

## PIECEWISE CONTINUOUS TOEPLITZ MATRICES AND OPERATORS: SLOW APPROACH TO INFINITY\*

ALBRECHT BÖTTCHER<sup>†</sup>, MARK EMBREE<sup>‡</sup>, AND LLOYD N. TREFETHEN<sup>‡</sup>

**Abstract.** The pseudospectra of banded finite dimensional Toeplitz matrices rapidly converge to the pseudospectra of the corresponding infinite dimensional operator. This exponential convergence makes a compelling case for analyzing pseudospectra of such Toeplitz matrices—not just eigenvalues. What if the matrix is dense and its symbol has a jump discontinuity? The pseudospectra of the finite matrices still converge, but it is shown here that the rate is no longer exponential in the matrix dimension—only algebraic.

**Key words.** Toeplitz matrix, piecewise continuous symbol, pseudospectra

**AMS subject classifications.** 47B35, 15A60

**PII.** S0895479800376971

Let  $T$  be a Toeplitz operator (singly infinite matrix) on  $\ell^2(\mathbf{N})$  with symbol  $a \in L^\infty(\mathbf{T})$ , where  $\mathbf{T}$  is the unit circle. Thus  $T = (a_{j-k})_{j,k=0}^\infty$ , where  $\{a_n\}_{n=-\infty}^\infty$  is the sequence of Fourier coefficients for  $a$ , a complex-valued function on  $\mathbf{T}$ . If  $a$  is continuous, then the spectrum  $\text{sp} T$  is the curve  $a(\mathbf{T})$  together with all of the points this curve encloses with nonzero winding number [6]. This result generalizes to piecewise continuous  $a$ : If  $a^\#(\mathbf{T})$  is the curve consisting of the components of  $a(\mathbf{T})$  connected by straight segments at points of discontinuity, then  $\text{sp} T$  is  $a^\#(\mathbf{T})$  together with all of the points it encloses with nonzero winding number; see [5, section 1.8].

A long-recognized anomaly is that the spectra of Toeplitz matrices  $T_N$  of finite dimension  $N$  look very different, typically consisting of points distributed along curves rather than across regions, even as  $N \rightarrow \infty$  [1, 5, 11, 12, 17]. Some kind of resolution of this anomaly was obtained with the discovery that, although the spectra of the matrix and the operator do not agree, the  $\varepsilon$ -pseudospectra may agree very closely [9, 10]. (The  $\varepsilon$ -pseudospectrum  $\text{sp}_\varepsilon A$  of a matrix or operator  $A$  is the set of points  $z \in \mathbf{C}$  satisfying  $\|(zI - A)^{-1}\| \geq \varepsilon^{-1}$ , where we write  $\|(zI - A)^{-1}\| = \infty$  when  $z \in \text{sp} A$ ; see, e.g., [13, 14].) In particular, if  $T_N$  is banded, then for each point  $z$  enclosed by  $a(\mathbf{T})$  with nonzero winding number,  $\|(zI - T_N)^{-1}\|$  grows exponentially as  $N \rightarrow \infty$  [3, 10]; the condition number  $\|V_N\| \|V_N^{-1}\|$  of any matrix  $V_N$  of eigenvectors of  $T_N$  is likewise exponentially large. As illustrated by numerical examples in [10], the result is that for small  $\varepsilon$ , the  $\varepsilon$ -pseudospectra of  $T_N$  typically look much like the spectrum of  $T$  for values of  $N$  on the order of hundreds.

A more general convergence result for  $\text{sp}_\varepsilon T_N$  has been proved in [2]. If  $a \in L^\infty(\mathbf{T})$  is piecewise continuous, then, for each  $\varepsilon > 0$ ,  $\text{sp}_\varepsilon T_N$  converges to  $\text{sp}_\varepsilon T$  as  $N \rightarrow \infty$ . The following question arises: If  $a$  is discontinuous, is the convergence still fast enough to be compelling for modest values of  $N$ ?

\*Received by the editors August 18, 2000; accepted for publication by P. Van Dooren May 3, 2002; published electronically November 6, 2002.

<http://www.siam.org/journals/simax/24-2/37697.html>

<sup>†</sup>Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, Germany (albrecht.boettcher@mathematik.tu-chemnitz.de).

<sup>‡</sup>Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK (LNT@comlab.ox.ac.uk). Current address of the second author: Department of Computational and Applied Mathematics, Rice University, 6100 Main Street—MS 134, Houston, TX 77005-1892 (embree@rice.edu). The research of this author was supported by UK Engineering and Physical Sciences Research Council grant GR/M12414.

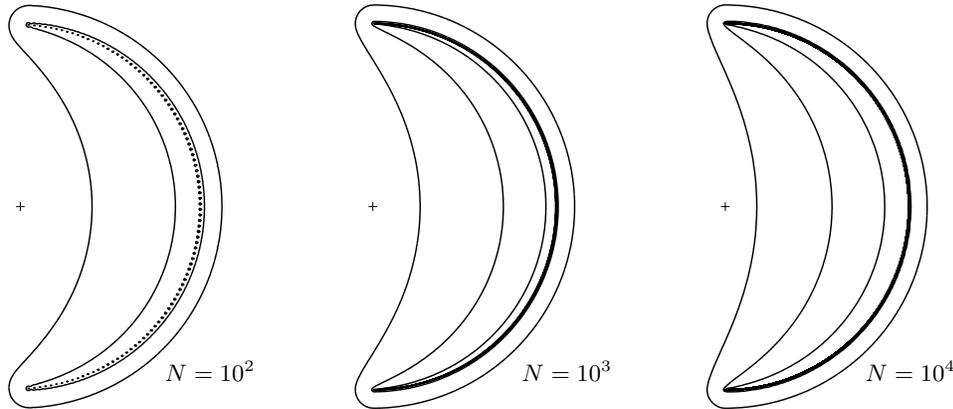


FIG. 1. Eigenvalues and  $\varepsilon$ -pseudospectra for the Toeplitz matrices  $T_N$  given by (1) for three values of  $N$  with  $\varepsilon = 10^{-1}, 10^{-2}$ , and  $10^{-3}$  (from the outside in). The cross (+) marks the origin. Except in the first image, the eigenvalues are so numerous that they appear fused into a curve. The thickness of this curve is actually due to the boundaries of the  $10^{-2}$ - and  $10^{-3}$ -pseudospectra; the boundary of the  $10^{-3}$ -pseudospectrum also affects the thickness of the middle eigenvalues in the first plot. We believe these images are correct to plotting accuracy.

We have found that the answer is no. If the symbol is discontinuous, the rate at which  $\|(zI - T_N)^{-1}\|$  and  $\|V_N\| \|V_N^{-1}\|$  increase as  $N \rightarrow \infty$  may drop from exponential to algebraic, changing the qualitative nature of the pseudospectra strikingly.

We consider the following simple example. Take  $a$  such that  $a(\mathbf{T})$  is the right half of the unit circle, specifically,  $a(e^{i\theta}) = ie^{-i\theta/2}$  for  $\theta \in [0, 2\pi)$ . Then  $\text{sp}T$  is the closed right half of the unit disk, and  $T_N$  is a dense Toeplitz matrix whose entries are given by the Fourier coefficients of the symbol

$$(1) \quad (T_N)_{jk} := \frac{1}{\pi(j - k + \frac{1}{2})}, \quad j, k = 1, \dots, N.$$

Figure 1 shows numerically computed  $\varepsilon$ -pseudospectra of  $T_N$  for  $N = 100, 1000$ , and  $10000$ , with  $\varepsilon = 10^{-1}, 10^{-2}$ , and  $10^{-3}$ . Note how far they are from  $\text{sp}T$  for the smaller values of  $\varepsilon$  and how the interior arcs approximate circles passing through  $\pm i$ . Figure 2 shows resolvent norms as a function of  $N$  for points on the real axis. For  $z = \frac{1}{2}$ , the bound  $\|(zI - T_N)^{-1}\|$  grows roughly like  $3.8N^{0.30}$ . At this rate, the resolvent norm will not exceed  $10^5$  until  $N \approx 10^{15}$ . For  $z = 0$ ,  $\|(zI - T_N)^{-1}\|$  grows roughly like  $0.4 \log N + 1.5$ ; it will not exceed  $10^5$  until  $N \approx 10^{108572}$ . This behavior is related to the ‘‘Moler phenomenon,’’ the observation that the norm of the matrix (1) approaches 1 spectacularly fast as  $N \rightarrow \infty$ , while the smallest singular value decays to 0 very slowly [5, section 4.5], [16].

Here is a mathematical foundation for these observations. Let  $a$  be a piecewise  $C^2$  function with at most one jump discontinuity, say, at  $e^{i\theta_0} \in \mathbf{T}$ . For  $z$  outside  $a(\mathbf{T})$ , let  $\arg(a - z)$  be any continuous argument of  $a - z$  on  $\mathbf{T} \setminus \{e^{i\theta_0}\}$ . Define  $\alpha_z$ , the Cauchy index of  $a$  with respect to  $z$ , by

$$\alpha_z = \frac{1}{2\pi} (\arg(a(e^{i(\theta_0+2\pi-0)}) - z) - \arg(a(e^{i(\theta_0+0)}) - z)),$$

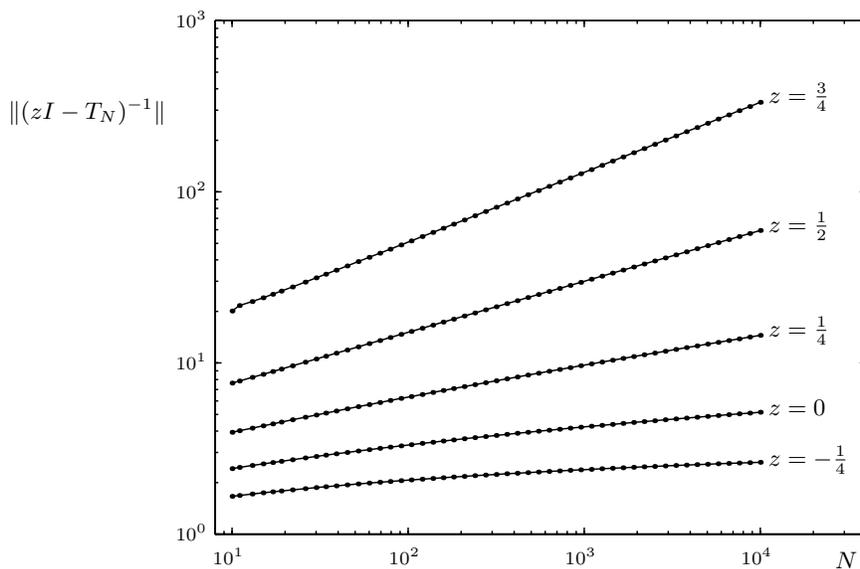


FIG. 2. The resolvent norm as a function of  $N$  for the class of matrices (1). The growth is algebraic for  $z = \frac{1}{4}, \frac{1}{2},$  and  $\frac{3}{4}$  and logarithmic for  $z = 0$ . For  $z = -\frac{1}{4}$ ,  $\|(zI - T_N)^{-1}\|$  is bounded by 4 (see Theorem 3.19 of [5]).

and put  $\beta_z = |\alpha_z|$ . If  $\beta_z < \frac{1}{2}$ , then  $zI - T_N$  is invertible for all sufficiently large  $N$ , and it is well known that  $\|(zI - T_N)^{-1}\| = O(1)$  in this case [7]. If  $\beta_z \geq 1$ , then  $\|(zI - T_N)^{-1}\|$  may grow exponentially, as trigonometric polynomials (i.e., banded matrices) with nonzero winding number about  $z$  show. The following result tells us that, for  $\frac{1}{2} \leq \beta_z < 1$ , we have just algebraic growth at a known rate.

**THEOREM.** *If  $\frac{1}{2} \leq \beta_z < 1$ , then, for every  $\delta > 0$ , there exist positive constants  $C_z$  and  $D_{z,\delta}$  such that*

$$(2) \quad C_z N^{2\beta_z - 1} \leq \|(zI - T_N)^{-1}\| \leq D_{z,\delta} N^{2\beta_z - 1 + \delta}$$

for all sufficiently large  $N$ .

In the example (1), we have  $\beta_z < \frac{1}{2}$  for all  $z$  outside  $\text{sp } T$  and  $\beta_z = \frac{1}{2}$  for  $z \in (-i, i)$ . For  $z$  in the interior of  $\text{sp } T$ , we have

$$(3) \quad \beta_z = 1 - \frac{1}{\pi} \arctan \frac{1}{x},$$

where  $x \in (0, 1)$  is the point at which the circular arc through  $-i, z, i$  intersects the real line. In particular,  $\frac{1}{2} < \beta_z < \frac{3}{4}$ , and hence, by our theorem, the resolvent norm increases at most like  $O(N^{1/2})$  for  $z$  in the interior of  $\text{sp } T$ , explaining the slow convergence seen in Figure 1. Moreover, formula (3) also reveals why the interior arcs of Figure 1 are close to circles passing through  $-i$  and  $i$ . Finally, our theorem explains Figure 2. For  $z = \frac{1}{2}$ , for example, we have  $2\beta_z - 1 = 0.295\dots$ , in good agreement with the growth  $3.8N^{0.30}$  estimated numerically.

*Sketch of the proof of the theorem.* The proof of the upper bound in (2) can be based on the argument used to prove Theorem 6.1(c) of [4]: A theorem by Verbitsky and Krupnik (see, e.g., Theorem 7.20 of [5]) states that the resolvent norm is uniformly bounded on certain weighted  $\ell^p$  spaces, and appropriate choice of these spaces,

together with Hölder’s inequality, gives the  $\ell^2$  estimate  $O(N^{2\beta_z-1+\delta})$ . To prove the lower bound in (2), assume that  $\frac{1}{2} \leq \alpha_z < 1$ . (The case  $-1 < \alpha_z \leq -\frac{1}{2}$  can be reduced to this case by passing to adjoints.) We can write  $a - z = c_z \varphi_{\gamma_z}$ , where  $c_z$  is a continuous and piecewise  $C^2$  function with no zeros on  $\mathbf{T}$  and with zero winding number and where  $\varphi_{\gamma_z}$  is a certain canonical piecewise continuous function with a single jump (see, e.g., pp. 170–171 and 182 of [5]). Here  $\gamma_z$  is a complex number whose real part equals  $\alpha_z$ . By Cramer’s rule, the  $(N, 1)$  entry of  $(zI - T_N)^{-1}$  is  $(-1)^{N+1}$  times the quotient of two Toeplitz determinants,

$$[(zI - T_N)^{-1}]_{N,1} = (-1)^{N+1} \frac{D_{N-1}(c_z \varphi_{\gamma_z-1})}{D_N(c_z \varphi_{\gamma_z})},$$

and, since  $|\operatorname{Re} \gamma_z| < 1$  and  $|\operatorname{Re} \gamma_z - 1| < 1$ , we can invoke Refinement 5.46 of [5] (which proves an important special case of the Fisher–Hartwig conjecture) to conclude that the absolute value of  $[(zI - T_N)^{-1}]_{N,1}$  is asymptotically equal to a nonzero constant times

$$\left| \frac{N^{-(\gamma_z-1)^2}}{N^{-\gamma_z^2}} \right| = |N^{2\gamma_z-1}| = N^{2\operatorname{Re} \gamma_z-1} = N^{2\beta_z-1}.$$

As the norm of  $(zI - T_N)^{-1}$  is greater than the modulus of its  $(N, 1)$  entry, we arrive at the lower bound of (2).  $\square$

For the matrix (1) at  $z = 0$ , the estimate (2) asserts that  $C \leq \|T_N^{-1}\| \leq D_\delta N^\delta$  for arbitrary  $\delta > 0$ . Using the Cauchy–Toeplitz structure of (1), Tyrtshnikov [16] showed that we actually have

$$C \log N \leq \|T_N^{-1}\| \leq D \log N.$$

We may summarize our observations as follows. Since the pseudospectra, or resolvent norms, converge,  $T_N$  must “behave” as if  $\operatorname{sp} T_N = \operatorname{sp} T$  for sufficiently large  $N$ . However, it is worth bearing in mind that a typical macroscopic physical system has on the order of  $10^8$  or  $10^{10}$  atoms or molecules in each direction (on the order of the cube root of Avogadro’s number or somewhat more). Thus, for  $T_N$  to behave like  $T$ , the dimension  $N$  will have to be larger than the numbers that usually pass for infinity in the physics of gases, liquids, and solids. Said another way, if one found a physical application governed by a matrix of the form (1), even if the dimension were very large, it is unlikely to be large enough to make approximation by the operator limit  $N = \infty$  physically appropriate for spectral analysis of the system.

As a further example, Figure 3 presents the Toeplitz matrices associated with the symbol  $a(e^{i\theta}) = \theta e^{i\theta}$ . The eigenvalues of these finite Toeplitz matrices have been studied by Basor and Morrison [1]. Our theorem provides us with the growth rate of the resolvent norm as  $N \rightarrow \infty$  in the regions where  $\beta_z < 1$ . Computational evidence suggests that the same rate is valid throughout the interior of the spectrum, although the values of  $\beta_z$  range up to  $\frac{3}{2}$ .

One could attempt to generalize our theorem and to raise conjectures suggested by our computations, but we will not pursue this here as our purpose is to point out the slow convergence phenomenon as briefly as possible.

**Note added in proof.** We wish to point out another class of problems where there exists a gap between algebraically and exponentially growing resolvent norms: certain nonsymmetric matrices related to the nonsymmetric Anderson models developed by Hatano and Nelson in the field sometimes known as nonhermitian quantum

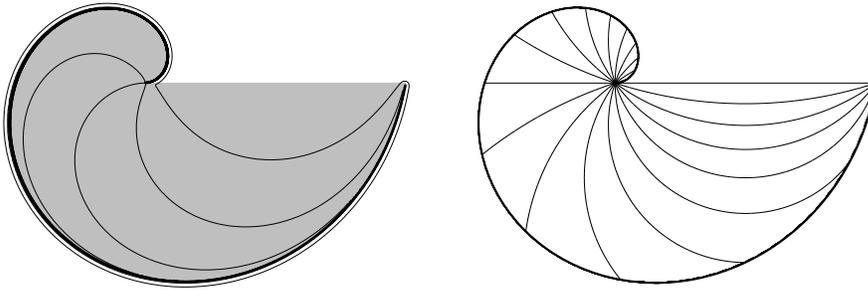


FIG. 3. Slow growth for the symbol  $a(e^{i\theta}) = \theta e^{i\theta}$ . On the left are computed eigenvalues and  $\epsilon$ -pseudospectra for the Toeplitz matrix of dimension  $N = 1000$  for  $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ . (The eigenvalues appear fused into a curve near the essential range of  $a$ .) The shaded region shows the spectrum of the corresponding infinite dimensional operator. On the right are contour lines of constant  $\beta_z$  for  $\beta_z = 0.5, 0.55, \dots, 1.45$  (clockwise from right).

mechanics [8]. In [15], it is shown that, for such matrices, the resolvent norm may grow algebraically in one part of the complex plane and exponentially in another. (An example is shown for a matrix of dimension one million, where the discrepancy in norms is between a few thousand and  $10^{99698}$ .) In these problems, as for Toeplitz matrices, it is likely that regions of exponentially large resolvent norm would “act like spectrum” in a physical application whereas other regions would not.

**Acknowledgments.** Some of our calculations were performed on the SGI Cray Origin2000 at the Oxford Supercomputing Center; the third plot of Figure 1 involved computation of the minimum singular value of  $10^4$  dense square matrices each of dimension  $10^4$ . We thank Richard Brent and Walter Gander for suggestions concerning fast Toeplitz algorithms and Nick Birkett, Jeremy Martin, and Tom Wright for advice concerning implementation and execution.

#### REFERENCES

- [1] E. L. BASOR AND K. E. MORRISON, *The Fisher-Hartwig conjecture and Toeplitz eigenvalues*, Linear Algebra Appl., 202 (1994), pp. 129–142.
- [2] A. BÖTTCHER, *Pseudospectra and singular values of large convolution operators*, J. Integral Equations Appl., 6 (1994), pp. 267–301.
- [3] A. BÖTTCHER AND S. GRUDSKY, *Toeplitz band matrices with exponentially growing condition numbers*, Electron. J. Linear Algebra, 5 (1999), pp. 104–125.
- [4] A. BÖTTCHER AND B. SILBERMANN, *Toeplitz operators and determinants generated by symbols with one Fisher-Hartwig singularity*, Math. Nachr., 127 (1986), pp. 95–124.
- [5] A. BÖTTCHER AND B. SILBERMANN, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, New York, 1999.
- [6] I. GOHBERG, *On the application of the theory of normed rings to singular integral equations*, Uspekhi Mat. Nauk, 7 (1952), pp. 149–156 (in Russian).
- [7] I. GOHBERG, *Toeplitz matrices composed of the Fourier coefficients of piecewise continuous functions*, Funktsional. Anal. i Prilozhen., 1 (1967), pp. 91–92 (in Russian).
- [8] N. HATANO AND D. R. NELSON, *Localization transitions in non-Hermitian quantum mechanics*, Phys. Rev. Lett., 77 (1996), pp. 570–573.
- [9] H. J. LANDAU, *On Szegő’s eigenvalue distribution theory and non-Hermitian kernels*, J. Anal. Math., 28 (1975), pp. 335–357.

- [10] L. REICHEL AND L. N. TREFETHEN, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices*, Linear Algebra Appl., 162–164 (1992), pp. 153–185.
- [11] P. SCHMIDT AND F. SPITZER, *The Toeplitz matrices of an arbitrary Laurent polynomial*, Math. Scand., 8 (1960), pp. 15–38.
- [12] P. TILLI, *Some results on complex Toeplitz eigenvalues*, J. Math. Anal. Appl., 239 (1999), pp. 390–401.
- [13] L. N. TREFETHEN, *Pseudospectra of matrices*, in Numerical Analysis 1991, D. F. Griffiths and G. A. Watson, eds., Longman Scientific and Technical, Harlow, UK, 1992, pp. 234–266.
- [14] L. N. TREFETHEN, *Pseudospectra of linear operators*, SIAM Rev., 39 (1997), pp. 383–406.
- [15] L. N. TREFETHEN, M. CONTEDINI, AND M. EMBREE, *Spectra, pseudospectra, and localization for random bidiagonal matrices*, Comm. Pure Appl. Math., 54 (2001), pp. 595–623.
- [16] E. E. TYRTYSHNIKOV, *Singular values of Cauchy–Toeplitz matrices*, Linear Algebra Appl., 161 (1992), pp. 99–116.
- [17] H. WIDOM, *Eigenvalue distribution for nonselfadjoint Toeplitz matrices*, in Toeplitz Operators and Related Topics, Oper. Theory Adv. Appl. 71, Birkhäuser, Basel, 1994, pp. 1–8.