L N TREFETHEN

Pseudospectra of matrices

1 Spectra

The notion of eigenvalues developed historically in connection with hermitian matrices and their infinite-dimensional counterparts, self-adjoint linear operators. Physically, eigenvalues may be the first thing one observes in such a system. The natural frequencies of oscillation of a string or a drum, for example, present themselves to the ear immediately. The great twentieth century example is quantum mechanics, with its remarkable discovery that atoms and molecules occupy energy states that can be interpreted as eigenfunctions of a self-adjoint Schrödinger operator. It was quantum mechanics that brought the ideas of matrices, operators, and eigenvalues to the central position in science and mathematics that they occupy today.

Speaking more generally, the matrices and operators that have been most studied over the years are *normal*. This means that their eigenvectors can be taken to be orthogonal, or equivalently, that they can be unitarily diagonalized. Along with the hermitian matrices, the normal matrices include all those that are skew-hermitian, unitary, or circulant, as well as others besides.

The history of spectral theory for non-normal matrices and operators is sparser. Most books on functional analysis make no assumption of normality in presenting the fundamental definitions. Yet when it comes to applications, it is surprising how little has been done with non-normal systems. The "small matrix" case, exemplified by such developments as the Jordan canonical form and its application to ordinary differential equations, is of course well understood and widely used. But applications of spectral ideas for strongly non-normal operators are rather rare, and in fact, among the largest consumers appears to be numerical analysis. Numerical analysts come at operators by way of "large matrices," that is, matrices whose dimensions are determined by a potentially unbounded discretization parameter rather than by the physics. Eigenvalues are typically applied to analyze stability ($|\lambda| \leq 1$?) and convergence rates ($|\lambda| \ll 1$?).

The point of these remarks is that the notion of eigenvalues, as applied to non-normal matrices and operators, is not so clearly sanctified by history as you may think. In the past few years I have become convinced that there is a reason for this. Eigenvalues and eigenvectors are an imperfect tool for analyzing non-normal matrices and operators, a tool that has often been abused. Physically, it is not always the eigenmodes that dominate what one observes in a highly non-normal system. Mathematically, eigenanalysis is not always an efficient means to the end that really matters: understanding *behavior*. The essence of the eigenvalue idea is a normal one, whose appropriateness in the non-normal

reprinted from:
D.F. Griffiths and G.A. Watson (Eds), Numerical analysis 1991,
Longman Scientific & Technical, Harlow, Essex, UK, 1992

case has been accepted largely by analogy. For non-normal systems eigenvalues may still be useful to a greater or lesser degree, like a nail file when a screwdriver can't be found, but they are rarely exactly right.

Here is an example. If u is a vector and A is a matrix of the same dimension, the differential equation $u_t = Au$ has the solution $u(t) = e^{tA}u(0)$. Consider in particular the 2×2 matrices

$$A' = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \qquad A'' = \begin{pmatrix} -1 & 5 \\ 0 & -2 \end{pmatrix}. \tag{1}$$

Figure 1 shows the 2-norms $\|e^{tA'}\|$ and $\|e^{tA''}\|$ as functions of t, which control the growth or decay that solutions $u(t)$ may exhibit. Quiz for the reader: which curve is which? Which of these matrices allows growth in norm during the transient phase?

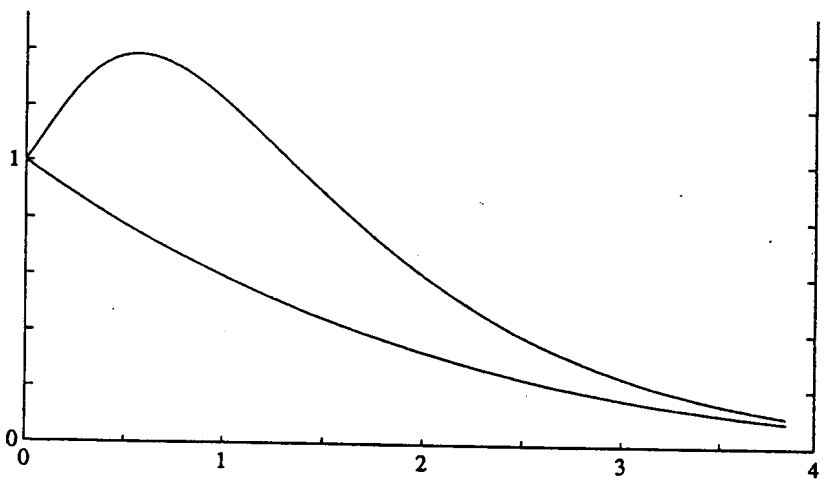


Figure 1. $\|e^{tA}\|$ vs. t for the matrices A' and A'' of (1). Which matrix corresponds the upper curve? For the answer, see the text.

The answer is: A''. The function $\|e^{tA'}\|$, the lower curve in the plot, decays monotonically, but $\|e^{tA''}\|$ grows for a time before eventually decaying. Now there is nothing deep about this result. Nevertheless, it may be surprising to some who are used to making predictions on the basis of eigenvalues. The eigenvalue -1 of A' is defective, according to one pattern of thinking, and therefore some growth must be expected in the transient phase before eventual geometric decay. On the other hand the eigenvalues of A'' are negative and distinct, so no growth should occur. Figure 1 proves that such reasoning is false. Defective eigenvalues are neither necessary nor sufficient for transient growth of $\|e^{tA}\|$. In fact eigenvalues of any kind, defective or not, do not necessarily say much about the behavior of e^{tA} or of A^n or of any other matrix process in the transient phase. Their significance is asymptotic as $t \to \infty$ or $n \to \infty$.

What makes this gap between the eigenvalues of A and the behavior of $\|e^{tA}\|$ possible is the fact that the eigenvectors are not orthogonal. When two eigenvectors are far from orthogonal, a linear combination of them may have large coefficients but small norm, thanks to cancellation. If the coefficients drift out of phase with increasing t so that the cancellation is lost, the norm of the linear combination may increase even though each individual eigenvector component is decaying monotonically. This is the explanation of the upper curve in Figure 1.

For the simplest quantitative approach to such phenomena, let V denote any matrix of eigenvectors of A and

$$\kappa(V) = \|V\| \, \|V^{-1}\|$$

its condition number (we define $\kappa(V) = \infty$ if A is not diagonalizable). Then it is easy to derive the bound

$$e^{t\alpha(A)} \leq \|e^{tA}\| \leq \kappa(V) e^{t\alpha(A)}, \tag{2}$$

where $\alpha(A)$ denotes the spectral abscissa of A, i.e., the maximum of the real parts of its eigenvalues. The analogous bound for matrix powers is

$$\rho(A)^n \leq \|A^n\| \leq \kappa(V)\rho(A)^n, \tag{3}$$

where $\rho(A)$ is the spectral radius of A, and the analogous bound for arbitrary functions $f(z)$ analytic in a neighborhood of the spectrum $\Lambda(A)$ is

$$\|f\|_{\Lambda(A)} \leq \|f(A)\| \leq \kappa(V)\|f\|_{\Lambda(A)}, \tag{4}$$

where $\|f\|_{\Lambda(A)} = \sup_{z\in\Lambda(A)} |f(z)|$. If A is normal, $\kappa(V)$ can be taken to be 1, and all three inequalities (2)–(4) coalesce into equalities. But if A is non-normal, there is always a gap. In general, any quantitative prediction about the behavior of a non-normal matrix, if based on eigenvalues, can be valid only up to a factor such as $\kappa(V)$ that quantifies the non-normality.

For the matrices A' and A'' we have $\kappa(V') = \infty$ and $\kappa(V'') = 10.1$; obviously (2) need not be sharp. For the matrices of Section 3 below, a typical value is $\kappa(V) \approx 10^{10}$. Here and throughout, eigenvector matrices V are taken to be normalized by columns $\|v\| = 1$, and decimal equalities such as "$\kappa = 10.1$" are to be interpreted as accurate to the precision given.

This paper is meant for those who are surprised by Figure 1. It is also meant for those who are not surprised. The fact is, many of us understand the pitfalls of eigenvalue analysis clearly enough when the spotlight is on, but tend to make mistakes as soon as the emphasis moves elsewhere. Nor are numerical analysts alone in such carelessness; one finds mistakes in the use of eigenvalues also, for example, in the literature of hydrodynamic stability in fluid mechanics. Our intuitions have been molded too greatly by the normal case.

2 Pseudospectra

If spectra may be misleading, what should one consider instead? Where in the complex plane does a non-normal matrix or operator "live"? I must state immediately that there is no fully satisfactory answer to this question. For a normal matrix, virtually any way of making the question precise leads to the spectrum. If A is not normal, however, no set in the complex plane has all the properties one would like. The spectrum is too small; the field of values (= numerical range) is too large; the disk about 0 of radius $\|A\|$ is even larger. See the discussion of "spectral sets" in [38].

The purpose of this paper is to present examples of *pseudospectra*, another imperfect answer that has the advantage of being a natural extension of the idea of spectra. The following ideas can be generalized to arbitrary closed operators in a Hilbert space [21,32,43], but for the sake of brevity, let us assume from now on that A is a matrix of dimension $N < \infty$ and $\|\cdot\|$ is the 2-norm.

Consider, as a function of $z \in \mathbf{C}$, the norm of the resolvent $(zI - A)^{-1}$. When z is an eigenvalue of A, $\|(zI - A)^{-1}\|$ can be thought of as infinite, and we shall use this convention. Otherwise, it is finite. How large? If A is normal, the answer is simple:

$$\|(zI - A)^{-1}\| = \frac{1}{\operatorname{dist}(z, \Lambda(A))}. \tag{5}$$

(Here $\Lambda(A)$ again denotes the spectrum of A, and $\operatorname{dist}(z, S)$ is the usual distance from the point z to the set S.) Thus in the normal case, the surface $\|(zI - A)^{-1}\|$ is determined entirely by the eigenvalues like a tent hanging from its poles. In the non-normal case, however, (5) is only a lower bound and the shape of the surface cannot be inferred from the eigenvalues. As our examples will show, $\|(zI - A)^{-1}\|$ may easily attain values as great as 10^{10} or 10^{20} even when z is far from $\Lambda(A)$.

With this in mind it is natural to define the **ϵ-pseudospectrum** of A, for each $\epsilon \geq 0$, by

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} \colon \|(zI - A)^{-1}\| \geq \epsilon^{-1}\}. \tag{6}$$

The ϵ-pseudospectra of A are closed, strictly nested sets with $\Lambda_0(A) = \Lambda(A)$. If A is normal, (5) implies that $\Lambda_\epsilon(A)$ is equal to the union of the closed ϵ-balls about the eigenvalues of A. In general, it may be much larger.

The norm of $(zI - A)^{-1}$ is its largest singular value, i.e., the inverse of the smallest singular value of $zI - A$. Therefore an equivalent definition of the pseudospectrum is:

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} \colon \sigma_N(zI - A) \leq \epsilon\}. \tag{7}$$

Another more interesting equivalent definition can be stated in terms of perturbations of A:

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} \colon z \text{ is an eigenvalue of } A + E \text{ for some } E \text{ with } \|E\| \leq \epsilon\}. \tag{8}$$

In other words, a pseudo-eigenvalue of A is an eigenvalue of a slightly perturbed matrix. (The equivalence of (6) and (8) is easily proved.) This equivalence suggests a new way

of regarding perturbations of eigenvalues: they may give information not only about how the properties of a matrix change when its entries change, but about properties it had already. In other words eigenvalue perturbations can be used as a *visualization device.*

Both matrices A' and A'' of (1) are non-normal, and their pseudospectra are accordingly bigger than the ϵ-neighborhoods about their spectra. Figure 2 plots $\Lambda_\epsilon(A')$ and $\Lambda_\epsilon(A'')$ for the values $\epsilon = 0.05, 0.15, 0.25, \ldots, 0.65$. The two pictures are decidedly different, both near the eigenvalues and far away. For $\epsilon = 0.05$, the innermost contour in each plot, $\Lambda_\epsilon(A')$ extends slightly further to the right in the complex plane than $\Lambda_\epsilon(A'')$. This is a reflection of the defective eigenvalue of A', and can be connected with the behavior of $\|e^{tA'}\|$ as $t \to \infty$. But for all of the larger values of ϵ, $\Lambda_\epsilon(A'')$ extends further to the right than $\Lambda_\epsilon(A')$, and this can be connected with the behavior of $\|e^{tA}\|$ for finite t. (Precise estimates can be derived by means of contour integrals; see [32,33,43].) As for the other limit $\epsilon \to \infty$, it turns out that it can be used to derive the field of values of A, and this is what controls behavior of $\|e^{tA}\|$ for $t \to 0$ [21]. The field of values is equal to the limit (or union) as $\epsilon \to \infty$ of the intersection of all half-planes $H \subseteq \mathbf{C}$ for which $H + \Delta_\epsilon \supseteq \Lambda_\epsilon(A)$, where Δ_ϵ denotes the disk about 0 of radius ϵ [43,25].

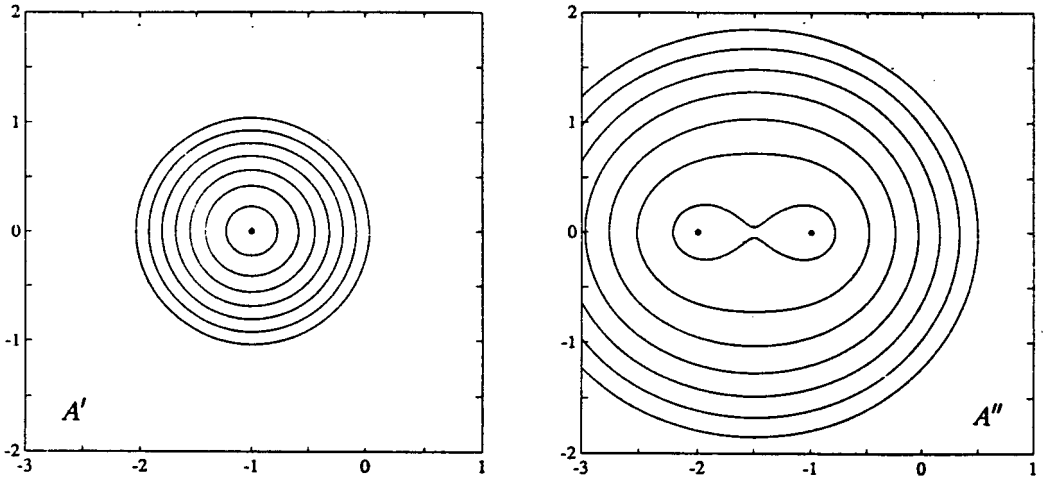


Figure 2. ϵ-pseudospectra of A' and A'' for $\epsilon = 0.05, 0.15, \ldots, 0.65$.

The study of non-normal matrices and operators is a large subject. Some applications of pseudospectral ideas within numerical analysis include:

- convergence of GMRES, CGS, and other nonsymmetric matrix iterations [29],
- design of hybrid iterations that avoid eigenvalue estimates [30],
- convergence of upwind vs. downwind Gauss-Seidel and SOR sweeps,
- backward error analysis of eigenvalue algorithms and polynomial zerofinders [28],
- pseudospectra of Toeplitz matrices [36],

- stiffness of ordinary differential equations [17],
- the Kreiss matrix theorem and generalizations [35],
- Lax-stability of the method of lines [23,24,34,35],
- stability of spectral methods [32,34,41].

Actual or potential applications of pseudospectra in other fields, usually involving operators instead of matrices, include:

- simple differential operators [32],
- convection-diffusion problems [31,32],
- the Fadle-Papcovitch operator (= a biharmonic operator on a half-strip) [39],
- Wiener-Hopf integral operators [1,32],
- the Orr-Sommerfeld operator for Poiseuille flow [33],
- Alfven waves in magnetohydrodynamics [22].

I am in the process of writing a book that will discuss most of these topics [43].

The remainder of this paper has a more limited ambition: to present graphical *examples* of pseudospectra of matrices. The accompanying discussion will give some indication of how pseudospectra may be useful in these application areas, but for full explanations, see the references.

The notion of the pseudospectrum is not new. Under other names, it has been defined previously by Varah [45], Demmel [5,6], Wilkinson [49], Godunov et al. [11], and undoubtedly others. Contour plots as in Figure 2 have appeared in [6] and [11], as well as in the publications [32,33,34] by myself and my colleagues. The recent work of Chatelin and her colleagues is also closely related [4]. Nevertheless this geometric side of linear algebra has received scant attention, and the emphasis has been on how eigenvalues change under perturbations rather than on the exploitation of information that goes beyond eigenvalues. In an era when eigenvalue and singular value computations are routine and graphics tools are universally available, this situation deserves to change.

3 Examples

Our examples consist of thirteen matrices of dimension $N = 32$, all presented in the same format. First the matrix is defined, usually by displaying its 6×6 analogue. Then two plots are presented together on a page. At the head of the page is listed $\kappa(V)$, the condition number of the normalized matrix of eigenvectors.

The upper plot shows 3200 dots representing eigenvalues of 100 perturbed matrices $A + E$, where each E is a random complex matrix with $\|E\| = 10^{-3}$. The entries of E are independent samples from a complex normal distribution with mean 0. To generate E, first a dense matrix $\tilde{E}$ is constructed whose entries are independent samples from the complex normal distribution of mean 0 and standard deviation 1. Then the norm $\|\tilde{E}\|$ is computed and we set $E := 10^{-3}\tilde{E}/\|\tilde{E}\|$.

The lower plot depicts the boundaries of $\Lambda_\epsilon(A)$ for $\epsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$, that is, the level curves $\|(zI-A)^{-1}\| = 10^2, 10^3, \ldots, 10^8$. To determine these curves, the smallest singular value $\sigma_{32}(z_{ij}I - A)$ is computed for each value z_{ij} on a grid of dimensions on the order of 100×100. This takes two or three minutes on a Cray-2. The resulting array of data is then fed to the Matlab contour plotter. The dashed curve on the same plot (sometimes off-scale and hence invisible) is the boundary of the field of values of A computed by the standard algorithm described, for example, in [19]. The thick solid dots are the eigenvalues.

In summary, the upper plot in each example represents a cheap and vivid approximation to the 10^{-3}-pseudospectrum of A. The lower plot depicts the whole surface $\|(zI-A)^{-1}\|$ instead of just a single slice.

Example 1. Jordan block [5,14,48]

There is only one place to begin: with the canonical non-normal matrix, the Jordan block

$$A_1 = \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}.$$

(Remember that despite appearances, the matrix under discussion is actually of dimension 32.) This is the best known matrix whose eigenvalues are sensitive to perturbations. The solitary eigenvalue 0 of A_1 is defective with multiplicity 32, and a perturbation of ϵ in the lower-left corner changes it into 32 distinct eigenvalues of magnitude $\epsilon^{1/32} \approx 1$. Consequently, the upper plot reveals a rather elegant halo of dots at a radius slightly less than $(10^{-3})^{1/32} \approx 0.8$. In the lower plot we see concentric circles. By construction, all of the dots in the upper plot lie on or inside the second-largest of these circles, i.e., the boundary of $\Lambda_{10^{-3}}(A_1)$.

It is striking that most of the dots in the upper plot lie near the outside edge. For this example, being defective, the density of dots at the origin is exactly 0, but even in nondefective cases the same edge-clustering phenomenon is commonly observed. It is *not* a result of taking random matrices E with $\|E\| = 10^{-3}$ instead of $\|E\| \le 10^{-3}$, but simply of choosing examples A that are interesting—that is, whose eigenvalues are sensitive to perturbations.

The fact that the eigenvalues of a Jordan block are highly sensitive to perturbations is widely familiar to numerical analysts. This is not entirely a good thing, for it is the *only* familiar example of this kind of sensitivity, and yet Jordan blocks are by no means representative of non-normal matrices. In general, the pseudospectra of a defective matrix need not look like concentric disks. Conversely, equally interesting pseudospectra can arise for matrices that are nondefective. In fact from the point of view of pseudospectra, Jordan

blocks have no special significance whatever. Though it is true that every matrix is similar to a direct sum of Jordan blocks, the similarity transformation involved is in general non-unitary and does not preserve $\Lambda_\epsilon(A)$. By a unitary similarity transformation one can make any matrix triangular (the Schur form), but not bidiagonal.

Example 2. Limaçon [2,32,42]

The first of the points just made can be illustrated by the "super Jordan block"

$$A_2 = \begin{pmatrix} 0 & 1 & 1 & & & \\ & 0 & 1 & 1 & & \\ & & 0 & 1 & 1 & \\ & & & 0 & 1 & 1 \\ & & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}.$$

Mathematically, A_2 is similar to A_1. However, the similarity transformation in question has condition number on the order of 10^{20}. This explains how it is possible that the figures are so different from those of Example 1. The pseudospectra are now bounded approximately by curves known as *limaçons.*

Notice the phenomenon of self-intersection in the curve of dots in the upper plot, which has no counterpart in the lower plot. Like the edge-clustering phenomenon, this is a matter of probability densities that vary greatly with z. Another kind of clustering phenomenon appears if one takes random matrices E that are real instead of complex (try it!). All of these effects are interesting, but for most applications, I believe they are unimportant.

These first two examples are both Toeplitz matrices, i.e., constant along diagonals. If A is an $N \times N$ Toeplitz matrix with entries $a_{1-N},\ldots,a_0,\ldots,a_{N-1}$, the *symbol* of A is the Laurent polynomial

$$f(z) = \sum_{j=1-N}^{N-1} a_j z^j.$$

For Examples 1 and 2 the symbols are $f_1(z) = z$ and $f_2(z) = z + z^2$. In general, it can be shown that the pseudospectra of a Toeplitz matrix approximate the region in the complex plane bounded by the image of the unit circle under the symbol. More precisely, $\Lambda_\epsilon(A)$ approximates the set of points $z \in \mathbf{C}$ that are enclosed by $f(|z| = 1)$ with nonzero winding number. As $N \to \infty$ and $\epsilon \to 0$ this approximation becomes an equality, while for finite N and ϵ, more careful estimates can be derived that involve the image under f of various curves $|z| = r$ with $r = O(\epsilon^{\pm 1/N})$. See [36]. The idea underlying such results is that ϵ-*pseudo-eigenvectors* for small ϵ can be constructed in the form $(1, z, z^2,\ldots,z^{N-1})^T$ when A is triangular, or as linear combinations of such vectors in the nontriangular case.

Example 3. Grcar matrix [15,30]

For a non-triangular Toeplitz example, consider the matrix

$$A_3 = \begin{pmatrix} 1 & 1 & 1 & 1 & & \\ -1 & 1 & 1 & 1 & 1 & \\ & -1 & 1 & 1 & 1 & 1 \\ & & -1 & 1 & 1 & 1 \\ & & & -1 & 1 & 1 \\ & & & & -1 & 1 \end{pmatrix}$$

investigated by Grcar, with symbol

$$f_3(z) = -z^{-1} + 1 + z + z^2 + z^3.$$

The pseudospectra now look like two pearls connected by a chain. All of the eigenvalues of A_3 are sensitive to perturbations, with condition numbers increasing exponentially with N (this is proved in [36]), but the eigenvalues at the end are especially sensitive.

This matrix is nondefective: its eigenvalues are distinct and $\kappa(V)$ is finite, though large. This illustrates the second point made in Example 1, that pseudospectra can be interesting even for diagonalizable matrices. From now on, all but two of our examples are diagonalizable, and all but one have eigenvalues even more sensitive than those of A_3.

An interesting feature of the Grcar example is that the region where the matrix "lives" surrounds the origin but does not contain it. For example, the field of values contains the origin, but it is clear from the pseudospectral contours that not much is happening there; in fact $\sigma_{32}(A)$ has the rather sizeable value 0.951. These properties become important in applications to iterative solution of nonsymmetric systems $Ax = b$, where the origin acquires a special significance. Many methods for such problems make use of an estimated field of values or the convex hull of a set of estimated eigenvalues [26], but for an example like this one, such methods cannot do very well. For better results one must consider pseudospectra or some other non-convex sets [42], or better yet, avoid the complex plane entirely [29].

Additional examples of pseudospectra of Toeplitz matrices are presented in [2], [36] and [46].

Example 4. Wilkinson matrix [47,48]

Many interesting matrices, though not Toeplitz, are approximately Toeplitz in one sense or another. An example is the "Wilkinson matrix"

$$A_4 = \begin{pmatrix} \frac{1}{N} & 1 & & & & \\ & \frac{2}{N} & 1 & & & \\ & & \frac{3}{N} & 1 & & \\ & & & \frac{4}{N} & 1 & \\ & & & & \ddots & 1 \\ & & & & & 1 \end{pmatrix}$$

(Wilkinson himself took $N = 20$ and scaled up the entries by the factor N). Here the diagonal element is not constant but varies smoothly. To a pure mathematician or a careless applied mathematician this may suggest that A_4, being diagonalizable, will behave more or less as if it is diagonal. The truth is that with $\kappa(V)$ greater than 10^{22}, this matrix behaves more like a Jordan block in most respects than like any diagonal matrix. This is certainly the impression suggested by the plots. The pseudospectra approximate ovals which one may think of as superpositions of disks with centers $1/N$, $2/N$,...,1. The fact that the eigenvalues happen to lie exactly where they do inside these ovals is analogous to the difference that distinguishes the carbon isotopes C_{12} and C_{13}: important perhaps in principle, and detectable by specialized experiments, but so deeply shielded by the surrounding electrons as to have negligible effect on the chemistry.

Example 5. Frank matrix [14,48]

The idea of condition numbers become famous in connection with *backward error analysis* of the effects introduced by machine arithmetic. In particular, suppose we want to compute the eigenvalues of a matrix A. Then the standard conclusion of backward error analysis has the following interpretation in terms of pseudospectra: unless A is triangular or has some other special structure, one cannot in general expect to do better on a computer than obtain a set of ϵ-pseudo-eigenvalues for some ϵ on the order of machine precision.

Our next example, the *Frank matrix,* is a classic example of a matrix whose eigenvalues are ill-conditioned and hence difficult to compute:

$$A_5 = \begin{pmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 5 & 5 & 4 & 3 & 2 & 1 \\ & 4 & 4 & 3 & 2 & 1 \\ & & 3 & 3 & 2 & 1 \\ & & & 2 & 2 & 1 \\ & & & & 1 & 1 \end{pmatrix}.$$

(With $N = 32$ the elements take the values 1,...,32.) By a similarity transformation it can be shown that the eigenvalues of A_5 are all real, but this is another similarity transformation whose condition number grows exponentially with N. As Wilkinson pointed out in [48], the lower eigenvalues have huge condition numbers and cannot be computed accurately except in high precision. In other words, the Frank matrix can be described as an example that illustrates that the pseudospectra of a matrix may be much bigger than the spectrum. The plots reveal this.

The philosophy of backward error analysis might be summarized as follows: *if the answer is highly sensitive to perturbations, you have asked a hard question.* However, the main reason for studying pseudospectra is a deeper principle: *if the answer is highly sensitive to perturbations, you have probably asked the wrong question!* When the eigenvalues of a highly non-normal matrix are troublesome to work with, they are probably irrelevant

anyway to whatever purpose one genuinely cares about. If the Frank matrix arose in an application, the last thing one would properly want to do is diagonalize it.

Example 6. Kahan matrix [13,20]

Another example of historical interest to numerical analysts is the matrix

$$A_6 = \begin{pmatrix} 1 & -c & -c & -c & -c & -c \\ & s & -sc & -sc & -sc & -sc \\ & & s^2 & -s^2c & -s^2c & -s^2c \\ & & & s^3 & -s^3c & -s^3c \\ & & & & s^4 & -s^4c \\ & & & & & s^5 \end{pmatrix}$$

with $s^{N-1} = 0.1$ and $c = \sqrt{1-s^2}$, which was devised by Kahan in the 1960s to illustrate that QR factorization with column pivoting is not a fail-safe method for determining the rank of a matrix. Rank determination is related to the question of *distance to singularity* of a matrix A [18]: how large a perturbation E is required to make A singular? In the 2-norm the answer is $\sigma_N(A)$, or equivalently, the smallest value of ϵ for which $0 \in \Lambda_\epsilon(A)$.

For the present example, the eigenvalues lie in the interval $[0.1, 1]$, which suggests that the matrix may be well-conditioned and hence unambiguously of full rank. But the plots tell a different story. In fact we have $\sigma_{32} = 1.04 \times 10^{-5}$, as can be seen in the plot by noting that the origin lies more or less on the curve $\|(zI-A)^{-1}\| = 10^{-5}$ (check it with a ruler!). This number shrinks at a rate $e^{-C\sqrt{N}}$ as $N \to \infty$.

Geometrically speaking, the point of Kahan's example is that the pseudospectra are *lopsided,* extending substantially to the left of the origin. It is easy to construct matrices with lopsided pseudospectra, but Kahan's ingenuity consisted in finding one that is triangular and whose columns satisfy the requisite pivoting condition.

Example 7. Demmel matrix [6,18]

Another "matrix nearness problem" is that of *distance to instability:* how large a perturbation E is required to move at least one eigenvalue of A into the right half-plane? Our next example,

$$A_7 = -\begin{pmatrix} 1 & B & B^2 & B^3 & B^4 & B^5 \\ & 1 & B & B^2 & B^3 & B^4 \\ & & 1 & B & B^2 & B^3 \\ & & & 1 & B & B^2 \\ & & & & 1 & B \\ & & & & & 1 \end{pmatrix},$$

with $B^{N-1} = 10^8$, was devised by Demmel in order to disprove a conjecture due to Van Loan [44]. The eigenvalues of A_7 are all -1, but under perturbations, they move away

from the real axis and first hit the right half-plane at a point with sizable imaginary part. This is evident in the plot. The Van Loan conjecture amounted to the assertion that the point of impact must be real.

In a sense Demmel's example is opposite to Kahan's: its purpose is to make $\sigma_N(A)$ as large as possible, not as small as possible. Certainly the plots are very different. In Example 7 the origin lies in a hole in the pseudospectra, entirely surrounded by regions of $\Lambda_\epsilon(A)$ for smaller ϵ. The reader may find these pictures a bit hard to interpret, for A has entries as large as 10^8, and some of the action, including many of the 3200 dots, is off-scale. The contour $\epsilon = 10^{-8}$ is the large oval in the left half-plane that encloses the eigenvalue -1, and the contour $\epsilon = 10^{-2}$ is the smallest circle in the right half-plane. The contours for other values of ϵ lie in between.

A plot like the lower one of Example 7 (with $N = 3$ and $B = 100$) appears in Demmel's 1987 paper [6]. This is the earliest plot I have seen in print of computed pseudospectra.

Example 8. Matrix of Lenferink and Spijker [24]

Our next matrix is tridiagonal:

$$A_8 = \begin{pmatrix} -5 & 2 & & & & \\ \frac{1}{2} & -7 & 3 & & & \\ & \frac{1}{3} & -9 & 4 & & \\ & & \frac{1}{4} & -11 & 5 & \\ & & & \frac{1}{5} & -13 & 6 \\ & & & & \frac{1}{6} & -15 \end{pmatrix}.$$

This elegant example was devised by Lenferink and Spijker to illustrate a point in the theory of numerical stability for discrete approximations of partial differential equations. In that field eigenvalues are often used for heuristic stability analysis, but they cannot be relied upon. Highly non-normal examples require a more careful analysis based on other ideas such as circle conditions [24], a generalized numerical range [24,25], or pseudospectra [34,35], and we shall comment further on such applications with Example 11.

It is obvious that A_8 can be symmetrized by the diagonal matrix $D = \mathrm{diag}(1, 2!, 3!, \ldots, N!)^T$, yielding a mathematically similar matrix in which all of the off-diagonal elements are replaced by 1. Therefore the spectrum of A_8 is real. What is interesting is the position-dependence of the similarity transformation. Near the upper-left corner, to speak loosely, A_8 is "close to normal," but the departure from normality grows steadily as one goes down the matrix. Each eigenvalue is more sensitive than the last, and in the complex plane this shows up as a family of pseudospectra shaped like wedges about the negative real axis.

Some of our examples become trivial in the limit $N \to \infty$. This is true in particular of the Toeplitz matrices, for which the pseudospectra for finite N approach the spectrum of the corresponding Toeplitz operator with $N = \infty$, padded by a border of width $O(\epsilon)$ [36]. The Lenferink-Spijker example is different. Computations indicate that the operator A_8 of

dimension $N = \infty$ has a discrete negative real spectrum, like its finite-dimensional sections, with pseudospectra spreading approximately in unbounded wedges of equal angles about that axis. For any finite ϵ, no matter how small, Λ_ϵ is approximately an infinite wedge, not a subset of the real axis. All of this structure is invisible if one looks just at the spectrum. It is invisible also, incidentally, to the device for analyzing families of finite-dimensional non-normal matrices known as the "spectrum of the family" [12,23].

Example 9. Companion matrix [27]

Example 9 is a companion matrix:

$$A_9 = \begin{pmatrix} 0 & 1 & & & & & \\ & 0 & 1 & & & & \\ & & 0 & 1 & & & \\ & & & 0 & 1 & & \\ & & & & 0 & 1 & \\ -c_0 & -c_1 & -c_2 & -c_3 & -c_4 & -c_5 \end{pmatrix},$$

with c_j chosen so that the polynomial

$$p(z) = c_0 + c_1 z + \cdots + c_{N-1} z^{N-1} + z^N$$

has zeros equally spaced in $[-2, 2]$: $\lambda_j = -2+4(j-1)/(N-1)$. The same numbers $\{\lambda_j\}$ are the eigenvalues of A_9; a proof consists of exhibiting the eigenvectors $(1, \lambda_j, \lambda_j^2, \ldots, \lambda_j^{N-1})^T$.

From the plots and the value $\kappa(V) \approx 10^9$ it is evident that companion matrices may be far from normal. The eigenvalues are highly sensitive to perturbations, and there is interesting structure in the pseudospectra. So far as I am aware, little is known about this structure. This is unfortunate, for such a knowledge would have application to the analysis of a familiar but poorly understood algorithm: computing zeros of polynomials by solving matrix eigenvalue problems [27]. In the language of the present paper, the essential question in evaluating this algorithm is: what is the relationship between the pseudospectra of a companion matrix and the sets of *pseudozeros* of the corresponding polynomial, that is, the sets of complex numbers obtainable as zeros of slightly perturbed polynomials? If the pseudozero sets are much smaller than the pseudospectra, the algorithm is unstable, but to the extent that they can be made comparable in size, the algorithm is stable.

It was Wilkinson who made the possible ill-conditioning of polynomial zerofinding problems famous, but he did not view the problem geometrically [47]. For a geometrical study leading to a surprisingly simple algebraic characterization of pseudozero sets in the ∞-norm, see [28]. Pseudozero sets may also be interpreted as *structured pseudospectra* of the companion matrix, defined as in (8) except with perturbations permitted only in the bottom row.

One thing apparent in the plots is that in contrast to the Wilkinson matrix A_4, whose eigenvalues are also equally spaced and whose behavior might consequently have been

expected to be analogous, the pseudospectra of A_9 exhibit a marked scale-dependence: the behavior is quite different inside and outside the unit circle. This is a consequence of having 1 in the super-diagonal. By a diagonal similarity transformation, this number 1 can be replaced by any $\alpha \neq 0$, if c_j is adjusted by a factor α^{N-j}; this transformation is essentially the scale change $z \mapsto \alpha z$ applied to the polynomial $p(z)$. More generally, the technique of matrix "balancing" employed, e.g., in Eispack and Matlab introduces an arbitrary diagonal similarity transformation. Little is known about how close such balancing operations may bring the pseudospectra of A_9 to the pseudozero sets of $p(z)$ [27], and this is a topic ripe for further research.

Example 10. Gauss-Seidel iteration matrix [8,9,36]

With the next example we turn to one of the icons of classical numerical analysis: the Gauss-Seidel iteration matrix $A_{10} = -(L+D)^{-1}U$ obtained from the symmetric tridiagonal matrix

$$B = L+D+U = \begin{pmatrix} -2 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & 1 & -2 & 1 & & \\ & & & 1 & -2 & 1 & \\ & & & & 1 & -2 & \end{pmatrix}.$$

Here L, D and U denote the lower-triangular, diagonal, and upper-triangular parts of B, respectively. Though B is symmetric, A_{10} is not. It is a highly non-normal matrix that is lower-Hessenberg and Toeplitz, except for zeros in the first column:

$$A_{10} = \begin{pmatrix} 0 & \frac{1}{2} & & & & \\ 0 & \frac{1}{4} & \frac{1}{2} & & & \\ 0 & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & & \\ 0 & \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & \\ 0 & \frac{1}{32} & \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{64} & \frac{1}{32} & \frac{1}{16} & \frac{1}{8} & \frac{1}{4} \end{pmatrix}.$$

The eigenvalues of A_{10} were determined analytically by Frankel in 1950 [9] and can also be derived from the more general theory of David Young, to be found in many textbooks.

The plots reveal pseudospectra in the shape of snowshoes. The largest eigenvalues of A_{10} are insensitive to perturbations, but the smaller eigenvalues are so sensitive as to be meaningless for practical purposes. Fortunately, they are never used for practical purposes! It is the largest eigenvalue that determines the convergence rate, conventionally speaking, and consequently, the non-normality of the Gauss-Seidel iteration matrix has been ignored over the years with no ill effects.

If B is replaced by a nonsymmetric matrix, however, such as might arise in the modeling of convection-diffusion equations, the situation changes. Now even the dominant eigenvalue of A become highly sensitive to perturbations, and predictions about

convergence based on eigenvalues may be misleading. In particular one encounters the curious anomaly that the spectral radius, hence the classical convergence rate, is unaffected by whether the Gauss-Seidel sweeps run "upwind" or "downwind," even though in actuality, the direction of sweep makes a great difference to convergence [16]. Such upwind-downwind anomalies become less surprising when one considers the pseudospectra for the two cases, which are utterly different. Realistic convergence bounds can be obtained through the use of the inequality [43]

$$\|A^n\| \le \epsilon^{-1}\rho_\epsilon(A)^{n+1}, \tag{9}$$

valid for any $\epsilon > 0$, where $\rho_\epsilon(A)$ denotes the *ϵ-pseudospectral radius* of A defined by $\rho_\epsilon(A) = \sup_{z\in\Lambda_\epsilon(A)}|z|$ (compare (3)). I hope to discuss these matters in a future paper.

Example 11. Chebyshev spectral [3,34,41]

The next example comes from the field of spectral methods for the numerical solution of partial differential equations, a source of many fascinating matrices. Let $x_0,\dots,x_N$ denote the Gauss-Lobatto-Chebyshev points $x_j = \cos(j\pi/N)$, $0 \le j \le N$. Given data $y_1,\dots,y_N$ at the points $x_1,\dots,x_N$, let f_j be the quantity obtained by (1) interpolating the values (x_j,y_j), together with the boundary condition $(x_0,0)$, by a polynomial $p_N(x)$ of degree N; (2) setting $f_j := p_N'(x_j)$, $1 \le j \le N$. Clearly $\{y_j\} \mapsto \{f_j\}$ defines a linear map from $\mathbf{R}^N$ to $\mathbf{R}^N$, hence a matrix which we denote A_{11}:

$$A_{11} = N^{-2} \times \text{Chebyshev spectral differentiation matrix with b.c. } u(1) = 0.$$

The entries of A_{11} can be expressed by explicit formulas [3]. A Matlab program to generate A_{11} can be derived either from these formulas or via the FFT; see [19].

The plots for A_{11} reveal a remarkable structure. The most striking feature is the set of nearly perfectly straight lines to the left of the origin. This cannot be an accident, and in fact, the explanation is that the differential operator that this matrix is approximating, namely the first derivative operator on $[-1,1]$ with boundary condition $u(1) = 0$, has pseudospectra exactly in the form of half-planes. The resolvent contours $\|(zI-A)^{-1}\| = \epsilon^{-1}$ for that operator are all vertical lines in the complex plane, and ϵ^{-1} grows exponentially as $\mathrm{Re}\, z \to -\infty$.

Finite difference discretizations yield less dramatic agreement of matrix pseudospectra with those of the underlying operator. But it is likely that a wide class of discretizations, spectral or finite-difference, satisfy at least the identity

$$\lim_{N\to\infty} \Lambda_\epsilon(\mathcal{L}_N) \to \Lambda_\epsilon(\mathcal{L}) \qquad (\forall \epsilon > 0) \tag{10}$$

if this limit is defined to mean, for example, pointwise convergence of $\|(zI-\mathcal{L}_N)^{-1}\|$ to $\|(zI-\mathcal{L})^{-1}\|$ as $N \to \infty$, or uniform convergence on compact subsets. See [32] and [36]. Convergence of this kind leaves open the possibility of spurious eigenvalues that have no

connection with the differential operator, so long as they move to ∞ as $N \to \infty$. Four of these "outlying eigenvalues" are visible in the plots for A_{11}. They are common but not universal in spectral discretization matrices [34].

It is essential to come to grips with the non-normality of spectral differentiation matrices if one is to make correct predictions about their numerical stability. Typically a matrix like A_{11} effects the space discretization of a time-dependent partial differential equation, while the time-discretization is handled by an o.d.e. formula with time step Δt; this is the *method of lines*. If the matrix were normal, it would be enough to fit its eigenvalues in the stability region of the o.d.e. formula to ensure Lax-stability of the overall computation. In the non-normal case, however, the restriction on Δt suggested by eigenvalue analysis can be too generous by a factor as large as N, the number of points in the grid [34,41]. It has been shown that an appropriate stability criterion in the non-normal case is that all points of the ϵ-pseudospectrum must lie within a distance $O(\epsilon)$ of the stability region as $\epsilon \to 0$ [34,35]. This result can be interpreted as a transplantation of the Kreiss matrix theorem [37] from the unit disk (for powers of matrices) to an arbitrary stability region (for more general recurrences).

Example 12. Random [7,10]

Our penultimate example is a simple one:

$$A_{12} = \text{random matrix.}$$

To be precise, A_{12} is an $N \times N$ matrix (with $N = 32$ as always) whose entries are random samples from the complex normal distribution of mean 0 and standard deviation $N^{-1/2}$. The standard deviation is normalized in this way so that the limit $N \to \infty$ will be well-behaved. In fact one has $\rho(A_{12}) \approx 1$, $\|A_{12}\| \approx 2$, where $\rho(A_{12})$ denotes the spectral radius, and these approximations can be shown to become equalities in various probabilistic senses as $N \to \infty$.

To clarify, A_{12} is a particular, fixed matrix constructed as described above. Another random matrix would look different in detail but probably not very different in general features.

The upper plot for A_{12} is unlike all those presented heretofore: it reveals just 32 dots instead of 3200! Of course each of these dots is a superposition of 100 copies, which look the same to this plotting resolution. The eigenvalues of A_{12} are insensitive to perturbations in comparison with the other examples we have presented. Still, the lower plot reveals that the eigenvalues of smaller modulus are somewhat sensitive. So far as I know, no theorems have been established to make these observations precise.

There is a lesson to be drawn from this example. If the matrices and operators that we cared about arose at random, it might be appropriate to say that eigenvalue analysis failed only in pathological cases of little importance. However, they do not arise at random. Highly non-normal systems are of special interest in the sciences, and we have a special interest in understanding them.

Example 13. Random upper-triangular [45]

Randomness alone does not ensure closeness to normality. For a final example we define A_{13} exactly like A_{12} except that all entries below the diagonal are set to zero:

$$A_{13} = \text{upper-triangular random matrix.}$$

The pictures change dramatically. Suddenly we have an exponential degree of non-normality again, to judge by the cloud of dots in the upper plot and the approximately circular contours, including all values down to 10^{-8}, in the lower plot. The condition number $\kappa(V)$ has increased by ten orders of magnitude.

I consider this example a particularly fascinating illustration of the subtle relationship between spectra and pseudospectra. Of course A_{13} has nonzero eigenvalues; they are the diagonal entries of the matrix, a set of N random numbers from the complex normal distribution of standard deviation $N^{-1/2}$. On the other hand the pictures suggest that the significance of these eigenvalues is open to question. The eigenvalue to the lower-right is a clear outlier, insensitive to perturbations; it will "behave like an eigenvalue" by any measure. The remaining eigenvalues become less and less sharply defined as one moves in towards the origin. It is impossible to draw a sharp line between meaningful eigenvalues and meaningless ones.

If A_{13} is replaced by a *strictly* upper-triangular random matrix, with zeros on the diagonal as well as the subdiagonals, the pictures change modestly. The curves become more nearly circular, but the scale $O(N^{-1/2})$ remains the same. The spectrum vanishes to the origin, but the pseudospectra remain.

4 Conclusion

I hope the examples in this paper have convinced the reader that there is a geometrical aspect to non-normality, and that it may be beautiful. The idea that it is also informative has been asserted but not argued in depth. For discussions of this point in various applications see [29,30,34,35,33] and the book [43], to appear, that will synthesize these and other developments.

My advice in practice is simple: if you find yourself computing eigenvalues of non-normal matrices, try perturbing the entries by a few percent and see what happens! If the effect on the eigenvalues is negligible, it is probably safe to forget about non-normality. If the effect is considerable, the time has come to be more careful.

Matrices are fascinating, but operators can be even more so [Reddy91]. That is where some of the most remarkable effects of non-normality appear, and in a few years, when I have studied more examples and found better ways to compute pseudospectra, I hope to be able to write a paper like this one with $N = \infty$.

Acknowledgments

The work described here was brought to its present form during sabbatical visits in 1990–91 to the University of New South Wales (Sydney), the Université Pierre et Marie Curie (Paris), and Oxford University. I am grateful for the support at these institutions of Ian Sloan, Philippe Ciarlet and Yvon Maday, and K. W. Morton. Additional support has been provided by an NSF Presidential Young Investigator award. Computations were carried out on the Cray-2 at MIT and with the invaluable aid of Matlab. The work of Godunov, et al. [11] was brought to my attention in May 1991 by Michael Eiermann at the University of Karlsruhe, and a German translation was graciously supplied by his colleague Elvira Winterfeld. Finally, my special thanks go to Satish Reddy for his many contributions to the theory of pseudospectra during several years of collaboration.

References

[1] P. M. Anselone and I. H. Sloan, *Spectral approximations for Wiener-Hopf operators,* J. Int. Eqs. Applics. 2 (1990), 237–261.

[2] R. M. Beam and R. F. Warming, *The asymptotic eigenvalue spectra of banded Toeplitz and quasi-Toeplitz matrices,* to appear.

[3] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics,* Springer-Verlag, New York, 1988.

[4] F. Chatelin, *Resolution Approchée D'Equations sur Ordinateur,* Lecture notes, Lab. de Statistique Théorique et Appliquée, Univ. Pierre et Marie Curie, Paris, 1989.

[5] J. W. Demmel, *A Numerical Analyst's Jordan Canonical Form,* PhD thesis, U. Calif. Berkeley, 1983.

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

[7] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices,* PhD thesis, MIT, 1989.

[8] G. Fairweather, *On the eigenvalues and eigenvectors of a class of Hessenberg matrices,* SIAM Review 13 (1971), 220–221.

[9] S. Frankel, *Convergence rates of iterative treatments of partial differential equations,* Math. Comp. 4 (1950), 65–75.

[10] V. L. Girko, *Theory of Random Determinants,* Kluwer Academic Press, Boston, 1990.

[11] S. K. Godunov, O. P. Kiriljuk, and W. I. Kostin, *Spectral portraits of matrices,* Preprint #3, Inst. of Math., Acad. Sci. USSR, Novosibirsk, 1990 (Russian; see Acknowledgements).

[12] S. K. Godunov and V. S. Ryabenki, *Theory of Difference Schemes,* North-Holland, Amsterdam, 1964.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 2nd ed., Johns Hopkins University Press, Baltimore, 1989.

[14] G. H. Golub and J. H. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form,* SIAM Review 18 (1976), 578–619.

[15] J. Grcar, *Operator coefficient methods for linear equations,* Sandia National Lab. Rep. SAND89-8691, November, 1989.

[16] H. Han, V. P. Il'in, R. B. Kellogg, and W. Yuan, *Analysis of flow directed iterations,* Tech. Note BN-1109, Inst. Phys. Sci. and Tech., U. Maryland, College Park, MD, February 1990.

[17] D. J. Higham and L. N. Trefethen, *Stiffness of ODEs,* in preparation.

[18] N. J. Higham, *Matrix nearness problems and applications,* in M. J. C. Gover and S. Barnett, eds., Applications of Matrix Theory, Oxford U. Press, 1989.

[19] N. J. Higham, *A collection of test matrices in MATLAB,* ACM Trans. Math. Soft., to appear.

[20] W. Kahan, *Numerical linear algebra,* Canad. Math. Bull. 9 (1966), 757–801.

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

[22] W. Kerner, *Large-scale complex eigenvalue problems,* J. Comp. Phys. 85 (1989), 1–85.

[23] V. Lakshimkantham and D. Trigiante, *Theory of Difference Equations: Numerical Methods and Applications,* Academic Press, New York, 1988.

[24] H. W. J. Lenferink and M. N. Spijker, *On the use of stability regions in the numerical analysis of initial value problems,* Rep. TW-89-07, Dept. Math. and Comp. Sci., U. Leiden, 1989.

[25] H. W. J. Lenferink and M. N. Spijker, *A generalization of the numerical range of a matrix,* Lin. Alg. Applics. 140 (1990), 251–266.

[26] T. A. Manteuffel, *Adaptive procedure for estimating parameters for the nonsymmetric Tchebychev iteration,* Numer. Math. 31 (1978), 183–208.

[27] C. Moler, *ROOTS—of polynomials, that is,* Mathworks Newsletter (MathWorks, Inc.) 5 (1991), 8–9.

[28] R. G. Mosier, *Root neighborhoods of a polynomial,* Math. Comput. 47 (1986), 265–273.

[29] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?,* SIAM J. Matrix Anal. Applics., to appear.

[30] N. M. Nachtigal, L. Reichel, and L. N. Trefethen, *A hybrid GMRES algorithm for nonsymmetric linear systems,* SIAM J. Matrix Anal. Applics., to appear.

[31] D. Pathria, S. C. Reddy, and L. N. Trefethen, *Eigenvalues and pseudo-eigenvalues of convection-diffusion operators,* in preparation.

[32] S. C. Reddy, *Pseudospectra of operators and discretization matrices and an application to stability of the method of lines,* PhD thesis, MIT, 1991.

[33] S. C. Reddy, P. Schmid and D. Henningson, *Pseudospectra of the Orr-Sommerfeld operator,* submitted to SIAM J. Appl. Math.

[34] S. C. Reddy and L. N. Trefethen, *Lax-stability of fully discrete spectral methods via stability regions and pseudo-eigenvalues,* Comp. Meth. Appl. Mech. Engr. 80 (1990), 147–164.

[35] S. C. Reddy and L. N. Trefethen, *Stability of the method of lines,* submitted to Numer. Math.

[36] L. Reichel and L. N. Trefethen, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices,* Lin. Alg. Applics., to appear.

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

[38] F. Riesz and B. Sz-Nagy, *Functional Analysis,* Frederick Ungar, New York, 1955.

[39] D. A. Spence, *A class of biharmonic end-strip problems arising in elasticity and Stokes flow,* IMA J. Numer. Anal. 30 (1989), 107–139.

[40] G. W. Stewart and Sun, *Matrix Perturbation Theory,* Academic Press, 1990.

[41] L. N. Trefethen, *Lax-stability vs. eigenvalue stability of spectral methods,* In: K. W. Morton and M. J. Baines (eds.), Numerical Methods for Fluid Dynamics III, 237–253. Oxford: Clarendon Press 1988.

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

[43] L. N. Trefethen, *Non-Normal Matrices and Pseudospectra,* book to appear.

[44] C. Van Loan, *How near is a stable matrix to an unstable matrix?,* in R. Brualdi, et al., eds., Contemporary Mathematics, v. 47, Amer. Math. Soc., 1985.

[45] J. M. Varah, *On the separation of two matrices,* SIAM J. Numer. Anal. 16 (1979), 216–222.

[46] R. F. Warming and R. M. Beam, *An eigenvalue analysis of finite-difference approximations for hyperbolic IBVPs II: the auxiliary Dirichlet problem,* in Proc. Third Intl. Conf. Hyperbolic Problems.

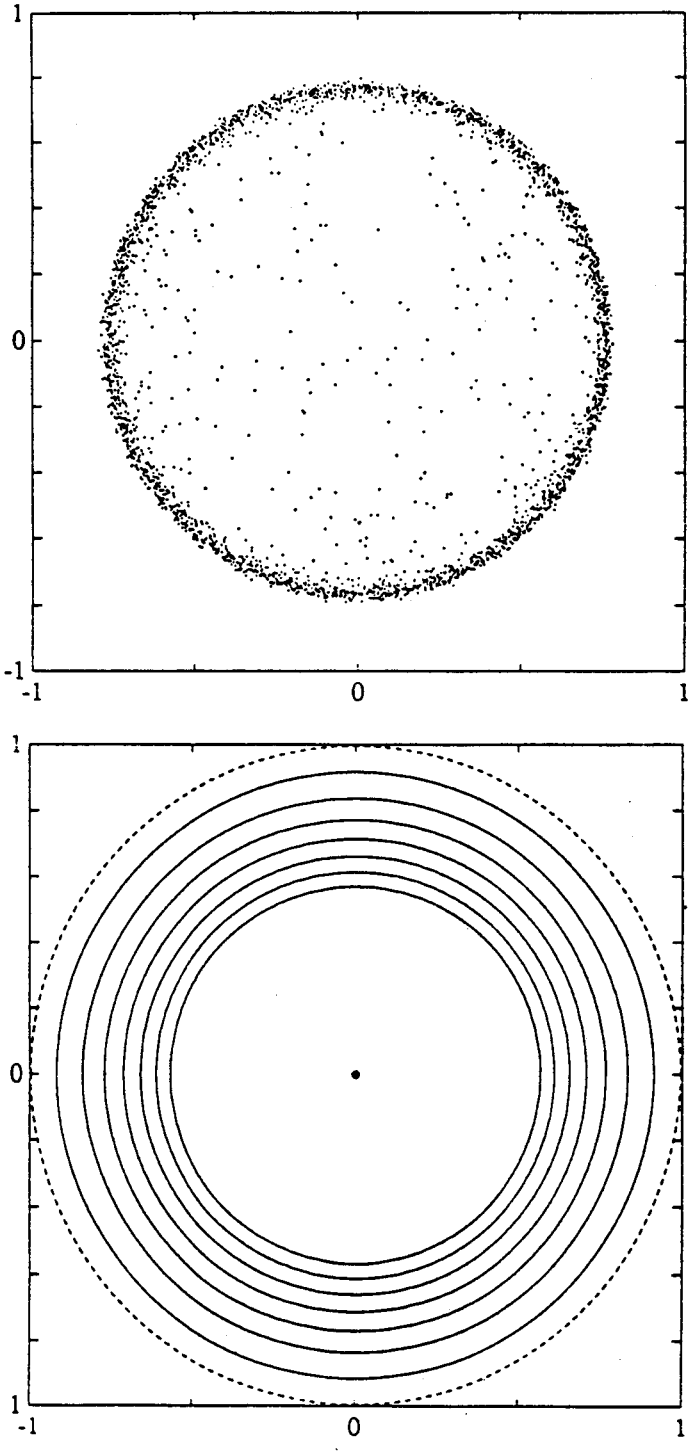
[47] J. H. Wilkinson, *Rounding Errors in Algebraic Processes,* Prentice-Hall, Englewood Cliffs, NJ, 1963.

[48] J. H. Wilkinson, *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford, 1965.

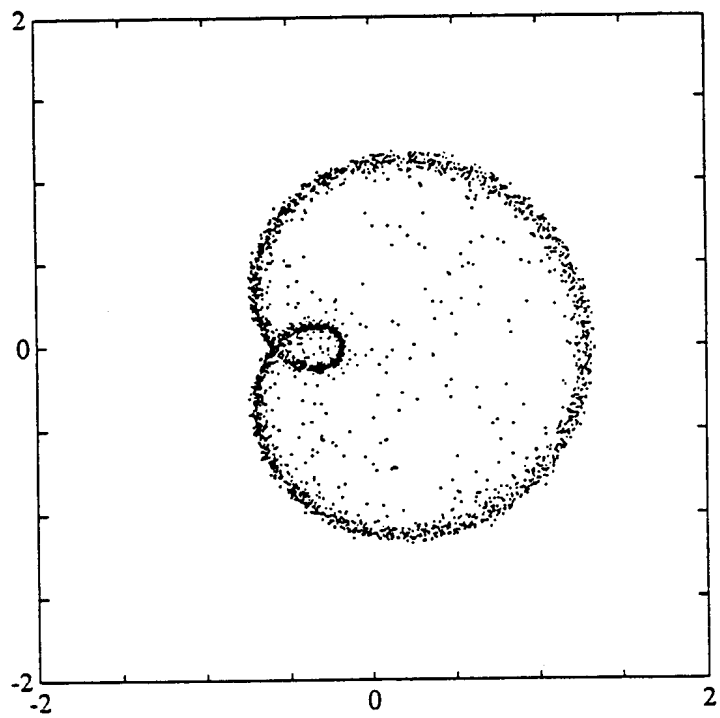
[49] J. H. Wilkinson, *Sensitivity of Eigenvalues II,* Utilitas Math. 30 (1986), 243–286.

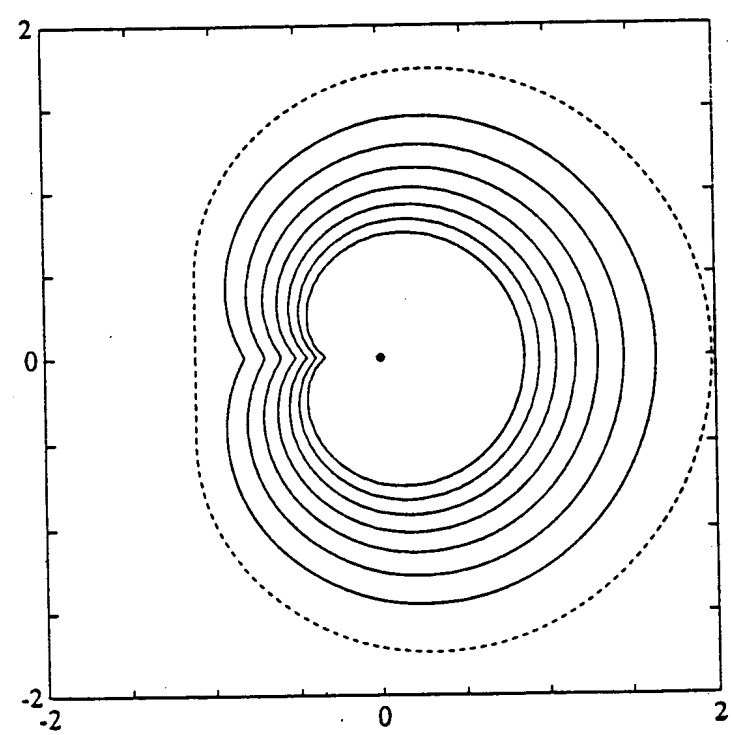
Department of Computer Science
Cornell University
Ithaca, NY 14853 USA
lnt@cs.cornell.edu

Example 1. Jordan block $\kappa(V) = \infty$

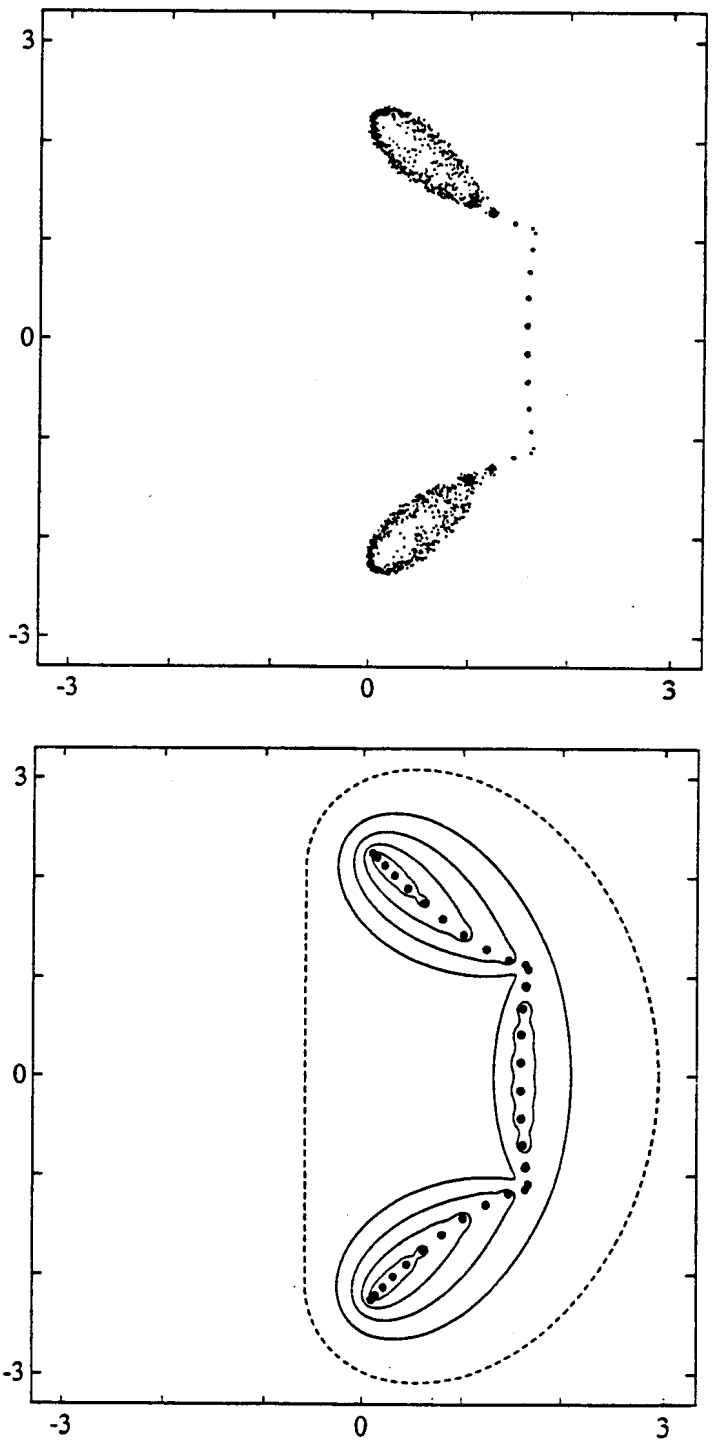


Example 2. Limaçon $\kappa(V)=\infty$

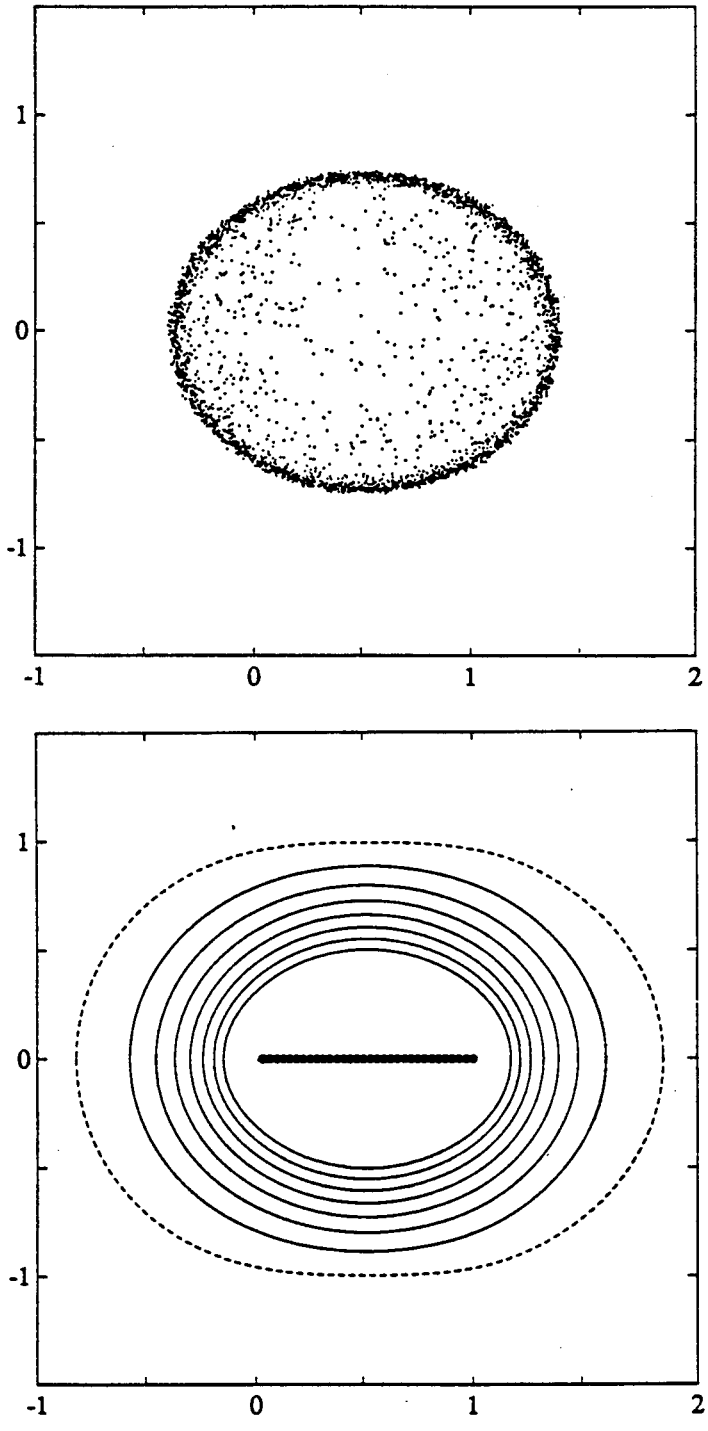




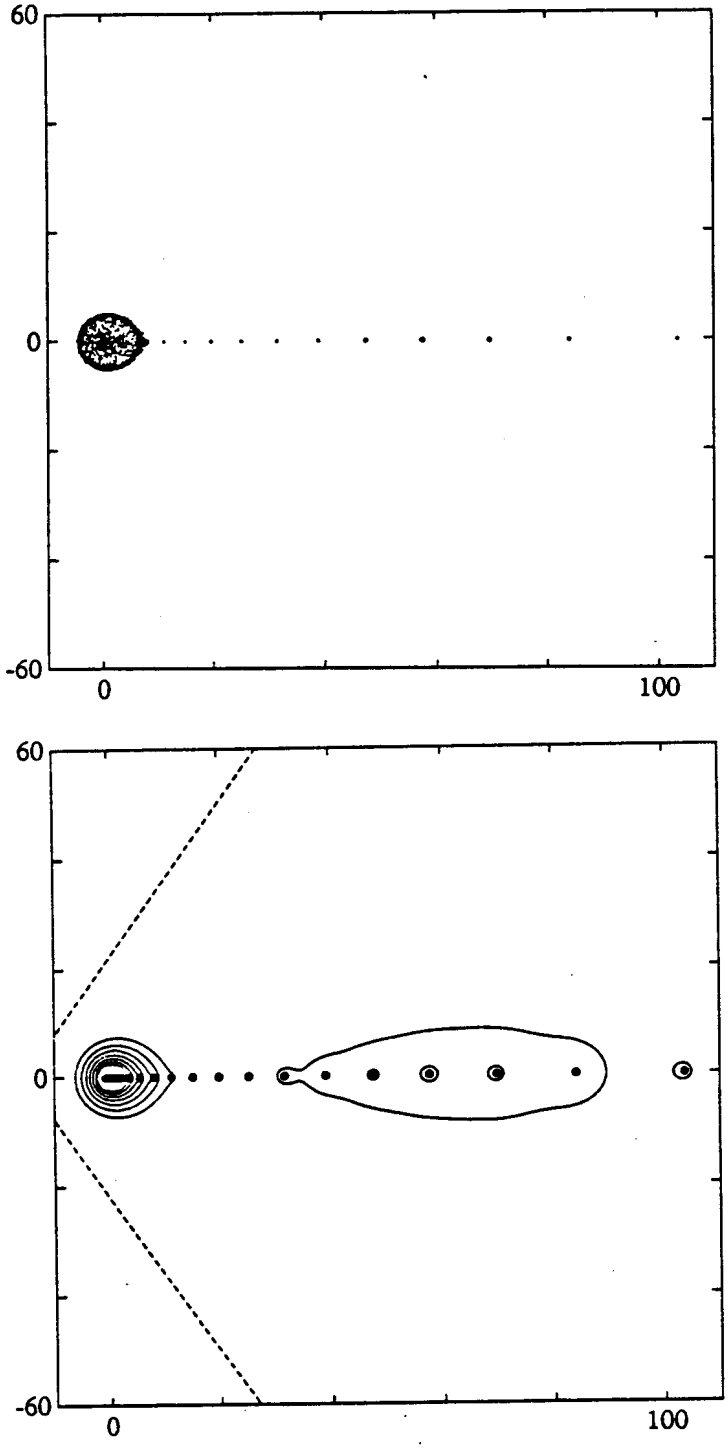
Example 3. Grcar matrix $\kappa(V) = 9.80 \times 10^4$



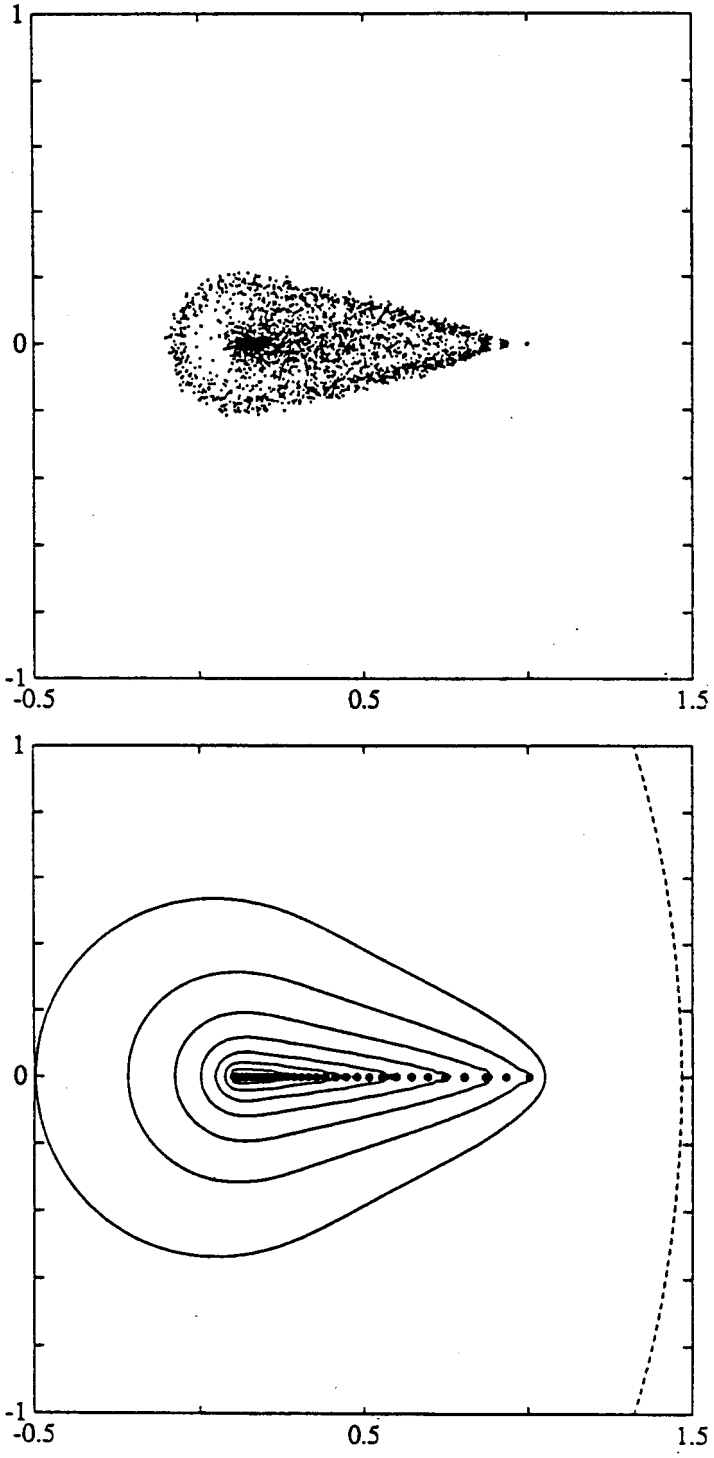
Example 4. Wilkinson matrix $\kappa(V) = 2.62 \times 10^{22}$



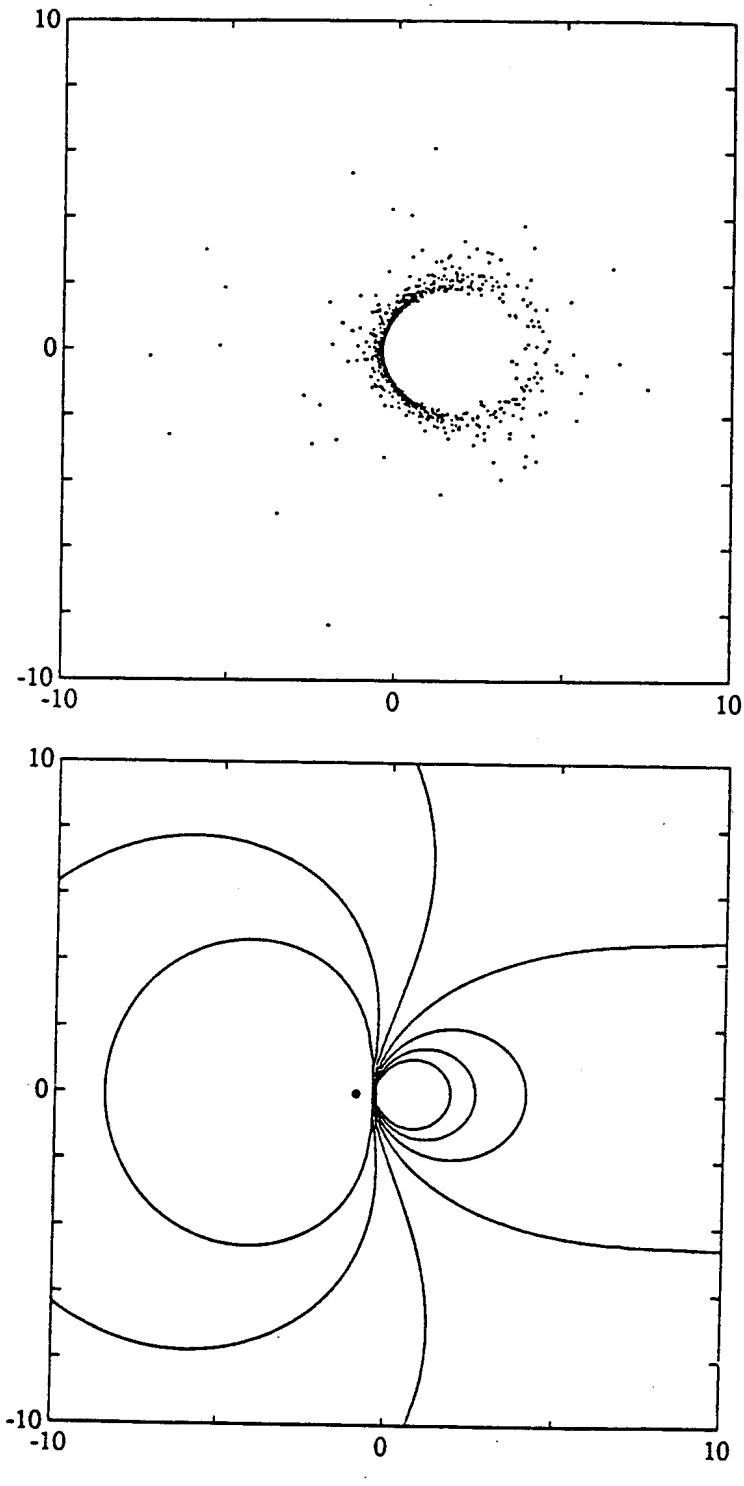
Example 5. Frank matrix $\kappa(V) = 7.81 \times 10^{11}$



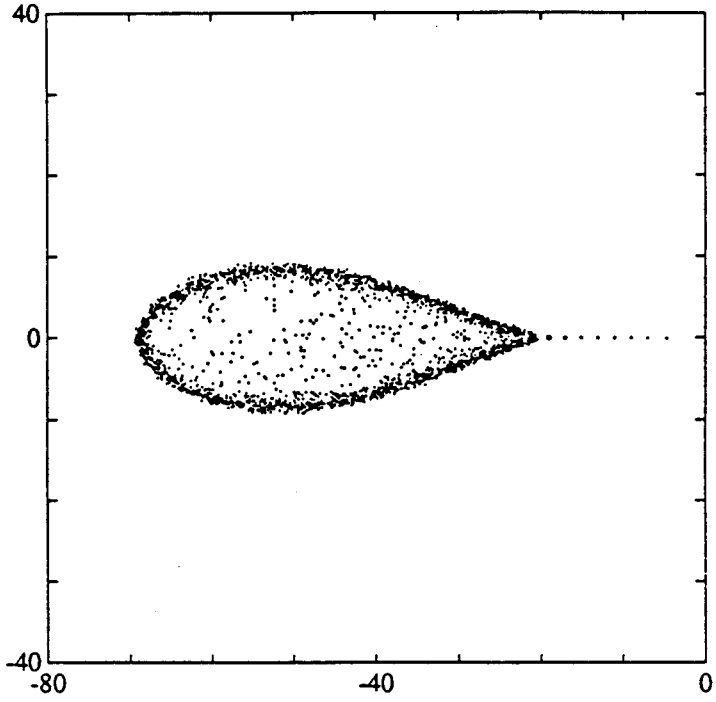
Example 6. Kahan matrix $\kappa(V) = 6.84 \times 10^8$

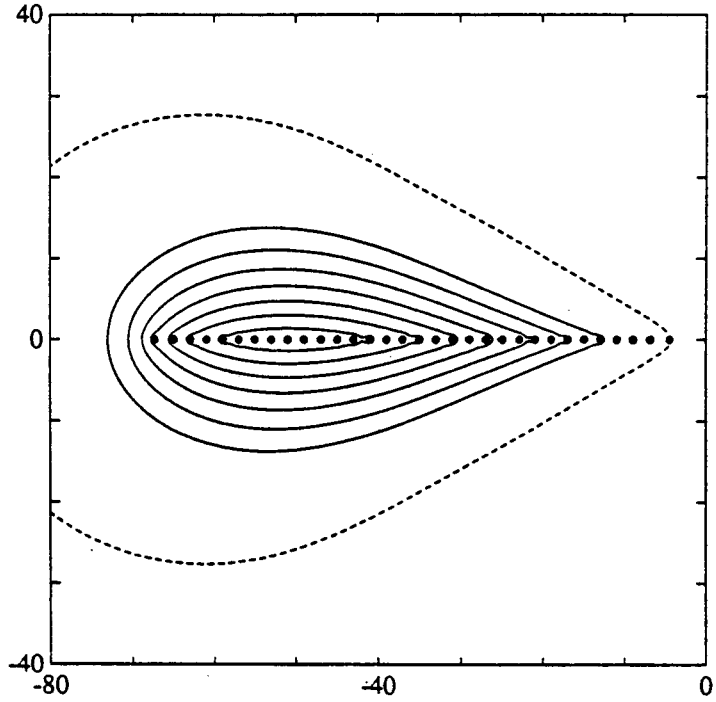


Example 7. Demmel matrix $\kappa(V) = \infty$

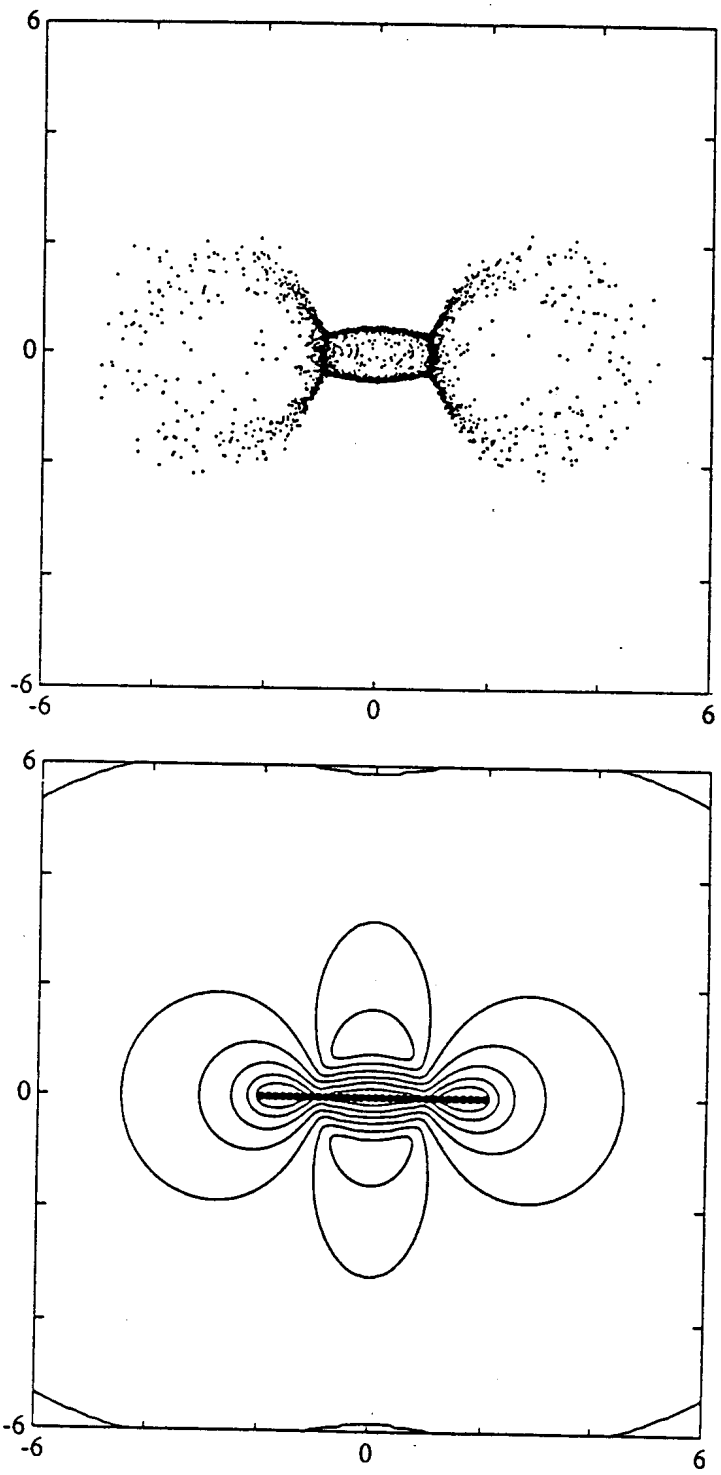


Example 8. Matrix of Lenferink and Spijker $\kappa(V) = 1.75 \times 10^9$

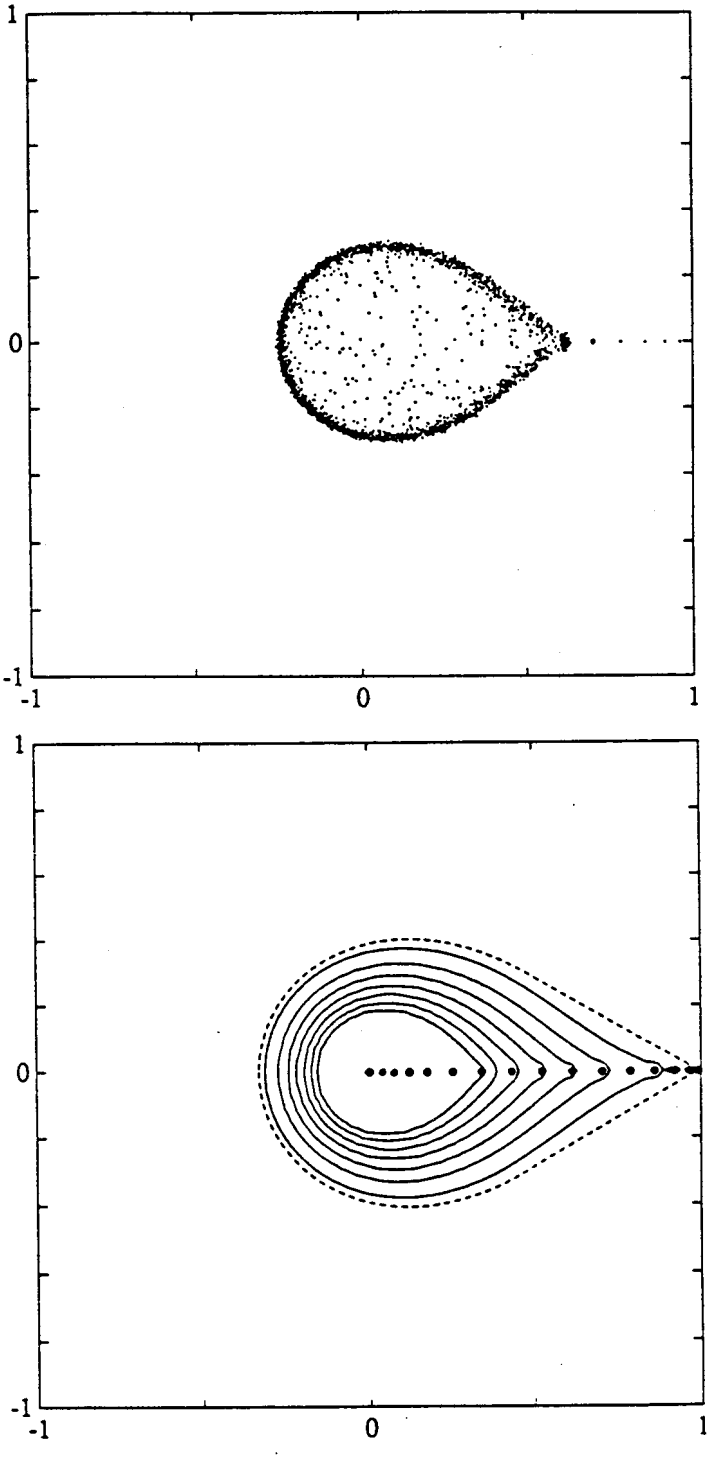




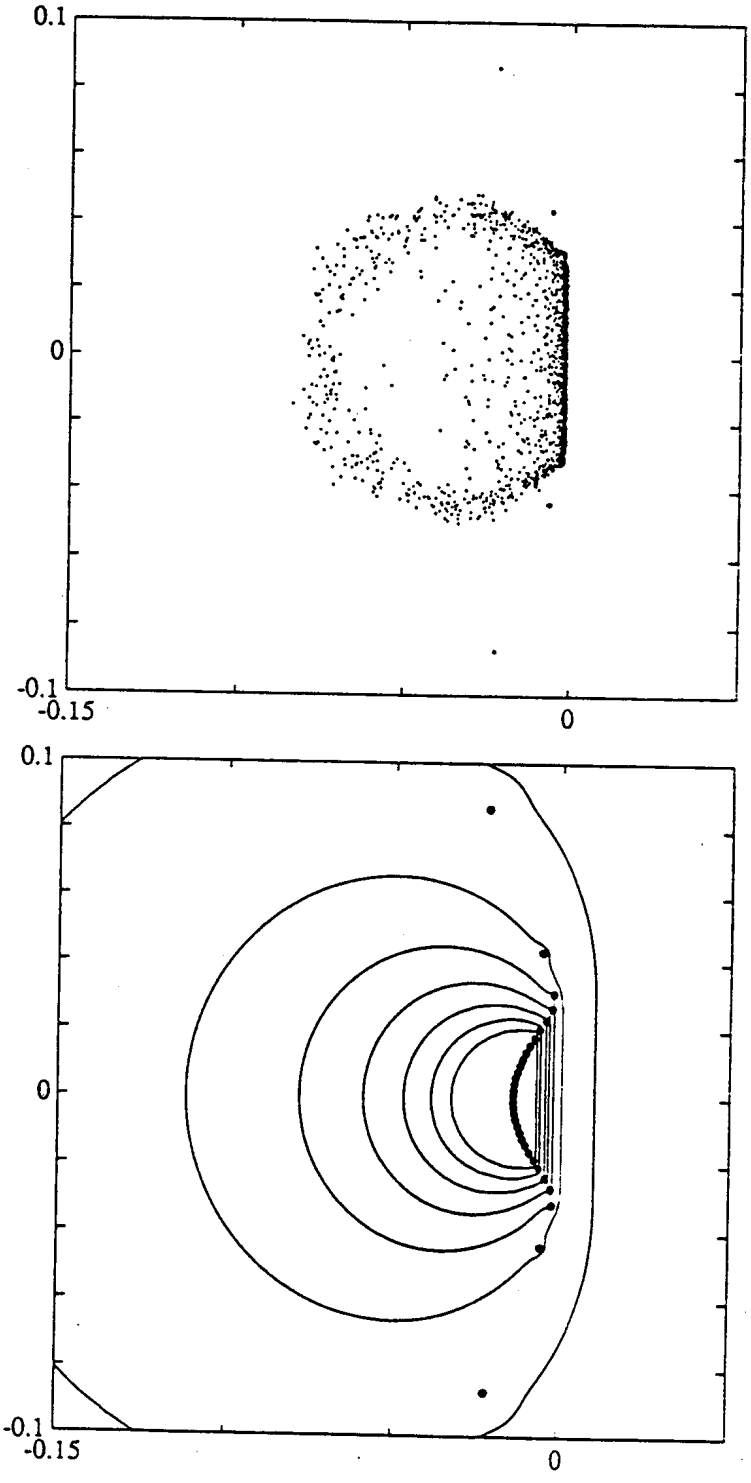
Example 9. Companion matrix $\kappa(V) = 1.55 \times 10^9$



Example 10. Gauss-Seidel $\kappa(V)=\infty$



Example 11. Chebyshev spectral $\kappa(V) = 2.69 \times 10^{14}$



Example 12. Random $\kappa(V) = 25.80$