Hydrodynamic Stability Without Eigenvalues

Lloyd N. Trefethen, Anne E. Trefethen, Satish C. Reddy, Tobin A. Driscoll

Fluid flows that are smooth at low speeds become unstable and then turbulent at higher speeds. This phenomenon has traditionally been investigated by linearizing the equations of flow and testing for unstable eigenvalues of the linearized problem, but the results of such investigations agree poorly in many cases with experiments. Nevertheless, linear effects play a central role in hydrodynamic instability. A reconciliation of these findings with the traditional analysis is presented based on the "pseudospectra" of the linearized problem, which imply that small perturbations to the smooth flow may be amplified by factors on the order of 10^5 by a linear mechanism even though all the eigenmodes decay monotonically. The methods suggested here apply also to other problems in the mathematical sciences that involve nonorthogonal eigenfunctions.

Hydrodynamic stability theory is the study of how laminar fluid flows become unstable, the precursor to turbulence. It is well known that turbulence is an unsolved problem, but not so well known that despite the efforts of generations of applied mathematicians, beginning with Kelvin, Rayleigh, and Reynolds, many of the presumably simpler phenomena of hydrodynamic stability also remain incompletely understood (*1*, *2*).

The traditional starting point of an investigation of hydrodynamic stability is eigenvalue analysis, which proceeds in two stages: (i) linearize about the laminar solution and then (ii) look for unstable eigenvalues of the linearized problem. An "unstable eigenvalue" is an eigenvalue in the complex upper half-plane, corresponding to an eigenmode of the linearized problem that grows exponentially as a function of time t. It is natural to expect that a flow will behave unstably if and only if there exists such a growing eigenmode, and over the years much has been learned about which flows possess such modes, a distinction that depends on the geometry, the Reynolds number, and sometimes other parameters.

For some flows, notably those with instabilities driven by thermal or centrifugal forces, the predictions of eigenvalue analysis match laboratory experiments. Examples are Rayleigh-Bénard convection (a stationary fluid heated from below) and Taylor-Couette flow (between a stationary outer and a rotating inner cylinder). For other flows, notably those driven by shear forces, the predictions of eigenvalue analysis fail to match most experiments. We consider the two most studied examples of this kind: (plane) Couette flow, the flow with a linear velocity profile between two infinite flat plates moving parallel to one another, and (plane) Poiseuille flow, the flow with a parabolic velocity profile between two stationary plates (Fig. 1). Other examples for which eigenvalue analysis fails include pipe Poiseuille flow (in a cylindrical pipe) and, to a lesser degree, Blasius boundary layer flow (near a flat wall).

For Poiseuille flow, eigenvalue analysis predicts a critical Reynolds number $R = 5772$ at which instability should first occur (*3*), but in the laboratory, transition to turbulence is observed at Reynolds numbers as low as $R \approx 1000$ (*4*). For Couette flow, eigenvalue analysis predicts stability for all R (*5*), but transition is observed for Reynolds numbers as low as $R \approx 350$ (*6*). These anomalies of "subcritical transition to turbulence" have been recognized for many years, and the explanation has traditionally been attributed to step (i) above. If linearization has failed, the reasoning goes, one must look more closely at the nonlinear terms or perhaps linearize about a solution other than the laminar one [the theory of "secondary instability" (*7*–*9*)].

Recently it has emerged, however, that the failure of eigenvalue analysis may more justly be attributable to step (ii). It is a fact of linear algebra that even if all of the eigenvalues of a linear system are distinct and lie well inside the lower half-plane, inputs to that system may be amplified by arbitrarily large factors if the eigenfunctions are not orthogonal to one another. A matrix or operator whose eigenfunctions are orthogonal is said to be "normal" (*10*), and the linear operators that arise in the Bénard and Taylor-Couette problems are in this category. By contrast, Reddy *et al.* (*11*) discovered in 1990 that the operators that arise in Poiseuille and Couette flow are in a sense exponentially far from normal. At about the same time, the startling discovery was made by Gustavsson (*12*), Henningson *et al.* (*13*), and Butler and Farrell (*14*) that small perturbations to these flows may be amplified by factors of many thousands, even when all the eigenvalues are in the lower half-plane (Fig. 2). The elegant paper by Butler and Farrell discusses many details omitted here and, together with a more recent paper by Reddy and Henningson (*15*), forms the foundation of the present work (*16*).

The shaded region in Fig. 2 has appeared in many publications (*17*) and corresponds to parameters for which unstable eigenmodes exist. The contours outside the shad-

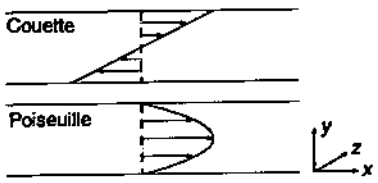

Fig. 1. Velocity profiles for two laminar flows (independent of x and z). The geometry is an infinite 3D slab of viscous incompressible fluid bounded by parallel walls. The laminar solutions satisfy the Navier-Stokes equations for all Reynolds numbers, but for higher R, the flows are unstable and rapidly become turbulent.

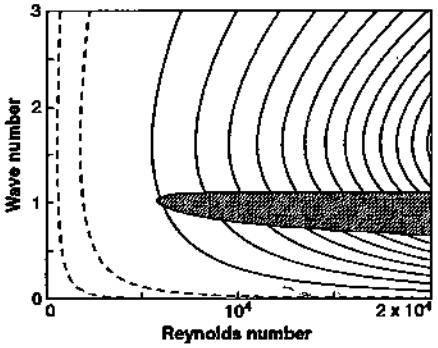

Fig. 2. Maximal resonant amplification of 3D perturbations in linearized Poiseuille flow as a function of Reynolds number R and xz wavenumber magnitude $k = \sqrt{\alpha^2 + \beta^2}$ (Eq. 6). In the shaded region, with leftmost point $R = 5772$, unstable eigenmodes exist and unbounded amplification is possible. The contours outside that region, from outer to inner, correspond to finite amplification factors of 10^3, 10^4 (dashed), 10^5, 2×10^5, . . ., 1.3×10^6. For example, amplification by a factor of 1000 is possible for all $R \geq 549$. In the laboratory, transition to turbulence is observed at $R \approx 1000$. The analogous picture for Couette flow looks qualitatively similar except that there is no shaded region.

L. N. Trefethen is in the Department of Computer Science, Cornell University, Ithaca, NY 14853 (lnt@cs.cornell.edu). A. E. Trefethen is in the Cornell Theory Center, Cornell University, Ithaca, NY 14853. S. C. Reddy is in the Courant Institute of Mathematical Sciences, New York University, New York, NY 10012. T. A. Driscoll is in the Center for Applied Mathematics, Cornell University, Ithaca, NY 14853.

ed region quantify the nonmodal amplification that may occur in these flows. The possibility of amplification of perturbations of viscous flows by nonmodal linear mechanisms has been recognized for a century (*18*), but until the recent developments, it was not known that the magnitudes involved were huge.

An essential feature of this nonmodal amplification is that it applies to three-dimensional (3D) perturbations of the laminar flow field (*19*). In much of the literature of hydrodynamic stability, attention has been restricted to 2D (xy) perturbations; in particular, the well-known Orr-Sommerfeld equation is an eigenvalue equation for 2D perturbations. A principal justification for this restriction has been Squire's theorem, which asserts that if a flow has an unstable 3D eigenmode for some R, then it has an unstable 2D eigenmode for some lower value of R (*20*). The new results indicate that this emphasis on 2D perturbations has been misplaced. When only 2D perturbations are considered, some amplification can still occur, but it is far weaker.

The growing attention to 3D linear, nonmodal phenomena represents a significant change in the traditional conception of problems of hydrodynamic stability.

Streamwise Vortices and Streaks

The flow features associated with this amplification process have a distinctive form: A perturbation to the velocity field in the form of a "streamwise vortex" evolves into a higher amplitude "streamwise streak" (Fig. 3). A streamwise vortex is an elongated region of vorticity approximately aligned with the x axis, and a streamwise streak is an elongated region of high or low velocity (relative to the mean flow) approximately in the x direction. Streamwise vortices and streaks are persistent features in laboratory experiments involving all kinds of internal and boundary layer shear flows (*21*). Physically, they are not hard to explain: In a shear flow, a small perturbation in the form of a streamwise vortex may move fluid from a region of higher to lower x velocity, or vice versa, where it will appear as a large, local perturbation in the x velocity (*22*). Because these features constitute 3D perturbations of the flow field, however, their prevalence has been difficult to reconcile with the predictions of eigenvalue analysis. Nonmodal analysis offers a linear explanation of why these structures are so common, for although streamwise streaks are not eigenmodes of the linearized flow problem, they are pseudomodes.

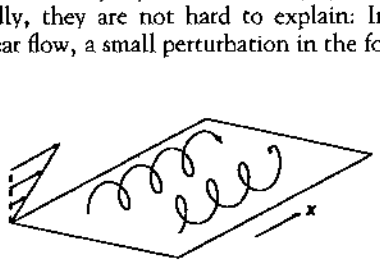

Fig. 3. Schematic illustration of a structure that appears in many shear flows. A low-amplitude vortex or counterrotating pair of vortices approximately aligned with the flow entrains fluid from regions of high to low x velocity and vice versa. The entrained fluid appears as a streak of locally high or low streamwise velocity (not shown). This process is linear but unrelated to eigenmodes.

To explain this term, we must define the evolution operator that is the mathematical basis of this article. Let $\mathbf{u}_0 = \mathbf{u}_0(x,y,z)$ denote the vector velocity field corresponding to Poiseuille or Couette laminar flow. Let $\mathbf{u}_0 + \mathbf{U}(t) = \mathbf{u}_0(x,y,z) + \mathbf{U}(x,y,z,t)$ be the velocity field of a slightly perturbed flow, that is, a nearby solution to the Navier-Stokes equations (uppercase letters distinguish quantities that are functions of t). If we take $\mathbf{U}$ to be infinitesimal, then it satisfies an equation

$$\frac{d\mathbf{U}}{dt}(t) = -i\mathcal{L}\mathbf{U}(t) \tag{1}$$

where $\mathcal{L}$ is a linear operator that we call the linearized Navier-Stokes evolution operator. As a measure of the size of solutions to this equation, we define

$$\|\mathbf{U}(t)\| = \left(\int |\mathbf{U}(x,y,z,t)|^2 dx\, dy\, dz \right)^{1/2} \tag{2}$$

which we call the energy norm because its square is sometimes interpreted as an energy.

Suppose the linearized fluid system is driven by a signal of the form $\mathbf{V}(t) = e^{-i\omega t}\mathbf{v}$ for some frequency $\omega \in \mathbb{C}$ ($\mathbb{C}$ denotes the set of complex numbers) and function $\mathbf{v} = \mathbf{v}(x,y,z)$

$$\frac{d\mathbf{U}}{dt}(t) = -i\mathcal{L}\mathbf{U}(t) + e^{-i\omega t}\mathbf{v} \text{ for } -\infty < t < \infty \tag{3}$$

It is easily verified that the response will be

$$\mathbf{U}(t) = ie^{-i\omega t}\mathbf{u} \text{ where } \mathbf{u} = (\omega\mathcal{I} - \mathcal{L})^{-1}\mathbf{v} \tag{4}$$

($\mathcal{I}$ denotes the identity operator). Thus, the operator $(\omega\mathcal{I} - \mathcal{L})^{-1}$, known as the resolvent of $\mathcal{L}$, transforms "inputs" $\mathbf{v}$ to the linearized fluid flow at frequency ω into corresponding "outputs" $\mathbf{u}$. The degree of amplification that may occur in the process is equal to the operator norm

$$\|(\omega\mathcal{I} - \mathcal{L})^{-1}\| \equiv \sup_{\mathbf{v}\neq 0} \frac{\|\mathbf{u}\|}{\|\mathbf{v}\|} \tag{5}$$

(sup denotes supremum or maximum). Because an arbitrary time-dependent perturbation of the laminar flow can be reduced to an integral over real frequencies by Fourier analysis, values of ω on the real axis $\mathbb{R}$ are of particular interest. The maximum possible amplification over all real frequencies (Fig. 2) is

$$\sup_{\omega \in \mathbb{R}} \|(\omega\mathcal{I} - \mathcal{L})^{-1}\| \tag{6}$$

An eigenvalue of $\mathcal{L}$ is a number $\omega \in \mathbb{C}$ such that $\mathcal{L}\mathbf{u} = \omega\mathbf{u}$ for some corresponding eigenfunction $\mathbf{u}$. Equivalently, it is a number ω with the property that perturbations with frequency ω can be amplified unboundedly: $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\| = \infty$. Generalizing this definition, for any $\epsilon \geq 0$, an "ϵ-pseudoeigenvalue" of $\mathcal{L}$ is a number ω such that $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\| \geq \epsilon^{-1}$, and a corresponding "$\epsilon$-pseudomode" is any function $\mathbf{u}$ with $\|\mathcal{L}\mathbf{u} - \omega\mathbf{u}\| \leq \epsilon\|\mathbf{u}\|$. If ϵ is small, then an ϵ-pseudomode $\mathbf{u}$ may be excited to a substantial amplitude by a small input perturbation, possibly including noise in an experimental apparatus. At $R = 5000$, for example, a streamwise streak is an ϵ-pseudomode of the linearized Poiseuille flow problem for $\epsilon \approx 1.2 \times 10^{-5}$ and thus can be excited by a streamwise vortex five orders of magnitude weaker in amplitude.

The set of ϵ-pseudoeigenvalues of an operator is the "ϵ-pseudospectrum" (*23*)

$$\Lambda_\epsilon(\mathcal{L}) = \{\omega \in \mathbb{C}: \|(\omega\mathcal{I} - \mathcal{L})^{-1}\| \geq \epsilon^{-1}\} \tag{7}$$

The pseudospectra $\{\Lambda_\epsilon(\mathcal{L})\}$ form a nested family of sets in the complex plane, with $\Lambda_0(\mathcal{L})$ equal to the spectrum $\Lambda(\mathcal{L})$. If $\mathcal{L}$ is normal, $\Lambda_\epsilon(\mathcal{L})$ is the set of all points at distance $\leq\epsilon$ from $\Lambda(\mathcal{L})$, but in the nonnormal case, it may be much larger.

Spectra and Pseudospectra

We numerically calculated spectra and pseudospectra for Couette flow with $R = 350$ and 3,500 and for Poiseuille flow with $R = 1{,}000$ and 10,000 (*24*). These Reynolds numbers roughly span the range in each case from occasional turbulence (in some experiments) to unavoidable turbulence (even in experiments under the most carefully controlled conditions). To perform our calculations, we first Fourier transform in x and z, reducing the calculation to one space dimension (y) and two real parameters α and β (wave numbers in x and z). The determination of the pseudospectra then requires a minimization in the $\alpha\beta$ plane (*25*).

In the case of Couette flow, the spectrum of $\mathcal{L}$ is a continuous region contained in the lower half-plane (Fig. 4) (for each α-β pair, the spectrum is discrete; the union over α and β is a continuum). This is true for all R, corresponding to the unconditional stability of Couette flow according to eigenvalue analysis (*5*). For Poiseuille flow, the spectrum lies in the lower half-plane for small R, but as R increases, two bumps appear that cross into the upper half-plane at $R = 5772$ (Fig. 5). These

bumps represent the mode that has received most of the attention in the literature, known as a Tollmien-Schlichting or TS wave. Judging by the spectra alone, one would conclude that Couette and Poiseuille flows are fundamentally different because one has unstable eigenmodes for certain values of R and the other does not.

The pseudospectra tell a different story. For both Couette and Poiseuille flow, they protrude significantly into the upper half-plane for all reasonably large values of R, implying that the corresponding evolution processes must feature prominent nonmodal effects (*26*). Qualitatively, the pseudospectra in the Couette and Poiseuille cases look more alike than different, suggesting that these flows should behave similarly, as is observed in experiments (*27*).

One cannot see in Figs. 4 and 5 what values of α and β are associated with various points in the spectrum and pseudospectra. In fact, the upper boundaries of the spectra correspond to modes with $\beta = 0$, in keeping with Squire's theorem, whereas the upper parts of the pseudospectra correspond to pseudomodes with $\beta \neq 0$, indicating that the effects of nonnormality are predominantly 3D. The highest points of the pseudospectra correspond to $\alpha = \mathrm{Re}\omega = 0$, hence purely streamwise structures, but approximately streamwise features with nonzero α and $\mathrm{Re}\omega$ are also strongly amplified (*28*).

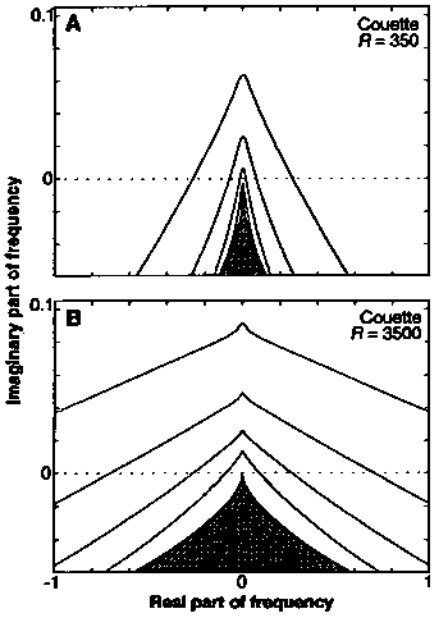

Fig. 4. Spectrum and ϵ-pseudospectra in the complex ω-plane of the linearized Navier-Stokes evolution operator for Couette flow. The spectrum is the shaded region, and the solid curves, from outer to inner, are the boundaries of the ϵ-pseudospectra for $\epsilon = 10^{-2}$, $10^{-2.5}$, 10^{-3}, and $10^{-3.5}$. The spectrum lies in the lower half-plane, but the pseudospectra extend significantly into the upper half-plane, revealing the nonnormality of this operator. Note that the real and imaginary axes are scaled differently.

Physical Implications

Pseudoresonance. One interpretation of nonnormality was described above in connection with Eq. 5. If a system is governed by a linear operator $\mathcal{L}$ that is normal or close to normal—familiar examples include musical instruments, vibrating structures, and molecules as described by quantum mechanics—then $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$ is large if and only if ω is close to an eigenvalue, and thus the frequencies at which the system resonates are determined by the spectrum. For a nonnormal system, however, $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$ may be large, and thus resonance or "pseudoresonance" may occur even when ω is far from the spectrum. A plot of pseudospectra can be interpreted as a plot of contours of equal resonance magnitude; the real axis is of particular interest because it corresponds to forcing at real frequencies.

Both Couette and Poiseuille flows exhibit strong 3D pseudoresonance for frequencies $\omega \approx 0$ (Fig. 6). The magnitude is $O(R^2)$, not $O(R)$ as one would have for a normal operator with spectra at the same distance $O(R^{-1})$ below the real axis. This and other asymptotic results of our calculations are summarized in Table 1.

Transient growth. A second interpretation of nonnormality involves the transient growth of flow perturbations that may evolve from certain initial conditions (*11–16*). Consider the initial-value problem

$$\frac{d\mathbf{U}}{dt}(t) = -i\mathcal{L}\mathbf{U}(t),\ t \geq 0,\ \mathbf{U}(0) = \mathbf{v} \quad (8)$$

The solution can be written $\mathbf{U}(t) = \exp(-it\mathcal{L})\mathbf{v}$, where $\exp(-it\mathcal{L})$ is the operator exponential. The factor by which such solutions can grow in time t is

$$\|\exp(-it\mathcal{L})\| = \sup_{\mathbf{v}\neq 0} \frac{\|\mathbf{U}(t)\|}{\|\mathbf{v}\|} \quad (9)$$

If $\mathcal{L}$ were a normal operator with spectrum in the lower half-plane, we would have $\|\exp(-it\mathcal{L})\| \leq 1$ for all $t \geq 0$, but in actuality, the growth is $\|\exp(-it\mathcal{L})\| = O(R)$ at times $t = O(R)$ (Fig. 7). This dependence on R can be explained as follows: The streamwise vortex-streak interaction is inviscid and operates on a time scale $O(R)$ before being shut off by the effects of viscosity (*13*). Such behavior is physically straightforward, appearing complicated only when interpreted in the basis of eigenmodes.

The $O(R^2)$ resonances of Figs. 2 and 6 are approximately equal to the integrals under the curves in Fig. 7. Mathematically, this can be seen from the formula

$$\mathcal{L}^{-1} = i\int_0^\infty \exp(-it\mathcal{L})dt \quad (10)$$

which implies

$$\|\mathcal{L}^{-1}\| \leq \int_0^\infty \|\exp(-it\mathcal{L})\|dt \quad (11)$$

In practice, this inequality is typically within a factor of 2 of equality. We can interpret the $O(R^2)$ figure physically by noting that the resonant amplification is a result of the combination of two effects: one (normal) factor $O(R)$ representing the time scale over which input energy can accumulate before it eventually decays, and another (nonnormal) factor $O(R)$ representing transient growth.

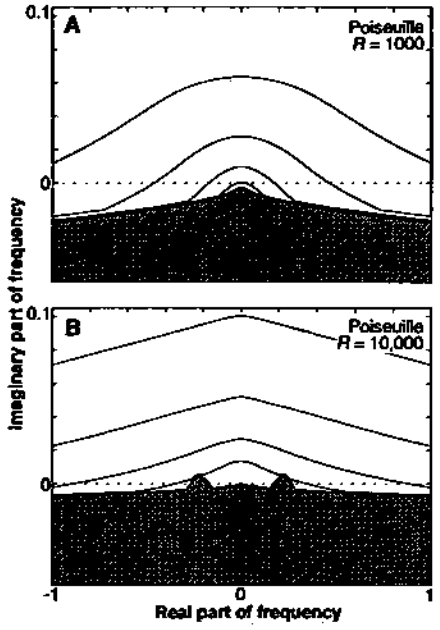

Fig. 5. Same as Fig. 4 but for Poiseuille flow. For $R > 5772$, two bumps in the spectrum extend into the upper half-plane.

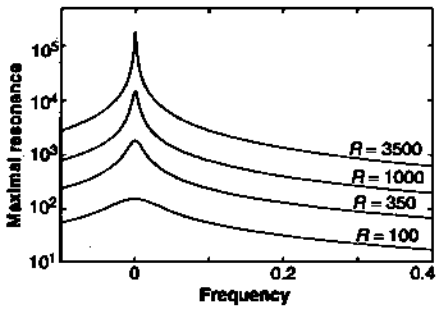

Fig. 6. Pseudoresonance curves for Couette flow at various Reynolds numbers. Each curve shows $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$, the maximum possible amplification of perturbations to the laminar flow field as a function of frequency ω. An analogous plot for Poiseuille flow looks qualitatively the same except that for $R > 5772$, an additional resonance of magnitude ∞ appears for $\omega \approx 0.2$.

A rigorous connection can be made between transient growth factors and the geometry of the pseudospectra. One can prove

$$\sup_{t\geq 0} \|\exp(-it\mathcal{L})\| \geq \sup_{\epsilon>0} \epsilon^{-1}\sigma_\epsilon(\mathcal{L}) \quad (12)$$

where $\sigma_\epsilon(\mathcal{L}) = \sup_{\omega\in\Lambda_\epsilon(\mathcal{L})} \mathrm{Im}\omega$ (*15*). In words, if the pseudospectra protrude far into the upper half-plane, then substantial transient growth must be possible. For the Poiseuille and Couette flows considered here, this estimate falls short of the actual growth by a factor ≈ 1.4.

Destabilizing perturbations. A third interpretation of nonnormality is based on an alternative, equivalent definition of the pseudospectra of an operator $\mathcal{L}$

$$\Lambda_\epsilon(\mathcal{L}) = \text{closure}\{\omega \in \mathbb{C}\colon \omega \in \Lambda(\mathcal{L} + \mathcal{E}) \text{ for some } \mathcal{E} \text{ with } \|\mathcal{E}\| \leq \epsilon\} \quad (13)$$

In words, the ϵ-pseudospectrum of $\mathcal{L}$ is the union of the spectra of all perturbed operators $\mathcal{L} + \mathcal{E}$ with $\|\mathcal{E}\| \leq \epsilon$, together with any limit points of this set (a technicality of little importance). It is interesting to reconsider Figs. 4 and 5 in the light of this alternative characterization. The pseudospectral contours in these figures imply that although the flows in question are eigenvalue stable (for $R < 5772$ in the Poiseuille case), exceedingly small perturbations to the evolution operator, of order $O(R^{-2})$, suffice to make them eigenvalue unstable (Fig. 8). The norm of the minimal destabilizing perturbation is $\epsilon_{min} = (\sup_{\omega\in\mathbb{R}}\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|)^{-1}$, the inverse of the maximal resonance of Eq. 6. For example, although Poiseuille flow with $R = 5000$ is eigenvalue stable, there exists a perturbation $\mathcal{E}$ of norm 1.2×10^{-5} that renders it unstable.

This raises the question of the physical meaning of operator perturbations and their relevance to instabilities observed in the laboratory. The minimal destabilizing perturbation $\mathcal{E}$ with $\|\mathcal{E}\| = \epsilon_{min}$ is easily characterized: It is the rank-1 operator $\epsilon_{min}\mathbf{v}\mathbf{u}^*$, where the asterisk denotes conjugate transpose and $\mathbf{v}$ and $\mathbf{u}$ are the functions that achieve the supremum in Eq. 5 for $\omega = 0$ [in matrix terminology, $\mathbf{u}$ and $\mathbf{v}$ are the principal singular vectors in the singular value decomposition (SVD) of $\mathcal{L}$ (*29*); the phrase "conjugate transpose" is also from matrix algebra]. This operator transforms streamwise streaks into streamwise vortices, closing the loop so that transient growth for finite t can feed back to become modal growth for all t. Of course, there is no reason to expect such a perfectly contrived perturbation to arise under natural conditions; yet our calculations show that even random perturbations of $\mathcal{L}$ often have much the same effect (*30*). Such perturbations might be introduced in the laboratory, for example, by imperfections in the boundary walls.

A different prospective application of the idea of destabilizing operator perturbations may be to the theory of "secondary instability" as an explanation of subcritical transition to turbulence (*7–9*). This theory is founded on the observation that when certain laminar shear flows are perturbed by certain physically motivated waves of large amplitude, the resulting problem is eigenvalue unstable. The great sensitivity of the spectrum of $\mathcal{L}$ to perturbations, however, raises the question of whether this effect might be better understood as a symptom of the nonnormality of the unperturbed problem than as a description of how transition to turbulence actually takes place under natural conditions (*31*).

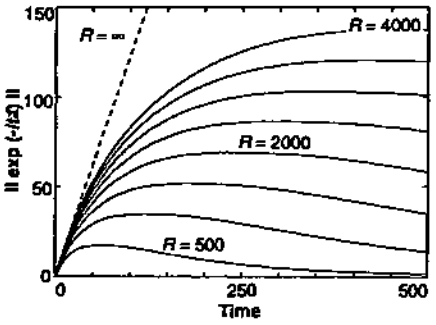

Fig. 7. Nonmodal transient growth of flow perturbations for Couette flow. For any finite R, viscous effects shut off the growth on a time scale $O(R)$.

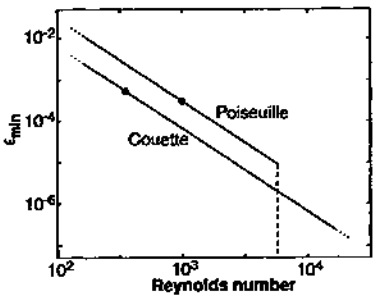

Fig. 8. Minimal norm $\epsilon_{min} = O(R^{-2})$ of an operator perturbation $\mathcal{E}$ that can destabilize an eigenvalue-stable shear flow at Reynolds number R. The dependence is so nearly quadratic that the curves appear straight to plotting accuracy. The dots mark the approximate Reynolds numbers at which Couette and Poiseuille flows are typically observed to undergo transition to turbulence (350 and 1000, respectively).

Favored Structures

Unlike eigenmodes, the pseudomodes of a linear operator are not uniquely determined; the precise structure excited in a highly nonnormal linear system will depend on the details of the excitation. But it is noteworthy that the three physical mechanisms considered above lead to similar predictions of what flow structures should be prominent in shear flows at high Reynolds numbers. Consider the initial-value problem of Eq. 8 with initial flow field $\mathbf{U}(0) = \mathbf{v}$ and solution $\exp(-it\mathcal{L})\mathbf{v}$. The amplitude history of this solution is given by $\|\exp(-it\mathcal{L})\mathbf{v}\|$, and the operator norm $\|\exp(-it\mathcal{L})\|$ is the upper envelope of all such curves corresponding to all initial functions $\mathbf{v}$. Figure 9 indicates how closely this envelope is approached by three physically interesting choices of $\mathbf{v}$. Each of these functions can be characterized mathematically as the function on which a certain linear operator attains its norm; the operators are $\mathcal{L}^{-1}$, $\exp(-it_{opt}\mathcal{L})$, and $\exp(-i0^+\mathcal{L})$, respectively, with the 0^+ notation indicating a limit as $t \to 0$. The first function, $\mathbf{v}_1$, is the one that excites a maximal resonant response, or equivalently, induces a minimal destabilizing perturbation. The second, $\mathbf{v}_2$, is the Butler-Farrell "optimal" that achieves maximal total growth at some time $t = t_{opt}$. The third, $\mathbf{v}_3$, is the perturbation with maximal growth rate at $t = 0$, which has been studied by Lumley and others (*14, 32*). We compute $\mathbf{v}_j$ numerically with the aid of the SVD applied to discrete approximations of the associated operators (*33*).

Although the physical ideas behind $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are different, $\mathbf{v}_1$ and $\mathbf{v}_2$ achieve comparable and near-maximal transient growth (and also comparable resonance,

Table 1. Leading-order behavior of various quantities as $R \to \infty$ (numbers accurate to 1% for $R > 100$) and the associated wave numbers α and β. The results for Poiseuille flow pertain to the highly nonnormal part of the problem, ignoring the $R > 5772$ mode (TS wave).

	Couette			Poiseuille		
	Value	α	β	Value	α	β
Distance of spectrum from real axis	$(R/2.47)^{-1}$	0	0	$(R/2.47)^{-1}$	0	0
Maximum resonance $\sup_{\omega\in\mathbb{R}}\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$	$(R/8.12)^2$	0	1.18	$(R/17.4)^2$	0	1.62
Transient growth $\sup_{t>0}\|\exp(-it\mathcal{L})\|$	$R/29.1$	$35.7/R$	1.60	$R/71.5$	0	2.04
Optimal time t_{opt} for above	$R/8.52$	$35.7/R$	1.60	$R/13.2$	0	2.04
Lower bound based on pseudospectra	$R/42.6$	0	1.62	$R/103.$	0	2.04
Transient growth ($\alpha = 0$)	$R/29.3$	0	1.66	$R/71.5$	0	2.04

not shown). These functions are also similar in structure, both having approximately the form of streamwise vortices that evolve into streamwise streaks.

Nonlinear Bootstrapping and Transition to Turbulence

Kelvin wrote in 1887 (34):

> It seems probable, almost certain indeed, that . . . the steady motion is stable for any viscosity, however small; and that the practical unsteadiness pointed out by Stokes forty-four years ago, and so admirably investigated experimentally five or six years ago by Osborne Reynolds, is to be explained by limits of stability becoming narrower and narrower the smaller is the viscosity.

This view of instability is still standard, but to this day, it has never been confirmed in detail. In this final section, we speculate about what the eventual confirmation may look like—about how nonlinear and linear mechanisms interact to bring about transition to turbulence.

Consider the 2 × 2 nonlinear model problem

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} + \|\mathbf{u}\|B\mathbf{u},\ A = \begin{pmatrix} -R^{-1} & 1 \\ 0 & -2R^{-1} \end{pmatrix},$$
$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad (14)$$

where R is a large parameter. The linear term, involving the nonnormal matrix A, amplifies energy transiently. The nonlinear term, involving the skew-symmetric matrix B, rotates energy in the $\mathbf{u}_1\mathbf{u}_2$ plane but does not create or destroy energy, for it acts orthogonally to the motion (35). Thus, we have linear amplification coupled with energy-neutral nonlinear mixing, a situation that holds also for the equations of fluid mechanics (36).

We consider the norms $\|\mathbf{u}(t)\|$ for solutions to Eq. 14 with $R = 25$ starting from eight different initial vectors $\mathbf{u}(0) = (0, \text{const})^T$ (Fig. 10), where the T denotes transpose. For $\|\mathbf{u}(0)\| \leq 10^{-4}$, the curves are approximately translates of one another on this log scale, indicating that the evolution is effectively linear. At $\|\mathbf{u}(0)\| = 4 \times 10^{-4}$, the nonlinearity has a pronounced effect. For $\|\mathbf{u}(0)\| = 5 \times 10^{-4}$ and higher amplitudes, the curves do not decay but blow up to a critical point of amplitude ≈ 1.

These calculations reveal a remarkable phenomenon: The amplitude growth is far greater than that of the linearized problem $d\mathbf{u}/dt = A\mathbf{u}$. We find that ϵ is of order R^{-3}, not R^{-1} (Table 2). This "bootstrapping" effect can be explained as follows. Suppose the solution at $t = 0$ consists of a vector of amplitude ϵ in a direction that excites growth of the linear problem $d\mathbf{u}/dt = A\mathbf{u}$ (the principal right singular vector of A). At a later time, of order R, the solution has

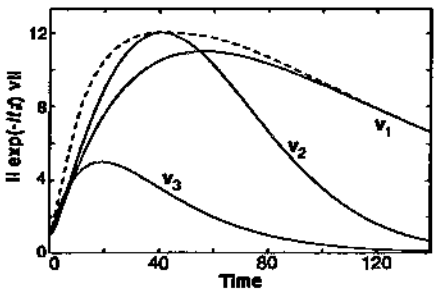

Fig. 9. Initial growth and eventual decay of $\|\exp(-it\mathcal{L})\mathbf{v}\|$ resulting from three initial perturbations $\mathbf{v}$ for Couette flow, $R = 350$. The dashed curve is the operator norm as in Fig. 7.

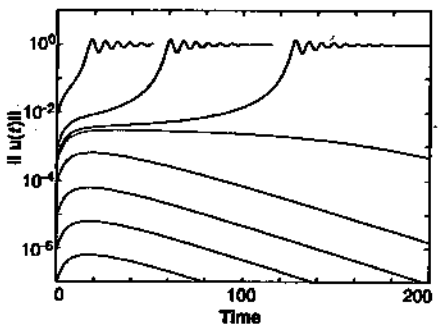

Fig. 10. $\|\mathbf{u}(t)\|$ for solutions to the nonlinear 2 × 2 model problem of Eq. 14 with initial amplitudes $\|\mathbf{u}(0)\| = 10^{-7}$, 10^{-6}, 10^{-5}, 10^{-4}, 4×10^{-4}, 5×10^{-4}, 10^{-3}, and 10^{-2}. The threshold amplitude is $\|\mathbf{u}(0)\| = 4.22 \times 10^{-4}$.

grown to order $R\epsilon$ by the linear growth mechanism but moved into a direction that no longer excites growth (the corresponding left singular vector). Meanwhile, however, the nonlinear term has had the effect of transferring some of this energy back to the original direction, with amplitude $R(R\epsilon)^2 = R^3\epsilon^2$ because the nonlinearity is quadratic and the time scale is $O(R)$. If $R^3\epsilon^2$ is of order less than ϵ, the process is not self-sustaining and the energy decays. On the other hand, if $R^3\epsilon^2$ is of order greater than ϵ, there is more energy than at the start and sustained growth may occur. Thus, the threshold amplitude is $\epsilon = O(R^{-3})$. A similar experiment shows that if the same nonlinear equation is driven by a forcing oscillation $e^{i\omega t}\mathbf{v}$ instead of an initial vector $\mathbf{u}(0)$, the threshold amplitude becomes $\epsilon = O(R^{-4})$.

It may appear that these results indicate the great power and importance of nonlinear effects. Yet in two senses, these energy growth scenarios are essentially linear. First, as mentioned above, the nonlinear term in Eq. 14 does not add energy but merely redistributes it. Second, the appearance of the bootstrapping phenomenon does not depend on the precise nature of the nonlinearity. Any quadratic nonlinear term that transfers energy from decaying to growing solution components has the potential for inducing a threshold amplitude $\epsilon = O(R^{-3})$ or $O(R^{-4})$ with respect to initial or forcing data, respectively; a random perturbation, for example, will often suffice. Higher order nonlinearities lead to similar effects, though the exponents may be lowered, for example, to $\epsilon = O(R^{-2})$ and $O(R^{-3})$ for a cubic nonlinearity.

Table 2. "Transition to turbulence" of a 2 × 2 nonlinear matrix model problem (Eq. 14). Although the growth factor M of the linearized problem is $O(R)$, nonlinear mixing effectively cubes this figure so that the threshold amplitude ϵ is $O(R^{-3})$.

R	M	ϵ	Ratio
12.5	3.18	3.41×10^{-3}	
			8.1
25	6.28	4.22×10^{-4}	
			8.0
50	12.5	5.27×10^{-5}	
			8.0
100	25.0	6.58×10^{-6}	

The Navier-Stokes equations are more complicated than our 2 × 2 model. One difference is that instead of 2-vectors, they act on functions with infinitely many degrees of freedom, most of which do not experience nonnormal linear growth. There will always be some energy in the growing pseudomodes, however, and in a pipe or channel of substantial extent, random fluctuations can be expected to raise the energy levels in such components locally well above the statistical average (37, 38). Another difference is that the nonlinear interactions in the Navier-Stokes equations act across different wave numbers α and β (12, 13), making the quadratic nonlinearity of Eq. 14 perhaps unrealistically strong. Notwithstanding these qualifications, we conjecture that transition to turbulence of eigenvalue-stable shear flows proceeds analogously to our model in that the destabilizing mechanism is essentially linear in the senses described above and the amplitude threshold for transition is $O(R^\gamma)$ for some $\gamma < -1$ (39).

Conclusion

We have discussed three linear approaches to the phenomenon of instability of shear flows: (i) pseudoresonance, (ii) transient growth of flow perturbations, and (iii) destabilizing operator perturbations. These ideas are by no means independent. Mathematically, all are related to the pseudospectra of the operator $\mathcal{L}$, and physically, all depend on the same mechanisms of extraction of energy from the mean flow by structures such as streamwise vortices. One should not expect that one of these ideas will prove to be right and the others wrong.

More likely, each may prove relevant to a particular class of experiments, for it is natural to speculate that transition to turbulence may be triggered in distinct circumstances by distinct causes, such as, respectively, (i) laboratory vibrations, (ii) initial or inlet disturbances, or (iii) deviations of the pipe or channel geometry from the Poiseuille or Couette ideal. In the coming decade, as numerical simulations of the nonlinear Navier-Stokes equations become routine, much progress will be made in the elucidation of these details and a fuller picture will emerge of the interaction of linear and nonlinear effects in fluid mechanics.

Besides hydrodynamic stability, there are other fields in which nonorthogonal eigenfunctions arise and eigenvalues may be misleading. Examples in fluid mechanics include the instability of magnetic plasmas (*40*) and the formation of cyclones (*41*). Examples in numerical analysis include the stiffness and numerical instability of discretizations of differential equations (*42*) and the convergence of iterative algorithms for nonsymmetric matrix problems (*43*). The recurring theme in these and other applications is that although the long-time behavior of an evolving system may be controlled by nonlinearities, some important phenomena are of a short-time nature and are essentially linear (*44*). If the linearized problem is far from normal, eigenvalues may be precisely the wrong tool for analyzing it, for eigenvalues determine the long-time behavior of a nonnormal linear process, not the transient.

REFERENCES AND NOTES

1. P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge Univ. Press, Cambridge, 1981).
2. This article is adapted from L. N. Trefethen, A. E. Trefethen, S. C. Reddy, *Tech. Rep. TR 92-1291* (Department of Computer Science, Cornell University, Ithaca, NY, 1992) and ———, T. A. Driscoll, *Tech. Rep. 92TR115* (Cornell Theory Center, Cornell University, Ithaca, NY, 1992). Readers interested in additional details are urged to contact the first author for copies of these reports.
3. S. A. Orszag, *J. Fluid Mech.* **50**, 689 (1971). The Reynolds number is defined by $R = LV\rho/\mu$, where L is the half-width of the channel, V is the maximum velocity, ρ is the density, and μ is the viscosity. The distance, velocity, and time variables are nondimensionalized by L, V, and L/V, respectively.
4. S. J. Davies and C. M. White, *Proc. R. Soc. London Ser. A* **119**, 92 (1928); V. Patel and M. R. Head, *J. Fluid Mech.* **38**, 181 (1969); D. R. Carlson, S. E. Widnall, M. F. Peeters, *ibid.* **121**, 487 (1982).
5. V. A. Romanov, *Funct. Anal. Appl.* **7**, 137 (1973).
6. A. Lundbladh and A. V. Johansson, *J. Fluid Mech.* **229**, 499 (1991); N. Tillmark and P. H. Alfredsson, *ibid.* **235**, 89 (1992).
7. P. S. Klebanoff, K. D. Tidstrom, L. M. Sargent, *ibid.* **12**, 1 (1962).
8. S. A. Orszag and A. T. Patera, *ibid.* **128**, 347 (1983).
9. B. J. Bayly, S. A. Orszag, T. Herbert, *Annu. Rev. Fluid Mech.* **20**, 359 (1988).
10. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, New York, 1976); A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer-Verlag, New York, 1983).
11. S. C. Reddy, P. J. Schmid, D. S. Henningson, *SIAM J. Appl. Math.* **53**, 15 (1993).
12. L. H. Gustavsson, *J. Fluid Mech.* **224**, 241 (1991).
13. D. S. Henningson, in *Advances in Turbulence*, A. V. Johansson and P. H. Alfredsson, Eds. (Springer-Verlag, New York, 1991), pp. 279–284; D. S. Henningson, A. Lundbladh, A. V. Johansson, *J. Fluid Mech.* **250**, 169 (1993).
14. K. M. Butler and B. F. Farrell, *Phys. Fluids A* **4**, 1637 (1992).
15. S. C. Reddy and D. S. Henningson, *J. Fluid Mech.* **252**, 209 (1993).
16. Additional related developments, both theoretical and experimental, are described in K. S. Breuer and J. H. Haritonidis, *J. Fluid Mech.* **220**, 569 (1990); Y. Bun and W. O. Criminale, *ibid.*, in press; B. G. B. Klingmann, *ibid.* **240**, 167 (1992)
17. This region has appeared in many places, including figures 4.11, 4.15, and 5.5 of (*1*). These and other figures in the literature have the x wave number α as the ordinate, not the xz wavenumber magnitude $k = (\alpha^2 + \beta^2)^{1/2}$, and thus correspond to 2D perturbations only. This changes the upper boundary in such a way that the width of the shaded region shrinks to zero as $R \rightarrow \infty$.
18. W. M'F. Orr, *Proc. R. Ir. Acad. Sect. A* **27**, 9 (1907); *ibid.*, p. 69; K. M. Case, *J. Fluid Mech.* **10**, 420 (1960) and *Phys. Fluids* **3**, 143 (1960); G. Rosen, *ibid.* **14**, 2767 (1971); T. Ellingsen and E. Palm, *ibid.* **18**, 487 (1975); M. T. Landahl, *J. Fluid Mech.* **98**, 243 (1980); D. J. Benney and L. H. Gustavsson, *Stud. Appl. Math.* **64**, 185 (1981); L. S. Hultgren and L. H. Gustavsson, *Phys. Fluids* **24**, 1000 (1981); A. D. D. Craik and W. O. Criminale, *Proc. R. Soc. London Ser. A* **406**, 13 (1986); B. F. Farrell, *Phys. Fluids* **31**, 2093 (1988).
19. By 3D, we mean z-dependent, as is conventional in this field. Thus, a perturbation that depends on y and z but not x is said to be 3D, not 2D.
20. H. B. Squire, *Proc. R. Soc. London Ser. A* **142**, 621 (1933).
21. R. Breidenthal, *J. Fluid Mech.* **109**, 1 (1981); L. P. Bernal and A. Roshko, *ibid.* **170**, 499 (1986); J. C. Lasheras and H. Choi, *ibid.* **189**, 53 (1988); M. Asai and M. Nishioka, *ibid.* **208**, 1 (1989); S. K. Robinson, *Tech. Memo. NASA TM-103859* (National Aeronautics and Space Administration, Washington, DC, 1991).
22. This process is sometimes called "lift-up"; see M. T. Landahl, *SIAM J. Appl. Math.* **28**, 735 (1975).
23. The term pseudospectrum was introduced by L. N. Trefethen [in *Numerical Analysis 1991*, D. F. Griffiths and G. A. Watson, Eds. (Longman, London, 1992), pp. 234–266]. Essentially the same idea has been applied over the years by various authors, including J. M. Varah, J. W. Demmel, and especially S. K. Godunov and his colleagues. Possibly its earliest application is by H. J. Landau [*J. Opt. Soc. Am.* **66**, 525 (1976)].
24. Spectra of $\mathcal{L}_{\alpha\beta}$ for particular values of α and β have been published by various authors (*3*), and pseudospectra of $\mathcal{L}_{\alpha\beta}$ appear in (*11*, *15*). Neither spectra nor pseudospectra of $\mathcal{L}$ itself have been plotted previously.
25. Our computations have been carried out on a Connection Machine-System CM-5 in Fortran and on Sun workstations in Matlab, with numerical methods adapted from those of Reddy and Henningson (*11–13*, *15*). First, the problem is Fourier transformed in x and z, and the working variables are taken as the y components of velocity and vorticity; a weighted norm is introduced for these variables that is equivalent to the energy norm $\|\cdot\|$. The 1D problem is discretized by a Chebyshev hybrid spectral method and then projected onto a space spanned by a set of dominant eigenmodes, leading to matrices typically of size about 80×80. In the Poiseuille case, the odd and even solutions with respect to y decouple, and we take advantage of this separation. The generation of most of our figures requires an optimization with respect to α or β or both. The largest calculations are required for the pseudospectra of Figs. 4 and 5, which are determined by evaluating $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$ on a grid in the ω-plane of typical size 64 by 64 and plotting contours of the resulting data array. For each value of ω on the grid, $\|(\omega\mathcal{I} - \mathcal{L})^{-1}\|$ is computed by numerical minimization of $\|(\omega\mathcal{I} - \mathcal{L}_{\alpha\beta})^{-1}\|^{-1}$ with respect to α and β, and each evaluation of this function entails a matrix singular value decomposition (SVD).
26. The pseudospectral contours in Fig. 5 bend around the TS bumps but come very close to them, revealing that the TS wave is oriented at a sizable angle to the remainder of the system. The angle is $\theta \approx 16.4°$ as measured in the energy inner product, corresponding to an eigenvalue condition number $\kappa = 1/\sin\theta \approx 3.52$. Thus the small bump in the $\epsilon = 10^{-3.5}$ pseudospectral contour of Fig. 5B lies at a distance $3.52 \times 10^{-3.5}$ from the spectrum.
27. One difference between the Couette and Poiseuille problems is that the spectra and pseudospectra are broader in the latter case. This is a consequence of the fact that the mean flow velocity is zero for Couette flow but nonzero for Poiseuille flow, so that a plane wave of wavenumber α traveling with the mean flow will have $\omega \approx 0$ for Couette flow but $\omega = O(\alpha)$ for Poiseuille flow. Figures 4 and 5 would look different if the Navier-Stokes problem were formulated in a moving reference frame, though the behavior along the imaginary axis would be essentially unchanged.
28. For details on structures with $\alpha \neq 0$ see (*14*) and B. F. Farrell and P. J. Ioannou, *Phys. Fluids A* **5**, 1390 (1993).
29. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins Univ. Press, Baltimore, MD, 1989); R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge Univ. Press, Cambridge, 1991).
30. At $R = 5000$, for example, approximately 90% of all random perturbations $\mathcal{E}$ to $\mathcal{L}$ with $\|\mathcal{E}\| = 10^{-3}$ render a Poiseuille flow eigenvalue unstable, where by a random perturbation we mean an elementwise random perturbation of the 20×20 matrix that represents the projection of $\mathcal{L}$ onto the space spanned by its 20 dominant eigenmodes (with fixed $\alpha = 0$ and $\beta = 1.62$).
31. The validity of the secondary instability model of natural transition to turbulence has been questioned for other reasons; alternatives are sometimes known as models of "bypass transition" [M. V. Morkovin, in *Viscous Drag Reduction*, C. S. Wells, Ed. (Plenum, New York, 1969)], pp. 1–31.
32. J. Serrin, *Arch. Ration. Mech. Anal.* **3**, 1 (1959); F. H. Busse, *Z. Angew. Math. Phys.* **20**, 1 (1969); J. L. Lumley, in *Developments in Mechanics*, L. H. N. Lee and A. H. Szewczyk, Eds. (Notre Dame Press, South Bend, IN, 1971), vol. 6, pp. 63–88; D. D. Joseph, *Stability of Fluid Motions I* (Springer-Verlag, New York, 1975).
33. To compute $\mathbf{v}_3$, we make use of its characterization as the principal eigenfunction of $(\mathcal{L} - \mathcal{L}^*)/2i$, with growth rate at $t = 0$ equal to the corresponding eigenvalue, or equivalently, to the distance that the numerical range of $\mathcal{L}$ protrudes into the upper half-plane. This is essentially the Hille-Yosida theorem of functional analysis (*10*).
34. Lord Kelvin (W. Thomson), *Philos. Mag.* **24**, 188 (1887).
35. This can be seen algebraically from the absence of B in the easily verified formula $d\|\mathbf{u}\|^2/dt = \mathbf{u}^T(A + A^T)\mathbf{u}$.
36. The equation for $d\|\mathbf{u}\|^2/dt$ in fluid mechanics is called the Reynolds-Orr equation (*1*). Our 2×2 model is not so far from the interaction of vortex "tilting" and "stretching" described in Orzsag and Patera (*8*); the first and second components of $\mathbf{u}$ are analogous to the normal components of velocity and vorticity, respectively.
37. We have not discussed the relation between spatially global and local results, which is somewhat obscured by modal or pseudomodal analysis, because different wave numbers α,β decouple orthogonally. A related matter is the investiga-

tion of spatial rather than temporal evolution of flow disturbances, which according to results of P. J. Schmid (personal communication), leads to results qualitatively analogous to those reported here.
38. Numerical experiments indicated that even random fluctuations tend to excite growing pseudomodes. A theory of the nonnormal response of linearized flows to stochastic forcing data has been developed by B. F. Farrell and P. J. Ioannou (in preparation).
39. The bounds $-21/4 \leq \gamma \leq -1$ are derived by G. Kreiss, A. Lundbladh, and D. S. Henningson [*Tech. Rep. TRITA-NA-9307* (Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, Sweden, 1993)].
40. W. Kerner, *J. Comp. Phys.* **85**, 1 (1989).
41. B. F. Farrell, *J. Atmos. Sci.* **46**, 1193 (1989).
42. S. C. Reddy and L. N. Trefethen, *Numer. Math.* **62**, 235 (1992); D. J. Higham and L. N. Trefethen, *BIT*, **33**, 285 (1993). The importance of a proper treatment of nonnormality in numerical analysis, and a corresponding proper definition of stability, was detailed by H.-O. Kreiss in the 1960s [*ibid.* **2**, 153 (1962); R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems* (Wiley-Interscience, New York, 1967)].
43. N. M. Nachtigal, S. C. Reddy, L. N. Trefethen, *SIAM J. Matrix Anal. Appl.* **13**, 778 (1992); N. M. Nachtigal, L. Reichel, L. N. Trefethen, *ibid.*, p. 796.
44. There is evidence that this statement applies even to fully developed turbulence; see M. J. Lee, J. Kim, P. Moin, *J. Fluid Mech.* **216**, 561 (1990) and K. M. Butler and B. F. Farrell, *Phys. Fluids A* **5**, 774 (1993).
45. Supported by an NSF grant [DMS-9116110 (L.N.T.)] and a graduate fellowship (T.A.D.). We acknowledge the assistance over several years of D. Henningson and P. Schmid and, more recently, of K. Butler, B. Farrell, and P. Ioannou. In addition, we acknowledge the helpful comments of B. Fornberg, P. Holmes, R. Scott, and E. Siggia, and also Thinking Machines Corporation for the donation of extensive CM-5 computer time.