25

Is nonnormality a serious computational difficulty in practice ?

Françoise Chaitin-Chatelin
Université Paris IX Dauphine and CERFACS
CERFACS, 42 av. G. Coriolis, 31057 Toulouse cedex, France.
Email : `chatelin@cerfacs.fr`

Abstract

The departure from normality of a matrix plays an essential role in numerical matrix computations since it rules the spectral instability. But this first consequence of high nonnormality was for long considered by practitioners as a mathematical oddity, since such matrices were not often encountered in practice. It appears now that more and more matrices, which have a possibly unbounded departure from normality, emerge in the modeling of physical problems at the edge of instability. They challenge many robust numerical codes because of a second and recently exposed consequence of nonnormality : the possible deterioration of the backward stability for algorithms (Chaitin-Chatelin and Frayssé (1996)).
In this paper, we address the following five questions :

(i) What is the connection between spectral instability and nonnormality ?
(ii) What is an appropriate measure of nonnormality ?
(iii) Where do highly nonnormal matrices come from ?
(iv) What is the influence of nonnormality on numerical stability in exact arithmetic ?
(v) What is its influence on the reliability of Numerical Software ?

Keywords

Spectral instability, nonnormality, exact arithmetic, finite precision, reliability of numerical software

R. F. Boisvert (ed.), *Quality of Numerical Software*

1 INTRODUCTION

It has long been known that nonnormal matrices can exhibit spectral instability (Henrici (1962), van der Sluis (1975), Chatelin (1988, 1993)). Despite this theoretical wisdom, nonnormality was considered in practice until recently as a mathematical curiosity only. However in the past few years, problems arising from high technology (Braconnier, Chatelin, and Dunyach (1995)) or theoretical physics (Reddy (1991), Kerner (1989), Trefethen, Trefethen, Reddy, and Driscoll (1993)) have emerged which display a departure from normality which can be exponentially growing with some parameter. The question of their computational treatment on computers with *finite precision arithmetic* requires therefore special attention. In this paper, we intend to review some of the computational consequences of high nonnormality for matrices. But first we investigate the connections between nonnormality and spectral instability.

2 MULTIPLE DEFECTIVE EIGENVALUES AND SPECTRAL INSTABILITY

A *diagonalizable* matrix A has eigenvalues λ which are stable under perturbations of the entries of A. Indeed the maps $A \longmapsto \lambda$ have continuous partial derivatives. If A is subjected to a normwise perturbation ΔA such that $\|\Delta A\| / \|A\| = \varepsilon$, then $|\Delta\lambda|$ is proportional to ε. Consequently, a matrix will display *spectral instability* if it is close, or equal, to a *defective* (i.e. nondiagonalizable) matrix A: there exists at least one eigenvalue λ of A which is multiple and defective (λ has less independent eigenvectors than required by its algebraic multiplicity m). Let $l > 1$ be the *ascent* of λ, that is the size of its largest Jordan block, $1 < l \leq m$. The map $A \longmapsto \lambda$ is not C^1 anymore, but remains continuous. It is in fact Hölder-continuous of order $1/l$ (see Chaitin-Chatelin and Frayssé (1996, pp. 26 and 59)): the perturbation ΔA, $\|\Delta A\| / \|A\| = \varepsilon$ creates a perturbation $|\Delta\lambda|$ which is proportional to $\varepsilon^{1/l}$. As a result, a double defective eigenvalue, for example, is computed, at best, with 7 to 8 digits instead of 15 to 16.

Eigenvalues which are simple or semisimple (i.e. multiple but not defective) are regular points because the maps $A \longmapsto \lambda$ are C^1. But a multiple defective eigenvalue λ is a *singularity* of the map $A \longmapsto \lambda$, which is not C^1 anymore (Chaitin-Chatelin and Frayssé (1996, pp. 24–27)). Singularities of a differentiable map are necessarily rare (Sard (1942)), and they are not generic: they vanish under almost any perturbation of the map. However, one should resist the temptation to underestimate the role of singularities in finite precision computations. Their computational influence can be dramatic, as we shall illustrate.

Such an influence can be already delineated by looking at the question: what is the measure of an eigenvalue ? Clearly, an eigenvalue has a topological dimension $d = 0$ in exact arithmetic. But in finite precision, it has a positive **fractal dimension** $D = 1 - 1/l$, $0 < D < 1$, if it is **defective with ascent** $l > 1$. The definition below of the fractal dimension of an eigenvalue λ should be compared to that of the fractal dimension of fractal sets familiar in physics (Mandelbrot (1983)). Our definition expresses the following approximation property of λ with a finite gauge (accuracy) ε:

- for a regular eigenvalue: $\Delta\lambda \propto \varepsilon$,

- for a singular eigenvalue: $\Delta\lambda \propto \varepsilon^{1/l}$.

For a regular (resp., singular) eigenvalue, $\lambda + \Delta\lambda$ lies in a disk of radius proportional to ε (resp., $\varepsilon^{1/l}$). $\varepsilon^D = \varepsilon^{1-1/l}$ is the ratio $\frac{\varepsilon}{\varepsilon^{1/l}}$ of the two radii. The exponent $D = 1 - 1/l$ characterizes the decrease of this ratio as $\varepsilon \to 0$. We give an example which enlightens the meaning of the fractal dimension of a singularity.

Example 21 Consider the family of matrices $A_n = QJ_nQ^*$, where J_n is a Jordan block of order n defined by

$$J_n = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \tag{1}$$

and Q is a unitary $n \times n$ matrix. J_n is the Jordan form of A_n which has the unique defective eigenvalue 0 with ascent n. Therefore, the sensitivity of this eigenvalue to perturbations of size ε in A_n is $\varepsilon^{\frac{1}{n}}$: it increases exponentially with n. We illustrate the high sensitivity of the zero eigenvalue by computing the eigenvalues of A_n by the classical QR algorithm (Chatelin (1993a)), for $n = 10$, 50, 200 and 500. The computed spectra are plotted in Figure 1. The difference that we see between the exact eigenvalue 0 and the n computed ones is only the result of the spectral instability, because the eigensolver QR is a reliable algorithm. The role of the computer arithmetic is to make this spectral instability visible, it does not create it.

It is clear that, as n increases, most of the computed eigenvalues for A_n tend to cluster first on a circle centered at 0, with radius converging to 1 as n increases, and then to gradually fill the interior for much larger values of n. For the defective eigenvalue 0 of A_n, one gets the fractal dimension $D_n = 1 - 1/n$, showing that $D_n \to 1$ as $n \to \infty$. This may explain why the computed eigenvalues tend to cluster, in their vast majority, on a circle which is a line of topological dimension 1. Figure 2 gives the computed spectra for $n = 1000$ and $n = 2000$. They form a fractal dust (Mandelbrot (1983)).

This example illustrates vividly that the computational influence of defective eigenvalues can be drastic. All the points inside the circles of computed eigenvalues are seen as eigenvalues equally by the computer. The exponential spectral instability exhibited above by A_n may seem overwhelming for large n. However, any situation is two-sided, and a more optimistic view can be derived from a look at the alternate perspective. It is known that the Toeplitz operator J which is the limit in l^2 of J_n as $n \to \infty$, has a spectrum which consists of the closed unit disk (Chatelin (1983)). Therefore, use the computed eigenvalues of A_n to approximate the spectrum of the Toeplitz operator J. In exact arithmetic, the task is hopeless for any finite n: 0 is always at distance 1 from the unit circle. But in finite precision, the spectral information delivered by a reliable software converges towards the border of the limit spectrum as $n \to \infty$. The spectral information computed from A_n is already qualitatively very good for n as low as 500 (which is small when compared to infinity !). $\triangle$

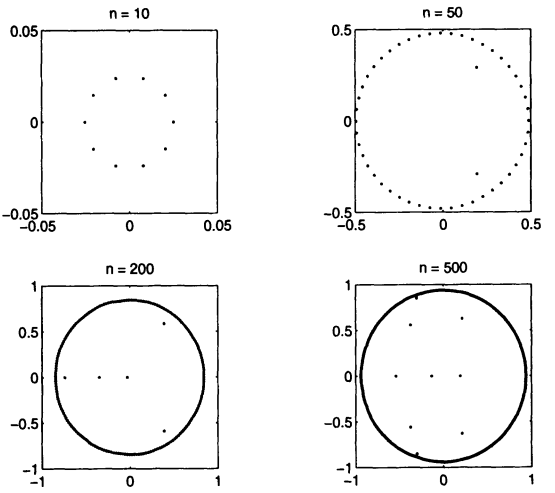

Figure 1 Eigenvalues of A_n computed with QR, $n = 10, 50, 200, 500$.

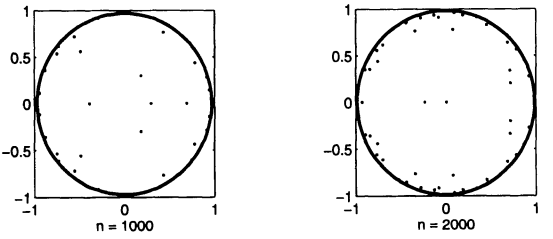

Figure 2 Computed spectra for A_n.

3 CONNECTION BETWEEN NONNORMALITY AND SPECTRAL INSTABILITY

A *normal* matrix A (such that $AA^* = A^*A$) is the most general matrix which has a diagonal Schur form. Therefore, all its eigenvalues and eigenvectors are well-conditioned : the spectral representation of a normal matrix is stable with respect to perturbations in A.

3.1 High nonnormality

Spectral instability, which requires defectiveness, implies nonnormality. However, the converse does not necessarily hold. A diagonalizable matrix with a very well conditioned basis of nonorthogonal eigenvectors is nonnormal but displays spectral stability. We consider an arbitrary defective (nondiagonalizable) matrix A with a Jordan form J and Jordan basis X. The Schmidt factorization $X = QR$ yields $A = XJX^{-1} = Q(RJR^{-1})Q^*$, which $S = RJR^{-1}$ represents the Schur form of A. We set $J = D + K$ and $S = D + N$, where

D is the diagonal of eigenvalues of A. It is easily checked that $\|K\|_2 = 1$. Therefore, the Jordan and Schur decompositions can be paralleled in the following way: for the Jordan (resp., Schur) decomposition, $\|K\|_2$ (resp., $\text{cond}_2(Q)$) is normalized at value 1 but $\text{cond}(X)$ (resp. $\|N\|$) can grow without bound. Moderate spectral instability therefore occurs when $\text{cond}(X)$ and the ascents l are bounded (resp. $\|N\|$ is bounded) when $A = XJX^{-1}$ (resp. $A = QSQ^*$ and $S = D + N$). Such a spectral instability is well understood and there exist reliable software to compute the spectral decomposition (Chatelin (1993a), Anderson, Bai, Bischof et al. (1995)). The matrices that occur in practice often depend, implicitly or explicitly, on one or several parameters which can be the order n, the matrix A itself, or a physical parameter such as the Reynolds or the Péclet number. Whenever the family of matrices under consideration is such that the ascent*of at least one eigenvalue is unbounded and/or the condition number of the Jordan basis is unbounded under the parameter variation, we shall say – in a somewhat loose sense– that the spectral instability of this family of matrices is pathological.

Definition 31 *A nonnormal matrix which displays a pathological spectral instability (in the meaning given above) is called* highly nonnormal.

Clearly, a nonnormal matrix such that $\|N\|$ can grow without limit is highly nonnormal. This remark leads to the difficult question of measures of nonnormality.

3.2 Measures of nonnormality

Nonnormal matrices with well-conditioned eigenvalues are computationally easy to handle.

> So a good measure of nonnormality should reflect how unstable the spectral decomposition is.

This is not easy. The two following indicators have been proposed by Henrici (1962):

(i) $\nu(A) = \|AA^* - A^*A\|$, directly computable from A,
(ii) $\Delta(A) = \|N\|$ where N is the strictly triangular part of the Schur form $S = D + N$, which is of theoretical interest.

They can be related to the conditioning of the eigenbasis when A is diagonalizable (see a review in Chaitin-Chatelin and Frayssé (1996) and Lee (1996)). The quantity $\nu(A)$ is computable but is not homogeneous in A. Since $\nu(A) \leq \|A^*A\| + \|AA^*\| \leq 2\|A\|^2$, only the homogeneous ratio $\nu(A)/\|A^2\|$ is possibly unbounded. We have introduced in Chatelin and Frayssé (1993) the *Henrici number* associated with A, defined by $\text{He}(A) = \dfrac{\nu(A)}{\|A^2\|}$.

*The ascent of an eigenvalue can grow without limit only for matrices of unlimited size n, $n \longrightarrow \infty$.

Whenever $\mathrm{He}(A)$ is large, then A exhibits spectral instability. Indeed, if A is diagonalizable, then (Smith (1967))

$$\mathrm{cond}(X) \geq \left(1 + \frac{1}{2}\mathrm{He}^2(A)\right)^{\frac{1}{4}}. \tag{2}$$

Examples of linear growth for $\mathrm{He}(A)$ are given in Chaitin-Chatelin and Frayssé (1996) with the Schur and Tolosa matrices. See also (Chatelin (1993a), Bennani and Braconnier (1994a)) for the Schur matrix and Bennani, Braconnier, and Dunyach (1994) for the Tolosa matrix.

However, the above indicators $\nu(A)$, $\Delta(A)$, and $\mathrm{He}(A)$ may fail to fully reflect the spectral instability as illustrated in the Example 21, where the key for spectral instability is that the ascent of the eigenvalue is not bounded.

Example 31 The matrix A_n in Example 21 has the eigenvalue 0 with ascent n which is unbounded. However with the 2-norm, $\nu_2(A_n) = \|A_nA_n^* - A_n^*A_n\|_2 = 1$ and $\Delta_2(A_n) = \|A_n\|_2 = 1$ for all n. Moreover $\|A_n^2\|_2 = 1$, so that $\mathrm{He}_2(A_n)$ is also constant at 1 as $n \longrightarrow \infty$. With the Frobenius norm, $\nu_F(A_n) = \sqrt{2}$, $\Delta_F(A_n) = \|A_n\|_F = \sqrt{n-1}$, $\|A_n^2\|_F = \sqrt{n-2}$, and therefore $\mathrm{He}_F(A_n) \to 0$ as $n \to 0$. The indicators do not reveal any pathological instability as $n \to \infty$. △

Example 32 A look at the Harwell/Boeing Collection.
In Table 1 are listed the values of the three quantities $\|A\|$, $\nu(A)$ and $\mathrm{He}(A)$ (computed with the Frobenius norm) for a sample of matrices with ***high spectral instability*** taken from the Harwell/Boeing Collection (Duff, Grimes, and Lewis (1992)). It should be clear from the above sample selected in the Harwell/Boeing Collection that the question of getting a reliable measure of the nonnormality, without computing the spectral decomposition, can be hard. Indeed for all but two matrices, the Henrici number $\mathrm{He}_F(A)$ is small or moderate, whereas all values of $\nu_F(A)$ and $\|A\|_F$ are significantly large. △

3.3 Nonnormality and singular value decomposition

We consider the Singular Value Decomposition of A, that is $A = U\Sigma V^*$ where U (resp., V) is a unitary eigenbasis for AA^* (resp., A^*A). We note that $W = UV^*$ is the unitary polar factor of A. The singular value decomposition can be written

$$Av_j = \sigma_j u_j, \quad j = 1, \ldots, n. \tag{3}$$

Therefore the singular values σ_i give metric information about A. But the angles between the right singular vectors v_j and their images which are the left singular vectors u_j, for $j = 1, \ldots, n$ should also contain information about A, when A is nonnormal. Let us first review the simplifications which occur when A is normal. We suppose, with increasing generality, that

1. A is hermitian positive definite, $\lambda = \sigma$, $V^*U = I$ and $V = U$.
2. A is hermitian, $\lambda = \pm\sigma$, $V^*U = \mathrm{sgn}\Lambda = \mathrm{diag}(\pm 1)$.

Table 1

Matrix Id	Size	$\|A\|_F$	$\nu_F(A)$	$\mathrm{He}_F(A)$
Young1C	841	$4.6\ 10^3$	$1.1\ 10^6$	1.3
MCCA	180	$2.3\ 10^{19}$	$4.5\ 10^{32}$	$0.1\ 10^{-5}$
†BCSSTK01	48	$7.3\ 10^9$	$1.9\ 10^{18}$	0.14
†BCSSTK02	66	$4.9\ 10^4$	$1.4\ 10^8$	0.29
†BCSSTK03	112	$3.4\ 10^{11}$	$3.7\ 10^{21}$	0.06
†BCSSTK06	420	$2.0\ 10^{10}$	$1.7\ 10^{19}$	0.36
†BCSSTK08	1083	$8.1\ 10^8$	$5.8\ 10^{15}$	0.30
IMPCOL E	225	$1.58\ 10^4$	$1.11\ 10^8$	95
WEST0497	497	$1.22\ 10^6$	$9.5\ 10^{11}$	281
WEST0655	655	$7.10\ 10^5$	$2.7\ 10^{11}$	1016
FS 183 1	183	$1.1\ 10^9$	$8.6\ 10^{13}$	$4.5\ 10^7$
†NOS2	957	$1.6\ 10^{12}$	$3.0\ 10^{21}$	0.02
LNS 511	511	$1.0\ 10^{11}$	$2.7\ 10^{21}$	$7.6\ 10^8$

3. A is normal, $\lambda = \sigma e^{i\theta}$, $V^*U = \mathrm{sgn}\Lambda = \mathrm{diag}(e^{i\theta})$.

Therefore, when A is normal, the eigendirections for AA^* and A^*A are identical, the eigenvectors are identical, up to a complex constant of unit modulus. What happens when A is nonnormal ? V^*U is not diagonal anymore: $V^*U = Q$ (where Q is unitary) instead of $V^*U = \mathrm{sgn}\Lambda$. That is $U = VQ$: an arbitrary *rotation* can be applied to V to get U.

Example 33 The SVD of A_n in Example 21 is $U_n\Sigma_n V_n^*$ with

$$U_n = I, \Sigma_n = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \\ & & & & 0 \end{pmatrix}, \text{ and } V_n^* = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}. \tag{4}$$

The singular values are 0 and 1. The left singular vectors $U_n = [e_1, \dots, e_n]$ are rotated of an angle of $\pi/2$ into $V_n = [e_2, \dots, e_n, e_1]$. This is interesting, yet this does not satisfactorily explains the spectacular exponential sensitivity of the eigenvalue 0 for A_n. △

The above discussion confirms that, when A is nonnormal, the Singular Value Decomposition contains information about A which is more reliable and more robust to

†For these symmetric matrices, only the lower triangular part is considered.

perturbations than the information provided by the spectral decomposition (Golub and Van Loan (1989), Trefethen (1992)). This is very valuable, but does not free us from the necessity of understanding the spectral instability of highly nonnormal matrices. Because in many computations, the algorithmic behaviour in finite precision is explained by the eigenvalues (and not singular values) which are "seen by the computer".

3.4 Perturbed spectra and spectral portraits

One reliable way to investigate computationally spectral instability is to test the sensitivity of the eigenvalues to perturbations ΔA on A, i.e. to compute the spectrum of $A + \Delta A$ by means of a *backward stable* algorithm, such as QR or Arnoldi-Tchebycheff as already illustrated in Example 21 above. Such techniques are available in the software environment PRECISE, presented in Chaitin-Chatelin and Frayssé (1996), where ΔA are real random perturbations of A. The perturbed spectra are related to the theoretically defined set

$$\sigma_\varepsilon(A) = \{z \in \mathbb{C};\ z \text{ is an eigenvalue of } A + \Delta A,\ \|\Delta A\| \leq \varepsilon \|A\|\}, \tag{5}$$

which is known as the *ε-pseudospectrum* of A (Trefethen (1992)). Clearly, the larger this set in $\mathbb{C}$, the more unstable the eigenvalues. It is easy to prove that

$$\sigma_\varepsilon(A) = \left\{ z \in \mathbb{C};\ \left\|(A - zI)^{-1}\right\| \geq \frac{1}{\varepsilon \|A\|} \right\}. \tag{6}$$

The contour line $\left\{ z \in \mathbb{C};\ \|(A - zI)^{-1}\| = \dfrac{1}{\varepsilon \|A\|} \right\}$ is then the border of the ε-pseudospectrum. And for nonnormal matrices, $\|(A - zI)^{-1}\|$ can be large even at points z which are far from exact eigenvalues. Indeed for a simple (or semi-simple) eigenvalue λ (ascent equal to 1), the condition number $c(\lambda)$ can be characterized (Chatelin (1993b)) by the ratio

$$\lim_{z \longrightarrow \lambda} \frac{\|(A - zI)^{-1}\|}{1/|z - \lambda|} = \lim_{z \longrightarrow \lambda} \left\|(A - zI)^{-1}\right\| |z - \lambda| = c(\lambda). \tag{7}$$

$c(\lambda)$ can be interpreted as the limit of the ratio of the absolute condition number of $A - zI$ to $1/|z - \lambda|$ which would represent the same condition number if A were normal. Hence, if $c(\lambda)$ is large, $\|(A - zI)^{-1}\|$ can be much larger than $1/|z - \lambda|$ in the neighborhood of λ, but also significantly further away. Such a fact is conveniently displayed by the *spectral portrait* of the matrix A (Godunov (1992)), that is the map $z \longmapsto \log_{10}(\|(A - zI)^{-1}\|\, \|A\|)$ (Toumazou (1996)).

In conclusion of this section, we can say that the question of defining a reliable measure of nonnormality is still largely open. It may be necessary in practice to turn to more expensive tools such as perturbed spectra, pseudo-spectra or spectral portraits, to analyze the spectral instability of highly nonnormal matrices. Such tools are developed for large matrices in the framework of the European project PINEAPL (1996).

It is time now to look at the fundamental question : where do the highly nonnormal matrices, that are encountered in Science and Technology, come from ?

4 NONNORMALITY IN PHYSICS AND TECHNOLOGY

4.1 Strong coupling

The matrices that can occur in Physics are often discretizations of a partial differential operator and their nonnormality can be inherited from the one of the operator. In the physical and mechanical examples cited below, the essential ingredient is that the *spectrum* of a family of operators depending on a parameter exhibits a *severe discontinuity* as the parameter varies. This can occur when the model describes a **strong coupling** between two physical phenomena.

Three such examples are described in Chaitin-Chatelin and Frayssé (1996): convection-diffusion in one dimension, controlled fusion of plasma and flutter. The modeling of the flutter phenomenon also gives rise to highly nonnormal matrices described in Bennani, Braconnier, and Dunyach (1994), Braconnier, Chatelin, and Dunyach (1995). One of such matrices, known as the Tolosa matrix, is in the Harwell/Boeing Collection (Duff, Grimes, and Lewis (1992)).

4.2 Convergence of numerical methods in exact arithmetic

The coupling between physical phenomena is often transferred in the numerical approximation of evolution equations as a necessity of a coupling between parameters such as time and space mesh sizes. Without a proper restriction on the mesh sizes, the numerical method can be unstable. So we expect that nonnormality in Physics can have an impact on the numerical stability of the approximation methods, in the absence of round-off.

This is indeed the case. It has been well-known for a long time that for fully discrete evolution equations, the condition which requires that the spectrum of spatial discretization lies in the stability region for the time-stepping formula is only a *necessary* condition for stability whenever the operator is nonnormal. Recently, Reddy and Trefethen (1990, 1992) have proposed a *necessary and sufficient* condition by means of the ε-pseudospectra.

5 INFLUENCE OF NONNORMALITY ON NUMERICAL SOFTWARE

The challenging influence of nonnormality on the numerical stability of approximations methods in *exact arithmetic* has been known for a long time (Godunov and Ryabenki (1964)) and recently emphasized by Trefethen, Trefethen, Reddy, and Driscoll (1993) and several other authors (see Higham and Owren (1995) for example).

In comparison, the influence of nonnormality on the reliability of numerical methods in *finite precision* seems to be much less widely appreciated. This should not be the case because the underlying phenomenon is essentially of the same nature in both cases. Its roots lie in the spectral instability of highly nonnormal matrices; and therefore in the sensitivity of the spectrum to perturbations in the data, that is in the matrix elements. There is theoretical and experimental evidence that both finite and iterative methods can be affected by high nonnormality when run in *finite precision* arithmetic : their reliabil-

ity can decrease severely. In addition, iterative methods can also be affected when the condition for convergence is not robust enough to perturbations generated by finite precision, resulting in a divergence in finite precision arithmetic, even though the condition for convergence is well-satisfied in exact arithmetic.

5.1 Reliability

The reliability or backward stability of finite and iterative methods can be analyzed by means of the backward error at the computed solution. The backward stability follows usually from a bound of the type

$$\text{Backward error } \leq C(n, A)\,\mathbf{u} \tag{8}$$

where $\mathbf{u}$ is the unit round-off and C is a constant which depends on the details of the arithmetic, the size n and possibly on the matrix A.
1) When the constant C does not depend on A, then the reliability of the method is unaffected by nonnormality. This is the case for :

(i) finite methods for $Ax = b$ such as LU with complete pivoting, QR and Hessenberg-Arnoldi implemented by Householder/Givens transformations and iterative modified Gram-Schmidt,
(ii) iterative methods for $Ax = \lambda x$ such as the QR algorithm.

2) When the constant C depends on A, the reliability can be affected by high nonnormality. This is the case for finite methods such as the QR factorization and Hessenberg-Arnoldi implemented by modified Gram-Schmidt, and for many iterative methods.

Numerical illustrations can be found in Bennani and Braconnier (1994a) and in Chaitin-Chatelin and Frayssé (1996). The importance of the quality of the orthogonality of the basis in Krylov methods is theoretically investigated in Braconnier (1995). In Bennani and Braconnier (1994b), the importance of the use of the backward error as a stopping criterion for eigensolvers is discussed and exemplified. This is all the more important in case of high nonnormality that $\|A\|$ is very large.

5.2 Convergence condition for iterative methods

The convergence condition often takes the form of a requirement on the spectrum of A to lie in some stability region. If A has a high spectral sensitivity, the condition may not remain satisfied under the perturbations on A generated by finite precision computations. An example of diverging successive iterations on a matrix A such that $\rho(A) < 0.41$ in exact arithmetic is given in Chatelin (1993b). The quality of convergence of successive iteration under high nonnormality is studied in Chaitin-Chatelin and Gratton (1996). Methods proved convergent in exact arithmetic may fail to converge in finite precision. There can exist an intermediate state between convergence and divergence, where computer results remain bounded (Chaitin-Chatelin and Frayssé (1996), Chaitin-Chatelin and Gratton (1996)). Simple computer simulations commonly found in Linear Algebra are described

in Braconnier, Chaitin-Chatelin, and Gratton (1996a, 1996b) which exhibit a chaotic behavior because of the computer arithmetic.

5.3 Conclusion

We have shown that high nonnormality can occur in Physics and Technology whenever there is a strong coupling of phenomena giving rise to physical instabilities. Therefore, high nonnormality should be taken seriously in Numerical Software since it may affect the reliability and the convergence of numerical methods when they are run on a computer with a finite precision arithmetic.

Acknowledgments

Parts of this paper were presented at various meetings, including "Sparse days in Saint-Girons" in July 1994, ILAS 94 in Rotterdam, August 1994 and the Workshop "Eigenvalues and Beyond", October 17-20 1995, of the International Linear Algebra Year at CERFACS. The author is grateful to Valérie, Serge, Thierry and Vincent, her past and present co-workers in the Qualitative Computing Group at CERFACS for their supportive enthusiasm and help.

REFERENCES

Anderson, E., Bai, Z., Bischof, C., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Ostrouchov, S., and Sorensen, D. (1995) *LAPACK User's Guide*. SIAM, Philadelphia, second ed.

Bennani, M. and Braconnier, T. (1994a) Comparative behaviour of eigensolvers on highly nonnormal matrices. Tech. Rep. TR/PA/94/23, CERFACS.

Bennani, M. and Braconnier, T. (1994b) Stopping criteria for eigensolvers. Tech. Rep. TR/PA/94/22, CERFACS.

Bennani, M., Braconnier, T. and Dunyach, J.-C. (1994) Solving large-scale nonnormal eigenproblems in the aeronautical industry using parallel BLAS, in *High-Performance Computing and Networking*, W. Gentzsch and U. Harms, eds., vol. 796, Springer-Verlag, 72–77.

Björck, Å. (1994) Numerics of Gram-Schmidt Orthogonalization. *Linear Algebra Appl.*, **197–198**, 297–316.

Braconnier, T. and Chaitin-Chatelin, F. (1996) Nonnormality: from Physics to Linear Algebra. In preparation.

Braconnier,, T., Chaitin-Chatelin F. and Gratton, S. (1996a) Entre convergence et divergence: exemples de chaos arithmétique en Algèbre Linéaire lié à la nonnormalité. Work in progress.

Braconnier, T., Chaitin-Chatelin, F. and Gratton, S. (1996b) Chaotic behavior for linear- and eigen- solvers applied on nonnormal matrices. In preparation.

Braconnier, T., Chatelin, F. and Dunyach, J.-C. (1995) Highly nonnormal eigenvalue problems in the aeronautical industry. *Japan J. Ind. Appl. Math.*, **12**, 123–136.

Braconnier, T. (1995) Influence of orthogonality on the backward error and the stopping criterion for Krylov methods. Numerical Analysis Report 281, Dept. of Mathematics, University of Manchester.

Chaitin-Chatelin, F. and Frayssé, V. (1996) *Lectures on Finite Precision Computations.* SIAM, Philadelphia.

Chaitin-Chatelin, F. and Gratton, S. (1996) Convergence of successive iteration methods in finite precision under high nonnormality. *BIT*, to appear.

Chaitin-Chatelin, F. (1994a) Is nonnormality a serious difficulty ? Sparse Days at St-Girons, France.

Chaitin-Chatelin, F. (1994b) Is nonnormality a serious difficulty ? Tech. Rep. TR/PA/94/18, CERFACS. Presented at ILAS 94, Rotterdam.

Chatelin, F. and Frayssé, V. (1993) Qualitative Computing : elements of a theory for finite precision computation. Lecture Notes for the Comett European Course, June 8-10, Thomson-CSF, LCR Corbeville, Orsay.

Chatelin, F. (1983) *Spectral approximation of linear operators.* Academic Press, New York.

Chatelin, F. (1988) *Valeurs propres de matrices.* Masson, Paris.

Chatelin, F. (1993a) *Eigenvalues of matrices.* Wiley, Chichester. Enlarged Translation of the French Edition with Masson.

Chatelin, F. (1993b) The influence of nonnormality on matrix computations, in *Linear Algebra, Markov Chains and Queueing Models*, R. J. Plemmons and C. D. Meyer, eds., Springer, New York, 13–19.

Duff, I. S. , Grimes, R. G. and Lewis, J. G. (1992) User's Guide for the Harwell-Boeing Sparse Matrix Collection. Tech. Rep. TR-PA-92-86, CERFACS.

Godunov, S. K. and Ryabenki, V. S. (1964) *Theory of Difference Schemes: an Introduction.* North-Holland, Amsterdam. Translation by E. Godfredsen.

Godunov, S. K. (1992) Spectral portraits of matrices and criteria of spectrum dichotomy, in *Computer arithmetic and enclosure methods*, J. Herzberger and L. Atanassova, eds., North-Holland and IMACS.

Golub, G. and Van Loan, C. (1989) *Matrix Computations.* Johns Hopkins University Press. Second edition.

Harrabi, A. (1995) Etude de la continuité des pseudo-spectres d'opérateurs bornés. Tech. Rep., CEREMADE, Université Paris IX Dauphine. in preparation.

Henrici, P. (1962) Bounds for iterates, inverses, spectral variation and field of values of nonnormal matrices. *Numer. Math.*, **4**, 24–40.

Higham, D. J. and Owren, B. (1995) Non-normality effects in a discretised nonlinear reaction-convection-diffusion equation. *J. Comp. Physics*, to appear.

Ilahi, A. (1996) Perturbations avec PRECISE et le théorème de Lidskii, Tech. Rep., CEREMADE, Université Paris IX Dauphine, in preparation.

Kerner, W. (1986) Computing the complex eigenvalue spectrum for resistive magneto-hydrodynamics, in *Large scale eigenvalue problems*, J. Cullum and R. A. Willoughby, eds., North-Holland, Amsterdam, 240–264.

Kerner, W. (1989) Large scale complex eigenvalues problems. *J. Comp. Phys.*, **85**, 1–85.

Lee, S. L. (1996) Best available bounds for departure from normality. *SIAM J. Matrix Anal. Appl.* To appear.

Mandelbrot, B. (1983) *The fractal geometry of Nature.* W. H. Freeman, New York.

Moro, J., Burke, J. V. and Overton, M. L. (1996) On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of Matrices with Arbitrary Jordan Structure. *SIAM J. Matrix Anal. Appl.* To appear.

PINEAPL (1996) Parallel NumErical Applications and Portable Libraries. http://www.nag.co.uk/projects/PINEAPL.html, Fourth Framework Project:20018.

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

Reddy, S. C. and Trefethen, L. N. (1992) Stability of the method of lines. *Numer. Math.*, **62**, 235–267.

Reddy, S. C. (1991) *Pseudospectra of operators and discretization matrices and an application to stability of the method of lines.* Ph.D. dissertation, Dept. of Mathematics, Massachusetts Institute of Technology.

Sard, A. (1942) The measure of the critical values of differentiable maps. *Bull. AMS*, **48**, 883–896.

Smith, R. A. (1967) The condition numbers of the matrix eigenvalue problem. *Numer. Math.*, **10**, 232–240.

Toumazou, V. (1996) *Portraits spectraux de matrices: un outil d'analyse de la stabilité.* Ph.D. dissertation, Université H. Poincaré, Nancy. In preparation.

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

Trefethen, L. N. (1992) Pseudospectra of matrices, in *Numerical Analysis 1991*, D. F. Griffiths and G. A. Watson, eds., Longman, Harlow.

van der Sluis, A. (1975) Perturbations of eigenvalues of nonnormal matrices. *Comm. ACM*, **18**, 30–36.

DISCUSSION

Speaker : F. Chaitin-Chatelin

J. Pryce : Please clarify whether your concept of "degree of non-normality" applies to a single matrix, or only to a parameterized family of matrices. In what context does it involve unbounded condition of the Jordan basis, or unbounded ascent, or both?

F. Chaitin-Chatelin : "High" is best understood by thinking of a family of matrices, which are typically a sequence of discretized differential operators for decreasing meshes. If the sequence of matrices is highly nonnormal then for small enough values of the parameter each individual matrix is also highly nonnormal.

B. Smith : Numerical software really assumes the input is inexact and therefore evaluation should be relative to this assumption. Professor Chatelin points out that actual problems (from Physics and elsewhere) often make this assumption but users, in fact, provide the software with exact data. The software provides an answer consistent with inexact data but the user assumes it is exact data. How do we inform/educate users of this?

F. Chaitin-Chatelin : By repeatedly explaining to the user that if the software is backward stable to machine precision, then the computed solution is the best one available from the computer. The forward error is a result of the instability of the problem to small perturbations in the data.

N. Higham : The Users' Guides for LAPACK and LINPACK both cover this point, by stressing that the best that can be expected in general is that the codes produce the right answer to a slightly perturbed problem, and they explain that merely storing the data on the computer introduces uncertainty. With reference to iterative refinement, the LINPACK Users' Guide says that "most problems involve inexact input data and obtaining a highly accurate solution to an imprecise problem may not be justified".

M. Wright : For certain kinds of problems, the data may have two different forms: exact numbers, such as those arising in simple geometric calculations; and numbers subject to error, such as those from observed data. There seems to be a shortage of error analysis that would apply to these situations, and available software may not provide a detailed understanding of the significance of the computed results.

M. Wright : In solving a problem for which the exact eigenvalues were known to be real, the computed eigenvalues turned out to be complex. This happened during the first step of a computation, and led to unexpected and confusing results. The algorithms were correct, and the software was reliable, but this would not be apparent to the typical user. Are there general ways for software developers to deal with this kind of issue?

N. Higham : My only comment is that if a stable method such as the QR algorithm was used, then the fact that the computed eigenvalues were complex shows that the original matrix was close to one with complex eigenvalues. If real eigenvalues were required for physical reasons then perhaps the original problem was not well formulated.

H. Stetter : The situation is more general. Users of numerical software should be able to specify qualitative properties of the problem and/or solution, and the software should be able to utilize that.

J. Pryce : It is desirable for users to be able to enter into software constraint information known on physical grounds. But it is also dangerous. For example, one may constrain concentrations ≥ 0 or force energy to be conserved, but (a) a misformulation of the problem may mean this is not true for the problem as posed, or (b) one may hide a long-term instability by a constraint which expresses what the user thinks is true about the solution.

W. Schiesser : Extending math software to allow engineers and scientists to define constraints or other properties of the solution in many cases will not help to produce better results since the constraints are already in the equations. That is, the set of equations, if solved correctly, will automatically conserve energy and momentum.

W.V. Snyder : In problems with exact parameters or implicit constraints, such as conservation of energy, where the answer produced by software are "different", is the solution to make codes more robust, perhaps causing unnecessary expense in solving simple problems, or to educate the user community better?

F. Chaitin-Chatelin : Certainly, educating the user community better is important, so that users are aware of the intrinsic difficulties of unstable computations due to physics.