The Circular Law and the Probability that a Random Matrix Has k Real Eigenvalues

Alan Edelman
Department of Mathematics
University of California
Berkeley, California 94720
edelman@math.berkeley.edu

March 23, 1993

Abstract

Let A be an n by n matrix whose elements are independent random variables with standard normal distributions. Girko's controversial circular law states that the distribution of appropriately normalized eigenvalues is asymptotically uniform in the unit disk in the complex plane. We derive the exact distribution of the complex eigenvalues for finite n, from which the circular law for normally distributed matrices may be derived.

Similar methodology allows us to derive a joint distribution formula for the real Schur decomposition of A. Integration of this distribution yields the probability that A has exactly k real eigenvalues. For example, we show that the probability that A has all real eigenvalues is exactly $2^{-n(n-1)/4}$.

1 Introduction

This paper investigates the eigenvalues of a real random n by n matrix of standard normals. We have two questions to explore:

- What is the probability $p_{n,k}$ that exactly k eigenvalues are real?
- Why do the complex eigenvalues when properly normalized roughly fall uniformly in the unit disk as in Figure 1?

Both questions can be answered by first factoring the matrix into some form of the real Schur decomposition, then interpreting this decomposition as a change of variables, and finally performing a wedge product derivation of the Jacobian of this change of variables. We demonstrate the power of these techniques by obtaining exact answers to these questions. Previous investigations of thse issues have lead to inconclusive results [10] and very cumbersome and probably erroneous results [12].

This paper may be thought of as a sequel to [8], where we answered the questions

- What is the expected number $E_n = \sum_k kp_{n,k}$ of real eigenvalues?
- Why do the real eigenvalues when properly normalized roughly fall uniformly in the interval $[-1, 1]$?

In fact, our investigation into $p_{n,k}$ preceded that of [12], but when we saw that $\sum kp_{n,k}$, always had a particularly simple form, we diverted our attention to understanding E_n and related issues. We felt that the derivation of $E_n = \sum kp_{n,k}$ must somehow be simpler than the derivation of the individual $p_{n,k}$.

Random eigenvalue experiments are irresistible given the availability of modern interactive computing packages. Numerical experiments beckon us for theoretical explanation as surely as any experiment in mechanics did centuries ago. If understanding is not sufficient motivation, random matrix eigenvalues arise

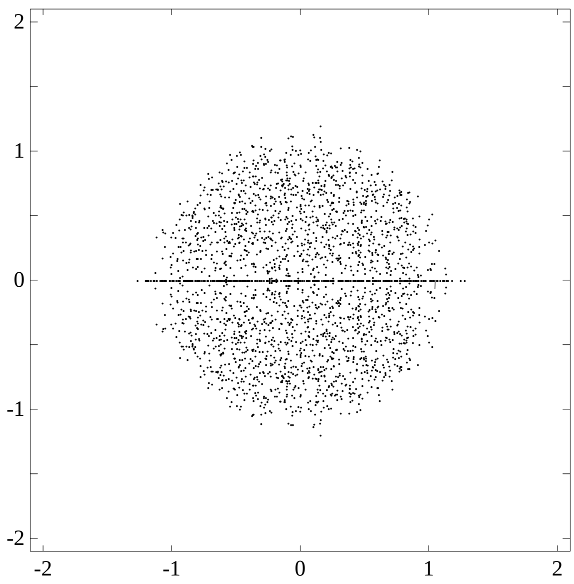

Figure 1: 2500 dots representing normalized eigenvalues of fifty random matrices of size $n = 50$. Clearly visible are the points on the real axis.

in models of nuclear physics [18], in multivariate statistics [20], and in other areas of pure and applied mathematics including numerical analysis [12, 9].[1]

Perhaps we are still at the tip of the iceberg in our understanding of eigenvalues of random matrices. Few papers contain exact distributional results for non-symmetric random matrices. Given an exact distribution formula for a normally distributed matrix, one expects such a formula to be asymptotically valid for matrices of elements of mean 0 and variance 1. This is a central limit effect. At the present time, however, there is no satisfying theory that seems to allow us to make the leap from normally distributed matrices to a wider class.

The contrast between the real and complex eigenvalues of a normally distributed matrix is illustrated in Figure 1. This figure plots the 2500 values of $\lambda/\sqrt{n}$ where λ is an eigenvalue of the random matrix of dimension $n = 50$. Notice that the complex normalized eigenvalues may appear to be roughly uniformly distributed in the unit disk. This is Girko's controversial circular law [11], which states that if the elements of a random matrix are independent with mean 0 and variance 1, then the distribution of the normalized eigenvalues is asymptotically uniformly distributed over the disk. Nobody doubts the overwhelming experimental evidence; the controversy is over whether this law has yet been proved mathematically.

Many have tried to understand Girko's "proofs," without success. This author has also not been able to verify Girko's proof, though we are not sure if the errors are in the translation from the original Russian. We have contacted Girko on this matter and eagerly await his response.

In this paper we rigorously prove a result that is at the same time weaker and stronger than Girko's law. We derive the exact distribution of the complex eigenvalues of a real normally distributed matrix. The limit of the distribution of the normalized eigenvalues is indeed the uniform distribution on the circular disk. This is weaker than Girko's law in the sense that our results only apply for the normal distribution. Intuitively, the normal distribution models all distributions with elements of mean 0 and variance 1, by central limit effects, but this sort of reasoning alone is not rigorous mathematics. Our result is stronger than Girko's law in that it gives exact distributions for finite n.

[1] Most modern papers on numerical linear algorithms, rightly or wrongly, contain numerical experiments on random matrices.

A much simpler case occurs when the random matrix comes from a complex normal distribution, i.e., the real and imaginary parts of each element are independent standard normals. In this case the exact distribution for the eigenvalue, distribution and radius can be found in Ginibre [10] and reported by Mehta [18, p.300] and Hwang [15].

Two recent books [18, 12] report on previous investigations concerning the eigenvalues of a random Gaussian matrix with no symmetry conditions imposed. Our approach improves on the work of Dyson [18, Appendix 35], Ginibre [10], and Girko[12, Theorem 3.5.1]. Ginibre and Girko compute expressions for the joint distribution by diagonalizing $A = X\Lambda X^{-1}$. Girko's form is in terms of the real eigenvalues and the real and imaginary parts of the complex eigenvalues. Dyson computed an expression for the special case when A has all real eigenvalues using the real Schur decomposition [21] $A = QRQ^T$, where R is upper triangular. Girko further expresses the probability c_k that A has k real eigenvalues as the solution to an n by n linear system of equations, where the coefficients are multivariate integrals. Girko's approach is rather cumbersome, and does not seem to be correct, at least in the English version.

After preparation of this manuscript, we learned that Theorem 6.1 was independently discovered by Lehmann and Sommers in [16] who consider the range of problems from the completely symmetric case to the completely antisymmetric case. Our derivation follows a different viewpoint.

Our approach is simplified by directly considering the real Schur decomposition [21] even when the matrix does not have all real eigenvalues. It is well known to researchers in matrix computations that orthogonal decompositions are of great value [13]. It is perhaps not surprising that such decompositions are also appropriate for random matrix theory.

2 Main results

Probability of exactly k real eigenvalues

The probability $p_{n,k}$ that a random A has k real eigenvalues has the form $r + s\sqrt{2}$, where r and s are rational. In particular, the probability that a random matrix has all real eigenvalues is

$$p_{n,n} = 1/2^{n(n-1)/4}.$$

Cor 7.1 and 7.2

In the tables below, we list the probabilities for $n = 1, \ldots, 9$ both exactly and to five decimal places.

n	k	$p_{n,k}$	
1	1	1	1
2	2	$\frac{1}{2}\sqrt{2}$	0.70711
	0	$1 - \frac{1}{2}\sqrt{2}$	0.29289
3	3	$\frac{1}{4}\sqrt{2}$	0.35355
	1	$1 - \frac{1}{4}\sqrt{2}$	0.64645
4	4	$\frac{1}{8}$	0.125
	2	$-\frac{1}{4} + \frac{11}{16}\sqrt{2}$	0.72227
	0	$\frac{9}{8} - \frac{11}{16}\sqrt{2}$	0.15273
5	5	$\frac{1}{32}$	0.03125
	3	$-\frac{1}{16} + \frac{13}{32}\sqrt{2}$	0.51202
	1	$\frac{33}{32} - \frac{13}{32}\sqrt{2}$	0.45673
6	6	$\frac{1}{256}\sqrt{2}$	0.00552
	4	$\frac{271}{1024} - \frac{3}{256}\sqrt{2}$	0.24808
	2	$-\frac{271}{512} + \frac{107}{128}\sqrt{2}$	0.65290
	0	$\frac{1295}{1024} - \frac{53}{64}\sqrt{2}$	0.09350

n	k	$p_{n,k}$	
7	7	$\frac{1}{2048}\sqrt{2}$	0.00069
	5	$\frac{355}{4096} - \frac{3}{2048}\sqrt{2}$	0.08460
	3	$-\frac{355}{2048} + \frac{1087}{2048}\sqrt{2}$	0.57727
	1	$\frac{4451}{4096} - \frac{1085}{2048}\sqrt{2}$	0.33744
8	8	$\frac{1}{16384}$	0.00006
	6	$-\frac{1}{4096} + \frac{3851}{262144}\sqrt{2}$	0.02053
	4	$\frac{53519}{131072} - \frac{11553}{262144}\sqrt{2}$	0.34599
	2	$-\frac{53487}{65536} + \frac{257185}{262144}\sqrt{2}$	0.57131
	0	$\frac{184551}{131072} - \frac{249483}{262144}\sqrt{2}$	0.06210
9	9	$\frac{1}{262144}$	0.00000
	7	$-\frac{1}{65536} + \frac{5297}{2097152}\sqrt{2}$	0.00356
	5	$\frac{82347}{524288} - \frac{15891}{2097152}\sqrt{2}$	0.14635
	3	$-\frac{82339}{262144} + \frac{1345555}{2097152}\sqrt{2}$	0.59328
	1	$\frac{606625}{524288} - \frac{1334961}{2097152}\sqrt{2}$	0.25681

Joint Eigenvalue Density given k real eigenvalues
The joint density of the ordered real eigenvalues λ_j and ordered complex eigenvalue pairs $x_j \pm iy_j$, $y_j > 0$ given that A has k real eigenvalues is

$$\frac{2^{l-n(n+1)/4}}{\prod_{i=1}^{n}\Gamma(i/2)}\Delta \exp\left(\sum(y_i^2 - x_i^2) - \sum\lambda_i^2/2\right)\prod \mathrm{erfc}(y_i\sqrt{2}),$$

where Δ is the magnitude of the product of the differences of the eigenvalues of A. Integrating this formula over the λ_j, x_j and $y_j > 0$ gives $p_{n,k}$. Theorem 6.1

Density of Non-real Eigenvalues
The density of a random complex eigenvalue of a normally distributed matrix is

$$\rho_n(x,y) = \sqrt{2/\pi}\, y e^{y^2-x^2}\mathrm{erfc}(y\sqrt{2})e_{n-2}(x^2+y^2),$$

where $e_n(z) = \sum_{k=0}^{n} z^k/k!$. Integrating this formula over the upper half plane gives half the expected number of non-real eigenvalues. Theorem 6.2

Notice the factor y in the density indicating a low density near the real axis. Readers may notice the white space immediately above and below the real axis in Figure 1. We think of the real axis as attracting these nearly real eigenvalues.

The Circular Law
Let z denote a non-real eigenvalue normalized by dividing by $\sqrt{n}$. As $n \to \infty$, z is uniformly distributed in the disk $|z| < 1$. Furthermore, as $n \to \infty$, each eigenvalue is almost surely non-real. Theorem 6.3

It is tempting to believe that the eigenvalues are almost surely not real, because the real line is a set of Lebesgue measure zero in the plane. However this reasoning is not correct. Indeed it is not true for finite n. When $n = 2$, we see that the probability of real eigenvalues is greater than non-real eigenvalues. The error arises because of the false intuition that the density of the eigenvalues is absolutely continuous with respect to Lebesgue measure.

3 Random Determinants, Permanents, Derangements, and Hypergeometric Functions of Matrix Argument

This section may be read independently from the rest of the paper. Here we relate random determinants with the theory of permanents and hypergeometric functions of matrix argument. We have not seen this connection in the literature.

We remind the reader that the permanent function of a matrix is similar to the usual definition of the determinant except that the sign of the permutation is replaced with a plus sign:

$$\mathrm{per}(A) = \sum_{\pi} a_{1\pi(1)}a_{2\pi(2)}\ldots a_{n\pi(n)}.$$

Generally the permanent is considerably more difficult to compute than the determinant.

Hypergeometric functions of a matrix argument are less familiar than permanents. They arise in multivariate statistics and more recently in specialized fields of harmonic analysis. Unlike the matrix exponential, for example, these functions take matrix arguments but yield scalar output. They are much more complicated than merely evaluating a scalar hypergeometric function at the eigenvalues of the matrix. Readers unfamiliar with the theory may safely browse or skip over allusions to this theory. Those wanting to know more should consult [20, Chapter 7].

Theorem 3.1 *If A is a random matrix with independent elements of mean 0 and variance 1, and if z is any scalar (even complex!) then*

$$\mathbf{E}\det(A^2+z^2I)=\mathbf{E}\det(A+zI)^2=\mathrm{per}(J+z^2I)=n!e_n(z^2)=n!\ {}_1F_1(-1;\frac{1}{2}n;-\frac{1}{2}z^2I)$$

where $\mathbf{E}$*denotes the expectation operator, J is the matrix of all ones,*

$$e_n(x)=\sum_{k=0}^{n}\frac{x^k}{k!}$$

is the truncated Taylor series for $\exp(x)$*, "per" refers to the permanent function of a matrix, and the hypergeometric function has a scalar multiple of the identity as argument.*

The alert reader might already detect our motivation for Theorem 3.1. It is the source of the term $e_{n-2}(x^2+y^2)$ in the density of the non-real eigenvalues that will appear again in Theorem 6.2.

Before we prove this theorem from a set of lemmas to follow, we remark on two special elementary cases. Taking $z=0$ is particularly simple. When $z=0$, the theorem states that the expected value of the determinant of A^2 is $n!$, and of course the permanent of a matrix of all ones is $n!$. Expanding and squaring the determinant of A^2, we see cross terms have expected value 0 and the other $n!$ squares have expected value 1.

Now consider $z^2=-1$. The theorem states that $\mathbf{E}(\det(A^2-I))=n!\sum_{k=0}^{n}(-1)^k/k!$, an expression which we recognize as D_n, the number of derangements on n objects, i.e. the number of permutations on n objects that leave no object fixed in place. Many combinatorics books point out that $\mathrm{per}(J)=n!$ and $\mathrm{per}(J-I)=D_n$ ([22, p. 28], [2, p. 161], and [19, p. 44]), but we have not seen a general formula for $\mathrm{per}(\lambda I-J)$, the *permanental characteristic polynomial* of J, in such books.

Theorem 3.1 is a synthesis of the following lemmas:

Lemma 3.1 *Let A,B, and C be matrices whose elements are the indeterminants a_{ij},b_{ij} and c_{ij}, respectively. Any term in the polynomial expansion of* $\det(A^2+B+C)$ *that contains a b_{ij} does not contain an a_{ij}.*

Proof Let X_{ij} denote the matrix obtained from A^2+B+C by removing the ith row and the jth column. Then

$$\det(A^2+B+C)=\pm b_{ij}\det(X_{ij})+\text{ terms independent of } b_{ij}.$$

Since a_{ij} only appears in the ith row or jth column of A^2, the lemma is proved. □

Corollary 3.1 *Let A be a matrix of independent random elements of mean 0, and let x and y be indeterminants. Then*

$$\mathbf{E}\det((A+xI)^2+y^2I))=\mathbf{E}\det(A^2+(x^2+y^2)I).$$
$$=\mathbf{E}\det\left(A+\sqrt{x^2+y^2}I\right)^2.$$

Proof We use Lemma 3.1 by letting $B=2xA$ and $C=(x^2+y^2)I$. Since,

$$\mathbf{E}\det((A+xI)^2+y^2I)=\mathbf{E}\det(A^2+2xA+(x^2+y^2)I),$$

it follows that any term multiplying $b_{ij}=2xa_{ij}$ in the expansion of the determinant is independent of a_{ij} and so has expected value 0. Therefore the term $2xA$ makes no contribution to the expected determinant, and may be deleted. The second equality may be verified similarly. □

Lemma 3.2 *If A is a random matrix of independent random elements of mean 0 and variance 1, then*

$$\mathbf{E}\det(A+zI)^2=\mathrm{per}(J+z^2I).$$

Proof Expand the determinant of $A + zI$ into its $n!$ terms and square. The expected value of any cross terms must be 0. The expected value of the square of any term is $(1 + z^2)^{s(\pi)}$, where π is the permutation associated with the term and $s(\pi)$ is the number of elements unmoved by π. The sum over all permutations is clearly equal to $\mathrm{per}(J + z^2 I)$ because $s(\pi)$ counts the number of elements from the diagonal in the expansion of the permanent. $\square$

Lemma 3.3 *If A is a random matrix with independent normally distributed elements then*

$$\mathbf{E}\det(A + zI)^2 = n!\,{}_1F_1(-1; \frac{1}{2}n; -\frac{1}{2}z^2 I).$$

Proof Let $M = A + zI$ and z a real scalar. Since the quantity of interest is clearly a polynomial in z the assumption that z is real can be later relaxed. The random matrix $W = MM^T$ has a non-central Wishart distribution [20, p. 441] with n degrees of freedom, covariance matrix I, and noncentrality parameters $\Omega = z^2 I$. The moments of the noncentral Wishart distribution were computed in 1955 and 1963 by Herz and Constantine in terms of hypergeometric functions of matrix argument. (See [20, Theorem 10.3.7, p. 447].) In general,

$$\mathbf{E}\det W^r = (\det\Sigma)^r 2^{mr} \frac{\Gamma_m(\frac{1}{2}n + r)}{\Gamma_m(\frac{1}{2}n)} {}_1F_1(-r; \frac{1}{2}n; -\frac{1}{2}\Omega),$$

where Γ_m denotes the multivariate gamma function. In our case, we care interested in $\mathbf{E}\det W$ which simplifies to

$$\mathbf{E}\det W = n!\,{}_1F_1(-1; \frac{1}{2}n; -\frac{1}{2}z^2 I).$$

$\square$

We remark that the identity

$${}_1F_1(-1; \frac{1}{2}n; -\frac{1}{2}z^2 I) = e_n(z^2)$$

can be obtained directly from a zonal polynomial expansion of the hypergeometric function of matrix argument using the homogeneity of zonal polynomials and the value of certain zonal polynomials at the identity matrix. (The key expressions are Formula (18) on p. 237 and Formulas (1) and (2) on p. 258 of [20].) The only partitions that play any role in the sum are those of the form $(1, 1, \ldots, 1, 0, \ldots, 0)$ with at most n ones.

We further remark that in light of Lemma 3.2, the assumption of normality may be relaxed to any distribution with mean 0 and variance 1. Thus we see our first example where the assumption of normality allows us to derive a formula which is correct in more general circumstances.

4 Orthogonal Matrix Decompositions

This section may also be read independently from the remainder of the paper. Here we review the real Schur decomposition in various forms that we will need to perform our calculations. We begin by pointing out that standard numerical linear algebra calculations may be best looked at geometrically with a change of basis.

4.1 Quaternions and Elementary Numerical Linear Algebra

We wish to study orthogonal (unit determinant) similarity transformations of 2 by 2 matrices:

$$M' = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} M \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, \tag{1}$$

where $c = \cos\theta$ and $s = \sin\theta$.

Researchers familiar with numerical linear algebra know that 2 by 2 matrices are the foundation of many numerical calculations.[2] However, we are not aware that the elementary geometry hidden in (1) is ever pointed out.

[2]Developers of LAPACK and its precursors spent a disproportionate amount of time on the software for 2 by 2 and other small matrix calculations.

Consider the following orthogonal basis for 2 by 2 matrices:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad L = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Any 2 by 2 matrix M can be written as $M = \alpha I + \beta J + \gamma K + \delta L$. This basis is a variation of the usual quaternion basis. (The usual quaternion basis is I, J, iK, and iL, where $i = \sqrt{-1}$. It is well known that real linear combinations of these latter four matrices form a noncommutative division algebra.)

In the I, J, K, L basis, it is readily calculated that the similarity in (1) may be expressed

$$\begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \\ \delta' \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & C & -S \\ & & S & C \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}, \tag{2}$$

where $C = \cos 2\theta$ and $S = \sin 2\theta$. In the language of numerical linear algebra,

> a 2 by 2 orthogonal similarity transformation is equivalent to a Givens rotation of twice the angle applied to a vector in $\Re^4$.

The set of matrices for which $\beta = 0$ is exactly the three dimensional space of real 2 by 2 symmetric matrices. Choosing an angle 2θ in the Givens rotation to zero out δ' corresponds to a step of Jacobi's symmetric eigenvalue algorithm. This is one easy way to see the familiar fact that only one angle $\theta \in [0, \pi/2)$ will diagonalize a symmetric 2 by 2 matrix, so long as the matrix is not a constant multiple of the identity.

A case of more concern to us is choosing an angle 2θ in the Givens rotation to set $\gamma' = 0$, in the general nonsymmetric case. Doing so we conclude:

Lemma 4.1 *Any 2 by 2 matrix M is orthogonally similar to a matrix with equal diagonals. If the matrix is not equal to $\alpha I + \beta J$, then there is only one angle $\theta \in [0, \pi/2)$ that transforms M into this form.*

Since the components in the I and J directions are invariant under the Givens rotation (2) we read right off that the trace of the matrix and the difference of the off-diagonal terms of the matrix is invariant under orthogonal similarities. We readily conclude

Lemma 4.2 *Let M be a 2 by 2 matrix with the non-real eigenvalues $x \pm yi$. There is a unique matrix orthogonally similar to M of the form*

$$Z = \begin{pmatrix} x & b \\ -c & x \end{pmatrix}, \quad bc > 0, \quad b \geq c.$$

Given the matrix M and its eigenvalues $x \pm yi$, b and c may be computed from the equations $bc = y^2$ and $b + c = M_{12} - M_{21}$. If M is not of the form $\alpha I + \beta J$, then the orthogonal matrix with determinant 1 is unique up to sign. Notice that adding π to θ just changes the sign of the orthogonal matrix. Adding $\pi/2$ to θ has the effect of interchanging the b and the c.

4.2 Real Schur Decomposition

The real Shur decomposition expresses A as orthogonally similar to an upper quasi-triangular matrix R [13, p.362]. To be precise $A = QRQ^T$ where

$$R = \begin{pmatrix} \lambda_1 & \cdots & R_{1k} & R_{1,k+1} & \cdots & R_{1m} \\ & \ddots & \vdots & \cdots & \cdots & \vdots \\ & & \lambda_k & R_{k,k+1} & \cdots & R_{km} \\ & & & Z_{k+1} & \cdots & R_{k+1,m} \\ & & & & \ddots & \vdots \\ & & & & & Z_m \end{pmatrix}. \tag{3}$$

R is an $n \times n$ matrix $(m = (n+k)/2)$ with blocks

$$R_{ij} \text{ of size } \begin{cases} 1 \text{ by } 1 & \text{if } i \le k, j \le k \\ 1 \text{ by } 2 & \text{if } i \le k, j > k \\ 2 \text{ by } 1 & \text{if } i > k, j \le k \\ 2 \text{ by } 2 & \text{if } i > k, j > k \end{cases}$$

Here R_{jj} is the real eigenvalue λ_j of A when $j \le k$ and R_{jj} is a 2 by 2 block

$$Z_j = \begin{pmatrix} x_j & b_j \\ -c_j & x_j \end{pmatrix}, \quad bc > 0, \quad b \ge c,$$

so that the complex eigenvalues of A are $x_j \pm y_j i$, where $y_j = \sqrt{b_j c_j}$, for $j > k$. Finally, as indicated in (3), R_{ij} is a zero block if $i > j$.

The block structure indicated in (3) is quite powerful in that it allows us to simultaneously handle all the possibilities rather cleanly. All that one must remember is that an index i refers to a size of one or two depending on whether $i \le k$ or $i > k$ respectively.

Since matrices with multiple eigenvalues form a set of Lebesgue measure 0, so long as our probability distribution is absolutely continuous with respect to Lebesgue measure (such as the normal distribution) we may disregard the possibility that A may have multiple eigenvalues. We can then take the Schur form to be unique if we make some further (arbitrary) requirements on R such as $\lambda_1 > \cdots > \lambda_k$, $x_{k+1} > \cdots > x_m$. Similarly the orthogonal matrix Q is generically unique if, for example, we assume the first row of Q is positive. For $j \le k$, it is easy to see by induction that the jth column of Q can be chosen to be of arbitrary sign. From Lemma 4.2, the next two columns of Q are unique up to sign, if the full matrix Q is to have determinant 1. Allowing Q to have determinant ± 1 allows us to simply specify that the first row of Q be generically positive.

4.3 Incomplete Schur Decomposition

Let a matrix A have a non-real eigenvalue pair $x \pm yi$. A numerical algorithm to compute the eigenvalues of A would deflate out these eigenvalues with an appropriate orthogonal transformation. We wish to do the same by performing an incomplete Schur decomposition.

Definition 4.1 *We say that $A = QMQ^T$ is an* **incomplete Schur decomposition** *for the matrix A with non-real eigenvalue pair $x \pm yi$ if*

$$A = Q \left(\begin{array}{cc|c} x & b & \\ -c & x & W \\ \hline & 0 & A_1 \end{array} \right) Q^T, \quad b \ge c, \ y = \sqrt{bc},$$

where A_1 is an $n-2$ by $n-2$ matrix, and Q is an orthogonal matrix that is the product of two Householder reflections as specified below.

The form of Q is important. We know that generically there is a unique n by 2 matrix H with positive first row for which

$$H^T A H = \begin{pmatrix} x & b \\ -c & x \end{pmatrix} \quad b \ge c.$$

Numerical linear algebra textbooks [13] describe how to construct Householder reflections Q_1 and Q_2 such that $Q_2 Q_1 H$ is the first two columns of the identity matrix. We take our Q in the definition above to be the matrix $Q_2 Q_1$. Notice that H is the first two columns of Q^T.

This decomposition will be used in the derivation of the circular law and complex eigenvalue density of normally distributed matrices.

5 Jacobian Computation

The symmetries built into the orthogonal decompositions allow us to compute the Jacobians of the various factorizations that are of interest to us.

We begin with a lemma that is familiar to numerical analysts who discuss the discrete Laplacian operator in more than one dimension:

Lemma 5.1 *Let X be an m by n matrix. Define the linear operator*

$$\Omega(X) = XA - BX,$$

where A and B are fixed square matrices of dimension m and n respectively. If λ_A is an eigenvalue of A and λ_B is an eigenvalue of B, then $\lambda_A - \lambda_B$ is an eigenvalue of the operator Ω.

Proof We remark that the operator Ω can be represented in terms of the Kronecker product as

$$\Omega = A^T \otimes I - I \otimes B.$$

If v_A is a left eigenvector of A, and v_B is a right eigenvector of B, then $v_B v_A^T$ is an eigenvector of Ω. So long as A and B have a full set of eigenvectors, we have accounted for all of the eigenvalues of Ω. This restriction is not needed by continuity. $\square$

We proceed to compute the Jacobians. Matrices and vectors of differential quantities are in bold face Roman letters so as to distinguish them from the notation for Lebesgue measure. Exterior products of differential quantities are expressed either in math italics or in wedge product notation. Math italics denote the wedge product over the independent elements of dX without regard to sign. Therefore if X is arbitrary, $\mathbf{dX}$ denotes the matrix with elements dx_{ij} and dX denotes the natural element of integration in $\Re^{n^2}$ which could also be expressed as $\wedge_{ij} dx_{ij}$ or $\wedge \mathbf{dX}$. (A wedge without range also means the exterior product over the independent entries.) If X is diagonal, then $dX = dx_{11} \dots dx_{nn}$ while if X is symmetric or upper triangular (antisymmetric), dX is a wedge product over the $n(n+1)/2$ $(n(n-1)/2)$ independent elements.

If Q is orthogonal, then the matrix $\mathbf{dH} = Q^T \mathbf{dQ}$ is antisymmetric, so

$$dH = \bigwedge_{i>j} q_i^T dq_j$$

is the natural element of integration (for Haar measure) over the space of orthogonal matrices.

5.1 Real Schur Decomposition

The Schur decomposition given in (3), indeed any matrix factorization, may be thought of as a change of variables. Let $l = (n-k)/2$ denote the number of complex conjugate eigenvalue pairs, and R^U denotes the strictly upper triangular part of R. (By this we mean R with the λ_i and Z_i replaced with zeros.) The n^2 independent parameters of A are expressed in terms of Q, $\Lambda = (\lambda_i)_{i=1}^k$, $Z = (Z_i)_{i=k+1}^m$ and R^U. For a matrix with k real eigenvalues and $l = (n-k)/2$ complex conjugate pairs, the n^2 independent parameters are found in the new variables as follows:

	parameters
Q	$n(n-1)/2$
Λ	k
Z	$3l$
R^U	$n(n-1)/2 - l$
A	n^2

To obtain the Jacobian of this change of parameters, we express dA in terms of $dH dR^U d\Lambda dZ$:

Theorem 5.1 *Let A be an $n \times n$ matrix written in real Schur form $A = QRQ^T$. The Jacobian of the change of variables is*

$$dA = 2^l \Delta_0 \prod_{i>k} (b_i - c_i)(dH\ dR^U\ d\Lambda\ dZ),$$

where

$$\Delta_0 = \prod_{i>j} |\lambda(R_{ii}) - \lambda(R_{jj})|.$$

Here Δ_0 denotes the absolute product of all the differences of an eigenvalue of R_{ii} and an eigenvalue of R_{jj}, where $i > j$. Every distinct pair of eigenvalues of A appears as a term in Δ_0 except for complex conjugate pairs. For reference,

$$dR^U = \bigwedge_{j>i} dR_{ij}$$

an exterior product over $n(n-1)/2 - 2l$ differentials of the strictly upper triangular part of R;

$$d\Lambda = d\lambda_1 \dots d\lambda_k;$$

and

$$dZ = \bigwedge_{j=k+1}^{m} dZ_j,$$

a product over $3l$ differentials.

Proof Since $A = QRQ^T$, we know that

$$\begin{aligned} \mathbf{dA} &= Q\mathbf{dR}Q^T + \mathbf{dQ}RQ^T + QR\mathbf{dQ}^T \\ &= Q(\mathbf{dR} + Q^T\mathbf{dQ}R + R\mathbf{dQ}^T Q)Q^T. \end{aligned}$$

Let

$$\mathbf{dH} = Q^T\mathbf{dQ}$$

which is antisymmetric and let

$$\mathbf{dM} = \mathbf{dR} + \mathbf{dH}R - R\mathbf{dH} = Q^T\mathbf{dA}Q.$$

It is evident that $dA = dM$ since the orthogonal matrices make no contribution to the Jacobian. Our goal is to compute

$$dM = \bigwedge_{i>j} dM_{ij} \bigwedge_{i=j} dM_{ij} \bigwedge_{i<j} dM_{ij}.$$

We use the same block decomposition for H and M as we did for R. We begin with the most complicated case. If $i > j$, we have that

$$\begin{aligned} \mathbf{dM}_{ij} = &\ \mathbf{dH}_{ij}R_{jj} - R_{ii}\mathbf{dH}_{ij} + \\ &\sum_{k<j} \mathbf{dH}_{ik}R_{kj} - \sum_{k>i} R_{ik}\mathbf{dH}_{kj}. \end{aligned} \tag{4}$$

The quantities in Equation (4) in bold face are matrices of differential quantities with either one or two rows and one or two columns. Following our notation convention, dM_{ij} denotes the exterior product of the one, two, or four independent differential quantities in $\mathbf{dM}_{ij}$. Notice that the $\mathbf{dH}_{ik}$ and $\mathbf{dH}_{kj}$ inside the summations have first index greater than i or second index smaller than j. Therefore, if we order the blocks by decreasing i and then increasing j, we see that

$$\bigwedge_{i>j} dM_{ij} = \bigwedge_{i>j} \wedge(\mathbf{dH}_{ij}R_{jj} - R_{ii}\mathbf{dH}_{ij}); \tag{5}$$

the differentials in the summation play no role.

Lemma 5.1 states that

$$\wedge(\mathbf{dH}_{ij}R_{jj} - R_{ii}\mathbf{dH}_{ij}) = \prod |\lambda(R_{ii}) - \lambda(R_{jj})| dH_{ij}.$$

There are either one, two,or four multiplicands in the product. Since R_{ii} and R_{jj} are either 1 by 1 or 2 by 2 matrices, we can explicitly write the product as

$$\prod |\lambda(R_{ii}) - \lambda(R_{jj})| = \begin{cases} |\lambda_j - \lambda_i| & \text{if } 1 \le j < i \le k \\ (\lambda_j - x_i)^2 + y_i^2 & \text{if } 1 \le j \le k < i \\ \left((x_j - x_i)^2 + (y_j - y_i)^2\right)\left((x_j - x_i)^2 + (y_j + y_i)^2\right) & \text{if } k \le j < i \end{cases} \tag{6}$$

Putting this all together we have that

$$\bigwedge_{i>j} dM_{ij} = \Delta_0 \bigwedge_{i>j} dH_{ij}.$$

We now consider $i = j$. In the following, we let an ellipsis $(\ldots)$ denote terms in dH_{ij} in which $i \neq j$. Such terms play no further role in the Jacobian. If $i \le k$, then

$$dM_{ii} = d\lambda_i + \ldots.$$

If $i > k$ then dM_{ii} is a bit more complicated. It is easy to see that in this case

$$dM_{ii} = \wedge(\mathbf{dZ}_i + \mathbf{dH}_{ii}Z_i - Z_i\mathbf{dH}_{ii}) + \ldots.$$

Notice that since $\mathbf{dH}$ is antisymmetric, then if $i > k$, $\mathbf{dH}_{ii}$ has the form

$$\begin{pmatrix} 0 & dh_i \\ -dh_i & 0 \end{pmatrix}.$$

It follows that

$$(\mathbf{dZ}_i + \mathbf{dH}_{ii}Z_i - Z_i\mathbf{dH}_{ii}) = \begin{pmatrix} dx_i + (b_i - c_i)dh_i & db_i \\ -dc_i & dx_i + (c_i - b_i)dh_i \end{pmatrix}. \tag{7}$$

With $dZ_i = db_i dc_i dx_i$, we have the exterior product of the elements in (7) is $2(b_i - c_i)dZ_i dh_i$. We therefore learn that

$$\bigwedge_{i=j} dM_{ij} = \prod_{i>k} 2^l (b_i - c_i) d\Lambda dZ \wedge_{i=k+1}^{m} dh_i + \ldots.$$

Finally, it is easy to verify that if $i < j$, $dM_{ij} = dR_{ij} + \ldots$. Therefore

$$\bigwedge_{i<j} dM_{ij} = dR^U + \ldots$$

completing the proof. $\square$

Given the 2 by 2 matrix

$$\begin{pmatrix} x & b \\ -c & x \end{pmatrix}, \; bc > 0, \; b \ge c,$$

let δ denote $b - c$. Then

Lemma 5.2 *The Jacobian of the change of variables from b and c to δ and y is*

$$db\ dc = \frac{2y}{\sqrt{\delta^2 + 4y^2}}\ dy\ d\delta.$$

Proof

Since

$$bc = y^2 \text{ and } b - c = \delta \ge 0,$$

it follows that

$$bdc + cdb = 2ydy \text{ and } db - dc = d\delta.$$

Therefore $(b + c)dbdc = 2yd\delta dy$. The conclusion is derived from the equation $(b + c)^2 = \delta^2 + 4y^2$. $\square$

5.2 Incomplete Schur Decomposition

The incomplete Schur decomposition is also a change of parameters with counts indicated in the table below.

	parameters
Q	$2n-3$
x, b, c	3
W	$2n-4$
A_1	$(n-2)^2$
A	n^2

Let H denote the first two columns of Q^T as we did in Section 5.2. The set of all possible H is one quarter of the $2n-3$ dimensional Stiefel submanifold of $\Re^{2n}$ [20]. The one quarter arises from that fact that we are assuming that the first row of H contains positive elements. The matrix of differentials $\mathbf{dH}$ is antisymmetric in its first two rows. Thus there is one differential element in the first two rows and $2n-4$ below the first two rows.

The natural element of integration on this submanifold is $dS = \wedge(Q^T\mathbf{dH})$. We use the notation dS to remind us that dS is a (higher dimensional) <u>s</u>urface element of the **<u>S</u>**tiefel manifold.

Theorem 5.2 *Let A be an $n \times n$ matrix written in incomplete Schur form $A = QMQ^T$. The Jacobian of the change of variables is*

$$dA = 2(b-c)\det\left((A_1 - xI)^2 + y^2 I\right)\ (db\ dc\ dx\ dA_1\ dW\ dS).$$

The proof of this theorem is very similar to that of Theorem 5.1, though simpler. The determinant comes from Lemma 5.1 with matrices of dimension 2 and $n-2$.

6 Applications to the Normal Distribution

The Jacobians computed in the previous section may be integrated to compute densities and probabilities. In the case of the real Schur decomposition we obtain the joint density of the eigenvalues conditioned on k eigenvalues being real. We further obtain the probability that k eigenvalues are real. The incomplete Schur decomposition gives us the density of a complex eigenvalue. Its integral is the expected number of complex eigenvalues.

6.1 Applications of the Real Schur Decomposition

In this section, we calculate the joint density of the eigenvalues of A and the probability that A has exactly k real eigenvalues If the elements a_{ij} of A are independent standard normals.

Theorem 6.1 *Let Θ_k denote the set of matrices A with exactly k real eigenvalues. Let $p_{n,k}$ denote the probability that $A \in \Theta_k$. The ordered real eigenvalues of A are denoted $\lambda_i, i = 1, \ldots, k$, while the $l = (n-k)/2$ ordered complex eigenvalue pairs are denoted $x_i \pm y_i\sqrt{-1}, i = k+1, \ldots, m$. Let*

$$c_{n,k} = \frac{2^{2l - n(n+1)/4}}{\prod_{i=1}^{n}\Gamma(i/2)}.$$

The joint distribution of the real and complex eigenvalues given that A has k real eigenvalues is

$$p_{n,k}^{-1} c_{n,k}\Delta_0 \exp\left(\sum(y_i^2 - x_i^2) - \sum\lambda_i^2/2\right)\prod\left[y_i \operatorname{erfc}(y_i\sqrt{2})\right], \tag{8}$$

where Δ_0 is as in Theorem 5.1. The probability that $A \in \Theta_k$ is

$$p_{n,k} = \frac{c_{n,k}}{k!l!}\int_{\substack{x_i \in \Re \\ y_i \in \Re^+ \\ \lambda_i \in \Re}} \Delta_0 \exp\left(\sum\lambda_i^2/2 + \sum(y_i^2 - x_i^2)\right)\prod\left[y_i\operatorname{erfc}(y_i\sqrt{2})\right] d\lambda_1 \ldots d\lambda_k dx_1 \ldots dx_l dy_1 \ldots dy_l. \tag{9}$$

Proof If the elements of A are independent standard normals then the joint probability density for A is

$$(2\pi)^{-n^2/2}\mathrm{etr}(-\frac{1}{2}AA^T)dA,$$

where

$$\mathrm{etr}(X) = \exp(\mathrm{trace}(X)).$$

Theorem 5.1 states that

$$(2\pi)^{-\frac{n^2}{2}}\mathrm{etr}(-\frac{1}{2}AA^T)dA \ =$$

$$(2\pi)^{-\frac{n^2}{2}}2^l\Delta_0\prod_{i>k}(b_i-c_i)\,(dH)\,\left(e^{-\sum\frac{r_{ij}^2}{2}}dR^U\right)\left(e^{-\sum\frac{\lambda_i^2}{2}}d\Lambda\right)\left(e^{-\sum\left(x_i^2+\frac{b_i^2}{2}+\frac{c_i^2}{2}\right)}dZ\right). \qquad (10)$$

The integral of (10) over Q, R^U, Λ and Z (with the restrictions on these variables specified in Section 4.2 to make the Schur factorization unique), counts every matrix in Θ_k once generically. This integral is

$$p_{n,k} = (2\pi)^{-n^2/2}\int_{A\in\Theta_k}\mathrm{etr}(-\frac{1}{2}AA^T)dA,$$

To obtain the joint density of the eigenvalues, we must integrate out all the variables other than the λ_i, the x_i and the y_i. To obtain the probability $p_{n,k}$, we must integrate over all the variables.

(i)The integral over the orthogonal matrices with positive elements in the first row is

$$\int dH = \mathrm{Vol}(O(n))/2^n. \qquad (11)$$

The derivation of the volume of the orthogonal group may be found in [20] to be

$$\mathrm{Vol}(O(n)) = \frac{2^n\pi^{n(n+1)/4}}{\prod_{i=1}^n\Gamma(i/2)}. \qquad (12)$$

The 2^n in (11) represents the volume of that part of the orthogonal group which has positive first row.

(ii) The integral over R^U is $(2\pi)^{n(n-1)/4-l/2}$. (13)

(iii) We make the change of variables from b_i and c_i to δ_i and y_i described in Lemma 5.2, and then we integrate out δ_i. Since $b_i^2+c_i^2=\delta^2+2y^2$, we see that integrating out $d\delta_i$ amounts to computing

$$2ye^{-y_i^2}\int_{\delta_i=0}^{\infty}\frac{\delta_i e^{-\delta_i^2/2}}{(\delta_i^2+4y^2)^{1/2}}d\delta_i.$$

This integral can be obtained from [14, 3.362.2, p. 315] with a change of variables. It equals

$$2\sqrt{2\pi}y_i e^{y_i^2}\mathrm{erfc}(y_i 2^{1/2}). \qquad (14)$$

We use horizontal braces to indicate the source of each term in the combined expression for the joint density of the real and complex eigenvalues:

$$p_{n,k}^{-1}\underbrace{\overbrace{(2\pi)^{-n^2/2}2^l\Delta_0\exp\left(-\sum x_i^2-\sum\lambda_i^2/2\right)}}_{(10)}\underbrace{\overbrace{\frac{\pi^{n(n+1)/4}}{\prod_{i=1}^n\Gamma(i/2)}}}_{(11)\text{ and }(12)}\underbrace{\overbrace{(2\pi)^{n(n-1)/4-l/2}}}_{(13)}\underbrace{\overbrace{2^l(2\pi)^{l/2}\prod[y_ie^{y_i^2}\mathrm{erfc}(y_i\sqrt{2})]}}_{(14)} \quad (15)$$

The $p_{n,k}^{-1}$ is the normalization constant so that the joint density integrates to 1. Therefore (15) is a probability density which simplifies to (8).

Equation (9) is an exact expression for $p_{n,k}$. Notice that we removed the ordering of the variables, and compensated for this by dividing by $k!$ and $l!$. The integral in (9) would separate into univariate integrals if it were not for the connecting term Δ_0.[3]

If we wish to compute $p_{n,k}$ explicitly we must integrate out the x_i, the y_i and the λ_i in (9). We now describe the approach we took to integrate out the x_i and y_i. We postpone discussion of the integration of the λ_i to Section 7.

From (6) we see that Δ_0 is a polynomial in the variables x_i and y_i. We can use the integration formula [14, 3.461.2]

$$\int_{-\infty}^{\infty} x^n e^{-x^2}\,dx = \frac{(n-1)!!}{2^{n/2}}\sqrt{\pi}, \quad n \text{ even} \tag{16}$$

and also

Lemma 6.1

$$\int_0^{\infty} y^{2n+1}\operatorname{erfc}(y2^{1/2})e^{y^2}\,dy = \frac{\Gamma(n+\frac{3}{2})}{\sqrt{\pi}(2n+2)2^{n+1}}\,{}_2F_1(n+1, n+\frac{3}{2}; n+2; \frac{1}{2}) \tag{17}$$

$$= \frac{n!(-1)^n}{2}\left(\sqrt{2}\sum_{k=0}^{n}(-1)^k\frac{(1/2)_k}{k!} - 1\right), \tag{18}$$

where the Pochhammer symbol $(x)_k$ denotes the product $(x)_k = x(x+1)\ldots(x+k-1)$.

Proof The expression in (17) for the integral as a hypergeometric function may be found in [14, 6.286.1]. We did not find this form particularly enlightening so we attempted to find another expression.

We outline our approach for obtaining the expression (18) leaving the details to the reader. Replacing the erfc in the integrand with $\frac{2}{\sqrt{\pi}}\int_{x=y\sqrt{2}}^{\infty} e^{-x^2}dx$ allows us to interchange the order of integration. A change of variables and a careful application of [14, 2.321.2] and standard Gamma integrals completes the task.

It is also possible, though tedious, to ignore the integral and directly show equality of the hypergeometric function expression and (18). The function ${}_2F_1(n, n+\frac{1}{2}; n+1; x)$ is obtained from ${}_2F_1(1, 3/2; 2; x)$ by taking $n-1$ derivatives [1, 15.2.2]. We may also show that

$${}_2F_1(1, 3/2; 2; x) = \frac{2}{x}\left((1-x)^{-1/2} - 1\right)$$

from [1, 15.2.21 and 15.1.8]. $\square$

6.2 Applications of the Incomplete Schur Factorization

We now turn to the distribution of the complex eigenvalues $x+yi$, $y>0$ of the real matrix A. Let $\rho_n(x,y)dxdy$ denote the expected number of complex eigenvalues of a random matrix A in an infinitesimal area $dx\,dy$ of the upper half plane. More rigorously $\rho_n(x,y)$ is the Radon-Nickodym derivative of the measure

$$\mu_n(\Omega) = \int_{\Omega} \rho_n(x,y)dxdy = \mathbf{E}_A\{\text{ Number of eigenvalues of } A \text{ contained in } \Omega\}$$

defined on measurable subsets Ω of the upper half plane.

Theorem 6.2 *The complex eigenvalue density is*

$$\rho_n(x,y) = \sqrt{\frac{2}{\pi}}\,y e^{y^2-x^2}\operatorname{erfc}(y\sqrt{2})e_{n-2}(x^2+y^2),$$

where $e_n(z) = \sum_{k=0}^{n} z^k/k!$.

[3] This reminds us of the electrostatic interactions in the Hamiltonian for a many electron system in quantum mechanics. Without these interactions Schrödinger's equation could be solved exactly.

Proof The techniques used in this proof are very similar to those used in Theorem 6.1. Theorem 5.2 states that

$$(2\pi)^{-\frac{n^2}{2}}\mathrm{etr}(-\frac{1}{2}AA^T)dA \ = \tag{19}$$
$$(2\pi)^{-n^2/2}2(b-c)\det\left((A_1-x)^2+y^2I\right)e^{-(x^2+\frac{b^2}{2}+\frac{c^2}{2})}\,db\,dc\,dx\ \mathrm{etr}(-\tfrac{1}{2}A_1A_1^T)\ dA_1\ \mathrm{etr}(-\tfrac{1}{2}WW^T)dW\ dS.$$

(i) The volume of the Stiefel manifold may be found in [20]. One quarter of the volume of the Stiefel manifold (because of the sign choice on the first row of H) is

$$\frac{(2\pi)^{n-1}}{2\Gamma(n-1)}. \tag{20}$$

(ii) The integral over W is $(2\pi)^{n-2}$. (21)

(iii) Exactly as before, the b and c components transform into (14).
(iv) We recognize the integral of $\det\left((A_1-x)^2+y^2I\right)\mathrm{etr}(-\frac{1}{2}A_1A_1^T)\ dA_1$. It is

$$(2\pi)^{(n-2)^2/2}\mathbf{E}_{A_1}\det\left((A_1-x)^2+y^2I\right).$$

This explains our motivation for Section 3. From Theorem 3.1, we learn that the value of the integral of the terms containing A_1 is

$$(2\pi)^{(n-2)^2/2}\Gamma(n-1)e_{n-2}(x^2+y^2). \tag{22}$$

Combining terms we have that $\rho_n(x,y) =$

$$\underbrace{\overbrace{(2\pi)^{-n^2/2}2e^{-x^2}}}_{(19)}\underbrace{\overbrace{\frac{(2\pi)^{n-1}}{2\Gamma(n-1)}}}_{(20)}\underbrace{\overbrace{(2\pi)^{n-2}}}_{(21)}\underbrace{\overbrace{2\sqrt{2\pi}ye^{y^2}\mathrm{erfc}(y\sqrt{2})}}_{(14)}\underbrace{\overbrace{(2\pi)^{(n-2)^2/2}\Gamma(n-1)e_{n-2}(x^2+y^2)}}_{(22)}. \tag{23}$$

This completes the proof. $\square$

Corollary 6.1 *The expected number of non real eigenvalues of a random matrix A is*

$$c_n = \sum_k (n-k)p_{n,k} = 2\int_{\substack{x\in\Re \\ y\in\Re^+}} \rho_n(x,y)dx\ dy.$$

The factor of 2 counts the complex conjugate pairs. We could proceed by integrating $\rho_n(x,y)$, to compute c_n, but this is not necessary because in [8] we computed

$$c_n = n - \frac{1}{2} - \sqrt{2}\frac{{}_2F_1(1,-1/2;n;\frac{1}{2})}{B(n,1/2)},$$

expressing c_n in terms of a hypergeometric function and an Euler Beta function. Other equivalent expressions for $n-c_n$ may be found in [8]. 28 We now turn to understanding this density in the context of Girko's circular law (or Figure 1) as $n\to\infty$. It is interesting to normalize the eigenvalues by dividing by $\sqrt{n}$. Thus we introduce

$$\hat{x} = x/\sqrt{n}, \quad \hat{y} = y/\sqrt{n}, \quad \hat{\rho}_n(\hat{x},\hat{y}) = n\sqrt{\frac{2n}{\pi}}\hat{y}e^{n(\hat{y}^2-\hat{x}^2)}\mathrm{erfc}(\hat{y}\sqrt{2n})e_{n-2}(n(\hat{x}^2+\hat{y}^2)).$$

$\hat{\rho}(\hat{x},\hat{y})$ is the density of the expected number of normalized eigenvalues in an infinitesimal area around $\hat{x}+i\hat{y}$.

Theorem 6.3 *The density of the normalized eigenvalues converges pointwise to a very simple form as $n \to \infty$:*

$$\lim_{n\to\infty} n^{-1}\hat{\rho}_n(\hat{x},\hat{y}) = \begin{cases} \pi^{-1} & \hat{x}^2+\hat{y}^2<1 \\ 0 & \hat{x}^2+\hat{y}^2>1 \end{cases} \tag{24}$$

Asymptotically, a random eigenvalue is almost surely complex. When normalized by dividing by $\sqrt{n}$, the eigenvalue converges in distribution to a random variable that is uniformly distributed in the unit disk $\hat{x}^2 + \hat{y}^2 < 1$.

Proof The formulas [1, 6.5.34] and [1, 7.1.13] may be used to verify (24). The former states that

$$\lim_{n\to\infty} \frac{e_n(n(\hat{x}^2+\hat{y}^2))}{e^{n(\hat{x}^2+\hat{y}^2)}} = \begin{cases} 1 \text{ for } \hat{x}^2+\hat{y}^2<1 \\ 0 \text{ for } \hat{x}^2+\hat{y}^2>1 \end{cases},$$

while the latter states that

$$\lim_{n\to\infty} \sqrt{n}\hat{y}e^{2n\hat{y}^2}\operatorname{erfc}(\hat{y}\sqrt{2n}) = (2\pi)^{-1/2}.$$

It is routine though tedious to show that the convergence is dominated by an integrable function on the plane. We need the inequality

$$e_n(nz)/\exp(nz) = \frac{1}{n!}\int_{nz}^{\infty} e^{-t}t^n\,dt \le \frac{n^n}{n!}z^n e^{-nz}\frac{z}{z-1}$$

and Stirling's inequality [1, 6.1.38]

$$n! \ge \sqrt{2\pi n}n^n e^{-n}.$$

We leave the details to the reader. By the Lebesgue dominated convergence theorem, we achieve convergence in distribution. □

7 Relationship with Symmetric Random Matrices

We now turn to the task of integrating the λ_i variables in (9). The relevant part of the integrand containing λ_i has the form

$$\prod q(\lambda_i)\prod|\lambda_i-\lambda_j|e^{-\sum\lambda_i^2/2}, \tag{25}$$

where $q(\lambda_i)$ is some polynomial function of λ_i originating in Δ_0.

This form is closely related to the joint density of the k real eigenvalues of a random k by k symmetric matrix $S = (A + A^T)/2$, where the elements of A are normally distributed. In the physics literature the probability distribution for S is known as the Gaussian orthogonal ensemble. The joint probability density for the eigenvalues $\lambda_1 \ge \ldots \ge \lambda_k$ of S is well known [18, 4, 20]:

$$\frac{2^{-k/2}}{\prod_{i=1}^k \Gamma(i/2)}\prod|\lambda_i-\lambda_j|e^{-\sum\lambda_i^2/2}.$$

Therefore the integral of (25) is essentially an expectation for the determinant of a polynomial in S. To be more precise:

Lemma 7.1

$$\frac{1}{k!}\int_{\lambda_i\in\Re}\prod q(\lambda_i)\prod|\lambda_i-\lambda_j|e^{-\sum\lambda_i^2/2} = \left\{2^{k/2}\prod_{i=1}^k\Gamma(i/2)\right\}\mathbf{E}_S\det(q(S)),$$

where E denotes expectation with respect to the symmetric random matrices defined above.

Dividing by $k!$ allows us to symmetrize the integrand so that we need not assume the λ_i are ordered.

If q is a polynomial with rational coefficients, then $\mathbf{E}\det(q(S))$ must be rational since all the moments of a standard normal are rational. We suspect that a better understanding of $\mathbf{E}\det(q(S))$ is possible, but we are not even aware of an exact formula for $\mathbf{E}\det S^2$. Such an understanding would help us simplify our formula for $p_{n,k}$. For now we are content with this application of Lemma 7.1.

Theorem 7.1

$$p_{n,k} = d_{n,k}^{-1} \int_{\substack{x_i \in \Re \\ y_i \in \Re^+}} \mathbf{E}_S \left\{ \prod_{i=1}^{k} \det\left((S - xI)^2 + y^2 I\right) \right\} \Delta_{xy} e^{\sum (y_i^2 - x_i^2)} \prod \left[y_i \mathrm{erfc}(y_i \sqrt{2}) \right] dx_1 \ldots dx_l dy_1 \ldots dy_l,$$

where

$$d_{n,k} = 2^{\frac{n(n-3)}{4} + k/2} l! \prod_{i=k+1}^{n} \Gamma(i/2)$$

and Δ_{xy} *denotes the multiplicands in* Δ_0 *that do not include a* λ *term.*

Corollary 7.1 *The probability of all real eigenvalues is* $p_{n,n} = 2^{-n(n-1)/4}$.

Proof When $k = n$ and $l = 0$, $d_{n,n} = 2^{n(n-1)/4}$ and the integrand is simply 1. □

Corollary 7.2 *The probability* $p_{n,k}$ *has the form* $r + s\sqrt{2}$ *where* r *and* s *are rational. Furthermore* r *and* s *may be expressed as an integer fractions with denominator equal to a power of* 2.

Proof The $\mathbf{E}_S$ term in the integrand is a polynomial in x and y with integer coefficients. Before taking expectations, the determinant is a polynomial in x, y, and the S_{ij} with integer coefficients. We recall that the nth moment of the normal distribution is $(n-1)!!$ if n is even and 0 if n is odd. Integrating out the x_i using Equation (16) gives a factor of $pi^{l/2}$ which is canceled by the $\pi^{l/2}$ in $\prod_{k+1}^{n} \Gamma(i/2)$. The integration of the y_i using (18) leads only to rational numbers of the proper form. □

Theorem 7.1 is the basis of our earliest Mathematica program (Apendix A) for computing $p_{n,k}$. We first evaluate $\mathbf{E}_S \prod_{i=1}^{k} \det\left((S - xI)^2 + y^2 I\right)$ by explicity computing the determinant symbollically in terms of the elements of S and the x_j and y_j. Then we replace the various powers of the elements of S with their expectation. We then integrate out the x_i and y_i by replacing powers using formulas (16) and (18).

We then modified the program to save some arithmetic. It is readily shown that a certain tridiagonal matrix with χ distributions on the three diagonals has the same eigenvalue distribution as that of S. Therefore, we may use T in place of S to compute the expectations of the determinants of polynomials in S. This is our current program in Appendix A. The program may be summarized by saying that a polynomial is computed in expanded form, then powers of variaous variables are symbolically replaced with the appropriate moments.

8 Eigenvalue Roulette and Numerical Simulations

Fifty five years ago Littlewood and Offord [17] proposed a game on random polynomials that they hardly had the technology to play:

> Let the reader place himself at the point of view of A in the following situation. An equation of degree 30 with real coefficients being selected by a referee, A lays B even odds that the equation has not more than r real roots. ... What smallest possible value of r will make A safe [from being ruined]?

With the benefit of easy to use interactive computer languages best exemplified by Matlab, it is trivial to play their game or our eigenvalue version of the game that we choose to call "eigenvalue roulette." As an example, when $n = 8$, the "wheel" contains the numbers $0, 2, 4, 6$ and 8 and the wheel may be spun simply by executing

```
>> rand('normal'); a=rand(8); sum(abs(imag(eig(a)))<.0001
```

The `.0001` is an arbitrary numerical tolerance that seems to be unimportant in practice.

As a more substantial check of our results, we executed LAPACK simultaneously on each of 128 processors of the CM-5 supercomputer. The experiments agreed perfectly with the theory.

9 Open Problems

This section contains a number of conjectures that we strongly suspect to be true.

Conjecture 9.1 *Let $\{M_n\}$ be a sequence of n by n random matrices. Let t_n denote the number of real eigenvalues of M_n. Then $\lim_{n\to\infty} t_n/n = \sqrt{2/\pi}$ almost surely.*

Intuitively this states that if we plot k/n against $np_{n,k}$ for $k = 0, \ldots, n$, the points "converge" to the degenerate distribution $\delta(x - \sqrt{2/\pi})$.

Conjecture 9.2 *An explicit formula for $p_{n,k}$ that is easier to compute from than 7.1 may be obtained.*

Our formula allowed us to compute $p_{n,k}$ explicitly when either $k = n$ or $n < 10$, but we have not yet succeeded in computing other probabilities due to the computational complexity of our Mathematica program. We suspect the program can be rewritten so as to compute values for say $n = 10$ or $n = 11$, but ultimately a better formulation will be needed.

Conjecture 9.3 *Girko's circular law for arbitrary random matrices of iid elements of mean 0 and variance 1 may be derived as a corollary of Theorem 6.3 using some kind of central limit theorem.*

Acknowledgements

The author would like to thank Eric Kostlan for many interesting discussions on these and related issues.

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1965.

[2] M. Aigner, *Combinatorial Theory*, Springer Verlag, New York, 1979.

[3] A. Edelman, Eigenvalues and condition numbers of random matrices, *SIAM J. Matrix Anal. Appl. 9* (1988), 543–560.

[4] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, PhD thesis, Department of Mathematics, MIT, 1989.

[5] A. Edelman, The distribution and moments of the smallest eigenvalue of a random matrix of Wishart type, *Linear Alg. Appl. 159* (1991) 55–80.

[6] A. Edelman, On the distribution of a scaled condition number, *Mathematics of Computation*, to appear.

[7] A. Edelman, Random matrix eigenvalues meet numerical linear algebra, *SIAM News 24* (November 1991), 11.

[8] A. Edelman, E. Kostlan, and M. Shub, How many eigenvalues of a random matrix are real?, *J. Amer. Math. Soc.*, to appear, 1994.

[9] A. Edelman, Bibliography of Random Eigenvalue Literature, currently available electronically by anonymous FTP from `math.berkeley.edu` in the directory `/pub/Alan.Edelman`.

[10] J. Ginibre, Statistical ensembles of complex, quaternion and real matrices, *J. Math. Phys. 6* (1965), 440–449.

[11] V.L. Girko, Circular law, *Theory Prob. Appl. 29* (1984), 694–706.

[12] V.L. Girko, *Theory of Random Determinants,* Kluwer Academic Press, Boston, 1990.

[13] G.H. Golub and C.F. van Loan, *Matrix Computations, Second Edition*, Johns Hopkins University Press, Baltimore, 1989.

[14] I.S. Gradshteyn and I.M. Ryzhik *Table of Integrals, Series and Products, Corrected and Enlarged Edition*, Academic Press, New York, 1980.

[15] C.R. Hwang, A brief survey on the spectral radius and the spectral distribution of large random matrices with i.i.d. entries, *Random Matrices and their Applications*, Contemporary Mathematics **50**, American Mathematical Society, Providence, 1986, pp. 145–152.

[16] N. Lehmann and H.-J. Sommers, Eigenvalue statistics of random real matrices, *Phys. Rev. Let. 67* (1991), 941–944.

[17] J.E. Littlewood, A.C. Offord. On the number of real roots of a random algebraic equation, *J. London Math. Soc. 13* (1938), 288–295.

[18] M.L. Mehta, *Random Matrices,* Academic Press, New York, 1991.

[19] H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, 6, Addison-Wesley, 1978.

[20] R.J. Muirhead, *Aspects of Multivariate Statistical Theory,* John Wiley & Sons, New York, 1982.

[21] F.D. Murnaghan and A. Wintner, A canonical form for real matrices under orthogonal transformations, *Proc. Nat. Acad. Sci., 17* (1931), 417–420.

[22] H.J. Ryser, *Combinatorial mathematics,* Carus mathematical monographs; no. 14, John Wiley & Sons, New York, 1963.

[23] H.-J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Spectrum of large random asymmetric matrices, *Phys. Rev. Let., 60* (1988), 1895–1898.

A Mathematica Program to compute $p_{n,k}$

```
(* Mathematica  program to compute the probability that an n x n random Gaussian matrix has k real eigenvalues.
              Alan Edelman -- UC Berkeley 12/92 *)

n=Input["n? "]; k=Input["k? "]; l=(n-k)/2;

(*A k x k Tridiagonal matrix with w on the diag and v off the diag *)
W=DiagonalMatrix[Table[w[i],{i,k}]]; Do[ (W[[i+1,i]]=v[k-i]; W[[i,i+1]]=v[k-i]), {i,k-1}];

(* Determinant involving x[i] and y[i] *)
II=IdentityMatrix[k];
z =If[ l==0 || k==0, 1, Det[(W-II*x).(W-II*x)+ II*y^2]];
z = Expand[Product[z/.x:>x[i]/.y:>y[i],{i,l}]];
z = z /. w[i_]^n_:>0 /; OddQ[n]; z = z /. w[i_]^n_:> (n-1)!!; z = z /. w[i_]->0;
z = z /. v[i_]^n_ :> 0 /; OddQ[n]; z = z /. v[i_]^n_ :> Gamma[(i+n)/2]/Gamma[i/2]; z = z /. v[i_]-> 0;

Do[(   (*Only introduce those x[ii] and y[ii] that we need so far*)
  z = z*y[ii]* Product[  ((x[j]-x[ii])^2+(y[j]-y[ii])^2)((x[j]-x[ii])^2+(y[j]+y[ii])^2), {j,ii+1,l}] ;
  z=Expand[z];
  (* Integrate x[ii] *)
  z = z /. x[ii]^n_ :> 0 /; OddQ[n];  z = z /. x[ii]^n_ :> (n-1)!!/(2^(n/2));  z = z /. x[ii]->0;
  (* Integrate y[ii] *)
  iy[n_]:=n!(-1)^n/2(Sqrt[2] Sum[ (-1)^k Pochhammer[1/2,k]/k!,{k,0,n}]-1);
  z = z /. y[ii]^n_ :> iy[(n-1)/2];  z = z /. y[ii]    :> iy[0];
),{ii,1,l}]
z=Simplify[ z Sqrt[Pi]^l /( 2^(n(n-3)/4+k/2) Product[Gamma[i/2],{i,k+1,n}] l!)];
```