Random $*$-Cosquare Matrices and Self-Inversive Polynomials

Ben Estevez

May 2018

1 Introduction

1.1 The notion of a random matrix:

Throughout both industry and the sciences, one finds matrices and matrix models ubiquitous. Their applications are varied: In physics, they can serve, for example, to encode the flexibility properties of a material under stress from an arbitrary direction; in population genetics, to describe the dynamics of allele frequencies under genetic drift [1].

But diverse as the applied contexts in which they appear may be, almost all matrices of physical relevance share in common that their form or entries are dependent on empirical observations taken on the system which they seek to describe. These observations, of course, can never be made with arbitrary precision, so we are immediately confronted with the problem that any matrices built from them will be subject to error which could invalidate them as instruments of study. Naturally, we then wish to find some precise means of coping with this error, and especially of testing whether or not an apparent feature of our matrices is actually significant and not due to random 'noise' introduced by factors exogenous to our system.

If one has a glancing familiarity with statistical theory, a move which might intuitively suggest itself as a means of resolving this difficulty is to make each entry of the matrix individually a random variable endowed with some probability distribution, and to study the random object which results. Such random objects are called *random matrices*, and, historically, this was essentially the context in which they were first introduced. The discovery was made by the statistician John Wishart, who used them to derive the Wishart matrix distribution, which can be used to compute the maximum likelihood estimator for the covariance matrix of a random vector with normally-distributed entries. [16]

Definition 1.1: An $n \times n$ random matrix M is a random variable which takes values in $M_{n\times n}(\mathbb{F})$, where $\mathbb{F}$ is some field.

1.2 Random matrix theory:

Wishart's innovation was of clear practical significance for scientists and statisticians, but did not, at least initially, provoke much study of random matrices as mathematical objects in themselves, and few advances occurred in the area for several decades. It was not until the 1950s that random matrix theory received its canonical problem – that of solving for the joint probability distribution on the eigenvalues of classes of random matrices – when Eugene Wigner, a physicist, found a way to apply random matrices to an otherwise intractable problem of physical importance in quantum mechanics. [9]

In order to describe the problem with which Wigner was faced, we here provide a very brief, heuristic overview of the relevant quantum mechanics:

In quantum mechanics, the state of a system of 'particles' is characterized by a wave function ψ, which is, formally, an equivalence class of vectors in a Hilbert Space H. The wavefunction has a physical interpretation through $||\psi||^2$, where $||\cdot||$ is the norm induced by the inner product on H, as the probability distribution on the possible locations in space and time of the particles that comprise the system. Because the state of a quantum mechanical system is given essentially by a probability distribution over classical states, it is not, in general, possible to compute measurable, physical quantities associated with the system – such as the position or momentum of a particle – exactly. Instead, we have to adopt an alternative procedure in order to compute what are effectively the average values of measurable quantities over the set of all possible states of the system. In doing this, we associate classical 'observables' – that is, functions of the state parameters of the system which output the value of a measurable quantity when the system is in some state – with linear operators which act on the space of states of the system and which satisfy a Hermitian symmetry property. The exact way in which we compute the average values of measurable quantities from these operators is not relevant to our aims, but it happens that their eigenvalues are of interest. This is because, due to the nature of the equations governing the evolution of the state of a quantum mechanical system, the eigenvalues of these operators correspond to the values of the associated observable that one can possibly measure during an experimental trial. [5]

The problem with which Wigner was concerned was that of determining the possible energies of the nuclei of 'heavy' atoms, which are simply atoms whose nuclei consist of large numbers of subatomic particles. More specifically, because it happens that the possible energy levels for a quantum mechanical system form a discrete set, Wigner wished to investigate the spacings between the 'nearest neighbor' energy levels of the described type of atom. As one might infer from the above, this computation is not in principle particularly complicated. In order to derive a theoretical prediction for width of these spacings, Wigner needed only solve the eigenvalue equation $\hat{H}\psi_i = E_i\psi_i$, where $\hat{H}$ is the operator (known as the Hamiltonian) corresponding with the total energy 'observable' of the system, for the energy eigenvalues $\{E_i\}$ associated with the eigenvectors $\{\psi_i\}$, and then compute differences. However, in practice, this problem proved to be impossible to resolve exactly not just due to the difficulty of actually computing the eigenvalues but of even constructing the Hamiltonian - which is already very complicated with just a few particles - for a system as elaborate as the ones he wished to consider. Wigner consequently decided to try to reduce the complexity of this problem by making a number of approximations.

Because the Hilbert Space in which the state vectors of his system were situated was infinite dimensional and Wigner knew from functional analysis that certain classes of operators on Hilbert spaces could be represented as limits of $n \times n$ matrices as $n \to \infty$, he began by replacing the Hilbert space with an n-dimensional vector space V, for n large. He then wanted to replace the operator $\hat{H}$ with an $n \times n$ matrix acting on V. Wigner did not, however, as noted, have any well-formed idea of what the entries of this matrix should look like. The only constraints which he was certain that he could place upon the entries were those imposed by the Hermitian symmetry requirements for observables discussed above. He therefore decided that he would simply make the matrix a Hermitian one and then, in accordance with the central limit theorem, choose those entries about which he still knew nothing – that is, those which were not constrained by the symmetries of the matrix – to be IID complex-valued random variables with Gaussian distribution $N_{\mathbb{C}}(0,1)$. This, of course, generates a random matrix, and random matrices constructed in this manner are said to be drawn from the *Gaussian Unitary Ensemble* (GUE). Having decided the form of the matrix, Wigner proceeded to compute the joint distribution of its eigenvalues, and from there the probability density of the nearest neighbor spacings of the energy levels.

Strikingly, although there was no a priori reason that this approach ought to have produced anything remotely in touch with reality, the resulting distribution compared well with distributions calculated from empirical data gathered in experiments on systems similar to those earlier described. The success of this model was subsequently responsible for a considerable amount of further interest and study of such ensembles of random matrices. [6]

1.3 The Gaussian Unitary Ensemble:

The computation of Wigner's which we have been discussing - that of deriving the JPDF of the eigenvalues of a random matrix in the GUE - is illustrative of many of the techniques by which the problem of solving for the eigenvalue distributions of random matrix ensembles is often approached. We therefore reproduce it in service of our later discussion:

We first introduce notation: For any matrix $A \in M_{n\times m}(\mathbb{C})$, we denote the exterior product $d\Re(A_{11}) \wedge d\Re(A_{12}) \ldots \wedge d\Re(A_{nm}) \wedge d\Im(A_{11}) \wedge d\Im(A_{12}) \ldots \wedge d\Im(A_{nm})$ of the differentials of the real and complex components of the entries of the matrix by the symbol (dA). [10]

Definition 1.2: Let $M_H^n = \{A \in M_{n\times n} : \overline{A^T} = A\}$ be the set of $n \times n$ Hermitian matrices. Formally, the Gaussian Unitary Ensemble is characterized by a probability distribution $P(A)$ on M_H^n with the following properties:

1. For any $n \times n$ unitary matrix U, the volume form $P(A)(dA)$ on M_H^n is invariant under the transformation $A \mapsto U^*AU$.
2. Where A_{ij} is the ij^{th} entry of A and f_{ij}^0, f_{ij}^1 are functions, the PDF $P(A)$ factors as $P(A) = \prod_{i\leq j} f_{ij}^0(\Re(A_{ij})) \prod_{i<j} f_{ij}^1(\Im(A_{ij}))$.

   (i.e. the entries of $P(A)$ on and below the diagonal are independent - with independent real and complex components - and those entries on the diagonal are strictly real) [9]

These constraints upon the form of the JPDF are sufficient to specify it completely. Furthermore, it is a fact that this characterization of the GUE is equivalent to the one given previously (as random Hermitian matrices with IID Gaussian entries). We use it instead for the purposes of succinctly deriving the form of $P(A)$.

We now prove the following theorem, using the proof presented in Mehta as guide (although substantially elaborated, since Mehta's book is sometimes rather terse):

Theorem 1.3:

$$P(A) = e^{-a\,Tr(A^2)+b\,Tr(A)+c} \quad (1)$$

where a, b, c $\in \mathbb{R}, a > 0$ [6]

Proof: We will need two lemmas, which we shall not prove:

Lemma 1.4: *Let f_1, f_2, f_3 be continuous, differentiable functions such that*

$$f_1(xy) = f_2(x) + f_3(y). \quad (2)$$

then $f_i = a\,ln(x) + b_i$, with $b_1 = b_2 + b_3$.

Lemma 1.5: *For any $B \in M_H^n$ and any function $f : M_H \longrightarrow \mathbb{R}$ which satisfies $f(TBT^*) = f(B)$, $\forall T$ unitary, $f(B)$ is expressible in terms of the traces of a finite number of powers of B. That is, $f(B) = f(Tr\,(B^{i_1}), Tr\,(B^{i_2}), ..., Tr\,(B^{i_k}))$.*

Consider the $n \times n$ matrix

$$U(\theta) = \begin{pmatrix} cos\,\theta & sin\,\theta & 0 & \cdots & 0 \\ -\,sin\,\theta & cos\,\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

We can see that this U is unitary, so $P(A)(dA)$ is invariant under $A \mapsto \widetilde{A} = UAU^*$, per the above.

Let us rewrite the relation $\widetilde{A} = UAU^*$ as $A = U^*\widetilde{A}U$ and differentiate with respect to θ. This results in

$$\begin{aligned} \frac{\partial A}{\partial \theta} &= \frac{\partial U^*}{\partial \theta}\widetilde{A}U + U^*\widetilde{A}\frac{\partial U}{\partial \theta} \\ &= \frac{\partial U^*}{\partial \theta}(UAU^*)U + U^*(UAU^*)\frac{\partial U}{\partial \theta} \\ &= \frac{\partial U^*}{\partial \theta}UA + AU^*\frac{\partial U}{\partial \theta} \\ &= \frac{\partial U^*}{\partial \theta}UA + A(\frac{\partial U^*}{\partial \theta}U)^*. \end{aligned}$$

We can compute that

$$\begin{aligned} \frac{\partial U^*}{\partial \theta}U &= \begin{pmatrix} -\,sin\,\theta & -\,cos\,\theta & 0 & \cdots & 0 \\ cos\,\theta & -\,sin\,\theta & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} cos\,\theta & sin\,\theta & 0 & \cdots & 0 \\ -\,sin\,\theta & cos\,\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \end{aligned}$$

Hence, substituting,

$$\frac{\partial A}{\partial \theta} = \begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & A_{n3} & \cdots & A_{nn} \end{pmatrix}$$

$$+ \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & A_{n3} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -A_{21} & -A_{22} & -A_{23} & \cdots & -A_{2n} \\ A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} + \begin{pmatrix} -A_{12} & A_{11} & 0 & \cdots & 0 \\ -A_{22} & A_{21} & 0 & \cdots & 0 \\ -A_{32} & A_{31} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -A_{n2} & A_{n1} & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -2A_{12} & A_{11} - A_{22} & -A_{23} & \cdots & -A_{2n} \\ A_{11} - A_{22} & 2A_{12} & A_{13} & \cdots & A_{1n} \\ -A_{23} & A_{13} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -A_{2n} & A_{1n} & 0 & \cdots & 0 \end{pmatrix}.$$

Now, since by property 1 of Def. 1.2 $P(A)$ is invariant under the transformation U, clearly $P(A)$ cannot depend on the parameter θ of the transformation $U(\theta)$, so

$$\frac{\partial P(A)}{\partial \theta} = 0 \qquad \forall A \in M_H. \tag{3}$$

On the other hand, taking the logarithm of $P(A)$ and using property 2 of Def. 1.2, we can write that

$$
\begin{aligned}
\frac{\partial\, ln(P(A))}{\partial\theta} &= \frac{\partial}{\partial\theta}(\, ln(\prod_{i\le j} f^{\,0}_{ij}(\Re(A_{ij}))\prod_{i<j} f^{1}_{ij}(\Im(A_{ij})))\,) \\
&= \frac{\partial}{\partial\theta}(\sum_{i\le j} ln(f^{\,0}_{ij}(\Re(A_{ij}))) + \sum_{i<j} ln(f^{1}_{ij}(\Im(A_{ij})))\,) \\
&= \sum_{i\le j}\frac{1}{f^{\,0}_{ij}(\Re(A_{ij}))}\frac{\partial f^{\,0}_{ij}(\Re(A_{ij}))}{\partial\Re(A_{ij})}\frac{\partial\Re(A_{ij})}{\partial\theta} + \sum_{i<j}\frac{1}{f^{1}_{ij}(\Im(A_{ij}))}\frac{\partial f^{1}_{ij}(\Im(A_{ij}))}{\partial\Im(A_{ij})}\frac{\partial\Im(A_{ij})}{\partial\theta} \\
&= 2\,\Re(A_{12})\left(\frac{1}{f^{0}_{22}}\frac{\partial f^{0}_{22}}{\partial\Re(A_{22})} - \frac{1}{f^{0}_{11}}\frac{\partial f^{0}_{11}}{\partial\Re(A_{11})}\right) + \frac{(\Re(A_{11})-\Re(A_{22}))}{f^{0}_{11}}\frac{\partial f^{0}_{11}}{\partial\Re(A_{12})} \\
&+ \sum_{j=3}^{n}\left(\frac{1}{f^{0}_{2j}}\frac{\partial f^{0}_{2j}}{\partial\Re(A_{2j})}\Re(A_{1j}) - \frac{1}{f^{0}_{1j}}\frac{\partial f^{0}_{1j}}{\partial\Re(A_{1j})}\Re(A_{2j})\right) \\
&+ \sum_{j=3}^{n}\left(\frac{1}{f^{1}_{2j}}\frac{\partial f^{1}_{2j}}{\partial\Im(A_{2j})}\Im(A_{1j}) - \frac{1}{f^{1}_{1j}}\frac{\partial f^{1}_{1j}}{\partial\Im(A_{1j})}\Im(A_{2j})\right)\,,
\end{aligned}
$$

where, again, the f^{i}_{jk} are as defined in property 2 under Def. 1.2. Since $ln\,(\cdot)$ is monotonic, we then have by (3) that the above expression is zero as a function of the $\Re(A_{ij}), \Im(A_{ij})$.

But notice that the first line and each of the summands on the second and third lines of the right hand side of the final equality above depends on a different set of (independent) variables than each line or summand. In order for these to sum to zero as functions as required, it is then necessary that each should be independent of the variables on which it nominally depends. In other words, each line or summand must sum to a constant.

The form of the distribution on each of the $\Re(A_{ij}), \Im(A_{ij})$ follows briefly from here, but it is apparent that there are multiple cases to consider, each of which is handled in essentially the same way. We therefore exhibit the proof for only one case explicitly.

By the aforesaid considerations, we know that for some constant c and $j \in \{1, 2, ..., n\}$, identically,

$$
\frac{1}{f^{0}_{2j}}\,\frac{\partial f^{0}_{2j}}{\partial\Re(A_{2j})}\,\Re(A_{1j}) - \frac{1}{f^{0}_{1j}}\,\frac{\partial f^{0}_{1j}}{\partial\Re(A_{1j})}\,\Re(A_{2j}) = \,c\,. \tag{4}
$$

We divide through by $\Re(A_{1j})\Re(A_{2j})$ to receive

$$
\frac{1}{f^{\,0}_{2j}\,\Re(A_{2j})}\,\frac{\partial f^{\,0}_{2j}}{\partial\Re(A_{2j})} - \frac{1}{f^{\,0}_{1j}\,\Re(A_{1j})}\,\frac{\partial f^{\,0}_{1j}}{\partial\Re(A_{1j})} = \frac{c}{\Re(A_{1j})\Re(A_{2j})}\,.
$$

We can see that this is of the form $g_1(\Re(A_{2j})) + g_2(\Re(A_{1j})) = g_3(\Re(A_{2j})\,\Re(A_{1j}))$. Hence, by applying Lemma 1.4, we have that

$$\frac{c}{\Re(A_{1j})\,\Re(A_{2j})} = a\,ln(\Re(A_{1j})\,\Re(A_{2j})) + b_1 + b_2\,.$$

Clearly, identity is possible here only if $a = c = 0$ and $b_1 = -b_2$. This means that

$$\frac{1}{f^0_{2j}\,\Re(A_{2j})}\,\frac{\partial f^0_{2j}}{\partial\Re(A_{2j})} = \frac{1}{f^0_{1j}\,\Re(A_{1j})}\,\frac{\partial f^0_{1j}}{\partial\Re(A_{1j})} = b_1\,.$$

We can then integrate to receive that

$$f^0_{ij} = e^{-b\,\Re(A_{ij})^2} \tag{5}$$

for some constant a, and a similar equation holds in the other cases (in particular, the density functions for the components of the off-diagonal elements are all exponentials of squares in elements of A).

Now observe that, by property 1 of Def. 1.2, $P(A)$ is a function from M_H to $\mathbb{R}$ which satisfies the relation noted in Lemma 1.5, so we may apply Lemma 1.5 to conclude that $P(A)$ is a function of the traces of a finite number of powers of A. Furthermore, since the components of the off-diagonal elements of A have PDFs which are exponentials of squares in the components, we have that $P(A)$ must be an exponential function of at most the trace of the second power of A. Consequently, the distribution is certainly of the form (1), but we do not know a priori that the coefficients a, b, c are real with $a > 0$. However, this follows simply from the fact that $P(H)$ is a probability distribution and so real-valued and normalizable.

This completes the proof. □

Having determined the form of the PDF on the set of Hermitian matrices, we now turn to the task of deducing from this the JDPF of the eigenvalues for matrices in the GUE. This proof - again following that found in Mehta - amounts, essentially, to changing variables and computing the Jacobian of the transformation, a theme which we will find recurrent in later discussions:

Theorem 1.6: *The joint probability distribution on the eigenvalues* $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ *of an* $n \times n$ *matrix drawn from the GUE is given by*

$$P_n = \frac{1}{Z} e^{-\sum_{i=1}^n \lambda_i^2} \prod_{i<j} |\lambda_i - \lambda_j|^2\,, \tag{6}$$

where Z is a normalization constant which can be computed (although we shall not do so) to be

$$Z = 2^{-\frac{n(n-1)}{2}}\pi^{n/2}\prod_{i=2}^{n+1}\Gamma(i)\,. \tag{7}$$

[9]

Proof:

Let A be an $n \times n$ Hermitian matrix and let $\{\lambda_i\}_{i=1}^n$ be its eigenvalue spectrum. By the spectral theorem, the eigenvalues of A are real and we can decompose A as $U\Lambda U^*$, where $\Lambda = diag(\lambda_1, ..., \lambda_n)$ and U is a unitary matrix whose columns are the eigenvectors of A. Since the eigenvalues of A are real, we may suppose that they are ordered as $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$. Furthermore, it is a fact that the set of Hermitian matrices having repeated eigenvalues has measure 0 in M_H^n, so in the context of probability theory (since it's necessary to integrate a density function to get a probability) we can ignore this set and thus choose the eigenvectors of A to be pairwise orthogonal. We can then fix the phase of the eigenvectors by demanding that the first nonvanishing entry be positive and real. Under these constraints, the decomposition $U\Lambda U^*$ of A is unique.

Consider now the distribution (1). Recall that due to symmetry, we need only choose the n real diagonal entries and $n(n-1)/2$ additional entries - for which we choose independently both the real and complex parts - in order to specify an $n \times n$ Hermitian matrix completely. This means that when we decompose an $n \times n$ Hermitian matrix via the spectral theorem, since n of our entries must be the (real) eigenvalues which are the nonvanishing components of the central diagonal matrix in the decomposition, we then have to specify $n(n-1)/2$ further entries, all of which must evidently be in the matrix of eigenvectors, in order to have determined our decomposition completely. For our convenience, we choose those entries whose components we will specify independently to be the entries below the diagonal of the eigenvector matrix. As we have that the decomposition $A = U\Lambda U^*$ is unique when carried out as specified above, the transformation

$$(\Re(U_{ij}), \Im(U_{ij}), \lambda_k) \mapsto f(\Re(U_{ij}), \Im(U_{ij}), \lambda_k) = U\Lambda U^* = A \in \mathbb{C}^{n^2}, \quad 1 \leq i < j \leq n$$

from the coordinates $(\Re(U_{ij}), \Im(U_{ij}), \lambda_k)$ to the coordinates $(\Re(A_{lh}), \Im(A_{lh}))$ is bijective, so we can apply the change of variables theorem and the fact that $Tr(A^k) = \sum_{i=1}^n \alpha_i^k$, for each $k \in \mathbb{N}$, in order to write

$$P(A)(dA) = e^{-a\sum_{i=1}^n \lambda_i^2 + b\sum_{i=1}^n \lambda_i + c}\, det(Df)\,(dU)(d\Lambda) \tag{8}$$

where Df is the Jacobian of the transformation.

Reindex the set of auxiliary parameters of the matrix U so that $\{\Re(U_{ij}), \Im(U_{ij}) : 1 \leq i < j \leq n\} = \{u_\ell : 1 \leq \ell \leq n(n-1)\}$. Now note that since U is unitary, we have the identity $U^*U = I$. We differentiate this with respect to the auxiliary parameter u_ℓ to find that

$$
\begin{aligned}
0 &= \frac{\partial}{\partial u_\ell}(U^*U) \\
&= \frac{\partial U^*}{\partial u_\ell}U + U^*\frac{\partial U}{\partial u_\ell}
\end{aligned}
$$

which implies that

$$
\frac{\partial U^*}{\partial u_\ell}U = -\left(\frac{\partial U^*}{\partial u_\ell}U\right)^* . \tag{9}
$$

On the other hand, we differentiate the identity $A = U\Lambda U^*$ with respect to u_ℓ to receive

$$
\frac{\partial A}{\partial u_\ell} = \frac{\partial U}{\partial u_\ell}\Lambda U^* + U\Lambda\frac{\partial U^*}{\partial u_\ell}
$$

which, left multiplying by U^*, right multiplying by U and setting -$\frac{\partial U^*}{\partial u_\ell}U = S^\ell$, gives us

$$
\begin{aligned}
U^*\frac{\partial A}{\partial u_\ell}U &= U^*\frac{\partial U}{\partial u_\ell}\Lambda + \Lambda\frac{\partial U^*}{\partial u_\ell}U \\
&= (\frac{\partial U^*}{\partial u_\ell}U)^*\Lambda + \Lambda\frac{\partial U^*}{\partial u_\ell}U \\
&= \Lambda\frac{\partial U^*}{\partial u_\ell}U - \frac{\partial U^*}{\partial u_\ell}U\Lambda \\
&= S^\ell\Lambda - \Lambda S^\ell .
\end{aligned}
$$

Written in terms of the entries of the matrices, this gives us the equations

$$
\sum_{i,j}\frac{\partial A_{ij}}{\partial u_\ell}\overline{U_{ik}}U_{jh} = S^\ell_{kh}(\lambda_h - \lambda_k) . \tag{10}
$$

We can then differentiate the same identity with respect to λ_q to find that

$$
\sum_{i,j}\frac{\partial A_{ij}}{\partial \lambda_q}\overline{U_{ik}}U_{jh} = \frac{\partial \Lambda_{kh}}{\partial \lambda_q} = \delta_{kh}\delta_{ki} , \tag{11}
$$

where δ_{pq} is the Kronecker delta.

Now we want to rewrite these equations in block matrix form as

$$
\begin{aligned}
&\begin{pmatrix} [\frac{\partial \Re(A_{ii})}{\partial \lambda_q}] & [\frac{\partial \Re(A_{ij})}{\partial \lambda_q}] & [\frac{\partial (A_{ij})}{\partial \lambda_q}] \\ [\frac{\partial \Re(A_{ii})}{\partial u_\ell}] & [\frac{\partial \Re(A_{ij})}{\partial u_\ell}] & [\frac{\partial (A_{ij})}{\partial u_\ell}] \end{pmatrix} \begin{pmatrix} v & w \\ L_0 & M_0 \\ L_1 & M_1 \end{pmatrix} \\
&= \begin{pmatrix} [\rho_{qk}] & [\sigma^0_{qkh}] & [\sigma^1_{qkh}] \\ [\epsilon^\ell_k] & [\Re(S^\ell_{kh})(\lambda_h - \lambda_k)] & [\Im(S^\ell_{kh})(\lambda_h - \lambda_k)] \end{pmatrix}
\end{aligned}
$$

where $1 \leq i < j \leq n$, $1 \leq k < h \leq n$, $1 \leq \ell \leq n(n-1)$ and $1 \leq q \leq n$, for $n \times n$ ρ, v and $[\frac{\partial \Re(A_{ii})}{\partial \lambda_q}]$, $n \times n(n-1)/2$ σ, $[\frac{\partial \Re(A_{ij})}{\partial \lambda_q}]$ and $[\frac{\partial \Im(A_{ij})}{\partial \lambda_q}]$, $n \times n(n-1)$ w, $n(n-1)/2 \times n(n-1)$ M^0 and M^1, $n(n-1) \times n(n-1)/2$, $[\frac{\partial \Re(A_{ij})}{\partial u_\ell}]$, $[\frac{\partial \Im(A_{ij})}{\partial u_\ell}]$, $[\Re(S^\ell_{kh})]$, and $[\Im(S^\ell_{kh})]$, $n(n-1) \times n$ ϵ^0, ϵ^1 and $[\frac{\partial \Re(A_{ii})}{\partial u_\ell}]$ and, lastly, $n(n-1)/2 \times n$ L^0 and L^1.

Moreover, we wish to write these equations in this form such that all of the aforementioned except for the matrices with derivatives of components of A explicitly appearing in their definitions do not depend on the eigenvalues and so that σ^0, σ^1 are matrices of all zeroes. Because the computations involved in showing this explicitly would consume several more pages without being particularly edifying, we simply assert that this can be done, and, noting that the first matrix on the left-hand side is the Jacobian Df of the above mentioned transformation, take the determinant of both sides of the matrix equation and separate out the factors of $(\lambda_k - \lambda_h)$ to establish that

$$det(Df) = \prod_{k<h} |\lambda_h - \lambda_k|^2 \, g(u_1, ..., u_{n(n-1)}) \tag{12}$$

where $g(u_1, ..., u_{n(n-1)})$ is some function which we do not need to determine. We then substitute into (8) and integrate with respect to the parameters u_ℓ (gathering the new constant into e^c) to receive

$$P(A)(dA) = e^{-a \sum_{i=1}^n \lambda_i^2 + b \sum_{i=1}^n \lambda_i + c} \prod_{k<h} |\lambda_h - \lambda_k|^2 \, (d\Lambda) \tag{13}$$

We can then linearly transform the eigenvalues as $\lambda_i = \frac{1}{\sqrt{2a}} \widetilde{\lambda}_i + \frac{b}{2a}$ and normalize in order to recover the density (6).

This completes the proof. □

Note the factor $\prod_{i<j}(\lambda_i - \lambda_j)$ that appears squared in (6). This is known as the *Vandermonde factor* or, alternately, the *Vandermonde determinant* since it arises as the determinant of a certain matrix which appears in the study of polynomials, and it is frequently seen in the the JPDFs of the eigenvalues of random matrix ensembles.

2 The Eigenvalue Statistics of Random $*$-Cosquare Matrices

In this section, we are concerned with an unsolved problem in random matrix theory, which we shall here describe:

Let M be an $n \times n$ random matrix with complex-valued IID entries distributed as $N_{\mathbb{C}}(0, 1)$, the standard complex Gaussian. Then M is said to be in the *Complex Ginibre Ensemble* (CGE), which (it can be shown) has distribution on $M_{n\times n}(\mathbb{C})$ given by

$$P(A) = \pi^{-n^2} e^{-\frac{1}{2}Tr\,(AA^*)} \tag{14}$$

over $M_{n\times n}(\mathbb{C})$. [6] We define a $*$-cosquare matrix to be an element of the image of the map $A \mapsto A^* A^{-1}$.

The problem in which we are interested is then to compute the distribution of the Gaussian $*$-Cosquare Ensemble, or in other words, to find the distribution of the random variable $M^* M^{-1}$ given that M is in the CGE. We further wish to compute the JPDF of the eigenvalues of $M^* M^{-1}$. This would enable us, ideally, to compute the joint intensities of the eigenvalues, which are quantities of importance that we shall shortly define. First, however, we contextualize our interest in this problem:

2.1 Context:

Broadly speaking, our interest in this problem has to do with its connections to a certain important type of random process known as a point process. Informally, a point process is a random set of distinct, separated points in some space. More exactly:

Let X be a metric space. A subset $E \subseteq X$ is said to be *discrete* if for each $x \in E$ there exists a neighborhood $V \subseteq X$ of x such that $V \cap E = \{x\}$. We denote by S_X the set of discrete subsets of X.

Definition 2.1: A point process on X is a random variable taking values in S_X.

A point process can be described in terms of a sequence of quantities known as the *joint intensities* of the process. The formal definition of the joint intensities of P is rather opaque, but in essence, they provide a measure of the conditional probability that a point of the point process' random discrete subset will fall into some small set given that we know that some number of other points of the random subset fall into other small sets which are mutually disjoint from one another. The utility of this is that it encodes information concerning how the locations of points in the random subset are correlated with one another, and can help us to determine if the points of a random subset tend to 'repel' or 'attract', in the sense that they exhibit clustering or dispersion at a rate not typical of a process where the locations of the points are independent.

Definition 2.2: Let P be a point process on $\mathbb{C}^m$, and let $\{D_1, D_2, ...\}$ be mutually disjoint subsets of $\mathbb{C}^m$. Denote by $N^P_{D_i}$ the random variable which counts the number of points of P in the set D_i. For each $k \in \mathbb{N}$, we define the the k^{th} joint intensity of P to be the function $\rho_k : \bigtimes_{i=1}^{k} \mathbb{C}^m \longrightarrow [0, \infty)$, if it exists, such that

$$\mathbb{E}\left[\prod_{i=1}^{k} N_{D_i}^{P}\right] = \int_{D_1} ... \int_{D_k} \rho_k(x_1, ..., x_k)\, dx_1 ... dx_k \,. \tag{15}$$

There are several classes of point process, each of which exhibits distinct characteristics, but we will discuss here only one: the class of *determinantal point processes.* Formally,

Definition 2.3: Let P be a point process on $\mathbb{C}^m$ with joint intensities ρ_k. If there exists a function $K : \mathbb{C}^m \times \mathbb{C}^m \longrightarrow \mathbb{C}$ (which we call the 'kernel' of the process) such that for each k we can express $\rho_k(x_1, x_2, ..., x_k)$ as the determinant of the $k \times k$ matrix

$$\begin{pmatrix} K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n) \\ K(x_2,x_1) & K(x_2,x_2) & \cdots & K(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n) \end{pmatrix}$$

then we say that P is a determinantal point process. [6]

As was the case with joint intensities, it is somewhat obscure from this definition what could be of interest about this particular sort of point process. Well, there are some considerations concerning the ease of performing certain computations of importance for determinantal point processes, but primarily we are interested in them because the individual points of determinantal point processes exhibit repulsion, in the sense noted above. To illustrate this, we first note that the Gaussian Unitary Ensemble actually is an example of a determinantal point process, and it can be shown that (for an $n \times n$ matrix) it has kernel

$$K(x,y) = \sum_{i=1}^{n} \frac{e^{\frac{-x^2-y^2}{4}} H_i(x) H_i(y)}{(2\pi)^{1/4} n!^{1/2}}, \tag{16}$$

where $H_i(\cdot)$ is the i^{th} Hermite polynomial. We can see that the GUE exhibits repulsion after the fashion of a determinantal point process in two ways. Firstly, by noting that if we fix some λ_i, the Vandermonde factor is zero when $\lambda_i = \lambda_j$ and so (by continuity) small in a neighborhood of λ_i, which in turn makes the probability of discovering another of the eigenvalues in that neighborhood small. Secondly, we can see this by observing that the density (6) is equivalent to

$$\frac{1}{Z} e^{-\frac{1}{T}(\frac{1}{2}\sum_{i=1}^{n} \lambda_i^2 - \sum_{i<j} log\,|\lambda_i - \lambda_j|)}$$

when $T = \frac{1}{2}$. But (with appropriate units) this corresponds, in the language of statistical mechanics, to the Boltzmann factor for a system of particles of like charge in two dimensions when confined by a harmonic potential, so it must be that the eigenvalues of matrices in the GUE obey the same statistics as such a system. Since we would obviously expect particles in a system of this sort to repel at non-zero temperatures (and clearly $T = 1/2$ is not zero), we would expect the same of the eigenvalues of matrices in the GUE. [6]

In light of the preceding discussion, it is apparent that determinantal point processes might be used to model any number of physical phenomena - for example, the distribution of trees in a forest - but they are also useful in the study of some important mathematical phenomena. One such phenomenon is the zero set of the Riemann zeta function. Remarkably, this again involves the eigenvalue statistics of the GUE. A conjecture has been made that the pair correlation function of the zeroes of the Riemann zeta function which are of the form $1/2 + bi,\ b \in \mathbb{R}$, is given by the second joint intensity of the eigenvalues of the Gaussian Unitary Ensemble, which, for one of the pair of eigenvalues fixed and taken as the origin and on a length-scale such that the expected distance between eigenvalues is 1, can be shown to be

$$\rho_2(x, 0) = 1 - \frac{sin^2(\pi x)}{(\pi x)^2}. \tag{17}$$

This conjecture, although unproven, has been strongly evidenced by precise numerical computations. [11]

This connection of the Gaussian Unitary Ensemble to the distribution of the roots of the Riemann zeta function relates strongly to our interest in the Gaussian $*$-Cosquare Ensemble. In particular, it relates by way of a property of the characteristic polynomials of $*$-Cosquare matrices, which we here discuss.

Definition 2.4: Let $p(z) = \sum_{i=0}^{n} c_i z^i \in \mathbb{C}[z]$ be a polynomial and let $U = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in the complex plane. Denote by $\widetilde{p}(z)$ the polynomial $\sum_{i=0}^{n} \overline{c_{n-i}} z^i$. If there exists $\omega \in U$ such that $p(z) = \omega\, \widetilde{p}(z)$, then we say that $p(z)$ is *self-inversive.* [14]

Note that, also, $\widetilde{p}(z) = z^n \overline{p(1/\overline{z})}$.

Proposition 2.5: For any $A \in M_{n\times n}(\mathbb{C})$, the characteristic polynomial of $A^* A^{-1}$ is self-inversive.

Proof:

Let $det(A) = re^{i\theta}$. Then

$$\begin{aligned} det(A^* A^{-1}) &= det(A^*) det(A^{-1}) \\ &= \overline{det(A)}/det(A) \\ &= re^{-i\theta}/re^{i\theta} = e^{-2\theta i}\,, \end{aligned}$$

which is on the unit circle in the complex plane. Furthermore, recall that for any matrices $A, B \in M_{n\times n}(\mathbb{F})$, the characteristic polynomial of AB is the same as that of BA (i.e., $det(AB-zI) = det(BA-zI)$). Now set $g(z) = det(A^*A^{-1}-zI)$. We can compute that

$$\begin{aligned}
g(z) &= det(A^*A^{-1} - z\,I) \\
&= det(-z\,A^*A^{-1}((A^*A^{-1})^{-1} - z^{-1}I)) \\
&= (-1)^n z^n det(A^*A^{-1}) det((A^*A^{-1})^{-1} - z^{-1}I) \\
&= e^{n\pi i - 2\theta i} z^n det(A(A^*)^{-1} - z^{-1}I) \\
&= e^{(n\pi\,-\,2\theta)i} z^n \overline{det((A(A^*)^{-1})^* - \overline{z^{-1}}I)} \\
&= e^{(n\pi\,-\,2\theta)i} z^n \overline{det((A^{-1}A^* - \overline{z^{-1}}I)} \\
&= e^{(n\pi\,-\,2\theta)i} z^n \overline{det((A^*A^{-1} - \overline{z^{-1}}I)} = e^{(n\pi\,-\,2\theta)i}\,\widetilde{g}(z)\,.
\end{aligned}$$

Since $e^{(n\pi\,-\,2\theta)i}$ is on the unit circle, we have that $g(z)$ is self-inversive, as required. □ [1]

From the definition, we can see that the roots of a self-inversive polynomial come in two varieties. If z is a root of a self-inversive polynomial, then either z is on the unit circle or $1/\overline{z}$ is a root also. The connection of self-inversive polynomials with the zeroes of the Riemann zeta function comes through a more general class of function, known as an L-function, of which the Riemann zeta function is a specific example. L-functions are required to satisfy a certain functional relation which can be written in a form which makes clear that they have a symmetry about the 'critical line' $\Re(z) = 1/2$ of the Riemann hypothesis which is analogous to the described symmetry that the zeroes of self-inversive polynomials possess. Since, as proved above, the characteristic polynomials of $*$-cosquare matrices are self-inversive, a major reason that we might wish to compute the JPDF of the eigenvalues of matrices in the Gaussian $*$-Cosquare Ensemble is then that this would facilitate the computation of the joint intensities of the underlying point process and these might provide additional information about the zeroes of some types of L-functions (which are conjectured to satisfy the Riemann hypothesis), as was the case with the joint intensities of the GUE. [2]

Additionally, random self-inversive polynomials have been appeared in the context of quantum mechanics, and it may be the case that the eigenvalue statistics of the $*$-Cosquare Ensemble resemble those of the roots of self-inversive polynomials, in which case this ensemble may also possess physical applications. [3]

[1] This proof, which is not at all challenging and has certainly been done before (in some variation), is 'original' work in that we were unable to find a source concerning $*$-cosquare matrices which proved this claim instead of simply asserting it and so had to come up with the proof ourselves - hence, no citation.

2.2 The Problem of Computing the JPDF of the Eigenvalues:

Aside from the above considerations, another reason that we take up this problem in particular is that it seems reasonably likely to be exactly solvable, which, perhaps unsurprisingly, is rather uncommon in dealing with arbitrary ensembles of random matrices. Our justification for believing that the JPDF for the eigenvalues may be exactly determinable in this case is founded in some interesting connections which the $*$-Cosquare Ensemble has to another ensemble of random matrices - the Spherical Ensemble (GSE) - for which the JPDF of the eigenvalues has been successfully computed. [4]

The Spherical Ensemble is the ensemble of random matrices of the form $A^{-1}B$, where A, B are drawn independently from the real Ginibre Ensemble, which is defined in the same way as the CGE discussed previously but with real-valued Gaussian entries. Much as with the GUE, one obtains the eigenvalue density of this ensemble by first computing the distribution on the elements of $A^{-1}B$ from the distributions on A and B, then performing a decomposition on $A^{-1}B$ which allows for it to be expressed in terms of the eigenvalues and a number of auxiliary parameters and, finally, integrating out the auxiliary parameters. The resulting distribution is rather unwieldy and doesn't bear repeating in this context, but there is an attribute of it which is relevant to the problem at hand. Namely, we find that when we stereographically project the eigenvalues of the GSE onto the Riemann sphere, they have a uniform density except on the great circle of the Riemann sphere which is the image of the real line under the projection. On this great circle, the density of eigenvalues is significantly higher than elsewhere, although this anomaly 'washes out' and the density becomes fully uniform on the sphere as the dimension of the random matrices becomes large. [8]

This feature of the JPDF of the eigenvalues of the GSE provokes our interest because nearly the same feature seems to be evident in the joint distribution of the eigenvalues for the Gaussian $*$-Cosquare Ensemble. We see in numerical simulations that, under stereographic projection onto the Riemann Sphere, the eigenvalues of a large number of realizations of random matrices in the $*$-Cosquare Ensemble seem to be distributed relatively uniformly on the sphere except on a great circle corresponding to the unit circle in the complex plane under projection.

We now briefly discuss an important aspect of random matrix theory: the phenomenon known as universality. Universality suggests essentially that, as the dimension of the matrices $n \to \infty$, the local statistics of the eigenvalues of matrices in particular ensembles will depend not on the actual distributions of the individual elements of the matrices but on their membership in particular classes of ensemble, called universality classes, which are characterized by certain overarching properties. Probably the most famous overarching 'law' which has been proven for a class of ensembles (in this case, the 'Wigner ensembles' of random Hermitian matrices with IID entries with mean 0 and variance 1) is the 'Wigner semi-circle law', which says that in the limit as n becomes infinite, the number of eigenvalues in an interval for an $n \times n$ Wigner ensemble matrix goes to the area under a semi-circular curve of radius 2. We state this result to illustrate:

Theorem 2.6: *Let M_n be an $n \times n$ random Hermitian matrix with IID entries have mean 0 and variance 1, let $I \subseteq \mathbb{R}$ be an interval and let N_I be the random variable which counts the number of eigenvalues of $\frac{1}{\sqrt{n}}M_n$ that are in I. Define $\rho(x)$ to be that function which is equal to $\frac{1}{2\pi}\sqrt{4-x^2}$ when $|x| \leq 2$ and 0 elsewhere. Then*

$$\lim_{n\to\infty} \frac{1}{n} N_I = \int_I \rho(x)dx \tag{18}$$

in the sense of probability. [15]

The apparent convergence of the eigenvalue JPDFs of the GSE and Gaussian $*$-Cosquare Ensemble under projection onto the Riemann sphere to a uniform density on the sphere suggests that these ensembles may belong to the same universality class. Moreover, since the GSE can be shown to be a determinantal point process, we then have reason to suspect that the $*$-Cosquare Ensemble will also be determinantal with similar statistics, and so potentially also have an eigenvalue JPDF which can be exactly determined, as noted.

We made some attempts to solve this problem, but quickly encountered computational difficulties which proved to be insurmountable within the confines of the time allotted for the composition of this thesis. We were initially hopeful that, due to the apparent formal similarity of the matrices in the GSE and Gaussian $*$-Cosquare Ensemble, we might be able to use the computational methods developed in Forrester and Mays' proof of the eigenvalue JPDF for the Spherical Ensemble as a guide. However, it turns out that the assumption of independence for the matrices A and B is quite crucial early on in this proofs, and the evident lack of independence between A^* and A^{-1} almost immediately precludes similar techniques from being used.

To be explicit, setting $A^{-1}B = M$, Forrester and Mays use independence to write the probability distribution on (A, B) as

$$\begin{aligned}
P(A,B)(dA)(dB) &= P(A)P(B)(dA)(dB) \\
&= (2\pi)^{-n^2} e^{-\frac{1}{2}Tr(AA^T)} e^{-\frac{1}{2}Tr(BB^T)}(dA)(dB) \\
&= (2\pi)^{-n^2} e^{-\frac{1}{2}Tr(AA^T+BB^T)}(dA)(dB) \\
&= (2\pi)^{-n^2} e^{-\frac{1}{2}Tr(AA^T(I+MM^T))}(dA)(dB)\,.
\end{aligned}$$

They then rearrange the identity $A^{-1}B = M$ to $B = AM$ and use the following Lemma:

Lemma 2.7: *For any matrix $X = \alpha Y \beta$, where Y is $P \times Q$ and has PQ independent entries, α is $P \times P$ and β is $Q \times Q$, we have that*

$$(dX) = |det(\alpha)|^Q |det(\beta)|^P (dY)\,. \tag{19}$$

[4]

Setting $X = B$, $\alpha = A$, $\beta = I$, this shows that

$$(dB) = |det(A)|^n (dM)\,. \tag{20}$$

From there, they are able to write

$$(2\pi)^{-n^2} e^{-\frac{1}{2}Tr(AA^T(I+MM^T))}(dA)(dB)$$
$$= (2\pi)^{-n^2} |det(AA^T)|^{n/2} e^{-\frac{1}{2}Tr(AA^T(I+MM^T))}(dA)(dM)$$

and use the symmetry of AA^T and a known integral (a 'Selberg' integral) to integrate out the dependency on A, leaving a distribution on M, as desired. [4]

However, in our case, although we can also set $B = A^{-1}A^*$ [2], rearrange to get, for example, $A^* = AB$ and use Lemma 2.7 to write that

$$(dA) = |det(A^*)|^n \overline{(dB)}\,, \tag{21}$$

this ultimately proves to be unhelpful because the density $P(A)$ cannot easily be rewritten in terms of B. This is due to the fact that $f(A) = A^{-1}A^*$ does not, to our knowledge, have a nice inverse function, and there is no possibility of separating the component matrices out via independence, as was the case in Forrester and Mays' computation. Furthermore, even if we were able to obtain the distribution on the elements, it's not clear, given the asymmetry of the matrix, what decomposition on B would usefully isolate the eigenvalues and allow us to integrate out the auxiliary parameters in order to obtain the JPDF on the eigenvalues. Hence, we were not able to progress with this problem.

3 Random Self-Inversive Polynomials:

Having been unable to carry out the computation outlined previously, but still being interested in the properties of self-inversive polynomials whose coefficients are somehow randomized, we proceeded to try to address an open question explicitly in the area of random self-inversive polynomials, rather than in random matrix theory proper. Preliminarily, let us note that

[2] It is a fact that the image of the map $A \mapsto A^{-1}A^*$ gives us the set of $*$-Cosquare matrices as well.

Definition 3.1: A random polynomial in $\mathbb{F}$, a field, is a random variable which takes values in the set $\mathbb{F}[x]$ of polynomials with coefficients in $\mathbb{F}$.

Similar to random matrices, we typically randomize these polynomials by making the coefficients individual random variables having some distribution. In our case, we wish to generate a random, monic self-inversive polynomial of degree n, which we will write $p(z) = z^n + \sum_{i=0}^{n-1} c_i z^i$. To do so, we rely on the fact that, from the definition of self-inversive polynomial, we have that the coefficients of the polynomial are constrained as $c_i = \omega \overline{c_{n-i}}$, where ω is the parameter such that $p(z) = \omega \widetilde{p}(z)$, which means that the random polynomial is completely specified (in the odd case[3]) by choices of only $c_0, ..., c_{(n-1)/2} \in \mathbb{C}$ and ω on the unit circle.[2]

The problem which we in this section consider is somewhat different than the previous one, although there are again extensive connections between them. Here, our aim was to determine the correlation functions on the *roots* of a random self-inversive polynomial of even degree whose independent coefficients are IID $N_{\mathbb{C}}(0,1)$ complex Gaussians, given the JPDF on the roots of the polynomial.

3.1 Context:

The analogy between the IID complex Gaussian entries of a matrix in the Gaussian $*$-Cosquare Ensemble and and the coefficients of a random self-inversive polynomial seems clear, but the fact that we are concerned with the roots of the polynomial may at first be surprising. We state a result:

Theorem 3.2: *Let $p(z) = z^n + \sum_{i=0}^{n-1} c_i z^i$ be a polynomial in $\mathbb{C}[z]$ having roots $\{\xi_1, ..., \xi_n\}$. Then the transformation $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ defined by $T(\xi_1, ..., \xi_n) = (c_{n-1}, ..., c_0)$ has Jacobian determinant*

$$det(DT) = \prod_{i<j} |\xi_i - \xi_j|^2 \,. \tag{22}$$

The details of the proof of this theorem are not important here, but, since these functions will recur shortly, we will note that it is instrumental to the proof to observe that the coefficients c_{n-i} of the polynomial can be expressed in terms of the zeroes as

$$c_{n-i} = (-1)^i \sum_{1 \leq i_i < ... < i_k \leq n} \xi_{i_1} ... \xi_{i_k} \,, \tag{23}$$

which are known (aside from the factor of $(-1)^i$ and considered as functions of the ξ_i) as *elementary symmetric functions.* [6]

[3]Replace $(n-1)/2$ with $n/2$ in the case that the degree of the polynomial is even, but note that for even degree, ω is necessarily 1 because the middle coefficient $a_{n/2}$ of the self-inversive polynomial must be real and so is not chosen.

The Jacobian determinant that appears in the above theorem is, once again, the square of the Vandermonde determinant. We have previously explained that when the Vandermonde factor appears in the JPDF of some number of objects, it causes those objects to 'repel' one another probabilistically. This means that, regardless of what the distribution placed on the coefficients of a random polynomial happens to be, when we change variables to study the statistics of the roots, we ultimately end up with a random process which exhibits repulsion. There is then a strong analogy between the roots of random polynomials and the eigenvalues of matrices in the random matrix ensembles which we have heretofore discussed.

3.2 The Problem of Computing the Correlation Functions:

Turning to the specific problem that we want to address, we find that, unlike with the eigenvalues of matrices in the Gaussian $*$-Cosquare Ensemble, the joint probability distributions on the roots of self-inversive polynomials have already been computed in the case that all the roots are on the unit circle. Since it can be shown that this set does not have measure zero (under a nonvanishing distribution) in the space of self-inversive polynomials, we will discuss self-inversive polynomials of this sort.

Theorem 3.3: *Let $p(z) = z^n + \sum_{i=0}^{n-1} c_i z^i$ be a random self-inversive polynomial with IID standard complex Gaussian coefficients. Denote by ξ_i, $i = 1, ..., n$, the roots of the polynomial, which we shall assume are all on the unit circle. Then for n odd, the JPDF of the roots, up to normalization, is*

$$P(\xi_1, ..., \xi_n) = \prod_{i<j} |\xi_i - \xi_j| \tag{24}$$

and for n even, the JPDF, up to normalization, is

$$P(\xi_1, ..., \xi_n) = |c_{n/2}| \prod_{i<j} |\xi_i - \xi_j| \,. \tag{25}$$

[2]

Since the problem of computing these distributions has been resolved, we may therefore proceed directly to what would have been the ultimate goal of the previous section: computing the correlation functions (which are the same as the joint intensities) for the roots of the polynomials. We find that even this problem has been resolved, but only for the case that n is odd. This is rather atypical. For most ensembles of random objects, the *even* case is the first dealt with, and the odd case is carried out subsequently and without much haste. However, in this case, not only has the odd case been resolved first, but the even case seems to be much more difficult. We can see why, heuristically, this would be the case, if we consider how the correlation functions are often computed. We define:

Definition 3.4: The Pfaffian of a $2n \times 2n$ matrix is defined as

$$Pf(A) = \frac{1}{2^n\, n!} \sum_{\sigma \in S_{2n}} sgn(\sigma) \prod_{i=1}^{n} a_{\sigma(2i-1),\sigma(2i)} \tag{26}$$

where S_{2n} is the Symmetric group of order $2n$.

One step by which one often proceeds to compute correlation functions is to express a certain integral known as the 'partition function' (essentially, the constant which normalizes a measure into a probability measure) as the Pfaffian of some matrix. Since the 'partition function' is something like a normalizing constant, it is then unsurprising in the case of the Gaussian ensemble of self-inversive polynomials that the partition function would involve an integral over the domain of definition of either the distribution (24) or (25), according to whether we are dealing with even or odd degree polynomials. In order to manipulate the partition function such that it can expressed as a Pfaffian, it is then necessary to express the integrand - that is, either distribution (24) or (25) - in a form which is amenable to this. Now, the form of the distributions (24) and (25) doesn't seem particularly complicated, but when we are integrating the distributions over a set, we encounter problems due to the presence of the absolute value, which can change sign in way which can be difficult to cope with. In the odd case, however, the distribution was put into a tractable form by factoring

$$\prod_{i<j} |\xi_i - \xi_j| = \prod_{i<j} (\xi_i - \xi_j)\, sgn(\xi_i - \xi_j)\,. \tag{27}$$

The Vandermonde factor $\prod_{i<j}(\xi_i - \xi_j)$ in the distribution can, as noted, be expressed as the determinant of the Vandermonde matrix, and it turns out that the product of the signs of the differences of the roots also can be conveniently expressed as the Pfaffian of the matrix $[sgn(\xi_i - \xi_j)]_{i,j=1}^{n}$. [12] As a consequence of this, some further manipulations using the definition of the determinant as a sum over a group of permutations and applying Fubini's theorem allow us to express the partition function as a Pfaffian. Ideally, we would like to do something similar for the even case, but we can clearly see that this would involve finding a matrix whose Pfaffian gives us the product of the signs of the differences of the roots multiplied by the sign of the middle coefficient $a_{n/2}$. But, as we noted above, the middle coefficient expressed in terms of the roots of the polynomial is an elementary symmetric function, for which we do not know a simple formula for the sign. Not knowing a formula for the sign of $c_{n/2}$, of course, renders it troublesome to determine what a matrix whose Pfaffian includes that as a factor. Our focus was thus to try to find a workable formula for $sgn(c_{n/2})$. [13]

We approached this from a few angles. One such angle, suggested by Dr. Sinclair, was to restrict to the case where we have a monic, self-inversive polynomial $f(z) = \sum_{k=0}^{2n} a_k z^k$ of order $2n$ with real coefficients, then change variables as $z = e^{2\pi i\theta}$ and due to the symmetries of the coefficients receive

$$\begin{aligned} z^{-n} f(z) &= a_n + 2\sum_{k=1}^{n} a_{n+k} cos(2\pi k\theta) \\ &= a_n + 2\sum_{k=1}^{n} a_{n+k} T_k(cos(2\pi\theta)) \,, \end{aligned}$$

where T_k is the k^{th} Chebyshev polynomial. [7] Since the Chebyshev polynomials are orthogonal with respect to the L^2 inner product with weight $\frac{1}{\sqrt{1-x^2}}$, we can then take the appropriate inner product of $T_1 = 1$ with the far left and right sides of the above identities in order to isolate the middle coefficient, and perhaps take its sign. That is, if $\langle \cdot, \cdot \rangle$ is the inner product with respect to which the Chebyshev polynomials are orthogonal, then

$$\langle z^{-n} f(z), T_1 \rangle = a_n \tag{28}$$

However, it's not clear to us that successfully computing this inner product would yield what is desired, if it is tractable.

References

[1] W. Bodmer. A migration matrix model for the study of random genetic drift. *Genetics*, 59:565–592, 1968.

[2] N. Snaith D. Farmer, F. Mezzadri. Random polynomials, random matrices and l-functions, ii. *Nonlinearity London*, 19:919–936, 2006.

[3] P. Lebeouf E. Bogomolny, O. Bohigas. Quantum chaotic dynamics and random polynomials. *Journal of Statistical Physics*, 85:639–679, 1996.

[4] P. Forrester and A. Mays. Pfaffian point process for the gaussian real generalised eigenvalue problem. *Probability Theory and Related Fields*, 154:1–47, 2011.

[5] D. Griffith. *Introduction to Quantum Mechanics*. American Mathematical Society, seventh edition, 2005.

[6] Y. Peres B. Virag J. Hough, M. Krishnapur. *Zeroes of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society, 2009.

[7] J. Vaaler M. Mossinghoff, C. Pinner. Perturbing polynomials with all their roots on the unit circle. *Mathematics of Computation*, 224:1707–1726, 1998.

[8] A. Mays. *A Geometrical Triumvirate of Random Matrices*. PhD thesis, University of Melbourne, 2011.

[9] M. Mehta. *Random Matrices*. Elsevier Academic Press, third edition, 2004.

[10] R. Muirhead. *Aspects of multivariate statistical theory*. Wiley-Interscience, first edition, 1982.

[11] A. Odlyzko. The 10 to the 20th zero of the riemann zeta function and 175 million of its neighbors.

[12] C. Sinclair. *Multiplicative Distance Functions*. PhD thesis, University of Texas at Austin, 2005.

[13] C. Sinclair. Correlation functions for =1 ensembles of matrices of odd size. *Journal of Statistical Physics*, 136:17–33, 2009.

[14] C. Sinclair and J. Vaaler. Self-inversive polynomials with all zeroes on the unit circle. *London Mathematical Society Lecture Note Series*, 352:312–321, 2008.

[15] V. Vu T. Tao. Random matrices: The universality phenomenon for wigner ensembles. *Proceedings of symposia in applied mathematics*, 72:121–172, 2014.

[16] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20:32–52, 1928.