A Hyperpfaffian Formulation of the Partition Function for the Circular Ensemble when β is an Even Square

Caleb H. Farrow

A thesis presented for the degree of

Bachelor of Science

Department of Mathematics

University of Oregon

Eugene, Oregon, United States of America

May 2024

Abstract

In random matrix theory, we seek to describe the probability density of the eigenvalues of random matrices. Random matrix ensembles appear in a wide variety of areas, including statistical mechanics, mathematical neuroscience, and the study of the Riemann zeta function. Here, we focus on the circular ensemble, providing an overview of modern techniques that show promise to expand our knowledge of these fascinating mathematical objects. We begin by establishing background material in algebra, measure theory, and measure-theoretic probability. Then, we introduce the foundations of random matrix theory with a focus on the circular ensemble. We present a hyperpfaffian formulation of the partition function for the circular ensemble when β is an even square, as Chris Sinclair demonstrated in 2018. This is a step towards solvability of the circular ensemble when β is an even square.

Acknowledgements

To my family and my friends, thank you for indulging me when going on my tangents about math. Additionally, I'd like to thank my coworkers at the Tutoring and Academic Engagement Center for listening to my explanations as I refined them. In particular, I want to extend my gratitude to Emily Elbaum and David Schmidt. Emily, thank you for your perspective regarding background material. David, thank you for asking the right questions to help me discover and fill in the gaps in my knowledge. To both of you, thank you for your unwavering support and kindness.

Thank you to the Kern-Ralph co-lab. You welcomed me with open arms and taught me how to engage in constructive academic discussions. I've thoroughly enjoyed learning about each of your projects, and the relevant tools in computational population genetics. I would not be where I am today were it not for the support of my graduate student mentor Victoria Caudill. Thank you for being confident in my abilities before I was, and for your advice about graduate school and beyond.

I'd also like to thank my committee members, Dr. Ben Young and Dr. Peter Ralph. Thank you both for supporting me through the writing process of this work, and for offering guidance as I'm learning how to navigate the world of academia. Special thanks to my advisor, Dr. Chris Sinclair for being my guide through the wide world of mathematics, and for helping me channel my curiosity about randomness into a tangible result. Thank you for being a great advisor, and for showing me a piece of what it's like to be a mathematician.

For all the students I've tutored, who inspire me to work as hard as they do.

Contents

1 Introduction and Background **5**

1.1 Introduction 5

1.2 Preliminaries 5

1.2.1 Constructing the exterior algebra 6

1.2.2 Determinants and Pfaffians 10

2 A Rigorous Perspective on Probability Theory **14**

2.1 Measure-Theoretic Foundations 14

2.1.1 σ-algebras 15

2.1.2 Measures 16

2.1.3 Measurable functions 17

2.1.4 A few special constructs to note 17

2.2 Foundations of Probability Theory 19

2.2.1 Probability spaces 20

2.2.2 Random variables 20

2.2.3 Examples 21

3 β Ensembles: An Overview **22**

3.1 Identities, Properties, and Useful Constructions 23

3.1.1 The Confluent Vandermonde 23

3.2 The Partition Function when β is an Even Square 23

3.3 Connection to Correlation Functions 27

A More preliminaries **29**

A.1 Set theory, mathematical foundations 29

A.1.1 Cartesian product 29

A.1.2 Equivalence relations 29
A.2 Linear algebra 30
A.2.1 Subspaces 30
A.2.2 Trace 30
A.3 Abstract algebra 30
A.3.1 Groups 30
A.3.2 Rings 31

Chapter 1

Introduction and Background

1.1 Introduction

The study of matrix-valued random variables is motivated in part by our continued fascination with the arrangement and interactions of particles in space. If our model is to describe the random positions of charged particles, we want the particles to be exchangeable and to repel each other. Hence, we study exchangeable random variables that repel, one of the simplest of such being the circular ensemble of random matrices.

Random matrix theory wishes to describe the behavior of a system when the number of particles is very large. Thus, it is natural to study the distribution of the eigenvalues of these random matrices. This research will focus on explicit computations when the number of particles is fixed and finite. The distribution depends heavily on the parameter β, which is related to the total energy in the system and often interpreted as inverse temperature. In particular, we will discuss the machinery needed to work towards solving the circular ensemble when β is an even square integer. Here, solvability means knowledge of the correlation functions.

1.2 Preliminaries

We assume the reader has a strong background in calculus and linear algebra. We present select topics from algebra, probability theory, and analysis so that such a reader may understand the computations later in the work. For additional background material, the interested reader may reference the Appendix.

1.2.1 Constructing the exterior algebra

The exterior algebra is a graded algebra with an alternating product. Here we construct the exterior algebra from the tensor algebra, rather than solely presenting it as the algebra generated by the appropriate relations. We do this to build intuition, and to provide some additional evidence regarding the existence of such an algebra. The machinery that the exterior algebra yields will be important in our mission moving forward, so it is important to understand its basic components.

Quotients

We will define V/U, the quotient of a vector space V by a subspace $U \subset V$. Such an object will be useful to us, as it will allow us to partition the vector space V into sets of elements of V we wish to view as equivalent. To do this, we should use an *equivalence relation* in our definition.

Definition. Let V be a vector space and $U \subset V$ be a subspace. Define an equivalence relation $\sim$ on V so that for $x, y \in V$, $x \sim y$ if $x - y \in U$. The *quotient space* of V by U is the set of orbits under the relation $\sim$, often denoted as $V/\sim$ or V/U.

For example, let V and W be vector spaces and let $T : V \to W$ be a linear transformation. Let $v_1, v_2 \in V$. If we take $U = \ker T$, then $v_1 \sim v_2$ whenever $T(v_1) - T(v_2) = 0$. This allows us to define a bijective map $\phi : V/U \to \mathrm{Im}(T)$. The existence and uniqueness of ϕ is referred to as the *First Isomorphism Theorem*. Quotients are particularly useful when we need such a bijective map, or when we wish to construct a new space around an interesting equivalence relation.

Products and tensors

The word product is quite ambiguous. For example, say we've defined multiplicative groups G and H. How do we define the product of these two groups? We may not know how to multiply elements of G with elements of H, which certainly causes problems. We alleviate this issue by defining some different notions of products that are distinct and useful in different contexts. Readers are likely familiar with the Cartesian product, which solves the product problem on sets X and Y by isolating elements of X in one coordinate and elements of Y in the other. This is a product of sets and so does not contain any algebraic information. We want our products to preserve the structure of the algebraic objects they are built from, so we need to modify this notion.

Definition. Let $\mathbb{F}$ be a field, and V, W be finite-dimensional vector spaces over the same field $\mathbb{F}$. We define the *direct product* $V \oplus W$ as the Cartesian product of V and W as sets, $V \times W$, together with the binary

operation + and scalar multiplication by elements of $\mathbb{F}$, defined as follows. Let $(v_1, w_1), (v_2, w_2) \in V \oplus W$, and $+_V, +_W$ denote addition in V and W respectively. Then we define the operation + so that $(v_1, w_1) + (v_2, w_2) = (v_1 +_V v_2, w_1 +_W w_2)$. Now let $\lambda \in \mathbb{F}$. We define $\lambda(v_1, w_1) \coloneqq (\lambda v_1, \lambda w_1)$.

We leave it to the interested reader to verify that with these operations, $V \oplus W$ does indeed define a vector space over the field $\mathbb{F}$.

While the direct product is one solution to the product problem, what if we want to define a way to "multiply" vectors in a vector space V? Readers are likely familiar with the dot product. Let's recall the formal definition of this construction.

Definition. Let $\vec{v} = (v_1, ..., v_n), \vec{w} = (w_1, ..., w_n) \in V$, a finite-dimensional vector space over a field $\mathbb{F}$. Then the *dot product* of $\vec{v}$ with $\vec{w}$, denoted $\vec{v} \cdot \vec{w}$, is a function $V \times V \to \mathbb{F}$ where $(\vec{v}, \vec{w}) \mapsto v_1 w_1 + v_2 w_2 + ... + v_n w_n$.

This product has some nice uses, like detecting orthogonality of vectors. There are lots of nice products like this one, called inner products.

Definition. An *inner product* on a vector space V over a field $\mathbb{F}$ is a function $\langle -, - \rangle : V \times V \to \mathbb{F}$ that is linear in the first argument, conjugate-symmetric, and positive-definite.

Notice that this product is not a binary operation on the vector space V. Some fun examples of inner products include the L^2 inner product, the expectation of a product of random variables, and the familiar dot product. Every inner product is also a *bilinear form*, a function $\beta : V \times V \to \mathbb{F}$ that is only required to be linear in both arguments. We will be particularly interested in bilinear forms, since they will help us understand tensor products.

The *tensor product* seeks to combine our two strategies for taking products. Given finite-dimensional vector spaces V and W over the same field $\mathbb{F}$, we want to define an object $V \otimes W$ that uniquely characterizes bilinear forms from $V \times W$ to Z, another vector space over $\mathbb{F}$. In other words, for each bilinear pairing $\beta : V \times W \to \mathbb{F}$, we want a corresponding map $\tilde{\beta} : V \otimes W \to Z$. We can express β as multiplication by some matrix with respect to some bases $b_1, ..., b_n$ for V and $c_1, ..., c_n$ for W. Each entry in this matrix is a pairing of basis vectors (b_i, c_j). We now give each of these pairings the notation $b_i \otimes c_j$, and declare these the basis vectors for $V \otimes W$. This is not quite enough to define the space $V \otimes W$, since we need $\otimes$ to be bilinear. This means we need to assure that for $r_1, r_2, s_1, s_2 \in \mathbb{F}$, $v_1, v_2, v \in V$, $w_1, w_2, w \in W$, $(r_1 v_1 + r_2 v_2) = r_1(v_1, w) + r_2(v_2, w)$ and $(v, s_1 w_1 + s_2 w_2) = s_1(v, w_1) + s_2(v, w_2)$. We have a tool that lets us accomplish this: quotients! Let $R = \text{span}\{(r_1 v_1 + r_2 v_2) - r_1(v_1, w) - r_2(v_2, w), (v, s_1 w_1 + s_2 w_2) - s_1(v, w_1) - s_2(v, w_2)\}$, and let $F\langle V \times W \rangle$ denote the free vector space over $V \times W$. This is the space of functions from $V \times W$ to the field $\mathbb{F}$ that are nonzero at only finitely many points. We construct $V \otimes W$ by taking the quotient of this free

vector space by R.

Defintion. $V \otimes W \coloneqq F\langle V \times W\rangle / R$.

This yields a vector space that satisfies only the relations necessary for bilinearity, with no extra redundant structure. Elements in this space are formal sums of *pure tensors*, elements that can be written as $v \otimes w$ if $v \in V$ and $w \in W$.

The tensor algebra

Before we define the tensor algebra, we provide a brief reminder about the linear algebra notion of an algebra over a field. Readers may be most familiar with the definition of an algebra from analysis or ring theory. In analysis, we often call a family of functions an *algebra* if it's closed under addition, multiplication, and scalar multiplication. In ring theory, an R-algebra is defined as a ring A together with a homomorphism from the commutative ring R to the center of A, all the elements of A that commute. These seem like distinct concepts, but in linear algebra, they coincide.

Definition. Let $\mathbb{F}$ be a field. An *algebra* $\mathcal{A}$ over $\mathbb{F}$ is a vector space together with a bilinear pairing.

In the previous section, we introduced the tensor product, which is a particularly useful tool for reasoning about bilinear forms. We have an equivalent formulation of this definition using the language of tensors.

Definition. An *algebra* over $\mathbb{F}$ is a vector space V with a multiplicative identity 1 and a linear transformation $L : V \otimes V \to V$ such that for $x, y, z \in V$, $L(x \otimes 1) = x - L(1 \otimes x)$ and $L(x \otimes L(y \otimes z)) = L(L(x \otimes y) \otimes z)$.

Keep in mind that L is just a bilinear pairing from $V \times V$ into V that satisfies some extra properties that make it algebraically compatible with the operation $\otimes$. This definition is certainly more complicated, but it allows us to reason about the space of bilinear pairings rather than each bilinear pairing on V independently. It proves useful as we move to define the tensor algebra on a vector space V. Before we do this, let's establish some notation. We will write $V^{\otimes i}$ to denote $V \otimes V \otimes ... \otimes V$, the tensor product of V with itself having i copies of V. To emphasize its role in the tensor algebra, we will also use $T^k(V)$ as another way to write $V^{\otimes k}$.

Definition. Let V be a vector space. Then *tensor algebra* on V, $T(V)$, is the set $\bigoplus_0^\infty T^k(V)$ with multiplication given by concatenation of tensors.

This is enough to make $T(V)$ an algebra. In fact, it is a *graded algebra*, an algebra that can be written as a direct sum of vector spaces. One interesting property of the tensor algebra over a vector space V of dimension 1 is that it is isomorphic to polynomials over $\mathbb{F}$. Indeed, if we choose a basis $t \in V$, we can

construct a map under which $1 \mapsto 1$, $t \mapsto t$, $t \otimes t \mapsto t^2$, $t \otimes t \otimes t \mapsto t^3$, and so forth. It is comforting that this construction agrees with the familiar space of polynomials in the case where $\dim V = 1$. Readers who are still uncomfortable with the tensor algebra are encouraged to understand this connection, and extend their knowledge of polynomials over a field to the tensor algebra.

The exterior algebra

We are now prepared to give the definition of the exterior algebra. Our goal is to create a graded algebra $\Lambda(V)$ in which for any element $v \in \Lambda(V)$, $v \wedge v = 0$, where $\wedge$ denotes the product that makes $\Lambda(V)$ an algebra and satisfies this extra property. Notice that the tensor algebra is a natural starting point here, since it is already a graded algebra.

Similar to when we defined the tensor product, we can use quotients to help us satisfy the additional properties that we want our algebra to have. In this case, we want to satisfy the relation $v \otimes v = 0$, where $v \in V$. We must be more careful in defining the set we quotient by than when we defined the tensor product. In that case, our underlying structure was a vector space, but here it is only an algebra. We need our algebraic operations to be well-defined on the orbits under the equivalence relation we quotient by. For this reason, we need to quotient by the *ideal* generated by the relation $v \otimes v = 0$. Recall that an ideal I in a ring $(R, +, \times)$ is a subgroup of $(R, +)$ that is closed under multiplication by any element of R. Notice that subspaces already satisfy these properties, and this is why we did not need to add additional constraints earlier.

Definition. Let V be a vector space, and I be the ideal generated by the relation $v \otimes v = 0$. The *exterior algebra*, $\Lambda(V)$ is the quotient $T(V)/I$ with multiplication $\wedge$ defined as follows. If $v, w \in \Lambda(V)$, then $v \wedge w := v \otimes w \bmod I$.

One interesting property of the wedge product $\wedge$ is that it is an alternating product. In fact, any product $\times$ on a ring R that satisfies the property that $r \times r = 0$ is alternating. See the appendix for a brief proof of this fact. These properties of the exterior algebra will be extremely important to our future calculations.

Exterior powers

The exterior algebra is graded, so it can be expressed as a direct sum of vector spaces. We will work in these vector spaces, the *exterior powers* of the vector space V.

Definition. Let V be a vector space. The k^{th} *exterior power* of V, denoted by $\Lambda^k(V)$ is a subspace of $\Lambda(V)$ that is spanned by elements that look like a wedge product of k vectors in V.

The elements of these subspaces are called *blades*, and if $\{e_1, ..., e_n\}$ is a basis for V, then $\Lambda^k(V)$ is spanned by $e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}$ for $1 \leq i_1 \leq i_k \leq n$. From this, it is not unexpected that the dimension of $\Lambda^k(V)$ is $\binom{n}{k}$, for an n-dimensional vector space V.

1.2.2 Determinants and Pfaffians

The study of determinants in part is the study of invariants of matrices under a change of basis. Beginning as a tool in solving systems of linear equations, determinants have become quintessential to the study of mathematics. Here, we present a more advanced perspective on the determinant using the exterior algebra.

Familiar formulas for the determinant

One important formula for the determinant is the *Leibniz formula*, given below for an arbitrary $n \times n$ matrix A.

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)},$$

where S_n is the symmetric group of order n, sgn is the sign homomorphism, and subscripts specify elements in the matrix. For example, when $n = 2$ and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

$$\begin{aligned} \det(A) &= \operatorname{sgn}(\mathrm{id}) A_{1,1} A_{2,2} + \operatorname{sgn}((12)) A_{1,2} A_{2,1} \\ &= ad - bc, \end{aligned}$$

which is the familiar formula for the determinant of a 2×2 matrix.

The determinant can also be expressed as the product of all eigenvalues of the matrix. This follows from the fact that eigenvalues are the roots of the characteristic polynomial of the matrix. So for an $n \times n$ matrix A,

$$\det(A) = \prod_{i=1}^{n} \lambda_n,$$

where λ_n are the n eigenvalues of the matrix A. Note that if an eigenvalue has multiplicity greater than one, we include this multiple times in our product.

The determinant in the exterior algebra

We can also use the exterior algebra to define the determinant. Let V be an n-dimensional vector space, $v_1, ..., v_n \in V$, and $T : V \to V$ be a linear transformation. This induces a map $T_n : \Lambda^n(V) \to \Lambda^n(V)$ where $\bigwedge_{i=1}^n v_i \mapsto \bigwedge_{i=1}^n T(v_i)$. But $\Lambda^n(V)$ is isomorphic to the ground field, so T_n must be multiplication by some scalar. We call this scalar the *determinant* of the matrix which represents the linear transformation T. The advantage of this method is that there are no coordinates. As the determinant is invariant under a change of basis, it is natural that there is a way to calculate it without needing to choose a basis.

The Wronskian

We will define a quantity called the *Wronskian* that will help us rewrite objects as determinants of matrices.

Definition. Let $f_1, ..., f_n$ be $(n-1)$-differentiable functions. For simiplicity, we assume these functions are real or complex valued. We will define an operator D^l on the space of functions on $\mathbb{R}$ or $\mathbb{C}$, which will give us the l^{th} derivative of the input function with a combinatorial factor. An *operator* is simply a function on a space of functions that maps functions to functions. With this clarification, we may define

$$D^l(f(x)) = \frac{1}{l!}\frac{d^l f}{dx^l}.$$

For our purposes, we will consider integer values of l in the range $[0, n)$. We can then define the *Wronskian* of a family of polynomials P_t by

$$\mathrm{Wr}(P_t; x) = \det \begin{bmatrix} p_1(x) & p_2(x) & \dots & p_n(x) \\ D^1 p_1(x) & D^1 p_2(x) & \dots & D^n p_n(x) \\ \vdots & \vdots & \ddots & \vdots \\ D^{(n-1)} p_1(x) & D^{(n-1)} p_2(x) & \dots & D^{(n-1)} p_n(x) \end{bmatrix}.$$

These Wronskians are applicable in many areas including differential equations, and take a pivotal role in our calculations.

Pfaffians

Let's examine the determinant formula once more, now assuming that A is a $2n \times 2n$ skew-symmetric matrix.

$$\det(A) = \sum_{\sigma \in S_{2n}} \operatorname{sgn}(\sigma) \prod_{i=1}^{2n} A_{i,\sigma(i)}.$$

Since A is skew-symmetric, we know that $A_{i,j} = -A_{j,i}$. One might expect that this special structure could result in some structure in the determinant, or perhaps simplification in the formula. In fact, the determinant of an anti-symmetric matrix can always be written as the square of a polynomial in the entries of the matrix. What can we learn from such a polynomial? We want to introduce a new function, called the *pfaffian* of a matrix, which we will denote by Pf, and we want our definition to satisfy the constraint that $(\operatorname{Pf} A)^2 = \det A$. We present two equivalent definitions of this object, and compare them to our definitions of determinants.

Definition. Let A be a $2n \times 2n$ skew-symmetric matrix. We define

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

The similarities between this formulation of the definition of the Pfaffian and the Leibniz formula for the determinant are clear, but it is not obvious that $(\operatorname{Pf} A)^2 = \det A$. There are many alternative proofs of this fact, many of which are worked through in Haber's 2015 notes [11].

Similar to determinants, we have a formula for the Pfaffian in the exterior algebra.

Definition. Let $A = [a_{ij}]$ be a $2n \times 2n$ skew-symmetric matrix, and $e_1, ..., e_{2n}$ be a basis of the field we are working over, typically $\mathbb{R}$ or $\mathbb{C}$. We define

$$\alpha = \sum_{i<j} a_{ij} e_i \wedge e_j.$$

This is on the determinantal line, so the Pfaffian of A is simply

$$\operatorname{Pf}(A) = \frac{\alpha^{\wedge N}}{N!},$$

Hyperpfaffians

The previous definition is particularly useful in extending this notion to the exterior algebra, where not every element may have a matrix representation. Note that if M is an even integer and $\omega \in \Lambda^M \mathbb{F}^{NM}$, then

$\frac{\omega^M}{M!}$ is on the determinantal line. This means that definition is also sensible in these spaces, so we define the *hyperpfaffian* as follows.

Definition.

$$\mathrm{PF}(\omega) = *\frac{\omega^{\wedge M}}{M!},$$

where $*$ denotes the hodge star operator.

Chapter 2

A Rigorous Perspective on Probability Theory

In this section, we review undergraduate probability and introduce some measure-theoretic notions necessary for our later constructions. The use of measure theory here allows us to use additional tools that make sense in probabilistic situations.

2.1 Measure-Theoretic Foundations

When working with probabilistic objects, we will almost surely encounter extremely rare events. A problem arises when ascribing probabilities to such events: what if we have infinitely many of them? For an example, let X be a uniformly-distributed random variable on $[0,1]$. In other words, X is a random number between 0 and 1, with the probability that X is in a subinterval of length c is c, where $c \in [0,1]$. For any point $x \in [0,1]$, consider the probability that $X = x$. We want this to be small because there are infinitely many numbers between 0 and 1. In fact, our only solution here is to take $\mathbb{P}(X = x) = 0$. Otherwise, our probability-ascribing function fails to be a distribution because for every $\varepsilon > 0$, $\sum_{x \in [0,1]} \varepsilon$ diverges.

Measure theory provides a natural framework for dealing with these probability 0 events. Here we remind the reader that events with probability 0 can occur, and do so quite frequently! Continuing the example, every time we sample from the uniform distribution on $[0,1]$, a probability zero event occurs. We will define σ-algebras, measures, and finally measure spaces. Then, we will discuss a few special σ-algebras and measures.

2.1.1 σ-algebras

We begin with a brief reminder of a topology $\mathcal{T}$ on some space T.

Definition. A *topology* $\mathcal{T}$ on a space X is a collection of subsets of X satisfying the following properties.

1. $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$.
2. If $\{X_\alpha\}_{\alpha \in \mathcal{A}} \in \mathcal{T}$, then $\bigcup \{X_\alpha\}_{\alpha \in \mathcal{A}} \in \mathcal{T}$.
3. If $X_1, ..., X_n \in \mathcal{T}$, then $\bigcap_i X_i \in \mathcal{T}$.

We take these sets in our topology to be the open sets in our space, and define the rest of our topological notions based on these open sets. The pair $(X, \mathcal{T})$ is referred to as a *topological space*. Measure theory has a similar foundational construction upon which we build the rest of the concepts: σ-algebras.

Definition. A *σ-algebra* $\sigma(X)$ on a set X is a collection of subsets of X satisfying the following properties.

1. $\varnothing \in \sigma(X)$
2. If $Y \subset X \in \sigma(X)$, then $Y^c \in \sigma(X)$, where Y^c denotes the complement of the set Y in X.
3. For any countable collection of elements in $\sigma(X)$, $E_1, E_2, ...$, their union $\bigcup_{i=0}^{\infty} E_i \in \sigma(X)$.

A pair $(X, \sigma(X))$ is called a *measure space*. Notice the similarities between this definition and the definition of a topology above. Both require the empty set and the whole set as elements, and they are closed under some operations on sets. They differ in which set-theoretic operations that they are closed under. Notice that a topology must be closed under arbitrary unions, but a σ-algebra need only be closed under countable unions. Additionally, σ-algebras are closed under complements, while topologies are not. These differences appear because of the difference in applications of these concepts.

We want to use topologies to construct a notion of open sets on a space. Topological spaces are most interesting when not all sets are open. The complement of an open set is closed, so if we required topologies to be closed under complements, the structure we induce on our space wouldn't be very interesting to study from the perspective of topology.

In contrast, we want to use σ-algebras to ascribe probabilities to events. Closure under arbitrary unions would cause our probability-ascribing structures we will define in this chapter to give us undesirable results. For example, if we wanted to define a probability structure on $[0, 1] \subset \mathbb{R}$, it would be necessary that each point has probability zero. But then we could also write $[0, 1]$ as the union of all of its points, and this would imply that all events in this space occur with probability 0.

It is often laborious to write down every element of a σ-algebra. We can also reason about σ-algebras that are *generated* by some collection of subsets of the ambient space X.

Definition. Let F be a collection of subsets of a set X. Define the σ-algebra *generated by* F to be the intersection of all σ-algebras containing F.

This is indeed a σ-algebra, and in fact is the smallest σ-algebra containing F. This is intuitive from the definitions we've given so far. For a proof of this fact, see the appendix.

2.1.2 Measures

Readers have likely seen before that our notion of distance can be generalized. Given a metrizable space X, there are many different functions $d : X \times X \to \mathbb{R}$ that satisfy the properties that we want a distance to have. We will not review metric spaces here, and instead will immediately transition to discussing the generalizations of areas.

Often, the first areas that we are introduced to in mathematics are areas of two-dimensional polygons, like squares or triangles. Then, we move to more complicated shapes in three dimensions, and then to integrals. For the purpose of this construction, we will focus on the properties of areas and volumes of shapes. Both areas and volumes of this variety are unsigned, and can be computed by adding up areas or volumes of smaller shapes. Additionally, we want the empty shape to have no area. We will formalize this property in our definition of a measure on a set.

Definition. Given a set X together with a σ-algebra $\sigma(X)$, the function $\mu : \sigma(X) \to \mathbb{R} \cup \infty$ is a *measure* if it satisfies the following properties.

1. $\mu(\varnothing) = 0$
2. For all sets $E \in \sigma(X)$, $\mu(E) \geq 0$
3. For any countable collection of disjoint sets $\{E_i\}_{i=0}^{\infty}$,

$$\mu\left(\bigcup_{i=0}^{\infty} E_i\right) = \sum_{i=0}^{\infty} \mu(E_i)$$

When we combine a set X, a σ-algebra $\sigma(X)$, and a measure μ, the resulting triple $(X, \sigma(X), \mu)$ is called a *measure space*. We will now present some examples of useful measure spaces.

2.1.3 Measurable functions

Each field in mathematics has its own collection of structure-preserving maps. Continuing our comparisons to topology, recall the following definition.

Definition. Given two topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{S})$, we say a map $f : X \to Y$ is *continuous* if for every open set $U \in Y$, $f^{-1}(U)$ is open in X.

In other words, continuous maps preserve the open set structure between topological spaces. Measure theory has a similar construction: measure-preserving functions.

Definition. Let $(X, \sigma(X)), (Y, \tau(Y))$ be measure spaces, and $f : X \to Y$. f is *measurable* if for every $E \in \tau(Y)$, $f^{-1}(E) \in \sigma(X)$.

The parallels between this definition and the definition of continuous maps between topological spaces are clear. The notable difference between the two is that continuous maps preserve open sets under taking preimages, while measurable functions preserve sets that belong to the σ-algebra under taking preimages.

2.1.4 A few special constructs to note

The Borel σ-algebra on $\mathbb{R}$

We saw earlier that topological spaces and σ-algebras are related, but distinct concepts. A natural question to ask is how we can construct a σ-algebra given some topological space $(X, \mathcal{T})$.

Since the foundational objects of topological spaces are open sets, a natural choice to study would be the σ-algebra generated by all open sets in $(X, \mathcal{T})$. We call this the *Borel* σ-algebra, and if $X = \mathbb{R}$, it is the same as the σ-algebra generated by all open intervals of the form (a, b) for $a, b \in \mathbb{R}$. We will write $\mathcal{B}(X)$ to denote the Borel σ-algebra on the topological space $(X, \mathcal{T})$.

In the real case, we can see why the Borel σ-algebra is a useful σ-algebra to study. If we want to put a measure on the real line, we want this to capture the idea of the length of an interval. It then is reasonable for us to take open intervals as the foundational structure in our σ-algebra.

Counting measure

Let $(X, \sigma(X))$ be a measure space. Perhaps the simplest example of a measure on $(X, \sigma(X))$ is the counting measure.

Definition. The *counting measure* $N : \sigma(X) \to \mathbb{Z}_{\geq 0}$ on $(X, \sigma(X))$ counts the number of elements of each set in $\sigma(X)$.

In other words, if $E \in \sigma(X)$, $N(E) = |E|$ if E is finite, and $N(E) = \infty$ if E has infinitely many elements. This construction will be useful when we define point processes, as we will need a way to count the number of points in a particular area.

Lebesgue measure on $\mathbb{R}$

We want to define a measure λ that tells us the length of any given set in $\mathbb{R}$. We understand how to do this when we want to find the length of an interval $I_1 = (a, b)$ or $I_2 = [a, b]$. We want λ to satisfy the property that $\lambda(I_1) = b - a = \lambda(I_2)$. It is less clear what the right action is when we encounter a set like the rational numbers. We know that this set is dense in $\mathbb{R}$, but how much of the real line is covered by rational numbers?

In analysis, many concepts can be traced back to convergence of sequences. We are also familiar with the concept of open covers in topology. It then seems reasonable that we may be able to use a sequence of open covers consisting of only open intervals to define the length of a more interesting subset of the real line.

Definition. Let (a, b) be an open interval in X and define the length of (a, b), $(a, b) = b - a$. The *Lebesgue outer measure* λ^* of a set $E \subset \mathbb{R}$ is defined as the greatest lower bound of the sum of the lengths of open covers of E that consist only of open intervals. In other words,

$$\lambda^*(E) = \inf\left\{ \sum_{k=1}^{\infty} l(I_k) : (I_k)_{k\in\mathbb{N}} \text{ is a sequence of open intervals such that} E \subset \bigcup_{k=1}^{\infty} I_k \right\}.$$

Notice that we only defined the Lebesgue outer measure. In order for λ to satisfy the length properties we want, for a set E we wish to measure the length of, we need $\lambda(A) = \lambda(A \cap E) + \lambda(A \cap E^c)$ for all $A \subset X$, which is not an immediate consequence of this definition. So, if the Lebesgue outer measure satisfies the condition that $\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$ for all $A \subset X$, we take $\lambda = \lambda^*$. Otherwise, we say that E is not Lebesgue-measurable. We claim that the set of all measurable subsets of $\mathbb{R}$ form a σ-algebra, called the Lebesgue σ-algebra.

The Lebesgue measure is particularly useful in constructing the Lebesgue integral, which is nicer than the Riemann integral because it allows for easier commuting of limits and integrals. We will not construct the Lebesgue integral in this text, but the interested reader can refer to Rudin's book [19].

Haar measure on the unit circle

While the Lebesgue measure is useful, the requirements for its existence are quite strict. We now want to generalize Lebesgue measure so that we are able to define a similar construction on the unit circle $\mathbb{T}$.

In our discussion of Lebesgue measure, our goal was to define a measure that captured the intuitive

idea of the length of intervals, and generalized this. Here, we instead begin with an axiomatic approach, where we list the properties that a well-behaved length, area, or volume should have and construct Haar measure from these. We focus only on the unit circle $\mathbb{T}$, though these ideas can be generalized to any locally compact topological group.

Definition. $\mu : \mathbb{T} \to \mathbb{R}$ is the *Haar measure* on $\mathbb{T}$ with the Borel σ-algebra if the following are satisfied.

1. $\mu(e^{i\theta}S) = \mu(S)$ for every $S \in \mathcal{B}(\mathbb{T})$.
2. $\mu(S) < \infty$ for every $S \in \mathcal{B}(\mathbb{T})$.
3. For all Borel subsets S, $\mu(S) = \inf\{\mu(U) : S \subset U \text{ and } U \text{ open}\}$.
4. For all open sets U, $\mu(U) = \sup\{\mu(K) : K \subset U, K \text{ compact}\}$.

Property (1) is translation-invariance. If we slide a set to a new position on the unit circle, we don't want its length to change. (2) guarantees that no set on our unit circle has infinite length. (3) and (4) together are analogous to the covering and filling properties of the Lebesgue measure. We want to be able to obtain the length of oddly-shaped sets by considering open covers of Borel sets, and alternatively can obtain this length by "filling" the set with compact subsets. Now that we understand why these properties are reasonable, let's introduce the Haar measure on the unit circle.

Theorem. The Haar measure μ on the unit circle $\mathbb{T}$ is given below. Let $f : [0, 2\pi] \to \mathbb{T}$ such that $f(t) = (\cos(t), \sin(t))$. Then if $S \in \mathcal{B}(\mathbb{T})$ and λ denotes Lebesgue measure,

$$\mu(S) = \lambda(f^{-1}(S)).$$

This measure is unique up to scaling. Often, this is normalized by $\frac{1}{2\pi}$ so that the measure of the whole unit circle is 1.

2.2 Foundations of Probability Theory

With the prerequisites from measure theory established, we are now ready to introduce their analogues in probability theory. This transition from measure theory to probability theory is mostly a change in terminology, with a small number of extra properties that need to be satisfied.

2.2.1 Probability spaces

Say we want to perform an experiment whose outcome may change each time we execute it. We first need to describe what information we will record during this experiment. If our experiment is flipping a coin, we might be interested in the number of times the coin landed on heads. If we are investigating the effects of fertilizer on plant growth, we might record the height of the plants. The possible outcomes of the experiment we are studying make up our *sample space*, which we will denote as S. Oftentimes, we are not just interested in these measurements, however, but how many of them fall in a certain range. These ranges are the subsets of our sample space, and we associate these with *events*. The collection of all possible events forms a σ-algebra, $\sigma(S)$. We are also interested in the probability that a particular event occurs. To do this, we define a measure $\mathbb{P} : \sigma(S) \to [0,1] \subset \mathbb{R}$ that satisfies the extra property that $\mathbb{P}(S) = 1$. We call such a function a *probability measure*.

This property, alongside the other properties that a measure must have, ensure that all possible probability-ascribing is sensible. An example of a sensible feature resulting from these definitions is that if E is any event, then $\mathbb{P}(E) = 1 - \mathbb{P}(E^c)$.

A *probability space* is simply the collection of a sample space, sigma algebra, and probability measure. This construction provides us with the possible outcomes of the experiment, the events we can observe, and the probabilities assigned to these events.

2.2.2 Random variables

Imagine that we are scientists studying the effect that anti-inflammatory medications have on human body temperature. We may hypothesize that anti-inflammatory medications reduce body temperature, but we don't expect to find one exact temperature that a human's body will reach after taking such a medication. We expect this measurement to be different each time we observe it, even with the same patient. Yet, we can expect this temperature to be consistently "close to" a value. We want to use a mathematical structure that captures this patterned randomness, and we want this structure to be fundamental to this field of probability.

We already have such a construction from measure theory: measurable functions. This is precisely the machinery we need here. The probabilistic term for a measurable function is a random variable.

Definition. Let $(S, \sigma(S), \mathbb{P})$ be a probability space, $(T, \tau(T))$ be a measurable space, and $X : S \to T$. X is a *random variable taking values in* T if for every event $E \in \tau(T)$, $X^{-1}(E)$ is an event in $\sigma(S)$.

We regard $\sigma(S)$ as subsets of possible outcomes, T as the values X can take, and $\tau(T)$ as subsets we can ascribe probabilities to.

2.2.3 Examples

These definitions of probability spaces and random variables are often unintuitive to those familiar with the usual undergraduate-level construction of probability theory. Here, we present some concrete examples to make these abstract concepts more tangible.

Tossing a coin

We begin with a simple example of a coin toss. The possible outcomes of this experiment are the coin landing on heads (H), or the coin landing on tails (T). Then our sample space $S = \{H, T\}$. We also need a σ-algebra of possible events. This is $\sigma(S) = \{\varnothing, H, T, S\}$. Lastly, we need a probability measure $\mathbb{P}$. Let's assume our coin is fair. Then $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$. We can define a random variable X to be the random outcome of a single coin toss. We want to express the outcome H as a numerical quantity, perhaps 1, and the outcome T as 0. Then we construct a measurable space $(E, \mathcal{E})$ where $E = \{0, 1\}$ and $\mathcal{E} = \{\varnothing, \{0\}, \{1\}, E\}$.

Chapter 3

β Ensembles: An Overview

With the preliminary concepts established, we now return to our goal: to study the distribution of eigenvalues of random matrices. To do this, we must establish some additional concepts and terminology. We begin with the following definition.

Definition. A $n \times n$ *random matrix* over the field $\mathbb{F}$ is a random variable that takes values in the space of $n \times n$ matrices, which we will call $\mathrm{Mat}_{n\times n}(\mathbb{F})$.

For our purposes, we will study random matrices over $\mathbb{R}$ or $\mathbb{C}$. We typically characterize random matrices by their entries. The Ginibre ensembles, for example, have independent, normally distributed entries. We can also create new ensembles to study by modifying an existing ensemble. The circular ensembles are a modification of the Gaussian ensembles, where we restrict the eigenvalues of the matrices to the unit circle.

We now present the joint probability densities of eigenvalues for the circular ensemble. A similar formula exists for other ensembles, the only difference being the inclusion of a weight function.

Definition. The joint probability density of eigenvalues $\lambda_1, ..., \lambda_N$ for the circular ensembles is given by

$$Q_N(\lambda_1, ..., \lambda_N) = \frac{1}{C_N} \prod_{1\le k<j\le n} |\lambda_k - \lambda_j|^\beta. \tag{3.1}$$

Here, β is a parameter usually interpreted as inverse temperature, and C_N is a normalization constant that we call the *partition function* of the ensemble.

This constant can be expressed as a product of N terms using Selberg's integral formula. However, it is useful to obtain a hyperpfaffian expression for C_N because this is a step towards *solving* the ensemble. An ensemble is called solved when we have knowledge of its correlation functions. Most often, knowledge here

means that we can express the correlation functions as a determinant or Pfaffian. We will discuss correlation functions in greater detail in future sections.

3.1 Identities, Properties, and Useful Constructions

Here, we present some identities, properties, and constructions that will be necessary to prove that the partition function of the circular ensemble when β is an even square can be expressed as a hyperpfaffian.

3.1.1 The Confluent Vandermonde

We will define the confluent Vandermonde matrix. Compare this to the definition of the Wronskian given in section 1.2.2.

Definition 3.1. The *confluent Vandermonde* is an $NL \times NL$ matrix. We will construct this object from smaller $NL \times L$ matrices. Let $P_n(x)$ be a family of polynomials where the degree of $p_n(x) \in P_n(x)$ is $n-1$ and define the matrix

$$\mathbb{V}(x) = [D^{l-1} p_n(x)]_{n,l=1}^{NL,L}.$$

Then, we arrange these matrices as columns in a larger matrix, like so.

$$\mathbb{V}(\vec{x}) = \begin{bmatrix} \mathbb{V}(x_1) & \mathbb{V}(x_2) & \dots & \mathbb{V}(x_N) \end{bmatrix}.$$

This matrix is the confluent Vandermonde.

Due to its special structure, the determinant of this matrix has a simple formula.

$$\det(\mathbb{V}(\vec{x})) = \prod_{m<n} (x_n - x_m)^{L^2}.$$

3.2 The Partition Function when β is an Even Square

Let $\beta = L^2$ for some $L = 2n$ with $n \in \mathbb{Z}_{\geq 1}$, and $e_1, e_2, ..., e_{NL}$ be a basis for the vector space $\mathbb{R}^{NL}$. We define the set $\mathcal{T}$ to be the increasing functions with domain $\{1, 2, ..., L\}$ and codomain $\{1, 2, ..., NL\}$. Then for each $t \in \mathcal{T}$, we define

$$\epsilon_t = e_{t(1)} \wedge e_{t(2)} \wedge ... \wedge e_{t(L)}.$$

The collection of these ϵ_t then forms a basis for $\Lambda^L(\mathbb{R}^{NL})$. This follows from our construction of exterior powers in section 1.2.1.

We also define $\omega(x) \in \Lambda^L(V)$ by

$$\omega(x) = \sum_{t \in \mathcal{T}} \mathrm{Wr}(P_t; x)\epsilon_t.$$

In 2018, Chris Sinclair showed that

$$C_N = \mathrm{PF}\left(\int_{-\pi}^{\pi} \omega(e^{i\theta}) d\mu(\theta) \right),$$

for any choice of monic family of polynomials P_t, and β and even square. Here, we present a proof following the same logic as section 4.1 in Sinclair's 2018 paper [22].

Definition 3.2. When $\beta = L^2$ is even, the partition function for the circular ensemble is given by

$$C_N = \mathrm{PF}\left(\int_{-\pi}^{\pi} \omega(e^{i\theta}) d\mu(\theta) \right).$$

Proof. The probability density of eigenvalues is given by equation (3.1), so for this to be a probability density it must hold that

$$C_N = \frac{1}{N!} \int_{[-\pi,\pi)^N} \prod_{m<n} |\lambda_m - \lambda_n|^\beta d\theta_1 d\theta_2 \dots d\theta_N.$$

Since we are working in the circular ensemble, all eigenvalues must lie on the unit circle. Let's rewrite these eigenvalues as eigenangles. Then

$$C_N = \frac{1}{N!} \int_{[-\pi,\pi)^N} \prod_{m<n} |e^{i\theta_n} - e^{i\theta_m}|^\beta d\theta_1 d\theta_2 \dots d\theta_N.$$

By 11.3.2 in Mehta's book, we know that

$$|e^{i\theta_n} - e^{i\theta_m}| = -ie^{-i/2(\theta_n+\theta_m)} \operatorname{sgn}(\theta_n - \theta_m)(e^{i\theta_n} - e^{i\theta_m}).$$

This identity can be proven using Euler's formula, some trig identities, and properties of the absolute value in the complex plane. We will not provide a formal proof here. Instead, we make a substitution to write

$$C_N = \frac{1}{N!} \int_{[-\pi,\pi)^N} \prod_{m<n} [-ie^{-i/2(\theta_n+\theta_m)} \operatorname{sgn}(\theta_n - \theta_m)(e^{i\theta_n} - e^{i\theta_m})]^\beta d\theta_1 d\theta_2 \dots d\theta_N.$$

We will now make a change of variables. We take

$$d\mu(\theta) = (-ie^{-i\theta/2})^{(N-1)}d\theta,$$

and remember that β is even so that

$$C_N = \frac{1}{N!}\int_{[-\pi,\pi)^N}\prod_{m<n}(e^{i\theta_n} - e^{i\theta_m})^\beta d\mu^N(\theta).$$

This is an integral of a determinant of a confluent Vandermonde matrix, which we defined in the previous section. We may then write

$$C_N = \frac{1}{N!}\int_{[-\pi,\pi)^N}\det\mathbb{V}(e^{i\theta})d\mu^N(\theta).$$

Using cofactor expansion, we can rewrite this determinant as a weighted sum of the determinants of minors. The minors that will be most useful to us are $L \times L$ blocks that are each in terms of exactly one of the variables $\theta_1, ..., \theta_N$. Then, given some function $t \in \mathcal{T}$, we can define

$$V_t(e^{i\theta}) = [D^{l-1}p_{t(n)}\theta]_{n,l=1}^L.$$

Notice that the only difference between this definition and the usual confluent Vandermonde $\mathbb{V}(\theta)$ is the arrangement of our polynomial family in the matrix. Instead of ordering from 1 to N, we order according to the increasing function t. This ensures that we are able to consider minors with non-adjacent rows or columns.

Then we can take increasing functions with disjoint ranges $t_1, ..., t_N$ and write this determinant as an alternating sum over products of the $\mathbb{V}_{t_j}(e^{i\theta_j})$. Here we take disjoint ranges of functions so that we avoid using the same minor more than once.

Since this will be an alternating sum, we must discuss its sign. We are permuting the order of rows and columns by increasing functions, so we need to assign these a sensible notion of sign. This should be determined by the appropriate wedge product of the ϵ_t as defined in the beginning of this section. More precisely, we define

$$\mathrm{sgn}(t_1, ..., t_N) = \frac{1}{\epsilon_{\mathrm{vol}}} * \epsilon_{t_1} \wedge \epsilon_{t_2} \wedge ... \wedge \epsilon_{t_N}, \tag{3.2}$$

where $*$ denotes the Hodge $*$ operator and ϵ_{vol} is the volume form of $\mathbb{R}^{NL}$ given by

$$e_1 \wedge e_2 \wedge ... \wedge e_{NL}.$$

With this established, we may now write

$$\det(\mathbb{V}(e^{i\theta})) = \sum_{(t_1,...,t_N)} \operatorname{sgn}(t_1,...,t_N) \prod_{j=1}^{N} \det(\mathbb{V}_{t_j}(e^{i\theta_j})).$$

Returning to the partition function,

$$C_N = \frac{1}{N!} \int_{[-\pi,\pi)^N} \sum_{(t_1,...,t_N)} \operatorname{sgn}(t_1,...,t_N) \prod_{j=1}^{N} \det(\mathbb{V}_{t_j}(e^{i\theta_j}))d\mu^N(\theta).$$

Then by using properties of integrals and Fubini's Theorem, we have that

$$C_N = \frac{1}{N!} \sum_{(t_1,...,t_N)} \operatorname{sgn}(t_1,...,t_N) \prod_{j=1}^{N} \int_{[-\pi,\pi)} \det(\mathbb{V}_{t_j}(e^{i\theta_j}))d\mu(\theta)$$

Notice that $\det(\mathbb{V}_{t_j}(e^{i\theta_j}))$ is a Wronskian, so that

$$C_N = \frac{1}{N!} \sum_{(t_1,...,t_N)} \operatorname{sgn}(t_1,...,t_N) \prod_{j=1}^{N} \int_{[-\pi,\pi)} \operatorname{Wr}(P_{t_n}; e^{i\theta})d\mu(\theta).$$

If we multiply this by ϵ_{vol}, we can use (3.2) to rewrite the product of $\operatorname{sgn}(t_1,...,t_N)\epsilon_{\text{vol}}$ as a wedge product of basis vectors.

$$C_N\epsilon_{\text{vol}} = \frac{1}{N!}\left[\sum_{(t_1,...,t_N)} \prod_{j=1}^{N} \int_{[-\pi,\pi)} \operatorname{Wr}(P_{t_n}; e^{i\theta})d\mu(\theta) \right] \operatorname{sgn}(t_1,...,t_N)\epsilon_{\text{vol}}.$$

Now using equation (3.2),

$$C_N\epsilon_{\text{vol}} = \frac{1}{N!} \sum_{(t_1,...,t_N)} \bigwedge_{j=1}^{N} \int_{[-\pi,\pi)} \operatorname{Wr}(P_{t_n}; e^{i\theta})d\mu(\theta)\epsilon_{t_n}.$$

Here, the sum and wedge product commute so that

$$C_N\epsilon_{\text{vol}} = \frac{1}{N!}\left[\sum_{t\in\mathcal{T}} \bigwedge_{j=1}^{N} \int_{[-\pi,\pi)} \operatorname{Wr}(P_{t_n}; e^{i\theta})d\mu(\theta)\epsilon_{t_n} \right]^{\wedge N}.$$

This is precisely

$$\mathrm{PF}\left(\int_{-\pi}^{\pi} \omega(e^{i\theta}) d\mu(\theta) \right),$$

which is the desired result.

□

3.3 Connection to Correlation Functions

Definition. Given an ensemble on W and pairwise disjoint sets $B_1, ..., B_m$, let N_{B_i} denote the number of particles in the set B_i. Then the function $R_m : W^m \to [0, \infty)$ is the m^{th} *correlation function* of the ensemble if

$$\mathbb{E}[N_{B_1} \cdot \cdots \cdot N_{B_m}] = \int_{B_1} \cdots \int_{B_m} R_m(y)\, d\mu^m(y) \ .$$

These correlation functions characterize the ensemble, since R_M is the probability density f. We now use the definition to verify that

$$R_m(y) = \frac{M!}{(M-m)!} \int_{W^{M-m}} f(y_1, ..., y_m, x_1, ..., x_{M-m})\, d\mu^{M-m}(x) \ .$$

Proof. We will show that

$$\mathbb{E}[N_{B_1} \cdot \cdots \cdot N_{B_m}] = \int_{B_1} \cdots \int_{B_m} \frac{M!}{(M-m)!} \int_{W^{M-m}} f(y_1, ..., y_m, x_1, ..., x_{M-m})\, d\mu^{M-m}(x) d\mu^m(y) \ .$$

First, observe that

$$\int_{W^{M-m}} f(x_1, ..., x_{M-m})\, d\mu^{M-m}(x) = 1$$

because it's a probability distribution. We then are left with the marginal distribution for $y_1, ..., y_m$ integrated over the regions $B_1, ..., B_m$,

$$\frac{M!}{(M-m)!} \int_{B_1} \cdots \int_{B_m} f(y_1, ..., y_m)\, d\mu^m(y).$$

But this is just the product of the number of ways to permute our m out of M particles and the probability that each of the coordinates lie in one of these sets. In other words, this expression is exactly $\mathbb{E}[N_{B_1} \cdot \cdots \cdot N_{B_m}]$,

as desired. □

Moving forward, we aim to use these techniques to solve the circular ensemble for β an even square. Our ability to write the partition function as a hyperpfaffian suggests that there is an expression for the correlation functions as hyperpfaffians. These expressions give us algebraic tools and identities to employ that may be the missing step towards solving the ensemble.

Appendix A

More preliminaries

A.1 Set theory, mathematical foundations

A.1.1 Cartesian product

Definition. The *Cartesian product* $X \times Y$ of two sets X and Y as the set of ordered pairs $\{(x, y) | x \in X, y \in Y\}$.

Notice that in this context, X and Y are simply sets, and do not have multiplications defined on them. Therefore, it is completely unreasonable to try to multiply anything in the traditional sense. This solves the "product problem" by isolating the elements of X in one coordinate and the elements of Y in the other, so no multiplication is required.

A.1.2 Equivalence relations

What makes equality special, and how do we generalize these properties? We know that for all numbers a, b, and c, $a = a$, $b = a$ if $a = b$, and $a = c$ if $a = b$ and $b = c$. We define a more general way of relating numbers that captures these properties, and equivalence relation.

Definition. An *equivalence relation* $\mathcal{R}$ on a set X is a set of ordered pairs of elements of the set X that satisfies the following properties.

1. Reflexivity, i.e. if $a \in X$, $(a, a) \in \mathcal{R}$.
2. Symmetry, i.e. if $a, b \in X$ and $(a, b) \in \mathcal{R}$, then $(b, a) \in \mathcal{R}$.
3. Transitivity, i.e. if $a, b, c \in X$ and $(a, b), (b, c) \in \mathcal{R}$, then $(a, c) \in \mathcal{R}$.

In practice, we often write $a \sim b$ to denote $(a, b) \in \mathcal{R}$.

A.2 Linear algebra

A.2.1 Subspaces

Definition.A *subspace* is a special subset of a vector space that inherits the same vector space structure as the original vector space.

In other words, a subspace must contain the zero vector and be closed under addition and scalar multiplication. One important example of a subspace is the kernel of a linear transformation. We invite the reader to verify that this is indeed a subspace.

Subspaces are special in a way that subgroups and subrings are not. Every subspace is the kernel of a linear transformation, whereas every subgroup or subring is not the kernel of a homomorphism. In group theory and ring theory, we need some additional constraints on our subgroups and subrings for this to be true. This property is particularly useful in defining quotients, and readers familiar with abstract algebra will appreciate the simplicity of the definition for vector spaces in comparison with those for groups or rings.

A.2.2 Trace

Perhaps the simplest interesting function of a matrix is its trace.

Definition. The *trace* of a matrix is the sum of its diagonal entries.

It is unsurprising that this is a linear map, but the following relationship between the trace of a matrix and its spectrum is less immediately clear. The trace of a matrix is equal to the sum of its eigenvalues. This is easier to see in the 2×2 case, where the middle coefficient of the characteristic polynomial is equal to the trace. Consequently, the spectrum of a matrix is invariant under a change of basis.

A.3 Abstract algebra

A.3.1 Groups

Definition. A *group* G is a set X together with a binary operation on it, often denoted $+$ or $\times$, that satisfies some special algebraic properties. A binary operation is simply a function $X \times X \to X$, where $X \times X$ denotes the Cartesian product of X with itself. We can think of groups as a collection of items with a rule for how to combine them. However, not any collection of items and any rule will do. Here are the special properties that a group must satisfy. We will let $a, b, c \in G$ and $+$ be our binary operation.

1. Associativity, i.e. $(a + b) + c = a + (b + c)$.

2. Identity element, i.e. there is some $e \in G$ such that $e + a = a$ and $a + e = a$.

3. Inverse elements, i.e. there is some $g \in G$ such that $a + g = e$ and $g + a = e$. We often denote g by $-a$, or if using the binary operation $\times$, by a^{-1}.

An important group in this work is the symmetric group on n elements, S_n. This is the group of permutations of a set with n elements. For simplicity, let's say the elements of this set are natural numbers. Then S_3, for example, acts on the set $\{1, 2, 3\}$. There are six different ways that we can permute these numbers. We can do nothing, swap 1 and 2, swap 2 and 3, swap 1 and 3, shift the position of each number to the right, and switch the position of each number to the left. These are the elements of S_3.

A.3.2 Rings

Definition. A *ring* is a set X together with two binary operations, often denoted $+$ and $\times$ such that $(R, +)$ is a group and $(R, \times)$ is associative and has an identity element.

Bibliography

[1] M. Adler, P. J. Forrester, T. Nagao, and P. van Moerbeke. Classical skew orthogonal polynomials and random matrices. *Journal on Statistical Physics*, 99(1-2):141–170, 2000.

[2] Richard Askey. Some basic hypergoemetric extensions of integrals of selberg and andrews. *SIAM Journal on Mathematical Analysis*, 11(6):938–951, 1980.

[3] Alexander I. Barvinok. New algorithms for linear k-matriod intersection and matroid k-parity problems. *Mathematical Programming*, 69(3):449–470, 1995.

[4] Alexei Borodin. Determinantal point processes. *Oxford Handbook of Random Matrix Theory*, 2010.

[5] Alexei Borodin and Christopher D. Sinclair. The ginibre ensemble of real random matrices and its scaling limits. *Communications in Mathematical Physics*, 291(1):177–224, 2009.

[6] N. G. de Brujin. On some multiple integrals involving determinants. *Journal of Indian Mathematical Society*, 19:133–151, 1955.

[7] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. *Journal of Mathematical Physics*, 3:140–156, 1962.

[8] Peter J. Forrester and S. Ole Warnaar. The importance of the selberg integral. *Bulletin of the American Mathematical Society*, 45(4):489–534, 2008.

[9] I. J. Good. Short proof of a conjecture by dyson. *Journal of Mahtematical Physics*, 11(6):1884–1884, 1970.

[10] J. Gunson. Proof of a conjecture by dyson in the statistical theory of energy levels. *Journal of Mahtematical Physics*, 3(4):752–753, 1962.

[11] Howard E. Haber. Notes on antisymmetric matrices and the pfaffian. Santa Cruz Institute for Particle Physics, 2015.

[12] J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Balint Virag. Zeros of gaussian analytic functions and determinantal point processes. *University Lecture Series, American Mathematical Society*, 51, 2009.

[13] Masao Ishikawa, Soichi Okada, HIroyuki Tagawa, and Jiang Zeng. Generalizations of cauchy's determinant and schur's pfaffian. *Advances in Applied Mathematics*, 36(3):251–287, 2006.

[14] Jean-Gabriel Luque and Jean-Yves Thibon. Pfaffian and hafnian identities in shuffle algebras. *Advances in Applied Mathematics*, 29(4):620–646, 2002.

[15] G. Mahoux and M. L. Mehta. A method of integration over matrix variables. *Journal of Physique I*, 1(8):1093–1108, 1991.

[16] Madan Lal Mehta. *Random matrices*. Pure and Applied Mathematics. Elsevier/Academic Press, Amsterdam, third edition edition, 2004.

[17] Madan Lal Mehta and Freeman J. Dyson. Statistical theory of the energy levels of complex systems. *Journal of Mathematical Physics*, 4:713–719, 1963.

[18] Eric M. Rains. Correlation functions for symmetrized increasing subsequences. 2000.

[19] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 3rd edition edition, 1976.

[20] Atle Selberg. Bemerkninger om et multipelt integral. *Norsk Mat. Tidsskr.*, 26:71–78, 1944.

[21] Christopher D. Sinclair. Correlation functions for $\beta = 1$ and n odd. *Journal of Statistical PHysics*, 136(1):17–33, 2009.

[22] Christopher D. Sinclair. Ensemble averages when β is a square integer. *Monatsh. Math*, 166(1), 2018.

[23] Terence Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.

[24] Craig A. Tracy and Harold Widom. Correlation functions, cluster functions, and spacing distributions for random matrices. *Journal of Statistical Physics*, 92(5-6):809–835, 1998.

[25] J. M. Wells. *On the solvability of β ensembles when β is a square integer*. PhD thesis, University of Oregon, 2019.

[26] Elisha D. Wolff. *Ensemble averages of assorted log-gas models*. PhD thesis, University of Oregon, 2022.

[27] Doron Zeilberger. A combinatorial proof of dyson's conjecture. *Discrete Mathematics*, 41(3):317–321, 1982.