Optimal excitation of three-dimensional perturbations in viscous constant shear flow

Brian F. Farrell
Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138

Petros J. Ioannou
Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 17 July 1992; accepted 29 January 1993)

The three-dimensional perturbations to viscous constant shear flow that increase maximally in energy over a chosen time interval are obtained by optimizing over the complete set of analytic solutions. These optimal perturbations are intrinsically three dimensional, of restricted morphology, and exhibit large energy growth on the advective time scale, despite the absence of exponential normal modal instability in constant shear flow. The optimal structures can be interpreted as combinations of two fundamental types of motion associated with two distinguishable growth mechanisms: streamwise vortices growing by advection of mean streamwise velocity to form streamwise streaks, and upstream tilting waves growing by the down gradient Reynolds stress mechanism of two-dimensional shear instability. The optimal excitation over a chosen interval of time comprises a combination of these two mechanisms, characteristically giving rise to tilted roll vortices with greatly amplified perturbation energy. It is suggested that these disturbances provide the initial growth leading to transition to turbulence, in addition to providing an explanation for coherent structures in a wide variety of turbulent shear flows.

I. INTRODUCTION

Transition from laminar to turbulent flow in experiments with sheared mean velocity profiles occurs for Reynolds numbers characteristically ranging from \( R \approx 1000 \) to \( R \approx 8000 \), depending on the level of noise in the experiment. With care to reduce background disturbances the transition Reynolds number can be greatly increased, reaching values as high as \( R = 10^5 \) in pipe Poiseuille flow. This observation and the fact that the behavior of small perturbations is described accurately by the linearized Navier-Stokes equations suggests that linear analysis should be sufficient to describe at least the early stages of the transition process in flows with sufficiently small initial perturbations. However, the search for an explanation of transition based on linear theory has been frustrated by the lack of an exponential modal instability in Couette, pipe Poiseuille, and in plane Poiseuille flow below \( R = 5772 \). This difficulty has led to extensive work on secondary instabilities of finite-amplitude perturbations to these canonical shear profiles. However, it has been appreciated more fully recently that the linear equations associated with these shear flows support a set of optimal perturbations that produce large transient growth with growth rates typically on the same advective time scale as inflectional instabilities. One subset of these optimal perturbations is related to the transient wave development first identified by Orr, while another is related to the transient streamwise streak growth extensively discussed by Landahl, although these mechanisms typically occur in combination in the 3-D optimals. When analyzed by expansion in normal modes of the linearized equations the growth of optimal perturbations is seen to occur despite the decay of all individual normal modes. In such a normal mode expansion growth arises from nonorthogonality of these modes due to non-normality of the associated dynamical operator.

Recent experiments conducted by Klingmann indicate that the initiating mechanism of transition is consistent with linear dynamics. Linear growth or decay in the asymptotic limit of long time is controlled by the largest real eigenvalue of the dynamical operator, which is often negative for Reynolds numbers at which transition is observed. However, because the level of background disturbance cannot be reduced to zero either in transitional or in turbulent flows, this asymptotic limit may not be deterministic of stability for real problems. If a subset of perturbations grows by as much as three orders of magnitude, as has been shown to be the case for an exponentially stable shear flow at moderate Reynolds number, then an alternative mechanism underlying turbulence transition and maintenance based on amplification of stochastic perturbations is suggested. In application of this linear theory to turbulent flow account must be taken of the disruption of the perturbations and the modification of the shear profile by the small-scale and uncorrelated disturbances; a mean field theory of this kind for streak production in boundary layer flow has been advanced recently.

While fully developed turbulence has traditionally been characterized as disordered, it is now widely recognized that a limited set of disturbances impart a considerable degree of order to the larger scales. One structure often associated with coherent motion in shear flow consists of inclined vortices described by Townsend as double roller eddies. Related structures observed in channel flows, called hairpin vortices by Klebanoff et al., can be

1390 Phys. Fluids A 5 (6), June 1993 0699-8213/93/061390-11$06.00 © 1993 American Institute of Physics 1390
described as consisting of inclined streamwise vortices similar to those of Townsend's rollers but connected by a spanwise oriented head region. Numerical simulations of homogeneous turbulent shear flow confirmed the presence of hairpin vortices and demonstrated the rapid emergence of these characteristic structures from uncorrelated forcing, strongly suggesting that these structures are characteristic of turbulent shear flows in general, and that they originate from a universal process not qualitatively dependent on the particular shear flow chosen. A Karhunen-Loeve decomposition of a direct simulation of turbulent Poiseuille flow revealed the dominant coherent structures to consist of streamwise rolls and ±71.5° oblique waves.16

While numerical and laboratory experiments have produced increasingly accurate depictions of the coherent structures responsible for observed spatial and temporal correlations, understanding the dynamic origin of coherent structures remains a challenge. A theory of the origin of coherent structures must account for the observations that they are similar in all shear flows, and that they arise rapidly and spontaneously from small random initial perturbations.14 This advective time scale development from small initial conditions implies the validity of linear theory, while the universality of structure implies a common instability mode, a mode that analysis of the normal mode spectrum fails to identify. Townsend attempted to comprehend these facts using rapid distortion theory, which consists of an application of linear theory to obtain an approximation to the initial distortion of perturbations in strong shear (it should be noted that the dynamical equations we will employ are identical to those of rapid distortion theory). Remarkably, the two-point velocity correlation functions associated with coherent structures observed in shear turbulence18 were quite accurately reproduced by this linear theory, thus providing further evidence for the approximate validity of linear theory in strong shear.

The general solution to the linear initial value problem in shear can be obtained using optimal excitation theory in which a complete set of perturbations ordered by energy growth is found as a solution to the variational problem for maximizing growth over a chosen interval of time. The linearized equations governing evolution of disturbances on an unbounded constant shear flow are

\[ L \Delta \psi = 0, \]  
\[ L \omega_z = -\alpha \partial \nu, \]  
\[ L \nu = \partial_y (\alpha \nu \partial_y) - 3\Delta, \]

where \( U = \partial y \) is the background velocity in the x direction, \( (u, v, w) \) denote the perturbation velocities in the x, y, z directions, respectively, \( \omega_z = \partial_x u - \partial_y v \) is the cross-stream component of vorticity, \( \nu \) is the coefficient of viscosity, and \( \Delta = \partial_x^2 + \partial_y^2 + \partial_z^2 \). The continuity equation,

\[ \partial_t \mu + \partial_x \nu + \partial_z \omega = 0, \]

has been used in the derivation of (1).

Trial form solutions of the form

\[ \psi = b(t) \exp \{ [k_1(t) x + k_2(t) y + k_3(t) z] \}. \]  
\[ \omega_z = \partial_y \psi \exp \{ [k_1(t) x + k_2(t) y + k_3(t) z] \}, \]

satisfy (1) for all \( x, y, z \) if

\[ k_1 = k_{01}, \quad k_2 = k_{02} - \nu t k_{01}, \quad k_3 = k_{03}, \]

with the subscript "0" denoting the value at the initial instant. The complex representation in (3) is used with the understanding that only the real parts are considered physical. According to (4) the planes of constant phase, which are perpendicular to the wave number vector \( (k_1, k_2, k_3) \), rotate clockwise under the influence of the mean shear, which is assumed to be positive. When \( k_1 k_{02} > 0 \) the phase planes become vertical at time \( t_p = k_{02}/\nu k_{01} \), while as \( t \to \infty \) the phase planes become nearly horizontal. Note that the continuity equation (2) constrains the velocity components to lie along planes of constant phase implying vanishing cross-stream velocity as

II. THE PLANE WAVE SOLUTION

The linearized equations governing evolution of disturbances on an unbounded constant shear flow are

\[ L \Delta \psi = 0, \]
\[ L \omega_z = -\alpha \partial \nu, \]
\[ L \nu = \partial_y (\alpha \nu \partial_y) - 3\Delta, \]
and the phase planes become horizontal. Also, the continuity equation implies that for each of these plane waves the nonlinear terms that have been neglected in \((1)\) vanish identically. Unfortunately, a superposition of plane wave solutions does not retain this property so that our analysis is restricted to linear validity when a superposition of waves is considered.

Solving for the time development of a plane wave perturbation requires specifying the initial cross-stream vorticity, \(\tilde{\omega}_0\), and cross-stream velocity, \(\tilde{u}_0\) from which the initial horizontal velocities \(\tilde{u}_0, \tilde{v}_0\) can be recovered with the aid of the continuity equation \((2)\) and the definition \(\tilde{\omega}_j = \text{Im} \tilde{u}_j - \text{Im} \tilde{v}_j\). The solution of \((1)\) can be considered as the sum of the solution of the homogeneous equations:

\[
L \Delta v = 0, \quad L \omega_{ph} = 0, \quad \partial_t u_h + \partial_t v + \partial_z u_h = 0, \quad (5a, b)
\]

with complex initial amplitudes:

\[
\tilde{\delta}(0) = \tilde{\delta}_0; \quad \tilde{\omega}_{ph}(0) = \tilde{\omega}_0, \quad (5c)
\]

and the solution of the inhomogeneous equation driven by \(v\) and with zero initial conditions:

\[
L \omega_{inh} = -\alpha \partial_t v, \quad (6a)
\]

\[
\partial_t u_{inh} + \partial_t \omega_{inh} = 0. \quad (6b)
\]

The significance of separating the homogeneous and inhomogeneous solutions will be discussed in the sequel. The time-dependent cross-stream velocity, \(v\), and cross-stream vorticity, \(\omega_p\), as given by the sum of the homogeneous and the inhomogeneous solutions is

\[
\tilde{\delta}(t) = \tilde{\delta}_0 \frac{K_x^2}{K_y^2} e^{-r}, \quad (7)
\]

\[
\tilde{\omega}_p(t) = \tilde{\omega}_{ph}(t) + \tilde{\omega}_{inh}(t), \quad (8a)
\]

with

\[
\tilde{\omega}_{ph}(t) = \tilde{\omega}_0 e^{-r}, \quad \tilde{\omega}_{inh}(t) = -i \tilde{\delta}_0 \frac{K_x^2}{K_y^2} [\theta] e^{-r}, \quad (8b)
\]

where \(K_x(t) = K_1 + K_2 + K_3\) is the total instantaneous wave number, \(K_y = K_1 + K_2\) is the total horizontal wave number, \(r = \gamma f K_2^2(t) d\tau\) is the cumulative dissipation factor, \(\theta = \tan^{-1}(\lambda/K_3)\), and \([f(t)] = f(t) - f(0)\) for any function of time \(f(t)\).

It is instructive to consider the 2-D limits of \((5)\) and \((6)\). First, consider motion with no spanwise variation \((\kappa_3 = 0)\). In that case the inhomogeneous solution vanishes and the homogeneous solution alone describes the evolution of the 2-D perturbation. The inviscid dynamics of this perturbation can be understood by noticing that it conserves spanwise vorticity, \(\omega_z = \partial_x v - \partial_y u\), leading, due to kinematic deformation by the shear flow, to transient growth of the cross-stream and streamwise velocity fields for waves that are initially inclined with constant phase surfaces oriented against the mean shear \(22\) (for \(\kappa_1 \kappa_2 \gg 1\)). This is the mechanism of growth in 2-D shear discussed by Orr \(7\) associated with conservation in a straining field of spanwise perturbation vorticity, \(\omega_z\), and it will be referred to as the Orr mechanism. The homogeneous solution of \((5)\) is the 3-D extension of this 2-D Orr mechanism. According to \((5)\) the homogeneous solution conserves cross-stream vorticity in the inviscid limit. When the planes of constant phase are parallel to the cross-stream axis, the cross-stream velocity reaches its maximum, while at later times the velocity fields decay, ultimately vanishing in the limit \(t \to \infty\). An example of a typical inviscid evolution of the vertical velocity for the 3-D Orr mechanism with \(\kappa_1 = 1, \kappa_2 = 10\), and various \(\kappa_3\) is shown in Fig. 1.

The case of no streamwise variation \((\kappa_1 = 0)\) is the other 2-D limit. In the absence of viscosity the homogeneous components of the solution remains constant and equal to their initial values as there is no time variation of the wave numbers as is. On the other hand, the inhomogeneous equation \((6)\) is continually forced by tilting of the background spanwise vorticity by the spanwise varying cross-stream velocity producing linear growth in time of the cross-stream vorticity, \(\partial_j\), as can be easily verified by taking \(\kappa_1 \to 0\) in \((8b)\). Equivalently, the spanwise varying cross-stream velocity can be regarded as lifting and depressing material parcels in the constant shear background flow producing velocity fields dominated by streamwise perturbation streaks, which are given by

\[
\tilde{u}(t) = (\tilde{u}_0 - \alpha t \tilde{\omega}_0) e^{-r}, \quad (9a)
\]

\[
\tilde{v}(t) = \tilde{v}_0 e^{-r}; \quad \tilde{u}(t) = \tilde{u}_0 e^{-r}. \quad (9b)
\]

While a single wave of the form \((9)\) is a nonlinear solution of the Navier-Stokes equations in an unbounded constant shear flow, to obtain a streamwise roll vortex, of the form discussed by Ellingsen and Palm \(23\) and Landahl \(8\), four plane wave solutions must be summed, leading to a solution strictly valid only in the linear limit. Again Eqs. \((6)\) provide the 3-D extension of the 2-D streamwise roll case \((\kappa_1 = 0)\).
In the general 3-D case in which both the Orr and tilting mechanism are operating, the expressions for the streamwise and spanwise velocities fields are

\[ \dot{u}(t) = \dot{u}_h + \dot{u}_{inh}, \quad \dot{w}(t) = \dot{w}_h + \dot{w}_{inh}, \]  

where

\[ \dot{u}_h = \dot{u}_0 e^{-\xi} \left( \frac{k_1^2 K_f^2}{2} \right) \dot{\theta}_0 e^{-\xi}, \]

\[ \dot{w}_h = \dot{w}_0 e^{-\xi} \left( \frac{k_2^2 K_f^2}{2} \right) \dot{\theta}_0 e^{-\xi}, \]

\[ \dot{u}_{inh} = \frac{k_3^2 K_f^2}{k_1^2} \left( \frac{\dot{\theta}}{\alpha} \right) e^{-\xi}, \]

\[ \dot{w}_{inh} = \frac{k_3^2 K_f^2}{k_1^2} \left( \frac{\dot{\theta}}{\alpha} \right) e^{-\xi}. \]

A typical evolution of the horizontal velocity components in the inviscid limit is also shown in Fig. 1 for the 3-D perturbation with initial wave numbers \( \kappa_1 = 1 \), \( \kappa_2 = 10 \), and various \( \kappa_3 \). The velocities induced by the inhomogeneous solution due to the tilting mechanism can be seen to asymptote to a constant in this inviscid limit, although with nonvanishing viscosity all perturbations would eventually decay to zero.

The interplay between the Orr mechanism and the tilting mechanism depends on the ratio \( r = \kappa_2 / \kappa_1 \). As we have discussed, when \( r = 0 \) only the Orr mechanism is present, while when \( r \to \infty \) only the tilting mechanism is at work. For intermediate values of \( r \) the induction of strong cross-stream velocities by the Orr mechanism produces greatly increased streamwise velocity. The maximum cross-stream velocity, attained at \( t_v = K_0/\alpha K_1 \), is given by

\[ \dot{v}_{max} = \dot{\theta}_0 \left( 1 + r^2 + \frac{\kappa_2^2}{\kappa_1^2} \right), \]

which for \( \kappa_2 / \kappa_1 = 10 \) and \( r = 1 \) produces a 52-fold increase of the vertical velocity from its initial value. This large increase in cross-stream velocity leads through the tilting mechanism to rapid development of streamwise streaks. The mean perturbation energy density, \( E(t) \), is defined as the mean perturbation kinetic energy per unit mass per unit fluid volume:

\[ E(t) = \lim_{L \to \infty} \frac{1}{2L} \int_{-L}^{L} \left( u^2 + v^2 + w^2 \right) dy. \]  

The mean perturbation energy density equation can be derived by multiplying each component of the momentum equation by its corresponding velocity and integrating over space to obtain

\[ \frac{dE}{dt} = - \lim_{L \to \infty} \frac{1}{2L} \int_{-L}^{L} (\alpha uv + v \partial u_j / \partial x_j) dy, \]

where the summation convention is used with the indices \( i,j \) denoting the coordinate components. Note that, as in the familiar 2-D case, the perturbation energy density waxes by upgradient Reynolds stresses, which can be recognized by the orientation of the planes of constant phase against the mean shear (for \( \kappa_2 / \kappa_1 > 0 \)). For a single plane wave the Reynolds stress is given by

\[ uu = u_h u + u_{inh} u, \]

in which \( * \) denotes a complex conjugation. There are two sources for the Reynolds stress: the stress due to the homogeneous term (15a), which reduces in the 2-D case to the term \( A \), and that due to the inhomogeneous term (15b), which reduces in the limit \( \kappa_1 \to 0 \) to the lift-up mechanism associated with the streamwise rolls. The time evolution of the two contributions in the inviscid limit for various \( \kappa_1 \) is shown in Fig. 2. The Reynolds stress due to the homogeneous contribution, given in (15a), is downgradient initially, at \( t_v = K_0 / \alpha K_1 \) it is zero, and for \( t > t_v \) it is upgradient. If this mechanism were operating alone, maximum energy density would occur at \( t_v \). Note that for
perturbations to the perturbation growth in the linear limit. The initial conditions and the value of the spanwise wave number $K_{03}$ are the same as in Fig. 1 ($\theta_{01} = 1$, $\theta_{02} = 0$, $K_{01} = 1$, $K_{03} = 10$). Note that $K_{01} = 0$ corresponds to the 2-D Orr case for which the energy maximum is reached at $t = 10$. For any $K_{03} \neq 0$ the energy eventually asymptotes in the inviscid limit to a finite constant. As $K_{03} \rightarrow \infty$ the energy as $t \rightarrow \infty$ diverges if viscosity is disregarded. In the presence of viscosity the energy as $t \rightarrow \infty$ vanishes.

For the general case the energy density growth of a single plane wave is given by

$$E(t) = \frac{|\partial_{\theta_{01}}|^2 + (K_{01}^2 + K_{03}^2)(t))|}{|\partial_{\theta_{01}}|^2 + K_{02}^2 |\partial_{\theta_{02}}|^2} e^{-2\gamma t}, \quad (16)$$

where $E_0$ is the initial energy density. This energy density growth is shown as a function of $K_{03}$ in Fig. 3 for the inviscid case. We note that as the lift-up mechanism associated with the inhomogeneous contribution to the Reynolds stress becomes important the energy density growth at large times is significantly larger than the energy density growth arising from the Orr mechanism operating alone.

### III. DETERMINATION OF THE OPTIMAL PERTURBATIONS

We have described the evolution of a plane wave perturbation in an infinite shear flow. Development was found to depend on the initial cross-stream velocity and vorticity. In a stochastically forced flow it is likely that for a given initial energy density all initial conditions are sampled, and consequently it is of interest to determine which of these initial conditions yield greatest energy growth in a specified time. The initial perturbations that maximize energy growth at a specific time, $T_{\text{opt}}$, and perhaps under specified additional constraints such as on the shape or the wavelength of the perturbation, will be called optimal perturbations. These optimal perturbations are the primary contributors to the perturbation growth in the linear limit. The success of rapid distortion theory in fully turbulent flow suggests that the linear approximation retains validity in turbulent flow.

The plane wave solutions are orthogonal in the inner product associated with either the $L_2$ or the energy norm. This orthogonality of the nonseparable plane waves should not be confused with the nonorthogonality of the normal mode solutions, which is necessary for the modal solutions to produce transient growth.\(^{25}\) A general initial perturbation can be considered as a Fourier synthesis of plane waves, and, because of the orthogonality property, determination of the optimal perturbation reduces to a search for the optimal plane wave. This optimal will be called the general optimal.

Note that the uniform shear flow has no intrinsic space scale in the inviscid limit. The expression for the energy growth (16) in this limit is therefore scale invariant. Consequently, if we do not limit the size of the perturbations, it can be verified that the optimal plane wave in the presence of viscosity will assume a large enough scale that the effects of viscosity can be disregarded for finite $T_{\text{opt}}$. We will start by confining our search to finding the inviscid optimals recognizing that viscosity will lead to the eventual decay of the perturbation.

The optimal plane wave was found by means of a downhill simplex method on the initial wave numbers, vorticity, and vertical velocity, i.e., given $T_{\text{opt}}$ find $K_{01}$, $K_{02}$, $K_{03}$, $\theta_{01}$, $\theta_{02}$, $\theta_{03}$, so that

$$E(T_{\text{opt}}; K_{01}, K_{02}, K_{03}, \theta_{01}, \theta_{02}, \theta_{03})$$

is maximized. The resulting energy density growth as a function of optimization time is shown in Fig. 4. Also, the initial orientation of the wave as a function of optimizing time is shown in Figs. 5 and 6. Note the monotonic in-
crease of the energy amplification with $T_{\text{opt}}$ shown in Fig. 4. The ratio of the spanwise to streamwise wave number, $k_{03}/k_{01} \approx 2$, indicates that the optimal initial perturbations are elongated in the streamwise direction by this factor. The cross-stream inclination of the plane wave is defined as the angle between the direction of the mean flow and the line of constant phase projected on the $(x,y)$ plane, i.e., $\phi = \tan^{-1} k_{02}/k_{01}$. The initial cross-stream wave number, $k_{02}$, that corresponds to the optimal perturbations is such that the plane wave assumes a cross-stream orientation ($\phi = \pi/2$) at a time $t_c < T_{\text{opt}}$ in order to benefit from the Orr intensification of the cross-stream velocity. This value of optimal $k_{02}$ arises because of the synergism between the tilting and the Orr mechanism, which is a general characteristic of the 3-D optimals. By contrast, in the 2-D case, where intensification of streamwise velocity by the tilting mechanism is absent and the Orr mechanism alone produces growth, $t_c > T_{\text{opt}}$.

It is of interest to further compare the full 3-D optima, for which both the tilting and the Orr mechanism operate in combination, with the 2-D optima, for which only one mechanism is operating. The optimal energy amplification of such purely 2-D cases is also presented in Fig. 4: for $k_{03}=0$, for which the tilting mechanism does not operate, and for $k_{01}=0$, the streamwise rolls for which the Orr mechanism does not operate. For both, due to algebraic coincidence, the energy density amplification of the optimals at $T_{\text{opt}}$ is

$$\frac{E(T_{\text{opt}})}{E(0)} = \frac{2(1+T_{\text{opt}}^2/4)^{1/2} + T_{\text{opt}}}{2(1+T_{\text{opt}}^2/4)^{1/2} - T_{\text{opt}}} ,$$

which for large optimizing times gives energy density amplification: $E(T_{\text{opt}})/E(0) \approx T_{\text{opt}}^2$. This relation is a consequence of the unboundedness of the flow that allows the perturbations for every $T_{\text{opt}}$ to assume large enough scale to evolve inviscidly during the growth stage. In contrast, the presence of boundaries sets a limit on the scale of the perturbations leading to a splitting for large $T_{\text{opt}}$ of the maximum growth attained by the Orr perturbations, constrained to $k_{03}=0$, and that attained by the streamwise rolls with $k_{01}=0$. To understand this splitting of the maximum growth attained by the structures in the presence of boundaries, consider the characteristics of the initial optimal configuration in the unbounded flow in these limits.

When $k_{03}=0$ the optimal plane wave has initial orientation:

$$k_{02} = \frac{T_{\text{opt}}}{2} + \left(1 + \frac{T_{\text{opt}}^2}{4}\right)^{1/2} ,$$

under the assumption that $(k_{02}^2 + k_{01}^2)^{1/2}$ is sufficiently small so that the cumulative dissipation factor $g$ in (7) can be neglected up to times $t=O(T_{\text{opt}})$. This condition can certainly be satisfied when the flow is unbounded, but not when $k_{02} / \sigma \geq D$ is enforced by the presence of boundaries $D$ units of distance apart. The maximum growth attained under this circumstance is consequently reduced.

The optimal initial excitation of the streamwise rolls occurs for

$$\omega = \frac{i}{2} \left[ T_{\text{opt}} - \left(1 + \frac{k_{02}^2}{k_{03}^2} + \frac{T_{\text{opt}}^2}{4}\right)^{1/2} \right] ,$$

and $k_{03}=0$ in the absence of boundaries. We find that this initial condition leads to reduced dissipation when boundary constraints are introduced compared to its counterpart in (18). Consequently, the streamwise rolls dominate the growth for large $T_{\text{opt}}$ in the presence of boundaries consistently with the numerical results of Butler and Farrell. Note in Fig. 4 that the energy growth attained by the 3-D optimal is always larger than the growth of the 2-D optimals, this difference increasing with $T_{\text{opt}}$. This is also consistent with the channel flow results of Butler and Farrell, who find that the global optima are associated with general 3-D disturbances.

These single plane waves, although general optimals, suffer from the disadvantage that they assume spatially unbounded initial excitation. To study the affect of limited spatial excitation we consider, with Taylor and Green, optimals constrained to be initially of the checkerboard form: $\cos(k_{01}x) \cos(k_{02}y) \cos(k_{03}z)$. The energy density amplification as a function of $T_{\text{opt}}$ for these bounded perturbations is shown in Fig. 4. Note that, in general, the
energy amplification of the checkerboard excitation is larger than that of 2-D single wave optimals but less than that of the general optimals. Again, the energy growth increases with $T_{opt}$.

With viscosity introduced we can define a Reynolds number, $R = \alpha l^2/\nu$, based on the scale, $l$, of the initial perturbation. In the presence of viscosity the monotonic increase of optimal energy growth with $T_{opt}$ (cf. Fig. 4) does not persist for large values of $T_{opt}$ because for large optimal times viscosity eventually damps the perturbation. Consequently, there will be an initial optimal perturbation and a $T_{max}$ for which maximum growth is attained for each $R$. This $T_{max}$ and the associated energy growth is a measure of the greatest possible growth for any perturbation for the specific $R$. For checkerboard initial perturbations with $R = 100, 1000, 10000$, maximum growth is achieved at $T_{opt} = 7, 15, 30$ advective units, respectively, with corresponding maximum energy growth of $E(T_{opt})/E(0) = 12.5, 109, 707$ (Fig. 7), approximately increasing as $R^2$. However, it can be shown using the above definition of $R$, that the minimum $R$ for which some perturbation growth can occur in the unbounded constant shear flow is $R = 19.7$. This result is interesting because it identifies the lowest Reynolds number for which growth can occur in the absence of boundary constraints (Appendix A).

In real flows ambient turbulent fluctuations provide a time scale that intercepts the growth of the perturbations by disrupting their coherent motion. This time scale is the eddy turnover time, $T_{eddy}$. Because of the monotonic increase of energy growth for small $T_{opt}$, the maximal growth of perturbations is attained for $T_{opt} \approx T_{eddy}$, which may be considerably smaller than $T_{max}$. Dimensional analysis provides an estimate of $T_{eddy}$ in the inertial subrange. An eddy of characteristic size $l$ has $T_{eddy} \approx \epsilon l / \epsilon^{3/2} \nu^{3/2}$ (with $\epsilon$ the rate of eddy energy dissipation), indicating the choice of a smaller $T_{opt}$ for smaller eddy scales $l$. When the inertial subrange comprises three decades of wave numbers, the expected $T_{opt}$ varies by a factor of 100 over the subrange, thus allowing for a range of optimal energy amplification of the order of $10^4$ between the small- and large-scale structures. In general, we do not have a priori knowledge of $T_{eddy}$, so that it has to be determined from observation. Typical of many experimental and numerical simulations is an eddy turnover time $O(10)^{24\theta^2}$ advective time units; consequently, we will choose $T_{opt} = 10$ for the examples to follow. The checkerboard optimal energy density growth as a function of $T_{opt}$ for $R = 1000$ and $R = 10000$ based on a disturbance with $\lambda = 1$. The abscissa is $\lambda = (k_{10}^2 + k_{20}^2)^{1/2}$ and the ordinate is $\Theta = \tan^{-1} k_{03}/k_{01}$. Note that the maximal growth occurs for $\Theta = 63^\circ$, that for small wave numbers viscosity does not affect the growth attained, and that for larger wave numbers viscosity affects least the structures neighboring the streamwise rolls ($\Theta \approx 90^\circ$).

**FIG. 7.** Evolution of energy density with viscosity for optimal checkerboard initial conditions. The optimizing times, $T_{opt}$, shown yield the maximal energy growth for the specific Reynolds number, $R$, where the Reynolds number is evaluated based on the streamwise wave number $k_{01} = 1$. For $R = 100$ maximal growth is attained for $T_{opt} = 7$. For $R = 1000$ maximal growth is attained for $T_{opt} = 15$, and for $R = 10000$ maximal growth is attained for $T_{opt} = 30$. The corresponding energy density growth is 12.5, 109, and 707, respectively.

**FIG. 8.** Contour plot of energy growth of checkerboard optimal perturbations as a function of the streamwise and spanwise wave number for $T_{opt} = 10$ and $R = 1000$ based on a disturbance with $\lambda = 1$. The abscissa is $\lambda = (k_{10}^2 + k_{20}^2)^{1/2}$ and the ordinate is $\Theta = \tan^{-1} k_{03}/k_{01}$. Note that the maximal growth occurs for $\Theta = 63^\circ$, that for small wave numbers viscosity does not affect the growth attained, and that for larger wave numbers viscosity affects least the structures neighboring the streamwise rolls ($\Theta \approx 90^\circ$).
FIG. 9. Evolution of the checkerboard optimal perturbation for $T_{opt}=10$ and $R=1000$. The volume containing perturbation energy greater than 75% of the maximum energy is displayed for different times. The fluid volume presented is one wavelength long in the streamwise direction ($x$), one wavelength in the spanwise direction ($z$), and the length of the cross-stream axis ($y$) is equal to that of the streamwise. The initial wave numbers of the optimal perturbation are $\kappa_2 = 0.44$, $\kappa_3 = 3.55$, $\kappa_1 = 0.9$. Panel (a) is for $t=0$, the maximum pointwise energy is normalized to 1. Panel (b) is for $t=5$, the maximum pointwise energy is 6.3. Panel (c) is for $t=10$, the maximum pointwise energy is 110. Panel (d) is for $t=15$, the maximum pointwise energy is 164. The disturbance decays for $t>15$.

Also note that viscosity damps least the roll-type solutions ($\kappa_2/\kappa_1 \to \infty$), consistent with the observation that in the viscous sublayer the coherent structures preferentially assume the form of streamwise rolls. Recently, Farrell and Ioannou compared this maximum growth achieved in the unbounded flow as a function of the spanwise and cross-stream wave number to that found by numerical solution in channel flows. It was found that the maximum growth spectra are similar.

While there is a large subspace of substantially growing perturbations, the morphology of their evolution is quite restricted. The evolution of a typical perturbation is shown in Fig. 9, where the volume containing perturbation energy greater than 75% of the maximum energy at each time is displayed. Note that the initially localized energy is sheared to produce structures of a double roller orientation. This structure is verified in the $x=0$ plane vector velocity field shown at $T_{opt}=10$ in Fig. 10. These double rollers generate energetic localized streamwise streaks that only slowly decay. This universal morphology, commonly observed in turbulent flows, is shown here to arise from the 3-D evolution of optimal perturbations.
The observation that transition to turbulence in shear flow can occur with very small perturbations and that the critical Reynolds number for transition increases monotonically with perturbation variance argues for a linear mechanism, underlying at least the initial stage of the transition process.  Similarly, the fact that rapid distortion theory accurately produces the velocity correlation functions observed in shear turbulence, which reveal the underlying coherent structures, also suggests the first-order validity of linear theory in that problem.  However, absence of modal instability in the canonical channel flow problems (excepting plane Poiseuille for $R > 5772$) has discouraged application of linear theory and encouraged development of alternative nonlinear and secondary instability theories.  Nevertheless, reanalysis of the linear problem demonstrated the potential for large transient growth of appropriately configured perturbations.

Because naturally occurring disturbances are not likely to assume the form of a single eigenmode, assessment of the growth potential of a general perturbation is important.  The least stable mode growth rate determines stability in the limit $t \to \infty$ for the linear problem, but this limit is not necessarily appropriate to many physical situations, even for flows that support modal instabilities because of neglected effects such as disruption of the eigenfunction on an eddy turnover time scale or limitations on the physical size of an experimental apparatus, which require that the mode also be absolutely unstable (have zero group velocity), in addition to being temporally unstable.  On the other hand, initial growth as determined by energy methods, or equivalently by the method of this work in the limit $T_{opt} \to 0$, reveals only the instantaneous initial tendency without determining the potential for finite growth.  The method of optimal growth includes both these limits, but, in addition, determines the potential for transient development and also the set of perturbations that accomplish this growth, given an appropriate development time and space scale from physical considerations for a given problem.

Previous work on optimal growth made use of the modal decomposition of the dynamical operator, $A$, for the associated dynamical system, $d\psi/dt = A\psi$, in which the possibility for transient growth depends on the non-normality of $A$ and the nonorthogonality of its eigenvectors in an appropriate inner product, generally that associated with energy.  While this method of analysis is general to all dynamical systems of the above form, in light of the universality of transition at subcritical Reynolds numbers and of the generality of coherent structures across shear flow problems we have chosen to adopt the unbounded constant shear problem for which analytic solutions exist.

An important result of previous work, verified here, is that the growth of 3-D optimals greatly exceeds the growth of 2-D optimals.  This serves to explain the observed much greater instability of 3-D shear flows.

We find that in limiting cases the optimal perturbations consist of either the Orr structure or the streamwise roll structure previously identified.  However, in general, the optimal perturbation combines these mechanisms in a synergistic manner in which the cross-stream velocity produced by the Orr mechanism enhances the streamwise rolls associated with streak production.  The resulting optimal structures are found to resemble the double roller eddies observed by Townsend.

There is no intrinsic scale imposed on the optimal structures in the limit of large Reynolds number based on the size of the perturbations and the shear and as a result the structures are scale invariant—the same intrinsic structure can undergo development at any scale between that effected by viscosity and the largest scale excited, or alternatively the largest scale not strongly disrupted by the background turbulent field over its growth.  It follows that the set of optimals is densely distributed in the field of perturbations, unlike the case of inflectional instabilities that select a restricted subset of unstable modes typically forming a highly restricted growing subspace of the perturbation field.

While modal instability proceeds from arbitrarily small initial perturbations, transient growth, in general, produces large but ultimately bounded growth.  It follows that the nature of the perturbation field has a more central role in transient development of optimal perturbations than it does in modal instability.  However, the optimal perturbations form a complete orthogonal set so that any perturbation field can be decomposed into contributions from optimal elements.  It follows that optimal growth structures dominate development from any perturbation field not contrived to exclude these growing perturbations and that optimal structures are likely to be universal features in observations of perturbed flows.

It is clear in the case of transition to turbulence in
shear flows that the perturbation field is essentially exogenous to the flow because the level of background perturbation controls the transition Reynolds number.\(^5,31\) However, in the case of fully developed shear turbulence, excitation of optimal structures must be endogenous to the flow, at least at sufficient distance from boundary contributions to the perturbation field. The great potential for growth found for optimal structures argues that the role of nonlinearity in scattering energy back into growing structures could be fairly weak, and yet the feedback could still be sufficient to transform the amplification into a self-sustaining oscillation maintaining turbulence in the manner discussed by Farrell.\(^5\) An explicit demonstration of this mechanism has been made by Schmidt and Henningson\(^34\) and Trefethen et al.\(^35\) Regardless of the success of such models, a particular mechanism underlying regeneration of optimal structures more specific than the appeal to "non-linearity" remains to be identified. For instance, it may be that the chance juxtaposition of the perturbation debris from previous growth episodes advected by the shear results in the local occurrence of near optimal structures and attendant bursts of perturbation energetics, a mechanism well known to give rise to highly intermittent cyclone formation in the atmosphere.\(^36,37\)

Optimal excitation theory, developed here for the simplest shear flow problem, provides a unified method for analysis of shear stability, including the potential for growth in the limits \(t \to 0\) and \(t \to \infty\), as well as in the physically important finite time domain, where large growth occurs, even in the absence of unstable normal modes. These perturbations identified as optimal in energy growth resemble observed recurrent structures in shear flow turbulence.

**ACKNOWLEDGMENTS**

We thank an anonymous referee for his comments, which led to improvement of our presentation.

Brian Farrell was supported by National Science Foundation Contract No. ATM-8912432. Petros Ioannou was supported by National Science Foundation Contract No. ATM-9216189. Computer time was provided by NCAR Contract No. 3512031; the National Center for Atmospheric Research is supported by the National Science Foundation.

**APPENDIX A: MINIMUM REYNOLDS NUMBER NECESSARY FOR PERTURBATION GROWTH**

We have shown that in unbounded shear flow single plane wave perturbations constrained to vary in the spanwise cross-stream plane \((y-z)\) plane, i.e., with \(\kappa_{01}=0\), and those constrained in the streamwise cross-stream plane \((x-y)\) plane, i.e., with \(\kappa_{03}=0\), achieve identical maximal energy density growth. It was shown that this result stems from the absence of a geometrically imposed scale limiting perturbation size. In this circumstance the effect of viscosity during the growth period may be made negligible by sufficiently increasing the perturbation scale. This result does not follow when the viscous effects necessarily enter, as in the case of bounded channel flows. When viscous effects are important the streamwise rolls grow more than perturbations constrained in the \(x-y\) plane. The increased maximum growth of streamwise rolls can be illustrated by determining the minimum Reynolds number for which growth is possible in an unbounded flow.

To consider the viscous effects in an unbounded flow we nondimensionalize time by \(1/\alpha\), where \(\alpha\) is the shear of the background flow and length by \(l=\pi/\lambda\), where \(\lambda\) is the initial total horizontal wave number. We can then define the Reynolds number as \(R=\alpha R^2/\nu\).

First consider perturbations with \(\kappa_{03}=0\). From (16) and (10) we obtain energy density growth \(G=\dot{E}(t)/E_0:\)

\[
G=\frac{1+\kappa_{02}^2}{1+(\kappa_{02}^2-b^2)} \exp\left(-\frac{2\pi^2}{R} \int_0^t dt' [1+(\kappa_{02}-\tau)^2]\right),
\]

where the variables are nondimensional and \(R\) is based on the cross-stream wave number \(\kappa_0\). At \(t=0\) we have

\[
\frac{d\ln G}{dt} = -2\pi^2 \frac{1+\kappa_{02}^2}{R} + \frac{2\kappa_{02}}{1+\kappa_{02}}.
\]

The minimum \(R_{\text{min}}\) for which growth is possible is given by

\[
R_{\text{min}} = \min_{\kappa_{02} > 0} \left(\frac{\pi(1+\kappa_{02}^2)}{\kappa_{02}}\right).
\]

The minimum is attained for a perturbation with \(\kappa_{02} = 3^{-1/2}\), leading to \(R_{\text{min}} = 3.08 \pi^2 = 30.4\). Orr et al \(27\) calculated for arbitrary 2-D perturbations meeting boundary conditions in a Couette flow \(R_{\text{min}} = 44.3\), where the Reynolds number is based on the channel width. As expected, the less constrained unbounded constant shear flow has a lower stability threshold.

Consider perturbations with \(\kappa_{01}=0\). The energy density growth can be derived from (16) and (9) to be

\[
G=\frac{(\bar{\omega}+\tau)^2}{\bar{\omega}^2 + 1 + \kappa_{02}^2} \exp\left(-\frac{2\pi^2}{R} \tau(1+\kappa_{02}^2)\right),
\]

where \(\bar{\omega}\) is the amplitude of \(\omega_{02}\) and \(R\) is based on the spanwise wave number \(\kappa_{03}\). Note that in (23) we have considered perturbations that have initial cross-stream vorticity having a \(\pi/2\) phase lag from the initial cross-stream velocity, as it can be easily seen that such initial conditions are favored. At \(t=0\) we obtain

\[
\frac{d\ln G}{dt} = -2\pi^2 \frac{1+\kappa_{02}^2}{R} + \frac{2\bar{\omega}}{1+\bar{\omega}^2 + \kappa_{02}^2}.
\]

The minimum \(R_{\text{min}}\) for which perturbation energy growth is possible is given by

\[
R_{\text{min}} = \min_{\bar{\omega}, \kappa_{02}} \left(\frac{\pi(1+\kappa_{02}^2)}{\bar{\omega}}\right).
\]

The minimum is attained for perturbations with \(\bar{\omega} = 1\) and \(\kappa_{02} = 0\). The minimum is \(R_{\text{min}} = 2\pi ^2 = 19.7\). Note that the stability threshold for streamwise rolls occurs for a lower value of \(R\) than it does for perturbations with \(\kappa_{03} = 0\). The critical \(R = 19.7\) for streamwise rolls in an unbounded flow.
is also smaller than the corresponding critical $R = 20.7$ calculated by Joseph for plane Couette flow, and $R = 49.6$ calculated by Busse and Joseph and Carmi for plane Poiseuille flow. These critical Reynolds numbers demonstrate the potential for growth of nonmodal perturbations.
