BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 44, Number 2, April 2007, Pages 277–284
S 0273-0979(06)01128-1
Article electronically published on October 20, 2006

Spectra and pseudospectra: The behavior of nonnormal matrices and operators, by Lloyd N. Trefethen and Mark Embree, Princeton Univ. Press, Princeton, NJ, 2005, xvii + 606 pp., US$65.00, ISBN 0-691-11946-5

1. Eigenvalues

Eigenvalues, latent roots, proper values, characteristic values—four synonyms for a set of numbers that provide much useful information about a matrix or operator. A huge amount of research has been directed at the theory of eigenvalues (localization, perturbation, canonical forms, ...), at applications (ubiquitous), and at numerical computation. I would like to begin with a very selective description of some historical aspects of these topics before moving on to *pseudo*eigenvalues, the subject of the book under review.

Back in the 1930s, Frazer, Duncan, and Collar of the Aerodynamics Department of the National Physical Laboratory (NPL), England, were developing matrix methods for analyzing flutter (unwanted vibrations) in aircraft. This was the beginning of what became known as matrix structural analysis [9] and led to the authors' book *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, published in 1938 [10], which was "the first to employ matrices as an engineering tool" [2]. Olga Taussky worked in Frazer's group at NPL during the Second World War, analyzing 6×6 quadratic eigenvalue problems (QEPs) $(\lambda^2 A_2 + \lambda A_1 + A_0)x = 0$ arising in flutter analysis of supersonic aircraft [25]. Subsequently, Peter Lancaster, working at the English Electric Company in the 1950s, solved QEPs of dimension 2 to 20 [12]. Taussky (at Caltech) and Lancaster (at the University of Calgary) both went on to make fundamental contributions to matrix theory and in particular to matrix eigenvalue problems [24], [17]. So aerodynamics provided the impetus for some significant work on the theory and computation of matrix eigenvalues. In those early days efficient and reliable numerical methods for solving eigenproblems were not available. Today they are, but aerodynamics and other areas of engineering continue to provide challenges concerning eigenvalues. The trend towards extreme designs, such as high-speed trains [15], micro-electromechanical (MEMS) systems, and "superjumbo" jets such as the Airbus 380, makes the analysis and computation of resonant frequencies of these structures difficult [21], [26]. Extreme designs often lead to eigenproblems with poor conditioning, while the physics of the systems leads to algebraic structure that numerical methods should preserve if they are to provide physically meaningful results.

Turning to the question of numerical methods for computing eigenvalues, we can get a picture of the state of the art in the 1950s by looking at the influential lecture notes *Modern Computing Methods* [16]. Its chapter "Latent Roots" describes the power method with deflation for nonsymmetric eigenvalue problems $Ax = \lambda x$ and the Jacobi method for symmetric problems. Nowadays the power method is regarded as the most crude and basic tool in our arsenal of numerical methods,

2000 *Mathematics Subject Classification.* Primary 65F15; Secondary 15A18.

though it proves very effective for computing Google's PageRank [18], [22]. The key breakthrough that led to modern methods was the idea of factorizing a matrix $A = BC$, setting $A \leftarrow CB = B^{-1}AB$, and repeating this process. With a suitable choice of B and C at each step, these similarity transformations force almost all A to converge to upper triangular form, revealing the eigenvalues on the diagonal. Rutishauser's LR algorithm (1958) was the first such method, followed in the early 1960s by the QR algorithm of Francis and Kublanovskaya. Here, A is factorized into the product of an orthogonal matrix (Q) and an upper triangular matrix (R). With appropriate refinements developed by Francis, such as an initial reduction to Hessenberg form and the use of shifts to accelerate convergence, together with more recent refinements that exploit modern machine architectures [1], [4], the QR algorithm is the state of the art for computing the complete eigensystem of a dense matrix. Nominated as one of the "10 algorithms with the greatest influence on the development and practice of science and engineering in the 20th century" [8], the QR algorithm has been the standard method for solving the eigenvalue problem for over 40 years. As Parlett points out [23], the QR algorithm's eminence stems from the fact that it is a "genuinely new contribution to the field of numerical analysis and not just a refinement of ideas given by Newton, Gauss, Hadamard, or Schur."

Returning to the flutter computations, what method would be used nowadays in place of the crude techniques available over 50 years ago to Taussky and Lancaster? The answer is again the QR algorithm, or rather a generalization of it called the QZ algorithm that applies to the more general problem $(\lambda X + Y)x = 0$ associated with the pencil $\lambda X + Y$. The quadratic problem $(\lambda^2 A_2 + \lambda A_1 + A_0)x = 0$, where the A_i are $n \times n$, is reduced to linear form, typically

$$(\lambda X + Y)\begin{bmatrix}\lambda x\\ x\end{bmatrix} := \left(\lambda\begin{bmatrix}A_2 & 0\\ 0 & I\end{bmatrix} + \begin{bmatrix}A_1 & A_0\\ -I & 0\end{bmatrix}\right)\begin{bmatrix}\lambda x\\ x\end{bmatrix} = 0\ , \tag{1}$$

and then the QZ algorithm is applied to the $2n \times 2n$ pencil. The pencil in (1) is called a companion linearization, and it is just one of infinitely many ways of linearizing the quadratic problem. Investigation of alternative linearizations and their ability to preserve structural properties of the QEP is an active area of research [13], [19], [20].

To give a feel for computational cost, on a typical PC all the eigenvalues of a 500×500 nonsymmetric matrix can be computed in about 1 second, while the eigenvalues of a 500×500 QEP require about 30 seconds. Computing eigenvectors as well increases these times by about a factor of 3.

2. Pseudoeigenvalues

The book under review gives food for thought for anyone who computes eigenvalues or uses them to make qualitative predictions. Its message is that if a matrix or operator is highly nonnormal then its eigenvalues can be a poor guide to its behavior, no matter what the underlying eigenvalue theorems may say. High nonnormality of a matrix is characterized by the property that any matrix of eigenvectors V has large condition number: $\|V\|\,\|V^{-1}\| \gg 1$. This message is easy to illustrate in the context of specific eigenvalues of specific matrices. A companion

matrix has the form

$$C = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \dots & \dots & -a_0 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \dots & \dots & 1 & 0 \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

(Note the connection with (1), in which $-X^{-1}Y$ is a block companion matrix.) The matrix C has the characteristic polynomial

$$\det(C - \lambda I) = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0).$$

Therefore with any degree n monic scalar polynomial is associated an $n \times n$ matrix, the companion matrix, whose eigenvalues are the roots of the polynomial. This connection provides one way of computing roots of polynomials: apply an eigensolver to the companion matrix. A companion matrix is nonnormal (unless its first row is the last row of the identity matrix) and so can have interesting pseudospectra. Figure 1 displays the boundaries of several pseudospectral contours of the balanced companion matrix $\widetilde{C}$ for the polynomial

$$p_{20}(z) = (z-1)(z-2)\dots(z-20).$$

(Here, $\widetilde{C} = D^{-1}CD$, with diagonal D chosen to roughly equalize row and column norms.) The *ϵ-pseudospectrum* of $A \in \mathbb{C}^{n\times n}$ is defined, for a given $\epsilon > 0$, to be the set[1]

$$\sigma_\epsilon(A) = \{\, z : z \text{ is an eigenvalue of } A + E \text{ for some } E \in \mathbb{C}^{n\times n} \text{ with } \|E\| < \epsilon \,\},$$

and it can also be represented, in terms of the resolvent $(zI - A)^{-1}$, as

$$\sigma_\epsilon(A) = \{\, z : \|(zI - A)^{-1}\| > \epsilon^{-1} \,\}.$$

Here, $\|\cdot\|$ can be any matrix norm, but the usual choice, and the one used in our examples, is the matrix 2-norm $\|A\|_2 = \max\{\, \|Ax\|_2 : \|x\|_2 = 1 \,\}$, where $\|x\|_2 = (x^*x)^{1/2}$. Any z in σ_ϵ is called an *ϵ-pseudoeigenvalue*. The curves in Figure 1 are the boundaries of $\sigma_\epsilon(\widetilde{C})$ for a range of ϵ, and the dots on the real axis are the eigenvalues (the numbers $1, \dots, 20$). The fact that the contour corresponding to $\epsilon = 10^{-13}$ surrounds the eigenvalues from 9 to 19 shows that these eigenvalues are extremely sensitive to perturbations: a perturbation of norm 10^{-13} can move them a distance $O(1)$. In fact, p_{20} is a notorious polynomial discussed by Wilkinson [27, pp. 41–43], [28]. Although it looks innocuous when expressed in factored form, its roots are extremely sensitive to perturbations in the coefficients of the expansion $p_{20}(z) = z^{20} - 210z^{19} + \cdots + 20!$. The pseudospectra of $\widetilde{C}$ display this sensitivity very clearly. The reader may have noticed a problem with this explanation: the definition of pseudospectra allows arbitrary perturbations to $\widetilde{C}$, even in the zero entries, so the perturbed matrices are not companion matrices of perturbations of p_{20}. As explained in Chapter 55, pseudospectra of a balanced companion matrix usually predict quite well the sensitivity of the roots of the underlying polynomial, but a complete understanding of this phenomenon is lacking. However, the notion of pseudospectrum is readily generalized to matrix polynomials $p(\lambda) = \sum_{i=0}^m \lambda^i A_i$

[1] Those familiar with pseudospectra may be surprised to see "$<$" rather than the more usual "$\le$" in the definitions. The authors decided to use strict inequality in the book because it proves more convenient for infinite dimensional operators.

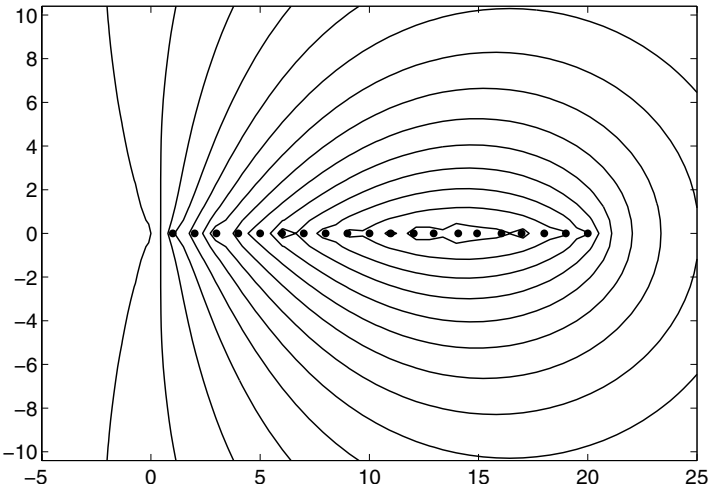

FIGURE 1. Boundaries of pseudospectra $\sigma_\epsilon(\widetilde{C})$ of 20×20 balanced companion matrix for $\epsilon = 10^{-14}$ (innermost curve), $10^{-13}, \ldots, 10^{-1}$ (outermost curve).

[14], and so pseudospectra theory can be applied directly to our scalar polynomial p_{20}.

Another example, from [11], shows in a different context how high nonnormality can lurk behind an easy-looking problem and produce surprising behaviour. Consider the linear system $Ax = b$ of order 100

$$\begin{bmatrix} 1.5 & & & & \\ 1 & 1.5 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 1.5 \end{bmatrix} x = \begin{bmatrix} 2.5 \\ 2.5 \\ \vdots \\ \vdots \\ 2.5 \end{bmatrix}.$$

Being lower bidiagonal, this system is trivial to solve by substitution, from first element to last, and in fact the solution is given explicitly by $x_i = 1 - (-2/3)^i$, $i = 1\colon 20$. We apply the successive overrelaxation (SOR) method in exact arithmetic with parameter $\omega = 1.5$, starting with an approximation to the exact solution correct to 16 significant decimal digits. Being a stationary iterative method, the SOR method produces a sequence of approximations $x_k \approx x$ with errors $e_k = x - x_k$ satisfying $e_k = G^k e_0$, for a certain matrix G. In this example, the spectral radius $\rho(G) = 1/2$, so we might expect the iteration to converge rapidly, with the errors roughly halving on each step. However, as Figure 2 shows, the errors $\|x - x_k\|_\infty / \|x\|_\infty$ (where $\|x\|_\infty = \max_i |x_i|$) initially grow rapidly, until they reach 10^{12}, only then starting to decrease to zero. (We stress that the computations underlying Figure 2 are essentially exact: rounding errors play no role.) Figure 3—computed in under a second on a PC—provides an explanation: even the 10^{-15} pseudospectrum of G lies partly outside the unit disk. Hence, although G has spectral radius $1/2$, tiny perturbations of it have spectral radius exceeding 1. Pseudospectral theory therefore *guarantees* that $\|A^k\|$ is very large for some k.

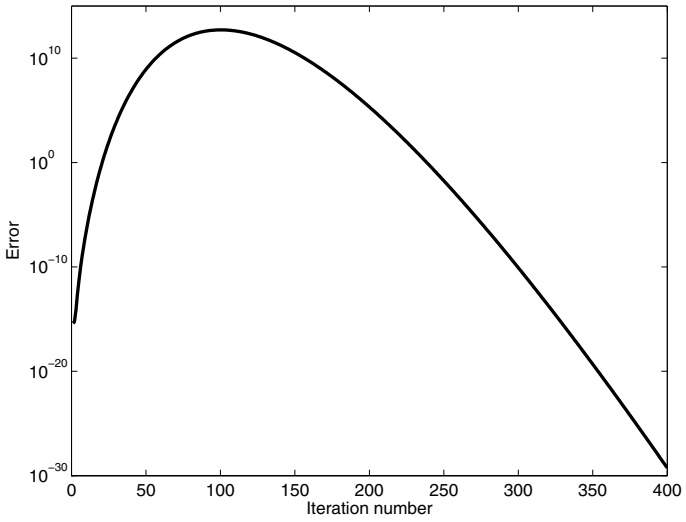

FIGURE 2. Convergence of SOR iteration.

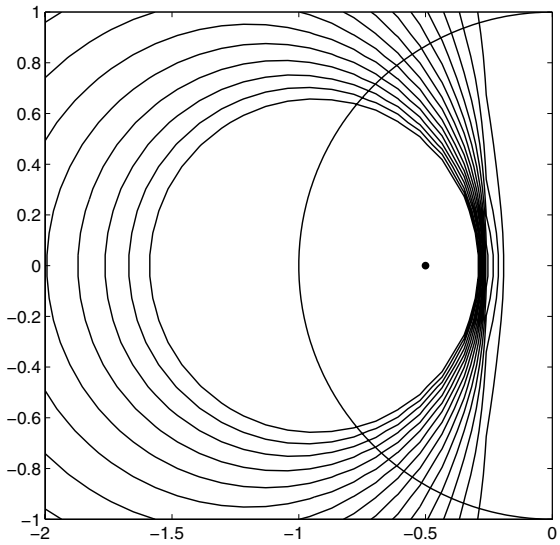

FIGURE 3. Pseudospectra of SOR iteration matrix G for $\epsilon = 10^{-15}$ (innermost curve), ..., 10^{-2} (outermost curve) along with the unit circle. All eigenvalues of the matrix are $-1/2$.

Indeed $\sup_{k\geq 0} \|A^k\| \geq \epsilon^{-1}(\rho_\epsilon(A) - 1)$ for all $\epsilon > 0$, where the pseudospectral radius $\rho_\epsilon(A) = \max\{\, |z| : z \in \sigma_\epsilon(A) \,\}$ (Theorem 16.4 of the book). With $\epsilon = 10^{-15}$ we see that $\sup_{k\geq 0} \|A^k\| \gtrsim 1.5 \times 10^{15}$. The key point is that Figure 3 instantly reveals that G is highly nonnormal and warns that G may not behave in a way that can be described by the eigenvalues alone.

3. Trefethen and Embree's book

The book under review is the first to cover pseudospectra in detail, and it provides a definitive treatment of the subject. Chapter 6 describes the history and reveals that pseudospectra have been invented at least five times, including for the first time in 1967 by Varah and for the fourth time by Trefethen in 1990. Trefethen has done the most to develop and popularize the subject and has written prolifically on it. Embree has worked on pseudospectra since obtaining his Ph.D. in 1999. Until the appearance of this book, no complete survey had been available. As explained in the preface, Trefethen had started to write the book in 1990, and the long gestation period is due to the rapid developments in the subject and the desire to produce a complete and unifying treatment.

The book comprises 60 chapters, each written as a self-contained essay. The necessary definitions and basic results are introduced as and when needed, so although it is a research monograph, the book makes relatively few assumptions about the reader's mathematical knowledge and is very easy to dip into. The writing is superb: eloquent, precise, and enjoyable to read. Ample reference is made to the literature via the 851 references. The many figures are works of art; only those who have tried to produce figures to book quality will appreciate the effort involved. In terms of the typesetting, it would be hard to find a better example of the use of LaTeX and the Computer Modern fonts.

The subject matter is broad. Among the problems and applications in which pseudospectra are studied and applied are random matrices, card shuffling, fluid mechanics, (twisted) Toeplitz matrices, stiff ordinary differential equations, population ecology, and lasers.

This review can give only a taste of pseudospectra, and the danger is that the two examples given above—or indeed any selective examples—fail to give a true impression of the depth and breadth of the subject and of the book. Let me therefore answer some questions that my two examples may suggest.

What do pseudospectra tell us that condition numbers don't? An eigenvalue condition number measures the worst-case change in an eigenvalue under sufficiently small perturbations of the matrix. The ϵ-pseudospectrum shows all possible eigenvalues under perturbations of size ϵ, no matter how large ϵ may be, and so gives a global perspective on the effects of perturbations.

Pseudospectra are based on complex, unstructured perturbations measured in an absolute way. Why not consider structured and/or relative perturbations, especially real perturbations when A is real? This is a common question, and one that is answered in many places in the book, not least in Chapter 50, "Structured Pseudospectra". In a nutshell the answer is that the unstructured absolute perturbation is remarkably successful at explaining matrix behaviour even for structured matrices.

Are pseudospectra difficult and expensive to compute? The computational aspects are one of the successes of pseudospectra, and clever algorithms are available that compute (or approximate) pseudospectra of dense (or very large and sparse) matrices. Moreover, a MATLAB function `eigtool` developed by Wright [29] makes producing plots such as Figures 1 and 3 (which were done with `eigtool`) extremely easy.

Pseudospectra produce qualititative information, but can they produce quantitative predictions? It is true that in the early days of pseudospectra most results

were more qualitative. However, the book amply demonstrates that the answer to the question is "yes", thanks to work in the last ten years or so: Chapters 14–19 give all sorts of quantitative bounds for transient behaviour of differential equations and difference equations in terms of pseudospectra.

Although the theory of pseudospectra originates in numerical analysis, it is starting to permeate other areas of mathematics. In particular, in spectral theory important results involving pseudospectra have been obtained by Davies, Simon, Dencker, Sjöstrand, Zworski and others; see [5], [6], [7], [30], [31], and the references therein. Pseudospectra also have much to say about the behaviour of Toeplitz matrices and operators, as the recent monograph of Böttcher and Grudsky [3] explains.

Finally, I return to aircraft flutter, which is still an issue in aerodynamic design and testing. Chapter 15 describes a problem arising in flutter computations for a Boeing 767, which had earlier been considered by Burke, Lewis, and Overton. Given matrices A (55×55) B (55×2) and C (2×55), the problem is to choose a 2×2 matrix K (representing feedback control) so that $A' = A + BKC$ is stable; that is, its eigenvalues all lie in the left half-plane. Burke, Lewis, and Overton found such a set of parameters using optimization techniques. However, plots of $\|e^{At}\|$ and $\|e^{A't}\|$ show that $e^{A't}$ has huge transient growth just as dangerous as that of e^{At}, even though $e^{At} \to \infty$ as $t \to \infty$ while $e^{A't} \to 0$. Importantly, pseudospectra explain this behaviour very well, in particular by providing lower bounds that accurately track the transients.

In 2006 we have excellent mathematical and computational tools to localize, bound, approximate, and compute eigenvalues of matrices and operators. Trefethen and Embree's book shows convincingly that pseudospectra enable us to look beyond these n numbers to understand better the behaviour of the underlying systems.

References

1. Zhaojun Bai and James W. Demmel, *On a block implementation of Hessenberg multishift QR iteration*, Int. J. High Speed Computing **1** (1989), no. 1, 97–112.
2. R. E. D. Bishop, *Arthur Roderick Collar. 22 February 1908–12 February 1986*, Biographical Memoirs of Fellows of the Royal Society **33** (1987), 164–185.
3. Albrecht Böttcher and Sergei M. Grudsky, *Spectral properties of banded Toeplitz matrices*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005. MR2179973
4. Karen Braman, Ralph Byers, and Roy Mathias, *The multishift QR algorithm. Part I: Maintaining well-focused shifts and level* 3 *performance*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 4, 929–947. MR1920926 (2003f:65061)
5. E. B. Davies, *Pseudo-spectra, the harmonic oscillator and complex resonances*, Proc. R. Soc. Lond. A. **455** (1999), no. 1, 585–599. MR1700903 (2000d:34183)
6. E. B. Davies and Barry Simon, *Eigenvalue estimates for non-normal matrices and the zeros of random orthogonal polynomials on the unit circle*, J. Approx. Theory **141** (2006), no. 2, 189–213.
7. Nils Dencker, Johannes Sjöstrand, and Maciej Zworski, *Pseudospectra of semiclassical (pseudo-)differential operators*, Comm. Pure Appl. Math. **57** (2004), no. 3, 384–415. MR2020109 (2004k:35432)
8. Jack J. Dongarra and Francis Sullivan, *Introduction to the top 10 algorithms*, Computing in Science and Engineering **2** (2000), no. 1, 22–23.
9. C. A. Felippa, *A historical outline of matrix structural analysis: A play in three acts*, Computers and Structures **79** (2001), 1313–1324.
10. R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary matrices and some applications to dynamics and differential equations*, Cambridge University Press, 1938, 1963 printing. MR0019077 (8:365d)

11. S. J. Hammarling and J. H. Wilkinson, *The practical behaviour of linear iterative methods with particular reference to S.O.R.*, Report NAC 69, National Physical Laboratory, Teddington, UK, September 1976.
12. Nicholas J. Higham, *An interview with Peter Lancaster*, SIAM News **38** (2005), no. 6, 5–6.
13. Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur, *Symmetric linearizations for matrix polynomials*, MIMS EPrint 2005.25, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, November 2005; to appear in SIAM J. Matrix Anal. Appl.
14. Nicholas J. Higham and Françoise Tisseur, *More on pseudospectra for polynomial eigenvalue problems and applications in control theory*, Linear Algebra Appl. **351–352** (2002), 435–453. MR1917486 (2003e:93035)
15. Ilse C. F. Ipsen, *Accurate eigenvalues for fast trains*, SIAM News **37** (2004), no. 9, 1–2.
16. National Physical Laboratory, *Modern computing methods*, Notes on Applied Science, no. 16, Her Majesty's Stationery Office, London, 1957. MR0088783 (19:579e)
17. Peter Lancaster, *Lambda-matrices and vibrating systems*, Pergamon Press, Oxford, 1966; reprinted by Dover, New York, 2002. MR0210345 (35:1238), MR1949393 (2003i:34016)
18. Amy N. Langville and Carl D. Meyer, *Google's PageRank and beyond: The science of search engine rankings*, Princeton University Press, Princeton, NJ, USA, 2006.
19. D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann, *Vector spaces of linearizations for matrix polynomials*, Numerical Analysis Report No. 464, Manchester Centre for Computational Mathematics, Manchester, England, April 2005; to appear in SIAM J. Matrix Anal. Appl.
20. ———, *Structured polynomial eigenvalue problems: Good vibrations from good linearizations*, MIMS EPrint 2006.38, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, March 2006; to appear in SIAM J. Matrix Anal. Appl.
21. Volker Mehrmann and Heinrich Voss, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods*, GAMM-Mitteilungen (GAMM-Reports) **27** (2004), 121–152. MR2124762 (2005m:65113)
22. Cleve B. Moler, *Cleve's Corner: The world's largest matrix computation: Google's PageRank is an eigenvector of a matrix of order 2.7 billion*, MATLAB News and Notes (October 2002).
23. Beresford N. Parlett, *The QR algorithm*, Computing in Science and Engineering **2** (2000), no. 1, 38–42.
24. Hans Schneider, *Olga Taussky-Todd's influence on matrix theory and matrix theorists: A discursive personal tribute*, Linear and Multilinear Algebra **5** (1977), 197–224. MR0460043 (57:39)
25. Olga Taussky, *How I became a torchbearer for matrix theory*, Amer. Math. Monthly **95** (1988), no. 9, 801–812. MR0967341 (90d:01077)
26. Françoise Tisseur and Karl Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev. **43** (2001), no. 2, 235–286. MR1861082 (2002i:65042)
27. James H. Wilkinson, *Rounding errors in algebraic processes*, Notes on Applied Science No. 32, Her Majesty's Stationery Office, London, 1963. Also published by Prentice-Hall, Englewood Cliffs, NJ, USA; reprinted by Dover, New York, 1994. MR0161456 (28:4661)
28. James H. Wilkinson, *The perfidious polynomial*, Studies in Numerical Analysis (G. H. Golub, ed.), Studies in Mathematics, vol. 24, Mathematical Association of America, Washington, D.C., 1984, pp. 1–28. MR0925210
29. Thomas George Wright, *Algorithm and software for pseudospectra*, Ph.D. thesis, Numerical Analysis Group, Oxford University Computing Laboratory, Oxford, UK, 2002, pp. vi+150.
30. Maciej Zworski, *A remark on a paper of E. B. Davies*, Proc. Amer. Math. Soc. **129** (2001), no. 10, 2955–2957. MR1840099 (2002e:35015)
31. ———, *Numerical linear algebra and solvability of partial differential equations*, Commun. Math. Phys. **229** (2002), no. 2, 293–307. MR1923176 (2003i:35008)

Nicholas J. Higham
The University of Manchester
E-mail address: higham@ma.man.ac.uk