BOHEMIAN UPPER HESSENBERG MATRICES

EUNICE Y. S. CHAN[*], ROBERT M. CORLESS[*], LAUREANO GONZALEZ-VEGA[†],
J. RAFAEL SENDRA[‡], JUANA SENDRA[§], AND STEVEN E. THORNTON[*]

Abstract. We look at Bohemian matrices, specifically those with entries from $\{-1, 0, +1\}$. More, we specialize the matrices to be upper Hessenberg, with subdiagonal entries ± 1. Many properties remain after these specializations, some of which surprised us. We find two recursive formulae for the characteristic polynomials of upper Hessenberg matrices. Focusing on only those matrices whose characteristic polynomials have maximal height allows us to explicitly identify these polynomials and give a lower bound on their height. This bound is exponential in the order of the matrix. We count *stable* matrices, normal matrices, and neutral matrices, and tabulate the results of our experiments. We prove a theorem about the only possible kinds of normal matrices amongst a specific family of Bohemian upper Hessenberg matrices.

1. Introduction. A matrix family is called **Bohemian** if its entries come from a fixed finite discrete (and hence bounded) set, usually integers. The name is a mnemonic for **Bo**unded **He**ight **M**atrix of **I**ntegers. Such populations arise in many applications (e.g. compressed sensing) and the properties of matrices selected "at random" from such families are of practical and mathematical interest. For example, Tao and Vu have shown that random matrices (more specifically real symmetric random matrices in which the upper-triangular entries $\xi_{i,j}$, $i < j$ and diagonal entries $\xi_{i,i}$ are independent) have simple spectrum [23]. An overview of some of our original interest in Bohemian matrices can be found in [16].

Bohemian families have been studied for a long time, although not under that name. For instance, Olga Taussky-Todd's paper "Matrices of Rational Integers" [24] begins by saying

> "This subject is very vast and very old. It includes all of the arithmetic theory of quadratic forms, as well as many of other classical subjects, such as latin squares and matrices with elements $+1$ or -1 which enter into Euler's, Sylvester's or Hadamard's famous conjectures."

The paper [19] by C. W. Gear is another instance. What is new here is the idea that these families are themselves interesting objects of study, and susceptible to brute-force computational experiments as well as to asymptotic analysis. These experiments have generated many conjectures, some of which we resolve in this paper. Others remain unsolved, and are listed on the Characteristic Polynomial Database [25]. Many of the conjectures have a number-theoretic or combinatorial flavour.

Typical computational puzzles arise on asking simple-looking questions such as "how many 6×6 matrices with the population[1] $\{-1, 0, +1\}$ are singular." The answer is not known as we write this, although we can give a probabilistic estimate (0.205 after 20,000,000 sample determinants[2]): brute computation seems futile because there are $3^{36} \doteq 1.7 \times 10^{17}$ such matrices. We do know the answers up to size five by five:

[*]Department of Applied Mathematics, Western University (echan295@uwo.ca, rcorless@uwo.ca, sthornt7@uwo.ca).

[†]Departamento de Matematicas, Estadistica y Computacion, Universidad de Cantabria (laureano.gonzalez@unican.es).

[‡]Research Group ASYNACS, Departamento de Física y Matemáticas, University of Alcalá (rafael.sendra@uah.es).

[§]Universidad Politécnica de Madrid (jsendra@etsist.upm.es).

[1]The population of a Bohemian family is the set of permissible entries.

[2]4103732 singular matrices out of twenty million sampled.

The number of n by n singular matrices with population $\{-1, 0, +1\}$ is, for $n = 1$, 2, 3, 4, and 5, just 1, 33, 7,875, 15,099,201, and 237,634,987,683. This represents fractions of their numbers (3^{n^2}) of 0.333, 0.407, 0.400, 0.351, and 0.280, respectively.

Yet such matrix families are both useful and interesting. For instance, one may use discrete optimization over a family to look for improved growth factor bounds [20]. Matrices with the population $\{-1, 0, +1\}$ have minimal height[3] over all integer matrices; finding a matrix in this family which has a given polynomial $p(\lambda) \in \mathbb{Z}[\lambda]$ as characteristic polynomial identifies a so-called "minimal height companion matrix", which may confer numerical benefits.

Recently the study of eigenvalues of structured Bohemian matrices (e.g. tridiagonal, complex symmetric) has been undertaken and several puzzling features are seen resulting from extensive experimental computations. For instance, some of the images at bohemianmatrices.com/gallery show common features including "holes".

Different matrix structures produce remarkably different pictures. One structure useful in eigenvalue computation is the upper Hessenberg matrix, which means a matrix $\mathbf{H}$ such that $h_{i,j} = 0$ if $i > j + 1$. These arise naturally in eigenvalue computation because the QR iteration is cheaper for matrices in Hessenberg form. Results on the determinants of Hessenberg matrices can be found in [21].

Remark 1.1. *on computing eigenvalues by first computing characteristic polynomials.* Numerical analysts are familiar with the superior numerical stability of computing eigenvalues iteratively, usually by the QR algorithm or some variant, rather than first computing characteristic polynomials and then finding roots. As is well-known, such an algorithm is *numerically unstable* because polynomials are usually badly-conditioned while eigenvalues are usually well-conditioned[4]. Somewhat surprisingly, for several families of Bohemian matrices, characteristic polynomials become valuable again: first because the matrix dimensions are typically small or at most moderate, the ill-conditioning does not matter much, and second because for some families (not all!) the number of distinct characteristic polynomials is vastly smaller than the number of matrices in the family. For instance, for the general five by five matrices with population $\{-1, 0, +1\}$, there are nearly one trillion such matrices, but fewer than two million characteristic polynomials. This compression is significant.

For other families of matrices, such as upper Hessenberg Toeplitz matrices, there is no compression at all because each matrix has a distinct characteristic polynomial. Circulant matrices fall between, having fewer characteristic polynomials but not vastly fewer. The lesson is that for some questions (though not others), prior computation of characteristic polynomials is valuable.

We begin our study in this paper by considering determinants of Bohemian upper Hessenberg matrices. We prove two recursive formulae for the characteristic polynomials of upper Hessenberg matrices[5]. For another recursive formula we refer to [17]. During the course of our computations, we encountered "maximal polynomial height" characteristic polynomials when the matrices were not only upper Hessenberg, but Toeplitz ($h_{i,j}$ constant along diagonals $j - i = k$); we have several results for such matrices, which will appear in [11]. Further restrictions to this class allowed identification of key results including explicit formulae for the characteristic polynomials

---

[3]height$(A) := ||\text{vec}(A)||_\infty$ is the largest absolute value of any entry in A.

[4]This has been well-known to the point of folklore since the work of Wilkinson. The well-conditioning of eigenvalues has only recently been quantified in some cases, but for instance the results of [3] do confirm the folklore.

[5]We do not claim originality; recursion relations for upper Hessenberg determinants are known.
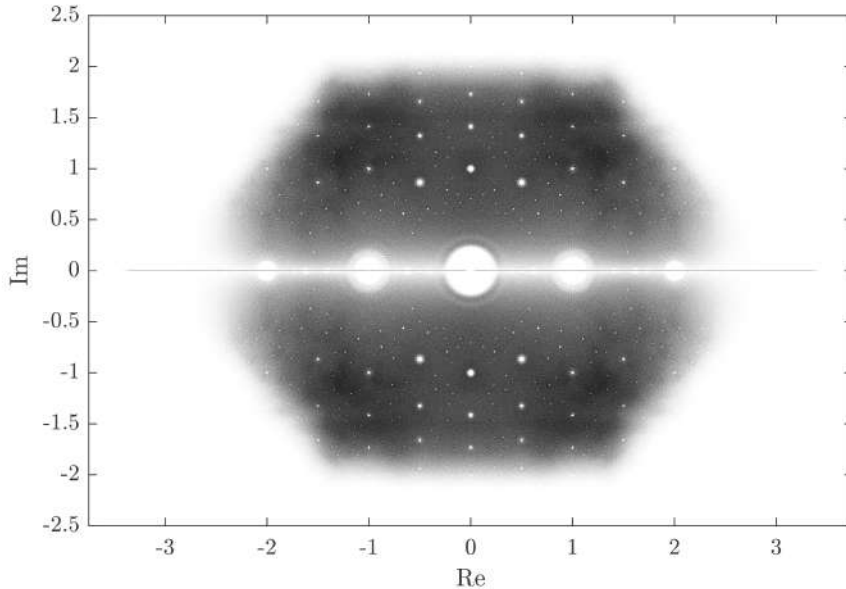
FIG. 1. *The set of eigenvalues of all 10,460,353,203 six by six upper Hessenberg matrices* **H** *with entries* $\mathbf{H}_{i,j} \in \{-1, 0, +1\}$, *and* $\mathbf{H}_{i+1,i} = -1$ *for* $1 \leq i \leq j \leq 6$. *A more detailed image can be found at* assets.bohemianmatrices.com/ gallery/ UH_6x6.png

of maximal height. In what follows, we lay out definitions and prove several facts of interest about characteristic polynomials and their respective height for these families.

In Figure 1 we see all eigenvalues of 6×6 upper Hessenberg matrices, subdiagonals fixed at -1, with population $P = \{-1, 0, +1\}$. We denote this set of matrices $\mathcal{H}^{6\times6}_{\{\pi\}}(P)$. There are $3^{21} = 10{,}460{,}353{,}203$ such matrices. We see a wide octagonal shape. The width of the figure reflects that some matrices might have diagonals -1, while some have diagonals 0, and others have diagonals 1. Of course mixed diagonals are also possible, but this should only tend to push things towards the centre.

This thinking motivates considering the subset of these matrices which has diagonal fixed at 0. We denote this set of matrices $\mathcal{Z}^{6\times6}_{\{\pi\}}(P)$. There are substantially fewer such matrices, only $3^{15} = 14{,}348{,}907$ to be exact, and their eigenvalues are pictured in Figure 2, with certain zones enhanced. We see that, roughly speaking, Figure 1 is partially explained by saying that, along with other eigenvalues, it contains three copies of Figure 2 placed with centres at -1, at 0, and at $+1$.

In this paper we seek to explain some of the features of these pictures, and to learn some things about these families of Bohemian matrices.

2. Prior Work. Visible features of graphs of roots and eigenvalues from structured families of polynomials and matrices have been previously studied. One well-known polynomial whose roots produce interesting pictures is the Littlewood polynomial,
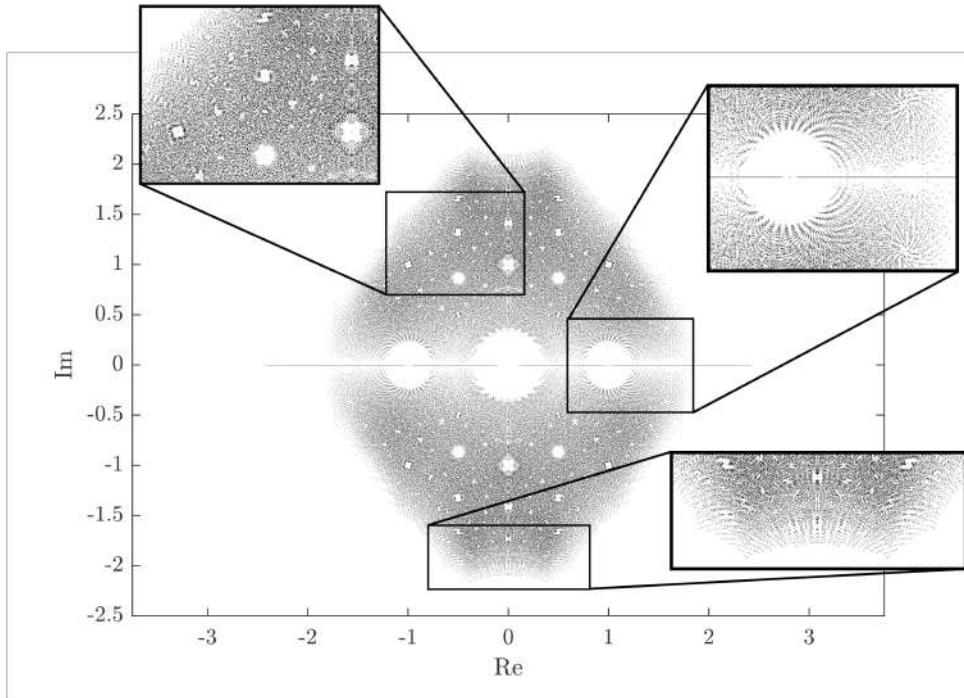
$$(2.1) \qquad\qquad p(x) = \sum_{i=0}^{n} a_i x^i \,,$$

FIG. 2. *The set of eigenvalues of all* 14,348,907 *matrices in* $\mathcal{Z}^{6\times 6}_{\{\pi\}}(\{-1, 0, +1\})$; *that is, six by six upper Hessenberg matrices* $\mathbf{H}$ *with entries* $\mathbf{H}_{i,j} \in \{-1, 0, +1\}$, *diagonal entries fixed as zero, and* $\mathbf{H}_{i+1,i} = -1$ *for* $1 \le i < j \le 6$. *A more detailed image can be found at* assets.bohemianmatrices. com/gallery/UH_0_Diag_6x6.png

where $a_i = \{-1, +1\}$. These polynomials have been studied in [1], [4], [5], and [6]. The image of their roots raises many questions, ranging from whether the set is (ultimately, as $n \to \infty$) a fractal and what the boundary of the set is, to questions about the holes in the image and its connection to various properties, such as degree and coefficients of the polynomial. Answers to some of these questions, particularly the ones involving the holes, have been shown to have some significance in number theory [2]. Roots of other polynomials have also been visualized; for more, see Christensen's[6] and Jörgenson's[7] web pages.

Corless used a generalization of the Littlewood polynomial (to Lagrange bases). In his paper [12], he gave a new kind of companion matrix for polynomials expressed in a Lagrange basis. He used generalized Littlewood polynomials as test problems for the algorithm.

"The Bohemian Eigenvalue Project" was first presented as a poster [15] at the East Coast Computer Algebra Day (ECCAD) 2015. The poster focused on preliminary results and many of the questions raised when visualizing the distributions of Bohemian eigenvalues over the complex plane. In particular, the poster focused on "eigenvalue exclusion zones" (i.e. distinct regions within the domain of the eigenvalues where no eigenvalues exist), computational methods for visualizing eigenvalues,

---

[6]https://jdc.math.uwo.ca/roots/
[7]http://www.cecm.sfu.ca/~loki/Projects/Roots/

and some results on eigenvalue conditioning over distributions of random matrices.

In Chan's Master's thesis [7], she extended Piers W. Lawrence's construction of the companion matrix for the Mandelbrot polynomials [13, 14] to other families of polynomials, mainly the Fibonacci-Mandelbrot polynomials and the Narayana-Mandelbrot polynomials. What is relevant here about this construction is that these matrices are upper Hessenberg and contain entries from a constrained set of numbers: $\{-1, 0\}$, and therefore fall under the category of being Bohemian upper Hessenberg. Both the Fibonacci-Mandelbrot matrices and Narayana-Mandelbrot matrices are also Bohemian upper Hessenberg, but the set that the entries draw from is $\{-1, 0, +1\}$. At the time of submission for Chan's Master's thesis, the largest number of eigenvalues successfully computed (using a machine with 32 GB of memory) were 32,767, 17,710, and 18,559 for the Mandelbrot, Fibonacci-Mandelbrot, and Narayana-Mandelbrot matrices, respectively. This makes the 16th Mandelbrot matrix the "largest" Bohemian matrix that we have solved at the time we write this paper.

These new constructions led Chan and Corless to a new kind of companion matrix for polynomials of the form $c(z) = za(z)b(z) + c_0$. A first step towards this was first proved using the Schur complement in [8]. Knuth then suggested that Chan and Corless look at the Euclid polynomials [9], based on the Euclid numbers. It was the success of this construction that led to the realization that this construction is general, and gives a genuinely new kind of companion matrix. Similar to the previous three families of matrices, the Euclid matrices are also upper Hessenberg and Bohemian, as the entries are comprised from the set $\{-1, 0, +1\}$. In addition, an interesting property of these companion matrices is that their inverses are also Bohemian with the same population, a property which we call "the matrix family having *rhapsody* [10]."

As an extension of this generalization, Chan et al. [10] showed how to construct linearizations of matrix polynomials, particularly of the form $z\mathbf{a}(z)\mathbf{d}_0 + \mathbf{c}_0$, $\mathbf{a}(z)\mathbf{b}(z)$, $\mathbf{a}(z) + \mathbf{b}(z)$ (when $\deg(\mathbf{b}(z)) < \deg(\mathbf{a}(z))$), and $z\mathbf{a}(z)\mathbf{d}_0\mathbf{b}(z) + \mathbf{c}_0$, using a similar construction.

3. Notation. In what follows, we present some results on upper Hessenberg Bohemian matrices of the form

$$(3.1) \qquad \mathbf{H}_n = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,n} \\ s & h_{2,2} & h_{2,3} & \cdots & h_{2,n} \\ 0 & s & h_{3,3} & \cdots & h_{3,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s & h_{n,n} \end{bmatrix}$$

with $s = \exp(i\theta_k)$, usually $s \in \{-1, +1\}$ (we do not allow zero subdiagonal entries, because that reduces the problem to smaller ones) and $h_{i,j} \in \{-1, 0, +1\}$ for $1 \leq i \leq j \leq n$. We denote the characteristic polynomial $Q_n(z) \equiv \det(z\mathbf{I} - \mathbf{H}_n)$.

DEFINITION 3.1. *The set of all $n \times n$ Bohemian upper Hessenberg matrices with upper triangle population P and subdiagonal population from a discrete set of roots of unity, say $s \in \{e^{i\theta_k}\}$ where $\{\theta_k\}$ is some finite set of angles, is called $\mathcal{H}^{n \times n}_{\{\theta_k\}}(P)$. In particular, $\mathcal{H}^{n \times n}_{\{0\}}(P)$ is the set of all $n \times n$ Bohemian upper Hessenberg matrices with upper triangle entries from P and subdiagonal entries equal to 1 and $\mathcal{H}^{n \times n}_{\{\pi\}}(P)$ is when the subdiagonals entries are -1.*

It will often be true that the average value of a population will be zero. In that

case, matrices with trace zero will be common. It is a useful oversimplification to look in that case at matrices whose diagonal is exactly zero.

DEFINITION 3.2. *For a population P such that $0 \in P$, let $\mathcal{Z}^{n \times n}_{\{\theta_k\}}(P)$ be the subset of $\mathcal{H}^{n \times n}_{\{\theta_k\}}(P)$ where the main diagonal entries are fixed at 0.*

4. Results of Experiments. The methods used for computing the characteristic polynomials and counting the number of eigenvalues presented in Tables 1–9 in this section will be discussed in detail in a forthcoming paper. Many of the smaller-dimension computations were done directly in MAPLE 2017; for instance, computation of the characteristic polynomials of all two million or so matrices in $\mathcal{H}^{5 \times 5}_{\{0\}}(\{0, +1\})$ took about six hours on a Surface Pro. The greater number of higher-dimension matrices, or matrices with larger populations, required special techniques and larger & faster machines. Eigenvalue computations were also done in MATLAB and in Python. The computed characteristic polynomials are available through the Characteristic Polynomial Database [25].

n	#matrices	#cpolys	#neutral polys	#neutral matrices
2	27	16	2	4
3	729	166	3	24
4	59,049	3,317	7	332
5	14,348,907	133,255	11	9,909
6	10,460,353,203	10,872,459	25	696,083

TABLE 1

Some properties of matrices in $\mathcal{H}^{n \times n}_{\{0\}}(\{-1, 0, +1\})$. The #matrices column reports the number of distinct matrices at each dimension. The #cpolys column reports the number of distinct characteristic polynomials at each dimension. The #neutral polys reports the number of characteristic polynomials where all roots have zero real part. The #neutral matrices column reports the number of matrices where all eigenvalues have zero real part.

n	#matrices	#cpolys	#neutral polys	#neutral matrices
2	3	3	2	2
3	27	15	3	6
4	729	140	7	66
5	59,049	2,297	11	1,069
6	14,348,907	67,628	25	45,375
7	10,460,353,203	3,606,225	45	4,105,977

TABLE 2

Some properties of matrices in $\mathcal{Z}^{n \times n}_{\{0\}}(\{-1, 0, +1\})$. The #matrices column reports the number of distinct matrices at each dimension. The #cpolys column reports the number of distinct characteristic polynomials at each dimension. The #neutral polys reports the number of characteristic polynomials where all roots have zero real part. The #neutral matrices column reports the number of matrices where all eigenvalues have zero real part.

Other questions than those answered in these tables can be asked of this data. For instance, one might be interested in the proportion of singular matrices. By asking which characteristic polynomials have zero constant coefficient, and counting the number of matrices that have that characteristic polynomial, one can answer such questions. In the case of six by six upper Hessenberg matrices with population $\{-1, +1\}$, there are 383,680 singular matrices, or about 18.3%. Recall that for "ran-

n	multiplicity 1	2	3	4	5	6
2	5	1				
3	35	0	1			
4	431	5	0	1		
5	9,497	9	3	0	1	
6	363,143	51	5	1	0	1

TABLE 3

Number of distinct eigenvalues of various multiplicities of matrices in $\mathcal{Z}_{\{0\}}^{n\times n}(\{-1,0,+1\})$. Most eigenvalues are simple. It turns out that every multiple eigenvalue also occurs as a simple eigenvalue for some other matrix. The only n-multiple eigenvalue of the class of n by n matrices is, of course, $\lambda = 0$.

n	#matrices	#cpolys	#distinct real λ	#neutrals polys	#neutral matrices
2	8	6	6	1	1
3	64	28	25	1	1
4	1,024	197	219	1	1
5	32,768	2,235	3,264	1	1
6	2,097,152	39,768	75,045	1	1
7	268,435,456	1,140,848	2,694,199	1	1

TABLE 4

Some properties of matrices in $\mathcal{H}_{\{0\}}^{n\times n}(\{0,+1\})$. The #matrices column reports the number of distinct matrices at each dimension. The #cpolys column reports the number of distinct characteristic polynomials at each dimension. The #distinct real λ column reports the number of distinct real eigenvalues in $\mathcal{H}_{\{0\}}^{n\times n}(\{0,+1\})$. The #neutral polys reports the number of characteristic polynomials where all roots have zero real part (here only z^n). We conjecture that this is always so (and that there is only one matrix for that neutral polynomial). The #neutral matrices column reports the number of matrices where all eigenvalues have zero real part.

dom" six by six matrices, where the entries are chosen perhaps uniformly over some real interval, the probability of singularity is *zero* because such matrices come from a set of measure zero. Yet in applications, the probability of singular matrices is often *nonzero* because of structure. By looking at Bohemian matrices, we get some idea of the influence of structure for finite dimensions n.

5. Upper Hessenberg Matrices. We can make sense of some of those experiments by theoretical results and proofs. We begin with a recurrence relation for the characteristic polynomial $Q_n(z) = \det(z\mathbf{I} - \mathbf{H}_n)$ for $\mathbf{H}_n \in \mathcal{H}_{\{\theta_k\}}^{n\times n}(P)$ where $s = \exp(i\theta_k)$. Later we will specialize the population P to contain only zero and numbers of unit magnitude, usually $\{-1, 0, +1\}$.

THEOREM 5.1.

$$(5.1) \qquad Q_n(z) = zQ_{n-1}(z) - \sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} Q_{n-k}(z)$$

with the convention that $Q_0(z) = 1$ ($\mathbf{H}_0 = []$, the empty matrix).

n	multiplicity 1	2	3	4	5	6
2	6	2				
3	43	2	2			
4	413	6	2	2		
5	6,920	6	3	2	2	
6	166,005	45	6	2	2	2

TABLE 5

Number of distinct eigenvalues of various multiplicities matrices in $\mathcal{H}^{n\times n}_{\{0\}}(\{0,+1\})$. Note that in this class of matrices, diagonal entries of the matrix need not be zero.

n	#matrices	#cpolys	#stables	#neutral polys	#neutral matrices	#distinct real λ
2	8	6	1	1	2	5
3	64	32	3	0	0	29
4	1,024	289	14	1	6	233
5	32,768	4,958	93	0	0	7,363
6	2,097,152	162,059	992	2	430	299,477
7	268,435,456	10,318,948		0	0	

TABLE 6

Some properties of matrices from $\mathcal{H}^{n\times n}_{\{0\}}(\{-1,+1\})$. The column #stables reports the number of characteristic polynomials with all roots in the left half plane; the corresponding number of matrices is 1, 4, 28, 424, and 11,613. Other columns are as in previous tables. Blank table entries represent unknowns.

Proof. We begin by proving the following equality:

$$(5.2) \qquad \det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{i-1} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{i-1,n} \end{matrix} \\ \hline 0 \quad \cdots \quad 0 \quad -s & -h_{i,n} \end{bmatrix} = -\sum_{k=1}^{i} s^{k-1} h_{i-k+1,n} Q_{i-k}(z)$$

for $1 \le i \le n$.

When $i = 1$ the left side of equation (5.2) reduces to $\det \begin{bmatrix} -h_{1,n} \end{bmatrix} = -h_{1,n}$, and the right side reduces to $-\sum_{k=1}^{1} s^{k-1} h_{1-k+1,n} Q_{1-k}(z) = -h_{1,n}$.

Assume inductively that

$$(5.3) \qquad \det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{j-1} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{j-1,n} \end{matrix} \\ \hline 0 \quad \cdots \quad 0 \quad -s & -h_{j,n} \end{bmatrix} = -\sum_{k=1}^{j} s^{k-1} h_{j-k+1,n} Q_{j-k}(z)$$

for $i = j - 1$. Then

$$\det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{j} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{j,n} \end{matrix} \\ \hline 0 \quad \cdots \quad 0 \quad -s & -h_{j+1,n} \end{bmatrix} = -h_{j+1,n} \det(z\mathbf{I} - \mathbf{H}_{j}) + s \det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{j-1} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{j-1,n} \end{matrix} \\ \hline 0 \quad \cdots \quad 0 \quad -s & -h_{j,n} \end{bmatrix}$$

$$(5.4) \qquad\qquad = -h_{j+1,n} Q_j(z) + s \left(-\sum_{k=1}^{j} s^{k-1} h_{j-k+1,n} Q_{j-k}(z) \right)$$

n	multiplicity 1	2	3	4	5	6
2	9	1				
3	65	0	0			
4	689	5	0	0		
5	20,565	3	0	0	0	
6	887,539	59	9	1	1	1

TABLE 7

Number of distinct eigenvalues of various multiplicities of matrices from $\mathcal{H}_{\{0\}}^{n\times n}(\{-1,+1\})$. The diagonal entries are not zero.

$$(5.5) \qquad = -h_{j+1,n}Q_j(z) - \sum_{k=1}^{j} s^k h_{j-k+1,n} Q_{j-k}(z)$$

$$(5.6) \qquad = -\sum_{k=0}^{j} s^k h_{j-k+1,n} Q_{j-k}(z)$$

$$(5.7) \qquad = -\sum_{k=1}^{j+1} s^{k-1} h_{(j+1)-k+1,n} Q_{(j+1)-k}(z) \,.$$

Next we prove the theorem. Performing Laplace expansion on the last row of $z\mathbf{I} - \mathbf{H}_n$ we get

$$(5.8) \quad Q_n(z) = \det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{n-1} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{n-1,n} \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 & -s \end{matrix} & z - h_{n,n} \end{bmatrix}$$

$$(5.9) \qquad = (z - h_{n,n})\det(z\mathbf{I} - \mathbf{H}_{n-1}) + s\det \begin{bmatrix} z\mathbf{I} - \mathbf{H}_{n-2} & \begin{matrix} -h_{1,n} \\ \vdots \\ -h_{n-2,n} \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 & -s \end{matrix} & -h_{n-1,n} \end{bmatrix}$$

$$(5.10) \qquad = zQ_{n-1}(z) - h_{n,n}Q_{n-1}(z) + s\left(-\sum_{k=1}^{n-1} s^{k-1} h_{n-1-k+1,n} Q_{n-1-k}(z)\right)$$

$$(5.11) \qquad = zQ_{n-1}(z) - h_{n,n}Q_{n-1}(z) - \sum_{k=1}^{n-1} s^k h_{n-k,n} Q_{n-1-k}(z)$$

$$(5.12) \qquad = zQ_{n-1}(z) - \sum_{k=0}^{n-1} s^k h_{n-k,n} Q_{n-1-k}(z)$$

$$(5.13) \qquad = zQ_{n-1}(z) - \sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} Q_{n-k}(z) \,.$$

THEOREM 5.2. *Expanding $Q_n(z)$ as*

$$(5.14) \qquad\qquad Q_n(z) = q_{n,n}z^n + q_{n,n-1}z^{n-1} + \cdots + q_{n,0},$$

we can express the coefficients recursively by

$$(5.15\text{a}) \qquad q_{n,n} = 1,$$

$$(5.15\text{b}) \qquad q_{n,j} = q_{n-1,j-1} - \sum_{k=1}^{n-j} s^{k-1} h_{n-k+1,n} q_{n-k,j} \quad for \quad 1 \le j \le n-1,$$

$$(5.15\text{c}) \qquad q_{n,0} = -\sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} q_{n-k,0} \quad for \quad n > 0, \quad and$$

$$(5.15\text{d}) \qquad q_{0,0} = 1 .$$

Proof. By Theorem 5.1

$$(5.16) \qquad Q_n(z) = z Q_{n-1}(z) - \sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} Q_{n-k}(z) .$$

The first term can be written

$$(5.17) \qquad z Q_{n-1}(z) = z \left[z^{n-1} + q_{n-1,n-2} z^{n-2} + \cdots + q_{n-1,0} \right]$$

$$(5.18) \qquad = z \left[z^{n-1} + \sum_{j=0}^{n-2} q_{n-1,j} z^j \right]$$

$$(5.19) \qquad = z^n + \sum_{j=0}^{n-2} q_{n-1,j} z^{j+1}$$

$$(5.20) \qquad = z^n + \sum_{j=1}^{n-1} q_{n-1,j-1} z^j$$

and the second term

$$s^{k-1} h_{n-k+1,n} Q_{n-k}(z) = s^{k-1} h_{n-k+1,n} \left[q_{n-k,n-k} z^{n-k} + q_{n-k,n-k-1} z^{n-k-1} + \cdots + q_{n-k,0} \right]$$

$$(5.21) \qquad = s^{k-1} h_{n-k+1,n} \sum_{j=0}^{n-k} q_{n-k,j} z^j .$$

Therefore,

$$Q_n(z) = z^n + \sum_{j=1}^{n-1} q_{n-1,j-1} z^j - \sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} \sum_{j=0}^{n-k} q_{n-k,j} z^j$$

$$= z^n + \sum_{j=1}^{n-1} q_{n-1,j-1} z^j - \sum_{j=0}^{n-1} \left(\sum_{k=1}^{n-j} s^{k-1} h_{n-k+1,n} q_{n-k,j} \right) z^j$$

$$= z^n + \sum_{j=1}^{n-1} \left(q_{n-1,j-1} - \sum_{k=1}^{n-j} s^{k-1} h_{n-k+1,n} q_{n-k,j} \right) z^j - \sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} q_{n-k,0} .$$

PROPOSITION 5.3. *All matrices in* $\mathcal{H}_{\{\theta_k\}}^{n \times n}(P)$ *are non-derogatory*[8].

---

[8]A non-derogatory matrix is a matrix for which its characteristic polynomial and minimal polynomial coincide (up to a factor of ± 1)

Proof. Let $\mathbf{H} \in \mathcal{H}^{n \times n}_{\{\theta_k\}}(P)$. Because $\mathbf{H}$ is upper Hessenberg

$$(5.22) \qquad \mathbf{H}^k_{i,j} = \begin{cases} f_{i,j,k} & \text{for} \quad i < j + k \\ s^k & \text{for} \quad i = j + k \\ 0 & \text{for} \quad i > j + k \end{cases}$$

for $0 \leq k \leq n - 1$ where $f_{i,j,k}$ are some functions of the entries of $\mathbf{H}$. Let

$$(5.23) \qquad \mathbf{A} = r(\mathbf{H}) = \sum_{k=0}^{n-1} c_k \mathbf{H}^k = \mathbf{0} \,.$$

We find $\mathbf{A}_{n,1} = s^{n-1} c_{n-1} = 0$ and therefore $c_{n-1} = 0$. Continuing recursively for k from $n-2$ to 1 we find $\mathbf{A}_{k+j,j} = s^k c_k = 0$ for $1 \leq j \leq n-k$ and therefore $c_k = 0$ (since $c_j = 0$ for $j > k$) for $1 \leq k \leq n-1$. We have $\mathbf{A} = c_0 \mathbf{H}^0 = \mathbf{0}$ and hence $c_0 = 0$. Thus, no non-zero polynomial of degree less than n exists that satisfies $r(\mathbf{H}) = \mathbf{0}$. Therefore, the minimal degree non-zero polynomial that satisfies $r(\mathbf{H}) = \mathbf{0}$ is the characteristic polynomial of $\mathbf{H}$. ♮

DEFINITION 5.4. *The characteristic height of a matrix is the height of its characteristic polynomial.*

Remark 5.5. The height of a polynomial is in fact a norm (the infinity norm of the vector of coefficients).

PROPOSITION 5.6. *For any matrix* $\mathbf{A}$, $-\mathbf{A}$ *has the same characteristic height as* $\mathbf{A}$.

PROPOSITION 5.7. *The maximal characteristic height of* $\mathbf{H}_n \in \mathcal{H}^{n \times n}_{\{0,\pi\}}(\{-1,0,+1\})$ *occurs when* $s^{k-1} h_{i,i+k-1} = -1$ *for* $1 \leq i \leq n-k+1$ *and* $1 \leq k \leq n$.

Proof. Since $s \in \{-1,+1\}$ and $h_{i,j} \in \{-1,0,+1\}$, $s^{k-1} h_{i,i+k-1} \in \{-1,0,+1\}$ and hence $\max |s^{k-1} h_{i,i+k-1}| = 1$. Let $s^{k-1} h_{i,i+k-1} = -1$. By Theorem 5.2

$$(5.24) \qquad q_{n,0} = -\sum_{k=1}^{n} s^{k-1} h_{n-k+1,n} q_{n-k,0}$$

$$(5.25) \qquad = \sum_{k=1}^{n} q_{n-k,0}$$

and

$$(5.26) \qquad q_{n,j} = q_{n-1,j-1} - \sum_{k=1}^{n-j} s^{k-1} h_{n-k+1,n} q_{n-k,j}$$

$$(5.27) \qquad = q_{n-1,j-1} + \sum_{k=1}^{n-j} q_{n-k,j} \,.$$

Since $q_{0,0} = 1$, and equations (5.25) and (5.27) are independent of s and $h_{i,j}$, all $q_{n,j}$ must be positive and the maximum characteristic height is attained. ♮

Remark 5.8. When $s = 1$ ($\theta = 0$) and $h_{i,j} = -1$ for all $1 \leq i \leq j \leq n$, $\mathbf{H}_n$ attains maximal characteristic height. By Proposition 5.6, $s = -1$ ($\theta = \pi$) and $h_{i,j} = 1$ will also attain maximal characteristic height. Both of these cases correspond to upper Hessenberg matrices with a Toeplitz structure as we explore in further detail in the paper [11].

DEFINITION 5.9. P *is invariant under multiplication by a fixed unit* $e^{i\theta}$ *if* $e^{i\theta}P = P$; *that is, each entry of* P, *say* p, *is such that* $e^{i\theta}p$ *is also in* P. *For instance,* $\{-1, 0, +1\}$ *is invariant under multiplication by* -1. *Note that invariance with respect to* $e^{i\theta}$ *implies invariance with respect to* $e^{-i\theta}$.

THEOREM 5.10. *Suppose* $\mathbf{H}_n \in \mathcal{H}_{\{\theta_k\}}^{n \times n}(P)$ *and* P *is invariant under multiplication by each* $e^{i\theta_k}$ *and by* $-e^{i\theta_k}$. *Then* $\mathbf{H}_n$ *is similar to a matrix in* $\mathcal{H}_{\{\pi\}}^{n \times n}(P)$, *and similar to a matrix in* $\mathcal{H}_{\{0\}}^{n \times n}(P)$.

Proof. We use induction. The case $n = 1$ is vacuously upper Hessenberg, though it is

$$\left[e^{i\theta_k}\right]\left[h_{11}\right]\left[e^{-i\theta_k}\right] = \left[h_{11}\right] \in \mathcal{H}_{\{\theta_k\}}^{1 \times 1}(P)\,.$$

For $n > 1$, partition the matrix as

$$\begin{bmatrix} h_{11} & h_{12}\cdots & h_{1n} \\ s & & \\ & & \mathbf{H}_{n-1} \end{bmatrix}$$

where $s = e^{i\theta_k}$ for some θ_k. Then conjugate by

$$\begin{bmatrix} 1 & & \\ -e^{i\theta_k} & & \\ & & \mathbf{I}_{n-2} \end{bmatrix}\begin{bmatrix} h_{11} & h_{12}\cdots & h_{1n} \\ s & & \\ & & \mathbf{H}_{n-1} \end{bmatrix}\begin{bmatrix} 1 & & \\ -e^{i\theta_k} & & \\ & & \mathbf{I}_{n-2} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} h_{11} & -e^{-i\theta_k}h_{12} & \cdots \\ -1 & & \\ & \tilde{\mathbf{H}}_{n-1} & \end{bmatrix}\,.$$

Clearly $\tilde{\mathbf{H}}_{n-1}$ is in $\mathcal{H}_{\{\theta_k\}}^{n-1 \times n-1}(P)$. By induction the proof is complete. ♮

Remark 5.11. For clarity, consider the case $n = 2$:

$$(5.28) \qquad \mathbf{H} = \begin{bmatrix} a & b \\ s & c \end{bmatrix},$$

where $a, b, c \in P$ and $s = e^{i\theta_k}$. Then, the following similarity transforms reduce the problem to one in $\mathcal{H}_{\{0\}}^{2 \times 2}(P)$ and one in $\mathcal{H}_{\{\pi\}}^{2 \times 2}(P)$.

$$(5.29) \qquad \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\theta_k} \end{bmatrix}\mathbf{H}\begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta_k} \end{bmatrix} = \begin{bmatrix} a & be^{i\theta_k} \\ 1 & c \end{bmatrix}$$

$$(5.30) \qquad \begin{bmatrix} 1 & 0 \\ 0 & -e^{-i\theta_k} \end{bmatrix}\mathbf{H}\begin{bmatrix} 1 & 0 \\ 0 & -e^{i\theta_k} \end{bmatrix} = \begin{bmatrix} a & -be^{i\theta_k} \\ -1 & c \end{bmatrix}.$$

6. Upper Hessenberg Toeplitz Matrices. Proposition 5.7 gives matrices in $\mathcal{H}_{\{0,\pi\}}^{n \times n}(\{-1, 0, +1\})$ with maximal characteristic height[9]. We noticed that they are Toeplitz matrices. This motivated our interest in upper Hessenberg Toeplitz matrices. We summarize some of the results of [11] here.

---

[9]We did not report the numbers of such matrices and polynomials that we found in our "results" section.

DEFINITION 6.1. *Consider matrices* $\mathbf{H}_n$ *where* $h_{i,i+k-1} = t_k$ *for* $1 \leq i \leq n-k+1$, $1 \leq k \leq n$ *and* $s = 1$. *We denote these matrices by* $\mathbf{M}_n$ *and they have a Toeplitz structure.*

Remark 6.2. The characteristic height of $\mathbf{M}_n$ is maximal when $t_k = -1$ for $1 \leq k \leq n$.

PROPOSITION 6.3. *Let* $F \subset \mathbb{R}$ *be a closed and bounded set with* $a = \min F$, $b = \max F$ *and* $\#F \geq 2$. *Let* $\mathbf{M}_n$ *be upper Hessenberg Toeplitz with* $t_k \in F$. *If* $|a| \geq |b|$, $\mathbf{M}_n$ *attains maximal characteristic height for* $t_k = a$ *for all* $1 \leq k \leq n$. *If* $|b| \geq |a|$, $\mathbf{M}_n$ *attains maximal characteristic height for* $t_k = a$ *for* k *even, and* $t_k = b$ *for* k *odd.*

PROPOSITION 6.4. *The maximum characteristic height grows at least exponentially in* n.

CONJECTURE 6.5. *The maximum characteristic height approaches* $C(1 + \varphi)^n$ *as* $n \to \infty$ *for some constant* C *where* φ *is the golden ratio.*

Remark 6.6. This limit is illustrated in Figure 3, motivating this conjecture.
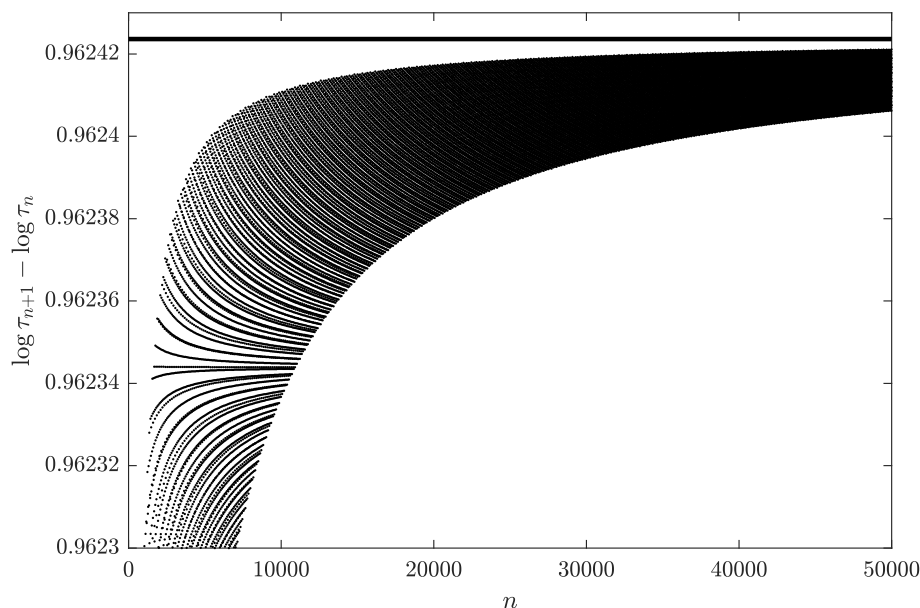

FIG. 3. *The points are* $\log \tau_{n+1} - \log \tau_n$ *for* n *from 0 to 50000 where* τ_n *is the maximal characteristic height of* $\mathbf{M}_n$ *(i.e. when* $t_k = 1$, *for example). The solid line is* $\log(1 + \varphi)$ *where* φ *is the golden ratio.*

7. Zero Diagonal Upper Hessenberg Matrices.

THEOREM 7.1. *Let* $\mathbf{A}_n \in \mathcal{Z}_{\{0\}}^{n \times n}(P)$ *for* $P = \{0, w_1, \ldots, w_m\}$ *for some fixed positive integer* m *and each* $|w_j| = 1$. *If* $\mathbf{A}_n$ *is normal, i.e.* $\mathbf{A}_n^* \mathbf{A}_n = \mathbf{A}_n \mathbf{A}_n^*$, *then for* $n \geq 3$, $\mathbf{A}_n$ *is symmetric,* w_j-*skew symmetric for some fixed* $1 \leq j \leq m$ *or* w_j-*skew circulant. These* $2m$ *matrices (*m *symmetric/*w_j-*skew symmetric, and* m w_j-*skew circulant matrices) are the only normal matrices in* $\mathcal{Z}_{\{0\}}^{n \times n}(P)$. *(For* $n = 1$, *this is only*

[0]; *for $n = 2$, the symmetric and circulant cases coalesce, so that there are only m such matrices.)*

Proof. To prove this theorem, we establish a sequence of lemmas. First, we partition $\mathbf{A}_n$. Put

$$
(7.1) \qquad \mathbf{A}_n = \begin{bmatrix} 0 & \mathbf{T}^* \\ \mathbf{e} & \mathbf{A}_{n-1} \end{bmatrix}
$$

where

$$
(7.2) \qquad \mathbf{e}^* = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}
$$

and

$$
(7.3) \qquad \mathbf{T}^* = \begin{bmatrix} t_{12} & t_{13} & \cdots & t_{1n} \end{bmatrix} = \begin{bmatrix} t_{21}^* & t_{31}^* \cdots t_{n1}^* \end{bmatrix}^* .
$$

Then the conditions of normality are

$$
(7.4) \qquad \mathbf{A}_n \mathbf{A}_n^* = \begin{bmatrix} \mathbf{T}^*\mathbf{T} & \mathbf{T}^*\mathbf{A}_{n-1}^* \\ \mathbf{A}_{n-1} & \mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-1}\mathbf{A}_{n-1}^* \end{bmatrix}
$$

must equal

$$
(7.5) \qquad \mathbf{A}_n^* \mathbf{A}_n = \begin{bmatrix} 1 & \mathbf{e}^*\mathbf{A}_{n-1} \\ \mathbf{A}_{n-1}^*\mathbf{e} & \mathbf{T}\mathbf{T}^* + \mathbf{A}_{n-1}^*\mathbf{A}_{n-1} \end{bmatrix} . \qquad\qquad ♮
$$

LEMMA 7.2. *The first row of $\mathbf{A}_n$ contains exactly one nonzero element, say τ in position j ($2 \le j \le n$).*

Proof.

$$
(7.6) \qquad \mathbf{T}^*\mathbf{T} = \sum_{j=2}^{n} |t_{ij}|^2 = 1
$$

from the upper left corner. Since each nonzero element of P has magnitude 1, exactly one entry must be nonzero. ♮

LEMMA 7.3. *If $\mathbf{A}_{n-1}$ is normal then $\mathbf{T} = \tau\mathbf{e}$ and $\mathbf{A}_n$ is τ-skew symmetric.*

Proof. If $\mathbf{A}_{n-1}$ is normal, then $\mathbf{T}\mathbf{T}^* + \mathbf{A}_{n-1}^*\mathbf{A}_{n-1}$ being equal to $\mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-1}\mathbf{A}_{n-1}^*$ implies $\mathbf{T}\mathbf{T}^* = \mathbf{e}\mathbf{e}^*$ so that $\mathbf{T}^* = \begin{bmatrix} \tau^* & 0 & \cdots & 0 \end{bmatrix}$ for some τ with $|\tau| = 1$. Then

$$
(7.7) \qquad \mathbf{T}^*\mathbf{A}_{n-1}^* = \mathbf{e}^*\mathbf{A}_{n-1} \Rightarrow \tau^* \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \mathbf{A}_{n-1}^* = \mathbf{e}^*\mathbf{A}_{n-1}^*
$$

and this says τ^* times the first row of $\mathbf{A}_{n-1}^*$ is the first row of $\mathbf{A}_{n-1}$.

But the first row of $\mathbf{A}_{n-1}^*$ is $\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$ because $\mathbf{A}_{n-1}$ is upper Hessenberg with zero diagonal. Thus the first row of $\mathbf{A}_{n-1}$ is $\begin{bmatrix} 0 & \tau^* & 0 & \cdots & 0 \end{bmatrix}$. Thus

$$
(7.8) \qquad \mathbf{A}_n = \left[\begin{array}{c|c|c} 0 & \tau^* & \\ \hline 1 & 0 & \tau^* \\ \hline & 1 & \\ & & \mathbf{A}_{n-2} \end{array} \right] \qquad \text{(remember } n \ge 3)
$$

and

$$(7.9) \qquad \mathbf{A}_{n-1} = \begin{bmatrix} 0 & \tau^* & \\ 1 & & \\ & & \mathbf{A}_{n-2} \end{bmatrix}$$

is normal. Because $\mathbf{A}_{n-1}$ is normal, and

$$(7.10) \qquad \mathbf{A}_{n-1}^* = \begin{bmatrix} 0 & 1 & & \\ \tau & 0 & 1 & \\ & \tau & & \\ & & & \mathbf{A}_{n-2}^* \end{bmatrix}$$

we have $\mathbf{A}_{n-1}^* \mathbf{A}_{n-1} = \mathbf{A}_{n-1} \mathbf{A}_{n-1}^*$ or

$$\begin{bmatrix} 0 & 1 & & \\ \tau & 0 & 1 & \\ & 1 & & \mathbf{A}_{n-2} \end{bmatrix} \begin{bmatrix} 0 & \tau^* & & \\ 1 & 0 & \tau^* & \\ & 1 & & \mathbf{A}_{n-2} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & \tau^* \\ 0 & 2 & \mathbf{e}_{n-2}^* \mathbf{A}_{n-2} \\ \tau & \tau \mathbf{A}_{n-2}^+ \mathbf{e}_{n-2} & \mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-2}^* \mathbf{A}_{n-2} \end{bmatrix}$$

must equal

$$\begin{bmatrix} 0 & \tau^* & & \\ 1 & 0 & \tau^* & \\ & 1 & & \mathbf{A}_{n-2} \end{bmatrix} \begin{bmatrix} 0 & 1 & & \\ \tau & 0 & 1 & \\ & 1 & & \mathbf{A}_{n-2} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & \tau^* \\ 0 & 2 & \mathbf{e}_{n-2}^* \mathbf{A}_{n-2} \\ \tau & \tau \mathbf{A}_{n-2}^+ \mathbf{e}_{n-2} & \mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-2}^* \mathbf{A}_{n-2} \end{bmatrix}.$$

The lower left block gives $\mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-2} \mathbf{A}_{n-2}^* = \mathbf{e}\mathbf{e}^* + \mathbf{A}_{n-2}^* \mathbf{A}_{n-2}$ so $\mathbf{A}_{n-2}$ must also be normal.

At this point, we see the outline of an induction:

$$(7.11) \qquad \mathbf{A}_n = \left[\begin{array}{c|c} 0 & \tau^* \\ \hline 1 & \mathbf{A}_{n-1} \end{array} \right]$$

being normal with $\mathbf{A}_{n-1}$ being normal implies that

$$(7.12) \qquad \mathbf{A}_{n-1} = \left[\begin{array}{c|c} 0 & \tau^* \\ \hline 1 & \mathbf{A}_{n-2} \end{array} \right]$$

where $\mathbf{A}_{n-2}$ is normal. Explicit computation of the $n = 3$ case shows the induction terminates. ♮

We now consider the harder case where

$$(7.13) \qquad \mathbf{A}_n = \begin{bmatrix} 0 & \mathbf{T}^* \\ \mathbf{e}_{n-1} & \mathbf{A}_{n-1} \end{bmatrix}$$

but where $\mathbf{A}_{n-1}$ is not itself normal. From Lemma 7.2 we know that $\mathbf{T}^*$ has only one nonzero element; call it τ^* as before. Then

$$(7.14) \qquad \mathbf{T}\mathbf{T}^* = \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

while

$$(7.15) \qquad \mathbf{e}\mathbf{e}^* = \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix},$$

and we may assume that the 1 in $\mathbf{T}\mathbf{T}^*$ does not occur in the first row and column (else we are in the previous case, and $\mathbf{A}_{n-1}$ will be normal). Here

$$(7.16) \quad \mathbf{A}_{n-1}\mathbf{A}_{n-1}^* - \mathbf{A}_{n-1}^*\mathbf{A}_{n-1} = \mathbf{T}\mathbf{T}^* - \mathbf{e}\mathbf{e}^* = \begin{bmatrix} -1 & & & & & & \\ & 0 & & & & & \\ & & \ddots & & & & \\ & & & 0 & & & \\ & & & & 1 & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{bmatrix}$$

is the departure of $\mathbf{A}_{n-1}$ from normality. We will establish that in fact

$$(7.17) \qquad \mathbf{T}^* = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & \tau^* \end{bmatrix}$$

and that

$$(7.18) \qquad \mathbf{A}_{n-1} = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix};$$

that is, the nonzero element can only occur in the last place. Notice that the upper left corner of 7.16 is, if the top row of $\mathbf{A}_{n-1}$ is $\begin{bmatrix} 0 & a_{1,2} & a_{1,3} & \cdots a_{1,n-1} \end{bmatrix}$,

$$(7.19) \qquad \sum_{j=2}^{n-1} |a_{1,j}|^2 - 1 \,.$$

Therefore, all $a_{1,j} = 0$ and the first row of $\mathbf{A}_{n-1}$ must be zero: i.e.

$$(7.20) \qquad \mathbf{A}_{n-1} = \begin{bmatrix} 0 & 0 & 0 & \cdots & & 0 \\ 1 & 0 & a_{3,3} & \cdots & & a_{2,n-1} \\ & 1 & 0 & \ddots & & \vdots \\ & & \ddots & \ddots & & a_{n-2,n-1} \\ & & & 1 & & 0 \end{bmatrix}$$

Then,

$$(7.21) \qquad \mathbf{A}_{n-1}\mathbf{T} = \mathbf{A}_{n-1}^{*}\mathbf{e} = \begin{bmatrix} 0 & 1 & & \\ 0 & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

If

$$(7.22) \qquad \mathbf{T} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \tau \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

then

$$(7.23) \qquad \mathbf{A}_{n-1}\mathbf{T} = \begin{bmatrix} 0 \\ \tau a_{2,j} \\ \vdots \\ \tau a_{j-1,j} \\ 0 \\ \tau \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

which is impossible unless $j = n$ (when the τ term is not present). Therefore,

$$(7.24) \qquad \mathbf{A}_{n-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & x & \cdots & x & 0 \\ & 1 & \ddots & \vdots & \vdots \\ & & \ddots & x & 0 \\ & & & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mathbf{U} & 0 \end{bmatrix},$$

and

$$(7.25) \qquad \mathbf{A}_{n-1}\mathbf{A}_{n-1}^* - \mathbf{A}_{n-1}^*\mathbf{A}_{n-1} = \begin{bmatrix} -1 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & 1 \end{bmatrix}.$$

Since

$$(7.26) \qquad \mathbf{A}_{n-1}^* = \begin{bmatrix} 0 & \mathbf{U}^* \\ 0 & 0 \end{bmatrix}$$

and

$$(7.27) \qquad \mathbf{A}_{n-1}\mathbf{A}_{n-1}^* = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{U}\mathbf{U}^* \end{bmatrix}$$

and

$$(7.28) \qquad \mathbf{A}_{n-1}^*\mathbf{A}_{n-1} = \begin{bmatrix} \mathbf{U}^*\mathbf{U} & 0 \\ 0 & 0 \end{bmatrix},$$

$$(7.29) \qquad \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{U}\mathbf{U}^* \end{bmatrix} - \begin{bmatrix} \mathbf{U}^*\mathbf{U} & 0 \\ 0 & 0 \end{bmatrix}$$

must be diagonal. Therefore, the first row of $\mathbf{U}\mathbf{U}^*$ must be zero except for the first element.

Remark 7.4. For $n = 4$, and $P = \{0, i, -i\}$ ($m = 2$) the following 4 matrices are normal:

w_j	w_j-skew symmetric	w_j-skew circulant
i	$\begin{bmatrix} 0 & i & 0 & 0 \\ 1 & 0 & i & 0 \\ & 1 & 0 & i \\ & & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & i \\ 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \end{bmatrix}$
$-i$	$\begin{bmatrix} 0 & -i & 0 & 0 \\ 1 & 0 & -i & 0 \\ & 1 & 0 & -i \\ & & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & -i \\ 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \end{bmatrix}$

8. Stable Matrices.

8.1. Type I Stable Matrices.
A *Type I stable matrix* $\mathbf{A}$ is a matrix with all of its eigenvalues strictly in the left half plane: if λ is an eigenvalue of A then $\Re(\lambda) < 0$. This nomenclature comes from differential equations, in that all solutions of the linear system of ODEs $dy/dt = \mathbf{A}y$ will ultimately decay as $t \to \infty$ if $\mathbf{A}$ is a type I stable matrix.

If the matrix $\mathbf{A}$ is not *normal*, then *pseudospectra* can play a role, in that even though all solutions y must ultimately decay, they might first grow large. See [18] for details.

By Theorem 7.1, only $2m$ of the zero diagonal upper Hessenberg matrices with population $P = \{-1, 0, +1\}$ are normal, where here $m = 2$. Similarly, when the population is $P = \{0, +1\}$ then $m = 1$ and only two matrices of every dimension are normal (the symmetric matrix with 1s on its upper diagonal, and the circulant matrix with a 1 in the last column of the first row).

THEOREM 8.1. *No* $\mathbf{A}_n \in \mathcal{Z}^{n \times n}_{\{\theta_k\}}(P)$ *is Type I stable, for any population* P.

Proof. Suppose $\mathbf{A}_n \in \mathcal{Z}^{n \times n}_{\{\theta_k\}}(P)$ has eigenvalues $\{\lambda_k\}_{k=1}^{n}$. Then

$$(8.1) \qquad \sum_{k=1}^{n} \lambda_k = \text{Trace}(\mathbf{A}_n) = 0\,.$$

Therefore, $\sum_{k=1}^{n} \text{Re}(\lambda_k) = 0$. This is n times the average, and so the average is zero. Since the maximum $\text{Re}(\lambda_k)$ must be larger than the average, this proves the theorem.

The proof of this theorem did not depend on the structure or population. Thus if we consider $\mathcal{H}^{n \times n}_{\{0\}}(P)$ instead of $\mathcal{Z}^{n \times n}_{\{0\}}(P)$, then we may simplify our search for stable matrices by restricting the computation to those with negative trace. This is in fact the first inequality of the Hurwitz criteria[10], which leads to an effective and efficient method to count stable matrices: start from the database of characteristic polynomials [25], decide using the Hurwitz criteria if all roots are in the left half-plane, and if so add its matrices to the count.

n	$\mathcal{H}^{n \times n}_{\{0\}}(\{-1, 0, +1\})$	$\mathcal{H}^{n \times n}_{\{0\}}(\{-1, +1\})$
2	4	1
3	44	4
4	1,386	28
5	130,735	424
6	35,217,156	11,613
7		617,619

TABLE 8
The numbers of Type I stable matrices for various populations and dimensions.

Remark 8.2. For stable matrices in $\mathcal{H}^{n \times n}_{\{0\}}(\{-1, 0, +1\})$ the maximum real part of any eigenvalue is, for $n = 2$, just -0.5 while for $n = 3$ it is $-1.226 \cdot 10^{-1}$. For $n = 4$ it is $-1.591 \cdot 10^{-2}$. For $n = 5$ it is $-5.176 \cdot 10^{-4}$. For $n = 6$ it is $-2.42 \cdot 10^{-5}$. The maximum real part of the eigenvalues seems to be approaching the real axis at least exponentially in n, for this population. It would be nice to have a good asymptotic estimate.

The sequence of maximum real parts of eigenvalues for $\mathcal{H}^{n \times n}_{\{0\}}(\{-1, +1\})$ gives at $n = 2$ $\Re(\lambda) = -1, -0.5, -2.168 \cdot 10^{-2}, -2.66 \cdot 10^{-3}, -1.70 \cdot 10^{-4}$, and $-2.62 \cdot 10^{-6}$ for $n = 7$.

---

[10] The Maple command `PolynomialTools[Hurwitz]` implements a well-known test to decide if $p \in \mathbb{C}[z]$ has all its roots strictly in the left half plane. Because that routine considers the complex case, and tests for pathological cases, it is too inefficient to use in this context. We unrolled the loops, essentially converting the code to specific tests of the principal minors of the Hurwitz matrix.
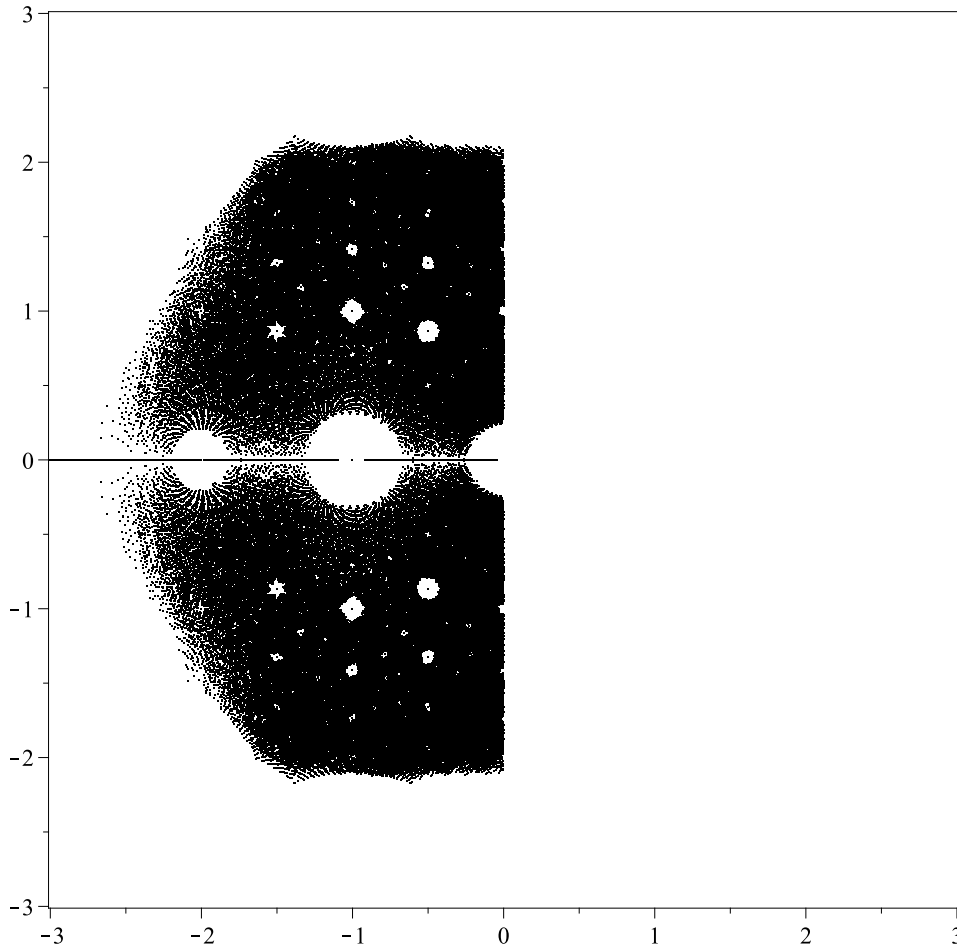
FIG. 4. *All eigenvalues of all 35,217,156 stable matrices from* $\mathcal{H}^{6\times6}_{\{0\}}(\{-1,0,+1\})$. *The maximum real part is approximately* $-2.42\cdot10^{-5}$. *There were only* 55,298 *distinct characteristic polynomials from all these matrices.*

8.2. Type II Stable matrices. A *Type II Stable Matrix* **A** has all its eigenvalues inside the unit circle. This class of matrices arises naturally on studying the simple linear recurrence relation $y_{n+1} = \mathbf{A}y_n$. Fairly obviously, all solutions of this difference equation will ultimately decay to 0 as $n \to \infty$ if and only if all eigenvalues of **A** are inside the unit circle (again, pseudospectra can play a role in the transient behaviour, sometimes significantly).

THEOREM 8.3. *If* **A** *is a Bohemian matrix with integer population P, then it is Type II stable if and only if it is nilpotent, in which case all its eigenvalues are* 0.

Proof. Suppose to the contrary that some eigenvalues are not zero.

The determinant of **A** must necessarily be an integer. If the integer is not zero, it is at least 1 in magnitude. The product of the eigenvalues is thus at least 1 in magnitude; hence there must be at least one eigenvalue that is at least 1 in magnitude.

If the matrix $\mathbf{A}$ has zero determinant but not all eigenvalues zero, then after factoring out z^m for the multiplicity of the zero eigenvalue, the product of the other eigenvalues becomes the constant coefficient (what was the coefficient of z^m in the original). This coefficient again must be an integer, and again at least one eigenvalue must be at least 1 in magnitude.

This proves the theorem, by contradiction. ♮

Remark 8.4. We did not, in fact, use that the matrix came from a Bohemian family; only that its entries were integers.

Searching for nilpotent matrices in various classes of Bohemian matrices turns up several puzzles. We give some preliminary results here in Table 9, but leave this mostly to future work. For instance, it seems clear from our experiments that the only nilpotent matrix in $\mathcal{H}_{\{0\}}^{n \times n}(\{0, +1\})$ is the (transpose of the) complete Jordan block of n zero eigenvalues; contrariwise the irregular behaviour for $\mathcal{H}_{\{0\}}^{n \times n}(\{-1, +1\})$ is very puzzling.

n	$\mathcal{Z}_{\{0\}}^{n \times n}(\{-1, 0, +1\})$	$\mathcal{H}_{\{0\}}^{n \times n}(\{0, +1\})$	$\mathcal{H}_{\{0\}}^{n \times n}(\{-1, +1\})$
2	1	1	2
3	3	1	0
4	21	1	0
5	271	1	0
6	9,075	1	324

TABLE 9
The numbers of nilpotent matrices for various populations and dimensions

Considering *general* Bohemian matrices with population $\{-1, 0, +1\}$, so that there are 3^{n^2} such matrices, we find that there are 1, 9, 481, 148,817, and 243,782,721 nilpotent matrices at dimensions 1 through 5 inclusive. We can fit this experimentally with the formula $\exp(0.5 + 0.38n + 0.23n^2)$, or something like 1.26^{n^2}, which vanishes very quickly compared to 3^{n^2}. This formula predicts that for $n = 6$ the probability of finding a nilpotent matrix is about 2.75×10^{-14}. It would be gratifying to have a better understanding of the number of nilpotent matrices in a family.

9. Concluding Remarks. The class of upper Hessenberg Bohemian matrices gives a useful way to study Bohemian matrices in general. This is an instance of Polya's adage "find a useful specialization." [22, p. 190] Because these classes are simpler than the general case, we were able to establish several theorems. Note that the three families $\mathcal{H}_{\{0\}}^{n \times n}(\{0, +1\})$, $\mathcal{H}_{\{0\}}^{n \times n}(\{-1, +1\})$, and $\mathcal{Z}_{\{0\}}^{n \times n}(\{-1, 0, +1\})$ are all subfamilies of $\mathcal{H}_{\{0\}}^{n \times n}(\{-1, 0, +1\})$.

In this paper we have introduced two new formulae for computing the characteristic polynomials of upper Hessenberg matrices. Our first formula, given in Theorem 5.1, also computes the characteristic polynomials recursively. Our second formula, given in Theorem 5.2, computes the coefficients recursively. We also explored some properties of zero diagonal Bohemian upper Hessenberg matrices. In Theorem 7.1, we show that the subset of these matrices that are normal are always symmetric, w_j-skew symmetric for some fixed $1 \leq j \leq m$, or w_j-skew circulant. In Theorem 8.1, we showed that no $\mathbf{H} \in \mathcal{Z}_{\{\theta_k\}}^{n \times n}(P)$ is stable.

Many puzzles remain. Perhaps the most striking is the angular appearance of the set $\mathbf{\Lambda}(\mathcal{H}_{\{0\}}^{n \times n}(P))$ of eigenvalues of $\mathcal{H}_{\{0\}}^{n \times n}(P)$, such as in Figures 1 and 2. General

matrices have eigenvalues asymptotic to a (scaled) disc [23]; our computations suggest that as $n \to \infty$, $\Lambda(\mathcal{H}_{\{0\}}^{n \times n}(P))/n^{1/2}$ tends to an irregular hexagonal shape, rather than a disk. More, the density does not appear to be approaching uniformity. Further, the boundary is irregular, with shapes suggestive of what is popularly known as the "dragon curve" (in reverse—these delineate where the eigenvalues are absent, near the edge). We have no explanation for this.

REFERENCES

[1] J. BAEZ, *The beauty of roots*, Available at: https://johncarlosbaez.wordpress.com/2011/12/11/the-beauty-of-roots/, (2011).

[2] F. BEAUCOUP, P. BORWEIN, D. W. BOYD, AND C. PINNER, *Multiple roots of* $[-1,1]$ *power series*, Journal of the London Mathematical Society, 57 (1998), pp. 135–147.

[3] C. BELTRÁN AND D. ARMENTANO, *The polynomial eigenvalue problem is well conditioned for random inputs*, arXiv preprint arXiv:1706.06025, (2017).

[4] P. BORWEIN, *Computational excursions in analysis and number theory*, Springer Science & Business Media, 2012.

[5] P. BORWEIN AND L. JÖRGENSON, *Visible structures in number theory*, The American Mathematical Monthly, 108 (2001), pp. 897–910.

[6] P. BORWEIN AND C. PINNER, *Polynomials with* {0,+1,-1} *coefficients and a root close to a given point*, Canadian Journal of Mathematics, 49 (1997), pp. 887–915.

[7] E. Y. S. CHAN, *A comparison of solution methods for Mandelbrot-like polynomials*, Electronic Thesis and Dissertation Repository, (2016). https://ir.lib.uwo.ca/etd/4028.

[8] E. Y. S. CHAN AND R. M. CORLESS, *A new kind of companion matrix*, Electronic Journal of Linear Algebra, 32 (2017), pp. 335–342.

[9] E. Y. S. CHAN AND R. M. CORLESS, *Minimal height companion matrices for Euclid polynomials*, Mathematics in Computer Science, (2018), https://doi.org/10.1007/s11786-018-0364-2.

[10] E. Y. S. CHAN, R. M. CORLESS, L. GONZALEZ-VEGA, J. R. SENDRA, AND J. SENDRA, *Algebraic linearizations of matrix polynomials*. Manuscript submitted for publication, 2017.

[11] E. Y. S. CHAN, R. M. CORLESS, L. GONZALEZ-VEGA, J. R. SENDRA, J. SENDRA, AND S. E. THORNTON, *Bohemian upper Hessenberg Toeplitz matrices*. In preparation., 2018.

[12] R. M. CORLESS, *Generalized companion matrices in the Lagrange basis*, in Proceedings EACA, Santander, Spain: Universidad de Cantabria, 2004, pp. 317–322.

[13] R. M. CORLESS AND P. W. LAWRENCE, *Mandelbrot polynomials and matrices*. In preparation.

[14] R. M. CORLESS AND P. W. LAWRENCE, *The largest roots of the Mandelbrot polynomials*, in Computational and Analytical Mathematics, Springer, 2013, pp. 305–324.

[15] R. M. CORLESS AND S. THORNTON, *Visualizing eigenvalues of random matrices*, ACM Communications in Computer Algebra, 50 (2016), pp. 35–39, https://doi.org/10.1145/2930964.2930969.

[16] R. M. CORLESS AND S. E. THORNTON, *The Bohemian eigenvalue project*, ACM Communications in Computer Algebra, 50 (2016), pp. 158–160.

[17] M. ELOUAFI AND A. D. A. HADJ, *A recursion formula for the characteristic polynomial of Hessenberg matrices*, Applied Mathematics and Computation, 208 (2009), pp. 177–179.

[18] M. EMBREE, *Pseudospectra*, in Handbook of Linear Algebra, L. Hogben, ed., Chapman and Hall/CRC, 2013, ch. 23.

[19] C. GEAR, *A simple set of test matrices for eigenvalue programs*, Mathematics of Computation, 23 (1969), pp. 119–125.

[20] N. Higham, *Bohemian matrices in numerical linear algebra.* Available at http://www.maths.manchester.ac.uk/~higham/conferences/bohemian/higham_bohemian18.pdf(June 20, 2018).

[21] K. Kaygisiz and A. Sahin, *Determinant and permanent of Hessenberg matrix and Fibonacci type numbers*, Gen, 9 (2012), pp. 32–41.

[22] G. Polya, *How to solve it: A new aspect of mathematical method*, Princeton University Press, 2014.

[23] T. Tao and V. Vu, *Random matrices have simple spectrum*, Combinatorica, 37 (2017), pp. 539–553.

[24] O. Taussky, *Matrices of rational integers*, Bulletin of the American Mathematical Society, 66 (1960), pp. 327–345.

[25] S. E. Thornton, *The characteristic polynomial database.* Available at http://bohemianmatrices.com/cpdb (Sept. 7, 2018).