Pseudospectra of linear operators[1]

Lloyd N. Trefethen

Abstract

The advent of ever more powerful computers has brought with it a new way of conceiving some of the fundamental eigenvalue problems of applied mathematics. If a matrix or linear operator A is far from normal, its eigenvalues or more generally its spectrum may have little to do with its behavior as measured by quantities such as $\|A^n\|$ or $\|\exp(tA)\|$. More may be learned by examining the sets in the complex plane known as the *pseudospectra* of A, defined by level curves of the norm of the resolvent, $\|(zI - A)^{-1}\|$. Five years ago, the author published a paper that presented computed pseudospectra of thirteen highly non-normal matrices arising in various applications. Since that time, analogous computations have been carried out for differential and integral operators. This paper, a companion to the earlier one, presents ten examples, each chosen to illustrate one or more mathematical or physical principles.

1 Introduction

Eigenvalues—more generally spectra—have been a standard tool of the mathematical sciences for a century and a half. Since the arrival of computers, they have also been a standard tool of scientific computing. For example, one of the most influential items of numerical software over the years has been EISPACK, a collection of Fortran subroutines for matrix eigenvalue computations used around the world since its introduction in the mid-1970s [Smi].

Eigenvalues are useful for three main reasons. The *algorithmic* reason is that if a matrix or linear operator can be diagonalized, transforming the problem to a basis of eigenfunctions, this may speed up the solution of various problems. The *physical* reason is that eigenvalues may give information about the behavior of an evolving system governed by a matrix or operator. In particular, they may give information about resonance, about instability, and about rates of growth

[1] This article will appear in *ICIAM '95: Proceedings of the Third International Congress on Industrial and Applied Mathematics,* to be published by Akademie-Verlag, Berlin. It is reproduced here, with minor modifications, by permission of Akademie-Verlag.

or decay as $t \to \infty$. Finally, there is a *psychological* reason for the usefulness of eigenvalues. Much of the human brain is specialized for the processing of visual information, and eigenvalues take advantage of this biological trait, supplementing the abstract notion of a matrix or operator by a picture in the complex plane. They give an operator a personality.

Here are some examples. In structural mechanics, eigenvalues may determine whether an automobile is too noisy or whether a building will collapse in an earthquake. In aeronautics, eigenvalues may determine whether the flow over a wing is laminar or turbulent. In quantum mechanics, eigenvalues may determine atomic energy levels and thus the frequency of a laser or the spectral signature of a star. In ecology, eigenvalues may determine whether a food web will settle into a steady equilibrium. In probability theory, eigenvalues may determine the rate of convergence of a Markov process. In electrical engineering, eigenvalues may determine the frequency response of an amplifier or the reliability of a national power system. In numerical analysis, eigenvalues may determine whether a discretization of a differential equation will get the right answer or how fast a conjugate gradient iteration will converge. And so on.

For these and other applications, eigenvalue analysis has proved highly successful. In the past decade, however, an awareness has been growing that for certain problems of science and engineering, the predictions suggested by eigenvalues do not always match what is actually observed. These are problems involving matrices or operators that are not normal, and indeed, are in some sense far from normal. A *normal* matrix is one that has a complete set of orthogonal eigenvectors. This is equivalent to the condition $AA^* = A^*A$, where A^* is the conjugate transpose. By contrast, the eigenvectors of a non-normal matrix, though they may form a complete set, are not orthogonal; the condition number of any matrix of eigenvectors is greater than 1 and possibly very large. For operators, the definitions are analogous, though the details become technical. A normal linear operator is again one that satisfies $AA^* = A^*A$, where A^* is now a suitably defined adjoint operator. In this paper we shall not discuss details of functional analysis; we only note that in all of what follows, we assume that A is a closed linear operator in a Hilbert space and that it generates a C_0 semigroup [Paz]. The Hilbert space norm is denoted by $\|\cdot\|$, and for our examples that come from physical problems, we always arrange things so that this norm corresponds to the square root of energy.

The difficulty with non-normal matrices and operators goes beyond the need for a Jordan canonical form instead of a diagonalization if a matrix lacks a complete set of eigenvectors. Any use of eigenvalues to derive physical predictions relies on an implicit transformation to eigenvector coordinates. If the matrix is normal, this transformation is unitary—a rotation or a reflection. If it is far from normal, however, the change to eigenvector coordinates may involve an extreme distortion of the state space. In the new coordinates, the physics of the system may become strangely complicated. A typical state of the system may be a superposition of huge eigenfunction components that nearly cancel, and the evolution over time intervals of scientific interest may be determined by how this pattern of cancellation evolves, rather than by the growth or decay of the individual eigenfunctions. In other words, there may be no good scientific reason for attempting to analyze the problem in terms of eigenvalues or eigenvectors.

For a fixed matrix or operator, the distortions associated with non-normality may be of finite magnitude and thus arguably of limited significance. In applications, however, one is often faced not with a single matrix or operator but with a family of them indexed by a parameter such as Reynolds number, Péclet number, or resistivity, and the degree of non-normality may increase unboundedly as this parameter approaches a limit of physical interest. In such circumstances it is impossible to get quantitative information of uniform validity from spectral information alone.

2 Resolvent Norms and Pseudospectra

Since the beginning of this subject, one of the principal tools for dealing with non-normality has been the matrix or operator known as the resolvent. We can motivate this idea as follows.

It is well known that in applied mathematics, it is not always enough to ask whether an operator A is singular or nonsingular; the norm of the inverse $\|A^{-1}\|$ may be important as well as the fact that it is finite. Thus, to an applied mathematician, problems are not just well-posed or ill-posed; they lie on a continuum from well- to ill-conditioned, with ill-posedness being the limit of infinite ill-conditioning. Now the definition of an eigenvalue also involves the singularity or nonsingularity of a matrix or operator. A number $z \in \mathbb{C}$ is an eigenvalue of a matrix A if $zI - A$ is singular, where I denotes the identity. In the operator case z is in the *spectrum* of A if a bounded inverse of $zI - A$ does not exist. It should not be surprising that for applications, it is sometimes desirable to extend these definitions too by asking not just whether $(zI - A)^{-1}$ exists but whether it is large or small. The matrix or operator $(zI - A)^{-1}$ is the *resolvent* of A at the point $z \in \mathbb{C}$.

From here, the following definition suggests itself. We have already defined the spectrum of a matrix or operator A; let us denote it by $\Lambda(A)$, a subset of the complex plane $\mathbb{C}$. (For matrices, $\Lambda(A)$ is finite, discrete, and nonempty; for operators, anything is possible so long as it is closed.) Now, for each $\epsilon \geq 0$, let us define a new subset of the complex plane, the *ϵ-pseudospectrum* of A, as follows:

$$\Lambda_\epsilon(A) \;=\; \{\, z \in \mathbb{C} : \; \|(zI - A)^{-1}\| \geq \epsilon^{-1} \,\}. \tag{1}$$

By convention we write $\|(zI-A)^{-1}\| = \infty$ if $z \in \Lambda(A)$. Thus the 0-pseudospectrum $\Lambda_0(A)$ is the same as the spectrum, but for $\epsilon > 0$, it can be shown, $\Lambda_\epsilon(\epsilon)$ is a larger set, and never empty. The ϵ-pseudospectra of A are a family of strictly nested closed sets, which grow to fill the whole complex plane as $\epsilon \to \infty$.

Pseudospectra can also be defined in other, equivalent ways. One is in terms of perturbations of the spectrum. For any matrix A we have

$$\Lambda_\epsilon(A) \;=\; \{\, z \in \mathbb{C} : \; z \in \Lambda(A + \Delta A) \text{ for some } \Delta A \text{ with } \|\Delta A\| \leq \epsilon \,\}, \tag{2}$$

and if A is an operator, the equivalence with (1) still holds if we take the closure of this set [Böt]. Thus a number z is in the interior of the ϵ-pseudospectrum of A if and only if it is in the spectrum of some perturbed operator $A + \Delta A$ with $\|\Delta A\| < \epsilon$.

A third equivalent definition, closer to computation, involves the singular value decomposition. For any matrix or operator A we have

$$\Lambda_\epsilon(A) \;=\; \{\, z \in \mathbb{C} : \; \sigma_{\min}(zI - A) \leq \epsilon \,\}, \tag{3}$$

where $\sigma_{\min}$ denotes the smallest singular value in the matrix case or the smallest s-number for an operator [GoKr]. In stating (3) we have made use of the assumption that A acts in a Hilbert space; in a Banach space setting we have (1) and (2) but not (3) [DKS].

If A is normal, then $\Lambda_\epsilon(A)$ is exactly the set of points in $\mathbb{C}$ at distance $\leq \epsilon$ from $\Lambda(A)$. If A is not normal, however, it may be much larger. Here is a physical interpretation of this observation [TTRD]. Consider a time-dependent driven system $du/dt = Au + e^{zt}f$, $u(0) = u_0$, where f is a fixed function in the Hilbert space under study. The solution to this problem is $u(t) = e^{zt}(zI - A)^{-1}f + \exp(tA)(u_0 - (zI - A)^{-1}f)$. If $z \in \Lambda_\epsilon(A)$, this means that for certain choices of f and u_0, $\|u(t)\|/\|e^{zt}f\|$ may come arbitrarily close to ϵ^{-1} (the norm is now an L^2 integral over t). In other words, whereas a system governed by a normal operator exhibits resonance only if the forcing frequency is close to the spectrum, a system governed by a non-normal operator may exhibit resonance or *pseudo-resonance* at frequencies far from the spectrum. This is one respect in which judging the behavior of non-normal systems solely by their spectra may be problematic.

3 Some History

The mathematical theory of non-normal operators originated with von Neumann in the 1920s. In a sense it has been one of the central topics of functional analysis ever since, whose elements are widely known. For example, the *Hilbert Space Problem Book* by Halmos contains numerous examples illustrating the curious behavior of the spectra of non-normal operators [Hal]. Treatments in monograph form include the book by Gohberg and Kreĭn [GoKr] and the massive work of Dunford and Schwartz [DuSc].

Specifically, the use of the resolvent has also been standard for many years, at least among the theoretically inclined. Of numerous contributions, I shall mention two that are especially relevant to this paper. One is the remarkable book *Perturbation Theory for Linear Operators* by Kato, in which all kinds of questions of matrix and operator theory are beautifully treated by resolvent techniques [Kat]. The other is the work by Kreiss over the years in the field of finite difference methods for partial differential equations. A landmark 1962 paper by Kreiss, containing what became known as the Kreiss matrix theorem, described compellingly the pitfalls of eigenvalue analysis and the uses of the resolvent as an alternative [Kre]. Some of Kreiss's ideas were presented shortly afterwards in the text by Richtmyer and Morton [RiMo].

A narrower question has to do with the history of pseudospectra—that is, of the explicit investigation of the sets bounded by level curves of the norm of the resolvent. The first mention of this idea in print that I have found is by H. J. Landau in 1975 [Lan], who used the term "ϵ-spectrum". The first sketch of a pseudospectrum that I know of is a hand-drawn figure by Kostin and Razzakov in 1985 [KoRa] ("spectral portrait"). These authors were members of a group led by S. K. Godunov in Novosibirsk, which pursued various ideas in this vein; others in this group include Bulgakov, Kirilyuk, and Malyshev. The first computer-generated plot of pseudospectra that I know of is by Demmel in 1987 [Dem]. Meanwhile, other related contributions were made in the 1970s and 1980s by various people, including Chatelin, Hinrichsen, Pritchard, Varah, and

Wilkinson. My own first use of pseudospectra (the idea has been proposed independently at least five times) was in 1990 [Tre1]. Since that year, publications related to pseudospectra have been numerous. Aside from myself, some of those involved have been Baggett, Borba, Böttcher, Braconnier, Brühl, Carpraux, Donato, van Dorsselaer, Driscoll, Eiermann, Erhel, Frayssé, Greenbaum, Henningson, D. Higham, N. Higham, Hochbruck, Huysmans, Jackiewicz, Kerner, Kraaijevanger, Lubich, Lui, Lumsdaine, Marques, Nachtigal, Nevanlinna, Ottaviani, Owren, Pathria, Plato, Reddy, Reichel, Riedel, Roch, Ruhe, Sadkane, Schmid, Silbermann, Spijker, Toh, Toumazou, A. Trefethen, Viswanath, Weideman, Wolf, and Wu.

This brings me, at last, to the subject of this paper. What has stimulated so much recent activity is the growing power of computers, which have made it feasible to plot pseudospectra. On the workstations of 1990, it was possible to compute pseudospectra of 32×32 matrices in a few minutes, and in 1992 I published a paper presenting thirteen examples of this kind [Tre2]. (The names given to the examples were Jordan block, Limaçon, Grcar, Wilkinson, Frank, Kahan, Demmel, Lenferink–Spijker, Companion, Gauss–Seidel, Chebyshev spectral, Random, and Random upper-triangular.) Each example was illustrated by two plots: first, a plot of boundaries of $\Lambda_\epsilon(A)$ for $\epsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$; second, a "plot of dots" displaying the superposition of the eigenvalues of 100 random matrix perturbations $A + \Delta A$, $\|\Delta A\| = 10^{-3}$.

Since 1992, various people have begun to compute pseudospectra of differential and integral operators, beginning with the outstanding paper on the Orr–Sommerfeld operator by Reddy, Schmid, and Henningson [RSH]. Such computations are often feasible on today's workstations, and they are certainly feasible on today's supercomputers. The purpose of this paper is to present ten examples of pseudospectra of operators. Along the way we shall present a few theorems that partially answer the question, what do pseudospectra tell us about the behavior of an operator? (The full answer is not known; see [GrTr] and [TrVi].)

This paper is in every way a sequel to [Tre2], to which I hope all readers have access. Copies can be obtained by contacting me by email.

4 Computation of Pseudospectra

Before presenting the examples, we must make a few remarks about how pseudospectra of operators can be computed. Of course the starting point is the matrix case.

There is an obvious algorithm for computing the pseudospectra of a matrix A: following (3), evaluate $\sigma_{\min}(zI - A)$ or its logarithm by standard singular value decomposition software at the points of a grid in the complex plane, then send the results to a contour plotter. This is what was done in [Tre2], and it remains the basic, simple technique. Various methods have been proposed recently, however, to speed up these computations. S.-H. Lui has made the elegant observation that if A is first reduced to Hessenberg or triangular form, then this form can be preserved in computations of $\sigma_{\min}(zI - A)$ for various values z by inverse iteration; the result is often a speedup by a factor of ten or more [Lui]. Half a dozen other studies of iterative methods for computing pseudospectra have also been carried in the past year or so. For one such approach, with references to several of the others, see [BrHi].

If A is a differential or integral operator, of course, we must discretize it. There is little general literature on this at present, but a variety of methods have proved successful. The obvious thing to do is simply to approximate the operator by a matrix and then compute the pseudospectra of the matrix. This procedure can be quite successful if the discretization is highly accurate, and in particular, spectral methods rather than finite differences or finite elements have been the basis of most of the computations so far. In large-scale applications, however, it is often best not to convert the whole problem to a single matrix. Instead, different discretization tolerances may be appropriate for different values of z; these tolerances can be chosen adaptively. In some applications it also happens that the linear operator is discretized by a matrix that depends on various parameters, in which case one must solve an optimization problem to minimize $\sigma_{\min}(zI - A)$ with respect to these parameters. Finally, for computations on parallel computers, one must divide the computation into subproblems. Fortunately, this is not hard, as different regions of the z-plane can be treated independently.

In the plots presented below, I believe the pseudospectra and other curves shown are generally accurate to within a millimeter or two, if not always to full plotting accuracy. For details on how the computations were carried out—sometimes a substantial task—see the original papers.

5 Ten Examples

We now begin the presentation of our examples.

For linear operators, unlike matrices, the idea of a "random perturbation" does not make sense, since the ball $\|u\| \le \epsilon$ is not compact. Thus in this paper we cannot repeat the device of [Tre2] of presenting eigenvalues of randomly perturbed operators as well as boundaries of pseudospectra. Instead, I have chosen a different second plot to accompany the first plot in each of these examples. This is a plot of $\|\exp(tA)\|$ as a function of t, usually on a logarithmic scale. The significance of $\exp(tA)$ is that it is the solution operator for the linear, autonomous problem $du/dt = Au$; in the theory of semigroups, $\{\exp(tA)\}$ is the semigroup and A is its infinitesimal generator [HiPh], [Paz]. However, it should be remembered that other functions of operators besides the exponential are also of interest in applications. Examples are A^n for discrete evolution processes, A^{-1} for systems of equations, and polynomial or rational functions $p(A)$ or $r(A)$ for iterations such as conjugate gradients, Lanczos, or GMRES. Though we do not discuss it here, pseudospectra are relevant to these problems too.

Let the *spectral abscissa* of A be defined by

$$\alpha(A) = \sup_{z \in \Lambda(A)} \mathrm{Re}\, z. \tag{4}$$

If A is normal, we have $\|\exp(tA)\| = e^{t\alpha(A)}$ for all $t \ge 0$. If A is not normal, we have the bounds

$$e^{t\alpha(A)} \le \|\exp(tA)\| \le \kappa(V) e^{t\alpha(A)} \tag{5}$$

for all $t \ge 0$. Here $\kappa(V) = \|V\|\,\|V^{-1}\|$ denotes the condition number of a "matrix of eigenvectors" V of A, if one exists (otherwise the upper bound should be replaced by ∞). For infinite-dimensional operators, the idea of a matrix of eigenvectors with bounded condition number can be made precise via the notion of a *Riesz basis* [GoKr]. We shall not give details, but our computed examples will report a number $\kappa(V)$; when it is finite, this means there exists a Riesz basis of eigenvectors with the indicated condition number.

A related quantity is the *numerical abscissa* of A, defined by

$$\omega(A) = \sup_{z \in W(A)} \mathrm{Re}\, z. \tag{6}$$

Here $W(A)$ is the *numerical range* (= field of values) of A, the set of all scalars u^*Au (= Rayleigh quotients) corresponding to functions u with $\|u\| = 1$. According to results related to the Hille–Yosida theorem [Paz], the numerical

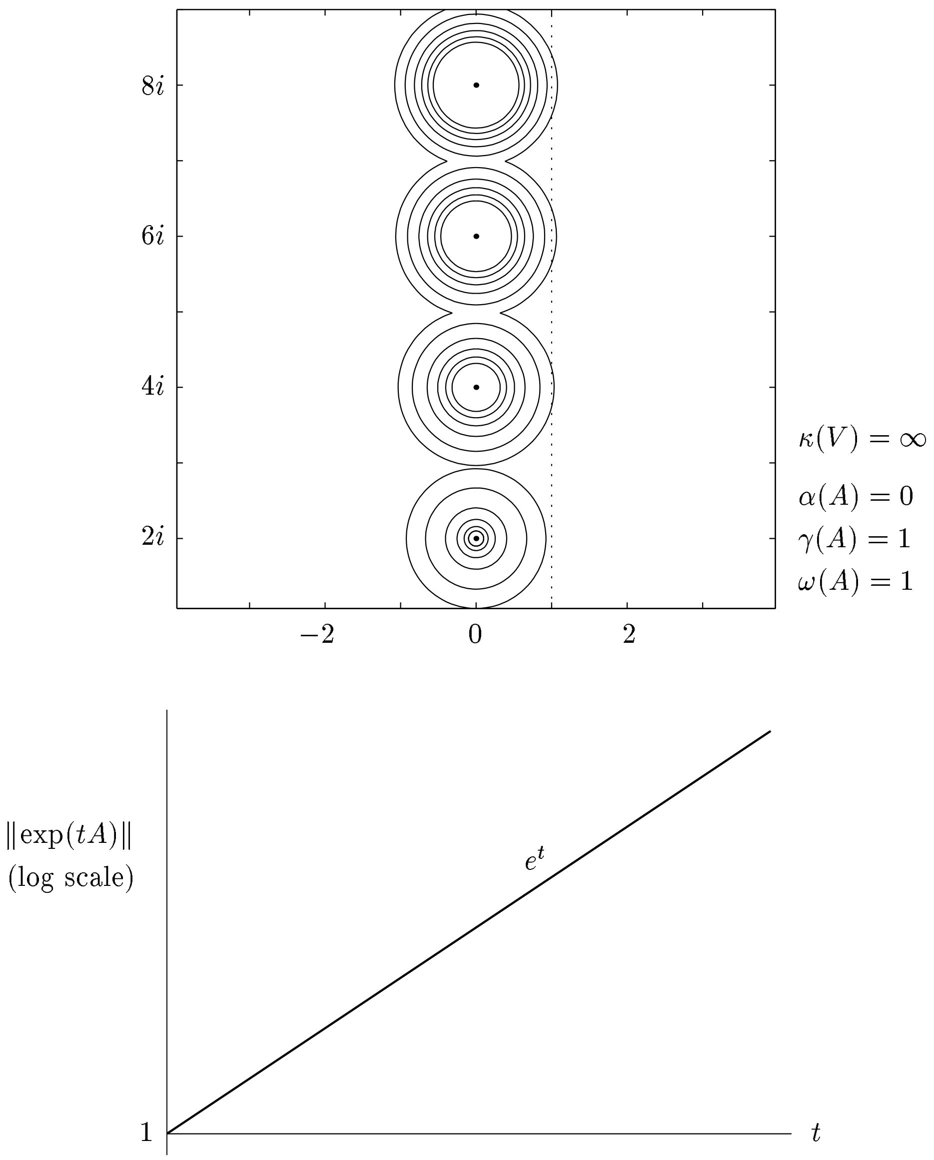


Figure 1: The Zabczyk operator, which illustrates that the growth abscissa $\gamma(A)$ of an unbounded operator may exceed the spectral abscissa $\alpha(A)$ (Theorem 1). The spectrum is the discrete set $\{2i, 4i, 6i, \ldots\}$, but the ϵ-pseudospectral abscissa is $\alpha_\epsilon(A) = 1 + \epsilon$ for each $\epsilon > 0$ and thus 1 in the limit $\epsilon \to 0$. The contours in the upper plot are boundaries of ϵ-pseudospectra $\Lambda_\epsilon(A)$ for $\epsilon = 10^{-1}, 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}$. (From Baggett [Bag], based on [Zab].)

abscissa of A determines the behavior of $\|\exp(tA)\|$ as $t \to 0$:

$$\frac{d}{dt}\|\exp(tA)\|_{t=0} = \omega(A). \tag{7}$$

Some of the upper plots of our figures include dashed curves corresponding to the boundary of $W(A)$; the numerical abscissa is the maximum real coordinate on such a curve.

Example 1. Zabczyk operator [Bag].

If A is non-normal, the spectrum $\Lambda(A)$ does not determine the behavior of $\|\exp(tA)\|$ for small t. Asymptotically as $t \to \infty$, however, one might expect the spectrum to be decisive. In particular, define the *growth abscissa* of A by

$$\gamma(A) = \lim_{t\to\infty} t^{-1} \log \|e^{tA}\| \tag{8}$$

(the limit always exists). For a matrix, we have $\gamma(A) = \alpha(A)$, the spectral abscissa of A, and the same is true for a bounded operator. For an unbounded operator, however, it was pointed out by Hille and Phillips that $\gamma(A) > \alpha(A)$ is possible.

This result may seem surprising, but it becomes easily understandable when interpreted in terms of pseudospectra. A simple example was devised by Zabczyk in 1975 [Zab], illustrated in Figure 1. Picking parameters arbitrarily for definiteness, consider a 10×10 Jordan block with eigenvalue $2i$. Adjoin to it a 20×20 Jordan block with eigenvalue $4i$. Adjoin again a 30×30 Jordan block with eigenvalue $6i$. Continue in this fashion to construct an operator A in ℓ^2 in the form of an infinite block diagonal matrix, where each block is bigger than the last and the eigenvalues are all distinct. The spectrum of the operator is a discrete subset of the imaginary axis: $\Lambda(A) = \{2i, 4i, 6i, 8i, \ldots\}$. For any fixed $\epsilon > 0$, however, the ϵ-pseudospectrum $\Lambda_\epsilon(A)$ is a countable union of disks centered at the eigenvalues whose radii rapidly approach $1+\epsilon$ as one moves further out the imaginary axis. It is hardly surprising that under such circumstances, $\|\exp(tA)\|$ never levels off to a constant, but maintains the value e^t for all $t \geq 0$.

The intuition suggested by this Zabczyk example is in fact valid for arbitrary operators A. For each $\epsilon \geq 0$, define the *ϵ-pseudospectral abscissa* of A by

$$\alpha_\epsilon(A) = \sup_{z\in\Lambda_\epsilon(A)} \operatorname{Re} z. \tag{9}$$

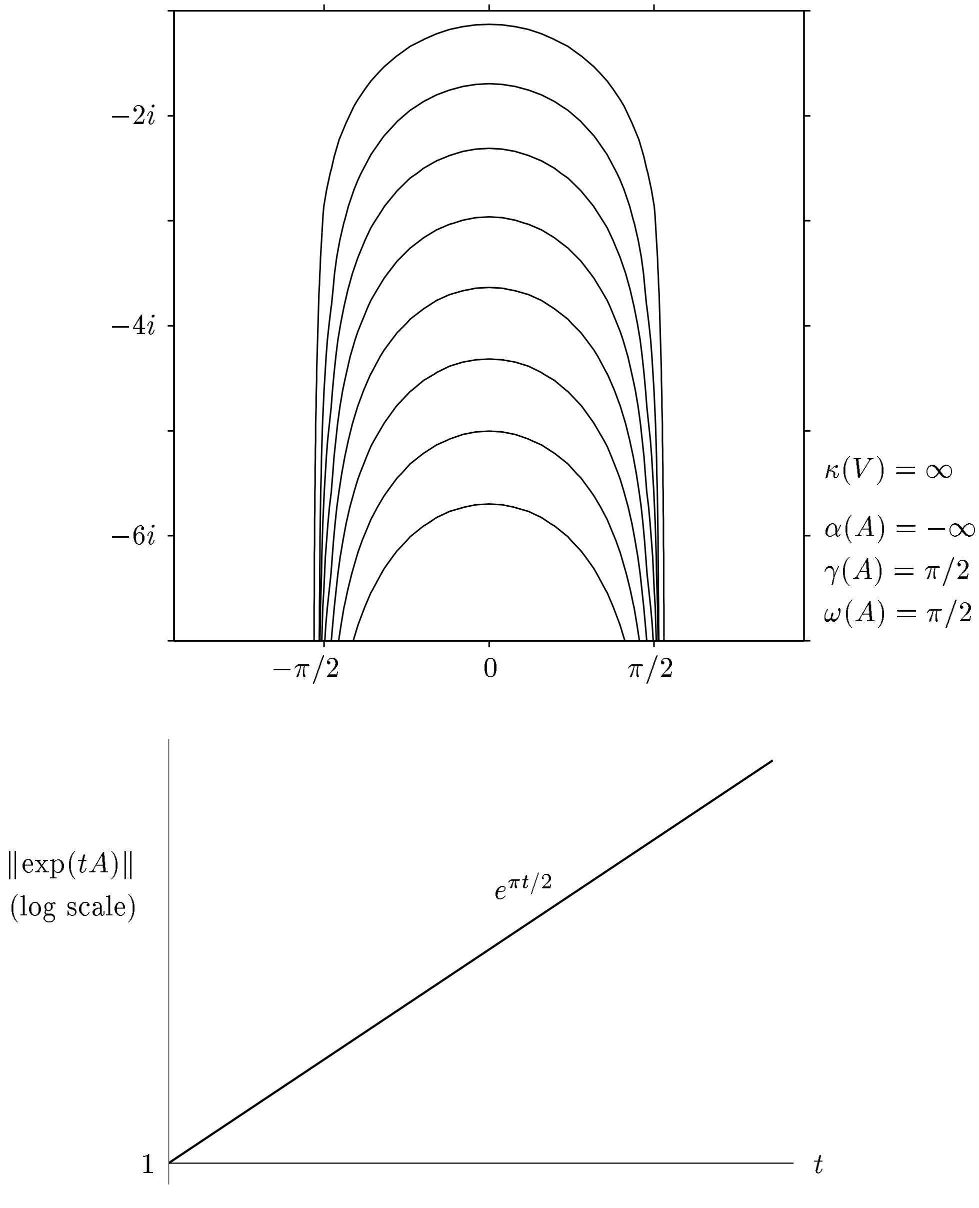


Figure 2: The Hille–Phillips operator, the original example illustrating that $\gamma(A)$ may exceed $\alpha(A)$. The spectrum is empty, but the resolvent norm $\|(zI - A)^{-1}\|$ grows doubly exponentially as z increases to ∞ along the negative imaginary axis. The ϵ-pseudospectral abscissa is $\alpha_\epsilon(A) = \pi/2 + \epsilon$ for each $\epsilon > 0$, and this explains why $\|\exp(tA)\| = e^{\pi t/2}$ for all t (Theorem 2). Contours at $\epsilon = 10^{-1}, 10^{-2}, 10^{-4}, \ldots, 10^{-128}$. (From Baggett [Bag], based on [HiPh].)

The following result may be called *Prüss's theorem,* as it was first established by Prüss in 1984 [Prü], generalizing earlier work by Gearhart and others.

Theorem 1. *For any A, the growth abscissa is equal to the limit $\epsilon \to 0$ of the ϵ-pseudospectral abscissas:*

$$\gamma(A) \;=\; \lim_{\epsilon\to 0} \alpha_\epsilon(A). \tag{10}$$

For the Zabczyk example of Figure 1, we have $\alpha(A)=0$ but $\alpha_\epsilon(A)=1+\epsilon$ for all $\epsilon>0$, and accordingly, $\gamma(A)=1$.

Example 2. Hille–Phillips operator [Bag].

Twenty years before Zabczyk, Hille and Phillips were the first to present an example of an operator with $\gamma(A)>\alpha(A)$ [HiPh]. It is a much less elementary example. They begin with the integral operator

$$[J(s)f](x) \;=\; \frac{1}{\Gamma(s)}\int_0^x (x-u)^{s-1} f(u)\,du \tag{11}$$

in the Hilbert space $L^2[0,1]$. If s is a positive integer, then $J(s)f$ is the $s^{\rm th}$ indefinite integral of f, and for arbitrary positive real s, $J(s)f$ can be interpreted as a fractional integral of f. Hille and Phillips show that $\{J(s)\}$ is in fact a semigroup not only with respect to s but also, if the problem is rotated through an angle $\pi/2$ by analytic continuation, with respect to $t=i^{-1}s$. Let A denote the infinitesimal generator of this semigroup (a rotation by $\pi/2$ of the generator of the fractional integration semigroup). Results for this operator A are presented in Figure 2. This time the spectrum of A is empty: no function $u(x)$ on $[0,1]$ with $u(0)=0$ is its own indefinite integral. The operator is certainly not nilpotent, however; we have $\|\exp(tA)\| = e^{\pi t/2}$ for all $t\ge 0$. Thus $\alpha(A)=-\infty$ and $\gamma(A)=\pi/2$.

Computing the pseudospectra of the Hille–Phillips operator is not easy. Though an explicit integral representation of the resolvent is known, the integrand involved is highly oscillatory and singular at the endpoints. Simple discretizations fail. By a sequence of reformulations of the problem, however, a successful calculation has been carried out by Jeffrey Baggett [Bag], and it is his results that are shown in the figure. The resolvent norm for this operator grows doubly exponentially along the negative imaginary axis: $\log\log\|(zI-A)^{-1}\| \sim |z|$ as ${\rm Im}z\to-\infty$ with ${\rm Re}z=0$.

To the right of the line ${\rm Re}z=\pi/2$, the Hille–Phillips operator "looks normal": we have $\alpha_\epsilon(A)=\pi/2+\epsilon$ for all $\epsilon>0$. This linear dependence is just like

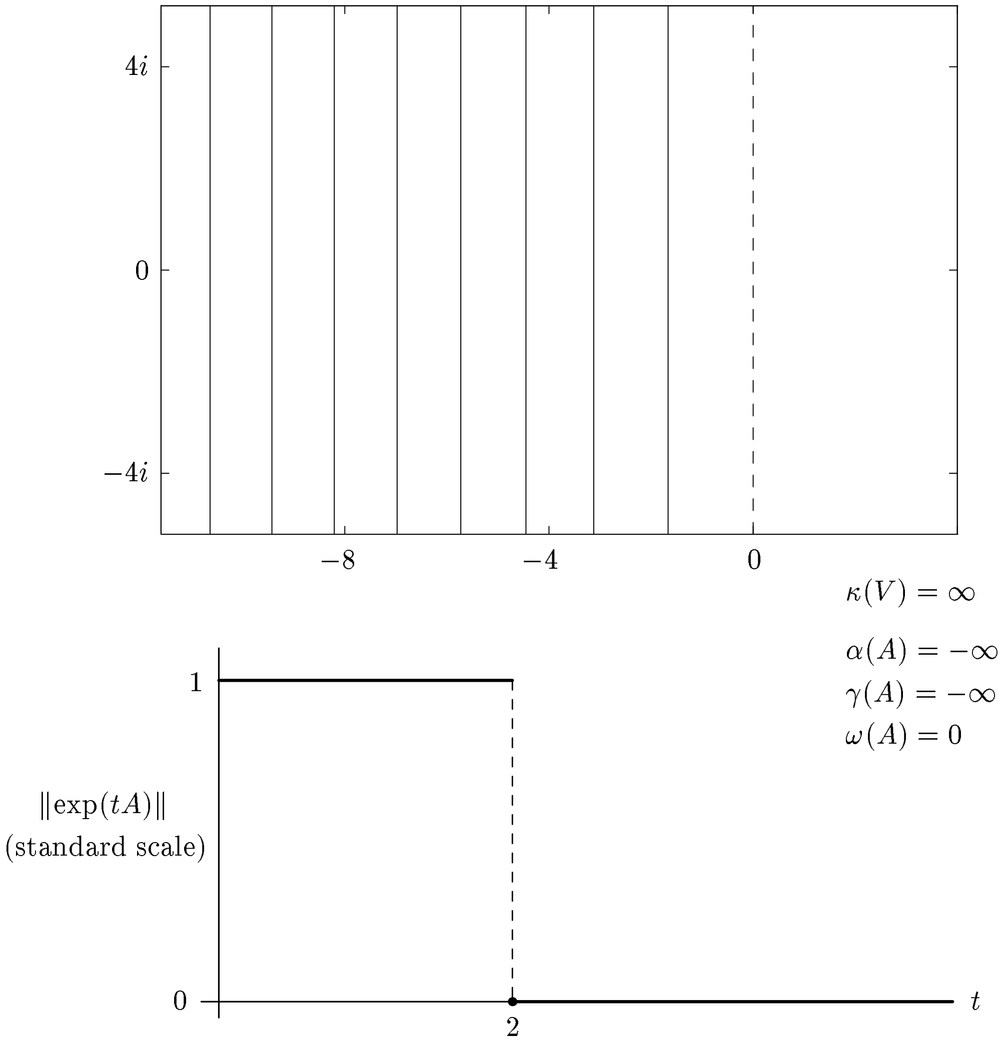


Figure 3: The first-order differentiation operator in $L^2[0,d]$ with boundary condition $u(d)=0$. The resolvent norm grows at the rate $e^{-d\mathrm{Re}z}$ as $\mathrm{Re}z \to -\infty$, which implies that the generator is nilpotent with period d (Theorem 3). The plots correspond to the case $d=2$, with contours at $\epsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-8}$; the pseudospectra are the half-planes to the left of these boundaries. The dashed line (the imaginary axis) is the boundary of the numerical range. (From Reddy [Red] and Driscoll and Trefethen [DrTr].)

what we observed for the Zabczyk example, where we had $\alpha_\epsilon(A) = 1+\epsilon$. As it happens, these examples are illustrative of a general principle: the ϵ-pseudospectral abscissas of an operator depend linearly on ϵ if and only if $\log\|\exp(tA)\|$ depends linearly on t. The following theorem is established in [Bag].

Theorem 2. $\|\exp(tA)\| = e^{\gamma t}$ *for all* $t \ge 0$ *if and only if* $\alpha_\epsilon(A) = \gamma + \epsilon$ *for all* $\epsilon > 0$.

Example 3. Differentiation operator [Red,DrTr].

Our third example, presented in Figure 3, is perhaps a more familiar one. Let A be the derivative operator d/dx acting in $L^2[0,d]$ with boundary condition $u(d) = 0$. The evolution process associated with A is the partial differential equation $\partial u/\partial t = \partial u/\partial x$ with $u(d) = 0$, whose solution is a leftward translation at speed 1; see §19.4 of [HiPh]. After time d, all solutions vanish out the left boundary; thus $\exp(tA) = 0$ for $t \ge d$, as shown in the figure.

This operator A is one of the simplest examples of a non-normal differential operator. An eigenfunction would have to have the form $e^{\lambda x}$ for some λ, but such a function could not satisfy the boundary condition (this differentiation operator is essentially the inverse of the integration (11)). Thus there are no eigenfunctions, and in fact, the spectrum of A is again empty. The resolvent norm, however, grows exponentially in the left half-plane, satisfying

$$\|(zI-A)^{-1}\| \;=\; \frac{e^{d|\mathrm{Re} z|}}{2|\mathrm{Re} z|} + O(|\mathrm{Re} z|^{-1}) \tag{12}$$

as $\mathrm{Re} z \to -\infty$ [DrTr]. (For fixed z with $\mathrm{Re} z < 0$, we also have exponential growth as $d \to \infty$.) The numerical range is exactly the left half-plane.

It is no coincidence that this nilpotent operator has a resolvent norm that grows exponentially with $|\mathrm{Re} z|$. In fact, these two properties are quantitatively related, as can be proved by the use of a Paley–Wiener theorem [DrTr]:

Theorem 3. *An operator* A *has* $\exp(tA) = 0$ *for* $t \ge \tau$ *if and only if* $\Lambda(A)$ *is empty and* $\|(zI-A)^{-1}\| = O(\exp(-\tau \mathrm{Re} z))$ *as* $\mathrm{Re} z \to -\infty$.

Example 4. Wave operator [DrTr].

As a variation on the last example, consider the second-order wave equation $\phi_{tt} = \phi_{xx}$ on $[0,\pi]$ with boundary conditions $\phi(0,t) = 0$ and

$$\phi_x(\pi,t) + \delta\phi_t(\pi,t) \;=\; 0, \tag{13}$$

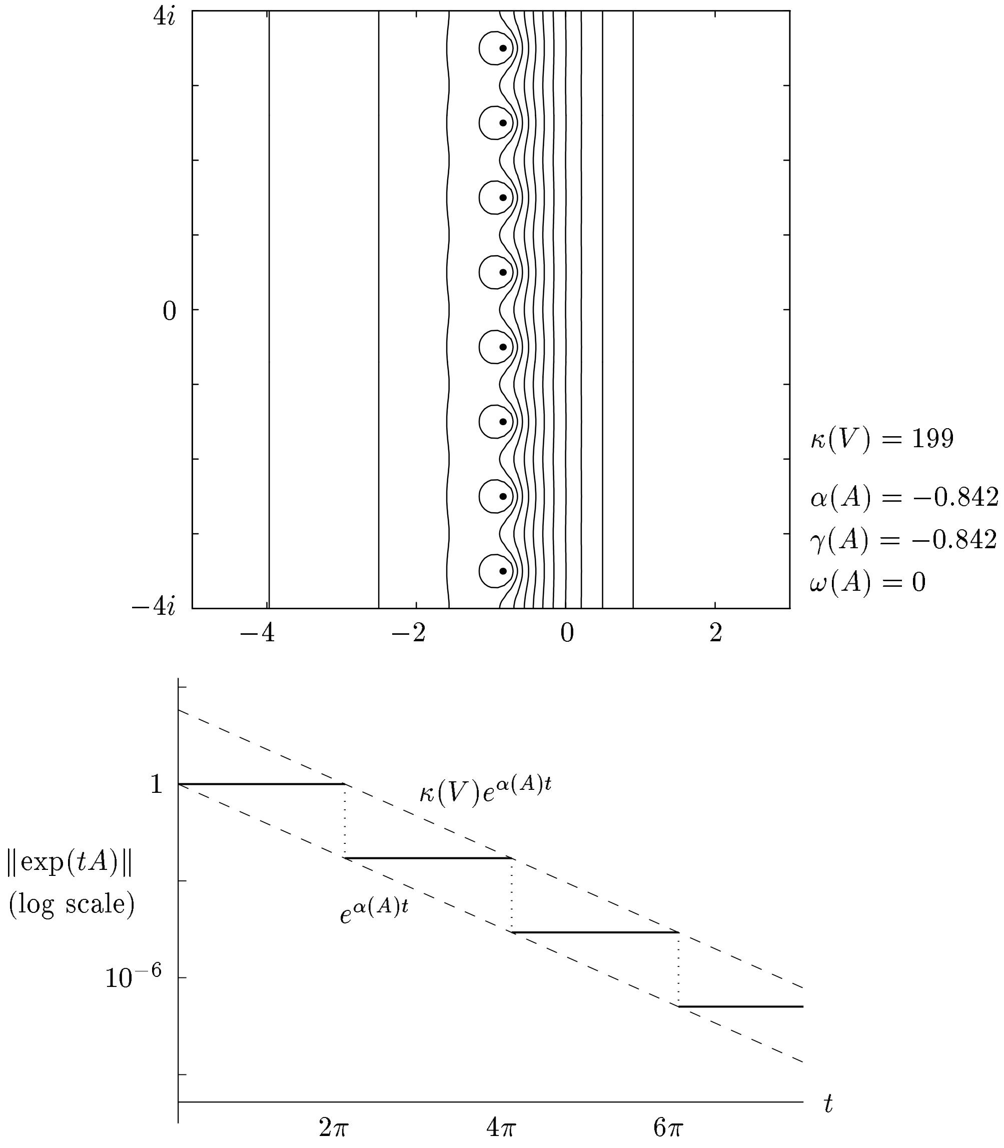


Figure 4: A wave operator with a nearly-absorbing boundary condition. The steps correspond to waves bouncing back and forth between the boundaries; their periodic structure is related to the periodicity of the pseudospectra in the imaginary direction, and their sharpness is possible because A has infinitely many eigenvalues of real part equal to the spectral abscissa. Contours at $\epsilon = 10^0, 10^{-0.2}, 10^{-0.4}, \ldots, 10^{-2}$; the left boundaries of the first seven pseudospectra are not visible on this scale. (From Driscoll and Trefethen [DrTr]; see also [CoZu], [Rid], [Ves].)

where δ is a real constant. (Subscripts denote partial derivatives.) The boundary condition at $x = 0$ has reflection coefficient -1, and the boundary condition at $x = \pi$ has reflection coefficient $R = (1-\delta)/(1+\delta)$. Physically, what is going on is this. Waves propagate left and right in the interval $[0, \pi]$ at speed 1, reflecting off the two boundaries. For $\delta = 1$, the reflection coefficient at the right is $R = 0$, so no energy is reflected and we have a nilpotent process of duration 2π. For $\delta \neq 1$, some energy is reflected; we have an imperfectly absorbing boundary condition. This problem has been studied by Cox and Zuazua [CoZu], Rideau [Rid], and Veselić [Ves], among others.

This second-order problem can be reduced to a first-order problem $du/dt = Au$ by introducing the variable $u = (\phi_x, \phi_t)^T$. The operator A takes the symmetric hyperbolic form

$$A = \begin{pmatrix} 0 & \partial/\partial x \\ \partial/\partial x & 0 \end{pmatrix}, \tag{14}$$

acting in $L^2[0,\pi] \times L^2[0,\pi]$. Figure 4 plots pseudospectra and evolution norms for the case $\delta = 0.99$, where the reflection coefficient is $R = 1/199$. The $\|\exp(tA)\|$ curve has the shape of a staircase, decreasing by the factor R whenever t is an integer multiple of 2π because of the reflections at the right-hand boundary. The spectrum is the regularly spaced set of points

$$\lambda_n = \frac{1}{2\pi}\log R + (n+\frac{1}{2})i, \qquad n \in \mathbf{Z}. \tag{15}$$

Note that the real part of this expression, the spectral abscissa $\alpha(A)$, matches the average rate of decrease of $\|\exp(tA)\|$ as illustrated in the figure. Between multiples of 2π, however, $\|\exp(tA)\|$ is constant. From the operator theory point of view, this discrepancy between the average and the local behavior is possible because of the non-normality of A. The height of the steps, R, is also equal to $\kappa(V)$, the condition number of a normalized infinite matrix (Riesz basis) of eigenfunctions.

Example 5. Convection–diffusion operator [ReTr].

Non-normal differential operators arise most familiarly in problems mixing first and second derivatives, e.g., convection and diffusion. Specifically, consider the operator $A = d/dx + d^2/dx^2$ acting in $L^2[0,d]$ with boundary conditions $u(0) = u(d) = 0$, where d is a positive constant. The spectrum $\Lambda(A)$ is the set of points

$$\lambda_n = -\frac{1}{4} - \frac{\pi^2 n^2}{d^2}, \qquad n \geq 1. \tag{16}$$

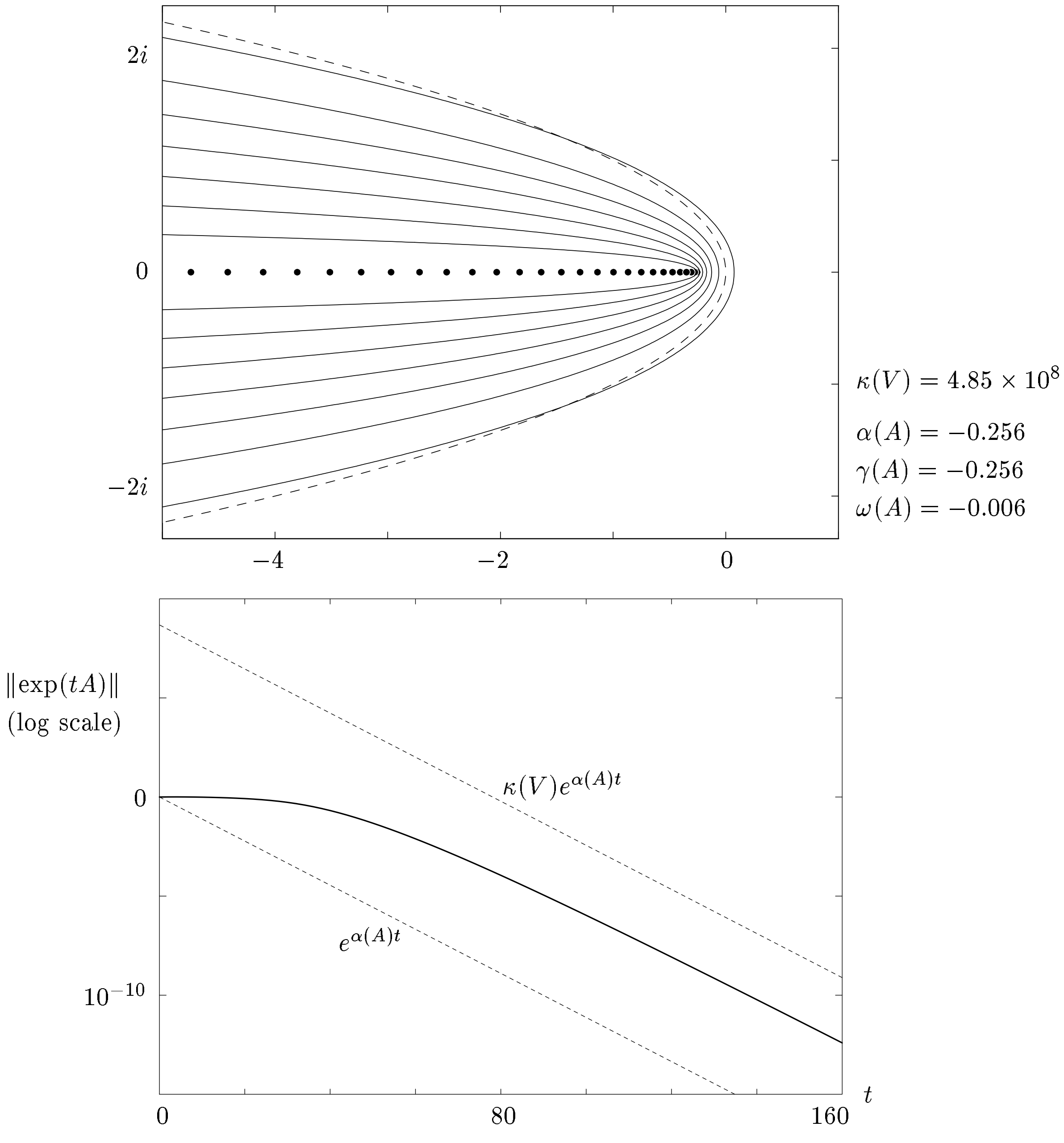


Figure 5: A convection–diffusion operator on the interval $[0, d]$, $d = 40$. Contours at $\epsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-7}$. As $d \to \infty$, the resolvent norm $\|(zI - A)^{-1}\|$ increases exponentially for every z inside the critical parabola $\mathrm{Re}z = -(\mathrm{Im}z)^2$. The dashed curve marks this critical parabola; the boundary of the numerical range is the same curve shifted left by $\pi^2/d^2 \approx 0.0062$ (T. Kato, private communication). (From Reddy and Trefethen [ReTr]; see also [Kat].)

However, any matrix of eigenfunctions has condition number at least $e^{d/2}$, and thus this is a highly non-normal problem for large d. Intuitively, the explanation is that for large d ($\approx$ large Péclet number in the more standard nondimensionalization), convection dominates diffusion. Physically, this means that information travels very inefficiently upstream (rightwards) in the interval $[0, d]$. For example, a smooth pulse beginning near $x = d$ will convect roughly without change to near $x = 0$ before diffusion finally begins to take hold and the shape begins to change to that of the dominant eigenfunction.

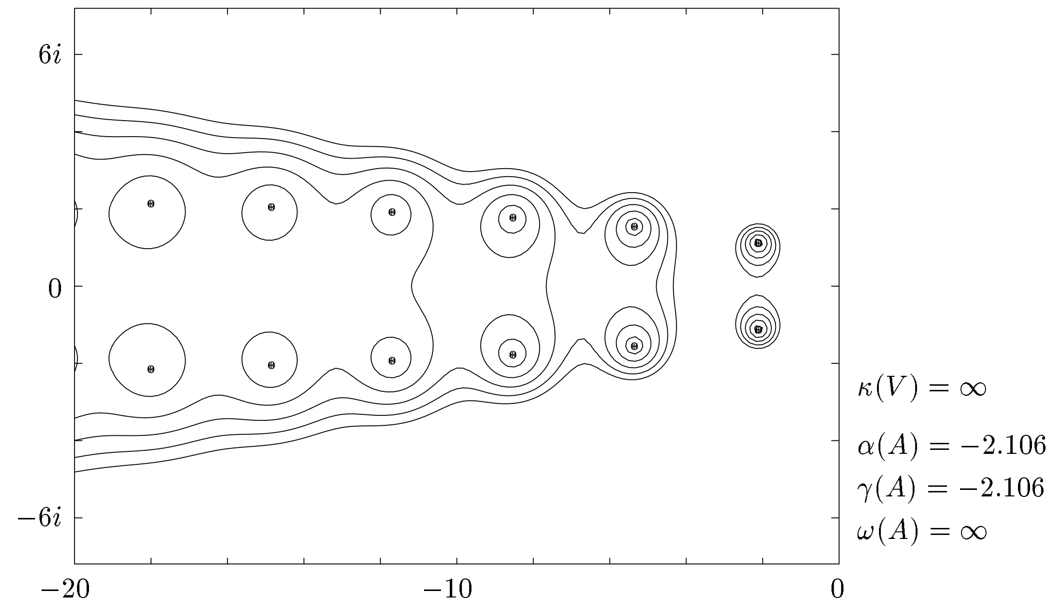
From these considerations one may expect that $\|\exp(tA)\|$ should exhibit a long, nearly-constant transient before eventually settling down to exponential decay. This behavior is apparent in Figure 5. As for the pseudospectra, they approximate parabolas, and it can be proved that, for each z inside the critical parabola $\mathrm{Re}z = -(\mathrm{Im}z)^2$, $\|(zI - A)^{-1}\|$ grows exponentially as $d \to \infty$. Precise estimates are given in [ReTr].

Example 6. Papkovich–Fadle operator.

The Papkovich–Fadle problem, named after two independent authors from around 1940 [Fad], [Pap], concerns the biharmonic operator Δ^2 in a semi-infinite strip. Specifically, we consider functions $u(t, y)$ in the strip $t \geq 0$, $|y| \leq 1$ satisfying $u_{tttt} + 2u_{ttyy} + u_{yyyy} = 0$ in the interior and $u(t, \pm 1) = u_y(t, \pm 1)$ along both infinite boundaries (subscripts denote partial differentiation). Along the end of the strip, data $u_y(0, y) = f(y)$ and $u_t(0, y) = g(y)$ are prescribed. This problem has been studied by a number of authors over the years for applications in solid mechanics, where u represents the shape of a semi-infinite clamped plate, and in fluid mechanics, where it is the stream function of an incompressible fluid flow at low Reynolds number (Stokes flow). Some of the many contributors to this literature have been Gregory [Gre], Joseph and Sturges [JoSt], and Spence [Spe].

The reason for the use of the variable t is that it is natural to view the Papkovich-Fadle problem as an evolution process with respect to that variable. How do conditions applied at the end of the strip determine the solution for larger t? One way to formulate the problem is in terms of the first derivatives, giving an evolution equation

$$\frac{d}{dt} \begin{pmatrix} u_y(t, \cdot) \\ u_t(t, \cdot) \end{pmatrix} = A \begin{pmatrix} u_y(t, \cdot) \\ u_t(t, \cdot) \end{pmatrix}, \tag{17}$$



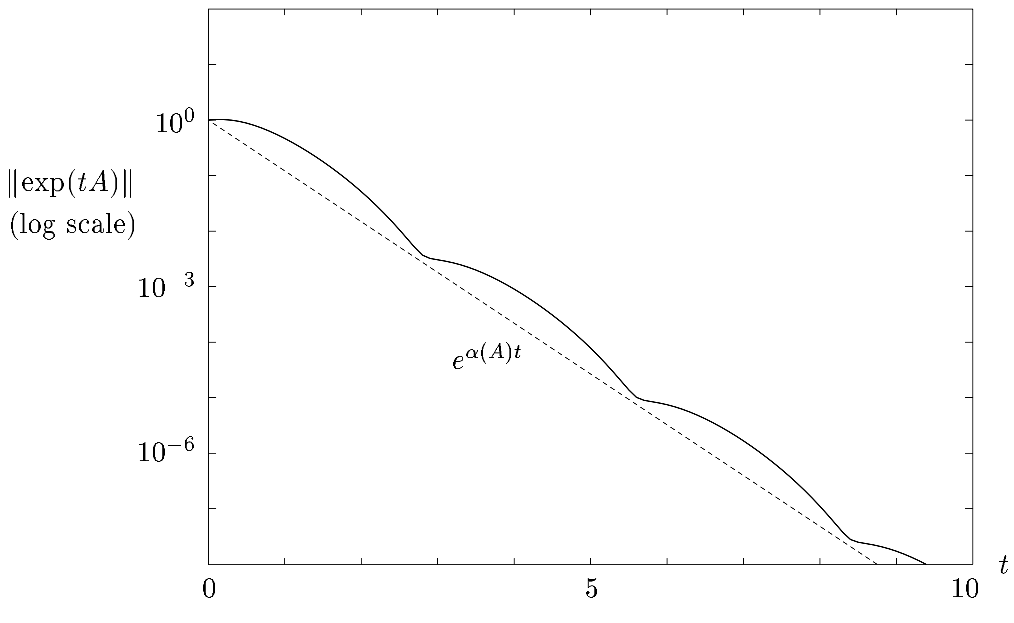


Figure 6: The Papkovich–Fadle operator derived from the biharmonic operator on a semi-infinite strip. The continued oscillations of $\|\exp(tA)\|$ as $t \to \infty$ are possible because A has more than one eigenvalue of real part equal to the spectral abscissa. Contours at $\epsilon = 0.25, 0.20, 0.15, 0.10, 0.05$. (From unpublished work by Weideman and Trefethen; see also [Fad], [Gre], [JoSt], [Pap], [Spe].)

with solution

$$\begin{pmatrix} u_y(t,\cdot) \\ u_t(t,\cdot) \end{pmatrix} = \exp(tA) \begin{pmatrix} f \\ g \end{pmatrix}. \tag{18}$$

Thus the data in the evolution problem are block 2-vectors, and the solution operators are block 2×2 matrices.

Figure 6 displays the results. Physically, the regular oscillations in the curve $\|\exp(tA)\|$ can be interpreted as follows. In a semi-infinite channel of fluid, they correspond to an infinite succession of counter-rotating vortices. (Since the amplitudes decay exponentially, only the first three or four of these can be observed in the laboratory.) In a semi-infinite solid plate, they correspond to back-and-forth pattern of bending one can induce (again with exponentially decaying amplitude) by twisting the plate at the end.

The degree of non-normality of this operator is in some respects mild; note the large values of ϵ chosen for the plot of pseudospectra. However, the non-normality becomes more pronounced as one moves deeper into the left half-plane. This corresponds to the singularity of this problem at the corners of the strip, a phenomenon of small space scales that can only be resolved by the higher eigenfunctions. In fact, the numerical abscissa of this operator is $\omega(A) = \infty$, and thus, by (7), though it is not visible in the figure, the slope of the $\|\exp(tA)\|$ curve is infinite at $t = 0$.

This example is the result of unpublished joint work with André Weideman. Further details will appear in [Tre3].

Example 7. Wiener–Hopf operator [Red].

Among non-normal matrices, the class whose pseudospectra are best understood are the Toeplitz matrices [ReiTr]. We can generalize these to operators of infinite dimension by making the domain unbounded, giving an infinite Toeplitz matrix, by making it continuous, giving a Wiener–Hopf integral operator, or both, giving a Wiener–Hopf integral operator on an unbounded domain. Pseudospectra of such operators have been studied by Reddy [Red] and by Böttcher and his colleagues in Chemnitz [Böt], [BöWo], [RoSi]. The latter group have proved a number of theorems regarding, for example, convergence of the pseudospectra of the finite sections of an infinite Toeplitz matrix to those of the infinite matrix. For related developments in the context of waveform relaxation, see [LuWu], and for Abel integral operators, see [Pla].

The behavior of Toeplitz and Wiener–Hopf operators can be quite wild, but pseudospectra of the wilder examples have not yet been computed. Figure 7

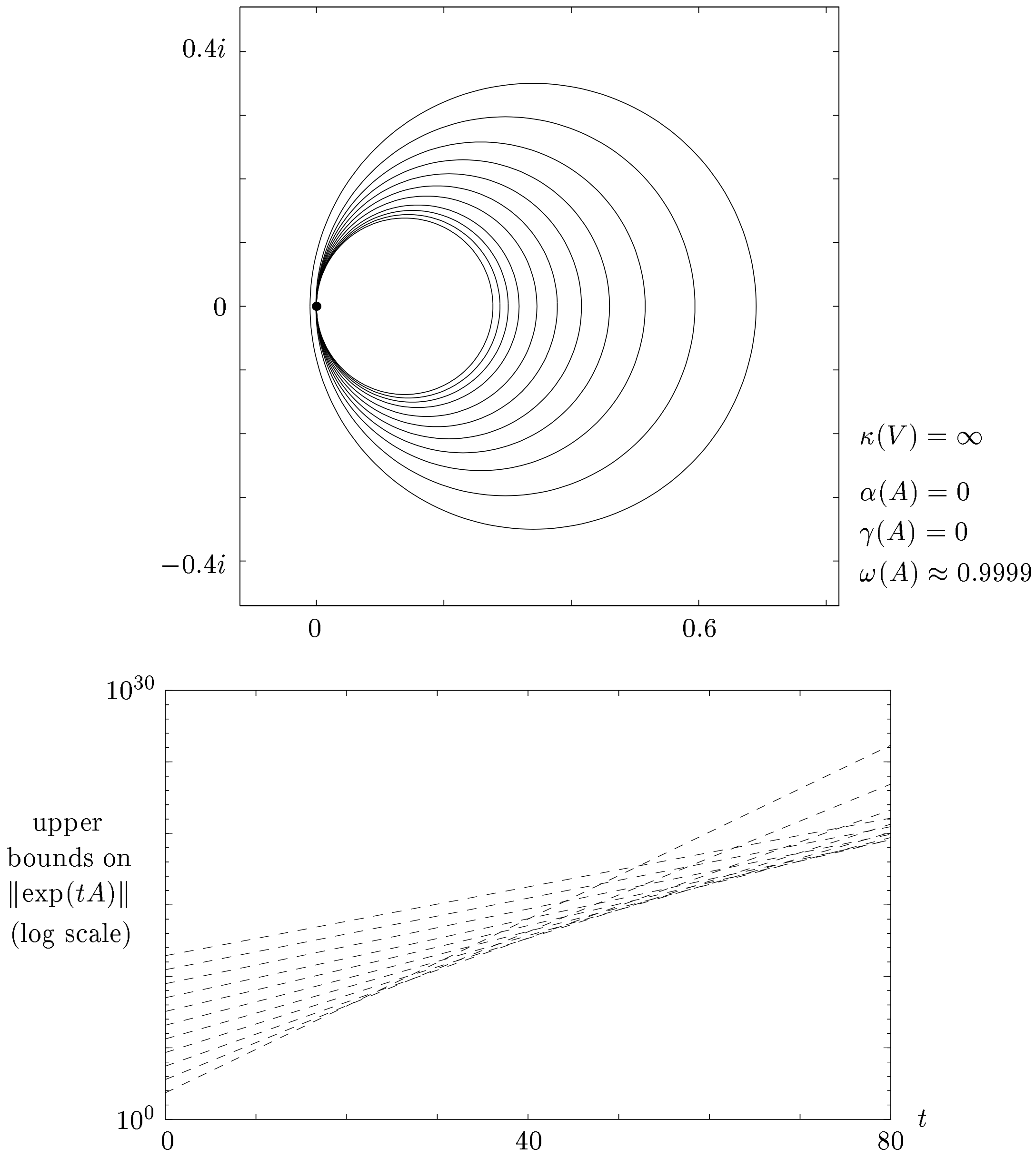


Figure 7: The Volterra Wiener–Hopf integral operator (19) with kernel $k(x) = e^x$ on $(-\infty, 0]$. The spectrum is just the origin. Contours at $\epsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-12}$. In the lower plot, since $\|\exp(tA)\|$ has not been computed for this example, eleven upper bounds from Theorem 4 are shown as dashed lines. (From Reddy [Red].)

shows an example from [Red] that is more sedate, since it is also a Volterra operator (analogous to a triangular Toeplitz matrix). In the space $L^2[0,d]$, A is here the integral operator

$$[Af](x) \;=\; \int_x^d e^{(x-y)} f(y)\, dy. \tag{19}$$

The spectrum is just the origin, but the pseudospectra approximate disks tangent to the origin in the right half-plane. The figure corresponds to $d = 10$.

For this example, $\|\exp(tA)\|$ has not been computed. In its place, let us take this opportunity to illustrate the simplest manner in which bounds on norms of functions of A can be obtained from pseudospectra. For any analytic function f, we have

$$f(A) \;=\; \frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} f(z)\, dz, \tag{20}$$

where Γ is any contour enclosing the spectrum of A [Kat]. One interesting choice of contour is the boundary $\partial\Lambda_\epsilon(A)$ of $\Lambda_\epsilon(A)$ for some $\epsilon > 0$, and from this choice, with $f(z) = e^{tz}$, we get the following bound:

Theorem 4. *For any $t \geq 0$ and $\epsilon > 0$ we have*

$$\|\exp(tA)\| \;\leq\; \frac{\epsilon^{-1}}{2\pi}\int_{\partial\Lambda_\epsilon(A)} e^{t\,\mathrm{Re}\, z}\, |dz|. \tag{21}$$

For each ϵ, this theorem gives us a function of t that is an upper bound for $\|\exp(tA)\|$. Suppose we take the eleven values of ϵ in Figure 7, approximate $\partial\Lambda_\epsilon(A)$ by a circle of diameter r_ϵ tangent to the origin, and bound the integrand above by $\exp(tr_\epsilon)$. This gives eleven bounds $\|\exp(tA)\| \leq \epsilon^{-1} r_\epsilon e^{tr_\epsilon}$ for the points (ϵ, r_ϵ). These are the dashed lines presented in the lower plot.

Example 8. Orr–Sommerfeld operator [RSH].

With our final three examples, we move to a topic in the field of fluid mechanics: the instability of incompressible flows in pipes and channels [DrRe]. In the past five years it has become clear that non-normality plays a crucial role in destabilizing these flows.

The problem of explaining why high-speed flows are invariably turbulent is more than a century old. The foundations were laid by Osborne Reynolds in 1883, who investigated flow of water through a long circular pipe [Rey]. At low speeds, a laminar flow is observed, corresponding to an easily identified solution

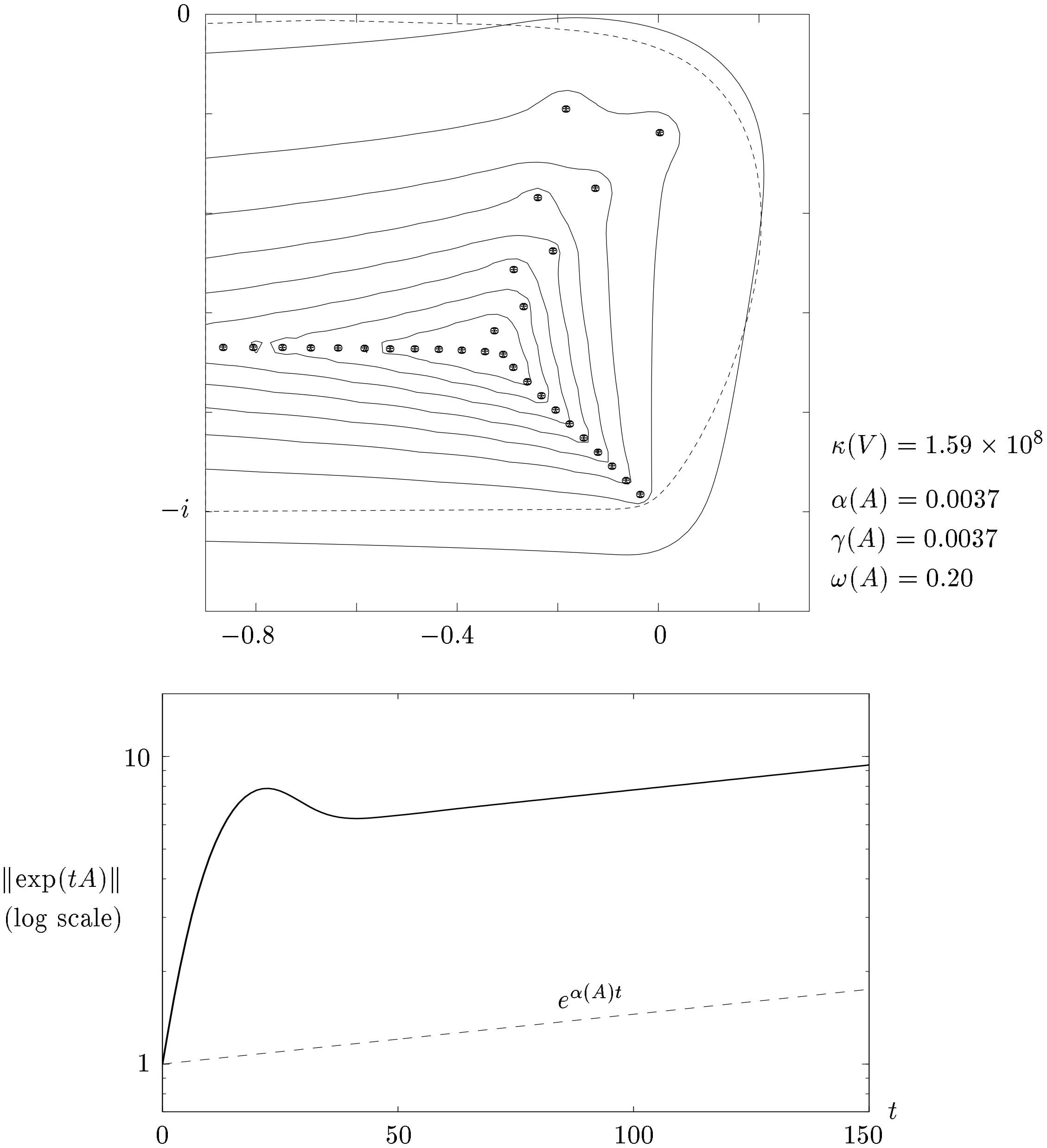


Figure 8: The Orr–Sommerfeld operator at Reynolds number $R = 10000$ with Fourier parameters $\alpha = 1$, $\beta = 0$. Contours at $\epsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-8}$; in the area of the cross, the resolvent norm $\|(zI - A)^{-1}\|$ increases exponentially as a function of R. The dashed curve marks the boundary of the numerical range. (From Reddy, Schmid and Henningson [RSH]; see also [DrRe], [Ors].)

of the Navier–Stokes equations (longitudinal flow with velocity profile given by a parabola, assuming the idealization of an infinitely long pipe). At higher speeds, however, instabilities invariably set in, and the laminar solution breaks down. Reynolds' careful experiments cast this phenomenon in a clear light, showing that the critical parameter is not just velocity but velocity times pipe radius divided by kinematic viscosity—the dimensionless quantity now known as the Reynolds number, denoted here by R.

Eigenvalue analysis is the obvious tool to explain these observations. The Naiver–Stokes equations are nonlinear, but one can linearize them by considering infinitesimal perturbations about the laminar flow. Formally speaking, this gives an evolution equation $du/dt = Au$ for some linear operator A (of course the details are nontrivial; this is the field of computational fluid dynamics). Does the spectrum of A extend into the right half-plane? If so, some perturbations will grow unstably until nonlinear effects become important, which may then lead to turbulence. If not, then all perturbations eventually decay, and the laminar flow should be stable and observable in the laboratory.

The literature on problems of hydrodynamic stability is enormous. Early in this century, the technique of eigenvalue analysis became established as the standard one, receiving little challenge from 1930 to 1990 [DrRe]. This is remarkable, since for the most basic flows in question, including flow in a pipe, it entirely fails to predict what is observed.

Our first fluids example, shown in Figure 8, corresponds to the flow of this kind that has been studied most by applied mathematicians. This is flow between two infinite parallel plates, known as *plane Poiseuille flow*—more difficult than a circular pipe in the laboratory, but easier analytically. The standard analysis runs as follows. The problem is Fourier transformed in the streamwise (x) and spanwise (z) variables, giving Fourier parameters α and β. What remains is an operator acting on functions of the cross-stream variable y, with a discrete spectrum. One investigates this spectrum for all α and β, looking for the smallest Reynolds number R at which an eigenvalue crosses into the right half-plane. It was proved by Squire in 1933 that this first crossing occurs for a two-dimensional flow, i.e., $\beta = 0$ — *"Squire's theorem"* [DrRe], [Squ]. The first high-accuracy computations were carried out by Orszag in 1971 [Ors], who found that the critical Reynolds number is $R \approx 5772$, with $\alpha \approx 1.02$. Thus, plane Poiseuille flow is mathematically stable for $R < 5772$ and unstable for $R > 5772$. The unstable mode for $R > 5772$ has been extensively studied and is known as a *Tollmien–Schlichting wave.*

However, the news from the laboratory is entirely different. Here, instability

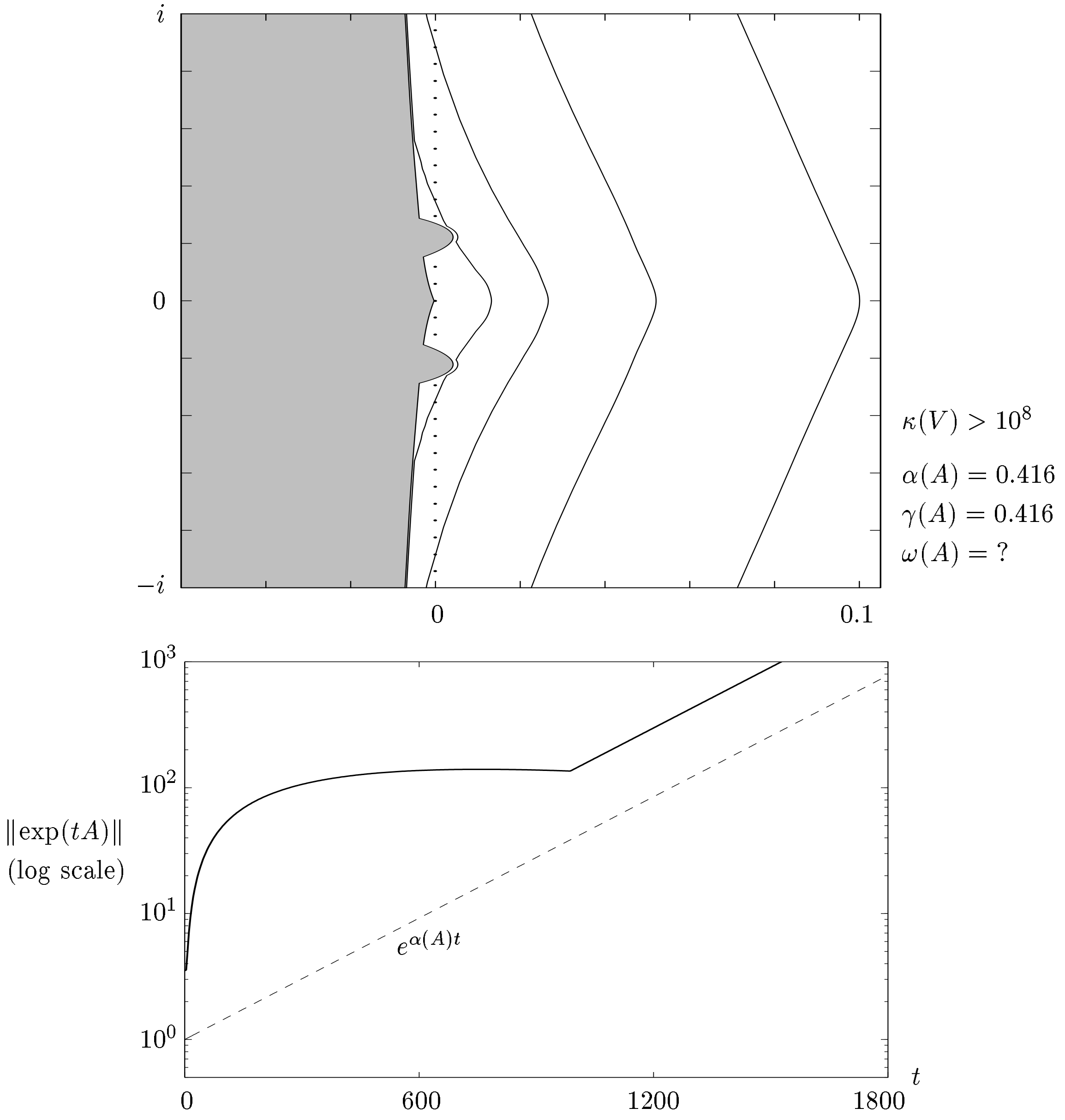


Figure 9: Plane Poiseuille flow operator at Reynolds number $R = 10000$. The shaded region is the spectrum, which is a 2D continuum. Note that the real and imaginary axes are scaled differently. Physically, the important feature of this example is the rapid transient amplification by a factor $\approx 10^2$, corresponding to a simple vortex $\rightarrow$ streak mechanism involved in transition to turbulence and in turbulence itself. Contours at $\epsilon = 10^{-2}, 10^{-2.5}, 10^{-3}, 10^{-3.5}$. (From Trefethen, Trefethen, Reddy and Driscoll [TTRD]; see also [BuFa], [Gus], [ReHe].)

and turbulence usually appear with $R \approx 1000$. The form and the time scale of the structures that appear bear no resemblance to Tollmien–Schlichting waves. Indeed, at any Reynolds number, Tollmien–Schlichting waves are all but unobservable except in experiments that take special steps to make them appear. Why?

Reddy, Schmid, and Henningson were among the first to investigate this problem from the point of view of non-normality. Figure 8, from [RSH], shows the pseudospectra of the "Orr–Sommerfeld operator" implicitly described above for $R = 10000$, $\alpha = 1$. At this high Reynolds number, one of the eigenvalues lies in the right half-plane—though only barely, with $\alpha(A) = 0.00374$. The rest of them form a Y-shape in the left half-plane that has been studied by a number of authors. Reddy et al. made the startling observation that the portion of the spectrum associated with the center of this Y is exceedingly non-normal. The basis of eigenfunctions for this example has condition number at least 1.59×10^8, a figure that rises with R at an exponential rate $C^{\sqrt{R}}$.

Based on Figure 8, then, one may conclude that the Orr–Sommerfeld operator is highly non-normal and that its eigenvalues near the center of the Y are so sensitive to perturbations that they are unlikely to be physically significant. The dominant eigenvalue, however, still appears robust enough. By itself, this example does not yet explain why transition at $R = 5772$ is not observed in the laboratory.

Example 9. Plane Poiseuille flow operator [TTRD].

What is missing from the analysis above is three-dimensionality. Eigenvalue analysis—Squire's theorem—predicts that only two-dimensional perturbations need to be considered to explain transition to turbulence. This prediction has never matched experiments, however, and in recent years it has become apparent what is wrong with it mathematically. Most of the action in the linearized Navier–Stokes problem is in three-dimensional structures that have nothing to do with eigenmodes.

Figure 9, based on large-scale computations reported in [TTRD], illustrates the situation. Here, instead of reducing the linearized Navier–Stokes problem by Fourier parameters α and β, we consider the whole operator acting on three-dimensional perturbations of the laminar flow. The spectrum changes from a discrete set to a 2D continuum, shown in the upper plot as a shaded region. Since $R > 5772$, two "Tollmien–Schlichting bumps" in the spectrum extend into the right half-plane, corresponding to weak exponential instability.

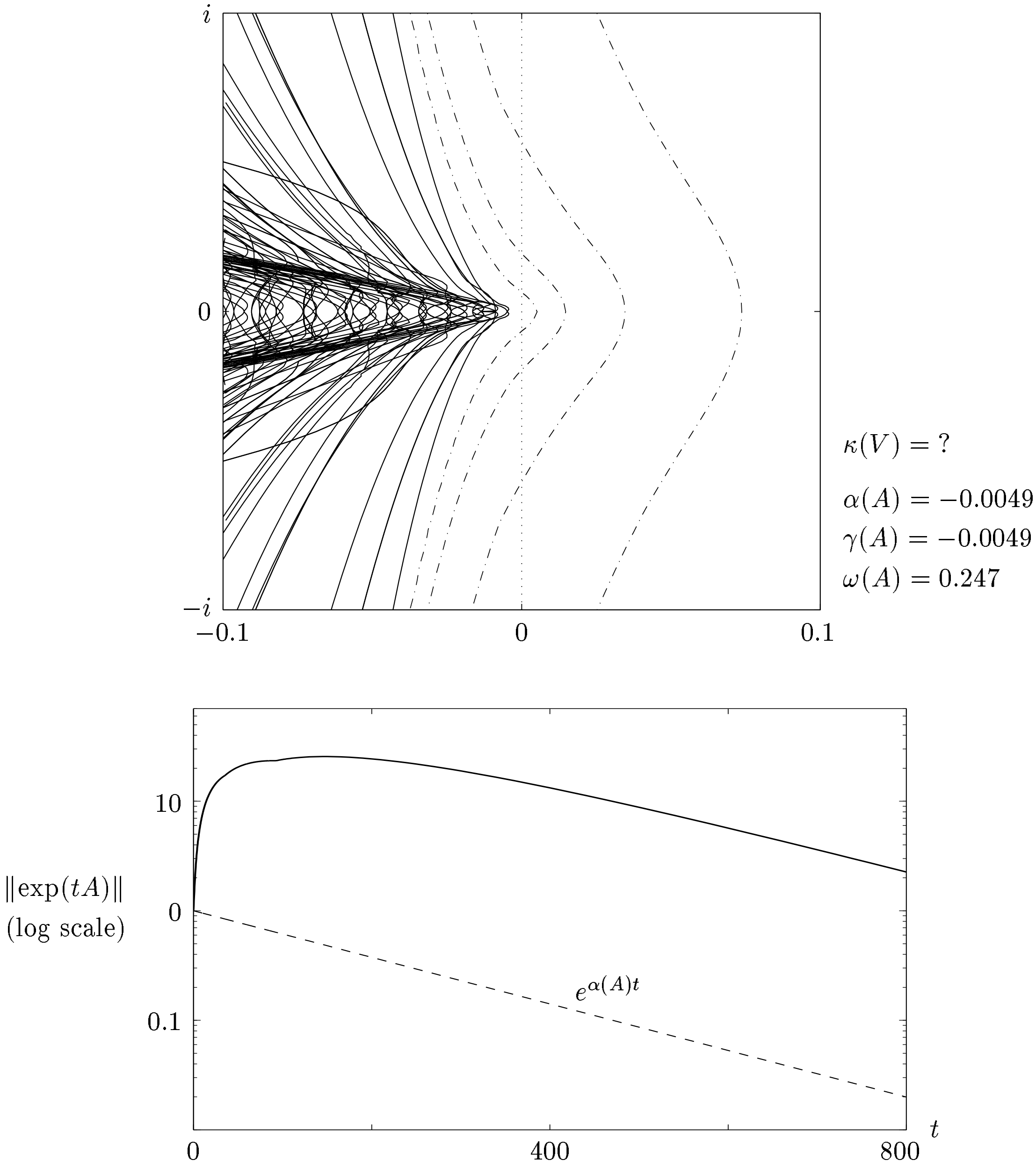


Figure 10: Pipe Poiseuille flow operator at Reynolds number $R = 3000$. The solid curves are the spectrum, which is very complicated; the gaps between them are genuine. The dot-dash curves are the boundaries of the ϵ-pseudospectra (just those portions to the right of the spectrum) for $\epsilon = 10^{-2}, 10^{-2.5}, 10^{-3}, 10^{-3.5}$. (From Trefethen, Trefethen and Schmid [TTS]; see also [Ber], [BoBr], [OSBr], [ScHe].)

But much more significant in the upper plot are the pseudospectra, which protrude substantially into the right half-plane. These indicate that the function $\|\exp(tA)\|$ must exhibit substantial transient growth before eventually settling down to behavior controlled by the spectral abscissa.

The lower plot of Figure 9 reveals this transient growth. On a rapid time scale, $\|\exp(tA)\|$ grows by a large factor of about 100 and lingers at a high value for a long time. (Both the amplification factor and the length of this time interval grow like $O(R)$ as $R \to \infty$.) Only for large t does the weak exponential growth associated with the eigenvalue begin to catch up. In the nonlinear problem governed by the Navier–Stokes equations, other effects are likely to have taken over before that eigenmode is ever seen. This, very briefly, appears to be the reason why eigenmodal effects are so hard to observe in the laboratory for these flows.

Physically, the transient amplification just described corresponds to a simple mechanism of streamwise vortices generating streamwise streaks. The physics is old and well understood; but until recent years, it was not appreciated how such effects could be reconciled with eigenvalue analysis. The first paper with a complete view of such a reconciliation was a beautiful 1988 work by Boberg and Brosa [BoBr]. That paper, however, went largely unnoticed for several years, leaving it to the independent and equally impressive 1992 paper by Butler and Farrell [BuFa] to communicate such ideas widely. Following closely after [BuFa] were the third and fourth members of what now appear to be a four-paper set, [ReHe] and [TTRD].

It was mentioned above that if the pseudospectra of an operator A protrude significantly into the right half-plane, then there must be a transient hump in the curve of $\|\exp(tA)\|$. One precise statement of this relationship is the following, derived from the Laplace transform:

Theorem 5. *For any $\epsilon > 0$ we have*

$$\sup_{t\geq 0} \|\exp(tA)\| \;\geq\; \epsilon^{-1}\alpha_\epsilon(A). \tag{22}$$

In particular, the inequality remains valid if the right-hand side is replaced by its supremum over $\epsilon > 0$.

For the present problem, we can see for example from the figure that for $\epsilon = 10^{-3.5}$, $\alpha_\epsilon(A)$ is about 0.012. Thus the $10^{-3.5}$-pseudospectrum extends about 38 times further into the right half-plane than would be expected for a normal operator with spectrum filling the left half-plane; by Theorem 5, this implies

$\sup_{t\geq 0}\|\exp(tA)\| \geq 38$. The actual size of the hump, before the exponential growth takes over, is about 140. By choosing the optimal value of ϵ in (22) rather than $10^{-3.5}$, we can improve 38 to a lower bound of about 97.

Theorem 5 corresponds to the "easy half of the Kreiss matrix theorem." The hard half, which derives an upper bound on $\|\exp(tA)\|$ from the pseudospectra, applies only to matrices, as the bound is proportional to the dimension m; see [DKS], [Kre], [RiMo], [WeTr]. In any particular case, however, upper bounds can generally be derived via contour integrals such as that of Theorem 4.

Example 10. Pipe Poiseuille flow operator [TTS].

Our final example looks very different and yet is physically almost the same. We now consider Reynolds' original problem of flow through a circular pipe instead of an infinite channel, following large-scale computations reported in [TTS]. This is often known as *pipe Poiseuille flow.*

In the channel, we had an eigenvalue in the right half-plane for $R > 5772$. In a pipe, however, the failure of eigenvalue analysis is plainer. There are no eigenvalues in the right half-plane at all, regardless of the size of R, yet in the laboratory, transition to turbulence is typically observed if R is greater than about 2000. The pseudospectra, illustrated in Figure 10, give some indication of how this can happen. They look much the same as in the last example, protruding substantially into the right half-plane; the effect is slightly less pronounced than before only because we have chosen a lower Reynolds number (to keep the complexity of the spectrum manageable). Again, these pseudospectra in the right-half place correspond to an $\|\exp(tA)\|$ curve that exhibits large transient growth, of amplitude and duration scaling with R as $R \to \infty$. The lower bound of Theorem 5 again captures more than half of the size of the hump. Thus, as in channels, it appears that turbulence in pipes typically comes about because small perturbations of the flow are amplified linearly into much larger persistent structures approximately in the form of streamwise streaks; these then undergo further evolution in a manner dependent on the nonlinear terms in the Navier–Stokes equations [BDT]. The elucidation of the details of this process is an active area of current research [RSB].

What about the bizarre spectrum depicted in Figure 10? For an infinite channel, there were two unbounded dimensions and thus two continuous Fourier parameters, giving rise to a spectrum that was a 2D continuum. For a pipe, however, there is only one unbounded dimension and only one continuous Fourier parameter. The spectrum becomes a collection of one-dimensional curves and is evidently of great complexity. The figure includes all of the components of the spectrum in the portion of the complex plane displayed; for values of z between these curves, $\|(zI - A)^{-1}\|$ is finite. The picture seems astonishingly complicated, considering that the operator in question describes nothing more than the evolution of infinitesimal perturbations of laminar flow in a pipe. But all of this excitement in the left half-plane is of no importance physically; what matter are the pseudospectra in the right half-plane.

6 Looking to the Future

Eigenvalue problems arise throughout computational science and engineering, most often in the context of differential or integral operators that are reduced by discretization to large matrices. Many of these problems are nonsymmetric, which usually means non-normal, and research into improved methods for computing their eigenvalues, especially iterative methods, is actively underway. To date, most such computations generate just numbers, not pictures, and most of the people who carry them out are not in the habit of investigating non-normality. As a result, it is likely that situations in which eigenvalues give misleading information have often gone unrecognized.

As computers continue to grow more powerful, however, and as the habit of interacting with them graphically becomes more strongly established, this state of affairs will change. I expect that within a few years, the plotting of eigenvalues will become a habitual practice among scientists and engineers and a standard feature offered by software packages. Once this happens, it will be natural to include information about pseudospectra in the plots, too, even if only crude estimates. We will begin to notice things that went unnoticed before, and this will change the way we think about eigenvalue problems.

Acknowledgments

This paper is the product of years of collaboration with a number of students and colleagues. I would like especially to thank Jeff Baggett, Toby Driscoll, Satish Reddy, Peter Schmid, Anne Trefethen, and André Weideman for bringing the operators shown on these pages to life. For comments on the present paper, my thanks are due to most of these and to others, including Anne Greenbaum, Kurt Riedel, and Kim-chuan Toh. This work has been supported by NSF Grant DMS-9500975CS and DOE Grant DE-FG02-94ER25199.

References

[Bag] J. S. Baggett: *Pseudospectra of an operator of Hille and Phillips,* Rep. 94-15, Interdisciplinary Project Center for Supercomputing, ETH-Zurich, August, 1994.

[BDT] J. S. Baggett, T. A. Driscoll, and L. N. Trefethen: *A mostly linear model of transition to turbulence,* Phys. Fluids A **7** (1995), 833–838.

[Ber] L. Bergström: *Optimal growth of small disturbances in pipe Poiseuille flow,* Phys. Fluids A **5** (1993), 2710–2720.

[BoBr] L. Boberg and U. Brosa: *Onset of turbulence in a pipe,* Z. Naturforschung **43a** (1988), 697–726.

[Böt] A. Böttcher: *Pseudospectra and singular values of large convolution operators,* J. Int. Eqs. Applics. **6** (1994), 267–301.

[BöWo] A. Böttcher and H. Wolf: *Spectral approximation for Segal–Bargmann space Toeplitz operators,* preprint, December 1994.

[BrHi] T. Braconnier and N. J. Higham: *Computing the field of values and pseudospectra using the Lanczos method with continuation,* manuscript, October, 1995.

[BuFa] K. M. Butler and B. F. Farrell: *Three-dimensional optimal perturbations in viscous shear flow,* Phys. Fluids A **4** (1992), 1637–1650.

[CoZu] S. Cox and E. Zuazua: *The rate at which energy decays in a string damped at one end,* Indiana Math. J. **44** (1995), 545–573.

[Dem] J. W. Demmel: *A counterexample for two conjectures about stability,* IEEE Trans. Aut. Control **AC-32** (1987), 340–342.

[DKS] J. L. M. van Dorsselaer, J. F. B. M. Kraaijevanger, and M. N. Spijker: *Linear stability analysis in the numerical solution of initial value problems,* Acta Numerica 1992, Cambridge U. Press, 1992, pp. 199–237.

[DrRe] P. G. Drazin and W. H. Reid: *Hydrodynamic Stability,* Cambridge U. Press, Cambridge, 1981.

[DrTr] T. A. Driscoll and L. N. Trefethen: *Pseudospectra of the wave operator with an absorbing boundary,* J. Comp. Appl. Math., in press.

[DuSc] N. Dunford and J. Schwartz: *Linear Operators I, II, III*, Interscience Publ., New York, 1958, 1973, 1971.

[Fad] J. Fadle: *Die Selbstpannungs-Eigenwertfunktionen der quadratischen Scheibe,* Ing.-Arch. **11** (1941), 125–149.

[GoKr] I. C. Gohberg and M. G. Kreĭn: *Introduction to the Theory of Linear Nonselfadjoint Operators,* Amer. Math. Soc., Providence, RI, 1969.

[GrTr] A. Greenbaum and L. N. Trefethen: *Do the pseudospectra of a matrix determine its behavior?*, Lin. Alg. Applics., to appear.

[Gre] R. D. Gregory: *The traction boundary value problem for the elastostatic semi-finite strip; existence of solution, and completeness of the Papkovich–Fadle eigenfunctions,* J. Elasticity **10** (1980), 295–327.

[Gus] L. H. Gustavsson: *Energy growth of three-dimensional disturbances in plane Poiseuille flow,* J. Fluid Mech. **224** (1991), 241–260.

[Hal] P. R. Halmos: *A Hilbert Space Problem Book*, 2nd. ed., Springer-Verlag, New York, 1982.

[HiPh] E. Hille and R. S. Phillips: *Functional Analysis and Semi-groups,* Amer. Math. Soc., Providence, RI, 1957.

[JoSt] D. D. Joseph and L. Sturges: *The convergence of biorthogonal series for biharmonic and Stokes flow edge problems: part II*, SIAM J. Appl. Math. **34** (1978), 7–26.

[Kat] T. Kato: *Perturbation Theory for Linear Operators,* Springer-Verlag, New York, 1976.

[KoRa] V. I. Kostin and S. I. Razzakov: *On convergence of the power orthogonal method of spectrum computing,* Trans. Inst. Math. Sib. Branch Acad. Sci., v. 6, 55–84, translated in *Software Optimization,* New York, 1986.

[Kre] H.-O. Kreiss: *Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren,* BIT **2** (1962), 153–181.

[Lan] H. J. Landau: *On Szegő's eigenvalue distribution theorem and non-Hermitian kernels,* J. d'Analyse Math. **28** (1975), 335–357.

[Lui] S.-H. Lui: *Computation of pseudospectra by continuation,* SIAM J. Sci. Comput., to appear.

[LuWu] A. Lumsdaine and D. Wu: *Spectra and pseudospectra of waveform relaxation operators,* TR-95-14, Dept. of Comp. Sci. and Engr., U. Notre Dame, June, 1995.

[Ors] S. A. Orszag: *Accurate solution of the Orr–Sommerfeld equation,* J. Fluid Mech. **50** (1971), 689–703.

[OSBr] P. O'Sullivan and K. Breuer: *Transient growth in circular pipe flow. Part I: linear disturbances,* Phys. Fluids **6** (1994), 3643–3651.

[Pap] P. F. Papkovich: *Über eine Form der Lösung des Byharmonischen Problems für das Rechteck,* C. r. (Dokl.) Acad. Sci. USSR **27** (1940), 334–338.

[Paz] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations,* Springer-Verlag, 1983.

[Pla] R. Plato: *On resolvent conditions for Abel integral operators,* preprint 459, Dept. of Math., Tech. U. Berlin, 1995.

[Prü] J. Prüss: *On the spectrum of C_0 semigroups,* Trans. Amer. Math. Soc. **284** (1984), 847–857.

[Red] S. C. Reddy: *Pseudospectra of Wiener–Hopf integral operators and constant-coefficient differential operators,* J. Integral Eqs. Appl. **5** (1993), 369–403.

[ReHe] S. C. Reddy and D. S. Henningson: *Energy growth in viscous channel flows,* J. Fluid Mech. **252** (1993), 209–238.

[RSB] S. C. Reddy, P. J. Schmid, and J. S. Baggett: *Stability of streamwise streaks and transition to turbulence in channel flows,* manuscript in preparation.

[RSH] S. C. Reddy, P. J. Schmid, and D. S. Henningson: *Pseudospectra of the Orr–Sommerfeld operator,* SIAM J. Appl. Math. **53** (1993), 15–47.

[ReTr] S. C. Reddy and L. N. Trefethen: *Pseudospectra of the convection–diffusion operator,* SIAM J. Appl. Math. **54** (1994), 1634–1649.

[ReiTr] L. Reichel and L. N. Trefethen: *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices,* Lin. Alg. Applics. **162–164** (1992), 153–185.

[Rey] O. Reynolds: *An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels,* Phil. Trans. Roy. Soc. **174** (1883), 935–982.

[RiMo] R. D. Richtmyer and K. W. Morton: *Difference Methods for Initial-Value Problems*, Wiley-Interscience, New York, 1967.

[Rid] P. Rideau: *Contrôle d'un assemblage de poudres flexibles par des capteurs actionneurs ponctuels: étude du spectre du système,* thesis, Ecole Nationale Supérieure des Mines de Paris, Sophia-Antipolis, France, November, 1985.

[RoSi] S. Roch and B. Silbermann: *C^*-algebra techniques in numerical analysis,* preprint, 1994.

[ScHe] P. J. Schmid and D. S. Henningson: *Optimal energy density growth in Hagen–Poiseuille flow,* J. Fluid Mech. **277**, 197–225.

[Smi] B. T. Smith, et al.: *Matrix Eigensystem Routines—EISPACK Guide,* Springer-Verlag, Berlin, 1976.

[Spe] D. A. Spence: *A class of biharmonic end-strip problems arising in elasticity and Stokes flow,* IMA J. Appl. Math. **30** (1983), 107–139.

[Squ] H. B. Squire: *On the stability for three-dimensional disturbances of viscous fluid flow between parallel walls,* Proc. Roy. Soc. A **142** (1933), 621–628.

[TTS] A. E. Trefethen, L. N. Trefethen, and P. J. Schmid: *Spectra and pseudospectra for pipe Poiseuille flow,* Rep. 94-16, Interdisciplinary Project Center for Supercomputing, ETH-Zurich, August, 1994.

[Tre1] L. N. Trefethen: *Approximation theory and linear algebra,* in: J. C. Mason and M. G. Cox, eds., Algorithms for Approximation II, Chapman and Hall, London, 1990.

[Tre2] L. N. Trefethen: *Pseudospectra of matrices,* in: D. F. Griffiths and G. A. Watson, Numerical Analysis 1991, Longman Sci. Tech. Publ., 1992, pp. 234–266.

[Tre3] L. N. Trefethen: *Spectra and Pseudospectra: The Behavior of Non-Normal Matrices and Operators,* book in preparation.

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

[TrVi] L. N. Trefethen and D. Viswanath: *Pseudospectra and matrix behavior,* manuscript in preparation.

[WeTr] E. Wegert and L. N. Trefethen: *From the Buffon needle problem to the Kreiss matrix theorem,* Amer. Math. Monthly **101** (1994), 132–139.

[Ves] K. Veselić: *On linear vibrational systems with one-dimensional damping,* Applicable Anal. **29** (1988), 1–18.

[Zab] J. Zabczyk: *A note on C_0-semigroups,* Bull. Acad. Polon. Sci. **23** (1975), 895–898.

Address:

Lloyd N. Trefethen, Department of Computer Science and Center for Applied Mathematics, Cornell University, Ithaca, NY 14853, USA (`LNT@cs.cornell.edu`, `http://www.cs.cornell.edu`)