# Hydrodynamic Stability and Turbulence: Beyond Transients to a Self-Sustaining Process

By Fabian Waleffe

Transition from laminar to turbulent flows has generally been studied by considering the linear and weakly non-linear evolution of small disturbances to the laminar flow. That approach has been fruitless for many shear flows and a last hope for its success has been the existence of transient growth phenomena. The physical origin of those linear transient effects is elucidated, revealing serious limitations both of previous analyses and of the phenomena themselves that preclude them from causing direct transition. Nonetheless, some transient effects are symptomatic of one element of a non-linear process that becomes self-sustaining at small enough dissipation. The process is identified and its description requires a departure from the traditional focus on the laminar flow. A theory is outlined in which the mean flow has an intrinsic spanwise variation. Evidence indicates this is also the central mechanism in the near wall-region of fully turbulent shear flows.

# 1. Introduction

The breakdown of fluid flows from a laminar to a turbulent state is a matter of everyday experience, yet our understanding of the processes involved remains far from satisfactory. The problem has two aspects, which have usually been addressed quite separately. One is the identification of the instability mechanisms responsible for the transition; the other is the development of a quantitative description of the turbulent state. The present work is related to both aspects but its connection to the problem of transition is the clearest. That connection is the main focus of this article.

The first step in a stability study is to add small enough perturbations to the laminar flow that their evolution is governed by linear equations. Instability of the laminar flow is established if there are exponentially growing disturbances. That approach is very helpful for many problems but has been unable to explain the observed instabilities in most basic shear flows such as plane Couette flow considered below [1]. A possible solution to that paradox was seen in the existence of weaker, and transient, algebraic growth phenomena, the most powerful of which became the object of systematic investigations in the late 70's [2-6]. Of course, the transients do not by themselves establish the instability of the flow. It is necessary to show that those transient amplifications can trigger nonlinear effects, which somehow prevent the eventual decay of the disturbances. The only work in that area is due to Benney and Gustavsson in 1981 [7], but their theory was later criticized and ran into difficulties [8]. Some of those difficulties have appeared in earlier non-linear theories involving 3D disturbances and are precisely due to the linear transients [9]. Sooner or later in the non-linear analysis a transient growth arises and invalidates the expansion. Benney [10] then suggested it was necessary to account directly for the large spanwise variation invariably introduced by the transient

Address: MIT, Department of Mathematics, Room 2-378, Cambridge, MA 02139

effects. That idea was a major inspiration for the non-linear process proposed herein. The other main inspiration was the conceptual pictures developed by experimentalists [11].

The linear transient phenomena have recently attracted intense renewed interest [12-15]. Although the mathematical language varies greatly from one paper to the next, they share a common underlying physical mechanism that is clarified later below. That discussion shows how linear analyses are misleading and seriously limited in range of validity. The most significant result is that the transient growths do not trigger any non-linear effects. A different point of view is thus needed, and although the many computations of maximum amplification are largely irrelevant, the transient amplification of spanwise perturbations to the mean flow indicates which direction to take.

The direction presented here is based on the identification of a non-linear process that becomes self-sustaining at a critical Reynolds number that seems to match the observations. Bifurcation of the flow results not from the instability of the laminar flow but rather from the the fact that this non-linear process acts as an "attractor" for most finite perturbations. Perhaps the scale of the process remains approximately constant as a function of the Reynolds number when nondimensionalized with respect to the shear rate and the viscosity. At high Reynolds numbers, the process would be localized near the wall, in the high shear region, and responsible for the observed streaks with the characteristic spacing of 100 wall units. These ideas have been tested against a series of numerical experiments [16]. This article outlines a theoretical description and a conceptual model of the self-sustaining mechanism, but first the linear transients are reviewed and interpreted.

## 2. Linear stability and transient growth

The basic flow considered is the simplest example of a shear flow and one of the simplest solutions of the incompressible Navier-Stokes equations:

$$\frac{\partial}{\partial t}\mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \frac{1}{R}\nabla^2 \mathbf{u},\tag{1}$$

where  $\mathbf{u}$  is the solenoidal velocity field  $\nabla \cdot \mathbf{u} = 0$ . The field  $\mathbf{u}$  has components (u, v, w) with respect to a Cartesian frame of reference (x, y, z). The fluid is bounded by two rigid parallel walls chosen perpendicular to the y-axis and located at  $y = \pm h$ . The flow is driven by the motion of the walls in their plane at velocity  $(\pm U_w, 0, 0)$ , respectively. The velocities are non-dimensionalized by  $U_w$  and lengths by h. The Reynolds number  $R = U_w h/\nu$  where  $\nu$  is the kinematic viscosity of the fluid. This is plane Couette flow and it admits the laminar solution:  $\mathbf{u} = (y, 0, 0)$ .

The laminar solution is observed experimentally only for Reynolds numbers R less than some critical value (about 350 for Couette flow) [16,17,18]. In order to study that instability of the laminar flow, one writes the full velocity field  $\mathbf{u}$  as the sum of the laminar solution (U(y), 0, 0) (with U(y) = y for plane Couette flow) plus a perturbation (u, v, w),

$$\mathbf{u} = (U(y), 0, 0) + (u, v, w). \tag{2}$$

Evolution equations for the perturbations (u, v, w) are deduced by substitution of (2) in the Navier-Stokes equations (1). The system of equations can be reduced to two coupled equations for the velocity and vorticity perpendicular to the walls, v and  $\eta$  respectively,

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} - \frac{1}{R} \nabla^2\right) \nabla^2 v - \frac{d^2 U}{dv^2} \frac{\partial v}{\partial x} = N_v(v, \eta) \tag{3}$$

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} - \frac{1}{R} \nabla^2\right) \eta + \frac{dU}{dy} \frac{\partial v}{\partial z} = N_{\eta}(v, \eta), \tag{4}$$

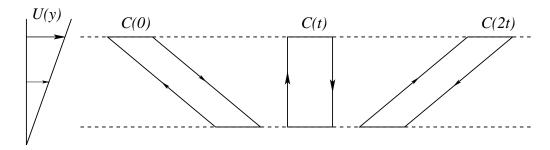
where  $\eta = \partial u/\partial z - \partial w/\partial x$ . Explicit expressions for the non-linear terms  $N_v(v, \eta)$ ,  $N_{\eta}(v, \eta)$  can be found in [7]. The procedure for arriving at these equations is standard and consists of taking

the y-component of the curl  $(\mathbf{j} \cdot \nabla \times (\mathbf{1}))$  and of the curl of the curl  $(\mathbf{j} \cdot (\nabla \times (\nabla \times (\mathbf{1}))))$  of the Navier-Stokes equations (1). The system (3), (4) must be supplemented by an equation for the average of the perturbation over x and z,  $\bar{u}(y,t)$ , which can not be kinematically determined from v and  $\eta$ ,

$$\left(\frac{\partial}{\partial t} - \frac{1}{R} \frac{\partial^2}{\partial y^2}\right) \bar{u} = -\frac{\partial}{\partial y} \overline{u} \overline{v} . \tag{5}$$

In the classical linear stability analysis, one considers only "small" perturbations to the laminar state and proceeds to ignore the non-linear terms on the right-hand side of (3), (4). The v-equation stands alone then and the focus of much work on stability theory has been the study of the eigenmodes of that equation (the famous Orr-Sommerfeld (OS) eigen-problem [1]). The homogeneous  $\eta$ -equation (i.e. with v=0) has only damped modes. This can be proven rigorously [7] but is also an intuitive result given that the homogeneous  $\eta$ -equation describes a simple advection-diffusion with  $\eta$  vanishing at the walls. For Couette flow, all eigenmodes of the v-equation are damped and this result appears inconsistent with the experimental observation of instability. Two suggestions have been proposed to resolve that conflict in the context of linear theory.

First, there is a possibility for transient algebraic growth in the v-equation due to the Orr mechanism [19, 20]. The physical mechanism is simple (Fig. 1) but the mathematical description has been involved (Continuous spectrum, non-orthogonality of eigenfunctions, pseudo-spectrum). It is only a transient effect even in the inviscid limit. At transitional Reynolds numbers ( $R \simeq 350$ ), the maximum amplification is quite modest [21]. This type of transient amplification is more pronounced for plane disturbances (for which  $\partial/\partial z = 0$  and w = 0) as can be gleaned from fig. 1.



**Fig. 1.** The Orr mechanism: Circulation around contour C is conserved (Kelvin's theorem). The velocity perturbation reaches a maximum when the contour is smallest then decays after that.

The second mechanism for transient algebraic growth is much more powerful and in the following "transient growth" will always refer to this second mechanism. It results from the forcing of  $\eta$  by v in (4) which induces a transient growth of  $\eta$  even if both modes are damped. The mathematical mechanism can be illustrated by the system

$$\frac{d}{dt} \begin{pmatrix} v \\ \eta \end{pmatrix} = \begin{pmatrix} -\lambda & 0 \\ 1 & -\mu \end{pmatrix} \begin{pmatrix} v \\ \eta \end{pmatrix}, \tag{6}$$

which has the solution

$$v(t) = v(0) e^{-\lambda t}$$
  

$$\eta(t) = -v(0) \frac{e^{-\lambda t} - e^{-\mu t}}{\lambda - \mu} + \eta(0) e^{-\mu t}.$$

The decay rates  $\lambda$ ,  $\mu > 0$  are proportional to 1/R. The first term on the right-hand side of  $\eta(t)$  grows algebraically like t at small times and decays exponentially at large times. The maximum

amplification is obtained at  $t^* = (\ln \lambda - \ln \mu)/(\lambda - \mu)$  and is given by  $\eta_{max} = (v(0)e^{-\mu t^*})/\lambda$ . In terms of the Reynolds number the maximum amplification is  $O(\epsilon R)$  where  $\epsilon$  is a measure of v(0) and the maximum is reached at a time t = O(R).

Gustavsson [12] studied the initial value problems for  $\eta$  forced by eigenmodes of the v equation. Butler and Farrell [13] formulated a variational problem to determine the structure of the initial perturbation v(0) leading to the largest amplification. Trefethen *et al.* [15] studied the problem using the concept of pseudospectrum. They computed curves of "pseudo-resonance" and studied the transient growth through the norm of the operator  $\exp(\mathbf{A}t)$  where  $\mathbf{A}$  is the matrix on the right-hand side of (6). The solution of (6) can be written  $(v(t), \eta(t))^T = \exp(\mathbf{A}t)(v(0), \eta(0))^T$ .

Originally it was thought that the amplification was largest when  $\lambda = \mu$ , hence the term "direct resonance," but closer inspection reveals that the amplification is ubiquitous and largest for small damping rates not necessarily equal. For the hydrodynamic problem, the growth is most significant when the damping rates are small and the imaginary part of the eigenvalues are close to each other. The result of Gustavsson [12] is that modes corresponding to streamwise-independent  $(\partial/\partial x = 0)$  disturbances lead to the largest transient amplifications. For x-independent modes, the eigenvalues of the v and  $\eta$  equations are purely real and negative, corresponding to weak viscous decay (in fact both sets of eigenmodes are orthogonal and analytically accessible, see [1], p. 159). That streamwise-independent disturbances are the most "efficient" is an earlier, and more general, result of energy stability and upper bound theories [22].

#### 3. Interpretation and limitations

The physical mechanism for transient growth and its limitations are best understood in its most favorable form, which is when the disturbances are streamwise-independent. In that case the Navier-Stokes equations with  $\partial/\partial x=0$  show that the spanwise motions v,w are independent of the streamwise velocity u

$$\frac{\partial}{\partial t}v + v\frac{\partial}{\partial y}v + w\frac{\partial}{\partial z}v = -\frac{\partial}{\partial y}p + \frac{1}{R}\nabla^2 v$$

$$\frac{\partial}{\partial t}w + v\frac{\partial}{\partial y}w + w\frac{\partial}{\partial z}w = -\frac{\partial}{\partial z}p + \frac{1}{R}\nabla^2 w$$

$$\frac{\partial}{\partial y}v + \frac{\partial}{\partial z}w = 0$$
(7)

and as a result the equation for the streamwise velocity u is linear:

$$\frac{\partial}{\partial t}u + v\frac{\partial}{\partial u}u + w\frac{\partial}{\partial z}u = \frac{1}{R}\nabla^2 u \tag{8}$$

Although one often reads the words "algebraic instability" to describe that passive advection, the word "instability" is something of a misnomer (Fig. 2). The confusion arises from the "overlinearization" around the laminar state (U(y), 0, 0) described in the previous section. Equation (8)

is already linear but the usual linear analyses (including the most recent work based on optimum perturbations and pseudo-spectra) further "linearize" the equation by writing u = U(y) + u'(y, z, t), v = v'(y, z), w = w'(y, z). Substituting in (8), dropping squares of primed quantities and neglecting viscosity yields

$$\frac{\partial}{\partial t}u' = -v'\frac{dU}{dy} \tag{9}$$

while the v, w motion is steady in the absence of viscosity. This equation shows that u' grows linearly with time u' = -v' (dU/dy) t. Equation (9) suggests that we are witnessing an algebraic instability on a time scale given by the mean shear dU/dy with the initial amplitude of the disturbance given by v. Those are the expected scalings for an instability of the mean shear. (The time scale for the Orr mechanism, Fig. 1, is clearly given by dU/dy also.) In fact, it is the opposite in this case. The time scale of the process is given by the advection time scale  $v/h = O(\epsilon)$ , while the maximum amplitude of the perturbation is given by dU/dy. The time scale for streak formation is the time scale of the downstream roll, not dU/dy.

Clearly, the linear analyses are misleading, and the numerous computations of maximum amplifications are largely irrelevant. The maximum amplitude attainable through shear-tilting is not  $O(\epsilon R)$  as suggested by the transient growth analyses (6), but  $O(U_{max} - U_{min}) = O(1)$  for u, or  $O(\gamma (dU/dy)_{max})$  for  $\eta$  with  $\gamma$  the spanwise wavenumber of the disturbance, and the time scale to reach this maximum is  $O(h/v) = O(1/\epsilon)$  and not O(R), where  $\epsilon$  is a measure of the amplitude of v in the present non-dimensionalization based on  $U_w$  and h. This maximum is realizable; it suffices to choose the amplitude of the downstream roll (v) large enough. "Large enough" is determined from the requirement that the dissipation time scale of the rolls be longer than the advection time scale,  $h^2/(\pi^2\nu) \ge h/v$  or  $\epsilon R \ge \pi^2$ , where the factor of  $\pi^2$  is introduced to estimate the effect of the boundary conditions  $v = \partial v/\partial y = 0$  at the walls. The structure of the maximum is also not well-determined from the "over-linear" analyses that neglect the  $w \frac{\partial u}{\partial z}$  term  $(O(\epsilon^2 R))$  and the advection of the modified mean shear  $v \partial u/\partial y$   $(O(\epsilon^3 R^2))$  in Eq.(8). Those terms are far from negligible near the real maximum (Fig. 3). The classical transient growth analysis is only valid when those terms can be neglected with respect to  $v \ dU/dy = O(\epsilon)$  which is the case if  $\epsilon R \ll 1$ . This makes it irrelevant because a larger perturbation would be able to create a more significant redistribution of streamwise momentum.

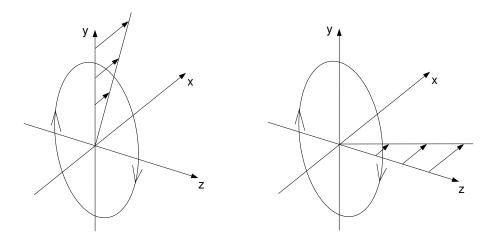


Fig. 2. The shear-tilting mechanism and the exact solution (10)

These points and the mechanism can be illustrated further through the following simple solu-

tion of the Navier-Stokes equations. Consider the flow

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 & -c(t) & b(t) \\ 0 & 0 & -a(t) \\ 0 & a(t) & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
(10)

(One should be cautioned that this flow does not satisfy realistic viscous boundary conditions. However, it satisfies inviscid boundary conditions for flow down a pipe whose axis is in the x direction.) The flow (10) has uniform time-dependent vorticity  $\boldsymbol{\omega} = \nabla \times \mathbf{u} = (2 \ a(t), b(t), c(t))$ . The v and w motions are that of a uniform solid body rotation around the x-axis, while the u component is the superposition of two uniform shear flows. Substituting this flow in the full vorticity equation  $D_t \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \nabla \mathbf{u} + \nu \nabla^2 \boldsymbol{\omega}$  yields

$$\dot{a} = 0 
\dot{b} = -a c 
\dot{c} = a b$$
(11)

If the initial conditions correspond to a Couette flow in the y-direction b(0) = 0,  $c(0) = c_0$ , with  $dU/dy = -c_0$ , on which a small downstream roll  $a(0) = \epsilon$  is introduced, the solution is

$$a(t) = \epsilon$$

$$b(t) = -c_0 \sin(\epsilon t)$$

$$c(t) = c_0 \cos(\epsilon t)$$
(12)

The time scale for the process is indeed given by the initial amplitude of the perturbation  $a_0 = \epsilon$ , while its amplitude is given by the initial mean shear  $c_0$ , quite in contrast with the scalings for an instability. For small times the vertical y-vorticity  $b \sim -c_0 \epsilon t$  shows algebraic growth. The growth saturates on a time scale of  $O(1/\epsilon)$ . The mechanism for the transient growth is the "lift-up" or shear-tilting effect (fig. 2).

The discussion in this section has other dramatic consequences. The interest in transient growth lies in the hope that non-linear effects can be triggered before the viscous decay of the disturbances, but it is clear that there are no significant non-linear effects. The component that is algebraically amplified (the streaks u) does not feedback on the original disturbances (the downstream rolls v, w) and the growth saturates on a time scale of order  $min(\epsilon^{-1}, R)$ . This is also the case for disturbances not purely streamwise. Benney and Gustavsson [7] recognized that fact and emphasized the importance of a multiple mode theory. They did not mention, however, that the algebraic growth is quickly shut off through the modification of the mean flow (Eq. (5)), on a time scale  $O(\epsilon^{-1})$  while the non-linear effects they analyzed occur on a time scale  $O(\epsilon^{-2})$ , even in the presence of multiple modes. Jang, Benney and Gran [27] showed that the non-linear interactions of oblique modes generated downstream rolls which then created large streaks by the shear-tilting mechanism. However, Waleffe *et al.* [8] showed that, at transitional Reynolds numbers, these non-linear effects came from the oblique rolls v and not from the transiently amplified streaks  $\eta$ , even though the latter are amplified by a factor about 10 (as in [15] Fig. 9).

Trefethen et al. [15] considered the  $2 \times 2$  non-linear model problem,

$$\frac{d}{dt} \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} -\lambda & 1 \\ 0 & -\mu \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} + (U^2 + V^2)^{1/2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix}$$
(13)

to show how some non-linear effects coupled to the transient linear amplification could lead to bifurcation of the system. This is an interesting example, although the non-linearity proportional to ||u|| is unphysical. A somewhat more realistic model that incorporates some of the elements discussed above is

$$\frac{d}{dt} \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} -\lambda & M \\ 0 & -\mu \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix}$$

$$\frac{d}{dt} M = -\sigma M + (\sigma - UV)$$
(14)

where M represents the amplitude of the mean shear  $(d\bar{u}/dy)$ , V the amplitude of streamwise vortices and U the amplitude of the streamwise streaks. The UV term is the Reynolds stress term. The rolls V decay viscously and there is no non-linear feedback from the streaks U onto V. The only significant "non-linear" effect is to reduce the mean shear M through the UV term thereby saturating the growth in U. The variables  $\lambda$ ,  $\mu$  and  $\sigma$  are positive and represent viscous decay rates. If  $\lambda = \mu = \sigma = 0$ , the model falls back on system (10) and a periodic solution. In the absence of U, V fluctuations, the mean shear relaxes to M=1 from viscous effects, this parallels the establishment of the mean shear  $d\bar{u}/dy=1$  by viscous diffusion from the walls in plane Couette flow. The model (14) illustrates that in the hydrodynamic context, transient growth does not directly lead to bifurcation.

# 4. Self-sustaining mechanism

The shear-tilting mechanism is capable of creating large modifications of the streamwise velocity, but that does not directly trigger transition. The appearance of strong streamwise streaks also occurs in weakly non-linear expansions involving oblique disturbances. Somewhere down the perturbation expansion, the interaction of oblique waves generates downstream rolls which then induce a significant redistribution of streamwise momentum. This invalidates the weakly non-linear theories because of the appearance of "secular" terms. Those secular terms are not artifacts of the non-linear expansion but rather represent some actual dramatic alteration of the streamwise velocity. That effect was probably first discovered by Benney and Lin [9].

What is needed is a theory which incorporates the strong spanwise distortion of the streamwise velocity ab initio. In essence, the study of a one-dimensional mean flow U(y) is somewhat artificial because very weak disturbances in the form of streamwise rolls create large spanwise fluctuations in the streamwise velocity that persist for a long time. The proper basic state has an intrinsic spanwise variation, U(y,z). One such theory has been proposed by Benney [10] in the form of a Mean Flow-First Harmonic theory, but subsequent work by Benney and Chow [28] seemed to lose the original focus. Nonlinear effects are critical in this theory because of the multitude of possible basic states U(y,z). The basic state under consideration is such that the nonlinear interaction of the instability developing on it sustains the spanwise variation. The theory is thus inherently nonlinear, and the search is for a self-sustaining mechanism.

Direct numerical simulations (DNS) have been used as a guide in the development and validation of these ideas and it has been possible to identify a generic mechanism in turbulent shear flows [16]. The skeleton of a theory for that mechanism is presented below. The spanwise fluctuations of the streamwise velocity are often referred to as "streaks," as a result of the streaky features they produce in visualization experiments. The complete process consists of the following three elements:

- (1) Spanwise modulation of the streamwise velocity by downstream rolls,
- (2) breakdown of the spanwise modulated flow from an inflectional wakelike instability.
- (3) regeneration of the rolls from the non-linear self-interaction of the growing instability.

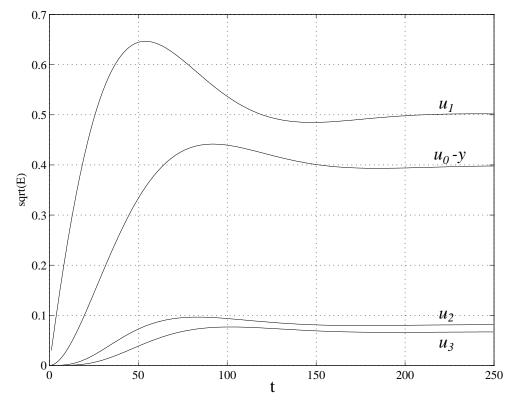


Fig. 3. Time development of the norms of streamwise velocity perturbations for down-stream rolls with  $\max(v) = 0.03$ .

#### 4.1. Streak generation

To illustrate the first step of the process, consider the introduction of downstream rolls

$$v(y,z) = V(y)\cos\gamma z$$
 and then  $W(y,z) = -\frac{1}{\gamma}\frac{dV(y)}{dy}\sin\gamma z,$  (15)

with  $\max[v(y)] = 0.03$  onto the laminar Couette flow profile U = y at R = 400, where v(y) is the slowest decaying downstream OS mode. The (even) downstream OS modes are given by  $V(y) = \cosh \gamma y / \cosh \gamma - \cos py / \cos p$ , with  $\gamma \tanh \gamma + p \tan p = 0$ , see Eq. (26.3) in [1] p.159. The amplitude of 0.03 is about that observed in the DNS and also corresponds to the estimate  $\epsilon R \simeq 10$  deduced in sect. 3 from the balance between dissipation and advection time scales. The streamwise velocity develops a spanwise variation that takes the form, from Eq. (8),

$$U(y,z) = u_0(y) + u_1(y)\cos(\gamma z) + u_2(y)\cos(2\gamma z) + u_3(y)\cos(3\gamma z) + \cdots$$
(16)

The development of the norms  $(\int u_1^2 dy)^{1/2}$ , etc...) of the streamwise perturbations is shown in Figure 3. Note that  $u_1$  reaches its maximum of O(1) after a time about  $\simeq \pi h/(2v_{\text{max}})$  which is approximately the time span needed to advect momentum from the walls to the middle of the channel. At that time, the modification of the mean  $u_0$  is far from negligible and there is a significant departure from the laminar solution U(y) = y. That modification of the mean flow was neglected in previous linear analyses. The downstream rolls are here artificially maintained against viscous decay, their amplitude fixed at its initial value, because the search is for a self-sustaining mechanism where the rolls are maintained on average. Thus the streamwise velocity reaches a steady state determined by a balance between viscous diffusion and advection by the rolls in a time scale of the order of one eddy-turnover time,  $2\pi h/v_{\text{max}}$ . The profiles  $u_i(y)$  and contours of u(y, z) at t = 50 are shown in Figures 4 and 5.

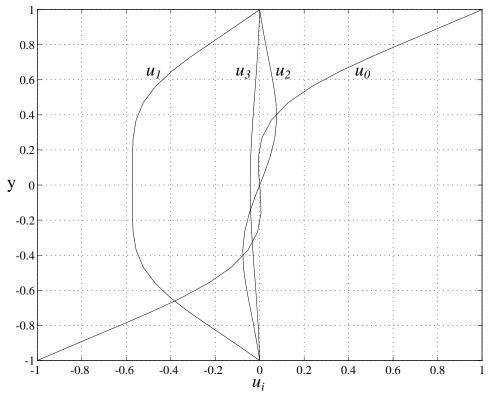


Fig. 4. Velocity profiles at t=50 for max(v)=0.03.

### 4.2. Streak breakdown

The next phase of the process consists of the breakdown of the spanwise modulated flow. To study that instability, linearize the Navier-Stokes equations about the spanwise modulated flow U(y,z), the resulting equations are

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x} - \frac{1}{R}\nabla^2\right)\nabla^2 v + \left(\frac{\partial^2 U}{\partial z^2} - \frac{\partial^2 U}{\partial y^2}\right)\frac{\partial v}{\partial x} + 2\frac{\partial U}{\partial z}\frac{\partial^2 v}{\partial x \partial z} = 2\frac{\partial^2}{\partial x \partial y}\left(w\frac{\partial U}{\partial z}\right),\tag{17}$$

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} - \frac{1}{R} \nabla^2\right) \eta = \left(\frac{\partial U}{\partial z} \frac{\partial}{\partial y} - \frac{\partial U}{\partial y} \frac{\partial}{\partial z}\right) v - \left(v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}\right) \frac{\partial U}{\partial z},\tag{17b}$$

with the kinematic relation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)w = -\frac{\partial\eta}{\partial x} - \frac{\partial^2v}{\partial u\partial z}.$$
 (17c)

These equations admit a general solution of the form  $\exp(i\alpha x + \lambda t)\mathbf{v}(y,z)$ , and there is no need for equ. (5) as long as  $\alpha \neq 0$ . With U(y,z) in (16) satisfying U(y,-z) = U(y,z) and  $U(-y,z+\pi/\gamma) = -U(y,z)$ , the equations have several symmetries and the following form of solution is selected

$$v = e^{i\alpha x + \lambda t} \sum_{n=1}^{\infty} v_n(y) \sin n\gamma z,$$

$$\eta = e^{i\alpha x + \lambda t} \sum_{n=0}^{\infty} \eta_n(y) \cos n\gamma z,$$

$$w = e^{i\alpha x + \lambda t} \sum_{n=0}^{\infty} w_n(y) \cos n\gamma z.$$
(18)

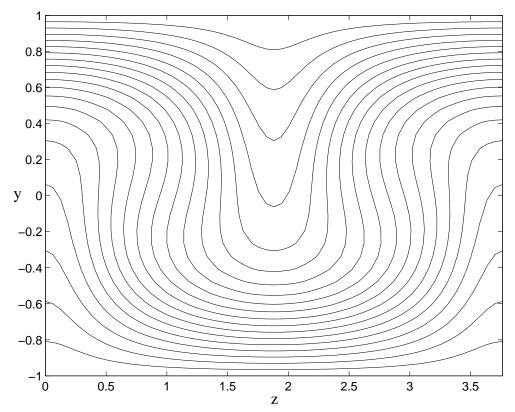


Fig. 5. Contours of streamwise velocity U(y,z) at t=50 and at 0.1 intervals from -1 at the lower wall to +1 at the upper wall, as induced by downstream rolls with  $\max(v)=0.03$ .

This particular form is chosen in anticipation of the fact that the instability is wake-like, resulting from the spanwise inflections in the velocity profile. The form (18) then corresponds to a fundamental (same period as the base flow) "sinusoidal" mode of instability, akin to the instability of flapping flags. The "varicose" mode has a sine expansion for w and  $\eta$  and cosine expansion for v. The subharmonic modes are a combination of varicose and sinusoidal modes and the expansion for w in that case has the form

$$w = e^{i\alpha x + \lambda t} \sum_{n=0}^{\infty} w_n(y) \cos \frac{2n+1}{2} \gamma z.$$

By symmetry, there is also a sine series with the same growth rate.

The growth rate of the first three fastest growing modes on the profile U(y,z) of Figure 5 are shown in Figure 6. There is only one unstable mode for this base flow, and the corresponding eigenvalue is purely real. The next two slowest decaying modes form a complex conjugate pair for most values of  $\alpha$  in the unstable range. The cutoff wavenumber above which there is no unstable modes is approximately  $\alpha \simeq 1.155$ . The maximum growth rate occurs approximately at  $\alpha \simeq 0.74$  and is equal to 0.1347, this is thus a powerful inertial instability given that 0.5 is an upper bound on the growth rate for the laminar profile (U=y). These results are obtained numerically using a Chebyshev expansion in y with 35 points and a Fourier series in z truncated after N terms, not counting zero. The results, the position of the cutoff wavenumber in particular, are fairly sensitive to the truncation.

For comparison, consider the inviscid inflectional instability of the purely spanwise profile  $u(z) = U_1 \cos \gamma z$  with  $U_1 = \max[u_1(y)]$ . The Rayleigh equation [1] for the inviscid stability of a

u(z) profile to normal modes of the form  $w(z) \exp(i\alpha x) \exp(\lambda t)$ , is

$$(\lambda + i\alpha u)(D^2 - \alpha^2) w - i\alpha (D^2 u) w = 0$$
(19)

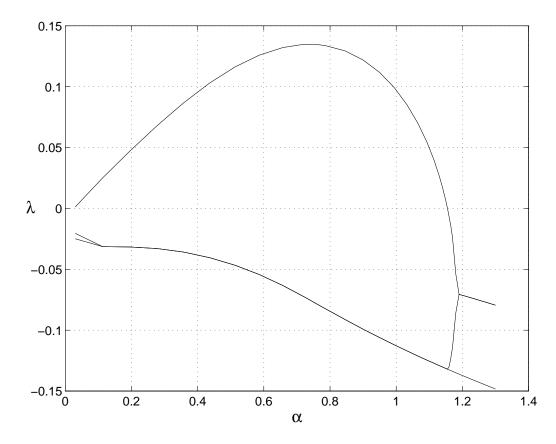
where D = d/dz and w is the spanwise velocity component (see [1] p. 130). For the profile  $u(z) = U_1 \cos(\gamma z)$ , the general form for w(z) is  $w(z) = \exp(i\beta z) \sum_n w_n \exp(in\gamma z)$ , but after some thought the sinusoidal wake-like instability ([1] p. 234) corresponds to a solution of the form  $w(z) = w_0 + w_1 \cos(\gamma z) + \cdots$ . Substituting this expansion truncated to the first 2 modes in the Rayleigh equation (19) yields

$$\lambda \alpha^2 w_0 = -i\alpha U_1 \alpha^2 w_1/2$$
$$\lambda(\alpha^2 + \gamma^2) w_1 = i\alpha U_1 (\gamma^2 - \alpha^2) w_0$$

which gives the growth rate

$$\lambda = \frac{|\alpha U_1|}{\sqrt{2}} \left( \frac{\gamma^2 - \alpha^2}{\gamma^2 + \alpha^2} \right)^{1/2}.$$
 (20)

This is shown in Figure 7. For this problem the exact growth rate can be expressed in terms of a continued fraction [29]. A two-mode truncation often gives a reasonable picture for parametric instabilities, as can be verified for Eq.(19) (Figure 7). However, for the present problem, up to twelve Fourier modes are necessary to obtain converged results. Although the results are qualitatively similar to those in Figure 6 the numerical values differ significantly (Figure 8). In particular, the position of the wavenumber cutoff is quite sensitive to the truncation.



**Fig. 6.** Growth rate (real part of  $\lambda$ ) of the first three least stable modes on the profile U(y,z) for  $\gamma = 1.67$ , R = 400.

## 4.3. Regeneration of downstream rolls

The third and final step of the self-sustaining process is the feedback on the original downstream rolls from the non-linear development of the streak instability. The equation governing the evolution of the downstream rolls  $v(y, z, t) = V(y, t) \cos \gamma z$  is obtained from (3) and has the form

$$\left[\frac{\partial}{\partial t} - \frac{1}{R}(D^2 - \gamma^2)\right] (D^2 - \gamma^2) V = f, \tag{21}$$

where the forcing term f is given by the projection of the non-linear interaction of the fastest growing eigenmodes of U(y, z) onto the  $\cos \gamma z$  mode,

$$f = \frac{1}{2} \left\langle \cos \gamma z, \frac{\partial^3}{\partial z^2 \partial y} (ww - vv) + \left( \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) \frac{\partial}{\partial z} (vw) \right\rangle.$$
 (22)

The brackets  $\langle \rangle$  denote an average over horizontal planes x, z, and v, w in this formula stand for the y and z velocity components of the most unstable mode only.

The shape of the first downstream OS mode and the steady solution to (21) are shown in Fig. 9. The downstream rolls are clearly regenerated. The structure of the sustained rolls is only slightly different from that of the first downstream OS mode. This new shape could be used as the form of new streamwise rolls and the whole sequence reiterated until convergence. However it is sufficient for the present purpose that the projection of the forced response onto the original shape is large. This is sufficient to demonstrate that the streamwise rolls are reenergized by the nonlinear development of the streak instability.

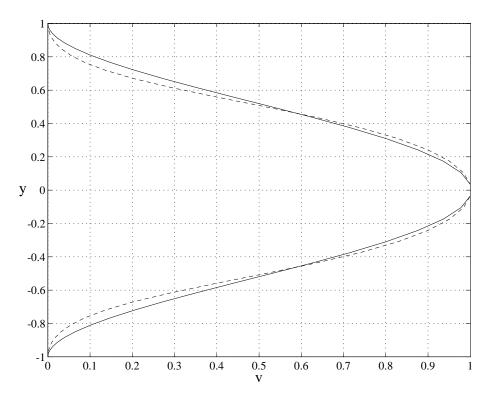


Fig. 9. Roll regeneration; the first downstream OS mode (solid) and the steady response to the non-linear forcing (dashed)

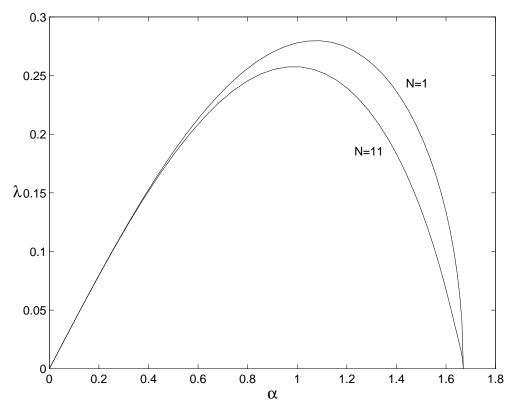
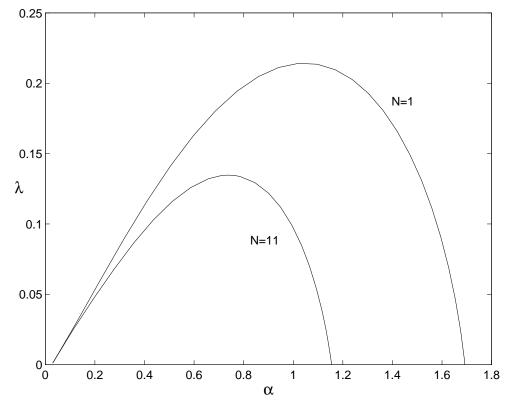


Fig. 7. Growth rate of the inviscid inflectional instability of  $U_1 \cos \gamma z$ , Eq.(19), with 2 Fourier modes (N=1) and the converged results at N=11.



**Fig. 8.** Same as 7 but for the full viscous instability of U(y, z), Eq.(17).

## 4.4. Spanwise varying mean field theory

The three phases can be incorporated into a mean field theory where the mean consists of a U(y,z) profile maintained by downstream rolls V(y,z), W(y,z). For a given Reynolds number, the amplitude of these rolls is adjusted so as to give a zero growth rate for the instability of the U(y,z) profile. The amplitude of that neutral mode would be determined from (21) in order to maintain the rolls at their proper amplitude. The amplitude of the instability is linked to the amplitude of the rolls through (21). One is left with the computationally intensive task of mapping out a solution hypersurface in the four dimensional space formed by the Reynolds number R, the spanwise wavenumber  $\gamma$ , the streamwise wavenumber  $\alpha$  and the amplitude of the perturbation. Hence, choosing  $R=400, \gamma=1.67$  and the amplitude of the rolls at  $\max[V(y)]=0.03$ , the streamwise wavenumber must be  $\alpha\simeq 1.155$  for marginality (Fig. 6). This gives one point on the solution hypersurface.

A mean field theory of a similar nature has been used with great success by Meksyn and Stuart [30]. Their work was limited to two dimensions and the interaction between the mean and the fluctuation was direct, while in the present proposal it takes place indirectly through the streamwise rolls. This proposed mean field theory has strong connections also with the work of Nagata [31] who discovered finite amplitude three-dimensional steady solutions in plane Couette flow. His technique is to track nonlinear solutions in rotating plane Couette flow as the rotation rate is reduced to zero. The mean field problem is essentially a special severe truncation of the steady Navier-Stokes equations. Although it might appear that Nagata's approach is more rigorous to find steady solutions, in practice a severe truncation must be employed there also. This probably explains some disturbing features of Nagata's solutions such as the very low Reynolds numbers at which they were discovered (150, while experiments and numerical simulations indicate a critical R near 350) and the fact that when perturbed those solutions decay back to the laminar state [32]. In the mean field theory, the truncation is inspired more by the physics and less by machine constraints. If successful, the present mean field theory would provide deep physical insight by isolating the fundamental interactions. One interesting feature of the proposed approach is that it can be readily extended to deal with oscillatory disturbances by allowing for a purely imaginary eigenvalue  $\lambda = i\omega$ . When  $\omega \neq 0$ , the mean field theory should be seen as a statistical theory for a mean and a fluctuating field. Non-zero  $\omega$  might be crucial to extend this theory to higher Reynolds number where the structures are closer to the walls and propagating.

#### 5. Model of the mechanism

The scenario sketched in the previous section, in which rolls create streaks that break down to maintain the rolls, can be illustrated by the following simple model,

$$\frac{d}{dt} \begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} -\lambda & M & 0 \\ 0 & -\mu & 0 \\ 0 & 0 & -\nu \end{pmatrix} \begin{pmatrix} U \\ V \\ W \end{pmatrix} + W \begin{pmatrix} 0 & 0 & -c \\ 0 & 0 & d \\ c & -d & 0 \end{pmatrix} \begin{pmatrix} U \\ V \\ W \end{pmatrix} 
\frac{d}{dt} M = -\sigma M + (\sigma - UV),$$
(23)

where U, V and M represent respectively the amplitude of the streamwise streaks, streamwise rolls and the mean shear as before, and W is the amplitude of the streak instability. The wake-like streak instability consists mostly of velocity in the spanwise direction. All constants  $\lambda, \mu, \nu, \sigma, c, d$  are positive, with the Greek symbols representing decay rates inversely proportional to the Reynolds number. The other two constants c and d are nonlinear interaction coefficients. They depend indirectly on the Reynolds number, through distortion of the mode structures. The instability W grows from the streaks U (cUW term) and feeds the rolls V by nonlinear self-interaction ( $dW^2$ 

term). The other nonlinear terms follow from conservation of energy and the meaning of the other terms is as previously.

In addition to the laminar state U = V = W = 0, M = 1, there may be other fixed points  $U_0, V_0, W_0, M_0$  determined by the roots of the cubic

$$C(X) = d^4 X^3 + \nu d^2 X^2 - \sigma (dc - \mu c^2 - \lambda d^2) X + \lambda \nu \sigma = 0, \tag{24}$$

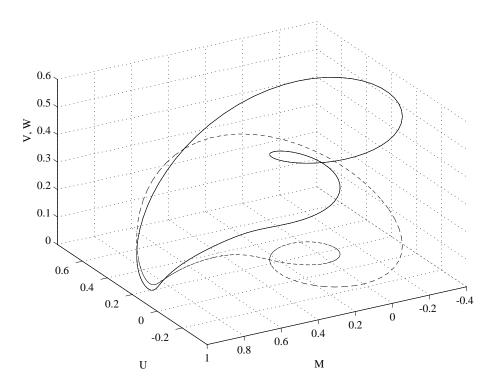
where  $X=W_0^2/\mu$ . The intercept at X=0 is positive and equal to  $\lambda\nu\sigma$ , the slope there is the coefficient of the X-term and must be negative to have positive roots. Thus a necessary but not sufficient condition to have a root is that  $dc>\mu c^2+\lambda d^2$ , which gives a lower bound on the critical Reynolds number. The exact equation for the critical Reynolds number can be derived from the criterion that  $\mathcal{C}(X_*) \leq 0$  at  $X_*$  such that  $d\mathcal{C}/dX=0$ . Fixing somewhat arbitrarily  $\lambda=\mu=\sigma=\nu/2$ , and c=0.8, d=1, the bifurcation occurs near  $\nu=0.25$ . Above that critical Reynolds number, there are two roots that emerge from a saddle-node bifurcation. It can be shown that the smallest one is always unstable. The stability analysis leads to finding the roots of a fourth order polynomial for the eigenvalues. All coefficients of that polynomial are positive except for the linear term that is indefinite and the constant term that has the sign of the slope of  $\mathcal C$  at the root. Hence, one can conclude that the lower amplitude root is always unstable as there will be one positive eigenvalue in that case.

For the parameter values  $\lambda = \mu = \sigma = \nu/2$ ,  $\nu = 0.1$ , c = 0.8, d = 1, the larger root corresponds to the fixed point  $U_0 = 0.3017$ ,  $V_0 = 0.1414$ ,  $W_0 = 0.0841$ ,  $M_0 = 0.1467$ , but it is observed numerically to be unstable at those parameter values. For the initial condition (U, V, W, M) = (0, 0.03, 0.01, 1), the system settles quickly onto a periodic orbit shown in Fig. 10. The period is about T = 140 and the system was integrated up to t = 5000. For larger values of c the flow converges to the fixed point after spending some time around the periodic orbit, while for smaller values of c the flow goes back to the laminar state.

The time evolution over one period is displayed in Fig. 11. It shows an interesting "burst-like" behavior with short-lived explosive growths of U, V and W which wipe out the mean shear M, followed by a slower viscous recovery in which W is very small. The growth of W is initiated by the instability of U from which it draws energy, but then the downstream rolls V are enhanced and augment the streaks U by extracting energy from the mean M, leading to an explosive nonlinear growth of U, V and W. Whether or not the model exhibits chaotic behavior is secondary to our purpose, which is to illustrate the nature of the self-sustained mechanism coexisting with the laminar state above some critical Reynolds number.

#### 6. Conclusion

A self-sustaining mechanism believed central to the instability of shear flows has been described. It consists of three elements: formation of streaks by downstream rolls, breakdown of the streaks and regeneration of the rolls from the streak breakdown. Each of the three elements of the self-sustaining process has been described and supported by numerical evidence. A simple four-equation model illustrates the mechanism and shows that it can coexist with the laminar state. This scenario appears not only responsible for the non-linear instability of shear flows but also occurs in the near-wall region of high Reynolds number turbulent shear flows. In a series of numerical experiments, Hamilton et al. [16] captured and analyzed the process as it takes place in the full Navier-Stokes equations at low Reynolds number. In those direct numerical simulations, a turbulent flow was obtained from initial random perturbations. The dimensions of the domain were then slowly reduced (thus reducing the Reynolds number) to suppress some of the disorder [33]. The flow settled on a near-periodic cycle consisting of the self-sustaining mechanism, lending support to the robustness and relevance of that process.

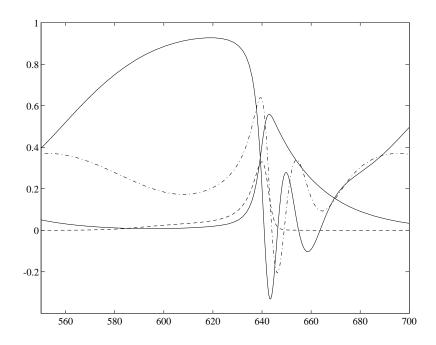


**Fig. 10.** Projection of the periodic orbit onto the U, M, V space (solid), and the U, M, W space (dashed).

There remains much work to be done. First it is necessary to solve the mean-field theory proposed herein in order to determine the favored spanwise scale and critical Reynolds number for the process. This work has focused on Couette flow but the process is believed to be generic and other shear flows must be considered to validate that point. One of the most exciting issues is the conjecture that the process maintains the same effective Reynolds number at any flow Reynolds number. This has been advanced as an explanation for the persistence of the observed streak spacing of 100 wall units [8,16].

The streak formation phase has been extensively studied before as a linear transient effect. A secondary aim of this article was to clarify the physical nature of those transient processes and point out some serious limitations of the classical linear analyses based on a unidimensional base flow. Most notably these analyses are valid only for  $\epsilon R \ll 1$  and the transient growth does not directly trigger transition.

No mention has yet been made of the theory of "secondary instability" [34,35,36], which has been very successful in explaining the development of three-dimensionality and the breakdown of spanwise rollers. This is very important especially for boundary [34,37] and mixing layers [38], where a primary instability introduces the spanwise rollers. In plane Couette and Poiseuille flows, the secondary instability theory loses its foundation because the primary instability is either nonexistent or inactive at the observed critical Reynolds number. It would be necessary to show that the spanwise rollers are sustained by the secondary instability but efforts in that direction have been unsuccessful. However, even when the primary instability is active, the secondary instability theory also describes only a transient effect, albeit more significant in that case because there is a definite departure from the laminar state. It identifies the major processes in the breakdown of the laminar flow but does not determine where the flow is going. That two stage development leads



**Fig. 11.** Time evolution of the variables over one period, U: dash-dot, V: lower solid, W: dashed, M: upper solid.

to "turbulence," which remains a mystery outside the scope of the secondary instability theory. In contrast, the mechanism explored in this article is complete and self-reproducing. The observations of the near-wall region of turbulent flows and the direct simulations, suggest that this could be the central part of the momentum exchange process with the disorder playing a lesser role.

## Acknowledgements

Early discussions and collaboration with John Kim and Jim Hamilton are gratefully acknowledged. John Kim also verified the results of Figure 8 for N=11 through Direct Numerical Simulations. The author thanks David Benney, Lou Howard, Joe Keller, Willem Malkus and Trevor Stuart for comments on the manuscript.

# References

- 1. P.G. Drazin and W.H. Reid, *Hydrodynamic stability* (Cambridge U. Press, 1981).
- 2. T. Ellingsen and E. Palm, "Stability of Linear flow," Phys. Fluids 18, 487-488 (1975).
- 3. M.T. Landahl, "Wave breakdown and turbulence," SIAM J. Applied Math. 28, 775 (1975).
- 4. L.H. Gustavsson and L.S. Hultgren, "A resonance mechanism in Plane Couette flow," J. Fluid Mech. 98, 149-159 (1980).
- 5. L.H. Gustavsson, "Resonant growth of three dimensional disturbances in plane Poiseuille flow," *J. Fluid Mech.* **112**, 253-264 (1981).
- 6. M.T. Landahl, "A note on an algebraic instability of inviscid parallel shear flows," J. Fluid Mech. 98, 243-251 (1980).

- 7. D.J. Benney and L.H. Gustavsson, "A New Mechanism for Linear and Non-linear Hydrodynamic Instability," Stud. Appl. Math. 64, 185-209 (1981).
- 8. F. Waleffe, J. Kim and J.M. Hamilton, "On the origin of streaks in turbulent shear flows," in *Turbulent Shear Flows* 8, 37-49 (1993) F. Durst, R. Friedrichs, B.E. Launder, F.W. Schmidt, U. Schumann, J.H. Whitelaw (eds.) (Springer-Verlag, 1993) (Proceedings of the Munich, 1991 symposium); and F. Waleffe *Ann. Res. Briefs-1990*, Center for Turbulence Research, Stanford U./NASA Ames, 159-168.
- 9. D.J. Benney and C.C. Lin, "On the secondary motion induced by oscillations in a shear flow," *Phys. Fluids* **3**, 656-657 (1961); D.J. Benney, "A non-linear theory for oscillations in a parallel flow", *J. Fluid Mech.* **10**, 209-236 (1961), see also ref. [1] p. 454; P.J. Schmid and D.S. Henningson, "A New mechanism for rapid transition involving a pair of oblique waves", *Phys. Fluids A* **4**, 1986 (1992).
- 10. D.J. Benney, "The evolution of disturbances in shear flows at high Reynolds numbers," Stud. Appl. Math. 70, 1-19 (1984).
- 11. M.S. Acarlar and C.R. Smith, "A study of hairpin vortices in a laminar boundary layer," J. Fluid Mech. 175, 1-41 and 43-83 (1987).
- 12. L.H. Gustavsson, "Energy growth of three-dimensional disturbances in plane Poiseuille flow," *J. Fluid Mech.* **224**, 241-260 (1991).
- 13. K.M. Butler and B.F. Farrell, "Optimal perturbations and streak spacing in wall-bounded shear flows," *Phys. Fluids A* 5, 774-777 (1993).
- 14. S.C. Reddy and D.S. Henningson, "Energy growth in viscous channel flow," J. Fluid Mech. 252, 57-70 (1993).
- 15. L.N. Trefethen, A.E. Trefethen, S.C. Reddy and T.A. Driscoll, "Hydrodynamic Stability without Eigenvalues," *Science* **261**, 578-584 (30 July, 1993).
- 16. J.M. Hamilton, J. Kim and F. Waleffe, "Regeneration mechanisms of near-wall turbulence structures," J. Fluid Mech. 287, (pp. 317-348) (1995)
- 17. N. Tillmark and P.H. Alfredsson, "Experiments on transition in plane Couette flow," J. Fluid Mech. 235, 89 (1992)
- 18. A. Lundbladh and A. Johansson, "Direct simulation of turbulent spots in plane Couette flow," J. Fluid Mech. 229, 499-516 (1991).
- 19. W.M.F. Orr, "The stability or instability of the steady motions of a perfect liquid and of a viscous liquid," *Proc. Roy. Irish Acad. A.* 27, 9-138 (1907).
- 20. K.M. Case, "Hydrodynamic instability and the inviscid limit," J. Fluid Mech. 10, 420 (1960) and "Stability of inviscid plane Couette flow," Phys. Fluids 3, 143 (1960).
- 21. B.F. Farrell, "Optimal excitation of perturbations in viscous shear flows", *Phys. Fluids* **31**, 2093-2102 (1988) and ref. [15].
- 22. F.H. Busse, "Bounds on the transport of mass and momentum by turbulent flow between parallel plates," Z. Angew. Math. Phys. 20, 1 (1969).
- 23. W.V.R Malkus "A proof of the non-existence of steady two-dimensional disturbances in unbounded plane parallel flow", *IGPP manuscript*, *University of California*, *Los Angeles*, 13 pages (1968) (unpublished).
- 24. J.T. Stuart, "The production of intense shear layers by vortex stretching and convection," NATO AGARD report 514, 1-29 (1965).

- 25. C.F. Pearson and F.H. Abernathy, "Evolution of the flow field associated with a streamwise diffusing vortex," *J. Fluid Mech.* **146**, 271-283 (1984).
- 26. D.W. Moore, "The interaction of a diffusing line vortex and an aligned shear flow," *Proc. R. Soc. Lond. A* **399**, 367-375 (1985).
- 27. P.S. Jang, D.J. Benney and R.L. Gran, "On the origin of streamwise vortices in a turbulent boundary layer," J. Fluid Mech. 169, 109-123 (1986).
- 28. D.J. Benney and K.A. Chow, "A mean flow first harmonic theory for hydrodynamic instabilities," Stud. Appl. Math. 80, 37-73 (1989).
- 29. L.D. Meshalkin and Ia.G. Sinai, "Investigation of the stability of a stationary solution of a system of equations for the plane movement of an incompressible viscous liquid," *J. Appl. Math. Mech.* (*PMM*) **25**, 1700-1705 (1961). (see also e.g. F. Waleffe, "On the three-dimensional instability of strained vortices", *Phys. Fluids A* **2**, 76-80 (1990).)
- 30. D. Meksyn and J.T. Stuart, "Stability of viscous motion between parallel planes for finite disturbances," *Proc. R. Soc. Lond. A* **208**, 517-526 (1951).
- 31. M. Nagata, "Three-dimensional finite-amplitude solutions in plane Couette flow: bifurcation from infinity," J. Fluid Mech. 217, 519-527 (1990).
- 32. R.M. Clever and F.H. Busse, "Three-dimensional convection in a horizontal layer subjected to constant shear," J. Fluid Mech. 234, 511-527 (1992).
- 33. J. Jimenez and P. Moin, "The minimal flow unit in near-wall turbulence," J. Fluid Mech. 225, 213-240 (1991).
- 34. B.J. Bayly, S.A. Orszag and T. Herbert, "Instability mechanisms in shear flow transition," *Ann. Rev. Fluid Mech.* **20**, 359-391 (1988)
- 35. S.A. Orszag and A.T. Patera, "Secondary instability of wall-bounded shear flows," *J. Fluid Mech.* **128**, 347-385 (1983).
- 36. T. Herbert, "Secondary instability of plane channel flow to subharmonic three-dimensional disturbances," *Phys. Fluids* **26**, 871-874 (1983).
- 37. T. Herbert, "Secondary instability of boundary layers," Ann. Rev. Fluid Mech. 20, 487-526 (1988).
- 38. M.M. Rogers and R.D. Moser, "Spanwise scale selection in plane mixing layers," *J. Fluid Mech.* **247**, 321-337 (1993).