CARD SHUFFLING AND THE POLYNOMIAL NUMERICAL HULL OF DEGREE K

ANNE GREENBAUM*

Abstract. The *polynomial numerical hull of degree k* is a set in the complex plane associated with a matrix A that is designed to give more information than the spectrum can provide about the behavior of A under the action of polynomials. We give an example of the application of the polynomial numerical hull of degree k to explain the cutoff phenomenon that is often observed in Markov processes. In particular, we consider the much-publicized result that it takes seven riffle shuffles to randomize a deck of cards [Bayer and Diaconis, *Ann. Appl. Prob.* 2, 294–313 (1992)]. While this result cannot be explained by eigenvalues, it is shown to be very well-predicted by the 1-norm polynomial numerical hulls of various degrees associated with a matrix derived from the probability transition matrix.

Key words. polynomial numerical hull, Markov chain, cutoff phenomenon

AMS subject classifications. 15A60, 65F15, 65F35

1. Introduction. Let A be an n by n matrix. The *polynomial numerical hull of A of degree k* was introduced by Nevanlinna [10, 11] and further studied by the author [5]. It is defined as

$$\mathcal{G}_k(A) = \{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \ \forall p \in \mathcal{P}_k\}, \tag{1}$$

where $\mathcal{P}_k$ is the set of polynomials of degree k or less. In many cases, the norm of interest is the 2-norm, but in this paper we will concentrate on the 1-norm polynomial numerical hull.

The polynomial numerical hull of any degree contains the spectrum of A since, if λ is an eigenvalue and v a corresponding normalized eigenvector of A, then $p(A)v = p(\lambda)v$ implies $\|p(A)\| \geq |p(\lambda)|$ for any matrix norm compatible with the given vector norm. For k greater than or equal to the degree of the minimal polynomial of A, the polynomial numerical hull of degree k consists precisely of the spectrum since the minimal polynomial of A has roots only at the eigenvalues but satisfies $\|p(A)\| = 0$. For values of k between one and the degree of the minimal polynomial, the polynomial numerical hull of degree k is a closed bounded set — it is a subset, for instance, of $\{z : \ |z| \leq \|A\|\}$ — that contains the spectrum.

These sets were identified to give more information than the spectrum alone can provide about the behavior of a matrix under the action of polynomials. In this paper we describe an application of the polynomial numerical hull in explaining the cutoff phenomenon that is often observed in Markov processes. In particular, we consider the much-publicized result that it takes seven riffle shuffles to randomize a deck of cards [1, 9]. While the eigenvalues of the probability transition matrix cannot explain this fact, the 1-norm polynomial numerical hulls of various degrees can.

A related set, that was perhaps first introduced by Landau [8] and Varah [13] and later popularized by Trefethen [12, 3], is the *ϵ-pseudospectrum* of A. It is defined as

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} : \ \|(zI - A)^{-1}\| \geq \epsilon^{-1}\}.$$

Again, while one often is interested in the 2-norm pseudospectra, the more relevant norm for the card shuffling problem is the 1-norm. Jónsson and Trefethen considered

[*] University of Washington, Mathematics Dept., Box 354350, Seattle, WA 98195. This work was supported by NSF contract DMS-9802919.

the 1-norm pseudospectra of the probability transition matrix associated with the riffle shuffle and demonstrated that the behavior of this system probably would not be governed by its eigenvalues; even for small values of ϵ, the ϵ-pseudospectrum was shown to contain much more than just the spectrum [6]. In [6], the authors also made the important contribution of presenting the problem in a linear algebra framework that made further analysis much easier, and that is the framework adopted here. Moreover, a MATLAB code was provided for computing the probability transition matrix and that code was used in the computations of this paper.

While examination of the pseudospectra of the transition matrix suggested that there might be a transient phase, the pseudospectral plots did not seem to provide a means by which the entire behavior of the system could be readily deduced; it was not suggested in [6] that one could look at the pseudospectra and easily determine the length of the transient phase or the exact behavior of the system during this phase. The lower bound (1) provided by the polynomial numerical hull enables one to estimate the early behavior of the system and the estimate turns out to be remarkably good in this case. For the card shuffling problem, as we will see in the next section (and as was noted in [6]), the distance from randomness is governed by the 1-norms of the powers of a matrix A that is derived from the probability transition matrix. It follows from (1) that these powers satisfy

$$\|A^k\|_1 \geq \sup\{|z|^k : z \in \mathcal{G}_{k,1}(A)\}, \tag{2}$$

where $\mathcal{G}_{k,1}(A)$ denotes the 1-norm polynomial numerical hull of degree k. Hence this distance can drop below one only when $\mathcal{G}_{k,1}(A)$ lies strictly inside the unit disk. As it turns out, this first happens for $k = 7$.

2. The Card Shuffling Problem and Polynomial Numerical Hulls.

A standard riffle shuffle can be modeled in the following way [1]: A deck of n cards is first cut into two heaps with the probability of a heap containing k cards being given by a binomial distribution: $\binom{n}{k}/2^n$. Cards from the left and right heaps are then riffled together into a single pile, with the probability of dropping the next card from a given heap being proportional to the number of cards in that heap. With this model, the process of shuffling cards is represented as a Markov chain. The states of the system are the $n!$ possible orderings of the cards, and, given the current ordering, one can write down the probability of achieving any other ordering on the next shuffle.

On the face of it, this simple model seems computationally intractable for even moderate size values of n since it deals with $n!$ possible states and requires the use of an $n!$ by $n!$ probability transition matrix. Bayer and Diaconis showed that the number of states of the system could be drastically reduced, however, by considering only the number of *rising sequences* in any particular ordering [1]. A rising sequence is a maximal list of consecutively numbered cards (perhaps interspersed with other cards) found during one pass through the deck. For example, the ordering 146235 has three rising sequences: $\{1, 2, 3\}$, $\{4, 5\}$, and $\{6\}$. By counting only the number of rising sequences and deriving the probabilities of moving from a given number of rising sequences to any other number with a single shuffle, the possible states of the system were reduced from $n!$ to n. Now the problem becomes easy to handle numerically.

The entries of the n by n probability transition matrix P are given by [6]:

$$P_{ji} = 2^{-n} \binom{n+1}{2i-j} \frac{\alpha_j}{\alpha_i},$$

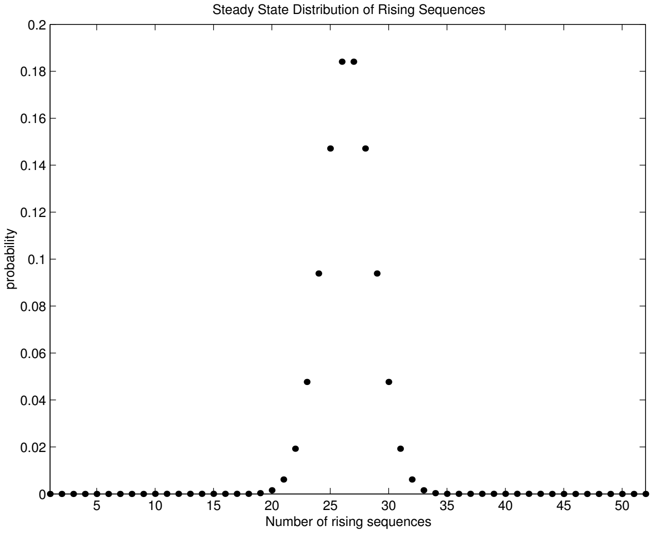

FIG. 1. *Steady-state distribution of rising sequences.*

where α_j denotes the number of permutations of $\{1, \ldots, n\}$ that have exactly j rising sequences. The numbers α_j are called *Eulerian numbers* and satisfy the recurrence

$$A_{r,k} = kA_{r-1,k} + (r-k+1)A_{r-1,k-1},$$

where $A_{11} = 1$, $A_{1k} = 0$ for $k \neq 1$, and $\alpha_j = A_{nj}$ [4, 7]. Here we have defined the probability transition matrix P so that a shuffle of the deck corresponds to applying P on the left to the current state vector.

The eigenvalues of P are known to be 2^{-j}, $j = 0, 1, \ldots, n-1$. The largest eigenvalue is 1 and the corresponding eigenvector,

$$v = (\alpha_1, \ldots, \alpha_n)^T / n!,$$

represents the steady-state to which the system converges after infinitely many shuffles. Figure 1 shows a plot of the steady-state distribution of rising sequences for a problem of size $n = 52$. It can be noted, for example, that with probability .9999 a well-shuffled deck has between 19 and 34 rising sequences.

The distance from randomness (the steady-state) after k shuffles is measured by the difference between $P^k s_0$ and $v = P^\infty s_0$, where s_0 is the initial state vector and $P^\infty = \lim_{j \to \infty} P^j = (v, \ldots, v)$. For an initially sorted deck, there is one rising sequence and $s_0 = (1, 0, \ldots, 0)^T$. If we define $A = P - P^\infty$, then $A^k = P^k - P^\infty$, and the distance from randomness after $k \geq 1$ shuffles is the norm of $A^k s_0$. An upper bound on this distance, which turns out to be exact [6] when the appropriate norm is used, is the norm of A^k.

The *asymptotic rate* of convergence toward the steady-state is determined by the spectral radius of A, which is the second largest eigenvalue of P: $1/2$. This means that after sufficiently many initial shuffles, each new shuffle will reduce the distance from the steady-state by about a factor of 2. At what point this asymptotic convergence rate starts to be achieved, however, *cannot* be determined from the eigenvalues, nor can the behavior of the system before this point.

The 1-norm of the difference between $P^k s_0$ and $v = P^\infty s_0$ is the sum of absolute values of the differences between the probability of each number of rising sequences

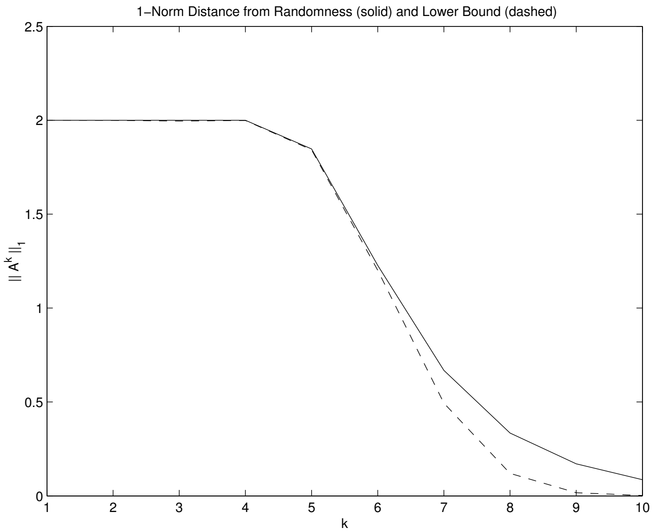

FIG. 2. *1-Norm distance from randomness (solid) and lower bound based on polynomial numerical hulls of various degrees (dashed).*

after k shuffles and that after infinitely many shuffles. This number is at most 2, and for $k = 0$ it is very nearly equal to 2 since the probability of having only one rising sequence after infinitely many shuffles is tiny; that is, $s_0 - v \approx (1, -v_2, \ldots, -v_n)^T$, where $v_2 + \ldots + v_n \approx 1$. If this number is large (e.g., greater than 1) then one probably can tell by counting the number of rising sequences after k shuffles that the deck has not been thoroughly mixed. The 1-norm is easily seen to be equal to twice the *total variation norm* used by Bayer and Diaconis [2].

The solid line in Figure 2 is a plot of $\|A^k\|_1$ versus k, again for the problem of size $n = 52$. After about $k = 6$ shuffles the asymptotic behavior of the system sets in, with $\|A^k\|_1$ being reduced by about a factor of two each time k is increased by one. Before this point, however, the behavior of the system is quite different. It can be seen from Figure 2 that $\|A^k\|_1$ remains almost constant until $k = 5$, where it begins to decrease. There are large (absolute) decreases at steps 6 and 7, and $\|A^k\|_1$ first drops below one at step $k = 7$.

The early behavior of the system is not hard to understand if one thinks about rising sequences and the steady-state distribution of Figure 1. Initially the deck has only one rising sequence and after one shuffle it has at most two. Since the probability of having only two rising sequences after many shuffles is extremely small, the 1-norm distance from the steady-state vector remains almost constant. After j shuffles, there are at most 2^j rising sequences so for $j \leq 4$ the number of rising sequences is less than or equal to 16. This is still highly unlikely in a well-shuffled deck, so the 1-norm distance to randomness does not decrease significantly.

How can one see this from properties of the matrix A? The eigenvalues give no hint that there is a (fairly) sudden cutoff. The 1-norm polynomial numerical hulls of various degrees do provide such insight, however. The 1-norm polynomial numerical hulls of degrees 1 through 8 are sketched in Figures 3 and 4, and Table 1 lists the values

$$\nu_k = \sup\{|z| : \ z \in \mathcal{G}_{k,1}(A)\}, \quad k = 1, \ldots, 10,$$

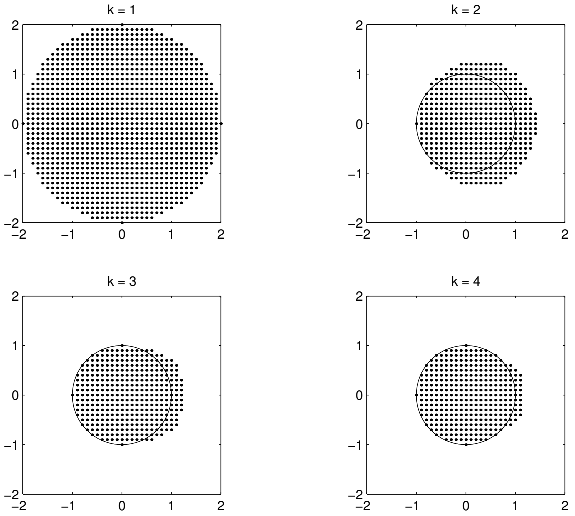

FIG. 3. *Polynomial numerical hulls of degrees $k = 1, 2, 3, 4$ for the card shuffling matrix. Points in the polynomial numerical hull of the given degree are marked with dots, and the unit circle is plotted for comparison in the $k = 2$, 3, and 4 figures.*

computed to three decimal places. Based on these numbers and using inequality (2) one obtains the third column of Table 1 and the dashed curve in Figure 2 as a lower bound on $\|A^k\|_1$. Note how closely the lower bound predicts the actual behavior of the system.

k	ν_k	ν_k^k	$\|A^k\|_1$
1	2.000	2.00	2.00
2	1.414	2.00	2.00
3	1.259	2.00	2.00
4	1.189	2.00	2.00
5	1.130	1.84	1.85
6	1.031	1.20	1.23
7	.904	.49	.67
8	.767	.12	.33
9	.635	1.7e-2	.17
10	.539	2.1e-3	8.6e-2

TABLE 1
ν_k versus k for the card shuffling matrix.

It should be noted that the computation of polynomial numerical hulls of various degrees presently is not an easy or exact task. A definition equivalent to (1) is

$$\mathcal{G}_k(A) = \{\zeta \in \mathbf{C} : \min_{p \in \mathcal{P}_k(0)} \|p(A - \zeta I)\| \geq 1\}, \tag{3}$$

where $\mathcal{P}_k(0)$ denotes the set of polynomials of degree k or less with value one at the origin. To see that (3) is equivalent to (1), note that $\zeta \in \mathcal{G}_k(A)$ (defined in (1)) if and only if $\|p(A)\| \geq |p(\zeta)|$ for all $p \in \mathcal{P}_k$, if and only if $\|q(A)\| \geq 1$ for all q of the form $q(z) = p(z)/p(\zeta)$, where $p \in \mathcal{P}_k$ and $p(\zeta) \neq 0$ (i.e., for all $q \in \mathcal{P}_k$ with $q(\zeta) = 1$), and

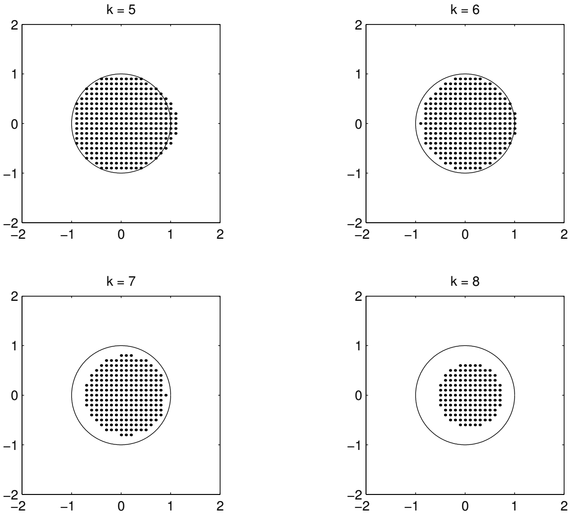

FIG. 4. *Polynomial numerical hulls of degrees $k = 5, 6, 7, 8$ for the card shuffling matrix. Points in the polynomial numerical hull of the given degree are marked with dots, and the unit circle is plotted for comparison. The first value of k for which $\nu_k < 1$ is $k = 7$: $\nu_7 = .904$.*

this is equivalent to $\|q(A)\| = \|I + c_1(A - \zeta I) + c_2(A - \zeta I)^2 + \ldots + c_k(A - \zeta I)^k\| \geq 1$ for all $c_1, \ldots, c_k$. This is the condition that ζ lie in the set defined in (3).

To compute the sets in Figures 3 and 4 and the values in Table 1, we tested many points ζ throughout a region known to contain $\mathcal{G}_{k,1}(A)$; for $k = 1$ we considered points in the disk of radius $\|A\|_1$ about the origin, and for $k > 1$ we tested points in $\mathcal{G}_{k-1,1}(A)$. For each point ζ on a lattice of spacing 0.1 throughout this region, an optimization code (initially FMINUNC in MATLAB) was run to determine the polynomial $p \in \mathcal{P}_k(0)$ for which $\|p(A - \zeta I)\|_1$ was minimal. When ζ is real, this optimization problem can be expressed as a linear programming problem: Find coefficients $c_1, \ldots, c_k$, an n by n matrix Z with nonnegative entries, and a number γ to satisfy

$$\min_{c_1, \ldots, c_k, Z} \gamma \quad \text{subject to}$$

$$-Z_{ij} \leq [I + \sum_{\ell=1}^{k} c_\ell (A - \zeta I)^\ell]_{ij} \leq Z_{ij}, \quad i, j = 1, \ldots, n,$$

$$\sum_{i=1}^{n} Z_{ij} \leq \gamma, \quad j = 1, \ldots, n.$$

This formulation (actually a modification of it, in which the matrices $(A - \zeta I)^\ell$, $\ell = 1, \ldots, k$ were replaced by a basis orthonormal in the Frobenius norm) was used as a check on the original optimization procedure, as was the use of different initial guesses for the solution. If a polynomial $p \in \mathcal{P}_k(0)$ was found for which $\|p(A - \zeta I)\|_1$ was less than one, then ζ was determined to lie outside $\mathcal{G}_{k,1}(A)$. If not, then several checks of the sort described above were run, to see if a different method or initial guess could lead to a polynomial with $\|p(A - \zeta I)\|_1 < 1$. In all cases, the only questionable points were those very very close to the boundary of the region, where the computed minimal value of $\|P(A-\zeta I)\|_1$ was very close to one, and a judgment was made as to whether the slight difference was due to roundoff or was real. If no polynomial could be found to make $\|p(A-\zeta I)\|_1 < 1$, then it was decided that $\zeta \in \mathcal{G}_{k,1}(A)$. To obtain the values in Table 1, we honed in on the largest value to an accuracy of .001 by using a bisection procedure. It is hoped that methods can be developed for computing polynomial numerical hulls of various degrees without simply testing many, many points.

It may seem surprising that the lower bound (2) provides such a good estimate of the actual behavior of this system at the early stages; there is no guarantee that this will always happen. It can be understood why the estimate works so well for $k = 1, \ldots, 4$ by noting that the vectors $s_0, Ps_0, \ldots, P^4 s_0$, and v each have their nonnegligible entries in different positions: if e_j denotes the jth unit vector, then $s_0 = e_1$, $Ps_0 \approx e_2$, $P^k s_0$ is an approximate linear combination of $e_{2^{k-1}+1}, \ldots, e_{2^k}$ for $k = 2, 3, 4$, and v is an approximate linear combination of $e_{19}, \ldots, e_{34}$. It follows that for any coefficients $c_0, c_1, \ldots, c_k$ we have

$$\|\sum_{j=0}^{k} c_j A^j\|_1 = \|\sum_{j=0}^{k} c_j A^j s_0\|_1 = \|\sum_{j=0}^{k} c_j P^j s_0 - (\sum_{j=1}^{k} c_j) v\|_1$$

$$\approx |c_0| + \sum_{j=1}^{k} |c_j| + |\sum_{j=1}^{k} c_j|. \tag{4}$$

The fact that the inequality $\|\sum_{j=0}^{k} c_j A^j\|_1 \geq \|\sum_{j=0}^{k} c_j A^j s_0\|_1$ is actually an equality as in (4) is established in [6].

Taking $k = 1$ in (4), it is seen that the condition that z be in $\mathcal{G}_{1,1}(A)$ defined in (1) is approximately the condition that $|c_0| + 2|c_1|$ be greater than or equal to $|c_0 + c_1 z|$ for all c_0, c_1. This clearly holds if and only if $|z| \leq 2$ and so $\mathcal{G}_{1,1}(A)$ is very close to a disk of radius 2 about the origin.

For k between 2 and 4, $\mathcal{G}_{k,1}(A)$ is not simply a disk about the origin, but one can argue using (4) that $z = 2^{1/k}$ lies in or very close to $\mathcal{G}_{k,1}(A)$; that is,

$$|c_0| + \sum_{j=1}^{k} |c_j| + |\sum_{j=1}^{k} c_j| \geq |\sum_{j=0}^{k} c_j 2^{j/k}|, \quad \forall\ c_0, c_1, \ldots, c_k. \tag{5}$$

To see that (5) holds, note that the left-hand side satisfies

$$\begin{aligned} |c_0| + \sum_{j=1}^{k} |c_j| + |\sum_{j=1}^{k} c_j| &= |c_0| + (2 - 2^{1/k}) \sum_{j=1}^{k} |c_j| + (2^{1/k} - 1) \sum_{j=1}^{k} |c_j| + |\sum_{j=1}^{k} c_j| \\ &\geq |c_0| + (2 - 2^{1/k}) \sum_{j=1}^{k} |c_j| + 2^{1/k} |\sum_{j=1}^{k} c_j|, \end{aligned}$$

while the right-hand side satisfies

$$|\sum_{j=0}^{k} c_j 2^{j/k}| \leq |c_0| + 2^{1/k} |\sum_{j=1}^{k} c_j 2^{(j-1)/k}|$$

$$\leq |c_0| + 2^{1/k} |\sum_{j=1}^{k} c_j| + |\sum_{j=1}^{k} c_j (2^{j/k} - 2^{1/k})|$$

$$\leq |c_0| + 2^{1/k} |\sum_{j=1}^{k} c_j| + (2 - 2^{1/k}) \sum_{j=1}^{k} |c_j|.$$

Combining this with the fact that $\nu_k^k \leq \|A^k\|_1 < 2$, for $k = 1, \ldots, 4$, it is established analytically that $\nu_k^k \approx 2$, for $k = 1, \ldots, 4$.

A number of other Markov processes (e.g., the Ehrenfests' urn [2, 6, 5]) go through a similar initial phase in which the 1-norm distance from the steady-state remains almost constant. For these problems as well, the lower bound (2) should provide a good estimate of the early behavior of the system.

3. Conclusions. While much work remains to be done to efficiently compute polynomial numerical hulls of various degrees and to understand their geometry and their relation to other sets such as the ϵ-pseudospectrum, it is clear that for this particular example the polynomial numerical hull of degree k is a valuable tool in understanding and predicting the behavior of the system. There are many other applications remaining to be explored; examples include stability analysis of finite difference schemes for time-dependent ordinary and partial differential equations and estimates of the convergence rate of the GMRES algorithm for solving linear systems. The polynomial numerical hulls of various degrees appear to provide important information about the behavior of a matrix that cannot be gleaned from the spectrum alone and so are deserving of further study.

Acknowledgments: The author thanks Nick Trefethen and Mark Embree for helpful comments on a draft and the first version of this paper.

REFERENCES

[1] D. Bayer and P. Diaconis, *Trailing the dovetail shuffle to its lair*, Ann. Appl. Prob., 2 (1992), pp. 294–313.

[2] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Natl. Acad. Sci., 93 (1996), pp. 1659–1664.

[3] M. Embree and L. N. Trefethen, *Pseudospectra Gateway*, www.comlab.ox.ac.uk/pseudospectra.

[4] R. Graham, D. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, MA., 1989.

[5] A. Greenbaum, *Generalizations of the field of values useful in the study of polynomial functions of a matrix*, Lin. Alg. Appl. 347 (2002), pp. 233–249.

[6] G. Jónsson and L. Trefethen, *A numerical analyst looks at the 'cutoff phenomenon' in card shuffling and other Markov chains*, in D. Griffiths, D. Higham, and G. Watson, eds., *Numerical Analysis 1997*, Addison Wesley Longman Ltd., 1998.

[7] D. Knuth, *The Art of Computer Programming, v. 3: Sorting and Searching*, Addison-Wesley, Reading, MA., 1973.

[8] H. J. Landau, *On Szegö's eigenvalue distribution theory and non-Hermitian kernels*, J. d'Analyse Math. 28 (1975), pp. 335–357.

[9] G. Kolata, *In shuffling cards, 7 is winning number*, NY Times, Sec. C, p. 1, Jan. 9, 1990.

[10] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

[11] O. Nevanlinna, *Hessenberg matrices in Krylov subspaces and the computation of the spectrum*, Numer. Funct. Anal. and Optimiz., 16 (1995), pp. 443–473.

[12] L. N. Trefethen, *Approximation theory and numerical linear algebra*, in Algorithms for Approximation II, J. Mason and M. Cox, eds., Chapman and Hall, London, 1990.

[13] J. M. Varah, *On the separation of two matrices*, SIAM J. Numer. Anal. 16 (1979), pp. 216–222.