BIT 24 (1984), 584–591

ON THE RESOLVENT CONDITION IN THE KREISS MATRIX THEOREM *

RANDALL J. LEVEQUE and LLOYD N. TREFETHEN

Department of Mathematics, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90024, U.S.A.

Department of Mathematics Massachusetts Institute of Technology Cambridge, MA 02139, U.S.A.

Dedicated to Germund Dahlquist
on the occasion of his sixtieth birthday.

Abstract.

The Kreiss Matrix Theorem asserts the uniform equivalence over all $N \times N$ matrices of power boundedness and a certain resolvent estimate. We show that the ratio of the constants in these two conditions grows linearly with N, and we obtain the optimal proportionality factor up to a factor of 2. Analogous results are also given for the related problem involving matrix exponentials e^{At}. The proofs make use of a lemma that may be of independent interest, which bounds the arc length of the image of a circle in the complex plane under a rational function.

AMS Subject Classification: primary 39A11; secondary 15A45, 30A10.

1. Introduction.

Let A be an $N \times N$ matrix that satisfies the *power boundedness condition*

$$p(A) = \sup_{n \geq 0} \|A^n\| < \infty, \tag{1}$$

where $\|\cdot\| = \|\cdot\|_2$. By a power series expansion it is readily verified that A then also satisfies the *resolvent condition*

$$r(A) = \sup_{|z| > 1} (|z|-1)\,\|(zI-A)^{-1}\| < \infty, \tag{2}$$

and moreover $r(A) \leq p(A)$. One of the assertions of the Kreiss Matrix Theorem [3, 4, 7] is that the converse is also valid: if $r(A) < \infty$, then $p(A) < \infty$ also, and $p(A)$ can be bounded in terms of N and $r(A)$ but otherwise independently of A.

* Research supported by NSF Mathematical Sciences Postdoctoral Fellowships, by the Courant Institute of Mathematical Sciences, and by the National Aeronautics and Space Administration under Contract No. NAS1-17070 while the authors were in residence at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665.

Received August 1983.

This result is useful in proofs of stability theorems for finite difference approximations to partial differential equations.

In this note we resolve an old question contributed to most recently by Tadmor [8]: given N and $r(A)$, how large can $p(A)$ be? According to Tadmor, Kreiss's original proof in [4] unwinds to give a far from sharp bound

$$p(A) \lesssim [r(A)]^{N^N}, \qquad (\forall A)$$

which subsequent improvements by Morton, Strang, and Miller lowered to

$$p(A) \lesssim 6^N(N+4)^{5N}r(A), \quad N^N r(A), \quad e^{9N^2}r(A) \qquad (\forall A).$$

A few years ago Strang (private communication) observed that a paper of Laptev [5] implicitly derives a much more reasonable estimate [3]

$$p(A) \leqslant (32e/\pi)N^2 r(A) \qquad (\forall A).$$

Finally Tadmor's proof, which makes use of an elegant Cauchy integral argument adapted from Laptev, yields a bound that is linear in N,

$$p(A) \leqslant (32e/\pi)Nr(A) \qquad (\forall A). \tag{3}$$

Tadmor conjectures that a linear dependence as in (3) is the best possible. However, up to now the strongest growth of $p(A)$ with $r(A)$ attained by an example has been logarithmic, i.e., $p(A) \approx r(A)\log N$ [6].

First we will show that Tadmor's conjecture is correct, by exhibiting a family of matrices $\{A_N\}$ for which $p(A_N) \sim eNr(A_N)$ as $N \to \infty$. By refining the Cauchy integral argument, we will then show that for arbitrary matrices (3) can be sharpened to $p(A) \leqslant 2eNr(A)$. (Our proof is essentially Tadmor's, but gains the factor $16/\pi$ over his by dealing with complex functions directly rather than taking real and imaginary parts.) Together these results establish that eN is the optimal constant of proportionality relating $p(A)$ to $r(A)$ except for a possible factor of 2. The final section will prove analogous results for the continuous problem involving matrix exponentials e^{At}.

2. Example with $p(A_N) \sim eNr(A_N)$.

Consider the $N \times N$ Jordan matrix

$$A = A_N = \begin{bmatrix} 0 & \gamma & & & \\ & 0 & \gamma & & \\ & & \ddots & \ddots & \\ & & & \ddots & \gamma \\ & & & & 0 \end{bmatrix}$$

with $N \geqslant 3$, $\gamma \geqslant N$ (these constraints could be relaxed considerably). For this matrix one has $\|A^n\| = \gamma^n$ for $n \leqslant N-1$ and $\|A^n\| = 0$ otherwise, so A is power bounded with

$$p(A) = \gamma^{N-1}. \tag{4}$$

On the other hand the resolvent matrix is

$$(zI-A)^{-1} = \frac{1}{z}\begin{bmatrix} 1 & \gamma/z & (\gamma/z)^2 & \cdots & (\gamma/z)^{N-1} \\ & 1 & \gamma/z & (\gamma/z)^2 & \vdots \\ & & \ddots & & \vdots \\ & & & & 1 \end{bmatrix}.$$

From the fact that $\|B\| \leqslant \sum |B_{1i}|$ for any upper-triangular Toeplitz matrix B, we obtain with a little calculation the estimates

$$\|(zI-A)^{-1}\| \leqslant \begin{cases} 2^N/|z| \text{ if } |z| \geqslant \gamma/2, \\ \gamma^{N-1}(1-|z|/\gamma)^{-1}|z|^{-N} \text{ if } |z| \leqslant \gamma/2. \end{cases}$$

By (2), one therefore has

$$r(A) \leqslant \max\left\{ \sup_{1 \leqslant \varrho \leqslant \gamma/2} (\varrho-1)\gamma^{N-1}(1-\varrho/\gamma)^{-1}\varrho^{-N},\ \sup_{\varrho \geqslant \gamma/2} (\varrho-1)2^N/\varrho \right\}.$$

This maximum is attained at a point $\rho = 1+N^{-1}+O(N^{-2})$, where the estimate becomes

$$r(A) \leqslant \frac{\gamma^{N-1}}{eN}(1+O(N^{-1})) \tag{5}$$

since $\gamma \geqslant N$. Comparing (4) and (5) shows that for this example one has

$$p(A_N) \leqslant (eN - const)\, r(A_N), \tag{6}$$

as required.

3. Proof of $p(A) \leqslant 2eNr(A)$ for all A.

THEOREM 1. *Let A be an $N \times N$ matrix with $r(A) < \infty$. Then*

$$p(A) \leqslant 2eNr(A). \tag{7}$$

REMARK. The factor of 2 is probably unnecessary; see the remark after the lemma in the Appendix.

PROOF. Suppose $r(A) < \infty$. The matrix A^n can be written in terms of the resolvent by means of a Cauchy integral (see [2], pp. 555–577)

$$A^n = \frac{1}{2\pi i} \int z^n (zI - A)^{-1} dz, \tag{8}$$

where the contour of integration is any curve enclosing the eigenvalues of A, which must all lie in $|z| \leqslant 1$ since $r(A) < \infty$. Let u and v be arbitrary unit N-vectors, i.e., $\|u\| = \|v\| = 1$. Then

$$v^* A^n u = \frac{1}{2\pi i} \int z^n q(z) dz$$

where $q(z) = v^*(zI - A)^{-1} u$. Integrating by parts gives

$$v^* A^n u = \frac{-1}{2\pi i(n+1)} \int_\Gamma z^{n+1} q'(z) dz.$$

Let the contour Γ of integration be taken as $\Gamma : |z| = 1 + 1/(n+1)$. On this path one has $|z^{n+1}| \leqslant e$, and there follows the bound

$$|v^* A^n u| \leqslant \frac{e}{2\pi(n+1)} \int_\Gamma |q'(z)|\,|dz|.$$

Now as verified on p. 155 in [8], q is a rational function of degree N. By the lemma in the Appendix, the integral above is accordingly bounded by $4\pi N$ times the supremum of $|q(z)|$ on Γ, and by (2) this supremum is at most $(n+1)\,r(A)$. Hence we obtain

$$|v^* A^n u| \leqslant 2eNr(A).$$

Since $\|A^n\|$ is the supremum of $|v^* A^n u|$ over all unit vectors u and v, this proves the theorem. ∎

4. Analogous results for e^{At}.

For problems that are continuous in time rather than discrete, stability depends on the boundedness of a family of matrix exponentials e^{At} $(t \geqslant 0)$ rather than of powers A^n. Correspondingly, the resolvent of A is of interest for z in the right half plane rather than outside the unit circle. Following (1) and (2), define

$$P(A) = \sup_{t \geq 0} \|e^{At}\| \quad \text{and} \tag{9}$$

$$R(A) = \sup_{\operatorname{Re} z > 0} \operatorname{Re} z \|(zI - A)^{-1}\|. \tag{10}$$

As before, one has $R(A) \leqslant P(A)$ readily, this time as a consequence of the Laplace transform formula $(zI - A)^{-1} = \int_0^\infty e^{-zt} e^{At} dt$. The continuous form of the Kreiss matrix theorem asserts that conversely, $P(A)$ can be bounded in terms of $R(A)$ and N, independently of A. We make this sharp by essentially the same argument used before:

THEOREM 2. *Let A be an $N \times N$ matrix with $R(A) < \infty$. Then*

$$P(A) \leqslant 2eNR(A). \tag{11}$$

REMARK. Again the factor of 2 is probably unnecessary. The Laptev/Tadmor estimate has a constant $32e/\pi$, as in (3).

PROOF. In analogy to (8), one has now

$$e^{At} = \frac{1}{2\pi i} \int e^{zt} (zI - A)^{-1} dz,$$

where the contour of integration can be taken as any line $\operatorname{Re} z = \mu > 0$. Integration by parts gives

$$v^* e^{At} u = -\frac{1}{2\pi i t} \int e^{zt} q'(z) dz$$

with $q(z) = v^*(zI - A)^{-1} u$. Taking the contour $\mu = 1/t$ now leads to the desired bound (11), again by the use of the lemma in the Appendix. ∎

Constructing an example to prove that (11) is sharp, on the other hand, is trickier than it was in the power-boundedness case. The following example achieves growth proportional to $N^{1/2}$, not N. Omitting details, define

$$A = A_N = \begin{bmatrix} -1 & \gamma & & & \\ & -1 & \gamma & & \\ & & \ddots & \ddots & \\ & & & -1 & \gamma \\ & & & & -1 \end{bmatrix}.$$

Then one has

$$e^{At} = e^{-t} \begin{bmatrix} 1 & \gamma t & \cdots & \frac{(\gamma t)^{N-1}}{(N-1)!} \\ & 1 & \gamma t & \vdots \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}.$$

For large γ, this matrix achieves maximum norm near $t = N$, where it is dominated by the upper-right entry, with mangitude approximately

$$P(A_N) \sim \frac{e^{-N} N^{N-1} \gamma^{N-1}}{(N-1)!} \sim \frac{\gamma^{N-1}}{(2\pi N)^{1/2}}. \tag{12}$$

For the second estimate we have used Stirling's formula. On the other hand the resolvent matrix is

$$(zI - A)^{-1} = \frac{1}{z+1} \begin{bmatrix} 1 & \frac{\gamma}{z+1} & \cdots & \left(\frac{\gamma}{z+1}\right)^{N-1} \\ & 1 & \frac{\gamma}{z+1} & \vdots \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

For large γ, $\operatorname{Re} z$ times the norm of this is maximized near $z = 1/N$, where again the upper-right entry dominates and one has

$$R(A_N) \sim \frac{\gamma^{N-1}}{eN}. \tag{13}$$

Comparing (12) and (13) shows that in this example one has $P(A_N)/R(A_N) \sim (N/2\pi)^{1/2} e$.

Appendix – Lemma on arc length of a rational function on a circle.

Let S be any circle or line in the complex plane, and define the L_1 and L_∞ norms over S by $\|f\|_1 = \int_S |f(z)|\,|dz|$, $\|f\|_\infty = \sup_S |f(z)|$. The following lemma provided the key argument in proving Theorems 1 and 2. For the case of a polynomial the result is a corollary of Bernstein's inequality [1], $\|q'\|_\infty \leqslant N\|q\|_\infty$ for $S = \{z : |z| = 1\}$, but the extension to rational functions appears to be new. Since $\|q'\|_1$ represents the arc length of the image of S under q, the lemma has a simple geometric meaning. The example $q(z) = b(z-s)$, where $b(z)$ is any finite Blaschke product of degree N (such as z^N) and s is the center of S, shows that it is sharp except for a factor of 2.

LEMMA. *Let q be a rational function of degree N with no poles on S. Then*

$$\|q'\|_1 \leqslant 4\pi N \|q\|_\infty.$$

REMARK. We believe that the bound is valid with a factor 2π instead of 4π, but have been unable to prove this.

PROOF. Since the composition of q with a Möbius transformation is again a rational function of type N, we can assume without loss of generality that S is the unit circle. Define $g(z)$ to be the angle of the tangent to $q(S)$ at $q(z)$, i.e.

$$g(z) = \arg[zq'(z)].$$

Let $TV[g]$ be the total variation of g over S, i.e. the "total rotation" of $q(S)$. The lemma is a consequence of the following two facts:

$$\text{(a)}\ \|q'\|_1 \leqslant TV[g]\,\|q\|_\infty,$$

$$\text{(b)}\ TV[g] \leqslant 4\pi N.$$

The proof of (a) is a matter of integration by parts:

$$\|q'\|_1 = \oint |q'(z)|\,|dz| = \oint zq'(z)e^{-ig(z)}\frac{dz}{iz}$$

$$= -i\oint q'(z)e^{-ig(z)}dz = \oint q(z)g'(z)e^{-ig(z)}dz$$

$$\leqslant \|q\|_\infty \oint |g'(z)|\,|dz| = \|q\|_\infty\, TV[g].$$

To prove (b), note that q' is of rational type $(2N-1, 2N)$, so $zq'(z)$ is of rational type $(2N, 2N)$ and can be written as a product

$$zq'(z) = \prod_{k=1}^{2N} \frac{a_k z + b_k}{c_k z + d_k}.$$

This implies

$$g(z) = \sum_{k=1}^{2N} \arg\left(\frac{a_k z + b_k}{c_k z + d_k}\right)$$

and therefore

$$TV[g] \leqslant \sum_{k=1}^{2N} TV\left[\arg\left(\frac{a_k z + b_k}{c_k z + d_k}\right)\right] \leqslant 4\pi N. \qquad \blacksquare$$

Acknowledgements and remark. We are grateful to Paul Garabedian and Eitan Tadmor for valuable discussions. Tadmor has pointed out that if A is an arbitrary but fixed bounded operator on l_2 (infinite matrix) with $r(A) < \infty$, then the arguments of [6] can be adapted to show that $\|A^n\|$ may grow as $n \to \infty$ in proportion to $\log n$, but it is not known if it can grow faster. Our example of §2 apparently sheds no light on this question, for if one seeks a family $\{A_N\}$ with $r(A_N)$ uniformly bounded in N, the numbers γ_N have to satisfy $\gamma_N \leqslant 1 + O(1/N)$, which implies that $p(A_N)$ is also uniformly bounded.

REFERENCES

1. N. de Bruijn, *Inequalities concerning polynomials in the complex domain*, Indag. Math. 9 (1947), pp. 591–598.
2. N. Dunford and J. Schwartz, *Linear Operators I*, Wiley-Interscience (1957).
3. D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM (1977).
4. H.-O. Kreiss, *Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren*, BIT 2 (1962), pp. 153–181.
5. G. I. Laptev, *Conditions for the uniform well-posedness of the Cauchy problem for systems of equations*, Soviet Math. Dokl. 16 (1975), pp. 65–69.
6. C. A. McCarthy and J. Schwartz, *On the norm of a finite Boolean algebra of projections and applications to theorems of Kreiss and Morton*, Comm. Pure Appl. Math. 18 (1967), pp. 389–396.
7. R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, Wiley-Interscience (1967).
8. E. Tadmor, *The equivalence of L_2-stability, the resolvent condition, and strict H-stability*, Lin. Alg. and its Applics. 41 (1981), pp. 151–159.