that is about 4 times smaller than the typical pipe flow Reynolds number. In other words, this plane Couette $R = 60\,000$ corresponds to $Re \approx 240\,000$ in pipe flow.

SSP, (1) the creation of streaks together with mean shear reduction, (2) the streak instability and (3) the direct nonlinear feedback from that instability onto the $x$-independent rolls (often called 'streamwise rolls'), have been explicitly verified for the fully resolved Navier-Stokes equations [33, 34, 36]. The SSP model (64), albeit simplistic, faithfully captures the essence of a process that appears to be fundamental for transition *and* turbulence in shear flows and that has been fully vetted through the construction of fully resolved unstable nonlinear states in the Navier-Stokes equations, not just in low order models (figs. 7,8).

## 4 Transition threshold: experiments

Figure 9 is from Hof, Juel & Mullin (2003) [14] and shows a threshold amplitude scaling like $R^{-1}$ for transition to turbulence in a pipe. The disturbance consists of one pulse of fluid injected tangentially through 6 equispaced holes of 0.5mm in a long pipe (15.7m) with a 20mm diameter. The transition amplitude is independent of the duration $\Delta t$ of the injection

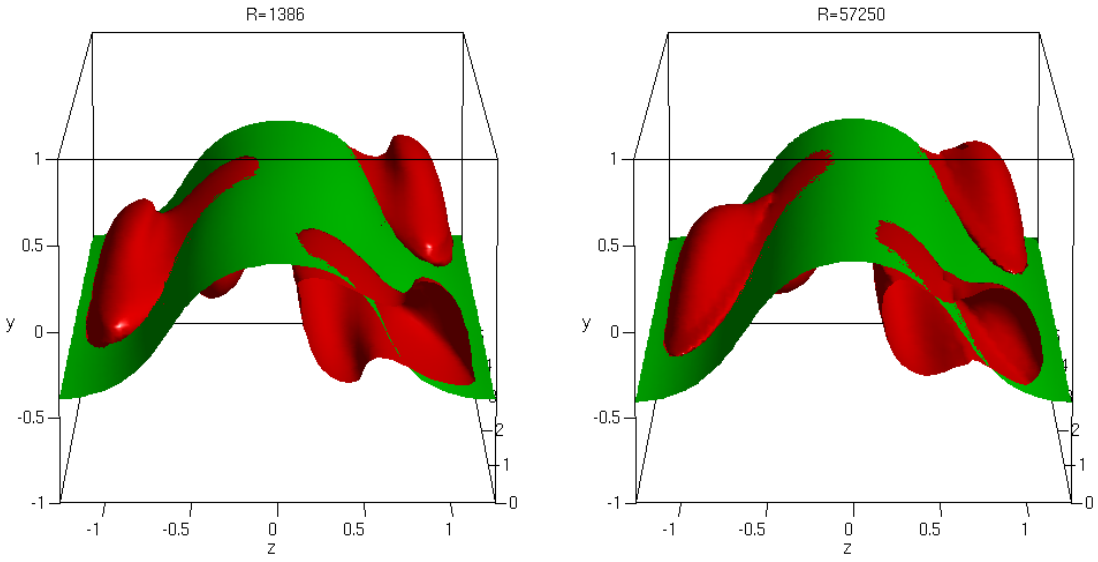


Figure 8: Visualization of the plane Couette flow lower branch steady state at $R = 1386$ (left) and $R = 57250$ (right) [38] whose scaling is shown in figure 7. The flow is in the $x$ direction (in and out of the page) and the total streamwise velocity $u = \hat{\mathbf{x}} \cdot \mathbf{v} = 1$ at $y = 1$ and $u = -1$ at $y = -1$. The green isosurface is $u = 0$ and would be a flat sheet at $y = 0$ for the laminar flow but here it is warped with an $O(1)$ deformation as a result of the $O(1)$ streaks $U$ in (64) and the $O(1)$ reduction of the mean shear $M$ in (64). The red isosurfaces are the level set $Q = 0.6 \max(Q)$ where $2Q = \nabla^2 p = \Omega_{ij}\Omega_{ij} - S_{ij}S_{ij}$ (see eqn. (4)), a standard but crude attempt to visualize vortices. These red isosurfaces correspond to a combination of the streamwise rolls and the fundamental wave, the 2nd and 3rd curves from the top in fig. 7 and have small magnitude. Note that the two figures are almost identical in spite of the huge difference in Reynolds number.

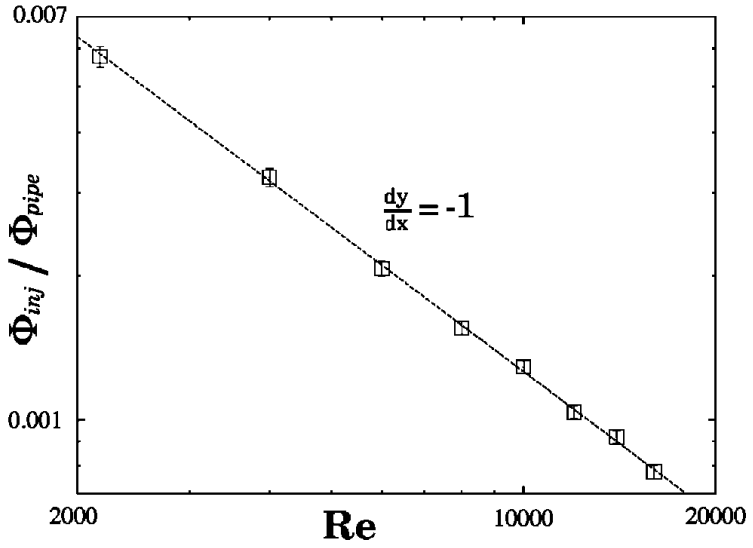


Figure 9: Transition threshold in pipe flow scaling as $R^{-1}$, from [14].

provided it is long enough, that is provided the length $\ell = U\Delta t$ is greater than about 3 diameters, where $U$ is the bulk velocity. The amplitude $\epsilon = \Phi_{inj}/\Phi_{pipe}$ is measured as the ratio of the disturbance mass flux to the pipe mass flux where $\Phi$ = velocity $\times$ area, and the Reynolds number is based on the bulk velocity and pipe diameter. The data is very well fitted by a power law $\epsilon \sim R^{-1}$. Other disturbances have since been used such as normal jets as well as fewer jets [21].

From the SSP point of view (Sect. 3.7), we may interpret these experiments as introducing streamwise[6] rolls $V$, sufficiently large in scale and amplitude and sufficiently long to develop streaks $U$ that can be unstable. The streak instability is inflectional in nature and requires sufficiently small streamwise wavenumbers $\alpha$. In practice we expect spanwise scales of the order of the pipe radius and streamwise scales about 2 to 3 times larger than that as discussed in SSP papers [31, 33, 36]. However the jets obviously introduce a range of scales from the 0.5mm holes to the pipe diameter 20mm so there is complex transient fluid behavior.

Figure 10 shows more recent data in pipe flow. The $Re^{-1}$ line on that plot is a fit to various data sets (not shown) corresponding to perturbations in the form of jets (one or more) through the pipe wall as in fig. (9). A second set of data is also shown for 'push-pull' disturbances, where fluid is injected and sucked from two neighboring holes with no net injected mass flux. That second set of data shows a $R^{-1.4}$ scaling. That scaling is associated with the development of trains of hairpin/horseshoe vortices before transition to turbulence, in contrast to the $R^{-1}$ disturbances that led to an abrupt transition. An example of the development of hairpin vortices is shown in figure 11. 'Hairpin' and 'horsehoe' are used to describe similar vortex structures. The term 'hairpin' is often used for small structures, especially those observed in turbulent shear flows, and 'horseshoe' for larger structures,

[6]'Streamwise rolls' have their axis in the streamwise direction, hence they are streamwise independent. Their wavenumber is actually spanwise.

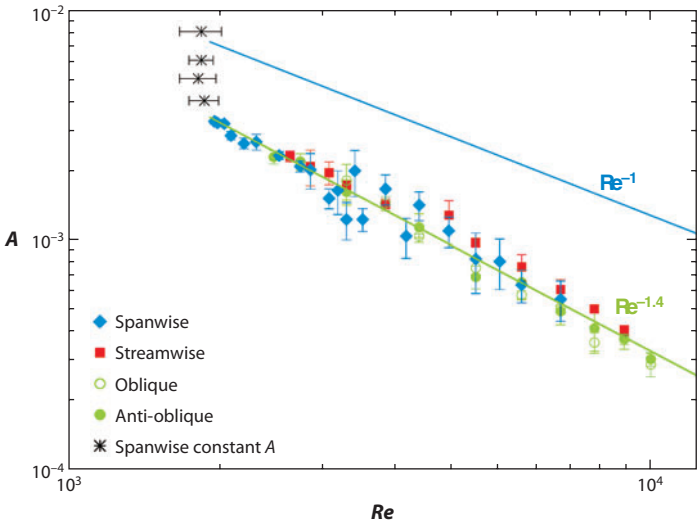


Figure 10: Transition threshold in pipe flow from [21, 22]. Here $Re$ is the Reynolds number based on bulk velocity and pipe diameter and $A$ is the threshold amplitude defined as the mass flux of the disturbance normalized by the mass flux in the pipe. The $Re^{-1}$ line is a fit to various data sets (not shown on this figure) corresponding to perturbations in the form of jets (one or more) through the pipe wall [22, Fig. 3]. The lower data sets correspond to smaller scale 'push-pull' disturbances, with the push-pull axis oriented in various ways (spanwise, streamwise, etc.), and that data is fitted by a $Re^{-1.4}$ line.

although this is qualitative and subjective, there is no precise definition of the structures associated with those terms.

The development of these hairpin structures is a transient effect but one should be careful not to quickly associate it with the linear transient growth of non-normal operators. The latter is associated in this context with the linearization of the Navier-Stokes equations about a laminar shear flow $\mathbf{v} = U(y)\hat{\mathbf{x}}$. The linearized equations are separable in the cross-flow coordinates $x$ and $z$ and time $t$. The result is a set of *dispersive* Fourier modes, $\hat{\mathbf{u}}(y)e^{i(\alpha x+\gamma z-\omega t)}$, that are not orthogonal, but do form a complete set. Each of these modes evolves and travels independently of the other modes and it is quite unlikely that linear evolution of such modes captures the highly coherent development of hairpin vortices shown in fig. 11. The development of hairpin vortices is undoubtedly a *nonlinear process* leading to such highly coherent structures formed from the roll-up and stretching of vortex sheets.

### 4.1 Theodorsen's horseshoes

Horseshoe vortices in turbulent shear flows were predicted in the 1950's by Theodorsen who suggested that they were the 'molecules' of [shear] turbulence [27, 28] and drew the sketch shown in figure 12. He based his prediction of horseshoe vortices on an analysis of the vorticity equation, the curl of the Navier-Stokes equations (3)

$$\partial_t \boldsymbol{\omega} + \mathbf{v} \cdot \nabla \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \nabla \mathbf{v} + \nu \nabla^2 \boldsymbol{\omega} \tag{74}$$

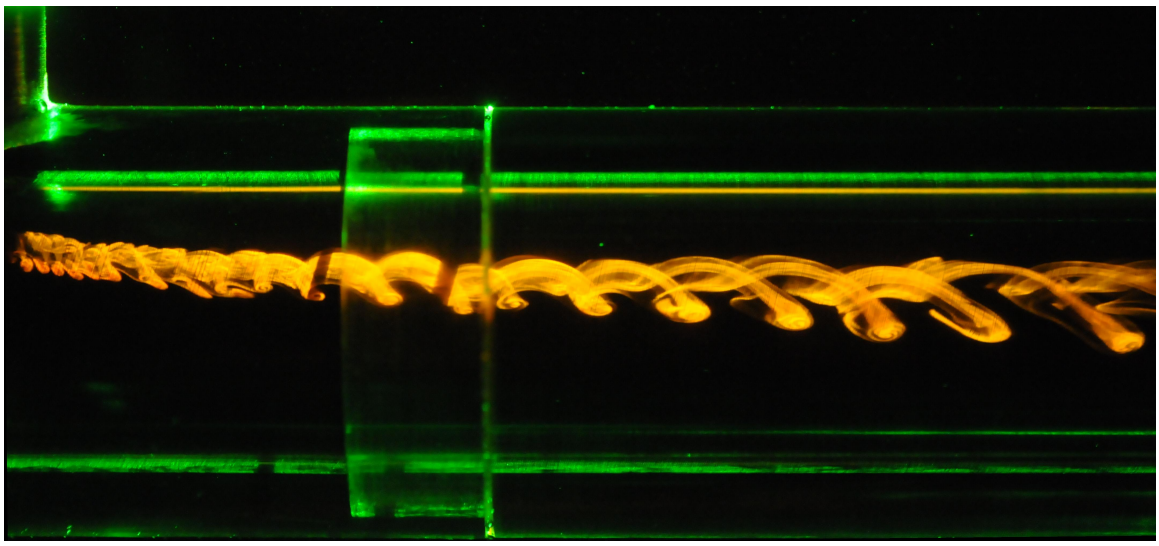

Figure 11: Hairpin vortices developing from a small jet in pipe flow (photograph by Finn Box in the Mullin Lab). This is the typical transient behavior of the disturbance leading to the $R^{-1.4}$ scaling in fig. 10.

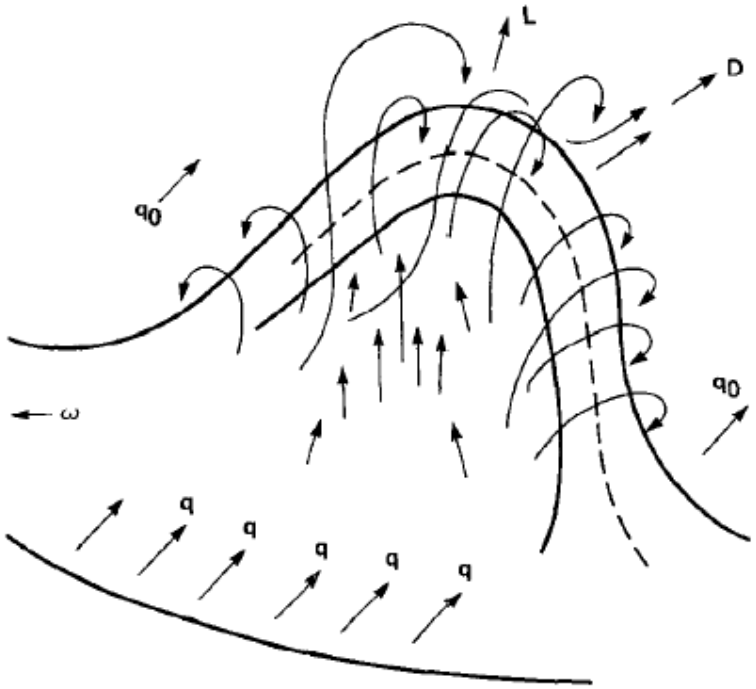


Figure 12: Hairpin/Horseshoe vortex postulated to be the 'molecule of shear turbulence' by Theodorsen in 1952. Here $\mathbf{q}$ is the fluid velocity and $L$ and $D$ are the lift and drag on the horseshoe structure [27].

where $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ is the curl of the velocity $\mathbf{v}$, thus $\nabla \cdot \boldsymbol{\omega} = 0$.

His picture, reproduced in fig. 12, is compelling and horseshoe structures have since been observed as ubiquitous features in turbulent shear flows, but his mathematical justification for those structures is mostly hand-waving and laced with not quite correct statements. Theodorsen considers the enstrophy equation, obtained by dotting (74) with $\boldsymbol{\omega}$,

$$\frac{D}{Dt}\frac{\omega^2}{2} = \boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v}) + \nu\, \boldsymbol{\omega} \cdot \nabla^2 \boldsymbol{\omega} \tag{75}$$

where $D/Dt = \partial_t + \mathbf{v} \cdot \nabla$ is the material derivative and $\omega^2 = \boldsymbol{\omega} \cdot \boldsymbol{\omega}$. He states that $\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v}) = \omega^2 dv_s/ds$ where $s$ is arclength along the local vortex line and $v_s = \mathbf{v} \cdot \boldsymbol{\omega}/\omega$ is the velocity component in the direction of the vortex line. This is not correct in general. If $\mathbf{s}$, $\mathbf{n}$, $\mathbf{b}$ represent an orthonormal Frenet-Serret frame along the vortex line, with $\mathbf{s} = \boldsymbol{\omega}/\omega$ and $\mathbf{v} = v_s\mathbf{s} + v_n\mathbf{n} + v_b\mathbf{b}$ then

$$\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v}) = \omega^2\, \mathbf{s} \cdot \frac{d}{ds}(v_s\mathbf{s} + v_n\mathbf{n} + v_b\mathbf{b}) = \omega^2 \left( \frac{dv_s}{ds} - \kappa v_n \right) \tag{76}$$

where $\kappa$ is the curvature of the vortex line. He argues that the viscous terms $\boldsymbol{\omega} \cdot \nabla^2 \boldsymbol{\omega}$ is negative and thus that $\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v})$ is positive 'on average' for stationary turbulent flows. The viscous term

$$\boldsymbol{\omega} \cdot \nabla^2 \boldsymbol{\omega} = \nabla^2 \frac{\omega^2}{2} - \nabla \boldsymbol{\omega} : \nabla \boldsymbol{\omega}^T \tag{77}$$

and the 2nd term is negative since $\nabla \boldsymbol{\omega} : \nabla \boldsymbol{\omega}^T = (\partial_j \omega_i)(\partial_j \omega_i) \geq 0$, but the first term can be positive. These considerations do not explain the horseshoe structure however.

It appears that the basic reason for Theodorsen's horseshoe proposal is to maximize enstrophy production *by the mean shear*, that is to maximize $\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v})$ for $\mathbf{v} = U(y)\hat{\mathbf{x}}$ in which case $\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v}) = \omega_x \omega_y\, dU/dy$ and the rate of enstrophy production, that is $\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \cdot \nabla \mathbf{v})\, /\omega^2 = (\omega_x \omega_y\, dU/dy)\, /(\omega_x^2 + \omega_y^2 + \omega_z^2)$ is maximized for $\omega_x = \omega_y$ with $\omega_z = 0$. There are two constraints on this optimum. First, since $\nabla \cdot \boldsymbol{\omega} = 0$, vortex lines cannot terminate at a point in the fluid. Second, very close to the wall, the vorticity must be $\boldsymbol{\omega} \approx \omega_z \hat{\mathbf{z}}$. Putting all these things together: (1) vortex lines starting in the spanwise $\hat{\mathbf{z}}$ direction near the wall, (2) turning in the 45° direction $\omega_x \approx \omega_y > 0$, $\omega_z \approx 0$ to maximize stretching by the mean shear, (3) turning back when $dU/dy \approx 0$ so that $\omega_x \approx \omega_y < 0$ again maximizes stretching and (4) smoothly connects back to the $\hat{\mathbf{z}}$ direction near the wall, leads to the very plausible horseshoe structure sketched in fig. 12. It seems clear from Theordorsen's depiction of *lift* and *drag* on the horseshoe — forces that are quadratic in the relative velocity — that this is implicitly a *coherent nonlinear structure*, not the linear dispersion of laminar flow eigenmodes.

### 4.2 An argument for $R^{-1.4}$ ?

Theodorsen's reasons for the horseshoe/hairpin vortex thus appear to be *quasi-linear*. The hairpin vortex must be nonlinear to hold itself together in a coherent packet, but the basic reason for the shape of the vortex appears to be merely stretching by the mean. If $\ell$ is the scale and $v$ the velocity amplitude of the initial perturbation (*i.e.* the small jet or push-pull disturbance in the pipe experiments), we require that the perturbation's Reynolds number

be of $O(1)$ for that perturbation to 'roll-up' and create a small coherent vortical structure, that is

$$\frac{v\ell}{\nu} \gtrsim O(1). \tag{78}$$

Now for 'optimal' vortex stretching by the mean shear, we require that $\omega dU/dy \gtrsim \nu\nabla^2\omega$, that is

$$\frac{U}{h} \gtrsim \frac{\nu}{\ell^2} \iff \frac{\ell^2}{h^2} \gtrsim \frac{\nu}{Uh} = \frac{1}{R}. \tag{79}$$

where $\ell$ is an estimate for the vortex core and $h$ is the pipe radius or the channel half-height. The meaning of this equation is that 'optimal' (*i.e.* 45° orientation) stretching of the vortex by the mean shear must at least balance the diffusion on the scale of the vortex core.

In the pipe experiments, the threshold amplitude is taken as the ratio of the perturbation mass flux $\approx v\ell^2$, to the pipe flux $\approx Uh^2$. Combining (78) and (79) yields

$$\frac{v\ell^2}{Uh^2} \gtrsim \frac{\nu}{Uh}\frac{\ell}{h} \gtrsim R^{-3/2}, \tag{80}$$

which is not $R^{-1.4}$ but close to it. Note that this is not an argument for transition but a quasi-linear argument for creation and amplification of hairpin vortices.

One issue with this argument is that $\ell$ is taken as both the size of the vortex core and of the initial perturbation. In the pipe experiments with $R^{-1.4}$ scaling [22], the scale of the push-pull disturbance is fixed at 1mm (two 1mm holes, 1mm apart) and the pipe radius is 20mm so $\ell/h \approx 0.05$ which seems large enough to satisfy (79), but these are mere scaling arguments and there may be significant hidden constant factors. For single jet disturbances with $R^{-1}$ scaling, Peixinho and Mullin [22, Fig. 2(b)] show a scaling $vd/\nu \gtrsim (D/d)^{1/2}$ for the *jet* Reynolds number at fixed $R = 2500$, where $d$ and $D$ are the hole and pipe diameters, respectively. This shows that small holes require stronger jets, perhaps to lead to vortical structures with a size $\ell$ large enough to satisfy (78), (79). This suggests that the steeper scaling $R^{-1.4}$ instead of $R^{-1}$ might be a small hole/low Reynolds number effect, although the Peixinho & Mullin $R^{-1.4}$ data is provided for a significant Reynolds number range.

Recent experimental studies of transition in plane Poiseuille flow by Lemoult, Aider and Wesfreid (2012) [19] show transition induced by a continuous small jet with $v/U \sim R^{-1}$ and a steeper scaling closer to $R^{-3/2}$ for low Reynolds number. The latter experiments document the development and persistence of hairpin vortices, especially in transitional states not laminar but not fully transitioned either. Thus the role of hairpin vortices and the threshold scaling remain to be clarified. In particular, small jets in cross-flows introduce multi-scale perturbations (the diameter of the jet and the thickness of the shear layer around that jet) that lead to the development of multiple vortex structures, such as a large scale counter-rotating vortex pair as well as smaller horseshoe (or necklace) vortices, possibly coupled to 'upright vortices' ([18] but the flow regimes are quite different from fig. 11). It may be that the visualizations draw our attention to the smaller scale trains of quasi-linear hairpin vortices but that it is really the larger scale counterrotating vortex pair that redistributes the mean to create streaks on the scale of the pipe radius and trigger the turbulent transition.

## References


[1] M. Acarlar and C. Smith, *A study of hairpin vortices in a laminar boundary layer*, J. Fluid Mech., 175 (1987), pp. 1–41 and 45–83.

[2] D. Acheson, *Elementary Fluid Dynamics*, Oxford University Press, 1990.

[3] J. S. Baggett and L. N. Trefethen, *Low-dimensional models of subcritical transition to turbulence*, Physics of Fluids, 9 (1997), pp. 1043–1053.

[4] G. Batchelor, *An Introduction to Fluid Dynamics*, Cambridge Mathematical Library, Cambridge University Press, 2000.

[5] D. Benney, *The evolution of disturbances in shear flows at high Reynolds numbers*, Stud. Applied Math., 70 (1984), pp. 1–19.

[6] D. J. Benney and L. H. Gustavsson, *A new mechanism for linear and nonlinear hydrodynamic instability*, Stud. Applied Math., 64 (1981), pp. 185–209.

[7] D. J. Benney and C. C. Lin, *On the secondary motion induced by oscillations in a shear flow*, Physics of Fluids, 3 (1960), pp. 656–657.

[8] S. J. Chapman, *Subcritical transition in channel flows*, Journal of Fluid Mechanics, 451 (2002), pp. 35–97.

[9] P. Drazin and W. Reid, *Hydrodynamic Stability*, Cambridge University Press, Cambridge, UK, 1981.

[10] T. Ellingsen and E. Palm, *Stability of linear flow*, Physics of Fluids, 18 (1975), pp. 487–488.

[11] L. H. Gustavsson, *Energy growth of three-dimensional disturbances in plane Poiseuille flow*, Journal of Fluid Mechanics, 224 (1991), pp. 241–260.

[12] L. H. Gustavsson and L. S. Hultgren, *A resonance mechanism in plane Couette flow*, Journal of Fluid Mechanics, 98 (1980), pp. 149–159.

[13] J. Hamilton, J. Kim, and F. Waleffe, *Regeneration mechanisms of near-wall turbulence structures*, J. Fluid Mech., 287 (1995), pp. 317–348.

[14] B. Hof, A. Juel, and T. Mullin, *Scaling of the turbulence transition threshold in a pipe*, Phys. Rev. Lett., 91 (2003), p. 244502.

[15] L. S. Hultgren and L. H. Gustavsson, *Algebraic growth of disturbances in a laminar boundary layer*, Physics of Fluids, 24 (1981), pp. 1000–1004.

[16] D. D. Joseph and W. Hung, *Contributions to the nonlinear theory of stability of viscous flow in pipes and between rotating cylinders*, Archive for Rational Mechanics and Analysis, 44 (1971), pp. 1–22. 10.1007/BF00250825.

[17] D. D. Joseph and L. N. Tao, *Transverse Velocity Components in Fully Developed Unsteady Flows*, Journal of Applied Mechanics, 30 (1963), pp. 147–148.

[18] R. M. Kelso, T. T. Lim, and A. E. Perry, *An experimental study of round jets in cross-flow*, Journal of Fluid Mechanics, 306 (1996), pp. 111–144.

[19] G. Lemoult, J.-L. Aider, and J. E. Wesfreid, *Experimental scaling law for the subcritical transition to turbulence in plane Poiseuille flow*, Phys. Rev. E, 85 (2012), p. 025303.

[20] A. Meseguer and L. N. Trefethen, *Linearized pipe flow to Reynolds number* $10^7$, J. Comput. Phys., 186 (2003), pp. 178–197.

[21] T. Mullin, *Experimental studies of transition to turbulence in a pipe*, Annual Review of Fluid Mechanics, 43 (2011), pp. 1–24.

[22] J. Peixinho and T. Mullin, *Finite-amplitude thresholds for transition in pipe flow*, Journal of Fluid Mechanics, 582 (2007), pp. 169–178.

[23] D. Rempfer, *On boundary conditions for incompressible Navier-Stokes problems*, Applied Mechanics Reviews, 59 (2006), pp. 107–125.

[24] 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.*, Proceedings of the Royal Society of London, 35 (1883), pp. 84–99.

[25] V. Romanov, *Stability of plane-parallel Couette flow*, Functional Anal. & its Applic., 7 (1973), pp. 137–146.

[26] P. Schmid and D. Henningson, *Stability and Transition in Shear Flows*, no. v. 142 in Applied Mathematical Sciences, Springer, 2001.

[27] T. Theodorsen, *Mechanism of turbulence*, in Proceedings of 2nd Midwestern Conference on Fluid Mechanics, Ohio State University, 1952, pp. 1–18.

[28] ——, *The structure of turbulence*, in Technical Note, vol. 31, University of Maryland, The Institute for Fluid Dynamics and Applied Mathematics, 1954, pp. 21–27.

[29] N. Trefethen, A. Trefethen, S. Reddy, and T. Driscoll, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.

[30] F. Waleffe, *Proposal for a self-sustaining process in shear flows*, Unpublished Center for Turbulence Research manuscript, available at `www.math.wisc.edu/~waleffe/ECS/sspctr90.pdf`, (1990).

[31] ——, *Hydrodynamic stability and turbulence: Beyond transients to a self-sustaining process*, Stud. Applied Math., 95 (1995), pp. 319–343.

[32] ——, *Transition in shear flows: Nonlinear normality versus non-normal linearity*, Phys. Fluids, 7 (1995), pp. 3060–3066.

[33] ——, *On a self-sustaining process in shear flows*, Phys. Fluids, 9 (1997), pp. 883–900.

[34] ——, *Three-dimensional coherent states in plane shear flows*, Phys. Rev. Lett., 81 (1998), pp. 4140–4148.

[35] ——, *Exact coherent structures in channel flow*, J. Fluid Mech., 435 (2001), pp. 93–102.

[36] ——, *Homotopy of exact coherent structures in plane shear flows*, Phys. Fluids, 15 (2003), pp. 1517–1543.

[37] F. Waleffe, J. Kim, and J. Hamilton, *On the origin of streaks in turbulent shear flows*, in Turbulent Shear Flows 8: selected papers from the Eighth International Symposium on Turbulent Shear Flows, Munich, Germany, Sept. 9-11, 1991, F. Durst, R. Friedrich, B. Launder, F. Schmidt, U. Schumann, and J. Whitelaw, eds., Springer-Verlag, Berlin, 1993, pp. 37–49.

[38] J. Wang, J. Gibson, and F. Waleffe, *Lower branch coherent states in shear flows: transition and control*, Phys. Rev. Lett., 98 (2007), p. 204501.