# Bohemian Matrices
# An Introduction and some Open Problems

Robert M. Corless
Joint work with many people

2023–10–16

Maple Transactions, ORCCA, Western University & University of Waterloo, Canada

# Announcement: Maple Transactions

**Maple Transactions**

an open access journal with no page charges

mapletransactions.org

We welcome expositions on topics of interest to the Maple community, including in computer-assisted research in mathematics, education, and applications. Student papers especially welcome.

# Example papers

For example, see

Peter J. Baddoo and Lloyd N. Trefethen. *Log-lightning computation of capacity and Green's function.* Maple Transactions Volume 1, Issue 1, Article 14124 (July 2021). https://doi.org/10.5206/mt.v1i1.14124

Richard P. Brent. *Some Instructive Mathematical Errors.* Maple Transactions Volume 1, Issue 1, Article 14069 (July 2021). https://doi.org/10.5206/mt.v1i1.14069

There is also a transcript of an interview with these authors, conducted by Annie Cuyt.

# Another announcement

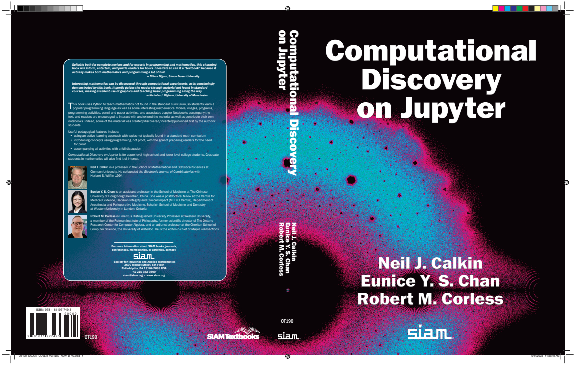


Figure 1: A new book from SIAM: Calkin, Chan, & Corless, “Computational Discovery on Jupyter”, hopefully available November

Details for parts of the talk can be found at *https://arxiv.org/abs/2202.07769*

The Maple Workbook that contains the source code partially implementing the Schmidt–Spitzer theorem can be found, together with all the images from our paper and the slides from this talk, at *https://github.com/rcorless/Bohemian-Matrix-Geometry*

Please download those images and look at them on your own devices. That gives higher resolution than this projection does.

(This is “Screen-sharing for in-person lectures” :)

# Rhapsodizing about Bohemian Matrices

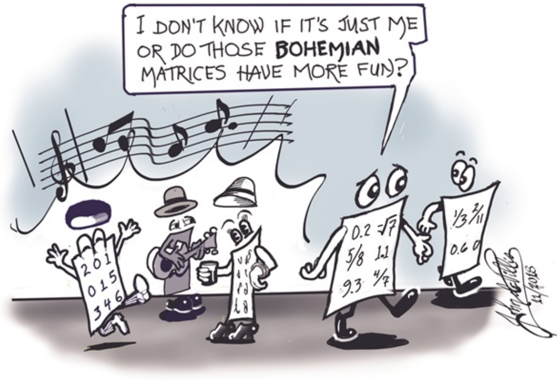


Figure 2: A cartoon by mathematician John de Pillis (UC Riverside), which appeared in Nick Higham's column in SIAM News

# Bohemian Matrices

A family of matrices is called "Bohemian" if all entries are all from a single finite population $P$. The name comes from BOunded HEight Matrix of Integers. See bohemianmatrices.com for instances.

See also the [link] London Mathematical Society Newsletter, November 2020, page 16.

Such matrices have been studied for quite a long time (e.g. by Olga Taussky–Todd), though the name "Bohemian" only dates to 2015. See also the Wikipedia entry at *https://en.wikipedia.org/wiki/Bohemian_matrices*.

# Earlier results by Terence Tao and Van Vu

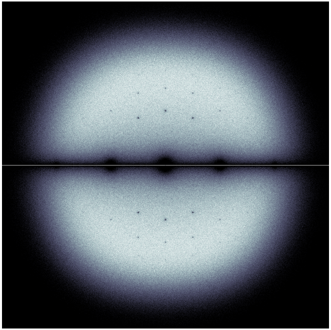

Figure 3: The distribution of eigenvalues of a family of matrices whose entries are drawn from a fixed finite population is *asymptotically uniform* in the scaled unit disk. [Terence Tao and Van Vu, 2006 & 2017]. This is visible already by dimension $m = 5$: the holes fill in as $m$ increases, and the real eigenvalues become negligible.

# Why we're interested

- Allows to study common properties of discrete structured matrices
- investigate extreme possibilities by exhaustive computation
- New look at some old problems (e.g. Hadamard conjecture). Instead of looking for matrices with largest determinant, look for matrices with largest coefficient in the characteristic polynomial? [link] Upper Hessenberg and Toeplitz Bohemians
- Generate new ideas and new conjectures

# Why we *were* interested

Our original motivation was simply the **construction of test problems for eigenvalue solvers**; Steven Thornton has by now solved several *trillion* eigenvalue problems, and uncovered low-dimension instances (10 by 10 matrices with complex entries, 20 by 20 matrices with real entries) for which 2018 Matlab's `eig` routine failed to converge. [Reported to the Mathworks, long since fixed.]

# Other uses

Nick Higham has used Bohemian matrices as a **class to optimize over** to look for improved lower bounds on such things as the growth factor in matrix factoring; Laureano Gonzalez-Vega has looked at *correlation matrices*. [link] David R. Nelson (Harvard) uses ideas like these to study **non-Hermitian quantum mechanics**. (Thanks to Nick Trefethen for making this connection). Matthew Lettington (Cardiff) is interested in these ideas for use with *magic squares* and related topics. Hence this visit!

# Structured matrices

We have used this idea to understand some things about simple matrix structures, such as [link] Skew-symmetric tridiagonal matrices and [link] Upper Hessenberg and Toeplitz Bohemians.

We will talk about some of these results today.

# Differences from more classical problems

If you are used to doing mathematical analysis, then you are typically interested in

- What happens as the dimension $m$ goes to infinity
- What the asymptotic measures or probabilities are
- Generic answers.

But if the dimension $m$ is only, say, 5 or 6, then there are likely to be noticeable effects of small size; and analytical tools are likely to be most helpful after experiments have generated conjectures.

# Skew-symmetric tridiagonal Bohemian matrices

Here is a $7 \times 7$ example of a **complex skew-symmetric tridiagonal** Bohemian matrix with population $[P]$. If $P$ has $\#P$ elements, then the number of such matrices is $\#P^6$.

$$\begin{bmatrix} 0 & u_1 & 0 & 0 & 0 & 0 & 0 \\ -u_1 & 0 & u_2 & 0 & 0 & 0 & 0 \\ 0 & -u_2 & 0 & u_3 & 0 & 0 & 0 \\ 0 & 0 & -u_3 & 0 & u_4 & 0 & 0 \\ 0 & 0 & 0 & -u_4 & 0 & u_5 & 0 \\ 0 & 0 & 0 & 0 & -u_5 & 0 & u_6 \\ 0 & 0 & 0 & 0 & 0 & -u_6 & 0 \end{bmatrix} \tag{1}$$

# A picture

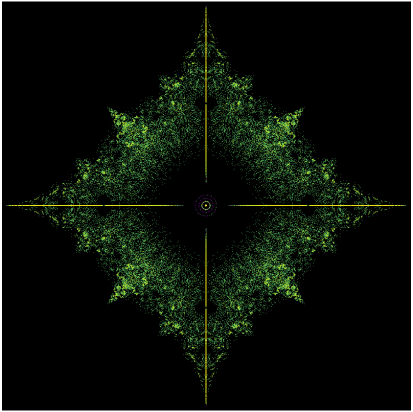

**Figure 4:** Density of eigenvalues of all $4^{14} = 268{,}435{,}456$ fifteen by fifteen skew-symmetric tridiagonal matrices with population $P = [\pm 1, \pm i]$ using $i = \sqrt{-1}$. Note the "rose" in the middle and its symmetries. Computed in Maple (10 seconds).

# Density of eigenvalues in $\mathbb{C}$

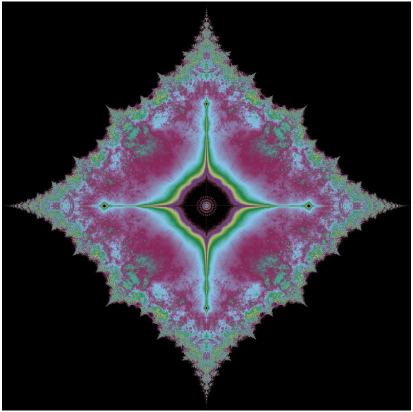

**Figure 5:** Density plot of eigenvalues of all $2^{30} = 1{,}073{,}741{,}824$ skew-symmetric tridiagonal matrices of dimension 31 with population $\{1, i\}$ with $i = \sqrt{-1}$. Hotter colours correspond to higher density. Picture by Aaron Asner.

# For this class, polynomials are almost useful

Figure 6 was computed using eigenvalues of only $2^{14} = 16{,}384$ matrices (thus explaining the mere 10 seconds taken), with $P = [1, i]$ not the $4^{14} > 2.68 \times 10^8$ matrices with $P = [\pm 1, \pm i]$. I might have done even better by using just the 8,146 unique characteristic polynomials of this family. The characteristic polynomials satisfy the recurrence relation

$$p_{k+1} = \lambda p_k + u_n^2 p_{k-1} \tag{2}$$

and $p_0 = 1$, $p_1 = \lambda$. **So there is no need for $-1$ or $-i$.**

Even so, I used Maple's `Eigenvalues` (NAG Library, LAPACK, comparable in speed to Matlab), because degree fifteen polynomials can still be ill-conditioned. [I could have used MPSolve by Bini and Robol.]

# The shape is not circular!

The Tao and Vu result does not apply to this class of matrices. Indeed the eigenvalues are uniformly bounded in $|\lambda| \leq 2$ by Gerschgorin; but why is the shape so angular? And is that edge better modelled by a fractal?

A referee brought to our attention the 2013 *Operators and Matrices* paper by Chandler-Wilde *et al* who explained this shape using so-called Kippenhahn polynomials; we explained it by an older theorem known as the Bendixon–Bromwitch–Hirsch theorem, which uses the spectra of the Hermitian and skew-Hermitian parts of a matrix to give a (sometimes tighter than Gerschgorin) bound. But the apparently fractal edge is still not well understood.

See the paper *Bohemian Matrix Geometry*.

# Our proof

The Bendixon–Bromwitch–Hirsch theorem: Write the matrix A as a sum of its Hermitian and skew-Hermitian parts:

$$\begin{aligned} \mathsf{A} &= \frac{1}{2}\left(\mathsf{A} + \mathsf{A}^*\right) + i\,\frac{1}{2i}\left(\mathsf{A} - \mathsf{A}^*\right) \\ &= \mathsf{H} + i\mathsf{S}\,. \end{aligned} \tag{3}$$

Both H and S are Hermitian so their eigenvalues $\mu_m \le \mu_{m-1} \le \ldots \mu_1$ and $\nu_m \le \cdots \le \nu_1$ are real.

The BBH theorem says that the eigenvalues of A lie inside the box $\mu_m \le \Re(\lambda) \le \mu_1$, $\nu_m \le \Im(\lambda) \le \nu_1$.

To get our needed diamond, one has to rotate by 90°; the details are important but left as an activity for people bored by the rest of this talk! Remember, the population of these skew-symmetric matrices is just $\{1, i\}$; try working with population $\pm 1 \pm i$ instead to start with.

# Fibonacci polynomials

The maximum **characteristic height** occurs when all $u_n = 1$ or (same height) when all $u_n = i$. Call that height $H_m$. The first few characteristic heights are

$$1, 1, 1, 2, 3, 4, 6, 10, 15, 21, 35, 56, 84, 126, 210, 330 \tag{4}$$

and $H_{15} = 330$. We have $F_{m+1}/(m+1) < H_m < F_{m+1}$ where $F_n$ is the $n$th Fibonacci number.

These maximal height polynomials are known as *Fibonacci polynomials* because $p_{m+1} = \lambda p_m + p_{m-1}$.

# Condition Number for polynomials

If

$$p(z) = \sum_{k=0}^{m} c_k \phi_k(z) \tag{5}$$

has its coefficients changed to $c_k(1 + \delta_k)$ then

$$\begin{aligned} |\Delta p(z)| &= \left| \sum_{k=0}^{m} c_k \delta_k \phi_k(z) \right| \\ &\leq \left( \sum_{k=0} |c_k| |\phi_k(z)| \right) \max_{0 \leq k \leq m} |\delta_k| \,. \end{aligned} \tag{6}$$

This is just the triangle inequality (or, if you prefer, a special case of Hölder's inequality). The term in brackets, which we denote $B(p)$, is called the *condition number* for evaluation of the polynomial $p(z)$. Usually the monomial basis $\phi_j(z) = z^j$ is used.

# Condition number of Fibonacci polynomials

The coefficients $c_k$ of Fibonacci polynomials $p_m(\lambda)$ are all positive: thus the condition number is

$$B(p_m)(\lambda) = \sum_{k=0}^{m} c_k |\lambda|^k \,. \tag{7}$$

Therefore the maximum condition number of Fibonacci polynomials on $|\lambda| \leq 2$ (which contains all eigenvalues) occurs at the boundary $|\lambda| = 2$.

Solving $p_{m+1}(2) = 2p_m(2) + p_{m-1}(2)$ needs $r^2 = 2r + 1$ or $r^2 - 2r + 1 = 2$ which has roots $r = 1 \pm \sqrt{2}$. Thus the condition number of Fibonacci polynomials on this interval grows like $(1 + \sqrt{2})^m$.

For $m = 15$ this is about $5 \cdot 10^5$ so double precision would have been enough for these polynomials.

# Use Eigenvalues

But the condition number of a random dimension $m$ eigen*value* problem is only $O(m^2)$, so about 225 for $m = 15$, so even for such a small dimension eigenvalues should be better. For this *real* skew-symmetric matrix the eigenvalue condition numbers are all 1, so it's even better.

# Polynomial condition can grow exponentially

To emphasize: the polynomial evaluation (and therefore rootfinding) condition numbers can grow exponentially with the dimension of the matrix, whereas we expect the eigenvalue condition number to grow only quadratically with the dimension.

# Symmetries

Since the coefficients are real, eigenvalues must occur in conjugate pairs. One can also deduce that $p_{2k+1}(\lambda)$ is odd and $p_{2k}(\lambda)$ is even. Therefore if $\lambda^*$ is an eigenvalue, so is $-\lambda^*$. *These are the only symmetries.* Let's look at that graph again.

## A picture

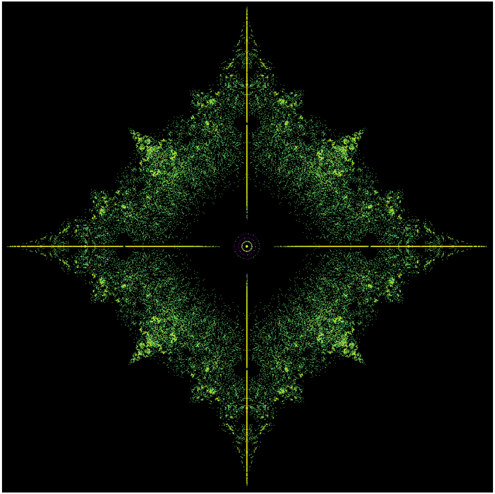

**Figure 6:** Density of eigenvalues of all $4^{14} = 268{,}435{,}456$ fifteen by fifteen skew-symmetric tridiagonal matrices with population $P = [\pm 1, \pm i]$. Note the "rose" in the middle and its symmetries. Computed in Maple (10 seconds).

# Zooming in

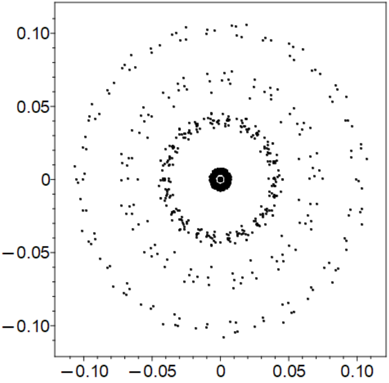


Figure 7: Zooming in on the rosette near zero: counting, we see a 15-fold symmetry in the outer ring, a 13-fold symmetry in the next smaller ring, then an 11-fold symmetry in the next smaller ring. *These symmetries are spurious* and therefore these eigenvalues are rounding errors.

# Nilpotent matrices

The reason for the rosette is *multiple eigenvalues at* 0. Indeed there are *nilpotent* matrices at dimension $m = 2^k - 1$, (and only at these dimensions). I found a recursive formula for a family[1] of such nilpotents: if $s = [u_1, u_2, \ldots, u_{m-1}]$ is the superdiagonal of a nilpotent matrix of dimension $m = 2^k - 1$, then both $[s, 1, i, \mathrm{rev}(s)]$ and $[s, i, 1, \mathrm{rev}(s)]$ are superdiagonals of nilpotent matrices of dimension $m = 2^{k+1} - 1$. Here "rev" means reverse the order of the list.

**Conjecture, experimentally checked to** $m = 31$: these are the **only** nilpotent skew-symmetric tridiagonal Bohemian matrices with population $\{1, i\}$. [This *ought* to be easily provable, but I failed on my first try, then got distracted.]

[1] I made a terrible pun about this family, too. Don't say you weren't warned.

# Jordan structure

It turns out that there is only one Jordan block for these nilpotent matrices, and the matrix $Q$ transforming these Bohemian matrices to Jordan form $AQ = QJ$ is such that it resembles a Sierpinski gasket (it is also Bohemian, as is $Q^{-1}$, so these matrices have *rhapsody*). This "Sierpinski"-ness is a kind of coincidence, considering what I will talk about next.

# One matrix Q when $m = 127$

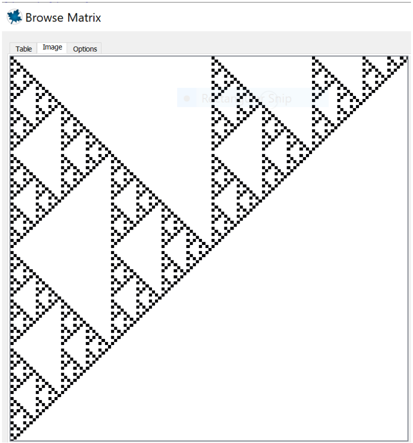


**Figure 8:** The structure of Q for one nilpotent A with $AQ = QJ$, for dimension $m = 127$. The nonzero entries of Q, pictured here simply as black squares, are $\pm 1$ and $\pm i$.

# A cleaner image

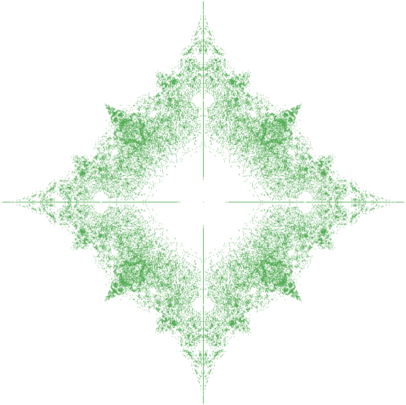

Figure 9: Computing and solving the characteristic polynomials removes the spurious rosette. It takes five times as long, in Maple, however.

# See the Maple Transactions paper

For more details, see *https://doi.org/10.5206/mt.v1i2.14360*. I'd like to move on to another class, which also mixes eigenvalues and polynomial computation.

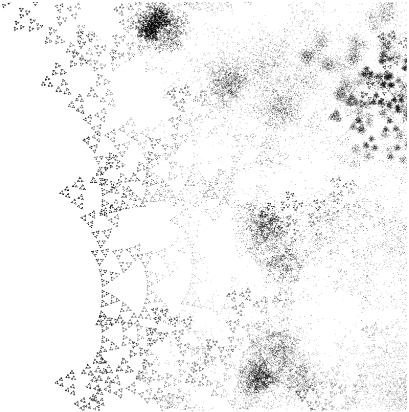

Figure 10: upper Hessenberg Toeplitz, $-1$ subdiagonal, zero diagonal, population cube roots of unity, dimension $m = 13$, all 531,441 matrices, zoomed in on an edge.

# Another population

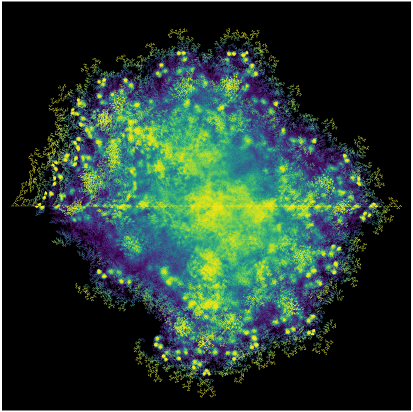

**Figure 11:** Population $\{-1, i, 1\}$, dimension $m = 15$, eigenvalues of all 4,782,969 UHTZD matrices. $-3 \leq x \leq 3$, $-3 \leq y \leq 3$

# Left edge

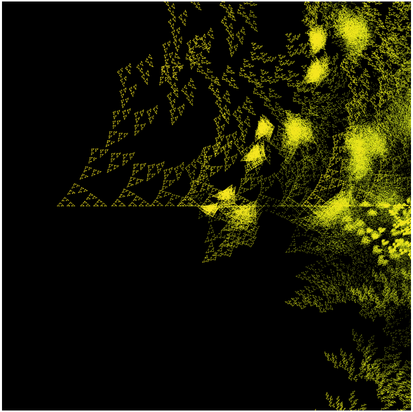

**Figure 12:** Population $\{-1, i, 1\}$, dimension $m = 15$, eigenvalues of all 4,782,969 UHTZD matrices. Zoomed in a bit to $-3 \le x \le -2$, $-1/2 \le y \le 1/2$

# Toeplitz matrix fun

Banded Toeplitz matrices are surprisingly “easy” to understand now (after work of Toeplitz, Szegő, Kac, Widom, Wiener, Schmidt & Spitzer, Böttcher et al., and many others).

$$\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & 0 & 0 \\ a_{-1} & a_0 & a_1 & a_2 & a_3 & 0 \\ a_{-2} & a_{-1} & a_0 & a_1 & a_2 & a_3 \\ 0 & a_{-2} & a_{-1} & a_0 & a_1 & a_2 \\ 0 & 0 & a_{-2} & a_{-1} & a_0 & a_1 \\ 0 & 0 & 0 & a_{-2} & a_{-1} & a_0 \end{bmatrix}$$

This matrix has “symbol” $\frac{a_{-2}}{z^2} + \frac{a_{-1}}{z} + a_0 + a_1 z + a_2 z^2 + a_3 z^3$. In general (for infinite dimension) it’s a Laurent series; for banded matrices, a Laurent polynomial.

# Bounds and patterns II

Theorem: (Toeplitz) The eigenvalues of an infinite-dimensional Toeplitz operator are related to* the image of the unit circle under the symbol: $a(e^{i\theta})$.

* Ok so I am not telling the whole story here. Which infinite matrix? And what about winding numbers? And for which class of symbols (functions) is this true for?

**Important Note**: The eigenvalues of finite-dimensional truncations of Toeplitz matrices do *not* necessarily converge to the spectrum of the corresponding infinite-dimensional Toeplitz operators (but their pseudospectra [link] do).

Another Theorem (Schmidt & Spitzer 1963): The eigenvalues of finite-dimensional *banded* Toeplitz matrices converge to semialgebraic curves (that can be determined by a simple algebraic computation) defined by the symbol.

# Our theorem, needed to explain the Sierpinski structure

The Schmidt–Spitzer curves (and therefore, we believe, eigenvalues) of finite-dimensional upper Hessenberg Toeplitz matrices converge to analogous computable curves defined by roots of convergent series.

This convergence allows us to explain the Sierpinski-like fractal structures in the Bohemian eigenvalue density plots.

# Eigenvalues of One Toeplitz matrix

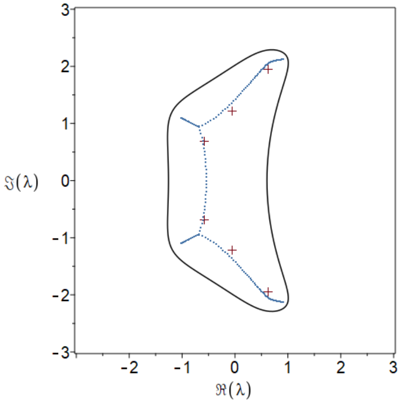


**Figure 13:** Eigenvalues of a single dimension $m = 6$ upper Hessenberg zero-diagonal Toeplitz matrix with entries from $\{-1,0,1\}$. The black curve is the image of the unit circle under the symbol; the dotted blue curve is the Schmidt–Spitzer curve for the infinite-dimensional banded Toeplitz matrix.

# The Schmidt–Spitzer semialgebraic curves

The curves are defined by *equal-magnitude* values of the so-called "symbol": $a(z) = a(e^{i\theta}z) = \lambda$. These are Laurent polynomials, so finding the zeros is just univariate polynomial rootfinding of $a(z) - a(e^{i\theta}z) = 0$, given $\theta$. However, $\lambda$ is in the curve *if and only if* the two equal-magnitude roots are the $q$th and $q+1$st smallest magnitude roots, where $q$ is the order of the pole in the Laurent polynomial (here $q = 1$). Combinatorics and complex analysis both!

Look at ToeplitzExperiments.maple

# What did we prove?

For *upper Hessenberg* matrices, a Laurent *polynomial* symbol

$$a(z) = -\frac{1}{z} + a_0 + a_1 z + \cdots + a_m z^m$$

is not very different to a (finite pole) Laurent *series* because the similarity transform by the diagonal matrix $D = \mathrm{diag}(1,\rho,\rho^2,\ldots)$ shows that the series

$$a(z) = -\frac{\rho}{z} + a_0 + a_1 \frac{z}{\rho} + \cdots + a_m \frac{z^m}{\rho^m} + \cdots$$

converges absolutely and uniformly for $|z| < \rho$ (where $|a_k| \leq B$ because Bohemian and so geometric). Everything follows from classical theorems afterwards: the equal-magnitude curves converge. In practice, they converge *rapidly* for the examples we tried.

# What does the theorem explain?

Using that theorem, we can look at upper Hessenberg Toeplitz Bohemian matrices with, say, a population with three elements. Then increasing the dimension by 1 gives us one new term in the symbol—$a_m$—which can have one of three values; this gives three new matrices and thus three new eigenvalues for each old eigenvalue, and *moreover these eigenvalues have to lie close to the semialgebraic curve from before*. This explains the “Sierpinski gasket” look of these images.

This is the *first* such explanation of the appearance of a fractal in a Bohemian context.

# Unsolved problems

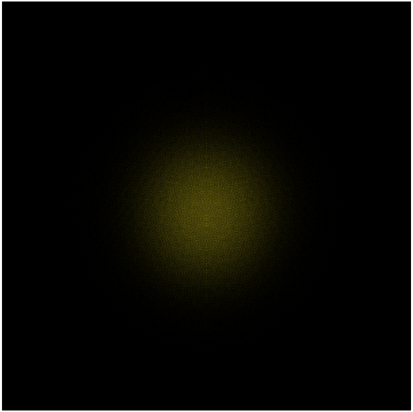

Figure 14: (Stupidly computed) eigenvalues of all $3^{12} = 531{,}441$ dimension $m = 13$ zero-diagonal circulant matrices with population $-1$, $i$, and 1. Why does this density plot look as it does? Circulant matrices are related to magic squares...

# More on Bohemians

You can find an older version of this talk at [YouTube link] a video on my YouTube channel.

You can find a related talk at

[YouTube link]"Skew Symmetric Tridiagonal Bohemians"

The (Maple Transactions!) papers that talk refers to are

[link] What can we learn from Bohemian Matrices? https://doi.org/10.5206/mt.v1i1.14039

and

[link] Skew-symmetric tridiagonal Bohemian matrices https://doi.org/10.5206/mt.v1i2.14360

See also chapter 5 of [link] the online version of my New Book, *Computational Discovery on Jupyter*, with Neil Calkin and Eunice Chan (to be published by SIAM physically … next month, perhaps!)

# Thank you

Thank you for listening!

MINISTERIO
DE CIENCIA
E INNOVACIÓN

This work was partially supported by NSERC grant RGPIN-2020-06438, and partially supported by the grant PID2020-113192GB-I00 (Mathematical Visualization: Foundations, Algorithms and Applications) from the Spanish MICINN. I also thank CUNEF Universidad for financial support.

This visit to Cardiff is funded in part by a grant from the Heilbronn Institute for Mathematical Research.