NEAR NORMAL DILATIONS OF NONNORMAL MATRICES AND LINEAR OPERATORS

ANNE GREENBAUM*, TREVOR CALDWELL†, AND KENAN LI‡

Abstract. Let A be a square matrix or a linear operator on a Hilbert space $\mathcal{H}$. A *dilation* of A is a linear operator M on a larger space $\mathcal{K} \supset \mathcal{H}$ such that $A = P_{\mathcal{H}} M|_{\mathcal{H}}$, where $P_{\mathcal{H}}$ is orthogonal projection onto $\mathcal{H}$. Often it is required additionally that M^k be a dilation of A^k for all or a range of positive integer powers k. While much work has been aimed at proving existence of dilations with various properties, there has been little study of the behavior of functions of these dilations and how it compares to that of the original operator. Is the original operator a major part of the dilated one or is it an insignificant piece? Does the larger operator represent some physical process, where the original operator might be an important component for certain times but not for others? In this paper we construct near normal dilations of nonnormal matrices, with the spectrum of the dilated operator around the boundary of the numerical range of the matrix. We compare the behavior of e^{tA} and e^{tM}, for $t > 0$. We find that the dilated operator takes on a life of its own, representing a wave that grows or decays but eventually dominates the part corresponding to the original operator. We derive other near normal dilations in which this behavior is less pronounced.

1. Introduction. Let A be a square matrix or a linear operator on a Hilbert space $\mathcal{H}$. A *dilation* of A is a linear operator M on a larger space $\mathcal{K} \supset \mathcal{H}$ such that $A = P_{\mathcal{H}} M|_{\mathcal{H}}$, where $P_{\mathcal{H}}$ is orthogonal projection onto $\mathcal{H}$. For instance, if A is an n by n matrix and we identify $\mathcal{H} = \mathbb{C}^n$ with the subset of $\mathcal{K} = \mathbb{C}^N$, $N > n$, consisting of vectors whose last $N - n$ components are 0, then any N by N matrix of the form

$$M = \begin{bmatrix} A & * \\ * & * \end{bmatrix}$$

is a dilation of A. Often it is required additionally that M^k be a dilation of A^k for all or a range of positive integer powers k, and when this holds for all $k \in \mathbb{N}^+$, M is said to be a *power dilation* of A.

A reason for studying dilations is that the operator M may have some nice properties that A does not have. For example, if A is highly *nonnormal*, then it is well-known that the spectrum $\sigma(A)$ may give little information about the 2-norm behavior of functions of A. If M is a normal or near normal power dilation of A, however, then $\|p(A)\| \leq \|p(M)\|$ for any polynomial p (since $p(A)$ is a block of $p(M)$), and if $\|\cdot\|$ denotes the 2-norm (as it will throughout this paper), then $\|p(M)\|$ is determined or approximately determined by $\sigma(M)$. That is, if M is similar to a normal operator $\mathcal{N}$ via a similarity transformation with a moderate condition number Q ($M = S\mathcal{N}S^{-1}$,

*University of Washington, Applied Mathematics Dept., Box 353925, Seattle, WA 98195. This work was supported in part by NSF grant DMS-1210886.

†University of Washington, Applied Mathematics Dept., Box 353925, Seattle, WA 98195.

‡University of Washington, Applied Mathematics Dept., Box 353925, Seattle, WA 98195.

where $\kappa(S) \equiv \|S\| \cdot \|S^{-1}\| = Q$), then $\|p(A)\| \leq \|p(M)\| \leq Q \sup_{z\in\sigma(M)} |p(z)|$.

In this paper, we construct near normal dilations of nonnormal matrices, with the spectrum of the dilated operator around the boundary of the *field of values*, or, *numerical range* of the matrix A that it dilates:

$$W(A) = \{(Aq, q) : q \in \mathcal{H},\ (q, q) = 1\}, \tag{1.1}$$

where $(\cdot, \cdot)$ denotes the inner product in $\mathcal{H}$. We compare the behavior of e^{tA} and e^{tM} for $t > 0$. While the ratio $\|e^{tM}\|/\|e^{tA}\|$ is close to 1 for small t, we find that other parts of the operator e^{tM} start to dominate for larger values of t and this ratio can become arbitrarily large. This is to be expected since asymptotically $\|e^{tA}\|$ behaves like a polynomial in t times $e^{t\gamma(A)}$ where $\gamma(A)$ is the *spectral abscissa* ($\sup\{\mathrm{Re}(z) : z \in \sigma(A)\}$), while, if $\sigma(M)$ is densely distributed around $\partial W(A)$, then $\|e^{tM}\|$ behaves, for all time, like a modest multiple (between $1/Q$ and Q) of $e^{t\alpha(A)}$, where $\alpha(A)$ is the *numerical abscissa* ($\sup\{\mathrm{Re}(z) : z \in W(A)\}$). If $\alpha(A) > \gamma(A)$, then $\|e^{tM}\|/\|e^{tA}\| \to \infty$ as $t \to \infty$.

What may be surprising is the mechanism by which this difference develops. Whatever physical process the original equation $\mathbf{y}' = A\mathbf{y}$ (whose solution is $\mathbf{y}(t) = e^{tA}\mathbf{y}(0)$) might represent, the dilated equation $\hat{\mathbf{y}}' = M\hat{\mathbf{y}}$ (whose solution is $\hat{\mathbf{y}}(t) = e^{tM}\hat{\mathbf{y}}(0)$) is seen to represent a wave moving to the left and growing or decaying at a rate determined by $\alpha(A)$. It will be shown in the next section that one step in forming the near normal dilation M is to use the Sz.-Nagy dilation theorem [7] (or a related finite dimensional version) to construct a unitary dilation of a contraction C derived from A. An early observation by Schreiber [17] was that *all* such dilations of strict contractions C are unitarily equivalent to the bilateral shift. Thus much information about the operator A is lost during this step and only information about $W(A)$ is recovered in the final dilation M. This suggests that to find a dilation whose behavior more closely resembles that of A, more information about A must be encoded into M. We discuss a possible way to do this.

The organization of this paper is as follows. In Section 2, we describe the construction of near normal dilations with spectrum on $\partial W(A)$, the boundary of $W(A)$. All of this material is known, but we have not found explicit constructions of such dilations elsewhere. In Section 3, we give some analytic examples and Section 4 describes numerical techniques for constructing such dilations. In Section 5, we illustrate the behavior of e^{tA} and e^{tM} for some example cases and we briefly discuss the behavior of powers of A and M. In Section 6, we consider a different near normal dilation whose spectrum is shifted as far to the left as possible for the given form of the dilation, although it now encloses a larger region than $W(A)$. Dilations of this sort show better agreement between $\|e^{tA}\|$ and $\|e^{tM}\|$ for a longer period of time. Section 7 describes still more possibilities for spectra of near normal dilations.

2. Near Normal Dilations with Spectrum Around $\partial W(A)$. It was shown by Sz.-Nagy [7] that every contraction C ($\|C\| \le 1$) has a unitary power dilation U. In matrix form, Sz.-Nagy's dilation is a doubly infinite block triangular matrix known as the Schäffer matrix [16]:

$$U = \begin{bmatrix} \ddots & \ddots & & & & & \\ & 0 & I & & & & \\ & & 0 & D_C & -C^* & & \\ & & & C & D_{C^*} & & \\ & & & & 0 & I & \\ & & & & & 0 & \ddots \\ & & & & & & \ddots \end{bmatrix}, \quad D_C = (I - C^*C)^{1/2}. \tag{2.1}$$

This is *necessarily* an infinite dilation and if it is truncated at some point, as in

$$\begin{bmatrix} 0 & I & & & \\ & 0 & D_C & -C^* & \\ & & C & D_{C^*} & \\ & & & 0 & I \\ & & & & 0 \end{bmatrix},$$

the truncated matrix is by no means unitary – its eigenvalues are 0 together with the eigenvalues of C, and it is not even diagonalizable.

Another form of unitary dilation is described, for example, in [9, p. 59]. Start with any unitary dilation of C; e.g.,

$$\begin{bmatrix} C & D_{C^*} \\ D_C & -C^* \end{bmatrix}. \tag{2.2}$$

Form a larger matrix U_K with the blocks of the matrix (2.2) in the $(1,1)$, $(1,2)$, $(K,1)$, and $(K,2)$ block positions and with block rows 2 through $K-1$ containing identity operators in blocks $(k,k+1)$, $k = 2,\ldots,K-1$, and zeros elsewhere:

$$U_K = \begin{bmatrix} C & D_{C^*} & & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \\ D_C & -C^* & & & \end{bmatrix}. \tag{2.3}$$

It is easy to check that this finite matrix is unitary and that U_K^k is a dilation of C^k, for $k = 1,\ldots,K-1$. In the limit as $K \to \infty$, this approaches Sz.-Nagy's minimal

isometric dilation [7].

If an operator T is similar to a contraction via a well-conditioned similarity transformation, say, $T = XCX^{-1}$ where C is a contraction and $\kappa(X) \le Q$, then the unitary dilations (2.1) and (2.3) of C become near normal dilations of T when appropriate similarity transformations are applied. Defining $S =$ block diag$(\ldots, I, X, I, \ldots)$ and $S_K =$ block diag$(X, I, \ldots, I)$, we find that SUS^{-1} and $S_K U_K S_K^{-1}$ are dilations of T, and, assuming that X has been normalized so that, say, $\|X\| = 1$, the condition numbers of S and S_K are the same as that of X.

A set $\Omega \subset \mathbb{C}$ is said to be a *complete Q-spectral set* for T if

$$\|P(T)\| \le Q \sup_{z\in\Omega} \|P(z)\|, \tag{2.4}$$

for all *matrix-valued polynomials* P. Here $P(z)$ is an ℓ by m matrix (where ℓ and m can be any positive integers) whose (i, j)-entry is $p_{ij}(z)$ for some polynomial p_{ij}, and the operator $P(T)$ is an ℓ by m array of operators whose (i, j)-block is $p_{ij}(T)$. It was shown by Paulsen [15, Theorem 9.11, p. 127] that if the unit disk $\mathcal{D}$ is a *complete Q-spectral set* for an operator T, then T is similar to a contraction via a similarity transformation with condition number Q: $T = XCX^{-1}$, where $\kappa(X) = Q$.

In 2004, Michel Crouzeix [3] made the fascinating conjecture that for any square matrix A and any polynomial p,

$$\|p(A)\| \le 2 \sup_{z\in W(A)} |p(z)|. \quad \text{(Crouzeix's conjecture)}$$

In 2007, he was able to *prove* this result with 2 replaced by 11.08 and even to establish a completely bounded version of the inequality [4]:

$$\|P(A)\| \le Q \sup_{z\in W(A)} \|P(z)\|, \quad \text{(Crouzeix's theorem)} \tag{2.5}$$

for some Q with $2 \le Q \le 11.08$, for all ℓ by m matrix-valued polynomials P and for all positive integers ℓ and m.

Assume that the interior of $W(A)$ is nonempty and not all of $\mathbb{C}$. Let g be a Riemann mapping (i.e., a bijective conformal mapping) from the interior of $W(A)$ to the unit disk $\mathcal{D}$ (extended continuously to the boundary). Then it follows from (2.5) that $\|P(g(A))\| \le Q \sup_{z\in\mathcal{D}} \|P(z)\|$; that is, the unit disk $\mathcal{D}$ is a complete Q-spectral set for $g(A)$. Hence from Paulsen's theorem, $g(A) = XCX^{-1}$ where C is a contraction and $\kappa(X) = Q$.

Now Sz.-Nagy's theorem implies that C has a unitary power dilation U, and $g^{-1}(U)$ is a normal power dilation of $g^{-1}(C)$ with spectrum on $\partial W(A)$. Letting $S =$ block diag$(\ldots, I, X, I, \ldots)$, we find that $Sg^{-1}(U)S^{-1}$ is a power dilation of A with spectrum on $\partial W(A)$ that is similar to the normal operator $g^{-1}(U)$ via a similarity

transformation S with condition number Q. Thus every linear operator A (whose numerical range has nonempty interior that is not all of $\mathbb{C}$) has a near normal power dilation with spectrum on $\partial W(A)$. [Note that if the interior of $W(A)$ is empty, then $W(A)$ is either a single point or a line segment, and A itself is normal.]

If C is a finite matrix (or, more generally, an *algebraic operator*; i.e., one for which $p(C) = 0$ for some polynomial p) and one considers a dilation U_K of the form (2.3), then one can construct a kth degree polynomial approximation f_k to g^{-1} that also satisfies $f_k(C) = g^{-1}(C)$. Assuming that $K \geq k+1$, then $f_k(U_K)$ will be a normal dilation of $g^{-1}(C)$ with spectrum *approximately* on $\partial W(A)$ and satisfying $[f_k(U_K)]^\ell$ is a dilation of $[g^{-1}(C)]^\ell$ for $\ell = 1, \ldots, K-k$. By increasing K, the degree of f_k can be taken to be large enough to approximate g^{-1} on $\overline{\mathcal{D}}$ to any desired level of accuracy and to make powers $[f_k(U_K)]^\ell$ be dilations of $[g^{-1}(C)]^\ell$ for any finite number of powers ℓ. Then defining $S_K = \text{block diag}(X, I, \ldots, I)$, we have $S_K f_k(U_K) S_K^{-1}$ is a dilation of A (whose powers up to a desired limit are dilations of the corresponding powers of A) with spectrum (approximately) on $\partial W(A)$, that is similar to the normal operator $f_k(U_K)$ via a similarity transformation S_K whose condition number is Q.

3. Examples. *Example 1. Jordan block.* Let J_λ be an n by n Jordan block with eigenvalue λ. The field of values of J_λ is the disk about λ of radius $r = \cos(\pi/(n+1))$. A Riemann mapping g from $W(J_\lambda)$ to the unit disk $\mathcal{D}$ that maps λ to 0 is $g(z) = (z-\lambda)/r$, and for this mapping we have

$$g(J_\lambda) = \begin{bmatrix} 0 & r^{-1} & & \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & r^{-1} \\ 0 & \ldots & \ldots & 0 \end{bmatrix} \tag{3.1}$$

Let $X = \text{diag}(1, r, \ldots, r^{n-1})$. Then the condition number of X is $r^{-(n-1)}$ which is easily seen to be less than or equal to 2 (with equality if $n = 2$ or $n = 3$) and decreasing towards 1 as $n \to \infty$. The matrix $X^{-1} g(J_\lambda) X$ has 1's on its superdiagonal and 0's elsewhere and hence is a contraction. If we denote this contraction by C, then (2.1) defines an infinite power dilation of C, and (2.3) defines a finite dilation of C whose powers up to $K-1$ are dilations of the corresponding powers of C. Applying $g^{-1}(z) = rz + \lambda$ to U and U_K, we obtain normal dilations of $g^{-1}(C)$ with spectrum on $\partial W(J_\lambda)$. Finally, if $S = \text{block diag}(\ldots, I, X, I, \ldots)$ and $S_K = \text{block diag}(X, I, \ldots, I)$,

then the matrices

$$
(3.2) \qquad Sg^{-1}(U)S^{-1} = \begin{bmatrix} \ddots & \ddots & & & & & \\ & \lambda I & rI & & & & \\ & & \lambda I & rD_C X & -rC^* & & \\ & & & J_\lambda & rX^{-1}D_{C^*} & & \\ & & & & \lambda I & rI & \\ & & & & & \lambda I & \ddots \\ & & & & & & \ddots \end{bmatrix}
$$

and

$$
(3.3) \qquad S_K g^{-1}(U_K)S_K^{-1} = \begin{bmatrix} J_\lambda & rX^{-1}D_{C^*} & & & \\ & \lambda I & rI & & \\ & & \ddots & \ddots & \\ & & & \lambda I & rI \\ rD_C X & -rC^* & & & \lambda I \end{bmatrix}
$$

are dilations of J_λ with spectrum on $\partial W(J_\lambda)$ that are similar to the normal operators $g^{-1}(U)$ and $g^{-1}(U_K)$, respectively, via similarity transformations with condition number $r^{-(n-1)} \leq 2$.

Example 2. Left shift operator. Let $S_L : \ell^2 \to \ell^2$ be the left shift operator: $S_L(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. Its Hermitian transpose is the right shift operator: $S_L^*(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$. The numerical range of S_L is

$$
\left\{ \sum_{j=1}^{\infty} \bar{x}_j x_{j+1} : \quad \sum_{j=1}^{\infty} |x_j|^2 = 1 \right\},
$$

which can be seen to be the open unit disk $\mathcal{D}$. Since $\|S_L\| = 1$, it follows that S_L has a unitary power dilation. The operators $(I - S_L S_L^*)$ and $(I - S_L^* S_L)$ satisfy

$$
(I - S_L S_L^*)(x_1, x_2, \ldots) = (x_1, x_2, \ldots) - S_L(0, x_1, x_2, \ldots) = (0, 0, \ldots),
$$

$$
(I - S_L^* S_L)(x_1, x_2, \ldots) = (x_1, x_2, \ldots) - S_L^*(x_2, x_3, \ldots) = (x_1, 0, 0, \ldots),
$$

and their square roots are 0 and $(I - S_L^* S_L)^{1/2} = (I - S_L^* S_L)$, respectively. Thus a

unitary dilation of S_L is the linear operator on $\ell^2 \oplus \ell^2$ that satisfies

$$\begin{bmatrix} S_L & 0 \\ (I - S_L^* S_L)^{1/2} & -S_L^* \end{bmatrix} \begin{bmatrix} (x_j)_{j=1}^\infty \\ (y_j)_{j=1}^\infty \end{bmatrix} = \begin{bmatrix} (x_{j+1})_{j=1}^\infty \\ \begin{pmatrix} x_1 \\ (-y_j)_{j=1}^\infty \end{pmatrix} \end{bmatrix}.$$

Since this is triangular it is also a power dilation of S_L, with spectrum on $\partial W(S_L)$.

Example 3. First derivative operator with a Dirichlet boundary condition. Let $L = \frac{d}{dx}$ be the first derivative operator whose domain is restricted to the set of absolutely continuous functions u on $[0,1]$ whose first derivative is in $\mathcal{L}^2(0,1)$ and which satisfy $u(1) = 0$ (cf. Example 2.7, p. 145 [10]). Then it can be shown that the spectrum of L is empty: $\sigma(L) = \phi$. The numerical range of L, however, is the entire left half-plane. Let $g(z) = (z+1)/(z-1)$. Then g is a Riemann mapping from the left half-plane to the unit disk $\mathcal{D}$, mapping the imaginary axis to the unit circle. The operator $g(L)$ is

$$g(L)u(x) = -(I-L)^{-1}(I+L)u(x) = u(x) - 2\int_x^1 e^{x-s}u(s)\,ds.$$

It follows from a theorem of von Neumann [13] that $\|g(L)\|_{\mathcal{L}^2(0,1)} \leq 1$. Hence if C is replaced by $g(L)$ in (2.1) and (2.3), we obtain unitary dilations of $g(L)$. The inverse operator g^{-1} turns out to be the same as g, so if we then apply this operator to U in (2.1), we obtain a normal dilation of L with spectrum on the imaginary axis. Since g^{-1} is not a polynomial in this case, the best one could obtain from (2.3) is a normal dilation of an operator $f_k(g(L))$ that in some sense approximates L and whose spectrum lies on the image of the unit circle under f_k.

4. Numerical Construction of Near Normal Dilations. Given an n by n matrix A, one can construct a finite near normal dilation M of A with spectrum (approximately) on $\partial W(A)$ as follows:

1. Compute the field of values $W(A)$. We did this using the `fov` command in `chebfun` [6]. This uses a standard algorithm [9, Sec. 1.5] for computing points around $\partial W(A)$, but makes use of the fact that the boundary is smooth (or, at least, piecewise smooth) to fit a Chebyshev or trigonometric series to the computed points and thereby obtain a level of accuracy near the machine precision.
2. Find a bijective conformal mapping from $W(A)$ to the unit disk $\mathcal{D}$. Here we used the Kerzman-Stein integral equation [11, 12] to find the boundary correspondence function. With points equally spaced in arclength around $\partial W(A)$, and using the super-algebraically convergent trapezoidal rule for integrals over $\partial W(A)$, we obtain accurate approximations to the images of these points on the unit circle. A (`chebfun`) trigonometric polynomial can then be

fit to the computed points and their images, mapping arclengths associated with points on $\partial W(A)$ to arclengths (or angles) associated with their images on the unit circle. This is the boundary correspondence function g_B.

To demonstrate the accuracy of the computed field of values and boundary correspondence function g_B, we applied the code to 2 by 2 matrices whose fields of values are elliptical disks and compared the results to known boundary correspondence functions for ellipses. Table 1 shows the maximum error over 5000 points equally spaced in arclength on $\partial W(A)$, when $\partial W(A)$ is an ellipse with ratio of major to minor axis equal to 2, 3, 5, and 10. With this approach we can accurately compute g_B even when $W(A)$ is a highly eccentric ellipse, provided that we use a reasonable number of points N_I in the integral equation. This requires solving a well-conditioned N_I by N_I system of linear equations. We do this directly, but for very large values of N_I an iterative method could easily be substituted.

N_I	2-1 ellipse	3-1 ellipse	5-1 ellipse	10-1 ellipse
16	6.5e-3	4.4e-2		
32	1.8e-5	8.6e-4	1.8e-2	
64	2.9e-8	1.4e-6	1.4e-4	
128		6.1e-9	6.9e-7	1.3e-4
256			9.2e-10	1.7e-5
512				2.7e-7
1024				1.2e-9

TABLE 1

Max error in boundary correspondence function g_B when $\partial W(A)$ is an ellipse with given ratio of major to minor axis and N_I points equally spaced in arc length around $\partial W(A)$ are used in the Kerzman-Stein integral equation.

3. Evaluate $g(A)$, where g is the Riemann mapping from $W(A)$ to $\mathcal{D}$ whose boundary values are determined by g_B. Here we used the Cauchy integral formula,

$$g(A) = \frac{1}{2\pi i} \int_{\partial W(A)} (zI - A)^{-1} g_B(z)\, dz,$$

again approximating the integral using the trapezoidal rule with the same equally spaced boundary points z used in the integral equation.

4. Once $g(A)$ has been formed, look for a similarity transformation X such that $C = X^{-1} g(A) X$ is a contraction and such that the condition number of X is as small as possible. This can be done using a constrained optimization package such as `fmincon` in Matlab. Although there is no guarantee that the optimal X will be found, we have generally been successful (for very small problems) at finding X with condition number less than or equal to 2 such

that $\|X^{-1}g(A)X\| \leq 1$.

5. Having found C, form a unitary dilation U_K of the form (2.3). We just guess at the number of blocks K that will be needed to ensure that functions of interest can be approximated using polynomials of degree at most $K-1$. This can be adjusted if necessary.
6. Evaluate $g^{-1}(U_K)$ and a polynomial approximation $f_k(U_K)$, where: (1) the degree k is significantly less than $K-1$, (2) $f_k(C) = g^{-1}(C)$ (so that the (1,1)-block of $f_k(U_K)$ will be exactly $X^{-1}AX$), and (3) $f_k(U_K) \approx g^{-1}(U_K)$ (so that the eigenvalues of $f_k(U_K)$ will lie approximately on $\partial W(A)$). If g^{-1} can be well-approximated by a polynomial of degree less than $K-1$, then the $(1,1)$-block of $g^{-1}(U_K)$ will be a close approximation to $g^{-1}(C)$, and other blocks that would be zero in $f_k(U_K)$ will be approximately zero in $g^{-1}(U_K)$. If this is not the case, then K must be increased. Typically, if long thin regions are mapped to the unit disk, then the degree of polynomial needed to approximate the inverse map is quite large, and the required value of K may be so large as to render the problem computationally infeasible.

   To compute $g^{-1}(U_K)$, we first do an eigendecomposition, $U_K = Q_K D_K Q_K^*$, and then evaluate $g^{-1}(D_K) = g_B^{-1}(D_K)$ using bisection. We locate each eigenvalue between a pair of mapped points on the unit circle (computed via the integral equation) and find the points on $\partial W(A)$ that mapped to this pair. The interval between this pair of points is bisected until we find a point whose image under g_B is sufficiently close to the eigenvalue. This could be accelerated with Newton's method, where a Newton step is accepted only if it lies within the interval known to contain the root. This is especially convenient with the Kerzman-Stein integral equation since it also produces an approximation to the derivative g_B'.

   Knowing $g^{-1}(D_K)$, we choose f_k to be of the form $q_1 + q_0 q_2$, where $q_1(C) = g^{-1}(C)$ (which is achieved by setting $q_1(z) = \sum_{j=1}^{n} \lambda_j \prod_{\ell \neq j} \frac{z - g(\lambda_\ell)}{g(\lambda_j) - g(\lambda_\ell)}$ if A is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$), $q_0(C) = 0$ (which is achieved by setting $q_0(z) = \prod_{\ell=1}^{n}(z - g(\lambda_\ell))$), and q_2 is the polynomial of degree $k-n$ that minimizes $\sum_{j=1}^{N} |q_0(d_j)q_2(d_j) - (g^{-1}(d_j) - q_1(d_j))|^2$ where $D_K = \mathrm{diag}(d_1, \ldots, d_N)$.
7. Finally, after computing $f_k(U_K)$, multiply its first block row on the left by X and its first block column on the right by X^{-1} to obtain the dilation M of A.

This is a rather involved procedure, and we have carried it out successfully only for very small matrices A.

If, instead of requiring M to have its spectrum on $\partial W(A)$, we require it to have its spectrum on a circle enclosing $W(A)$, the computation becomes *much* easier, and we have a guaranteed algorithm for computing such a dilation M having eigenvectors with condition number at most 2. We first compute $W(A)$ and the smallest circle

(or some other circle) enclosing it. This is mapped to the unit disk by shifting and scaling: If the circle about $W(A)$ has center λ and radius r, then $g(z) = (z - \lambda)/r$. Now the field of values of $g(A)$ is just the scaled and shifted field of values of A – $W(g(A)) = (W(A) - \lambda)/r$ – so $W(g(A))$ is a subset of $\mathcal{D}$. A fixed point iteration described in [2], based on work of Ando [1] and Okubo and Ando [14], can be used when the numerical radius of $g(A)$ is less than or equal to 1 to find a matrix X with $\kappa(X) \leq 2$ such that $C = X^{-1}g(A)X$ is a contraction, and the algorithm is guaranteed to converge. With C thus constructed, we form the unitary dilation U_K in (2.3), and since $g^{-1}(z) = rz + \lambda$ is just a first degree polynomial, there is no problem with evaluating $g^{-1}(U_K)$. The matrix $M = S_K g^{-1}(U_K) S_K^{-1}$, where $S_K =$ block diag$(X, I, \ldots, I)$, will be a dilation of A with spectrum on the disk enclosing $W(A)$ and with eigenvectors having condition number at most 2. Its powers M^k will be dilations of the corresponding powers A^k, for $k = 1, \ldots, K-1$.

5. Behavior of e^{tA} and e^{tM}. Consider the pair of coupled ordinary differential equations $\mathbf{y}' = A\mathbf{y}$, where

$$A = \begin{bmatrix} -1 & 4 \\ 0 & -1 \end{bmatrix}. \tag{5.1}$$

For any given initial vector $\mathbf{y}(0) = [u_0, v_0]^T$, the solution is $\mathbf{y}(t) = e^{tA}\mathbf{y}(0)$, or, $v(t) = e^{-t}v_0$, $u(t) = e^{-t}u_0 + 4te^{-t}v_0$. The only eigenvalue of A is -1 and the field of values is the disk of radius 2 about -1. Asymptotically as $t \to \infty$, $\|\mathbf{y}(t)\|$ decays like te^{-t}. For all t, it is known (see, e.g., [18, p. 138]) that $\|\mathbf{y}(t)\| \leq e^{t\alpha(A)}$, where $\alpha(A) = 1$ is the numerical abscissa. For small $t > 0$, $\|\mathbf{y}(t)\|$ tends to behave like $e^{t\alpha(A)} = e^t$, but for larger time this is not the case.

Let M be a near normal dilation of A. Using the construction of the previous section, we produce a finite dilation M with eigenvalues around $\partial W(A)$ having the form:

$$M = \left[\begin{array}{cc|cc|ccc|cc}
-1 & 4 & 0 & 0 & & & & & \\
0 & -1 & 0 & 1 & & & & & \\
\hline
 & & -1 & 0 & 2 & 0 & & & \\
 & & 0 & -1 & 0 & 2 & & & \\
\hline
 & & & & \ddots & & & \ddots & \\
 & & & & & \ddots & & & \ddots \\
\hline
2 & 0 & 0 & 0 & & & & -1 & 0 \\
0 & 0 & -2 & 0 & & & & 0 & -1
\end{array}\right], \tag{5.2}$$

where the diagonal dots in the matrix indicate many more blocks identical to the ones above.

Suppose $\hat{\mathbf{y}}' = M\hat{\mathbf{y}}$. Then $\hat{\mathbf{y}}(t) = e^{tM}\hat{\mathbf{y}}(0)$ and $\|e^{tM}\| \le 2e^{t\alpha(A)} = 2e^t$ since M has an eigenvector matrix with condition number 2 and the spectral abscissa of M is at most $\alpha(A)$. Since e^{tM} (or a polynomial approximation to it, where the degree of the polynomial can be taken to be arbitrarily large by including enough blocks in place of the diagonal dots in (5.2)) is a dilation of e^{tA} (or the same polynomial approximation to e^{tA}), it follows that $\|e^{tA}\| \le \|e^{tM}\|$. This is slightly weaker (by a factor of 2) than the known bound $\|e^{tA}\| \le e^{t\alpha(A)}$ but this technique of bounding $\|f(A)\|$ by $\|f(M)\|$ can be used for other functions as well.

Suppose $\hat{\mathbf{y}}(0) = [u_0, v_0, 0, \ldots, 0]^T$. Figure 1 shows the behavior of $\mathbf{y}(t)$ and $\hat{\mathbf{y}}(t)$ at various times $t > 0$, where $\mathbf{y}(0) = [u_0, v_0]^T$ was chosen randomly. The asterisks represent the components of $\mathbf{y}(t)$ and the solid line goes through the components of $\hat{\mathbf{y}}(t)$. The horizontal axis is just the index of the components. Here we inserted 37 blocks where the diagonal dots are in (5.2) so the dimension of M was 80 by 80. For early time (e.g., $t = 0.1$ and $t = 0.2$), the major action is in the first two components. A small disturbance forms at the right end of the plots. By times $t = .5$ and $t = 1$, this disturbance at the right has grown to be about the same size as the first two components. At $t = 2$, the first two components have decreased in size and are now significantly smaller than the rightmost components, which have grown. By $t = 4$, the wave at the right has not only grown in size but its peak has shifted a bit to the left, and it is now so much larger than the first two components that it is clear that the behavior of $\hat{\mathbf{y}}(t)$ has almost nothing to do with that of $\mathbf{y}(t)$. While this figure shows the behavior of $e^{tM}\hat{\mathbf{y}}(0)$ for a special vector $\hat{\mathbf{y}}(0)$ that has nonzeros only in its first two components, the same behavior is reflected in the norm ratios, $\|e^{tM}\|/\|e^{tA}\|$, which are printed at the top of each plot. For $t = 0.1$ and $t = 0.2$, the norm ratio is close to 1, by $t = 1$ it has grown to almost 2, and by $t = 4$ the norm of e^{tM} is so much larger than that of e^{tA} that the bound $\|e^{tA}\| \le \|e^{tM}\|$ would no longer be of interest for most purposes.

While the original equation $\mathbf{y}' = A\mathbf{y}$ was a pair of coupled ordinary differential equations, the dilated equation $\hat{\mathbf{y}}' = M\hat{\mathbf{y}}$ might be thought of as a spatially differenced approximation to a pair of coupled *partial differential equations*. Define

$$\hat{\mathbf{y}}(t) = \begin{bmatrix} u_1(t) \\ v_1(t) \\ \vdots \\ u_K(t) \\ v_K(t) \end{bmatrix},$$

where $u_j(t)$ and $v_j(t)$ are approximations to certain functions $u(x,t)$ and $v(x,t)$,

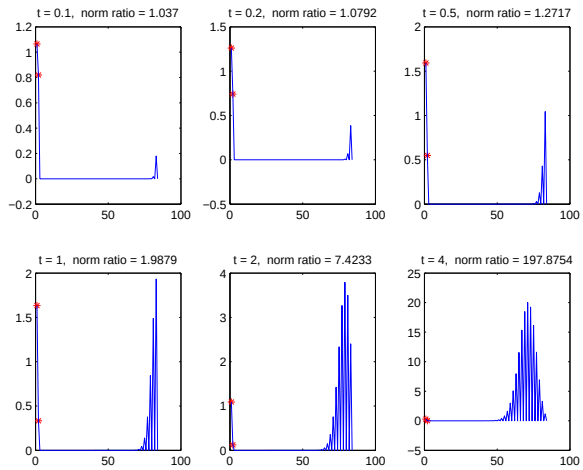

FIG. 1. *Behavior of* $\mathbf{y}(t) = e^{tA}\mathbf{y}(0)$ *(asterisks) and* $\hat{\mathbf{y}}(t) = e^{tM}\hat{\mathbf{y}}(0)$ *(solid line), where* $\mathbf{y}(0) = [u_0, v_0]^T$ *is random and* $\hat{\mathbf{y}}(0) = [u_0, v_0, 0, \ldots, 0]^T$*. A small disturbance develops on the right and soon overshadows the action on the left. (Note change in vertical scale of the plots.)*

respectively, at $x = j$. Then the middle blocks of the dilated equation $\hat{\mathbf{y}}' = M\hat{\mathbf{y}}$ are:

$$\begin{aligned} \frac{du_j}{dt} &= -u_j + 2u_{j+1} = u_j + 2(u_{j+1} - u_j) \approx u_j + 2\left.\frac{\partial u}{\partial x}\right|_{x=j} \\ \frac{dv_j}{dt} &= -v_j + 2v_{j+1} = v_j + 2(v_{j+1} - v_j) \approx v_j + 2\left.\frac{\partial v}{\partial x}\right|_{x=j}. \end{aligned} \tag{5.3}$$

This can be thought of as a method of lines approach using forward differences in space for the partial differential equations $u_t = u + 2u_x$ and $v_t = v + 2v_x$, where u and v are coupled through the boundary conditions at the ends of the domain. The equation $w_t = aw + bw_x$ is the equation of a wave moving to the left with speed b and growing/decaying like e^{at}; its solution is $w(t) = e^{at}w_0(x + bt)$.

In fact, the equations for $u_j(t)$ and $v_j(t)$, $j = 1, \ldots, K$, can be solved exactly, assuming that $u_2(0) = v_2(0) = \ldots = u_K(0) = v_K(0) = 0$. From the formula for M in (5.2), we have

$$v_K' = -v_K - 2u_2 \Rightarrow v_K(t) = e^{-t}v_K(0) - 2\int_0^t e^{-(t-s)}u_2(s)\,ds,$$

which is 0 as long as $u_2(s) = 0$ for $s \le t$. Similarly, for $j = K-1, \ldots, 2$, we have $v_j' = -v_j + 2v_{j+1} \Rightarrow v_j(t) = e^{-t}v_j(0) + 2\int_0^t e^{-(t-s)}v_{j+1}(s)\,ds$, which is 0 as long as $v_{j+1}(s) = 0$ for $s \le t$. Thus, $v_2, \ldots, v_K$ are 0 until the disturbance at the right in Figure 1 travels to block 2, but at this point e^{tM} is no longer a dilation of e^{tA} and we realize that if we wish to study times as large as this then we need to include more blocks in the matrix M of (5.2).

Looking at the equation for u_K, we find

$$u_K' = -u_K + 2u_1 \Rightarrow u_K(t) = e^{-t}u_K(0) + 2\int_0^t e^{-(t-s)}u_1(s)\,ds = 2\int_0^t e^{-(t-s)}u_1(s)\,ds.$$

Since $u_1(t)$ and $v_1(t)$ are the same as for the original problem; that is, $v_1(t) = e^{-t}v_0$ and $u_1(t) = e^{-t}u_0 + 4te^{-t}v_0$, we can substitute the expression for $u_1(s)$ to obtain

$$u_K(t) = 2\int_0^t e^{-(t-s)}[e^{-s}u_0 + 4se^{-s}v_0]\,ds = 2te^{-t}u_0 + 4t^2e^{-t}v_0.$$

The equation for u_{K-1} is

$$u_{K-1}' = -u_{K-1} + 2u_K \Rightarrow u_{K-1}(t) = 2\int_0^t e^{-(t-s)}u_K(s)\,ds,$$

and substituting the expression for $u_K(s)$, this becomes

$$u_{K-1}(t) = e^{-t}\left[2t^2u_0 + \frac{8}{3}t^3v_0\right].$$

Continuing in this way to compute $u_{K-2}, u_{K-3}, \ldots$, we find

$$u_{K-j}(t) = e^{-t}\left[\frac{2^{j+1}}{(j+1)!}t^{j+1}u_0 + \frac{2^{j+3}}{(j+2)!}t^{j+2}v_0\right], \quad j = 0, \ldots, K-2. \tag{5.4}$$

Figure 2 shows a plot just of $u_j(t)$, $j = 1, \ldots, K$ for some later times t, where the wave is seen to grow at about the rate e^t while traveling to the left about 2 spaces for each unit increment in t.

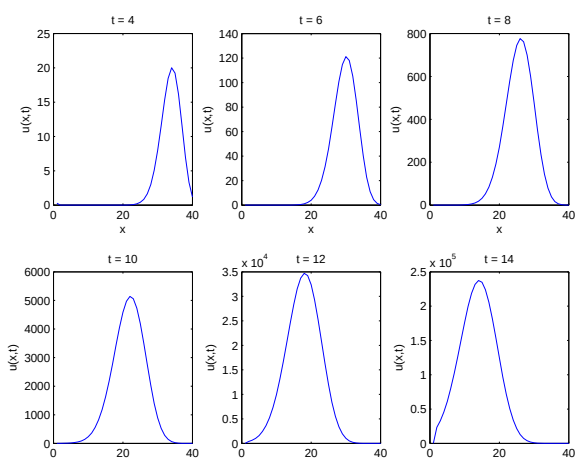

FIG. 2. *Behavior of $u_j(t) \approx u(x,t)|_{x=j}$. The wave grows roughly like e^t as it moves to the left at about speed 2. (Note change in vertical scale of the plots.)*

It is not surprising that the behavior of $\|e^{tM}\|$ differs from that of $\|e^{tA}\|$. The eigenvalues of M are shown in Figure 3, where it can be seen that they are spread all around $\partial W(A)$; thus the spectral abscissa of M is indeed close to $\alpha(A) = 1$, and $\|e^{tM}\|$ will be approximately between $\frac{1}{2}e^t$ and $2e^t$ for all t, since an eigenvector matrix of M has condition number 2. On the other hand $\|e^{tA}\|$ must eventually decrease with t. What *is* perhaps surprising is just how little the two have to do with each other. It is mainly the middle blocks of (5.2) that determine the behavior of $\|e^{tM}\|$ or of $e^{tM}\hat{\mathbf{y}}(0)$. These were generated by first inserting blocks into a unitary dilation of a contraction C to which $g(A)$ was 2-similar (where $g(z) = \frac{1}{2}(z+1)$ was a conformal mapping from $W(A)$ to $\mathcal{D}$) in order to make powers of the unitary matrix dilations of the corresponding powers of C, and then applying $g^{-1}(z) = 2z - 1$ to this unitary matrix. The result was the pair of equations (5.3), which do not directly involve the matrix A at all but are simply determined by the mapping g^{-1} from $\mathcal{D}$ to $W(A)$.

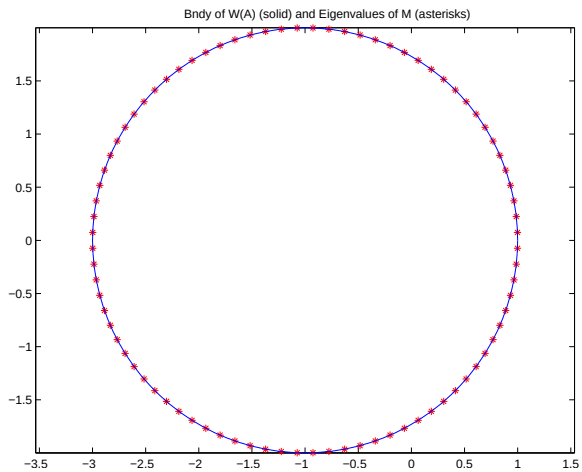

FIG. 3. *$W(A)$ and eigenvalues of M. Eigenvalues of M are all around $\partial W(A)$, so the spectral abscissa of M is approximately equal to the numerical abscissa of A.*

This problem corresponds to a very special matrix A whose field of values is a disk. It is for this reason that the structure of its dilation M in (5.2) is so simple. If $W(A)$ is not a disk, then the mapping g^{-1} from $\mathcal{D}$ to $W(A)$ will be more complicated, and the matrix $g^{-1}(U)$ or $f_k^{-1}(U_K)$, where U and U_K are the unitary matrices defined in (2.1) and (2.3), will not have such a simple structure. One might ask whether the same sort of phenomenon as seen in Figures 1 and 2 would occur in this case.

From our experiments, the answer appears to be "yes". For example, Figure 4 shows the eigenvalues and field of values of an 8 by 8 Chebyshev spectral approximation to an advection-diffusion operator, $Ly = \eta y_{xx} + y_x$, $y(0) = y(1) = 0$, where $\eta = 1/30$. This example can be found in [18, p. 410-411]. The $\circ$'s in the plot are the eigenvalues of a dilation M constructed using the numerical technique described

in the previous section. Because we used a finite matrix U_K (with $K = 100$ blocks) and constructed a polynomial approximation f_k (of degree $k = 18$) to the mapping g^{-1} from $\mathcal{D}$ to $W(A)$, the eigenvalues of M do not lie *exactly* on $\partial W(A)$, but the difference is barely detectable in the plot. The solid line parabola in the figure that nearly matches the boundary of $W(A)$ on the right but becomes wider to the left is the boundary of the numerical range of the differential operator: Its equation is $\mathrm{Re}(z) = -\eta(\mathrm{Im}(z))^2 - \eta\pi^2$ [18, p. 121]. While this is a very coarse matrix approximation, the numerical abscissa of the matrix matches that of the differential operator to 4 decimal places; it is about -0.3290. The eigenvalues of the matrix are quite different from those of the differential operator, but the spectral abscissa of A is -4.2. This means that aymptotically $\|e^{tA}\|$ will decrease like $e^{-4.2t}$, while for all time $\|e^{tM}\|$ will decrease like $e^{-0.3290t}$.

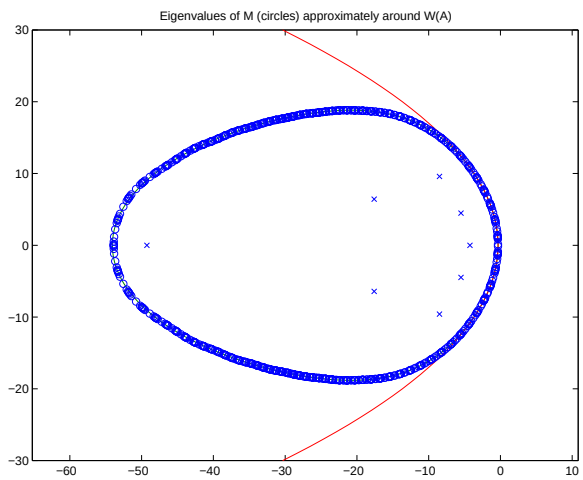

FIG. 4. *$W(A)$ and eigenvalues of M for a Chebyshev spectral approximation to $Lu = \eta u_{xx} + u_x$, $u(0) = u(1) = 0$, where $\eta = 1/30$. Eigenvalues of M ($\circ$'s) are approximately around $\partial W(A)$. Eigenvalues of A are marked with $\times$'s. Parabola that matches $W(A)$ on the right and becomes wider to the left is boundary of numerical range of the differential operator.*

Figure 5 shows the behavior of $e^{tA}\mathbf{y_0}$ (asterisks) and of $e^{tM}\hat{\mathbf{y}}_\mathbf{0}$ (with real and imaginary parts each plotted as solid lines), where $y(x, 0)$ was taken to be the function $x(1-x)(x+2)$, and the vector $\mathbf{y_0}$ consisted of the values of $y(x, 0)$ at the Chebyshev points shifted to the interval $[0, 1]$, while $\hat{\mathbf{y}}_\mathbf{0}$ consisted of these values appended with 0's.

The difference in behavior of matrix powers mimics that of the matrix exponentials. If the equations $\mathbf{y}' = A\mathbf{y}$ and $\hat{\mathbf{y}}' = M\hat{\mathbf{y}}$ are differenced with a backward Euler method, for instance, then the resulting systems of equations take the forms $\mathbf{y}^{(m+1)} = (I - (\Delta t)A)^{-1}\mathbf{y}^{(m)}$ and $\hat{\mathbf{y}}^{(m+1)} = (I - (\Delta t)M)^{-1}\hat{\mathbf{y}}^{(m)}$. For $\Delta t = 0.05$, the matrix powers $(I - (\Delta t)A)^{-k}$ and $(I - (\Delta t)M)^{-k}$ both decrease in norm monotonically with k, with the norm ratios being similar to those of $\|e^{tM}\|/\|e^{tA}\|$ when

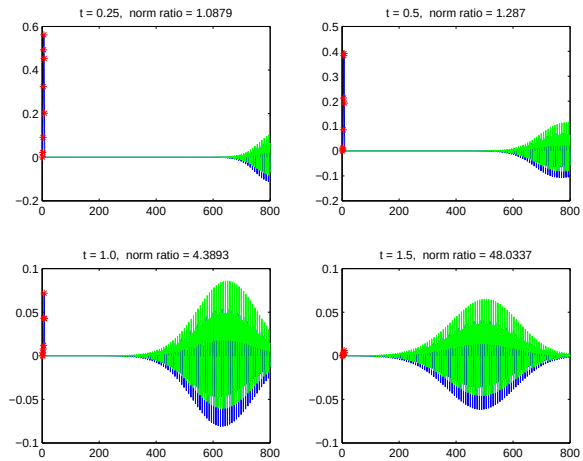

FIG. 5. *Behavior of* $\mathbf{y}(t) = e^{tA}\mathbf{y_0}$ *(asterisks) and* $\hat{\mathbf{y}}(t) = e^{tM}\hat{\mathbf{y}}_\mathbf{0}$ *(solid lines for real and imaginary parts), where* $\mathbf{y_0}$ *consists of values of* $y(x,0) = x(1-x)(x+2)$ *at Chebyshev points and* $\hat{\mathbf{y}}_\mathbf{0} = [\mathbf{y_0}^T, 0, \ldots, 0]^T$. *A small disturbance develops on the right, and although its magnitude is decreasing over time (note change in vertical scale of the plots), it soon overshadows the action on the left, whose magnitude eventually decreases at a much faster rate.*

$t = k\Delta t$. Figure 6 shows $(I - 0.05A)^{-10}\mathbf{y}^{(\mathbf{0})}$ (asterisks) and $(I - 0.05M)^{-10}\hat{\mathbf{y}}^{(\mathbf{0})}$ (o's for first 8 entries, which match those of $(I - 0.05A)^{-10}\mathbf{y}^{(\mathbf{0})}$, a line through real parts and one through imaginary parts of every tenth entry of the remaining components) when $\mathbf{y}^{(\mathbf{0})}$ and $\hat{\mathbf{y}}^{(\mathbf{0})}$ are the same as above, except that here we used a matrix M with 200 blocks. Note that since we used a polynomial f_k of degree 18 to approximate the mapping g^{-1}, only polynomials of degree 11 or less in M are guaranteed to be dilations of the corresponding polynomials in A (since $11 \cdot 18 < 199$); neither e^{tM} nor $(I - (\Delta t)M)^{-k}$ is such a polynomial, but they can be well-approximated by such polynomials; this is why their upper left blocks match the corresponding functions of A, to a close approximation.

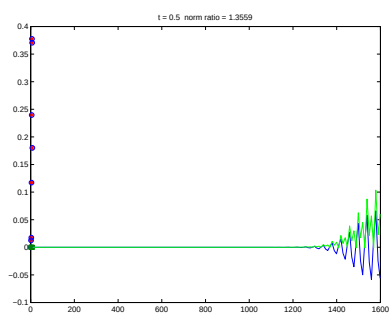

FIG. 6. *Behavior of* $(I-(\Delta t)A)^{-k}\mathbf{y}^{(\mathbf{0})}$ *(asterisks) and* $(I-(\Delta t)M)^{-k}\hat{\mathbf{y}}^{(\mathbf{0})}$ *($\circ$'s for first 8 entries, a line through real parts and one through imaginary parts of every tenth entry of the remaining components),* $\Delta t = 0.05$, $k = 10$.

6. Other Near Normal Dilations. Note that if M is a power dilation of A, then so is $\tilde{M} = \varphi(M)$ when $\varphi(z)$ is of the form $z+\chi_A(z)h(z)$, where χ_A is the minimal polynomial of A and h is any analytic function. Then $\varphi(A) = A$ but the spectrum of $\varphi(M)$ is the image under φ of the spectrum of M. Given a near normal dilation M with spectrum on $\partial W(A)$, one can attempt to choose h to minimize the largest real part of $\varphi(\sigma(M)) \approx \varphi(\partial W(A))$. We tried this for A in (5.1), allowing h to be any polynomial of degree at most 5 and using an optimization code to perform the minimization over the coefficients of h. There is no guarantee that we have found the global minimum, but we were able to move the spectrum of the dilation to the left. The resulting curve $\varphi(\partial W(A))$ is shown in Figure 7, along with the original circle $\partial W(A)$. Although the new curve encloses a larger region than $W(A)$, the rightmost point in that region has been moved left from 1 to about 0.07. Since $\|e^{tA}\| \le \|e^{t\tilde{M}}\|$ and $\tilde{M}$ has an eigenvector matrix with condition number 2, we obtain the improved estimate:

$$\|e^{tA}\| \le \min\{e^t, 2e^{0.07t}\}.$$

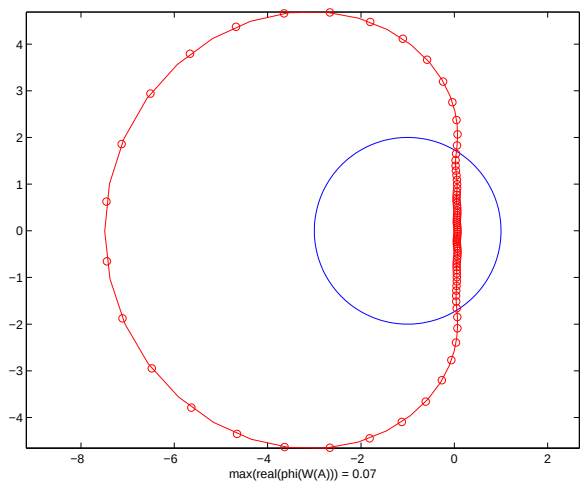

FIG. 7. *$\partial W(A)$ (circle about -1 of radius 2) and $\varphi(\partial W(A))$ (curve enclosing larger region but shifted left) for A in (5.1). The rightmost point has been moved left from 1 to about 0.07. $\circ$'s are eigenvalues of $\tilde{M} = \varphi(M)$, which is another near normal dilation of A.*

It was noted in Section 4 that while it may be computationally difficult to construct a near normal dilation of a given matrix A with spectrum around $\partial W(A)$, it is much easier to construct a dilation with spectrum around a circle enclosing $W(A)$. Consider again the advection-diffusion problem discussed in the previous section. There we used a Chebyshev discretization matrix of size $n = 8$ because, with larger n, the optimization problem of minimizing $\kappa(X)$ subject to $\|X^{-1}g(A)X\| \le 1$ becomes more difficult, as does the problem of mapping between $W(A)$ and the unit

disk and approximating the inverse mapping by a polynomial, since $W(A)$ becomes much longer and narrower (relative to its length) as n increases. But if we instead find a dilation of A with spectrum around a circle enclosing $W(A)$, then the mappings g and g^{-1} between the enclosing disk and the unit disk are simple scalings and translations, and there is a guaranteed algorithm [2] for finding X with $\kappa(X) \leq 2$ and $\|X^{-1}g(A)X\| \leq 1$. To illustrate, we again consider the Chebyshev spectral discretization of the advection-diffusion operator $\eta y_{xx} + y_x$ on $[0,1]$ with $y(0) = y(1) = 0$, $\eta = 1/30$. Now, however, we use $n = 60$ points in the discretization, giving a much more accurate approximation to the differential operator. The eigenvalues and field of values of A are plotted in Figure 8, as is the smallest circle enclosing $W(A)$ and the eigenvalues of a computed near normal dilation of A, when the dilation is chosen so that its spectrum lies on the enclosing circle instead of on $\partial W(A)$. In this case, the condition number of an eigenvector matrix for A was 1.8×10^6, while the condition number of an eigenvector matrix for the dilation M was 1.5. The spectral abscissa of M was almost equal to the numerical abscissa of A, which is -0.3290, so $\|e^{tM}\|$ will decay like $e^{-0.3290t}$, as in the smaller problem of Section 5. The spectral abscissa of A was -7.8, so asymptotically $\|e^{tA}\|$ decays much faster, like $e^{-7.8t}$ but, because A is highly nonnormal, very different things may happen early on.

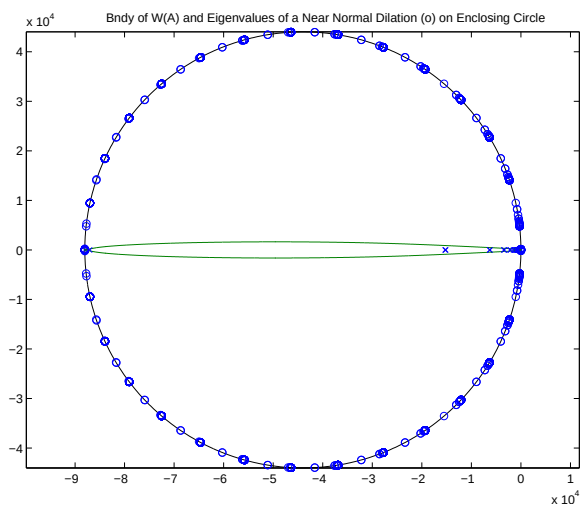

FIG. 8. *$\partial W(A)$ and enclosing circle, when A is a 60 by 60 Chebyshev spectral discretization of $\eta y_{xx} + y_x$ on $[0,1]$ where $y(0) = y(1) = 0$, $\eta = 1/30$. $\times$'s are eigenvalues of A; $\circ$'s are eigenvalues of a near normal dilation of A.*

7. Dilations with Spectra on the Boundary of Other Regions. The following theorem, proved independently by Herrero [8] and Voiculescu [19], can be found, for example, in [15, Theorem 9.13, p. 129]

THEOREM 7.1. **(Herrero-Voiculescu).** *Let T be a bounded linear operator on a Hilbert space $\mathcal{H}$, and let Ω be an open set in the complex plane such that $\sigma(T)$ is*

contained in Ω and such that $\partial\Omega$ consists of a finite number of simple closed rectifiable curves. Then T is similar to an operator that has a normal power dilation with spectrum on $\partial\Omega$, and the similarity may be chosen to satisfy

$$\|S\| \cdot \|S^{-1}\| \leq \frac{1}{2\pi} \int_{\partial\Omega} \|(zI - T)^{-1}\| \; |dz|.$$

The *ϵ-pseudospectrum* is a set in the complex plane that has been extensively studied in connection with the behavior of nonnormal matrices and linear operators [18]. The ϵ-pseudospectrum $\Lambda_\epsilon(A)$ of a matrix or linear operator A is the set of points $z \in \mathbb{C}$ such that z is in the spectrum of $A+E$ for some E with $\|E\| < \epsilon$, and it also can be characterized as the set of points $z \in \mathbb{C}$ for which $\|(zI - A)^{-1}\| > \epsilon^{-1}$. It contains the spectrum of A and its boundary consists of a finite number of simple closed curves. From the above theorem it follows that A is similar, via a similarity transformation with condition number at most $\mathcal{L}(\partial\Lambda_\epsilon(A))/(2\pi\epsilon)$ where $\mathcal{L}(\cdot)$ denotes the length, to an operator that has a normal power dilation with spectrum on $\partial\Lambda_\epsilon(A)$. From this it follows that A itself has a power dilation M with spectrum around $\partial\Lambda_\epsilon(A)$, and such that M is similar to a normal operator via a similarity transformation with condition number at most $\mathcal{L}(\partial\Lambda_\epsilon(A))/(2\pi\epsilon)$. Unfortunately, while this theorem establishes the existence of such a dilation, it does not provide a method for constructing it when the set Λ_ϵ has multiple components. (For ϵ large enough so that Λ_ϵ is simply connected, the procedure of this paper can be used.) It would be interesting to construct such a dilation M corresponding to a multi-component Λ_ϵ and compare the norm behavior of functions of M to that of the same functions of A.

REFERENCES

[1] T. Ando, *Structure of operators with numerical radius one*, Acta. Sci. Mat. (Szeged) 34, pp. 11-15.

[2] D. Choi and A. Greenbaum, *An algorithm for finding a 2-similarity transformation from a numerical contraction to a contraction*, to appear in SIAM J. Matrix Anal. Appl.

[3] M. Crouzeix, *Bounds for analytical functions of matrices*, Integr. Equ. Oper. Theory 48 (2004), pp. 461-477.

[4] M. Crouzeix, *Numerical range and functional calculus in Hilbert space*, J. Functional Analysis 244 (2007), pp. 668-690.

[5] M. Crouzeix, private communication.

[6] T. A. Driscoll, N. Hale, and L. N. Trefethen, editors, Chebfun Guide, Pafnuty Publications, Oxford, 2014.

[7] B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland, New York, 1970.

[8] D. Herrero, *A Rota universal model for operators with multiply connected spectrum*, Rev. Roum. Math. Pures et Appl. 21 (1976), pp. 15-23.

[9] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge, 1991.

[10] T. Kato, *Perturbation Theory for Linear Operators*, 2nd Edition, Springer-Verlag, 1976.

[11] N. Kerzman and E.M. Stein, *The Cauchy kernel, the Szegö kernel, and the Riemann mapping function*, Math. Ann. 236 (1978), pp. 85-93.

[12] N. Kerzman and M. Trummer, *Numerical conformal mapping via the Szegö kernel*, Jour. of Comp. Appl. Math 14 (1986), pp. 111-123.

[13] J. von Neumann, *Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes*, Mach. Nachrichten 4 (1951), pp. 258-281.

[14] K. Okubo and T. Ando, *Constants related to operators of class C_p*, Manuscripta Math, 16 (1975), pp. 385-394.

[15] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, 2002.

[16] J. Schäffer, *On unitary dilations of contractions*, Proc. Amer. Math. Soc. 6 (1955), p. 322

[17] M. Schreiber, *Unitary dilations of operators*, Duke Math. J., 23 (1956), pp. 579-594.

[18] L. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press, 2005.

[19] D. Voiculescu, *Norm-limits of algebraic operators*, Rev. Roum. Math. Pures et Appl. 19 (1974), pp. 371-378.