Random commuting matrices

John E. McCarthy *

May 13, 2025

Abstract

We define a *random commuting d-tuple of n-by-n matrices* to be a random variable that takes values in the set of commuting d-tuples and has a distribution that is a rapidly decaying continuous weight on this algebraic set. In the Hermitian case, we characterize the eigenvalue distribution, and derive the limit as n tends to infinity for a Gaussian weight. In the non-Hermitian case, we get a formula that holds if the set is irreducible. We show that there are qualitative differences between the single matrix case and the several commuting matrices case.

1 Introduction

1.1 Overview

The study of random matrices, and in particular their eigenvalue distribution, is a well-established field, dating back to the pioneering work of Wigner [45, 46, 47], Dyson [16, 17, 18] and Ginibre [21]. For a comprehensive treatment, see e.g. the books [5, 6, 13, 14, 32, 34, 38, 42]. The field has many applications in mathematics and theoretical physics. Typically the entries of the matrix are chosen independently from each other (or for Hermitian matrices the upper triangular entries are chosen independently) according to some probabilitly distribution, and then one tries to characterize the distribution of specific properties of the matrix, such as the eigenvalues, or the norm, or the singular values, both in the finite case and in the limit as the size of the matrix goes to infinity. The best understood case in random matrix theory is when the entries of the matrix are chosen i.i.d. from a Gaussian distribution.

What would it mean to choose random matrices subject to some algebraic constraint? For a specific example, what would it mean to choose a random pair (X, Y) of commuting n-by-n matrices? One approach would be to first choose X by some random method, and then to choose something in its commutant. Generically, X will have n different eigenvalues, and Y will then have the same eigenvectors, so it would only be necessary to describe the distribution of the eigenvalues of Y. But what should that distribution be? Should it depend on the eigenvalues of X? And should the matrices X be chosen in such a way that ones with a 'larger' commutant are more heavily weighted? If the possible entries of X and Y

*Partially supported by National Science Foundation Grant DMS 2054199
AMS Subject Class: 60B20, 15B52.

came from a large but finite set, one could in principle choose a pair independently, and then keep it if and only if the pair commutes. But if the distribution is continuous, then the probability that any given pair would commute is zero, and one does not want to condition on a zero-probability event.

The purpose of this note is first to describe a systematic way to make sense of the notion of a random d-tuple of matrices that satisfies certain algebraic relations, and which can reduce to the i.i.d. Gaussian model if the relations are absent. Secondly, we shall explore this theory in the particular case of random commuting d-tuples. In the Hermitian case, we can derive the limiting distribution for a Gaussian weight; in the non-Hermitian case our results are far less complete.

1.2 What is a random d-tuple of matrices that satisfy algebraic relations?

Let $\mathbb{M}_n$ denote the algebra of n-by-n complex matrices, and $\mathbb{M}_n^d$ the set of d-tuples $X = (X^1, \ldots, X^d)$ of n-by-n matrices. We shall use superscripts to denote the components of X, subscripts to denote sequences, and paired subscripts the matrix entries (so X_{ij}^r is the (i,j) element of the r^{th} matrix in the d-tuple X).

Let $\mathfrak{V}_n \subseteq \mathbb{M}_n^d$ be a set that is unitarily invariant, by which we mean that if U is in the n-by-n unitary group, denoted $\mathcal{U}(n)$, and we define UXU^* componentwise:

$$UXU^* \ := \ (UX^1U^*, \ldots, UX^dU^*),$$

then if $X \in \mathfrak{V}_n$, so is UXU^*. Identifying $\mathbb{M}_n^d$ componentwise with $\mathbb{R}^{2n^2d}$, we will always assume that $\mathfrak{V}_n$ is locally a smooth submanifold of $\mathbb{M}_n^d$ except on a singular set of lower dimension.

One collection of examples is given by non-commutative algebraic sets, by which we mean there are non-commutative polynomials $p_1, \ldots, p_N$ in the $2d$ variables $\{x^1, (x^1)^*, \ldots, x^d, (x^d)^*\}$ so that

$$\mathfrak{V}_n \ = \ \{X \in \mathbb{M}_n^d : p_j(X) = 0 \ \forall \ 1 \leq j \leq N\}. \tag{1.1}$$

(Note that the polynomials p_j are applied to the matrices, not the entries of the matrices.) Any set $\mathfrak{V}_n$ defined as in (1.1) is unitarily invariant. We could also look at other conditions that are unitarily invariant, like having determinant one, or trace zero.

Example 1.1. Our first example is the set of commuting d-tuples, which we shall call $\mathfrak{C}_n^d$.

$$\mathfrak{C}_n^d \ = \ \{X \in \mathbb{M}_n^d : X^rX^s = X^sX^r \ \forall \ 1 \leq r, s \leq d\}.$$

The commuting Hermitian d-tuples we shall call $\mathfrak{C}_n^{d,H}$:

$$\mathfrak{C}_n^{d,H} \ = \ \{X \in \mathfrak{C}_n^d : (X^r)^* = X^r \ \forall \ 1 \leq r \leq d\}.$$

Example 1.2. A second example is the set of anti-commuting d-tuples:

$$\begin{aligned} \mathfrak{A}_n^d &= \{X \in \mathbb{M}_n^d : X^rX^s + X^sX^r = 0 \ \forall \ 1 \leq r \neq s \leq d\} \\ \mathfrak{A}_n^{d,H} &= \{X \in \mathfrak{A}_n^d : (X^r)^* = X^r \ \forall \ 1 \leq r \leq d\}. \end{aligned}$$

Example 1.3. We can take $d = 1$, and consider the set of nilpotent matrices:

$$\mathfrak{N}_n \;=\; \{X \in \mathbb{M}_n : (X)^n = 0\}.$$

We want to define a probability measure on $\mathfrak{V}_n$, and we want the measure to be unitarily invariant, since we do not want to privilege any particular choice of coordinates. Let the real dimension of $\mathfrak{V}_n$, which we will think of as a subset of $\mathbb{C}^{dn^2}$ with the usual Euclidean metric, be $d_{\mathfrak{V}_n}$. Let $dX|_{\mathfrak{V}_n}$, which we shall just write as dX when $\mathfrak{V}_n$ is understood, be Hausdorff measure of dimension $d_{\mathfrak{V}_n}$ on $\mathfrak{V}_n$. Let w_n be a unitarily invariant weight on $\mathfrak{V}_n$, by which we mean that $w_n : \mathfrak{V}_n \to [0, \infty)$ is a non-negative measurable function that satisfies

$$\begin{aligned} w_n(UXU^*) &= w_n(X) \qquad \forall\, U \in \mathcal{U}(n) \\ \int_{\mathfrak{V}_n} w_n(X)dX &= 1. \end{aligned}$$

Let μ_n be the probability measure $w_n(X)dX$ on V.

Definition 1.4. *A random d-tuple of n-by-n matrices in $\mathfrak{V}_n$ is a random variable with values in $\mathfrak{V}_n$ and distribution μ_n.*

The most important example for us is when the weight is Gaussian, which means it is of the form

$$w_n(X) \;=\; c_n e^{-\gamma \mathrm{tr}[\sum_{r=1}^d (X^r)^* X^r]} \;=\; c_n e^{-\gamma \sum_{r=1}^d \sum_{i,j=1}^n |X^r_{ij}|^2}, \tag{1.2}$$

where c_n is a normalizing constant.

Example 1.5. If $d = 1$ and we take $\mathfrak{V}_n$ to be the Hermitian matrices $\mathbb{M}_n^H$, then with the weight (1.2) we will recover the Gaussian Unitary Ensemble. With $\mathfrak{V}_n = \mathbb{M}_n$ we would get the complex Gaussian matrix ensemble. With $\mathfrak{V}_n = \mathcal{U}(n)$, we get Haar measure.

1.3 Why are random tuples of matrices interesting?

When $d = 2$, the study of random pairs of commuting Hermitian matrices is the same as the study of random normal matrices. These have been studied extensively—see e.g. [2, 3, 10, 19, 44, 33, 4, 26]—both for their mathematical interest and their applications to physics. To the best of my knowledge, the eigenvalue distribution for random commuting d-tuples has not been studied in the Hermitian case for $d > 2$ or the general case for $d > 1$. So I shall give some reasons I think the subject is worth pursuing, and list some open questions for the Gaussian case ((1.2) with the normalization $\gamma = \frac{1}{2}$).

If $X = (X^1, \ldots, X^d)$ is a commuting d-tuple of n-by-n matrices, then an eigenvector of X is a non-zero vector $v \in \mathbb{C}^n$ such that $X^j v = \lambda^j v$ for each $1 \le j \le d$. We call the d-tuple $(\lambda^1, \ldots, \lambda^d)$ an eigenvalue of X; it is a point in $\mathbb{C}^n$ (or $\mathbb{R}^n$ if each X^j is self-adjoint). If X has n distinct eigenvalues, $\lambda_1, \ldots, \lambda_n$, then the collection of all eigenvalues is a point in $(\mathbb{R}^d)^n$.

1. Large random Hermitian matrices were studied by Wigner and Dyson as models for the Hamiltonian of a complicated system with many weakly interacting parts, such as heavy atoms, and the eigenvalues were the possible energy levels. If one wished to study a collection of d such objects whose Hamiltonians commuted then tuples of random Hermitian matrices would provide a model.

2. Compact Lie groups have Haar measure as a natural probability measure. The construction in Subsection 1.2 allows one to define a probability measure on Lie algebras that are unitarily invariant. It is not so canonical, but at least for the Lie algebra of the unitary group (which is the skew-Hermitian matrices) one does get the Gaussian Unitary Ensemble (up to a factor of i).

3. If one thinks of random d-tuples of commuting Hermitian matrices as some sort of 'bosonic' model, then one can think of random d-tuples of anti-commuting Hermitians as a 'fermionic' model. For the case of a Hermitian anti-commuting pair (X, Y), the eigenvalues come naturally as pairs $(\pm\lambda, \pm\mu)$ where λ and μ are non-negative eigenvalues of X and Y respectively, so can be identified with points in $(\mathbb{R}_+)^2$. They are not joint eigenvalues; there is a two-dimensional subspace on which X and Y look like

$$\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}, \quad \begin{pmatrix} 0 & \mu \\ \mu & 0 \end{pmatrix}.$$

For finite n, the distribution was studied in [31], and the pairs (λ, μ) were called the skew spectrum.

Question 1.6. *What is the limiting distribution of the skew spectrum in $\mathfrak{A}_n^{2,H}$? Is it area measure on a quarter disk?*

Corollary 4.11 says that if X is chosen randomly in $\mathfrak{C}_n^{d,H}$ and one knows the distribution of eigenvalues for one matrix in the d-tuple, X^1 say, then one can deduce the value of d, provided n is large enough. One could imagine using this to count the number of elements in the set by studying one randomly chosen element very carefully. Is there a similar phenomenon in the anti-commuting case?

Question 1.7. *Let X_n be chosen randomly in $\mathfrak{A}_n^{d,H}$. Does the limiting distribution of eigenvalues of X_n^1 depend on d? Can it be explicitly described?*

4. The non-Hermitian case is, not surprisingly, much harder. When $d > 2$, it is no longer true that a generic[1] d-tuple of commuting n-by-n matrices will have n linearly independent eigenvectors [23], though this is true for $d = 2$ [35]. So if one wants to study the joint eigenvalues it would make sense to restrict one's attention to the component of $\mathfrak{C}_n^d$ that do have n linearly independent eigenvectors. Some partial progress is in Section 5 below, but it does not address the limit.

Question 1.8. *What is the limiting distribution of the joint eigenvalues of $(\frac{1}{\sqrt{n}}X, \frac{1}{\sqrt{n}}Y)$ when (X, Y) is chosen randomly in $\mathfrak{C}_n^2$?*

The limiting distribution of the eigenvalues of a single random matrix is area measure on a disk [32]; this is also the limiting distribution for a pair of commuting Hermitian matrices [19]. Is it possible that a pair of commuting matrices behaves like a 4-tuple of commuting Hermitian matrices? Theorem 5.5 shows that this is not the case when $d = 2$, but it might still be true in the limit.

[1] We use generic to mean a condition on elements of a set that only fails to hold on a subset of strictly smaller dimension.

Question 1.9. *Is the answer to Question 1.8 area measure on a sphere in* $\mathbb{C}^2$*?*

5. Multivariable operator theory studies d-tuples of commuting operators on separable Hilbert spaces. For many questions, it is sufficient to understand the finite dimensional case, if one can get a bound on the constants that is independent of the dimension. The random case may be more accessible and give new insights. For example, it is an open question whether Andô's inequality holds with a constant in 3 variables, i.e. whether there is a constant C so that for any n the inequality

$$\sup\{\|p(X)\| : X \in \mathfrak{C}_n^3, \max_{1\le r\le 3}\|X^r\| \le 1\} \;\le\; C\|p\|_{\mathbb{D}^3} \tag{1.3}$$

holds for every polynomial p in 3 variables. The fact that C cannot be chosen to be 1 was proved in [43] and [11];

Question 1.10. *Does* (1.3) *hold most of the time, in a probabilistic sense?*

6. There are deep ties between random matrices, free probability, and the von Neumann algebras of free groups—see e.g. [34]. If G is a group that is not free, one could hope that the group von Neumann algebra could be connected to the theory of random tuples that satisfy appropriate algebraic relations.

   To frame a specific question, suppose G is a discrete group with d generators $g_1, \ldots, g_d$ and N relations. For each relation R_j of the form $g_{i_1}^{\varepsilon_1} \ldots g_{i_k}^{\varepsilon_k} = e$, where each $\varepsilon_i = \pm 1$, let p_j be the non-commutative polynomial $p_j(x) = x_{i_1}^{[\varepsilon_1]} \ldots x_{i_k}^{[\varepsilon_k]} - 1$, where we write $x_i^{[\varepsilon]}$ to mean x_i if $\varepsilon = 1$ and x_i^* if $\varepsilon = -1$. Let $\mathfrak{V}_n$ be the set of d-tuples of n-by-n unitaries U satisfying $p_j(U) = 0$ for $1 \le j \le N$. On the group von Neumann algebra of G, let τ be the trace and let L_g be the image of g under the left regular representation on $\ell^2(G)$.

   Question 1.11. *Let* U_n *be a random tuple in* $\mathfrak{V}_n$*. For any word of the form* $g_{i_1}^{\varepsilon_1} \ldots g_{i_k}^{\varepsilon_k}$*, does* $\frac{1}{n}\mathrm{tr}[(U_n^{i_1})^{[\varepsilon_1]} \ldots (U_n^{i_k})^{[\varepsilon_k]})]$ *converge almost surely to* $\tau[L_{g_{i_1}}^{[\varepsilon_1]} \ldots (L_{g_{i_k}})^{[\varepsilon_k]}]$ *?*

   More generally, does the theory of random tuples connect to a theory analogous to free probability, but with some extra algebraic constraints, and also to the theory of von Neumann algebras of non-free groups?

7. Some algebraic relations cannot be attained on finite dimensional matrix algebras. For example, if $U, V \in \mathcal{U}(n)$ and $UV = e^{2\pi i\theta}VU$, then taking determinants we see that $e^{2\pi in\theta} = 1$, and in particular θ must be rational. However, if θ is irrational, Pimsner and Voiculescu [39] showed that if p_n/q_n is the sequence of best rational approximants to θ, then the irrational rotation C*-algebra A_θ can be embedded into an AF algebra that is the inductive limit of $\mathbb{M}_{q_n} \oplus \mathbb{M}_{q_{n-1}}$ under suitable connecting maps, and there are unitaries U_n, V_n on $\mathbb{M}_{q_n}$ satisfying $U_nV_n = e^{2\pi ip_n/q_n}V_nU_n$. (For more background and definitions of the terms used, see [12]). Let $\mathfrak{V}_n = \{(U, V) \in \mathbb{M}_{q_n} : UV = e^{2\pi ip_n/q_n}VU\}$.

   Question 1.12. *Does sampling random pairs in* $\mathfrak{V}_n$ *give information about* A_θ*?*

8. Our proposed framework can be used even when $d = 1$ to study single matrices with algebraic constraints. In [15], Dykema and Haagerup studied strictly upper triangular matrices, where all the elements were chosen to be i.i.d. $N(0, \frac{1}{\sqrt{n}})$. They showed that the limit converges in *-moments to a circular operator. Is this true if one loooks at all nilpotent matrices?

   Question 1.13. *Let X_n be chosen randomly in $\mathfrak{N}_n$. Does X_n converge in *-moments to a circular operator?*

   Every nilpotent matrix is unitarily equivalent to an upper triangular matrix, so choosing X_n is the same as choosing a unitary U_n and an upper-triangular T_n so that $X_n = U_n^* T_n U_n$. (This representation is not unique, but that doesn't matter for this discussion). The unitaries will be equally distributed with respect to Haar measure, since all our measures are unitarily invariant. But if X_n is chosen as in Definition 1.4 with a Gaussian weight on $\mathfrak{N}_n$, the distribution of $T_n = U_n^* X_n U_n$ will be different from the one in which all the entries are i.i.d. normal.

9. There are numerous quantities whose distribution one could seek associated with a d-tuple of matrices in addition to the eigenvalues, such as the joint spectral radius, the joint numerical range, various notions of singular numbers, etc. Let us just focus on the norm $\|X\| := \max_{1 \leq r \leq d} \|X^r\|$.

   Question 1.14. *If X_n is chosen randomly in $\mathfrak{C}_n^{d,H}$ does $\frac{1}{\sqrt{n}}\|X_n\|$ tend almost surely to R_d, where R_d is as in Theorem 2.1? If so, is there some version of the Tracy-Widom law describing how quickly it converges?*

   The numbers R_d decrease from 2 to 1 as d increases from 1 to 4, but then stay constant. Do the norms of random commuting non-Hermitian matrices also tend to a limit that is the same for all d above some threshold?

   Question 1.15. *If X_n is chosen randomly in $\mathfrak{C}_n^d$ does $\frac{1}{\sqrt{n}}\|X_n\|$ tend almost surely to some number T_d? Is T_d constant for $d \geq d_0$?*

10. The rate at which the eigenvalue distribution of a random n-by-n matrix converges to the limit has been extensively studied—see e.g. [30] for the case of a random Hermitian matrix, and [2, 3] for a random normal matrix. In both cases the fluctuation from the limit is shown to be a Gaussian variable. When $d \geq 4$, Theorem 4.10 says that the eigenvalue distribution converges to surface area measure on the sphere S^{d-1}. On the interior of the ball, since the limiting distribution is 0, the fluctuation must be non-negative.

    Question 1.16. *At what rate does the distribution of joint eigenvalues of a random d-tuple in $\frac{1}{\sqrt{n}}\mathfrak{C}_n^{d,H}$ on an open subset of the unit ball in $\mathbb{R}^d$ tend to 0?*

2 Main results

In this note, we shall consider random commuting d-tuples. There are two natural cases to consider, the Hermitian and the general case, which we denote by $\mathfrak{C}_n^{d,H}$ and $\mathfrak{C}_n^d$. Since commuting matrices have common eigenvectors, the eigenvalues of an element of $\mathfrak{C}_n^d$ are points in $\mathbb{C}^d$, and for an element of $\mathfrak{C}_n^{d,H}$ they are points in $\mathbb{R}^d$. We would like to understand how they are distributed, both for fixed n and as n tends to infinity.

For an alternative approach, where the matrices are not required to commute exactly but are penalized for having large commutators, see the recent paper by Guionnet and Maurel-Segala [22].

2.1 Main results: Hermitian case

We only consider weights which are unitarily invariant, i.e. $w(UXU^*) = w(X)$. We shall call unitarily invariant functions *class functions*. By Proposition 4.2, with probability 1 an X in $\mathfrak{C}_n^{d,H}$ will have exactly n distinct eigenvalues. Since the unitary orbits of X in $\mathfrak{C}_n^{d,H}$ are uniquely determined by their eigenvalues, we can define a function $\tilde{w}$ on $(\mathbb{R}^d)^n$ by $\tilde{w}(\lambda) = w(X)$, where X is any element of $\mathfrak{C}_n^{d,H}$ that has λ as its eigenvalues. Writing $\lambda = \lambda_j^r$ for $1 \le j \le n$ and $1 \le r \le d$, the order of the j's does not matter, but the order of the r's does.

Our first theorem is a d-dimensional analogue of the Ginibre formula. We write $|\lambda_i - \lambda_j|$ to denote the Euclidean distance between λ_i and λ_j in $\mathbb{R}^d$.

Theorem 4.5: *Let $w : \mathfrak{C}_n^{d,H} \to [0,\infty)$ be a continuous integrable class function, and assume $\int_{\mathfrak{C}_n^{d,H}} w(X)dX = 1$. Let random matrices in $\mathfrak{C}_n^{d,H}$ be chosen with distribution $w(X)dX$. Then the density of their eigenvalues is given by*

$$\rho_n(\lambda) \;=\; \beta_n \; \tilde{w}(\lambda) \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^2. \tag{2.1}$$

Note that the repulsion between the eigenvalues is the square of the Euclidean distance in $\mathbb{R}^d$, not of the components.

Using Johansson's large deviation estimate [30], we show that as $n \to \infty$ the distribution tends to the equilibrium measure for the logarithmic energy. These have been described explicitly for $d \ge 3$ by Chafai, Saff and Womersley [8, 9], and for $d = 1$ and 2 in [47] and [19]. This yields the following result (see Theorem 4.10 for a precise statement).

Theorem 2.1. *Let X_n be a random variable in $\mathfrak{C}_n^{d,H}$, and let the weights be probabilities given by $w_n(X) \;=\; c_n e^{-\frac{1}{2}\mathrm{tr}[\sum_{r=1}^d (X^r)^* X^r]}$. As $n \to \infty$, the eigenvalues of $\frac{1}{\sqrt{n}} X_n$ tend to the following equilibrium distribution on the closed ball of radius R_d, where σ^d is normalized area*

measure on the sphere, and $d\varrho$ means radial measure.

$$\begin{aligned}
\frac{1}{2\pi}\sqrt{4-x^2}\,dx, &\quad R_1 = 2 \quad d=1;\\
\frac{1}{2\pi}1_{|x|<\sqrt{2}}\,dx^1dx^2, &\quad R_2=\sqrt{2} \quad d=2;\\
\frac{3}{4\pi^2}\frac{1}{\sqrt{\frac{4}{3}-|x|^2}}1_{|x|<R_3}\,d\varrho d\sigma^2, &\quad R_3=\sqrt{\tfrac{4}{3}} \quad d=3;\\
\sigma^{d-1}_{R_d}, &\quad R_d := 1, \quad d\geq 4.
\end{aligned}$$

A corollary of this result (see Corollary 4.11) is that if one chooses a random commuting Hermitian d-tuple in $\mathfrak{C}^{d,H}_n$ and then chooses randomly one of the d matrices in the d-tuple, then the normalized eigenvalues of this matrix will tend, as $n\to\infty$, to a semi-circular distribution on $[-R_d, R_d]$ for $d\leq 4$ but will be more heavily concentrated near 0 for any $d\geq 5$.

2.2 Main results: General case

In the non-Hermitian case, the set $\mathfrak{C}^d_n$ is irreducible if $d\leq 2$ or $n\leq 3$ or for a certain range of n when $d=3$ (see Subsection 5.1 for details). We restrict our attention to the irreducible case, which happens when the elements with n eigenvalues are generic. In this case we can write a generic $X\in\mathfrak{C}^d_n$ as

$$X \;=\; UQU^* \;=\; UADA^{-1}U^*,$$

where U is a unitary in $\mathcal{U}(n)$, Q is an upper triangular commuting d-tuple, D is a diagonal d-tuple and A is an upper triangular unipotent matrix.

In Proposition 5.3 we describe how to take advantage of the unitary invariance of w_n to decompose $w_n(X)dX$ as $w_n(Q)\kappa(Q)dQd\nu$, where ν is the invariant measure on $\mathcal{U}(n)/\mathbb{T}^n$ and κ is a function described in Subsection 5.2.

In Theorem 5.4 we describe how to get the eigenvalue distribution of a random element in $\mathfrak{C}^d_n$ by integrating with respect to both U and A. The formula is unfortunately complicated. In Theorem 5.5 we calculate the distribution explicitly for commuting d-tuples of 2-by-2 matrices. We get:

Theorem 5.5: *Let X be a random element of $\mathfrak{C}^d_2$ chosen with distribution*

$$w(X) \;=\; Ce^{-\gamma\|X\|^2_F}.$$

Then the eigenvalues λ of X in $(\mathbb{C}^d)^2$ have distribution

$$\rho^d(\lambda) \;=\; C'_d\; e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)}|\lambda_2-\lambda_1|^2\sum_{j=0}^{d-1}\frac{1}{(d-1-j)!}\frac{2^j}{\gamma^{j+1}}\frac{1}{|\lambda_2-\lambda_1|^{2j}}.$$

For $d = 1, 2, 3$ we get respectively

$$\begin{aligned}
\rho^1(\lambda) &= C_1' e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{\gamma}|\lambda_2 - \lambda_1|^2\right] \\
\rho^2(\lambda) &= C_2' e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{\gamma}|\lambda_2 - \lambda_1|^2 + \frac{2}{\gamma^2}\right] \\
\rho^3(\lambda) &= C_3' e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{2\gamma}|\lambda_2 - \lambda_1|^2 + \frac{2}{\gamma^2} + \frac{4}{\gamma^3}|\lambda_2 - \lambda_1|^{-2}\right].
\end{aligned}$$

Notice that, unlike in the Hermitian case or the case $d = 1$, the negative powers that appear for $d \geq 3$ show that there is an *attraction* between the eigenvalues, not a repulsion.

3 Known results

3.1 The Gaussian Unitary Ensemble

The Gaussian Unitary Ensemble (or GUE) is the matrix ensemble in which the strictly upper triangular entries $(\xi_{ij})_{i<j}$ are chosen i.i.d. $N_{\mathbb{C}}(0, 1)$ (ie with a complex normal distribution with mean 0 and variance 1), the diagonal entries ξ_{ii} are chosen independently from $N_{\mathbb{R}}(0, 1)$, and the lower triangular entries are defined by $\xi_{ji} = \overline{\xi_{ij}}$. This gives an ensemble of Hermitian matrices, parametrized by $\mathbb{R}^{n^2}$. The probability measure is given by

$$C_n e^{-\frac{1}{2}\sum_{i,j=1}^n |X_{ij}|^2} dm(X) \tag{3.1}$$

where m is Lebesgue measure on $\mathbb{R}^{n^2}$, and C_n is the appropriate normalizing constant.

The Ginibre formula [21] asserts that the density function of the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$ in $\mathbb{R}^n$ is given by

$$\rho_n(\lambda) \;=\; C_n' e^{-|\lambda|^2/2} \prod_{1 \leq 1 < j \leq n} |\lambda_i - \lambda_j|^2. \tag{3.2}$$

The Wigner semi-circular law says that if one defines a measure by putting a point mass of weight $1/n$ at each eigenvalue of $\frac{1}{\sqrt{n}}X$ then these measures converge weakly to the semi-circular distribution on $[-2, 2]$ given by

$$\frac{1}{2\pi}\sqrt{4 - x^2}dx.$$

See e.g. [13, Thm. 6.96 and Prop. 6.156] for a proof.

If instead of choosing each matrix entry independently one chooses Hermitian matrices with a distribution that replaces (3.1) by

$$C_n e^{-\mathrm{tr}(Q(X))} dX,$$

where $Q(x) = \gamma|x|^\alpha + O(|x|^{\alpha-1})$ for some positive constants α, γ, then the Ginibre formula still holds, becoming

$$\rho_n(\lambda) \;=\; C_n' e^{Q(\lambda)} \prod_{1 \leq 1 < j \leq n} |\lambda_i - \lambda_j|^2$$

[13, Thm. 5.22]. The semi-circle distribution is replaced by the equilibrium measure for Q [13, Thm. 6.96] (see Subsection 3.2 for a definition of equilibrium measure).

3.2 Equilibrium measures for the logarithmic potential

Let $Q(x)$ be an external field on $\mathbb{R}^d$, a non-negative continous function that goes to infinity like some positive power of $|x|$. For μ a probability measure on $\mathbb{R}^d$, define the logarithmic energy by

$$I^Q(\mu) \;=\; \int\int \log \frac{1}{|x-y|} d\mu(x) d\mu(y) + \int Q(x) d\mu(x). \tag{3.3}$$

The equilibrium measure for Q is the unique probability measure μ^Q that minimizes $I^Q(\mu)$ (for existence and uniqueness, see [7, Thm. 4.4.14]).

Let us restrict to the case $Q(x) = \gamma|x|^\alpha$. When $\alpha = 2$, the equilibrium measures are known for all d. Let σ^d be area measure on the unit sphere S^d, normalized to have total mass 1, and σ_R^d be normalized area measure on the sphere of radius R.

Theorem 3.1. *Let $Q(x) = \gamma|x|^2$. Then the equilibrium measure is given by*

$$\frac{2}{\pi R_1^2}\sqrt{(R_1^2 - x^2)_+}\; dx, \qquad R_1 := \sqrt{\tfrac{2}{\gamma}}, \quad d = 1; \tag{3.4}$$

$$\frac{1}{\pi R_2^2} 1_{|x|<R_2}\; dx^1 dx^2, \qquad R_2 := \tfrac{1}{\sqrt{\gamma}}, \quad d = 2; \tag{3.5}$$

$$\frac{1}{\pi^2 R_3^2}\frac{1}{\sqrt{R_3^2 - |x|^2}} 1_{|x|<R_3}\; dr d\sigma^2, \quad R_3 := \sqrt{\tfrac{2}{3\gamma}}, \quad d = 3; \tag{3.6}$$

$$\sigma_{R_d}^{d-1}, \qquad R_d := \tfrac{1}{\sqrt{2\gamma}}, \quad d \geq 4. \tag{3.7}$$

Result (3.4) is the semi-circle law, first obtained by Wigner [47]. For other values of α when $d = 1$, see [13, Sec. 6.7]. Result (3.5) is the circular law for random normal matrices; it was proved by Elbau and Felder in [19]; for another proof see [42]. Results (3.6) and (3.7) are due to Chafai, Saff and Womersley [8, 9]. Among other results they also show that the equilibrium for all $\alpha \geq 2$ and all $d \geq 4$ is given by $\sigma^{d-1}_{(1/\gamma\alpha)^{1/\alpha}}$.

3.3 Other results

The Hoffman-Weilandt inequality [27] says that if we let $\mu_j(A)$ denote the j^{th} largest eigenvalue of the self-adjoint matrix A, then for any self-adjoint n-by-n matrices A and B we have

$$\sum_j |\mu_j(A) - \mu_j(B)|^2 \leq \|A - B\|_F^2. \tag{3.8}$$

4 Random commuting Hermitian matrices

4.1 Generic d-tuples

In this section, we shall study random d-tuples in $\mathfrak{C}_n^{d,H}$. All matrices in this section are Hermitian, and dimension always means real dimension. First we shall prove that the set of X for which any of the matrices has a repeated eigenvalue is of lower dimension than $\mathfrak{C}_n^{d,H}$, so will have measure 0.

Definition 4.1. *Let Y be a self-adjoint matrix. We shall say Y has banner $(r_1, r_2, \ldots, r_p)$, where $r_1 \geq r_2 \geq \cdots \geq r_p \geq 1$ are integers, if Y has p distinct eigenvalues, with multiplcities $r_1, \ldots, r_p$.*

Let $\mathcal{G}(d, n)$ denote the complex Grassmanian of all d-dimensional linear subspaces of $\mathbb{C}^n$. It has dimension $2d(n-d)$.

Proposition 4.2. *For $d \geq 2$, the set $\{X \in \mathfrak{C}_n^{d,H} : \text{ banner}(X^1) = (r_1, \ldots, r_p)\}$ has dimension $n^2 + (d-2)n + p$.*

Proof. We shall prove this by induction on d. First, let $d = 1$. To choose a matrix with banner $(r_1, \ldots, r_p)$, one can first choose an r_1 dimensional subspace, and the first eigenvalue, then an r_2 dimensional subspace of $\mathbb{C}^{n-r_1}$ and the second eigenvalue, and so on. This yields a space of dimension

$$\begin{aligned} 2r_1(n-r_1) + 2r_2(n-r_1-r_2) + \cdots + 2r_{p-1}(n-r_1-\cdots-r_{p-1}) + p & \\ = \; 2n(n-r_p) - 2\sum_{i=1}^{p-1}\sum_{j=1}^{i} r_i r_j + p & \\ = \; 2n(n-r_p) - \left(\sum_{i=1}^{p-1} r_i\right)^2 - \sum_{i=1}^{p-1} r_i^2 + p & \\ = \; 2n(n-r_p) - (n-r_p)^2 - \sum_{i=1}^{p} r_i^2 + r_p^2 + p & \\ = \; n^2 - \sum_{i=1}^{p} r_i^2 + p. & \end{aligned} \tag{4.1}$$

Note that $n^2 - \sum_{i=1}^p r_i^2 + p$ is always $\leq n^2$, and equals n^2 if each r_i is 1 (which means $p = n$).

Now assume $d \geq 2$, and that we have shown that if $d' < d$, the dimension of $\mathfrak{C}_n^{d',H}$ is $n^2 + (d'-1)n$. Note that (4.1) proves this assertion for $d' = 1$, and if $d' > 1$ it follows from the inductive hypothesis, since the maximum value that $n^2 + (d'-2)n + p$ can attain is $n^2 + (d'-1)n$, which occurs when $p = n$. To calculate the dimension of $\{X \in \mathfrak{C}_n^{d,H} : \text{banner}(X^1) = (r_1, \ldots, r_p)\}$, note that once X^1 has been chosen, then the dimension of $\{X^2, \ldots, X^d\}$ is $\sum_{i=1}^p \dim(\mathfrak{C}_{r_i}^{d-1,H})$. By the inductive hypothesis, this gives

$$\begin{aligned} \dim\{X \in \mathfrak{C}_n^{d,H} : \text{ banner}(X^1) = (r_1, \ldots, r_p)\} &= n^2 - \sum_{i=1}^{p} r_i^2 + p + \sum_{i=1}^{p} (r_i^2 + (d-2)r_i) \\ &= n^2 + (d-2)n + p. \end{aligned}$$

This proves the proposition for d, and in particular shows that the dimension of $\mathfrak{C}_n^{d,H}$ is $n^2 + (d-1)n$. □

It follows from Proposition 4.2 that the set of $X \in \mathfrak{C}_n^{d,H}$ for which each X^r has n distinct eigenvalues is of full measure. Let us call these the generic elements, and write

$$\mathfrak{C}_{n,\text{gen}}^{d,H} = \{X \in \mathfrak{C}_n^{d,H} : \text{ each } X^r \text{ has } n \text{ distinct eigenvalues}\}.$$

Note that at any point where the eigenvalues of X are distinct (in particular at the generic elements) the algebraic set $\mathfrak{C}_n^{d,H}$ is smooth, since coordinates can be given by specifying an orthonormal frame for the eigenvectors and a choice of corresponding eigenvalues.

We could choose to order the eigenvalues of X, which are points in $\mathbb{R}^n$, by, for example, putting the eigenvalues in dictionary order. However, we will not do this, and so the eigenvalues of elements of $\mathfrak{C}_{n,\mathrm{gen}}^{d,H}$ will be invariant under the symmetric group $\mathfrak{S}(n)$, and when we list them as $(\lambda_1, \ldots, \lambda_n)$ this is with some random ordering.

4.2 Ginibre formula

Next we prove the d-dimensional version of the Ginibre formula (3.2). Both the proof in [42, Sec. 2.6.1] and in [13, Thm. 5.22] of the one dimensional case can be adapted to d dimensions, though some care must be taken since the tangent space to $\mathfrak{C}_n^{d,H}$ can no longer be naturally identified with $\mathfrak{C}_n^{d,H}$. In the d-variable case, the eigenvalues of an elements of $\mathfrak{C}_n^{d,H}$ will be points in $(\mathbb{R}^d)^n$, and we will write them as $\lambda = (\lambda_1, \ldots, \lambda_n)$ where each λ_i is a d-tuple $\lambda_i = (\lambda_i^1, \ldots, \lambda_i^d)$.

If w is a real-valued function on $\mathfrak{C}_{n,\mathrm{gen}}^{d,H}$ that is invariant under the action of the unitary group $\mathcal{U}(n)$ (ie. $w(X) = w(U^*XU)$ for all $U \in \mathcal{U}(n)$) then $w(X)$ can only depend on the eigenvalues of X, since they completely determine the unitary equivalence class of X. Such functions are called *class functions.* If w is a class function, there is a function $\tilde{w} : (\mathbb{R}^d)^n \to \mathbb{R}$ that is invariant with respect to the action of $\mathfrak{S}(n)$ so that

$$w(X) \;=\; \tilde{w}(\lambda),$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$ is some ordering of the eigenvalues of X. Moreover, if w is continuous, so is $\tilde{w}$.

For $x \in \mathbb{R}^d$, we shall use $|x|$ to mean the Euclidean norm $\sqrt{\sum_{r=1}^d |x^r|^2}$. We shall write dX to mean $d\mathcal{H}_{\dim \mathfrak{C}_n^{d,H}}(X)$. We shall use the subscript F for the Frobenius (or Hilbert-Schmidt) norm on $\mathfrak{C}_n^d$, viz.

$$\|X\|_F^2 \;=\; \sum_{r=1}^{d} \sum_{i,j=1}^{n} |X_{ij}^r|^2.$$

Let $v(k)$ denote the volume of the unit ball in $\mathbb{R}^k$.

Conjugating a diagonal d-tuple by a unitary gives the most general Hermitian d-tuple, but diagonal unitaries have no effect, so we must understand $\mathcal{U}(n)/\mathbb{T}^n$, where we identify a point $t \in \mathbb{T}^n$, the set of n-tuples of unimodular complex numbers, with the diagonal unitary D_t that has the entries of t on the diagonal. We shall write 1_n for the identity in $\mathbb{T}^n$. The tangent space at the identity of $\mathcal{U}(n)$ is the Lie algebra $\mathfrak{u}(n)$ which consists of the skew-Hermitian matrices (matrices satisfying $S^* = -S$). Let $\mathcal{A}_{0,n}$ denote the subspace of skew-Hermitian matrices that are 0 on the diagonal.

Lemma 4.3. *Let $g : \mathcal{A}_{0,n} \times \mathbb{T}^n \to \mathcal{U}(n)$ be defined by $g(S,t) = e^S D_t$. Then $dg : \mathcal{A}_{0,n} \times \mathbb{R}^n \to \mathfrak{u}(n)$ is an isomorphism between the tangent space of $g : \mathcal{A}_{0,n} \times \mathbb{T}^n$ at $(0, 1_n)$ and the tangent space of $\mathcal{U}(n)$ at the identity.*

Proof. Note that

$$dg|_{(0,1_n)} : (S, \theta) \;\mapsto\; S + D_{i\theta},$$

and that this map is an isomorphism. □

In particular, by the inverse function theorem one can parametrize a neighborhood of $[\mathbb{T}^n]$ in $\mathcal{U}(n)/\mathbb{T}^n$ as $\{e^S\}$ where S ranges over a neighborhood of 0 in $\mathcal{A}_{0,n}$. Define a constant β_n by

$$\beta_n \;=\; \lim_{\varepsilon \to 0^+} \frac{\mathrm{vol}_{\mathcal{A}_{0,n}}\{S : S \in \mathcal{A}_{0,n}, \|S\|_F < \varepsilon\}}{\mathrm{vol}_{\mathcal{U}(n)/\mathbb{T}^n}\{e^S : S \in \mathcal{A}_{0,n}, \|S\|_F < \varepsilon\}}. \tag{4.2}$$

The number $\frac{1}{\beta_n}$ is the Jacobian determinant of e^S, normalized by the requirement that the total Haar measure of $\mathcal{U}(n)$ is 1.

Lemma 4.4. *Let $X \in \mathfrak{C}^{d,H}_{n,\mathrm{gen}}$. The tangent space $T_X\mathfrak{C}^{d,H}_n$ is the set $Z = (Z^1, \ldots, Z^d)$ of d-tuples of self-adjoint matrices that satisfy*

$$[X^r, Z^s] \;=\; [X^s, Z^r] \qquad \forall\ 1 \le r \ne s \le d. \tag{4.3}$$

If e_j is an orthonormal basis of eigenvectors of X with corresponding eigenvalues λ_j, then with respect to this basis equations (4.3) *become*

$$Z^s_{ij}(\lambda^r_j - \lambda^r_i) \;=\; Z^r_{ij}(\lambda^s_j - \lambda^s_i) \qquad \forall\ 1 \le r \ne s \le d. \tag{4.4}$$

This result has been known for a long time—we do not know where it first appeared. For the convenience of the reader, we include the proof from [1, Lemma 7.3]. (One could also adapt the proof of Lemma 5.2 below).

Proof. We want to show that there exists a C^1 curve $X(t)$ of commuting self-adjoints with $X(0) = X$ and $X'(0) = Z$ if and only if (4.4) holds.

($\Rightarrow$): If $X(t) = X + tZ + o(t)$ is commutative, calculate

$$[X^r(t), X^s(t)] \;=\; t\,([X^r, Z^s] - [X^s, Z^r]) + o(t).$$

The coefficient of t must vanish, giving (4.4).

($\Leftarrow$): Suppose (4.4) holds. This means

$$Z^s_{ij}(\lambda^r_j - \lambda^r_i) \;=\; Z^r_{ij}(\lambda^s_j - \lambda^s_i) \qquad \forall\, r \ne s,$$

so

$$Z^r_{ij}\frac{1}{\lambda^r_j - \lambda^r_i} \;=\; Z^s_{ij}\frac{1}{\lambda^s_j - \lambda^s_i} \qquad \text{if } \lambda^r_j - \lambda^r_i \ne 0 \ne \lambda^s_j - \lambda^s_i. \tag{4.5}$$

Define a skew-selfadjoint matrix Y by

$$Y_{ij} \;=\; Z^r_{ij}\frac{1}{\lambda^r_j - \lambda^r_i} \qquad \text{for any } r \text{ such that } \lambda^r_j - \lambda^r_i \ne 0. \tag{4.6}$$

For any $i \ne j$, there is some r with $\lambda^r_j - \lambda^r_i \ne 0$, so (4.6) defines Y_{ij}; and (4.5) says it doesn't matter which r we choose. Let all the diagonal terms of Y be 0.

Define

$$X^r(t) \;=\; e^{tY}(X^r + t\ \mathrm{diag}(Z^r))e^{-tY}.$$

Since e^{tY} is a unitary matrix and $X^r + t\ \mathrm{diag}(Z^r)$ is diagonal, $X(t) \in CXAM_n^d$ and

$$\frac{d}{dt}X^r(t)|_{t=0} = [Y, X^r] + \mathrm{diag}(Z^r) = Z^r.$$

□

Theorem 4.5. *Let $w : \mathfrak{C}_n^{d,H} \to [0,\infty)$ be a continuous integrable class function, and assume $\int_{\mathfrak{C}_n^{d,H}} w(X)dX = 1$. Let random matrices in $\mathfrak{C}_n^{d,H}$ be chosen with distribution $w(X)dX$. Then the density of their eigenvalues is given by*

$$\rho_n(\lambda) \;=\; \beta_n\ \tilde{w}(\lambda) \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^2. \tag{4.7}$$

where β_n is given by (4.2).

The theorem asserts that if we wish to integrate some class function with respect to $w(X)dX$, we can first integrate over $\mathcal{U}(n)/\mathbb{T}^n$, and then integrate over the eigenvalues. This gives the decomposition

$$w(X)dX \;=\; \rho_n(\lambda)d\nu d\lambda \tag{4.8}$$

where $d\lambda$ is Lebesgue measure on $(\mathbb{R}^d)^n$ and $d\nu$ is volume measure on $\mathcal{U}(n)/\mathbb{T}^n$, with ρ_n given by (4.7). Notice that the repulsion between the eigenvalues is based on their Euclidean separation in $\mathbb{R}^d$.

Let E_{ij} denote the matrix which is 1 in the (i,j) entry and 0 elsewhere.

Proof. The idea is to decompose $w(X)dX$ as in (4.8).

As every element in $\mathfrak{C}_n^{d,H}$ is unitarily equivalent to a diagonal tuple, it is sufficient to look at a neighborhood of a diagonal tuple. Fix some orthonormal basis of $\mathbb{C}^n$ and choose a diagonal tuple D in $\mathfrak{C}_{n,\mathrm{gen}}^{d,H}$ with diagonal entries $\lambda = (\lambda_1, \ldots, \lambda_n)$ in $(\mathbb{R}^d)^n$. Let ε be a sufficiently small positive number, and consider $\{X : \|X - D\|_F < \varepsilon\}$.

By the Hoffman-Wielandt inequality (3.8), the eigenvalues of X can be ordered as $\kappa = (\kappa_1, \ldots, \kappa_n)$ where κ_j is the eigenvalue of X closest to λ_j, and

$$\sum_{r=1}^{d}\sum_{j=1}^{n} |\kappa_j^r - \lambda_j^r|^2 \;\le\; \|X - D\|_F^2 \;<\; \varepsilon^2.$$

By Lemma 4.3, we can parametrize a small neighborhood of D by

$$X \;=\; G(\theta, S) \;=\; e^S(D + D_\theta)e^{-S}$$

where S ranges over a neighborhood of 0 in $\mathcal{A}_{0,n}$ and θ over a neighborhood of 0 in $(\mathbb{R}^d)^n$. The derivative dG at $(0,0)$ is the map between the tangent spaces

$$\begin{aligned} dG : (\mathbb{R}^d)^n \times \mathcal{A}_{0,n} &\to T_D\mathfrak{C}_n^{d,H} \\ (\theta, S) &\mapsto SD - DS + D_\theta. \end{aligned}$$

To calculate the Jacobian determinant of G, let $R_{ij} = \frac{1}{\sqrt{2}}(E_{ij} - E_{ji})$ and $S_{ij} = \frac{i}{\sqrt{2}}(E_{ij} + E_{ji})$. Then we can use the real coordinates θ for $(\mathbb{R}^d)^n$ and $\{R_{i,j}, S_{ij} : i < j\}$ for $\mathcal{A}_{0,n}$. We have

$$\begin{aligned} dG : (\theta, 0) &\mapsto \theta \\ dG : (0, R_{ij}) &\mapsto -i(\lambda_j^r - \lambda_i^r) S_{ij} \\ dG : (0, S_{ij}) &\mapsto i(\lambda_j^r - \lambda_i^r) R_{ij}. \end{aligned}$$

Therefore the Jacobian of G is

$$\sqrt{|\det(dG^* dG)|} \;=\; \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^2.$$

This gives

$$dX \;=\; \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^2 \; d\theta dS.$$

By Lemma 4.3, dS is locally $\beta_n d\nu$. Thus we have shown that

$$w(X)dX = \beta_n \tilde{w}(\lambda) \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^2 d\lambda d\nu,$$

which gives (4.7). □

Remark 4.6. As the constant β_n depends only on the unitary group $\mathcal{U}(n)$ and not on d or w, it can be calculated in a number of ways; see e.g. [21, 42].

4.3 Density of eigenvalues

In Theorem 4.5 we gave a formula for the joint distribution of eigenvalues in $(\mathbb{R}^d)^n$. Now we ask what is the density of eigenvalues in $\mathbb{R}^d$, as $n \to \infty$. This will depend on the equilibrium measure of $\log \tilde{w}$, and choosing the correct scaling. We shall restrict ourselves to weights of the following form, where $\alpha, \gamma > 0$:

$$Q(x) \;=\; \gamma(\sum_{r=1}^{d} |x^r|^2)^{\alpha/2}. \tag{4.9}$$

The following theorem is a special case of [7, Thm. 4.4.14].

Theorem 4.7. *Let Q be as in* (4.9)*, and let the logarithmic energy be defined as in* (3.3)*. Then there is a unique probability measure μ^Q on $\mathbb{R}^d$ where* $\inf I^Q(\mu)$ *is attained, and moreover μ^Q is compactly supported.*

We call μ^Q the equilibrium measure (for Q) and set

$$E^Q \;=\; I^Q(\mu^Q).$$

We define w_n on $\mathfrak{C}_n^{d,H}$ by

$$w_n(X) \;=\; C_n \; e^{-\mathrm{tr}\; Q(X)},$$

where C_n is a normalizing constant to ensure $\int_{\mathfrak{C}_n^{d,H}} w_n(X)dX = 1$. If $X \in \mathfrak{C}_{n,\mathrm{gen}}^{d,H}$ and λ is some ordering of the eigenvalues, then

$$\tilde{w}_n(\lambda) \;=\; C_n e^{-\sum_{j=1}^n Q(\lambda_j)}.$$

Let $\rho = \rho_n$ be as in (4.7), so

$$\rho_n(x) \;=\; \beta_n C_n \; e^{-\sum_{j=1}^n Q(x_j)} \prod_{1\le i<j\le n} |x_i - x_j|^2.$$

To find the distribution of a single eigenvalue, we must integrate out the other $n-1$ eigenvalues. First we rescale by $y = n^{-\frac{1}{\alpha}}x$, defining a new probability density on $(\mathbb{R}^d)^n$ by

$$\mathcal{P}_n(y) \;=\; D_n \, e^{-n\sum_{j=1}^n Q(y_j)} \prod_{1\le i<j\le n} |y_i - y_j|^2. \tag{4.10}$$

Here D_n is a normalizing constant to make $\int_{(\mathbb{R}^d)^n} \mathcal{P}_n(y)dy = 1$. It is related to the previous constants by

$$D_n \;=\; \left(n^{\frac{1}{\alpha}}\right)^{n^2+(d-1)n} \; \beta_n C_n.$$

If ϕ is a bounded continuous function on $\mathbb{R}^d$, we define an expectation by

$$\mathcal{E}_n(\phi) \;:=\; \int_{(\mathbb{R}^d)^n} \frac{1}{n}\sum_{i=1}^n \phi(y_i)\mathcal{P}_n(y)dy.$$

The principal result of this section is the following theorem.

Theorem 4.8. *Let Q be as in 4.9, and let ϕ be a bounded and continuous function on $\mathbb{R}^d$. Then*

$$\lim_{n\to\infty} \mathcal{E}_n(\phi) \;=\; \int_{\mathbb{R}^d} \phi(x)d\mu^Q(x). \tag{4.11}$$

Theorem 4.8 was proved in the $d = 1$ case by K. Johansson [30], and his proof generalizes to the d-dimensional case. For the convenience of the reader we shall give a complete proof, following the exposition in Section 6.4 of the beautiful monograph [13], but we emphasize that generalizing to d dimensions requires little more than notational changes.

We start with a large deviation estimate, which for $d = 1$ is Lemma 4.2 in [30] or Lemma 6.67 in [13]. Our proof is a straightforward generalization, but since the proof that Claim 4.13 holds must be modified for $d \ge 2$, we include the whole proof for completeness.

Define, for $s, t \in \mathbb{R}^d$,

$$k(s,t) \;:=\; \log\frac{1}{|s-t|} + \frac{1}{2}(Q(s)+Q(t)),$$

and for $y = (y_1, \ldots, y_n) \in (\mathbb{R}^d)^n$,

$$\begin{aligned} K_n(y) \;&=\; \sum_{i\ne j} k(y_i, y_j) \\ &=\; \sum_{1\le i<j\le n} \log|y_i - y_j|^{-2} + (n-1)\sum_{i=1}^n Q(y_i). \end{aligned} \tag{4.12}$$

Let

$$A_{n,\eta} \ := \ \{y \in (\mathbb{R}^d)^n : \frac{1}{n^2} K_n(y) \leq E^Q + \eta\}.$$

Lemma 4.9. *Let $\mathcal{P}_n(y)$ be the probability density given by (4.10). Let $\eta, a > 0$. Then there exists n_0 which depends on η but not on a so that*

$$\int_{(\mathbb{R}^d)^n \setminus A_{n,\eta+a}} 1 \ d\mathcal{P}_n \ \leq \ e^{-an^2}, \quad n \geq n_0.$$

Proof. Assume $d \geq 2$, as the $d = 1$ case is proved in [30]. Let $0 < \varepsilon < \eta/2$. For $\delta > 0$, let

$$\psi_\delta(y) \ = \ \frac{1}{v(d)\delta^d} \int_{B(y,\delta)} d\mu^Q(x),$$

where $v(d)$ denotes the volume of the unit ball in $\mathbb{R}^d$, and $B(y, \delta)$ means the ball centered at y of radius δ. Then ψ_δ is a compactly supported function, and $\psi_\delta(y)dy$ tends vaguely to μ^Q as $\delta \to 0$. Warning: ψ_δ need not be continuous, for example if μ^Q puts mass on a sphere of radius δ.

Claim: For δ sufficiently small, we have

$$I^Q(\psi_\delta) \ \leq \ E^Q + \frac{\varepsilon}{2}. \tag{4.13}$$

Proof of Claim:

$$\begin{aligned} &\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \log|x-y| \psi_\delta(x)\psi_\delta(y)dxdy \\ &= \int\int \left[\frac{1}{v(d)^2\delta^{2d}} \int_{B(0,\delta)} \int_{B(0,\delta)} \log|x-y+s-t| \ dsdt \right] d\mu^Q(x) d\mu^Q(y). \end{aligned}$$

Since $d \geq 2$, the function $\log|x|$ is subharmonic. Applying subharmonicity of $\log|x|$ twice, we get

$$\frac{1}{v(d)^2\delta^{2d}} \int_{B(0,\delta)} \int_{B(0,\delta)} \log|x-y+s-t| \ dsdt \ \geq \ \log|x-y|.$$

Therefore we conclude

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \log \frac{1}{|x-y|} \psi_\delta(x)\psi_\delta(y)dxdy \ \leq \ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \log \frac{1}{|x-y|} d\mu^Q(x) d\mu^Q(y).$$

As

$$\lim_{\delta \to 0} \int Q(x)\psi_\delta(x)dx \ = \ \int Q(x) d\mu^Q(x),$$

we have proved the claim.

The rest of the proof proceeds exactly as in [13]. Fix δ so that (4.13) holds. Let E be the compact support of ψ_δ. We have

$$\begin{aligned} \frac{1}{D_n} &= \int_{(\mathbb{R}^d)^n} e^{-K_n(y) - \sum_{i=1}^n Q(y_i)} dy \\ &\geq \int_{E^n} e^{-K_n(y) - \sum_{i=1}^n Q(y_i) - \sum_{i=1}^n \log \psi_\delta(y_i)} \prod_{i=1}^n \psi_\delta(y_i) dy. \end{aligned}$$

Fix some positive δ for which (4.13) holds. Applying Jensen's inequality to the probability measure $\prod_{i=1}^n \psi_\delta(y_i)dy$ we get

$$\begin{aligned}
\log D_n &\leq \int \left(K_n(y) + \sum_{i=1}^n Q(y_i) + \sum_{i=1}^n \log \psi_\delta(y_i) \right) \prod_{i=1}^n \psi_\delta(y_i) dy \\
&= n(n-1) I^Q(\psi_\delta) + n \int_{\mathbb{R}^d} Q(x)dx + n \int_{\mathbb{R}^d} (\log \psi_\delta(x)) \psi_\delta(x) dx \\
&\leq n^2 I^Q(\psi_\delta) + O(n) \\
&\leq n^2 (E^Q + \frac{\varepsilon}{2}) + O(n).
\end{aligned}$$

Therefore there exists n_1 so that if $n \geq n_1$, then

$$\frac{1}{n^2} \log D_n \ \leq \ E^Q + \varepsilon.$$

So for $n \geq n_1$, and recalling that $\varepsilon < \eta/2$, we have

$$\begin{aligned}
\int_{(\mathbb{R}^d)^n \setminus A_{n,\eta+a}} 1 \, dy &= D_n \int_{\{K_n(y) \geq n^2(E^Q+\eta+a)\}} e^{-K_n(y) - \sum Q(y_i)} dy \\
&\leq \int_{(\mathbb{R}^d)^n} e^{-\sum Q(y_i)} e^{n^2(E^Q+\varepsilon)} e^{-n^2(E^Q+\eta+a)} dy \\
&\leq \left[\left(\int_{\mathbb{R}^d} e^{-Q(x)} dx \right)^n e^{-\eta n^2/2} \right] e^{-an^2}.
\end{aligned}$$

Now choose $n_0 \geq n_1$ so that the term in brackets is bounded by 1, and the lemma is proved. □

With Lemma 4.9 in hand, the proof of Johansson's theorem [30, Thm. 2.1] generalizes to $\mathbb{R}^d$. In particular, the proof of [13, Thm. 6.96] applies with only notational changes.

Proof of Theorem 4.8. First, we will show that

$$\limsup_{n\to\infty} \mathcal{E}_n(\phi) \ \leq \ \int_{\mathbb{R}^d} \phi(x) d\mu^Q(x). \tag{4.14}$$

Fix $\eta > 0$. Since ϕ is bounded, it follows from Lemma 4.9 that the left-hand side of (4.14) is the same as

$$\limsup_{n\to\infty} \int_{A_{n,2\eta}} \frac{1}{n} \sum_{i=1}^n \phi(y_i) \mathcal{P}_n(y) dy.$$

Let $y* = (y_1^*, \ldots, y_n^*)$ be a point where the function $\frac{1}{n}\sum_{i=1}^n \phi(y_i)$ attains its maximum on the compact set $A_{n,2\eta}$. Let ν_n be the probability measure on $\mathbb{R}^d$ that has mass $\frac{1}{n}$ at each point y_i^*. We have

$$\int_{A_{n,2\eta}} \frac{1}{n} \sum_{i=1}^n \phi(y_i) \mathcal{P}_n(y) dy \ \leq \ \int_{\mathbb{R}^d} \phi(t) d\nu_n(t).$$

If $s, t \in \mathbb{R}^d$ then $|s-t| \le \sqrt{(1+|s|^2)(1+|t|^2)}$, so

$$\begin{aligned}(n-1)\sum_{i=1}^{n} Q(y_i^*) - (n-1)\sum_{i=1}^{n}\sum \log(1+|y_i^*|^2) &\le K_n(y^*) \\ &\le n^2(E^Q + 2\eta).\end{aligned}$$

Therefore

$$\int_{\mathbb{R}^d} Q(t) - \log(1+|t|^2)\; d\nu_n(t) \;\le\; \frac{n}{n-1}(E^Q + 2\eta). \tag{4.15}$$

It foilows from (4.15) that the sequence ν_n is tight. (Recall that a sequence of probability measures ν_n is called tight if for every $\varepsilon > 0$ there exists a compact set whose complement has measure less than ε for all sufficiently large n). Since the sequence is tight, there is some subsequence ν_{n_j} that converges weakly to some probability measure ν^η for which

$$\int \phi d\nu^\eta \;=\; \limsup_{n\to\infty} \int \phi d\nu_n.$$

(First extract a subsequence whose limit is the limsup, then extract a further subsequence that converges weakly).

Observe that for any $L > 0$ we have

$$\begin{aligned}\frac{1}{n^2}K_n(y^*) &= \frac{1}{n^2}\sum_{i\neq j} k(y_i^*, y_j^*) \\ &\ge \frac{1}{n^2}\left[\sum_{i,j=1}^{n} \min(L, k(y_i^*, y_j^*) - nL\right] \\ &= \int\int \min(L, k(s,t)) d\nu_n(s) d\nu_n(t) - \frac{L}{n}.\end{aligned}$$

Since $y^* \in A_{n,2\eta}$, we get

$$\int\int \min(L, k(s,t)) d\nu_n(s) d\nu_n(t) \;\le\; \frac{L}{n} + E^Q + 2\eta. \tag{4.16}$$

Apply (4.16) with $n = n_j$, let $j \to \infty$ and then let $L \to 0$. This gives

$$\int\int k(s,t) d\nu^\eta(s) d\nu^\eta(t) \;\le\; E^Q + 2\eta. \tag{4.17}$$

Now let η_j be a sequence that decreases to 0. From (4.15) we see that ν^{η_j} is tight, so it has a subsequence that converges weakly to some measure ν. From (4.17), we see that ν must be the equilibrium measure μ^Q. So we have

$$\limsup_{n\to\infty} \mathcal{E}_n(\phi) \;\le\; \int \phi \; d\nu^{\eta_j}$$

for all j, and letting $j \to \infty$ we get (4.14).

Repeating the argument with y_* where $\frac{1}{n}\sum \phi(y_i)$ attains its minimum, we get the reverse inequality for the liminf, and hence we get (4.11).

4.4 Gaussian Case

Let $Q(x) = \gamma|x|^2$. (The positive constant γ is just a simple scaling, and plays no important rôle. But as some authors default to $\gamma = \frac{1}{2}$ and others to $\gamma = 1$, we include it to make the formulas seem more familiar.) Let X_n be chosen in $\mathfrak{C}_n^{d,H}$ with distribution $C_n e^{-\mathrm{tr}(Q(X))}$, with C_n the normalizing constant. Let E_n denote expectation with respect to this probability. Then for ϕ a continuous bounded function on $\mathbb{R}^d$, we have

$$\begin{aligned} E_n\left[\frac{1}{n}\,\mathrm{tr}\,(\phi(\frac{1}{\sqrt{n}}\,X_n))\right] &= \int_{\mathfrak{C}_n^{d,H}} \frac{1}{n}\,\mathrm{tr}\,(\phi(\frac{1}{\sqrt{n}}X))w_n(X)dX \\ &= \frac{1}{n}\int_{(\mathbb{R}^d)^n} \sum_{i=1}^n \phi(\frac{1}{\sqrt{n}}\lambda_i)\rho_n(\lambda)d\lambda \\ &= \frac{1}{n}\int_{(\mathbb{R}^d)^n} \sum_{i=1}^n \phi(y_i)\mathcal{P}_n(y)dy \\ &= \int_{\mathbb{R}^d} \phi(x)R_{1,n}(x)dx. \end{aligned}$$

So combining Theorem 4.8 with Theorem 3.1, we get the following result in the Gaussian case.

Theorem 4.10. *Let $Q(x) = \gamma|x|^2$. Let X_n be chosen in $\mathfrak{C}_n^{d,H}$ with distribution $C_n e^{-\mathrm{tr}(Q(X))}$. Then for all ϕ that are continuous and bounded on $\mathbb{R}^d$, we have*

$$\lim_{n\to\infty} E_n\left[\frac{1}{n}\,\mathrm{tr}\,(\phi(\frac{1}{\sqrt{n}}\,X_n))\right] = \int_{\mathbb{R}^d} \phi(x)d\mu^Q(x),$$

where the equilibrium measure μ^Q is given by Theorem 3.1.

Let us consider what happens when we project the equilibrium measures for the Gaussian case onto the x_1 axis. For $d = 2$, the density is

$$\frac{1}{\pi R_2^2}\int_{-\sqrt{R_2^2-x_1^2}}^{\sqrt{R_2^2-x_1^2}} 1dx_2 = \frac{2}{\pi R_2^2}\sqrt{R_2^2 - x_1^2}.$$

For $d = 3$, it is the weighted area of a disk of radius $\sqrt{R_3^2 - x_1^2}$, where the weight at radius s is $\frac{1}{\pi^2 R_3^2}\frac{1}{\sqrt{R_3^2-x_1^2-s^2}}$. Switching to polar coordinates on this disk, the density is

$$\frac{2}{\pi R_3^2}\int_0^{\sqrt{R_3^2-x_1^2}} \frac{s}{\sqrt{R_3^2 - x_1^2 - s^2}}ds = \frac{2}{\pi R_3^2}\sqrt{R_3^2 - x_1^2}.$$

For $d \geq 4$, we project normalized area measure for the $(d-1)$ sphere S^{d-1} with radius R_d. Let us change variables to $x_1 = R_d\cos\theta$ for $0 \leq \theta \leq \pi$, where θ represents the angle with the positive x_1 axis. The area of the shell between θ and $\theta + \Delta\theta$ is approximately $\Delta\theta$ times the $(d-2)$-dimensional area of a sphere of radius $\sin\theta$, which is $cR_d^{d-2}\sin^{d-2}\theta$. As

$\Delta x = -R_d \sin\theta \Delta\theta$, we get that the area of the shell between x and $x + \Delta x$ is approximately Δx times

$$cR_d^{d-3} \sin^{d-3}\theta \;=\; c\left(\sqrt{R_d^2 - x_1^2}\right)^{d-3}.$$

Letting $\Delta x \to 0$, and plugging in the normalizing value for c which comes from the identity

$$\int_{-1}^{1} (\sqrt{1-x^2})^{d-3} dx \;=\; \frac{\sqrt{\pi}\;\Gamma(\frac{d-1}{2})}{\Gamma(\frac{d}{2})},$$

which is valid for $d > 1$, we get that the density is

$$\frac{\Gamma(\frac{d}{2})}{\sqrt{\pi} R_d^{d-2}\;\Gamma(\frac{d-1}{2})}\;\left(\sqrt{R_d^2 - x_1^2}\right)^{d-3}.$$

For $d = 2, 3, 4$ we still get a semi-circular distribution, though the radius R_d shrinks as d increases from 1 to 4. For $d > 4$ the radius R_d stays $\frac{1}{\sqrt{2\gamma}}$, but the eigenvalues of X^1 cluster more around 0. We have proved the following.

Corollary 4.11. *Let X_n be chosen in $\mathfrak{C}_n^{d,H}$ with distribution $C_n e^{-\gamma \operatorname{tr}(\sum_{j=1}^d (X^j)^2)}$. Then for all ϕ that are continuous and bounded on $\mathbb{R}$, we have*

$$\lim_{n\to\infty} E_n\left[\frac{1}{n}\;\operatorname{tr}\left(\phi(\frac{1}{\sqrt{n}}\;X_n^1)\right)\right] \;=\; \int_{-R_d}^{R_d} \phi(x) f_d(x) dx,$$

where R_d are as in Theorem 3.1 and

$$\begin{aligned}
f_1(x) &= \frac{2}{\pi R_1^2}\sqrt{R_1^2 - x^2} \\
f_2(x) &= \frac{2}{\pi R_2^2}\sqrt{R_2^2 - x^2} \\
f_3(x) &= \frac{2}{\pi R_3^2}\sqrt{R_3^2 - x^2} \\
f_d(x) &= \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi} R_d^{d-2}\;\Gamma(\frac{d-1}{2})}\;\left(\sqrt{R_d^2 - x^2}\right)^{d-3}, \qquad d \geq 4.
\end{aligned}$$

5 Random Commuting Matrices: the non-Hermitian case

5.1 Generic d-tuples

In this section, we shall drop the assumption that the matrices are Hermitian, and work with random matrices from $\mathfrak{C}_n^d$.

There is a qualitative difference between the algebraic sets $\mathfrak{C}_n^d$ for different values of d, n. In particular, there are major differences between the reducible and irreducible cases. Motzkin and Taussky proved that $\mathfrak{C}_n^2$ is irreducible and of dimension $n^2 + n$ for all n [35].

Gerstenhaber reproved this, and proved that $\mathfrak{C}_4^4$ is reducible [20]. Guralnick proved that $\mathfrak{C}_n^d$ is irreducible for all d if $n \leq 3$ [23].

When $d = 3$, Guralnick proved that $\mathfrak{C}_{32}^3$ is reducible [23]. It is now known that $\mathfrak{C}_n^3$ is reducible if $n \geq 29$ [28, 36], and irreducible if $n \leq 10$ (see [24, 28, 37, 25, 40, 41] for n going from 4 to 10.)

Let us call $X \in \mathfrak{C}_n^d$ *generically diagonalizable* if each component X^r has n distinct eigenvalues, and define

$$\mathfrak{C}_{n,\text{gen}}^d \;=\; \{X \in \mathfrak{C}_n^d : \text{ each } X^r \text{ has } n \text{ distinct eigenvalues}\}.$$

When are the generically diagonalizable d-tuples dense in $\mathfrak{C}_n^d$? In [28], Holbrook and Omladic showed that the answer is yes if and only if $\mathfrak{C}_n^d$ is irreducible. In the irreducible case, the non-diagonalizable elements are of smaller dimension than the diagonalizable ones. Guralnick proved that $\dim \mathfrak{C}_n^3 > \dim \mathfrak{C}_{n,\text{gen}}^3$ if $n \geq 32$. For a study of the components in the non-irreducible case, see [29].

Proposition 5.1. *Let $\mathfrak{C}_n^d$ be irreducible. Then*

$$\dim_{\mathbb{C}}(\mathfrak{C}_n^d \setminus \mathfrak{C}_{n,\text{gen}}^d) \;<\; \dim_{\mathbb{C}}(\mathfrak{C}_{n,\text{gen}}^d) = n^2 + (d-1)n. \tag{5.1}$$

Proof. Let us write $\mathcal{W}^r = \{X \in \mathfrak{C}_n^d : X^r \text{ does not have } n \text{ distinct eigenvalues}\}$. Let q^r be the discriminant of the characteristic polynomial of X^r, thought of as a polynomial in the dn^2 entries of an element of $\mathbb{M}_n^d$, and let $q = q^1 \cdots q^d$. Then

$$\mathcal{W}^r \;=\; \{X \in \mathfrak{C}_n^d : q^r(X) = 0\},$$

and

$$\mathfrak{C}_n^d \setminus \mathfrak{C}_{n,\text{gen}}^d \;=\; \cup_{r=1}^d \mathcal{W}^r \;=\; \{X \in \mathfrak{C}_n^d : q(X) = 0\}.$$

As q does not vanish identically on $\mathfrak{C}_n^d$, assumed irreducible, it follows that

$$\dim_{\mathbb{C}}(\mathfrak{C}_n^d \setminus \mathfrak{C}_{n,\text{gen}}^d) \;<\; \dim_{\mathbb{C}}(\mathfrak{C}_n^d).$$

When $\mathfrak{C}_n^d$ is irreducible, it is shown in [24] that $\dim_{\mathbb{C}}(\mathfrak{C}_n^d) = n^2 + (d-1)n$. It is obvious that $\dim_{\mathbb{C}}(\mathfrak{C}_{n,\text{gen}}^d) = n^2 + (d-1)n$, so we get (5.1). □

5.2 Ginibre formula

Throughout the remainder of Section 5 we shall assume that $\mathfrak{C}_n^d$ is irreducible.

Assumption I: $\mathfrak{C}_n^d$ is irreducible.

(If this is not the case, then we must replace $\mathfrak{C}_n^d$ by its principal component, that is the closure of the set of jointly diagonalizable matrices). Let $w : \mathfrak{C}_n^d \to [0, \infty)$ be a continuous function that satisfies $w(U^*XU) = w(X)$ for all unitary $U \in \mathcal{U}(n)$, and is normalized so

$$\int_{\mathfrak{C}_n^d} w(X)dX \;=\; 1,$$

where dX is Hausdorff measure of the appropriate dimension restricted to $\mathfrak{C}_n^d$ (from (5.1) the appropriate dimension is $2n^2 + 2(d-1)n$).

Let $X \in \mathfrak{C}_n^d$. There is a unitary U in $\mathcal{U}(n)$ so that the d-tuple U^*XU is jointly upper triangular. Once such a U has been chosen, let $\lambda_j \in \mathbb{C}^d$ be the point whose r^{th} component is the j^{th} diagonal entry of U^*X^rU. We shall call the multi-set $\{\lambda_1, \ldots, \lambda_n\}$ the *multi-spectrum* of X. It is straightforward to see that it is independent of the particular choice of U, that it consists of the eigenvalues of X, and that the multiplicity of a particular eigenvalue λ is

$$\dim_{\mathbb{C}} \left(\cap_{r=1}^d \ker(X^r - \lambda^r)^n \right).$$

By Proposition 5.1, on a set of full measure each eigenvalue will have multiplicity one, so we shall just talk about the spectrum.

Let $\mathcal{T}_n^d$ denote the commuting d-tuples of n-by-n matrices that are upper triangular with respect to a fixed basis. Any generic element Q of $\mathcal{T}_n^d$ can be written uniquely as ADA^{-1}, where D is a diagonal d-tuple and A is an upper triangular n-by-n matrix with 1's on the diagonal. Let us write $F_{n,1}$ to denote the set of upper triangular n-by-n matrices with 1's on the diagonal, and Ad_A the linear operator on $\mathbb{M}_n$ of conjugation by A; so $\mathrm{Ad}_A(X) = AXA^{-1}$.

Lemma 5.2. *Let Q be in $\mathcal{T}_n^d \cap \mathfrak{C}_{n,\text{gen}}^d$. Then the tangent space to $\mathfrak{C}_n^d$ at Q is*

$$\{Z \in \mathbb{M}_n^d : [Q^r, Z^s] = [Q^s, Z^r], \quad \forall\ 1 \le r \ne s \le d\}. \tag{5.2}$$

The tangent space to $\mathcal{T}_n^d$ at Q is $\{Z \in \mathcal{T}_n^d : (5.2) \text{ holds}\}$. The real dimension of the first tangent space is $2n^2 + 2(d-1)n$, and of the second is $n^2 + (2d-1)n$.

Proof. ($\Rightarrow$) Suppose $\gamma(t)$ is a C^1 curve in $\mathfrak{C}_n^d$ with $\gamma(0) = Q$ and $\gamma'(0) = Z$. Then $\gamma^r(t) = Q^r + tZ^r + o(t)$. As

$$0 \;=\; \gamma^r(t)\gamma^s(t) - \gamma^s(t)\gamma^r(t) \;=\; t\left([Q^r, Z^s] - [Q^s, Z^r]\right) + o(t),$$

we conclude that $[Q^r, Z^s] = [Q^s, Z^r]$.

($\Leftarrow$) Suppose Z is in (5.2). Let A be an upper triangular matrix that diagonalizes Q: $Q = ADA^{-1}$. Let $Y = A^{-1}ZA$. Then

$$[D^r, Y^s] \;=\; [D^s, Y^r] \quad \forall\ 1 \le r \ne s \le d. \tag{5.3}$$

Let λ_j^r be the diagonal entries of D^r. Then (5.3) says

$$Y_{ij}^s(\lambda_j^r - \lambda_i^r) \;=\; Y_{ij}^r(\lambda_j^s - \lambda_i^s) \quad \forall\ 1 \le r \ne s \le d. \tag{5.4}$$

Define a matrix B by

$$B_{ij} \;=\; \begin{cases} \frac{1}{\lambda_j^r - \lambda_i^r} Y_{ij}^r, & i \ne j \\ 0 & i = j \end{cases}.$$

By (5.4) the definition does not depend on r, and since Q is generically diagonalizable we never divide by zero. Let E^r be the diagonal of Y^r. Define a curve $\gamma(t)$ by

$$\gamma^r(t) \;=\; Ae^{tB}(D^r + tE^r)e^{-tB}A^{-1}.$$

Then $\gamma(t) \in \mathfrak{C}_n^d$, and if Z (and hence Y) is in $\mathcal{T}_n^d$, then so is $\gamma(t)$ for all t. Moreover, $\gamma(0) = Q$ and

$$\gamma'(0) \;=\; Ad_A([B, D] + E) \;=\; Z.$$

So the tangent space is as claimed.

To calculate the dimensions of the tangent spaces, note that Ad_A is an isomorphism from the tangent space at D to the tangent space at Q, so it suffices to calculate the dimension of the tangent space at D. In $\mathfrak{C}_n^d$, one can choose the diagonal entries freely, which gives $2dn$ dimensions, and the off diagonal entries for $r = 1$, which by (5.4) then determines all the other off-diagonal entries, for an additional $2(n^2 - n)$ dimensions. In the upper-triangular case, one can again choose the diagonal entries freely, and one has $n^2 - n$ off diagonal dimensions. □

Let us fix $Q_0 \in \mathcal{T}_n^d \cap \mathfrak{C}_{n,\mathrm{gen}}^d$. We shall use T to mean tangent space. Let ε be small and positive, and consider $\{X : \|X - Q_0\|_F < \varepsilon\}$. For ε sufficiently small, the eigenvalues of X will be so close to the eigenvalues of Q_0 that it makes sense to order the eigenvalues of X so that the j^{th} eigenvalue is closest to the j^{th} eigenvalue of Q_0. Let $X = UQU^*$ be an upper triangular factoring of X that preserves this order. The unitary U is unique up to multiplication by a diagonal unitary.

As in Theorem 4.5 we can parametrize a neighborhood of Q_0 by

$$\begin{aligned} g : \mathcal{T}_n^d \times \exp(\mathcal{A}_{0,n}) &\to \mathfrak{C}_n^d \\ (Q, U) &\mapsto UQU^* \;=\; \mathrm{Ad}_U(Q). \end{aligned}$$

We want to decompose

$$dX \;=\; \kappa(Q) dQ d\nu$$

where ν is the probability measure on the homogeneous space $\mathcal{U}(n)/\mathbb{T}^n$ obtained by taking the quotient of Haar measure on $\mathcal{U}(n)$, and dQ is Hausdorff measure of the appropriate dimension (which is $n^2 + (2d-1)n$) on $\mathcal{T}_n^d$. For $d = 1$ see the treatment by Tao in [42, Sec. 2.6.2].

Decompose $T_{Q_0}\mathfrak{C}_n^d = T_{Q_0}\mathcal{T}_n^d \oplus (T_{Q_0}\mathcal{T}_n^d)^\perp$. Let P be orthogonal projection from $T_{Q_0}\mathfrak{C}_n^d$ onto $T_{Q_0}\mathcal{T}_n^d$. Let $\mathrm{ad}_{Q_0}(S)$ be the d-tuple $([Q_0^1, S], \ldots, [Q_0^d, S])$. Then we can write dg in block form

$$dg_{(Q_0,1)} \;=\; \begin{array}{c} \\ T_{Q_0}\mathcal{T}_n^d \\ (T_{Q_0}\mathcal{T}_n^d)^\perp \end{array} \begin{array}{c} \begin{array}{cc} T_{Q_0}\mathcal{T}_n^d & \mathcal{A}_{0,n} \end{array} \\ \begin{pmatrix} \mathrm{id} & -P\ \mathrm{ad}_{Q_0} \\ 0 & -P^\perp\ \mathrm{ad}_{Q_0} \end{pmatrix} \end{array} \tag{5.5}$$

The Jacobian of g is the square root of the absolute value of the determinant of the real linear map $(dg)^t dg$. From the structure of (5.5), this depends only on the $(2,2)$ entry. Define a real linear map $\Lambda_Q : \mathcal{A}_{0,n} \to \mathcal{A}_{0,n}$ by

$$\langle \Lambda_Q S_1, S_2 \rangle \;=\; \langle P^\perp \mathrm{ad}_Q(S_1), \mathrm{ad}_Q(S_2) \rangle.$$

Then the Jacobian of g at $(Q_0, 1)$ is $\sqrt{|\det(\Lambda_{Q_0})|}$. Summarizing we have proved the following.

Proposition 5.3. *Assume $\mathfrak{C}_n^d$ is irreducible. With $X = UQU^*$ as above, we have*

$$w(X) dX \;=\; \kappa(Q) w(Q) dQ d\nu,$$

where $\kappa(Q) = \beta_n \sqrt{|\det(\Lambda_Q)|}$.

5.3 Calculating κ

We can calculate κ explicitly in the following special cases: $d = 1$ (which was done by Ginibre [21]); when Q is diagonal; and when $n = 2$.

Case: the d-tuple Q equals D, a diagonal d-tuple. Let λ_i denote the eigenvalues of D. Now the tangent space $T_D\mathfrak{C}_n^d$ decomposes into the direct sum of $T_D\mathcal{T}_n^d$ and the strictly lower triangular d-tuples Z that satisfy (5.2). We have

$$\mathrm{ad}_D(E_{ji}) \;=\; (\lambda_j^r - \lambda_i^r)E_{ji}.$$

With R_{ij}, S_{ij} as in the proof of Theorem 4.5, we have that for $i < j$,

$$\begin{aligned} -P^\perp \mathrm{ad}_D : R_{ij} &\mapsto -\frac{1}{\sqrt{2}}(\lambda_j^r - \lambda_i^r)E_{ji} \\ S_{ij} &\mapsto \frac{i}{\sqrt{2}}(\lambda_j^r - \lambda_i^r)E_{ji}. \end{aligned}$$

Therefore the Jacobian of g is

$$\kappa(D) \;=\; c_n \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^2,$$

where $c_n = 2^{-n(n-1)/2}\beta_n$.

Case: $d = 1$. When $d = 1$ the conditions in (5.2) are vacuous, and $P^\perp$ is projection onto the strictly lower triangular matrices. As shown in [42, p. 190], if one orders the pairs (j, i) with $j > i$ in the order $(n, 1), (n, 2), (n-1, 1), (n, 3), (n-1, 2), (n-2, 1), \ldots$, then the $(2, 2)$ entry of (5.5) becomes triangular, with diagonal entries $Q_{ii} - Q_{jj}$. Consequently

$$\kappa(Q) \;=\; c_n \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^2, \tag{5.6}$$

where $\{\lambda_i\}$ are the eigenvalues of Q.

Case: $n = 2$. The space $T_Q\mathfrak{C}_2^d \ominus T_Q\mathcal{T}_2^d$ is one dimensional. If the eigenvalues of Q are λ_1 and λ_2, then $T_Q\mathfrak{C}_2^d \ominus T_Q\mathcal{T}_2^d$ is of complex dimension one, and is spanned by the d-tuple

$$\omega \;=\; \frac{1}{|\lambda_2 - \lambda_1|}\left[\begin{pmatrix} 0 & 0 \\ (\lambda_2^r - \lambda_1^r) & 0 \end{pmatrix}\right].$$

Indeed,

$$\begin{aligned} [Q^s, \omega^r] &= \frac{1}{|\lambda_2 - \lambda_1|}\begin{pmatrix} 0 & 0 \\ (\lambda_2^r - \lambda_1^r)(\lambda_2^s - \lambda_1^s) & 0 \end{pmatrix} \\ &= [Q^r, \omega^s], \end{aligned}$$

so by Lemma 5.2 ω is in $T_Q\mathfrak{C}_2^d$. Moreover it is clearly orthogonal to $T_Q\mathcal{T}_2^d$.

As

$$\mathrm{ad}_{Q_0}(E_{21}) \;=\; \left[\begin{pmatrix} 0 & 0 \\ (\lambda_2^r - \lambda_1^r) & 0 \end{pmatrix}\right],$$

we get that

$$\begin{aligned} -P^{\perp}\mathrm{ad}_{Q_0} : R_{12} &\mapsto \frac{1}{\sqrt{2}}|\lambda_2-\lambda_1|\omega \\ S_{12} &\mapsto \frac{-i}{\sqrt{2}}|\lambda_2-\lambda_1|\omega. \end{aligned}$$

Therefore

$$\kappa(Q) \;=\; c_2|\lambda_2-\lambda_1|^2. \tag{5.7}$$

5.4 From upper triangular to diagonal

We are interested in determining the distribution $\rho_n(\lambda_1,\ldots,\lambda_n)$ of the eigenvalues. We shall mainly consider Gausssian weights on $\mathfrak{C}_n^d$, that is weights of the form

$$w(X) \;=\; Ce^{-\gamma \mathrm{tr}\sum_{r=1}^d (X^r)^* X^r} \;=\; Ce^{-\gamma\|X\|_F^2}. \tag{5.8}$$

The important property of Gaussian weights that we shall use, in addition to their unitary invariance, is that for upper triangular d-tuples Q they factor as

$$w(Q) \;=\; w_1(\|D\|_F^2)w_2(\|Q-D\|_F^2), \tag{5.9}$$

where D is the diagonal of Q.

When $d=1$, ρ can be derived from κ in a straightforward way. Indeed, the off-diagonal terms can be chosen independently, and since the weight factors as in (5.9), one can integrate out all the off-diagonal terms, and since κ from (5.6) depends only on the eigenvalues, we get that the density of the eigenvalues for $d=1$ is

$$\rho_n(\lambda) \;=\; C\prod_{1\le i<j\le n}|\lambda_i-\lambda_j|^2 e^{-\gamma\sum_{j=1}^n|\lambda_j|^2}.$$

See [21] or [42, 2.6.2].

For $d>1$, we proceed as follows. Let Diag_n^d denote the d-tuples of diagonal n-by-n matrices. To make reading easier, we shall drop the subscript 0 and fix an element $Q\in\mathcal{T}_n^d$. We have $Q=\mathrm{Ad}_A(D)$ for some $A\in F_{n,1}$ and D the diagonal with the same eigenvalues as Q (in the same order), which we write $\lambda_1,\ldots,\lambda_n$. We can parametrize a neighborhood of Q in $\mathcal{T}_n^d$ by

$$\begin{aligned} G : \mathrm{Diag}_n^d \times F_{n,1} &\to \mathcal{T}_n^d \\ (\widetilde{D},\widetilde{A}) &\mapsto \widetilde{A}\widetilde{D}\widetilde{A}^{-1} \;=\; \mathrm{Ad}_{\widetilde{A}}(\widetilde{D}). \end{aligned}$$

Let us calculate dG at (D,A). The tangent space of $F_{n,1}$ at A is the space of strictly upper triangular matrices, which we shall denote by $F_{n,0}$. We shall work in complex coordinates, so the Jacobian will be $\det(dG)^*dG$ without a square root. We have

$$\begin{aligned} dG_{(D,A)} : (\mathbb{C}^d)^n\times F_{n,0} &\to T_Q\mathcal{T}_n^d \\ dG_{(D,A)}(E,B) &= \lim_{t\to 0}\frac{1}{t}[(A+tB)(D+tE)(A+tB)^{-1}-Q] \\ &= \mathrm{Ad}_A([A^{-1}B,D])+\mathrm{Ad}_A(E). \end{aligned} \tag{5.10}$$

As a basis for $(\mathbb{C}^d)^n$ we can choose $\{\xi_p^r : 1 \le p \le n, 1 \le r \le d\}$, where ξ_p^r is the d-tuple that is E_{pp} in the r^{th} slot and 0 elsewhere. As a basis for $F_{n,0}$ we can choose $\{E_{ij} : 1 \le i < j \le n\}$.

We have $dG(\xi_p^r, 0)$ is the d-tuple which in the r^{th} slot is the upper triangular d-tuple that has (k,l) entry $A_{kp}A_{pl}^{-1}$ in the r^{th} slot and 0 elsewhere. We have $dG(0, E_{ij})$ is the d-tuple whose (k,l) entry in the r^{th} slot is

$$\delta_{ik}\lambda_j^r A_{jl}^{-1} - \sum_q A_{kq}\lambda_q^r A_{qi}^{-1}A_{jl}^{-1}.$$

Define Γ_Q to be the complex linear endomorphism of $(\mathbb{C}^d)^n \times F_{n,0}$ given by

$$\Gamma_Q \;=\; dG^*_{(D,A)}dG_{(D,A)}. \tag{5.11}$$

We have

$$\begin{aligned}
&\Gamma_Q((\xi_{p_1}^{r_1}, E_{i_1j_1}), (\xi_{p_2}^{r_2}, E_{i_2j_2})) \\
&= \sum_{k,l=1}^n \left[\delta_{r_1,r_2} A_{kp_1}A_{p_1l}^{-1}\overline{A_{kp_2}}\,\overline{A_{p_2l}^{-1}} \;+\; \sum_{q_2=1}^n A_{kp_1}A_{p_1l}^{-1}\overline{A_{kq_2}}\,\overline{\lambda_{q_2}^{r_1}}\,\overline{A_{q_2i_2}^{-1}}A_{j_2l}^{-1}\right. \\
&+ \left. \sum_{q_1=1}^n A_{kq_1}\lambda_{q_1}^{r_2}A_{q_1i_1}^{-1}A_{j_1l}^{-1}\overline{A_{kp_2}}\,\overline{A_{p_2l}^{-1}} \;+\; \sum_{s=1}^d\sum_{q_1,q_2=1}^n A_{kq_1}\lambda_{q_1}^s A_{q_1i_1}^{-1}A_{j_1l}^{-1}\overline{A_{kq_2}}\,\overline{\lambda_{q_2}^s}\,\overline{A_{q_2i_2}^{-1}}A_{j_2l}^{-1}\right].
\end{aligned}$$

Then the Jacobian of G is $|\det(\Gamma_Q)|$. Thus we have proved the following (recall that the determinant of Λ is calculated for a real linear map, and for Γ it is complex linear, which explains why the first has a square root and the second does not).

Theorem 5.4. *Assume $\mathfrak{C}_n^d$ is irreducible. With $X = UADA^{-1}U^*$ as above, we have*

$$w(X)dX \;=\; \beta_n w_1(\|D\|_F^2)w_2(\|Ad_A(D) - D\|_F^2)\sqrt{|\det(\Lambda_{Ad_A(D)})|}\;|\det(\Gamma_{Ad_A(D)})|d\lambda dAd\nu. \tag{5.12}$$

The eigenvalue distribution of X is given by

$$\rho(\lambda) \;=\; \beta_n w_1(\|D\|_F^2)\int_{F_{n,1}} w_2(\|Ad_A(D) - D\|_F^2)\sqrt{|\det(\Lambda_{Ad_A(D)})|}\;|\det(\Gamma_{Ad_A(D)})|dA. \tag{5.13}$$

5.5 Density of eigenvalues

We would like to find an explicit formula for the density of the eigenvalues, which would come from evaluating the right-hand side of (5.13). We shall do this first in the known case $d = 1$, as an illustration. Then we shall do it in the case of d arbitrary, but with $n = 2$.

Case: $d = 1$. We have from (5.10)

$$\begin{aligned}
dG_{(D,A)} : \mathbb{C}^d \times F_{n,0} &\to T_Q\mathcal{T}_n^1 \\
(E, B) &\mapsto Ad_A([A^{-1}B, D] + E).
\end{aligned}$$

Both the domain and codomain can be identified with the upper triangular matrices $\mathcal{T}_n^1$. As A is upper triangular with 1's on the diagonal, multiplication by A and A^{-1} are unipotent

linear operators on $\mathcal{T}_n^1$, and hence have determinant 1. So the Jacobian of G is the same as the Jacobian of the map

$$(E, B) \ \mapsto \ [A^{-1}B, D] + E.$$

Decomposing $\mathcal{T}_n^1 = \mathbb{C}^d \oplus F_{n,0}$, this has the same determinant as

$$B \ \mapsto \ [A^{-1}B, D].$$

As multiplication by A^{-1} is unipotent on $F_{n,0}$, this in turn has the same Jacobian as

$$B \ \mapsto \ [B, D],$$

which is

$$\Delta(\lambda) \ := \ \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^2.$$

So if we take (5.12) and write it as

$$w(X)dX \ = \ w_1(\|D\|_F^2) w_2(\|ADA^{-1} - D\|_F^2) \kappa(ADA^{-1}) \Delta(\lambda) d\lambda dA d\nu,$$

integrate with respect to ν and A and use (5.6), we get

$$\rho(\lambda) \ = \ c_n w_1(\|D\|_F^2) \Delta(\lambda) \int_{F_{n,1}} w_2(\|ADA^{-1} - D\|_F^2) \Delta(\lambda) dA. \tag{5.14}$$

(We can pull $\kappa(ADA^{-1}) = c_n \Delta(\lambda)$ out of the integral since it does not depend on A.) As

$$ADA^{-1} - D \ = \ [A, D]A^{-1} \ = \ ad_A(D)A^{-1},$$

we can make the change of variables

$$\begin{aligned} B &= [A, D]A^{-1} \\ dB &= \Delta(\lambda) dA \end{aligned}$$

to conclude that the integral in (5.14) equals $\int_{F_{n,0}} w_2(\|B\|_F^2) dB$, which is some constant. So we get Ginibre's density formula [21] for some constant C':

$$\rho(\lambda) \ = \ C' \Delta(\lambda) w_1(\|D\|_F^2),$$

which if we assume w is Gaussian as in (5.8) becomes

$$\rho(\lambda) \ = \ C' \Delta(\lambda) e^{-\gamma \sum_{j=1}^n |\lambda_j|^2}.$$

Case: $n = 2$, d arbitrary. We shall restrict to the weight w being as in (5.8). We can write

$$A \ = \ \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}.$$

From (5.11), the matrix for Γ_Q has the form

$$\begin{array}{c} \\ \xi_1^s \\ \xi_2^s \\ E_{12} \end{array} \begin{array}{c} \begin{array}{ccc} \xi_1^r & \xi_2^r & E_{12} \end{array} \\ \begin{pmatrix} \delta_{rs}(1+|\alpha|^2) & -\delta_{rs}|\alpha|^2 & -\overline{\alpha}(\lambda_2^s-\lambda_1^s) \\ -\delta_{rs}|\alpha|^2 & \delta_{rs}(1+|\alpha|^2) & \overline{\alpha}(\lambda_2^s-\lambda_1^s) \\ -\alpha(\bar{\lambda}_2^r-\bar{\lambda}_1^r) & \alpha(\bar{\lambda}_2^r-\bar{\lambda}_1^r) & |\lambda_2-\lambda_1|^2 \end{pmatrix} \end{array} \tag{5.15}$$

To calculate the determinant of (5.15) we shall use the Schur complement. The first $2d$-by-$2d$ part of (5.15) has determinant $(1+2|\alpha|^2)^d$. The Schur complement with respect to the bottom right-hand entry is

$$|\lambda_2-\lambda_1|^2 - \sum_{r=1}^{d} \langle \begin{pmatrix} 1+|\alpha|^2 & -|\alpha|^2 \\ -|\alpha|^2 & 1+|\alpha|^2 \end{pmatrix}^{-1} \begin{pmatrix} -\overline{\alpha}(\lambda_2^r-\lambda_1^r) \\ \overline{\alpha}(\lambda_2^r-\lambda_1^r) \end{pmatrix}, \begin{pmatrix} -\overline{\alpha}(\lambda_2^r-\lambda_1^r) \\ \overline{\alpha}(\lambda_2^r-\lambda_1^r) \end{pmatrix} \rangle \;=\; \frac{|\lambda_2-\lambda_1|^2}{1+2|\alpha|^2}.$$

Multiplying these together, we get

$$\det(\Gamma_Q) \;=\; (1+2|\alpha|^2)^{d-1}|\lambda_2-\lambda_1|^2.$$

Using this and (5.7), Theorem 5.4 gives

$$w(X)dX \;=\; Cc_2\; e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} e^{-\gamma|\alpha|^2|\lambda_2-\lambda_1|^2} |\lambda_2-\lambda_1|^4 (1+2|\alpha|^2)^{d-1} d\lambda d\alpha d\nu.$$

Integrate first with respect to ν,then use polar coordinates $\alpha = re^{i\theta}$, and we get

$$\rho(\lambda) \;=\; C'\; e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} |\lambda_2-\lambda_1|^4 \int_0^\infty e^{-\gamma|\lambda_2-\lambda_1|^2 r^2} (1+2r^2)^{d-1} r dr. \tag{5.16}$$

The following identity is elementary:

$$I_d \;:=\; \int_0^\infty r^{2d+1} e^{-kr^2} dr \;=\; \frac{d!}{2k^{d+1}}. \tag{5.17}$$

Using(5.17) in (5.16) with $k = \gamma|\lambda_2-\lambda_1|^2$, we have proved the following result, where we absorb some factors depending on d into C'_d.

Theorem 5.5. *Let X be a random element of $\mathfrak{C}_2^d$ chosen with distribution*

$$w(X) \;=\; Ce^{-\gamma\|X\|_F^2}.$$

Then the eigenvalues λ of X in $(\mathbb{C}^d)^2$ have distribution

$$\rho^d(\lambda) \;=\; C'_d\; e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} |\lambda_2-\lambda_1|^2 \sum_{j=0}^{d-1} \frac{1}{(d-1-j)!} \frac{2^j}{\gamma^{j+1}} \frac{1}{|\lambda_2-\lambda_1|^{2j}}. \tag{5.18}$$

For $d = 1, 2, 3$ we get respectively

$$\begin{aligned} \rho^1(\lambda) &= C'_1 e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{\gamma}|\lambda_2-\lambda_1|^2\right] \\ \rho^2(\lambda) &= C'_2 e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{\gamma}|\lambda_2-\lambda_1|^2 + \frac{2}{\gamma^2}\right] \\ \rho^3(\lambda) &= C'_3 e^{-\gamma(|\lambda_1|^2+|\lambda_2|^2)} \left[\frac{1}{2\gamma}|\lambda_2-\lambda_1|^2 + \frac{2}{\gamma^2} + \frac{4}{\gamma^3}|\lambda_2-\lambda_1|^{-2}\right]. \end{aligned}$$

Remark 5.6. The negative powers in (5.18) are integrable, since λ_j are in $\mathbb{C}^d$ and the smallest negative power is $4-2d$. However, once $d \geq 3$, the negative powers show that, unlike in the Hermitian case, there is an *attraction* between the eigenvalues, not a repulsion.

6 Notation

For the convenience of the reader, we will list our notation here.

$\mathfrak{A}_n^d = \{X \in \mathbb{M}_n^d : X^r X^s = -X^s X^r \ \forall\ r \neq s\}$
$\mathfrak{A}_n^{d,H} = \{X \in \mathfrak{A}_n^d : (X^r)^* = X^r \ \forall\ r\}$
$A_{n,\eta} = \{y \in (\mathbb{R}^d)^n : \frac{1}{n^2} K_n(y) \leq E^Q + \eta\}$
$\mathrm{Ad}_A(B) = ABA^{-1}$
$\mathcal{A}_{0,n}$ the skew-Hermitian matrices that have 0's on the diagonal
$\mathfrak{C}_n^d = \{X \in \mathbb{M}_n^d : X^r X^s = X^s X^r \ \forall\ r, s\}$
$\mathfrak{C}_n^{d,H} = \{X \in \mathfrak{C}_n^d : (X^r)^* = X^r \ \forall\ r\}$
$\mathfrak{C}_{n,\mathrm{gen}}^d$ and $\mathfrak{C}_{n,\mathrm{gen}}^{d,H}$ denote the generically diagonalizable tuples (resp. the Hermitian ones)
$\mathrm{diag}(A)$ the diagonal matrix whose entries are the diagonal entries of A
E_{ij} is the elementary matrix that has entry 1 in the (i,j) slot and 0 elsewhere
$\|X\|_F^2 = \sum_{r=1}^d \sum_{i,j=1}^n |X_{ij}^r|^2$ Frobenius norm
$F_{n,0}$ upper triangular n-by-n matrices with 0's on the diagonal
$F_{n,1}$ upper triangular n-by-n matrices with 1's on the diagonal
$\mathcal{G}(d,n)$ the complex Grassmanian of all d-dimensional linear subspaces of $\mathbb{C}^n$
$GL(n,\mathbb{C})$ is the invertible matrices
$I^Q(\mu) = \int\int \log \frac{1}{|x-y|} d\mu(x) d\mu(y) + \int Q(x) d\mu(x)$
$k(s,t) = \log \frac{1}{|s-t|} + \frac{1}{2}(Q(s) + Q(t))$
$K_n(y) = \sum_{1 \leq i < j \leq n} \log |y_i - y_j|^{-2} + (n-1) \sum_{i=1}^n Q(y_i)$
$\mathbb{M}_n$ n-by-n matrices over $\mathbb{C}$
$\mathbb{M}_n^d$ d-tuples in $\mathbb{M}_n$
$\mathfrak{N}_n = \{X \in \mathbb{M}_n : (X)^n = 0$
$\mathcal{P}_n(y) = D_n\, e^{-n \sum_{j=1}^n Q(y_j)} \prod_{1 \leq i < j \leq n} |y_i - y_j|^2$
R_d are characteristic radii of a Gaussian distribution in d variables (see Thm. 3.1)
$\mathfrak{S}(n)$ the symmetric group on n elements.
$\mathcal{T}_n^d$ the commuting d-tuples of n-by-n matrices that are upper triangular with respect to a fixed basis
$\mathcal{U}(n)$ is the n-by-n unitary matrices
$v(d)$ denotes the volume of the unit ball in $\mathbb{R}^d$
$\tilde{w}$ is the function of the eigenvalues λ of $X \in \mathfrak{C}_n^{d,H}$ that satisfies $w(X) = \tilde{w}(\lambda)$
dX is Haussdorf measure on $\mathfrak{V}_n$
μ^Q, the equilibrium measure for Q, is the unique probability measure μ^Q that minimizes $I^Q(\mu)$ (see Subsection 3.2)
σ_R^{d-1} is normalized surface area measure on the sphere RS^{d-1}

Acknowledgements. This work was done while the author was visiting the Department of Mathematics at the University of California at Berkeley. The author would like to thank the Department for its hospitality. The author would also like to thank the anonymous referee for many helpful suggestions that improved the readability of the paper.

References

[1] Jim Agler, John E. McCarthy, and N. J. Young. Operator monotone functions and Löwner functions of several variables. *Ann. of Math. (2)*, 176(3):1783–1826, 2012.

[2] Yacin Ameur, Hå kan Hedenmalm, and Nikolai Makarov. Fluctuations of eigenvalues of random normal matrices. *Duke Math. J.*, 159(1):31–81, 2011.

[3] Yacin Ameur, Haakan Hedenmalm, and Nikolai Makarov. Random normal matrices and Ward identities. *Ann. Probab.*, 43(3):1157–1201, 2015.

[4] Yacin Ameur, Nam-Gyu Kang, and Nikolai Makarov. Rescaling Ward identities in the random normal matrix model. *Constr. Approx.*, 50(1):63–127, 2019.

[5] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[6] Gordon Blower. *Random matrices: high dimensional phenomena*, volume 367 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2009.

[7] Sergiy V. Borodachov, Douglas P. Hardin, and Edward B. Saff. *Discrete energy on rectifiable sets*. Springer Monographs in Mathematics. Springer, New York, [2019].

[8] Djalil Chafai, Edward B. Saff, and Robert S. Womersley. On the solution of a Riesz equilibrium problem and integral identities for special functions. *J. Math. Anal. Appl.*, 515(1):Paper No. 126367, 29, 2022.

[9] Djalil Chafai, Edward B. Saff, and Robert S. Womersley. Threshold condensation to singular support for a Riesz equilibrium problem. *Anal. Math. Phys.*, 13(1):Paper No. 19, 2023.

[10] Ling-Lie Chau and Oleg Zaboronsky. On the structure of correlation functions in the normal matrix model. *Comm. Math. Phys.*, 196(1):203–247, 1998.

[11] Michael J. Crabb and Alexander M. Davie. Von Neumann's inequality for Hilbert space operators. *Bull. London Math. Soc.*, 7:49–50, 1975.

[12] Kenneth R. Davidson. *C^*-algebras by example*, volume 6 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 1996.

[13] P. A. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, volume 3 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

[14] Percy Deift and Dimitri Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2009.

[15] K. Dykema and U. Haagerup. Invariant subspaces of Voiculescu's circular operator. *Geom. Funct. Anal.*, 11(4):693–741, 2001.

[16] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Mathematical Phys.*, 3:140–156, 1962.

[17] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. II. *J. Mathematical Phys.*, 3:157–165, 1962.

[18] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. III. *J. Mathematical Phys.*, 3:166–175, 1962.

[19] Peter Elbau and Giovanni Felder. Density of eigenvalues of random normal matrices. *Comm. Math. Phys.*, 259(2):433–450, 2005.

[20] Murray Gerstenhaber. On dominance and varieties of commuting matrices. *Ann. of Math. (2)*, 73:324–348, 1961.

[21] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Mathematical Phys.*, 6:440–449, 1965.

[22] Alice Guionnet and Édouard Maurel-Segala. Matrix models at low temperature. https://arxiv.org/abs/2210.05239.

[23] R. Guralnick. A note on commuting pairs of matrices. *Lin. Multilin. Alg.*, 31:71–75, 1992.

[24] Robert M. Guralnick and B. A. Sethuraman. Commuting pairs and triples of matrices and related varieties. *Linear Algebra Appl.*, 310(1-3):139–148, 2000.

[25] Yongho Han. Commuting triples of matrices. *Electron. J. Linear Algebra*, 13:274–343, 2005.

[26] Håkan Hedenmalm and Aron Wennman. Planar orthogonal polynomials and boundary universality in the random normal matrix model. *Acta Math.*, 227(2):309–406, 2021.

[27] A. J. Hoffman and H. W. Wielandt. The variation of the spectrum of a normal matrix. *Duke Math. J.*, 20:37–39, 1953.

[28] John Holbrook and Matjaz Omladic. Approximating commuting operators. *Linear Algebra Appl.*, 327(1-3):131–149, 2001.

[29] Joachim Jelisiejew and Klemen Sivic. Components and singularities of Quot schemes and varieties of commuting matrices. *J. Reine Angew. Math.*, 788:129–187, 2022.

[30] Kurt Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.

[31] John E. McCarthy and Hazel T. McCarthy. Random anti-commuting Hermitian matrices. *Random Matrices Theory Appl.*, 13(4):Paper No. 2450019, 13, 2024.

[32] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[33] Mark Mineev-Weinstein, Mihai Putinar, and Razvan Teodorescu. Random matrices in 2D, Laplacian growth and operator theory. *J. Phys. A*, 41(26):263001, 74, 2008.

[34] James A. Mingo and Roland Speicher. *Free probability and random matrices*, volume 35 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017.

[35] T. S. Motzkin and Olga Taussky. Pairs of matrices with property L. II. *Trans. Amer. Math. Soc.*, 80:387–401, 1955.

[36] Nham V. Ngo and Klemen Sivic. On varieties of commuting nilpotent matrices. *Linear Algebra Appl.*, 452:237–262, 2014.

[37] Matjaz Omladic. A variety of commuting triples. *Linear Algebra Appl.*, 383:233–245, 2004.

[38] Leonid Pastur and Mariya Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[39] M. Pimsner and D. Voiculescu. Imbedding the irrational rotation C^*-algebra into an AF-algebra. *J. Operator Theory*, 4(2):201–210, 1980.

[40] Klemen Sivic. On varieties of commuting triples. *Linear Algebra Appl.*, 428(8-9):2006–2029, 2008.

[41] Klemen Sivic. On varieties of commuting triples III. *Linear Algebra Appl.*, 437(2):393–460, 2012.

[42] Terence Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[43] N.Th. Varopoulos. On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory. *J. Funct. Anal.*, 16:83–100, 1974.

[44] P. B. Wiegmann and A. Zabrodin. Conformal maps and integrable hierarchies. *Comm. Math. Phys.*, 213(3):523–538, 2000.

[45] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, 62:548–564, 1955.

[46] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. II. *Ann. of Math. (2)*, 65:203–207, 1957.

[47] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.