Lectures on Random Matrix Theory

Govind Menon

April 14, 2015

Contents

Chapter 1

Overview

1.1 What is a random matrix?

There are two distinct points of view that one may adopt. On one hand, our intuitive ideas of randomness are intimately tied to the notion of sampling a realization of a random variable. Thus, given a random number generator, one may build a random Hermitian matrix, $M \in \mathsf{Her}(n)$, by choosing its real diagonal and complex upper-triangular entries independently at random. It is conventional to assume further that all the diagonal entries have the same law, and that all the upper-triangular entries have the same law. For example, we may assume that the entries on the diagonal are ± 1 with probability $1/2$, and that the upper-triangular entries are $\pm 1 + i \pm 1$ with probability $1/4$. Random matrices of this kind, are said to be drawn from *Wigner ensembles*.

On the other hand, one may adopt a more analytic view. The Hilbert-Schmidt inner product of two Hermitian matrices, $\mathrm{Tr}(M^2) = \sum_{j,k=1}^{n} |M_{jk}|^2$, defines a natural metric $\mathrm{Tr}(dM^2)$ and volume form dM on $\mathsf{Her}(n)$ (see Lecture **??**). Thus, each positive function $p : \mathsf{Her}(n) \to [0, \infty)$ that decays sufficiently fast as $\|M\| \to \infty$, may be normalized to define a probability measure. A fundamental example is the law of the *Gaussian Unitary Ensemble* (GUE)

$$p_{\mathrm{GUE}}(M)dM = \frac{1}{Z_n} e^{-\frac{1}{2}\mathrm{Tr}(M^2)} dM. \tag{1.1.1}$$

Here Z_n is a normalization constant that ensures p_{GUE} is a probability density (we use the same notation for different ensembles; thus the numerical value of Z_n must be inferred from the context). The term GUE was introduced by Freeman Dyson [5], and refers to an important invariance property of p_{GUE}. Each $U \in \mathsf{U}(n)$ defines a transformation $\mathsf{Her}(n) \to \mathsf{Her}(n)$, $M \mapsto UMU^*$. It is easily checked that the volume form dM is invariant under the map $M \mapsto UMU^*$, as is the measure $p_{\mathrm{GUE}}(M)dM$. More generally, a probability measure on $\mathsf{Her}(n)$ is said to be *invariant* if $p(M)\,dM$ remains invariant under the map $M \mapsto UMU^*$, for each $U \in \mathsf{U}(n)$. Important examples of invariant ensembles

are defined by polynomials in one-variable of the form

$$V(x) = a_{2N}x^{2N} + a_{2N-1}x^{2N-1} + \ldots + a_0, \quad a_j \in R, j = 0, 1, \ldots, 2N, \ a_{2N} > 0. \tag{1.1.2}$$

Then the following probability measure is invariant

$$p(M)dM = \frac{1}{Z_n}e^{-\mathrm{Tr}V(M)}. \tag{1.1.3}$$

More generally, any potential $V : \mathbb{R} \to \mathbb{R}$ such that $V(x) \to \infty$, sufficiently rapidly as $|x| \to \infty$ defines an invariant ensemble.

We have assumed that all matrices are Hermitian simply to be concrete. The above notions extend to ensembles of matrices from $\mathsf{Symm}(n)$ and $\mathsf{Quart}(n)$. The notion of invariance in each case is distinct: for $\mathsf{Symm}(n)$, the natural transformation is $M \mapsto QMQ^T$, $Q \in \mathsf{O}(n)$; for $\mathsf{Quart}(n)$ it is $M \mapsto SMS^D$, $S \in \mathsf{Sp}(2n, \mathbb{R})$. The standard Gaussian ensembles in these cases are termed GOE (the *Gaussian Orthogonal Ensemble*) and GSE (the *Gaussian Symplectic Ensemble*), and they are normalized as follows:

$$p_{\mathrm{GOE}}(M) = \frac{1}{Z_n}e^{-\frac{1}{4}\mathrm{Tr}(M^2)}dM, \quad p_{\mathrm{GSE}}(M) = \frac{1}{Z_n}e^{-\mathrm{Tr}(M^2)}dM. \tag{1.1.4}$$

The differing normalizations arise from the different volume forms on $\mathsf{Symm}(n)$, $\mathsf{Her}(n)$ and $\mathsf{Quart}(n)$ as will be explained in Lecture **??**. For now, let us note that the densities for all the Gaussian ensembles may be written in the unified form

$$Z_n(\beta)^{-1}e^{-\frac{\beta}{4}\mathrm{Tr}(M^2)} \tag{1.1.5}$$

where $\beta = 1, 2$ and 4 for GOE, GUE and GSE respectively. While it is true that there are no other ensembles that respect fundamental physical invariance (in the sense of Dyson), many fundamental results of random matrix theory can be established for all $\beta > 0$. These results follow from the existence of ensembles of *tridiagonal* matrices, whose eigenvalues have a joint distribution that interpolates those of the $\beta = 1, 2$ and 4 ensembles to all $\beta > 0$ [4].

1.2 The main limit theorems

The basic question in random matrix theory is the following: what can one say about the statistics of the eigenvalues of a random matrix? For example, what is the probability that the largest eigenvalue lies below a threshold? Or, what is the probability that there are no eigenvalues in a given interval? The difficulty here is that even if the entries of a random matrix are independent, the eigenvalues are strongly coupled.

Gaussian ensembles play a very special, and important, role in random matrix theory. These are the only ensembles that are both Wigner and invariant (see Theorem **??** below). Pioneering, ingenious calculations by Dyson [5], Gaudin and Mehta [17, 16], on the Gaussian ensembles served to elucidate the

fundamental limit theorems of random matrix theory. In this section we outline these theorems, assuming always that the ensemble is GUE. Our purpose is to explain the form of the main questions (and their answers) in the simplest setting. All the results hold in far greater generality as is briefly outlined at the end of this section.

By the normalization (1.1.1), a GUE matrix has independent standard normal entries on its diagonal (mean zero, variance 1). The off-diagonal entries have mean zero and variance $1/2$. We denote the ordered eigenvalues of the GUE matrix by $\lambda_1 \leq \lambda_2 \leq \ldots \lambda_n$. A fundamental heuristic for GUE matrices (that will be proven later, and may be easily simulated) is that the largest and smallest eigenvalues have size $O(\sqrt{n})$. In fact, $\lambda_1 \approx -2\sqrt{n}$ and $\lambda_n \approx 2\sqrt{n}$ as $n \to \infty$. Since there are n eigenvalues, the gap between these eigenvalues is typically $O(1/\sqrt{n})$. There are thus two natural scaling limits to consider as $n \to \infty$:

1. Rescale $M \mapsto n^{-1/2}M$ so that the spectral radius is $O(1)$. In this scaling limit, n eigenvalues are contained within a bounded interval, and we obtain a deterministic limit called the *semicircle law*.

2. Rescale $M \mapsto n^{1/2}M$ so that the gaps between eigenvalues are $O(1)$. In this scaling limit, we expect a random limiting point process. The limiting point process is a *determinantal point process* called the Sine_2 process.

In fact, the situation is more subtle. While the expected value of the gap between eigenvalues for a GUE matrix is indeed $O(1/n)$, the gaps are $O(n^{-2/3})$ about the edge of the spectrum. There is an an entirely different scaling limit called the Airy_2 process obtained by rescaling the spectrum of $M \pm 2\sqrt{n}I$.

In all that follows, we consider a sequence of random matrices of size n sampled from GUE(n). To make this explicit, the matrix is denoted M_n, and its ordered eigenvalues are denoted $\lambda_n^1 \leq \lambda_2^n, \ldots, \lambda_n^n$.

1.2.1 The semicircle law

Definition 1. The probability density and distribution function

$$p_{\mathrm{sc}}(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbf{1}_{|x|\leq 2}, \quad F_{\mathrm{sc}}(x) = \int_{-\infty}^{x} p_{\mathrm{sc}}(x'),\, dx' \qquad (1.2.1)$$

are called the *semicircle density* and the *semicircle distribution* respectively.

Theorem 2. *Let M_n be a sequence of GUE matrices of size n. The rescaled empirical spectral measures*

$$\mu_n(dx) = \frac{1}{n}\sum_{j=1}^{n} \delta_{n^{-1/2}\lambda_j^n}(dx) \qquad (1.2.2)$$

converge weakly to the semicircle density almost surely.

Theorem 2 may also be interpreted as the statement that the empirical spectral distribution of the matrices $M_n/\sqrt{n}$ converges to the semicircle distribution. The shortest proof of Theorem (2) uses the following integral transform.

Definition 3. Assume μ is a measure on $\mathbb{R}$ that satisfies the finiteness condition

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\mu(dx) < \infty. \tag{1.2.3}$$

The Stieltjes transform of μ is the function

$$\hat{\mu}(z) = \int_{-\infty}^{\infty} \frac{1}{x-z}\mu(dx), \quad z \in \mathbb{C}\backslash\mathbb{R}. \tag{1.2.4}$$

The Stieltjes transform is of fundamental importance in the theory of orthogonal polynomials and spectral theory. This is because there are natural Stieltjes transforms associated to the resolvent $(M-z)^{-1}$, such as

$$\text{Tr}(M-z)^{-1}, \quad \text{and} \langle v^*(M-z)^{-1}v, \quad v \in \mathbb{C}^n, |v| = 1. \tag{1.2.5}$$

The proof of Theorem 2 uses a recursive expression for the law of $\text{Tr}(z-M_n)^{-1}$. As $n \to \infty$, the fixed point of this recursion, R_{sc} solves the quadratic equation

$$R^2 - zR + 1 = 0. \tag{1.2.6}$$

It is then easy to verify that

$$R_{\text{sc}}(z) = \frac{1}{2}\left(-z + \sqrt{z^2-4}\right), \quad z \in \mathbb{C}\backslash[-2,2]. \tag{1.2.7}$$

We recover the semicircle law from $R_{\text{sc}}(z)$ by evaluating the jump in $\text{Im}(R_{\text{sc}}(z))$ across the branch cut $[-2,2]$.

1.2.2   Fluctuations in the bulk: the sine process

We now rescale so that the gaps between eigenvalues is $O(1)$, and the scaling limit is a random process. This random process will always be denoted Sine_2 (and Sine_β for the general β-ensembles). Each realization of the Sine_2 process is a countable set of points $\{x_k\}_{k=-\infty}^{\infty}$. One of the fundamental statistics associated to a point process is the probability of having k points in an interval. In order to state a typical fluctuation theorem that describes these probabilities, we must define the sine-kernel and its Fredholm determinants.

Definition 4. The sine-kernel is the integral kernel on $\mathbb{R} \times \mathbb{R}$ given by

$$K_{\text{sine}}(x,y) = \frac{\sin\pi(x-y)}{\pi(x-y)}, \quad x \neq y, \tag{1.2.8}$$

and $K_{\text{sine}}(x,x,) = 1$.

In the following theorem we will assume that x and y are restricted to a finite interval (a, b). The sine-kernel defines an integral operator on $L^2(a, b)$ that we denote by $K_{\mathsf{sine}} \mathbf{1}_{(a,b)}$. The kernel $K_{\mathsf{sine}}(x, y)$ is clearly continuous, thus bounded, for $x, y \in (a, b)$. Thus, $K_{\mathsf{sine}} \mathbf{1}_{(a,b)}$ defines an integral operator on $L^2(a, b)$ that is trace-class, and it has a well-defined Fredholm determinant

$$\det\left(1 - K_{\mathsf{sine}} \mathbf{1}_{(a,b)}\right) = 1 + \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \int_{(a,b)^m} \det\left(K_{\mathsf{sine}}(x_j, x_k)_{1 \le j, k \le m}\right) dx_1 dx_2 \ldots dx_m.$$
(1.2.9)

Though perhaps mysterious at first sight, the origin of this formula is rather simple. Integral operators with some smoothness and boundedness (in particular, continuous integral operators K whose trace $\int_a^b |K(x, x)| dx$ is finite) may be approximated on a discrete-grid of size h by a finite-dimensional discretization K_h. The determinant $(I - K_h)$ is then the usual determinant of a matrix and we may use the definition of the determinant to expand $\det(I - K_h)$ in a finite series, which is nothing but the infinite series above in the instance when all terms beyond $m = \mathrm{rank}(K_h)$ vanish. This approach was pioneered by Fredholm in 1900 before the development of functional analysis.

Theorem 5 (Gaudin-Mehta [17]). *For each finite interval $(a, b) \subset \mathbb{R}$,*

$$\lim_{n \to \infty} \mathbb{P}\left(\sqrt{n} \lambda_k^n \notin (a, b), \ 1 \le k \le n\right) = \det\left(1 - K_{\mathsf{sine}} \mathbf{1}_{(a,b)}\right).$$
(1.2.10)

The probabilities of the Sine_2 process can be expressed without reference to the matrices M_n. For each interval (a, b) let $N_{(a,b)} = \sum_{k=-\infty}^{\infty} \mathbf{1}_{x_k \in (a,b)}$. Then,

$$\mathbb{P}\left(N_{(a,b)} = 0\right) = \det\left(1 - K_{\mathsf{sine}} \mathbf{1}_{(a,b)}\right).$$
(1.2.11)

For comparison, if we had a Poisson process $\{\tilde{x}_k\}_{k=-\infty}^{\infty}$ with rate $\lambda(dx)$, the associated count $\tilde{N}_{(a,b)}$ would satisfy

$$\mathbb{P}\left(\tilde{N}_{(a,b)} = 0\right) = 1 - \exp\left(\int_a^b \lambda(dx)\right).$$

1.2.3 Fluctuations at the edge: the Airy point process

Theorem 5 reveals that the gaps between consecutive eigenvalues λ_j^n and λ_{j+1}^n is of $O(1)$. However, the fluctuations at the edge are much smaller, of $O(n^{-2/3})$. The point process of shifted and scaled eigenvalues converges in distribution to a limiting point process, $\{y_k\}_{k=1}^{\infty}$ called the Airy_2 process. In order to describe the law of this process, we must define the Airy function and the Airy kernel.

Definition 6. The Airy function is defined by the oscillatory integral

$$\mathrm{Ai}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} e^{-ik^3/3} \, dk.$$
(1.2.12)

The Airy function is one of the classical special functions [1]. It admits several alternative definitions. For instance, the oscillatory integral in (1.2.12) may be deformed into an absolutely convergent integral in the complex plane. This argument allows us to establish that the Airy function is entire and to determine its asymptotic expansions as $x \to \pm\infty$.

These properties may also be established using the theory of ordinary differential equations in the complex plane [10]. It is easy to verify from (1.2.12) that $\mathrm{Ai}(x)$ satisfies the differential equation

$$\varphi''(x) = x\varphi, \quad -\infty < x < \infty. \tag{1.2.13}$$

Equation (1.2.13) admits two linearly independent solutions, only one of which decays as $x \to \infty$. Upto a (fixed by convention, but otherwise arbitrary) normalization constant, the decaying solution is $\mathrm{Ai}(x)$.

Definition 7. The Airy kernel is the continous integral kernel on $\mathbb{R} \times \mathbb{R}$ given by

$$K_{\mathsf{Airy}}(x,y) = \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y}, \quad x \neq y,$$

and by continuity at $x = y$.

Observe that both the sine and Airy kernel have the form

$$K(x,y) = \frac{f(x)f'(y) - f'(x)f(y)}{x-y}, \quad x \neq y \tag{1.2.14}$$

where f solves a second-order linear differential equation. Similar kernels arise in various limiting models in random matrix theory. For instance, the Bessel kernel – corresponding to $f(x) = J_\alpha(x)$, the Bessel function with parameter α – describes fluctuations about the *singular values* of random positive definite Hermitian matrices.

Theorem 8. *For each interval $(a,b) \subset \mathbb{R}$, $-\infty < a < b \leq \infty$,*

$$\lim_{n\to\infty} \mathbb{P}\left(n^{2/3}\left(\frac{\lambda_k^n}{\sqrt{n}} - 2 \right) \notin (a,b),\, 1 \leq k \leq n \right) = \det\left(1 - K_{\mathsf{Airy}}\mathbf{1}_{(a,b)} \right). \tag{1.2.15}$$

As in the remarks following Theorem 5, the expression $\det\left(1 - K_{\mathsf{Airy}}\mathbf{1}_{(a,b)} \right)$ gives the probability that no points of a realization of the Airy_2 point process lie in (a,b).

1.2.4   Fredholm determinants, Painlevé equations, and integrable systems

It is immediate from Theorem 5 and Theorem 8 that the Fredholm determinants $\det\left(1 - K_{\mathsf{sine}}\mathbf{1}_{(a,b)} \right)$ and $\det\left(1 - K_{\mathsf{Airy}}\mathbf{1}_{(a,b)} \right)$ are positive for all (a,b). This is astonishing, if one treats (1.2.9) as a starting point, since it is by no means clear that the signed infinite series sums to a positive number! It is in fact, rather

challenging to extract meaningful information, such as the asymptotics of tails, from the expression of probabilities as Fredholm determinants. A crucial piece of the puzzle lies in the connection between Fredholm determinants and the theory of integrable systems. More precisely, the Fredholm determinants satisfy differential equations in a and b (or more generally in endpoints of intervals, when one considers the obvious extensions of Theorem 5 and Theorem 8 to a collection of intervals $\prod_{j=1}^{m}(a_m, b_m)$). These ordinary differential equations have a special, integrable structure, that allows their analysis. The following theorems illustrate this aspect of random matrix theory.

Theorem 9 (Jimbo-Miwa-Mori-Sato [11]). *For all $t > 0$,*

$$\det\left(1 - K_{\mathsf{sine}}\mathbf{1}_{\left(-\frac{t}{2},\frac{t}{2}\right)}\right) = \exp\left(\int_0^t \frac{\sigma(s)}{s}\, ds\right), \qquad (1.2.16)$$

where $\sigma(t)$ is the solution to the Painlevé-5 equation

$$\left(t\sigma''\right)^2 + 4\left(t\sigma' - \sigma\right)\left(t\sigma' - \sigma + \sigma^2\right) = 0, \qquad (1.2.17)$$

which satisfies the asymptotic condition

$$\sigma(t) = -\frac{t}{\pi} - \frac{t^2}{\pi^2} - \frac{t^3}{\pi^3}, \quad t \downarrow 0. \qquad (1.2.18)$$

Theorem 10 (Tracy-Widom distribution [24]). *For all real t,*

$$F_2(t) := \det\left(1 - K_{\mathsf{Airy}}\mathbf{1}_{(t,\infty)}\right) = \exp\left(-\int_t^\infty (s-t)q^2(s)\, ds\right), \qquad (1.2.19)$$

where q is the solution to the Painlevé-2 equation

$$q'' = tq + 2q^3, -\infty < t < \infty \qquad (1.2.20)$$

which satisfies the asymptotic condition

$$q(t) \sim \mathrm{Ai}(t), \quad t \to \infty. \qquad (1.2.21)$$

We will discuss the basic properties of Painlevé equations and integrable systems in Lecture **??**. Here is a brief preview.

The Painlevé differential equations are a special family of *nonlinear* ordinary differential equations that generalize the classical theory of *linear* differential equations in the complex plane and the associated theory of special functions [10]. For example, the Painlevé-2 equation (1.2.20) may be viewed as a nonlinear analogue of the Airy differential equation (1.2.13). The theory of Painlevé equations was developed in the early years 1900's, by Boutroux and Painlevé, but fell into obscurity [1] . It was reborn in the 1970s with the discovery

---

[1]Paul Painlevé was rather restless: he began in mathematics, became an early aviation enthusiast, and then turned to politics. He rose to become the Prime Minister of France for part of World War I, and was later the designer of the disastrous Maginot line.

of their importance in integrable systems and exactly solvable models in statistical mechanics, such as the Ising model in 2D [15]. We illustrate these links with a fundamental integrable system: the Korteweg-De Vries (KdV) equation

$$u_t + 6uu_x + u_{xxx} = 0, \quad -\infty < x < \infty, \ t \geq 0. \tag{1.2.22}$$

Despite the fact that KdV is nonlinear, it may be solved explicitly through the inverse scattering transform. We will not discuss this method in detail. But in order to make the connection with random matrix theory, let us note that if one seeks self-similar solutions to KdV of the form

$$u(x,t) = \frac{1}{(3t)^{2/3}} q\left(\frac{x}{(3t)^{2/3}}\right) \tag{1.2.23}$$

then q satisfies the Painlevé-2 equation (1.2.20). In fact, it is in this context that Hastings and McLeod established the existence of a solution to (1.2.20) that satisfies the asymptotic condition (1.2.21) [9]. It is remarkable that it is exactly this solution that describes the Tracy-Widom distribution $F_2(t)$!

1.2.5   Universality

We have restricted attention to matrices from GUE to present some of the central theorems in the subject in an efficient manner. One of the main achievements of the past decade has been the establishment of *universality* – informally, this is the notion that the limiting fluctuations in the bulk and edge described by the Sine_2 and Airy_2 processes, hold for both Wigner and invariant ensembles which satisfy natural moment assumptions. The idea of universality is of clear practical importance (we need understand only a few universal limits). It also appears to hold the key to some of the connections between random matrix theory and other areas of mathematics. The explanation of these connections may lie in the fact that determinantal point processes, such as the Sine_2 and Airy_2 process, have the simplest structure of *strongly interacting* point processes. By contrast, Poisson processes, while universal, describe *non-interacting* points.

1.3   Connections to other areas of mathematics

Random matrix theory has deep connections with many areas of mathematics, many of which are still poorly understood. A brief overview of some of these connections is presented below. While some of these notions, such as the connections with stochastic PDE require more background than we assume, some other connections (e.g. with quantum gravity) are in fact more elementary (and fundamental) than one may naively expect. Our purpose here is to present a small sample of the rich set of ideas that make the subject so attractive.

1.3.1 Number theory

The Riemann zeta function is defined by the infinite sum

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathsf{Re}(s) > 1. \tag{1.3.1}$$

The function $\zeta(s)$ is central to number theory, since it provides a generating function for the distribution of the prime numbers via Euler's product formula

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}, \quad \mathsf{Re}(s) > 1. \tag{1.3.2}$$

For instance, the divergence of the harmonic series at $s = 1$ provides a proof that there are infinitely many primes. The study of $\zeta(s)$ by complex analysis is the cornerstone of analytic number theory. The basic facts are as follows. The function $\zeta(z)$ extends to a meromorphic function on $\mathbb{C}$ by analytic continuation, which has a simple pole at $s = 1$ where the residue is 1. A closely related function is the Riemann ξ-function

$$\xi(s) = \frac{1}{2\pi^{s/2}} s(s-1) \Gamma\left(\frac{s}{2}\right) \zeta(s). \tag{1.3.3}$$

Recall that the Γ function is a 'continuous interpolation' of the factorial, defined by the integral

$$\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1}\, dx, \quad \mathsf{Re}(s) > 0. \tag{1.3.4}$$

The Γ-function extends to a meromorphic function $\mathbb{C}$, which has simple poles at $\ldots, -2, -1, 0$ where the residue is 1. These poles cancel the 'trivial' zeros of the ζ function, and the essential difficulties related to the study of the ζ function are more transparent for the ξ function. It satisfies the functional equation

$$\xi(s) = \xi(1 - s), \quad s \in \mathbb{C} \tag{1.3.5}$$

The celebrated *Riemann Hypothesis* is the conjecture that all zeros of the ξ function lie on the critical line $\mathsf{Re}(s) = 1/2$ (this line is the symmetry axis for the functional equation above). In his fundamental paper on the distribution of prime numbers (translated in [7] and [19]) Riemann presented a series of asymptotic expansions that would imply rigorous bounds on the distribution of primes if the Riemann Hypothesis is true.

The connection between random matrix theory and the Riemann Hypothesis is two-fold. First, if one could construct a Hermitian operator with point spectrum whose eigenvalues coincide with the zeros of $\xi(i(s - 1/2))$ then the Riemann Hypothesis would follow immediately (since all eigenvalues of a Hermitian operator are real). The catch, of course, is to determine such an operator. Nevertheless, as we discuss below, random matrix theory has shed new lie on the spectral theory of several operators, deterministic and random. Thus, the

theory provides a catalog of 'guesses'. Second, if one assumes the Riemann hypothesis, the *fluctuations* in the zeros of $\zeta(s)$ are described by the sine-kernel! Under the Riemann hypothesis, the non-trivial zeros of $\zeta(s)$ may be written $\gamma_n = \frac{1}{2} \pm it_n$ with $0 < t_1 < t_2 < \dots$. Let

$$w_n = \frac{t_n}{2\pi} \log\left(\frac{t_n}{2\pi}\right), \quad \text{and} \quad N(x) = \sum_{k=1}^{\infty} \mathbf{1}_{w_n \leq x}. \qquad (1.3.6)$$

This rescaling is chosen so that $\lim_{x\to\infty} N(x)/x = 1$ in accordance with the Prime Number Theorem.

Despite the fact that the zeros w_n are deterministic, we may introduce probabilistic notions by counting the (rescaled) zeros upto a level $x > 0$. For example, we may define the empirical probability measure

$$\mu_1(dw; x) = \frac{1}{N(x)} \sum_{k=1}^{N(x)} \delta_{w_k}(dw). \qquad (1.3.7)$$

In order to study the gaps between eigenvalues, we must consider instead the empirical measures

$$\mu_2(dl; x) = \frac{1}{x} \sum_{1 \leq j,k \leq N(x); j \neq k} \delta_{w_j - w_k}(dl). \qquad (1.3.8)$$

The expectation of a continuous function with respect to $\mu_2(dl; x)$ is denoted

$$R_2(f; x) = \int_{-\infty}^{\infty} f(l)\mu_2(dl; x) = \frac{1}{x} \sum_{1 \leq j < k \leq N(x)} f(w_j - w_k). \qquad (1.3.9)$$

Under the assumption that f is band-limited, i.e. that its Fourier transform has compact support, Montgomery established the following

Theorem 11 (Montgomery). *Assume the Riemann Hypothesis. Assume f is a Schwartz function whose Fourier transform $\hat{f}$ is supported in $[-1, 1]$. Then*

$$\lim_{x\to\infty} R_2(f; x) = \int_{-\infty}^{\infty} f(l)\mu_2(dl), \quad \mu_2(dl) = \left(1 - \left(\frac{\sin \pi l}{\pi l}\right)^2\right) dl. \qquad (1.3.10)$$

The point here is that the right hand side of (1.3.10) is precisely the 2-point function for the sine-process. More generally, Montgomery's theorem is now known to hold for the distribution of n-consecutive gaps. That is, the rescaled fluctuations converge to the Sine_2 process in distribution. Bourgade's thesis is an excellent review of the state of the art [2].

1.3.2   Combinatorics and enumerative geometry

We will present two problems of enumerative combinatorics that connect with random matrix theory. As a first example, we note that the $2m$-th moment of

the semicircle law

$$\int_{-2}^{2} x^{2m} p_{\text{sc}}(x)\, dx = \frac{1}{m+1}\binom{2m}{m} = C_m, \tag{1.3.11}$$

the m-the Catalan number. An analytic proof of this identity follows from a comparison between the Stieltjes transform $R_{\text{sc}}(z)$, and the generating function

$$\hat{C}(x) = \sum_{m \ge 0} C_m x^m = \frac{1 - \sqrt{1 - 4x}}{x}. \tag{1.3.12}$$

The Catalan numbers describe the solution to many combinatorial problems [2]. For example, C_m enumerates the number of *Bernoulli excursions* or *Dyck paths* of length $2m$: these are walks S_k, $1 \le k \le 2m$ such that $S_0 = S_{2m} = 0$, $S_k \ge 0$, $0 \le k \le 2m$, and $|S_{k+1} - S_k| = 1$.

A deeper set of connections between integrals on $\mathsf{Her}(n)$ and geometry was first noticed by the physicist 't Hooft [23]. Ignoring (for now), physicists' motivation, let us illustrate a particular computational technique that underlies their work. Consider a matrix integral of the form

$$Z_n(z) = \int_{\mathsf{Her}(n)} e^{\text{Tr}(-zM^4)} p_{\text{GUE}}(M)\, dM, \quad \mathsf{Re}(z) > 0. \tag{1.3.13}$$

The quartic nonlinearity prevents us from expressing this integral in closed form. Nevertheless, this integral may be expanded in a Taylor series

$$Z_n(z) = \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} \int \left(\text{Tr}(M^4)\right)^k p_{\text{GUE}}(M)\, dM, \quad \mathsf{Re}(z) > 0. \tag{1.3.14}$$

A fundamental lemma on Gaussian integrals (on $\mathbb{R}^N$) (Wick's lemma) allows us to reduce each integral above to a sum over pairings of indices. It is convenient to keep track of these pairings with a graphical description, called a *Feynman diagram*. 't Hooft observed that when $\mathbb{R}^N \equiv \mathsf{Her}(n)$ the Feynman diagram associated to each term in (1.3.14) enumerates embedded graphs on a Riemann surface. This characterization was independently discovered by mathematicians.

Lemma 1 (Harer-Zagier [8]). *Let $\varepsilon_g(m)$ denote the number of ways to pair the edges of a symmetric $2m$-gon to form an orientable surface with genus g. Then*

$$f(m, n) = \sum_{g=0}^{\infty} \varepsilon_g(m) n^{m+1-2g} = \int_{\mathsf{Her}(n)} \text{Tr}(M^{2m}) p_{\text{GUE}}(M)\, dM. \tag{1.3.15}$$

Note that only finitely many terms in the sum are non-zero. The series above is an instance of a *genus*-expansion. It illustrates the beautiful fact that matrix integrals serve as the generating functions for Riemann surfaces with a given combinatorial decomposition!

---

[2]Stanley lists 66 examples in [21, Exercise 6.19].

1.3.3   Spectral and inverse spectral theory of operators

While Theorem 2–Theorem 8 associate limits to the spectrum of the operators M_n, it is natural to ask if there are limiting operators that may be naturally associated to the limiting spectra. Thus, for Theorem 2 we ask for a 'natural' operator that has spectral density given by the semicircle law, p_{sc}, and for Theorem 5 and Theorem 8 we seek 'natural' random operators that have pure point spectra with the law of the Sine_2 and Airy_2 point processes. What is a 'natural' operator is, of course, a subjective idea, but convincing candidates operators are suggested by inverse spectral theory.

We say that a matrix $T \in \mathsf{Symm}(n)$ is a Jacobi matrix if all its off-diagonal entries are strictly positive. The spectral measure of a Jacobi matrix is the measure whose Stieltjes transform is $e_1^T(T-z)^{-1}e_1$. There is a $1-1$ correspondence between the space of $n \times n$ Jacobi matrices and probability measures on the line with n atoms. This correspondence extends naturally to semi-infinite Jacobi matrices. The essence of this theory (due to Stieltjes) is that the entries of T may be determined from the *continued fraction expansion* of $e_1^T(T-z)^{-1}e_1$. This correspondence will be considered in detail in Lecture **??**, but here is a concrete example. By applying Stieltjes' procedure to the semicircle law, we discover that $p_{sc}(x)$ is the spectral density for the seminfinite tridiagonal matrix that is 1 on the off-diagonal, and 0 in all other entries. This follows from the continued fraction expansion

$$R_{sc}(-z) = \cfrac{1}{z - \cfrac{1}{z - \cfrac{1}{z - \ldots}}} \tag{1.3.16}$$

Ensembles of tridiagonal matrices are of practical important in numerical linear algebra. For instance, a key pre-processing step while solving symmetric linear systems is to transform the matrix to tridiagonal form by Householder's procedure (Lecture 3). Dumitriu and Edelman pushed forward the Gaussian measures under this procedure to obtain a family of tridiagonal ensembles, known as the general-β ensembles [4]. Further, Edelman and Sutton made a formal expansion of these operators, and observed that as $n \to \infty$, the tridiagonal operators appeared to converge to the *stochastic Airy operator* [6]:

$$H_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}}\dot{b}, \quad 0 < x < \infty \tag{1.3.17}$$

with Dirichlet boundary conditions at $x = 0$. Here $\dot{b}$ denotes (formally) white noise (it is not hard to define H_β rigorously).

Theorem 12 (Ramirez-Rider-Virag [18]). *The spectrum $\sigma(H_\beta)$ of the operator H_β is almost surely a countably infinite number of eigenvalues $\mu_1 < \mu_2 < \mu_3 < \ldots$. Moreover, $\sigma(H_\beta)$ has the same law as the Airy_β point process.*

In particular, for $\beta = 2$, the spectrum of the stochastic Airy operator describes the limiting fluctuations at the edge of the spectrum of GUE matrices. Despite the simplicity of this characterization, it is not know how to recover the explicit determinantal formulas of Tracy and Widom from this formulation.

Chapter 2

Integration on spaces of matrices

In this section, we review the geometry of the classical Lie groups, as well as the spaces $\mathsf{Symm}(n)$, $\mathsf{Her}(n)$ and $\mathsf{Quart}(n)$ and explain how to integrate over these groups and spaces. Our main goal is the following

Theorem 13 (Weyl's formula).

$$DM = |\triangle(\Lambda)|^{\beta} \, D\Lambda \, DU \qquad (2.0.1)$$

where $\triangle(\Lambda)$ is the Vandermonde determinant

$$\triangle(\Lambda) = \prod_{1 \leq j < k \leq n} (\lambda_j - \lambda_k), \qquad (2.0.2)$$

$D\Lambda$ is Lebesgue measure on $\mathbb{R}^n$, and DU denotes (unnormalized) Haar measure on $\mathsf{O}(n)$, $\mathsf{U}(n)/\mathbb{T}^n$, and $\mathsf{USp}(n)/\mathbb{T}^n$ in the cases $\beta = 1, 2$ and 4 respectively.

Remark 14. It is common to normalize the Haar measure such that it is a probability measure. We have ignored this constant here, though is is explored in the exercises. The essential aspect of (2.0.1) is that the Jacobian for diagonalization is given by $|\triangle(\Lambda)|^{\beta}$. This has far-reaching consequences for random matrix theory and has the interesting physical interpretation of *eigenvalue repulsion*.

In what follows, we first present a detailed description of integration on $\mathsf{O}(n)$ and $\mathsf{Symm}(n)$. The ideas are then extended to $\mathsf{Her}(n)$ and $\mathsf{Quart}(n)$.

2.1   Integration on $\mathsf{O}(n)$ and $\mathsf{Symm}(n)$

All matrices in this section lie in $\mathbb{M}_{n \times n}^{\mathbb{R}}$.

There is a natural volume form on each finite-dimensional inner-product space of dimension p. For example, on $\mathbb{R}^p$, the standard inner product defines the metric with infinitesimal length element $ds^2 = \sum_{j=1}^{p} dx_j^2$ and the volume form $Dx = dx_1 dx_2 \ldots dx_p$ (we follow the notation of [25] for volume forms). More generally, each $g \in \mathsf{Symm}_+(p)$ defines an inner-product and metric on $\mathbb{R}^p$.

$$\langle x, y \rangle_g = \sum_{j,k=1}^{p} g_{jk} x_j y_k, \quad ds^2 = \sum_{j,k=1}^{p} g_{jk} dx_j dx_k. \tag{2.1.1}$$

The associated p-dimensional volume form is

$$Dx = \sqrt{\det(g)}\, dx_1 \ldots dx_p. \tag{2.1.2}$$

A (linear) isometry of $\mathbb{R}^n$ is a linear transformation that leaves the inner-product invariant [1]. The Lie group $\mathsf{O}(n)$ is the group, under composition, of linear transformations of $\mathbb{R}^n$ that preserve the standard metric $g = I$. For each $O \in \mathsf{O}(n)$ and each $x \in \mathbb{R}^n$ we must have $(Ox)^T(Ox) = x^T x$. Thus, $\mathsf{O}(n)$ is equivalent to the group of matrices O such that $O^T O = I$. The group operation is matrix multiplication. It is easy to check that the group axioms are satisfied, but a little more work is required to check that $\mathsf{O}(n)$ is a differentiable manifold, and that the group operation is smooth.

We now introduce the natural volume forms on $\mathsf{Symm}(n)$ and $\mathsf{O}(n)$. We first note that the space $\mathsf{Symm}(n)$ is isomorphic to $\mathbb{R}^p$, $p = n(n+1)/2$ via the map

$$M \mapsto \xi = (M_{11}, \ldots, M_{nn}, M_{12}, \ldots, M_{n-1,n}). \tag{2.1.3}$$

Thus, all that is needed to define integrals over $\mathsf{Symm}(n)$ is a choice of inner-product. We will always use the Hilbert-Schmidt inner product

$$\mathsf{Symm}(n) \times \mathsf{Symm}(n) \to \mathbb{R}, \quad (M, N) \mapsto \mathrm{Tr}(M^T N) = \mathrm{Tr}(MN). \tag{2.1.4}$$

The associated infinitesimal length element is

$$ds^2 = \mathrm{Tr}(dM^T dM) = \sum_{j=1}^{n} M_{jj}^2 + 2 \sum_{j<k} M_{jk}^2. \tag{2.1.5}$$

In ξ coordinates on $\mathbb{R}^p$, the associated metric tensor g is diagonal and takes the value 1 for the first n coordinates (diagonal terms), and the value 2 for all the other coordinates (off-diagonal terms). Thus, the metric tensor $g \in \mathsf{Symm}_+(p)$ has determinant $2^{n(n-1)/2}$. We apply formula (2.1.2) to find the following volume form on $\mathsf{Symm}(n)$,

$$DM = 2^{n(n-1)/4} \prod_{j=1}^{n} dM_{jj} \prod_{1 \le j < k \le n} dM_{jk}. \tag{2.1.6}$$

---

[1]It is not necessary to assume that the transformation is linear, but it takes more work to prove that an isometry must be affine.

Each $O \in \mathsf{O}(n)$ defines a map $\mathsf{Symm}(n) \to \mathsf{Symm}(n)$, $M \mapsto OMO^T$. This map is an isometry of $\mathsf{Symm}(n)$ with the metric above. It is in this sense that (2.1.8) is the natural inner-product. Since this map is an isometry, the volume element DM is also invariant.

$\mathsf{O}(n)$ is a differentiable manifold. Thus, in order to define a volume form on $\mathsf{O}(n)$, we must identify its tangent space $T\mathsf{O}(n)$, and then introduce an inner-product on $T\mathsf{O}(n)$. Further, the 'natural' inner-product must be invariant under the group operations. The tangent space at the identity to $\mathsf{O}(n)$, $T_I\mathsf{O}(n)$, is isomorphic to the Lie algebra, $\mathfrak{o}(n)$, of $\mathsf{O}(n)$. In order to compute $\mathfrak{o}(n)$ we consider smooth curves $(-1, 1) \to \mathsf{O}(n)$, $t \mapsto Q(t)$ with $Q(0) = I$, differentiate the equation $Q(t)^T Q(t) = I$ with respect to t, and evaluate at $t = 0$ to find

$$\dot{Q}(0)^T = -\dot{Q}(0). \tag{2.1.7}$$

Thus, each matrix in $\mathfrak{o}(n)$ is antisymmetric. Conversely, given an antisymmetric matrix A, the curve $t \mapsto e^{tA}$ gives a smooth curve in $\mathsf{O}(n)$ that is tangent to I at $t = 0$. Thus,

$$T_I\mathsf{O}(n) = \mathfrak{o}(n) = \{A \,|\, A = -A^T\}. \tag{2.1.8}$$

The tangent space at arbitrary $O \in \mathsf{O}(n)$ is obtained by replacing (2.2.2) with the condition that $O^T\dot{O}$ is antisymmetric. Thus,

$$T_O\mathsf{O}(n) = \{OA \,|\, A \in \mathfrak{o}(n)\}. \tag{2.1.9}$$

Finally, given $A, \tilde{A} \in \mathfrak{o}(n)$, we define their inner product $\mathrm{Tr}(A^T\tilde{A}) = -\mathrm{Tr}(A\tilde{A})$. This inner-product is natural, because it is invariant under left-translation. That is, for two vector $OA, O\tilde{A} \in T_O\mathsf{O}(n)$ we find $\mathrm{Tr}\left((OA)^T(O\tilde{A})\right) = \mathrm{Tr}(A^T\tilde{A})$. The associated volume form on $\mathsf{O}(n)$ is called *Haar measure*. It is unique, upto a normalizing factor, and we write

$$DO = 2^{n(n-1)/4} \prod_{1 \leq j < k \leq n} dA_{jk}. \tag{2.1.10}$$

2.2   Weyl's formula on $\mathsf{Symm}(n)$

Let us now recall some basic facts about $\mathsf{Symm}(n)$. Each matrix $M \in \mathsf{Symm}(n)$ has n real eigenvalues and an orthonormal basis of real eigenvectors. We write Λ for the matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ of eigenvalues, and Q for a matrix whose k-th column is a normalized eigenvector of M associated to the eigenvalue λ_k, $1 \leq k \leq n$. Since the columns of Q are orthogonal and normalized to length 1, it is immediate that $Q \in \mathsf{O}(n)$. Thus,

$$MQ = Q\Lambda \quad \text{and} \quad M = Q\Lambda Q^T. \tag{2.2.1}$$

In what follows, we will view the transformation $M \mapsto (\Lambda, Q)$ as a change of variables, from $\mathsf{Symm}(n) \to \mathbb{R}^n \times \mathsf{O}(n)$. Strictly speaking, this change of variables is not well-defined since (2.2.1) is unaffected if we replace the k-th

column Q_k of Q by $-Q_k$. This issue is considered more carefully in Lemma 3 and Lemma 4 below. In a loose sense, diagonalization is analogous to polar coordinates in $\mathbb{R}^n$,

$$\mathbb{R}^n \to [0, \infty) \times S^{n-1}, \quad x \mapsto (r, u), \quad r = |x|, u = \frac{x}{r}. \qquad (2.2.2)$$

Polar coordinates are natural for rotation invariant probability density on $\mathbb{R}^n$. For example, the standard Gaussian measure on $\mathbb{R}^n$ may be written

$$e^{\frac{-|x|^2}{2}} Dx = C_n e^{\frac{-r^2}{2}} r^{n-1} \, dr \, Du, \qquad (2.2.3)$$

where Du denotes the normalized $n-1$-dimensional measure on S^{n-1} and C_n is a universal constant. The factor r^{n-1} is the Jacobian of this transformation. Weyl's formula shows that the Jacobian for (2.2.1) is $|\triangle(\Lambda)|$. The proof of Weyl's formula relies on an orthogonal decomposition of $T_M\mathsf{Symm}(n)$.

Lemma 2.
$$T_M\mathsf{Symm}(n) \cong \mathbb{R}^n \oplus \mathfrak{o}(n), \qquad (2.2.4)$$

and these subspaces are orthogonal.

Proof. We first assume that M is diagonal. Consider a smooth curve $(-1, 1) \to \mathsf{Symm}(n)$, $t \mapsto M(t) = Q(t)\Lambda(t)Q(t)^T$ such that $M(0) = \Lambda(0) = \Lambda$, and $Q(0) = I$. We differentiate this expression with respect to t and evaluate it at $t = 0$ to find the following expression for a tangent vector in $T_\Lambda\mathsf{Symm}(n)$:

$$\dot{M} = \dot{\Lambda} + [\dot{Q}, \Lambda]. \qquad (2.2.5)$$

Here $\dot{\Lambda}$ can be an arbitrary diagonal matrix, and $\dot{Q}$ an arbitrary antisymmetric matrix. Since the diagonal terms in the commutator $[\dot{Q}, \Lambda]$ vanish, these two matrices are orthogonal. Thus,

$$T_\Lambda\mathsf{Symm}(n) \cong \mathbb{R}^n \oplus \mathfrak{o}(n). \qquad (2.2.6)$$

When M is not diagonal, we consider a curve $M(t)$ as above, with $M(0) = M$, $\Lambda(0) = \Lambda$ and $Q(0) = Q$. Now equation (2.2.5) is replaced by

$$\dot{M} = Q\left(\dot{\Lambda} + [Q^T\dot{Q}, \Lambda]\right)Q^T. \qquad (2.2.7)$$

The matrices $Q^T\dot{Q}$ are antisymmetric and span $\mathfrak{o}(n)$. Moreover, the matrices $[Q^T\dot{Q}, \Lambda]$ and $\dot{\Lambda}$ are orthogonal as before. $\qquad \square$

Remark 15. The proof of Lemma 2 reveals that all matrices of the form $Q\left([Q^T\dot{Q}, \Lambda]\right)Q^T$ lie on an *isospectral manifold* – i.e. a manifold of matrices in $\mathsf{Symm}(n)$ with the same spectrum as Λ.

Proof of Weyl's formula for $\beta = 1$. We now have two coordinate systems on $T_M \mathsf{Symm}(n)$. The coordinates ξ^α, $1 \leq \alpha \leq p$ give the metric (2.1.5). The second coordinate system is $\left(\dot{\Lambda}, \dot{A} \right)$ where $\dot{\Lambda}$ is a diagonal matrix and $\dot{A}$ is an antisymmetric matrix. We use (2.2.7) to see that the infinitesimal length element in this coordinate system is

$$ds^2 = \mathrm{Tr}(dM^2) = \sum_{j=1}^{n} d\lambda_j^2 + 2 \sum_{1 \leq j < k \leq n} (\lambda_j - \lambda_k)^2 dA_{jk}^2. \tag{2.2.8}$$

Thus, the metric tensor in these coordinates is a diagonal matrix in $\mathsf{Symm}_+(p)$ that takes the value 1 on the first n coordinates, and the value $2(\lambda_j - \lambda_k)^2$ for each term A_{jk}. By (2.1.2), the volume form is

$$DM = 2^{n(n-1)/4} \prod_{j=1}^{n} d\lambda_j \prod_{1 \leq j < k \leq n} |\lambda_j - \lambda_k| \, dA_{jk} = |\triangle(\Lambda)| \, D\Lambda \, DQ. \tag{2.2.9}$$

$\square$

2.3 Diagonalization as a change of coordinates

Some care is needed when treating the map $M \to (\Lambda, Q)$ as a change of variables. First, the map is not even well-defined in general, since the sign of each normalized eigenvector is arbitrary. Second, even if we fix the signs, the choice of eigenvectors is degenerate when M has repeated eigenvalues. The following lemmas address this issue.

Lemma 3. *Assume $M_0 \in \mathsf{Symm}(n)$ has distinct eigenvalues. Then there exists $\varepsilon > 0$ such that for each $s \in \{\pm 1\}^n$, there is a C^∞ map*

$$h^{(s)} : B_\varepsilon(M_0) \to \mathbb{R}^n \times \mathsf{O}(n), \quad M \mapsto \left(\Lambda, \, Q^{(s)} \right)$$

that is a C^∞ diffeomorphism onto its image.

Proof of Lemma 3. An outline of the proof is presented. The details are left to the exercises.

The choice of s corresponds to fixing the signs of the eigenvectors as follows. Let a basis of normalized eigenvectors of M_0 be fixed. Call the associated matrix of eigenvectors Q_0. For each s, let $Q_0^{(s)} = \mathrm{diag}(s_1, \ldots, s_n) Q_0$. Each $Q_0^{(s)}$ is also an eigenvector matrix for M_0. Since the eigenvalues of M are distinct, we may use the implicit function theorem to solve the algebraic equations that determine the eigenvalues and eigenvectors, in a way that is consistent with the choice of s. $\square$

Lemma 4. *Assume that M_0 has a repeated eiegenvalue. Then for every $\varepsilon > 0$ there exists M_ε, such that $\|M_0 - M_\varepsilon\| < \varepsilon$ and M_ε has distinct eigenvalues.*

Proof. Exercise.   □

Lemma 3 show that the map $M \mapsto (\Lambda, Q)$ provides a local coordinate system near each matrix with distinct eigenvalues. Lemma 4 shows that set of such matrices is dense. In fact, Weyl's formula shows that the set of $M \in \mathsf{Symm}(n)$ with repeated eigenvalues has measure zero with respect to DM.

2.4   Independence and Invariance implies Gaussian

Fix $M \in \mathsf{Symm}(n)$ with spectrum $\sigma(M)$. Fix an interval $(a, b) \subset \mathbb{R}$ and let $\mathsf{Symm}(n)_{(a,b)}$ denote the set of $M \in \mathsf{Symm}(n)$ with spectrum $\sigma(M) \subset (a, b)$. Each function $f : (a, b) \to \mathbb{R}$ extends naturally to a map $\mathsf{Symm}(n)_{(a,b)} \to \mathsf{Symm}(n)$ as follows:

$$f(M) = Q f(\Lambda) Q^T, \quad M = Q \Lambda Q^T, \quad f(\Lambda) = \mathrm{diag}(f(\lambda_1), \dots, f(\lambda_n)). \quad (2.4.1)$$

Clearly, $\mathrm{Tr}(f(M)) = \mathrm{Tr}(f(\Lambda)) = \sum_{j=1}^n f(\lambda_j)$. Each $f : \mathbb{R} \to \mathbb{R}$ that grows sufficiently fast as $x \to \pm\infty$ defines an invariant distribution on $\mathsf{Symm}(n)$

$$\mu(DM) = \frac{1}{Z} \exp\left(-\mathrm{Tr}(f(M))\right) DM. \quad (2.4.2)$$

This is the most general form of an invariant probability distribution.

By contrast, a Wigner distribution relies on independence of the entries of M. This means that if a Wigner distribution has a density, then it must be of the form

$$\mu(DM) = \frac{1}{Z} \left(\prod_{j=1}^n f_j\left(M_{jj}\right) \prod_{1 \le j < k \le n} f_{jk}\left(M_{jk}\right) \right) DM. \quad (2.4.3)$$

Theorem 16. *Assume a probability measure μ on $\mathsf{Symm}(n)$ is both a Wigner distribution and an invariant distribution. Assume further that $\mu(DM)$ has a strictly positive, smooth density of the form (2.4.2) and (2.4.3). Then $\mu(DM)$ is a Gaussian ensemble,*

$$\mu(DM) = \frac{1}{Z} e^{-\frac{1}{2\sigma^2} \mathrm{Tr}(M - \gamma I)^2}, \quad (2.4.4)$$

with variance σ^2 and mean γI, for some $\gamma \in \mathbb{R}$.

Proof. We first illustrate the essential calculation for 2×2 matrices. Suppose

$$\mu(DM) = p(M) \, DM = \frac{1}{Z} f(M_{11}) g(M_{22}) h(M_{12}) dM_{11} dM_{12} dM_{22}. \quad (2.4.5)$$

We compute the variation in μ along an isospectral curve (see Remark 15). Consider the curve $M(t) = Q(t) M Q(t)^T$ with

$$Q(t) = e^{tR}, \quad R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (2.4.6)$$

The matrix R spans $\mathfrak{so}(2)$. We differentiate $M(t)$ with respect to t to obtain

$$\dot{M}(0) = [M, R] = \begin{pmatrix} -2M_{12} & M_{11} - M_{22} \\ M_{11} - M_{22} & 2M_{12} \end{pmatrix} \qquad (2.4.7)$$

Thus, the infinitesimal change in the density $p(M(t))$ is

$$\begin{aligned} \frac{1}{p}\frac{dp}{dt}\bigg|_{t=0} &= \frac{f'(M_{11})}{f(M_{11})}\dot{M}_{11} + \frac{g'(M_{22})}{f(M_{22})}\dot{M}_{22} + \frac{h'(M_{12})}{h(M_{12})}\dot{M}_{12} \qquad (2.4.8) \\ &= -2M_{12}\left(\frac{f'(M_{11})}{f(M_{11})} - \frac{g'(M_{22})}{f(M_{22})}\right) + (M_{11} - M_{22})\frac{h'(M_{12})}{h(M_{12})}. \end{aligned}$$

On the other hand, since $\mu(DM)$ is invariant, $p(M(t)) = p(M)$ and

$$\frac{dp}{dt}\bigg|_{t=0} = 0. \qquad (2.4.9)$$

We equate (2.4.8) and (2.4.9), and separate variables to obtain

$$\frac{1}{M_{11} - M_{22}}\left(\frac{f'(M_{11})}{f(M_{11})} - \frac{g'(M_{22})}{f(M_{22})}\right) = c = \frac{1}{2M_{12}}\frac{h'(M_{12})}{h(M_{12})}, \qquad (2.4.10)$$

for some constant $c \in \mathbb{R}$. Equation (2.4.10) immediately implies that

$$h(M_{12}) = h(0)e^{cM_{12}^2}. \qquad (2.4.11)$$

Separating variables again in (2.4.10), we find with a second constant $b \in \mathbb{R}$,

$$\frac{f'}{f} = cM_{11} + b, \quad \frac{g'}{g} = cM_{22} + b, \qquad (2.4.12)$$

which integrates to

$$f(M_{11}) = f(0)e^{\frac{cM_{11}^2}{2}}e^{bM_{11}}, \quad g(M_{22}) = g(0)e^{\frac{cM_{22}^2}{2}}e^{bM_{22}}. \qquad (2.4.13)$$

We combine all the terms to obtain

$$p(M) = f(0)g(0)h(0)e^{c\frac{\mathrm{Tr}(M^2)}{2}}e^{b\mathrm{Tr}(M)}. \qquad (2.4.14)$$

Since $p(M)$ integrates to 1, we must have $c < 0$, say $c = -1/\sigma^2$. The scalar b is arbitrary and contributes a shift in the mean that is a scalar multiple of I. The combination of constants $f(0)g(0)h(0)$ may be absorbed into the normalization constant Z^{-1}. We have thus proved Theorem 16 for $n = 2$.

In order to prove Theorem 16 for arbitrary n we generalize the above argument as follows. Fix a pair of off-diagonal indices $1 \le l < m \le n$. We consider a rotation in $\mathbb{R}^n$ that rotates the $x_l x_m$ plane as above, and leaves the other coordinates invariant. This entails replacing the matrix R in the argument above with the matrix $R^{lm} \in \mathfrak{so}(n)$ with coordinates $R_{jk}^{lm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$. The

argument above now shows that the density of p in the M_{ll}, M_{lm} and M_{mm} coordinates is a Gaussian distribution of the form (2.4.14):

$$p(M^{lm}) = e^{c\frac{\text{Tr}((M^{lm})^2)}{2}} e^{b\text{Tr}(M^{lm})},$$

where M^{lm} denotes the 2×2 matrix

$$M^{lm} = \begin{pmatrix} M_{ll} & M_{lm} \\ M_{lm} & M_{mm} \end{pmatrix}.$$

At this stage, the constants c and b depend on l and m. But now note that since the same argument applies to *every* pair of indices $1 \le l < m \le n$, the constants c and b must be independent of l and m. $\square$

2.5   Integration on $\mathsf{Her}(n)$ and $\mathsf{U}(n)$

The space of Hermitian matrices $\mathsf{Her}(n)$ is a vector-space of real dimension n^2, as may be seen by the isomorphism $\mathsf{Her}(n) \to \mathbb{R}^{n^2}$,

$$M \mapsto \xi = (M_{11}, \ldots, M_{nn}, \mathsf{Re}M_{12}, \ldots, \mathsf{Re}M_{n-1,n}, \mathsf{Im}M_{12}, \ldots, \mathsf{Im}M_{n-1,n}).$$
$$\text{(2.5.1)}$$

The Hilbert-Schmidt inner product on $\mathsf{Her}(n)$ is

$$\mathsf{Her}(n) \times \mathsf{Her}(n) \to \mathbb{C}, \quad (M, N) \mapsto \text{Tr}(M^*N). \qquad \text{(2.5.2)}$$

The associated infinitesimal length element is

$$ds^2 = \text{Tr}(dM^2) = \sum_{j=1}^{n} dM_{jj}^2 + 2 \sum_{1 \le j < k \le n} d\,\mathsf{Re}M_{jk}^2 + d\,\mathsf{Im}M_{jk}^2. \qquad \text{(2.5.3)}$$

Thus, in the coordinates ξ, the metric is an $n^2 \times n^2$ diagonal matrix whose first n entries are 1 and all other entries are 2. We apply (2.2.1) to obtain the volume form on $\mathsf{Her}(n)$

$$DM = 2^{n(n-1)/2} \prod_{j=1}^{n} dM_{jj}^2 \prod_{1 \le j < k \le n} d\,\mathsf{Re}M_{jk}\,d\,\mathsf{Im}M_{jk}. \qquad \text{(2.5.4)}$$

The unitary group, $\mathsf{U}(n)$ is the group of linear isometries of $\mathbb{C}^n$ equipped with the standard inner-product $\langle x, y \rangle = x^* y$. Thus, $\mathsf{U}(n)$ is equivalent to the group of matrices $U \in \mathbb{M}_{n \times n}^{\mathbb{C}}$ such that $U^*U = I$. The inner-product (2.5.3) and volume form (2.5.4) are invariant under the transformation $M \mapsto UMU^*$.

The Lie algebra $\mathfrak{u}(n)$ is computed as in Section 2.1. We find

$$\mathfrak{u}(n) = T_I\mathsf{U}(n) = \left\{ A \in \mathbb{M}_{n \times n}^{\mathbb{C}} \,|\, A = -A^* \right\}, \quad T_U\mathsf{U}(n) = \{UA, A \in \mathfrak{u}(n)\}.$$
$$\text{(2.5.5)}$$

The transformation $M \mapsto iM$ is an isomorphism between Hermitian and anti-Hermitian matrices. In fact, the map $\mathsf{Her}(n) \to \mathsf{U}(n)$, $M \mapsto e^{iM}$ is onto and

locally one-to-one. The inner-product $A, \tilde{A} \to -\text{Tr}(A\tilde{A})$ is invariant under left-translation. Thus, we obtain the volume form for Haar measure on $\mathsf{U}(n)$

$$D\tilde{U} = 2^{n(n-1)/2} \prod_{j=1}^{n} dA_{jj}^2 \prod_{1 \leq j < k \leq n} d\mathsf{Re}A_{jk}\, d\mathsf{Im}A_{jk}. \qquad (2.5.6)$$

However, when viewing diagonalization $M \mapsto U\Lambda U^*$ as a change of variables on $\mathsf{Her}(n)$, it is necessary to quotient out the following degeneracy: For each $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{T}^n$, the diagonal matrix $D = \text{diag}\left(e^{i\theta_1}, \ldots, e^{i\theta_n}\right)$ is unitary and $M = U\Lambda U^*$ if and only if $M = UD\Lambda D^* U^*$. Thus, for $\mathsf{Her}(n)$, the measure DU must be replaced by Haar measure on the quotient Lie algebra $\mathsf{U}(n)/\mathbb{T}^n$. The form of Haar measure on $\mathsf{U}(n)/\mathbb{T}^n$ follows from the following assertion, which is proved as in Section 2.1.

Lemma 5. *Each matrix $\dot{M} \in T_M\mathsf{Her}(n)$ is of the form*

$$\dot{M} = U\left(\dot{\Lambda} + [U^*\dot{U}, \Lambda]\right)U^*, \quad \dot{\Lambda} \in T_\Lambda\mathbb{R}^n, \quad \dot{U} \in T_U\mathsf{U}(n), \text{diag}(\dot{U}) = 0. \quad (2.5.7)$$

The matrices $\text{diag}(\dot{\Lambda})$ *and* $[U^*\dot{U}, \Lambda]$ *are orthogonal under the inner-product (2.5.2).*

Thus, the volume form on the quotient $\mathsf{U}(n)/\mathbb{T}^n$ is locally equivalent to a volume form on the subspace of $\mathsf{Her}(n)$ consisting of matrices with zero diagonal:

$$DU = 2^{n(n-1)/2} \prod_{1 \leq j < k \leq n} d\mathsf{Re}A_{jk}\, d\mathsf{Im}A_{jk}. \qquad (2.5.8)$$

2.6 Integration on Quart(*n*) and USp(*n*)

The field of quaternions, $\mathbb{H}$, is the linear space

$$x = c_0 + c_1\mathfrak{e}_1 + c_2\mathfrak{e}_2 + c_3\mathfrak{e}_3, \quad c_i \in \mathbb{R}, \ i = 0, 1, 2, 3, \qquad (2.6.1)$$

equipped with the non-commutative rules of multiplication

$$\mathfrak{e}_1^2 = \mathfrak{e}_2^2 = \mathfrak{e}_3^2 = \mathfrak{e}_1\mathfrak{e}_2\mathfrak{e}_3 = -1. \qquad (2.6.2)$$

These rules ensure that the product of any two quaternions is again a quaternion. Each $x \in \mathbb{H}$ has a complex conjugate $\bar{x} = c_0 - c_1\mathfrak{e}_1 - c_2\mathfrak{e}_2 - c_3\mathfrak{e}_3$, and its absolute value $|x|$ is determined by

$$|x|^2 = \bar{x}x = c_0^2 + c_1^2 + c_2^2 + c_3^2. \qquad (2.6.3)$$

Each non-zero $x \in \mathbb{H}$ has a multiplicative inverse $1/x = \bar{x}/|x|^2$. Thus, $\mathbb{H}$ is indeed a field.

The normed linear vector space $\mathbb{H}^n$ consists of vectors $x = (x_1, \ldots, x_n)^T$ with inner product $\langle x, y \rangle = \sum_{j=1}^{n} \bar{x}_j y_j$. The adjoint, $M^\dagger$ of a linear transformation $M : \mathbb{H}^n \to \mathbb{H}^n$ is defined by the inner-product

$$\langle M^\dagger x, y \rangle := \langle x, My \rangle. \qquad (2.6.4)$$

It follows that the entries of $M^\dagger$ are $M_{jk} = \bar{M}_{kj}$. We say that an operator is self-adjoint if $M = M^\dagger$. It is anti self-adjoint if $M = -M^\dagger$. The space of self-adjoint operators is denoted $\mathsf{Quart}(n)$. We equip this space with the Hilbert-Schmidt norm as before.

The group $\mathsf{USp}(n)$ is the set of linear transformations of $\mathbb{H}^n$ that preserve this inner product. We thus require that for each $x, y \in \mathbb{H}^n$

$$\langle x, y \rangle = \langle Ux, Uy \rangle = \langle U^\dagger U x, y \rangle. \tag{2.6.5}$$

Thus, $\mathsf{USp}(n)$ is equivalent to $U \in \mathbb{M}_{n \times n}^{\mathbb{H}}$ such that $U^\dagger U = I$. As for $\mathsf{U}(n)$ we find that its Lie algebra $\mathfrak{usp}(n)$ is the space of anti self-adjoint matrices. The inner-product on $\mathfrak{usp}(n)$ and Haar measure are defined exactly as in Section 2.5, as is the analogue of Lemma 5 and the Weyl formula.

2.7   Exercises

1. Show that

$$\triangle(\Lambda) = \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \dots & \lambda_n^{n-1}. \end{pmatrix}. \tag{2.7.1}$$

2. The *Pauli matrices*,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.7.2}$$

allow a representation of the quarternions in terms of Hermitian matrices.
(a) Show that the Pauli matrices together with the identity matrix span $\mathsf{Her}(2)$.
(b) Show that the matrices $\{i\sigma_1, i\sigma_2, i\sigma_3\}$ form a basis of $\mathfrak{su}(2)$. (This is the subalgebra of $\mathfrak{u}(2)$ consisting of trace-free matrices).
(c) Verify that if $\mathfrak{e}_j = i\sigma_j$, the rules (2.6.2) hold (replace 1 by I_2).

3. The *canonical symplectic matrix* of size $2n \times 2n$ denoted J_n, or simply J, is the matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{2.7.3}$$

where 0 and I denote the $n \times n$ zero and identity matrices. The *symplectic group* $\mathsf{Sp}(2n, \mathbb{R})$ (not to be confused with the *unitary* symplectic group $\mathsf{USp}(n)$!) is

$$\mathsf{Sp}(2n, \mathbb{R}) = \left\{ S \in \mathbb{M}_{n \times n}^{\mathbb{R}} \, \middle| \, S^T J S = J \right\}. \tag{2.7.4}$$

Verify that $\mathsf{Sp}(2n, \mathbb{R})$ is a group and compute its Lie algebra $\mathfrak{sp}(2n, \mathbb{R})$.

4. Use the Gaussian integral

$$\int_{\mathbb{R}^n} e^{-\frac{|x|^2}{2}} \, dx_1 \dots dx_n.$$

to compute the $n-1$-dimensional volume ω_{n-1} of the unit sphere S^{n-1}. Determine the asymptotic behavior of ω_{n-1} as $n \to \infty$.

Hint: Do the integral two ways– once in Cartesian coordinates, and once in polar coordinates.

5.   Assume given a C^1 function $f : (a, b) \to \mathbb{R}$, and extend it to a function $f : \mathsf{Symm}(n) \to \mathsf{Symm}(n)$ as in (2.4.1). Compute the Jacobian of this transformation. Apply this formula to the function $f(x) = e^{ix}$ to compute the analogue of Weyl's formula on $\mathsf{U}(n)$ (note that each $U \in \mathsf{U}(n)$ is of the form e^{iM} for some $M \in \mathsf{Her}(n)$).

6.   Prove Lemma 4.

7. Assume $f : \mathbb{R} \to (0, \infty)$ satisfies the *functional equation*

$$f(x + y) = f(x)f(y), \quad x, y \in \mathbb{R}. \tag{2.7.5}$$

It is easy to check that for each $a \in \mathbb{R}$ functions of the form $f(x) = e^{ax}$ solve (2.7.5). Show that these are the only solutions to (2.7.5) assuming only that f is continuous. (Do *not* assume that f is differentiable).

Remark 17. The use of row operations in Problem (1) underlies the introduction of orthogonal polynomials. Problems (2) and (3) may be combined to show that $\mathsf{Sp}(2n, \mathbb{C}) \cap \mathsf{U}(n) \cong \mathsf{USp}(n)$. The approach in Problem (4) yields the volume of $\mathsf{O}(n)$, $\mathsf{U}(n)$ and $\mathsf{USp}(n)$ when applied to GOE, GUE and GSE. The assumptions of Problem (7) may be weakened further – measurability is enough! You could try to develop a similar approach for the functional equation implicit in the proof of Theorem 16. That is, can you establish a stronger form of Theorem 16 that does not assume differentiability ?

5. Every $V \in \mathsf{U}(n)$ is of the form $V = \exp(iM)$ for $M \in \mathsf{Her}(n)$ (you can assume this fact, but try and prove it). Thus, show that $V = U\mathrm{diag}(e^{i\alpha_1}, \ldots, e^{i\alpha_n})U^*$ for real numbers $\alpha_1, \ldots, \alpha_n$. Combine this with problem (4) to derive

Chapter 3

Jacobi matrices and tridiagonal ensembles

3.1 Jacobi ensembles

The space of real $n \times n$ tridiagonal matrices is denoted $\mathsf{Tridiag}(n)$. A typical matrix in $\mathsf{Tridiag}(n)$ is written

$$T = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & & \cdots \\ b_1 & a_2 & b_2 & & & \cdots \\ 0 & b_2 & a_3 & \ddots & & 0 \\ \vdots & \vdots & \ddots & \ddots & & b_{n-1} \\ 0 & 0 & \cdots & b_{n-1} & & a_n \end{pmatrix}. \tag{3.1.1}$$

Jacobi matrices, and their closure within the space $\mathsf{Tridiag}(n)$ are the manifolds

$$\begin{aligned} \mathsf{Jac}(n) &= \{T \in \mathsf{Tridiag}(n) \mid b_j > 0,\, 1 \le j \le n\}, \tag{3.1.2} \\ \overline{\mathsf{Jac}(n)} &= \{T \in \mathsf{Tridiag}(n) \mid b_j \ge 0,\, 1 \le j \le n\}. \end{aligned}$$

Jacobi matrices, or more generally Jacobi operators, are of fundamental importance in spectral theory. They also play an important role in approximation theory, the theory of orthogonal polynomials, and in numerical linear algebra. An essential step in the symmetric eigenvalue problem is the reduction of a full symmetric matrix to an isospectral tridiagonal matrix (*tridiagonalization*) by a sequence of orthogonal reflections. Under this procedure, the Gaussian ensembles push forward to ensembles of tridiagonal matrices whose laws have the following simple description.

Definition 18 (Dumitriu-Edelman [4])**.** For each $\beta > 0$, the Hermite(β) ensemble consists of $T \in \mathsf{Tridiag}(n)$ such that a_k, $1 \le k \le n$, are iid normal random variables with mean zero and variance β, and b_k, $1 \le k \le n-1$ are independent $\chi_{k\beta}$ random variables.

The point here is that while the Hermite(β) ensembles are the push-forwards of the Gaussian ensembles when $\beta = 1$, 2 or 4, they interpolate Dyson's classification of ensembles to every $\beta > 0$. When combined with classical spectral theory, they provide a distinct, and important, perspective on the limit theorems of random matrix theory. Our immediate goal in this chapter is the following

Theorem 19. *Fix $\beta > 0$ and assume $T \sim$ Hermite(β). Then the marginal distribution of its eigenvalues is*

$$p_{\mathrm{Hermite}(\beta)} D\Lambda = \frac{1}{Z_{n,\beta}} e^{-\mathrm{Tr}(\Lambda)^2} |\triangle(\Lambda)|^\beta \, D\Lambda. \qquad (3.1.3)$$

The chapter concludes with a more refined version of Theorem 19 that includes the distribution of the *spectral measure* of matrices $T \sim$ Hermite(β).

3.2 Householder tridiagonalization on $\mathsf{Symm}(n)$

Each $M \in \mathsf{Symm}(n)$ may be diagonalized $M = Q\Lambda Q^T$. However, the computation of Λ depends on the solvability of the characteristic polynomial $\det(zI - M) = 0$. For $n \geq 5$, there is no general closed form solution for the characteristic polynomial. [1] Nevertheless, every matrix always admits the following reduction.

Theorem 20. *For every $M \in \mathsf{Symm}(n)$ there exists a tridiagonal matrix T and $Q \in \mathsf{O}(n)$ such that*

$$M = QTQ^T. \qquad (3.2.1)$$

The transformation (3.2.1) is given by a change of variables

$$\mathsf{Her}(n) \to \overline{\mathsf{Jac}(n)} \times \left(S^{n-2} \times S^{n-3} \times \ldots S^1 \right). \qquad (3.2.2)$$

under which the volume form DM on $\mathsf{Symm}(n)$ transforms as follows:

$$DM = C_n \prod_{j=1}^{n} da_j \prod_{k=1}^{n-1} b_k^{n-k-1} db_k \prod_{l=1}^{n-2} d\omega_l \qquad (3.2.3)$$

where $d\omega_l$ denotes uniform measure on the sphere S^l, and C_n is a universal constant.

Remark 21. The space $\mathsf{Tridiag}(n)$ clearly inherits the inner-product $\mathrm{Tr}(T^2) = \sum_{j=1}^{n} a_j^2 + 2 \sum_{j=1}^{n-1} b_j^2$ from $\mathsf{Symm}(n)$. However, the volume form obtained from this metric is *not* the same as the volume form in (3.3.3) above.

Remark 22. (For algebraists!) The proof will also show that T and Q may be computed with a finite number of the following algebraic operations: addition, multiplication and square-roots.

---

[1] Practical numerical schemes for eigenvalue decomposition are unaffected by this algebraic obstruction, since they rely on iteration.

Definition 23. Suppose $v \in \mathbb{R}^n$ is a unit vector. The *Householder reflection* in v is the matrix

$$P_v = I - 2vv^T. \tag{3.2.4}$$

Lemma 6. *The matrix P_v has the following properties:*

(a) $P_v^2 = I$.

(b) $P_v \in \mathsf{O}(n)$.

Proof. Decompose $\mathbb{R}^n$ into the orthogonal subspaces span$\{v\}$ and $v^\perp$. Then $P_v v = -v$ and $P_v|_{v^\perp} = I$. Thus, $P_v^2 = I$. This proves (a). By construction $P_v^T = P_v$. Thus, by (a), we also have $P_v^T P_v = I$.   □

Proof of Theorem 20. 1. The proof relies on a sequence of Householder reflections that progressively introduce zeros in a sequence of matrices similar to M. The first such matrix is the following. Let $w_1 = (M_{21}, \ldots, M_{n1})^T \in \mathbb{R}^{n-1}$ denote the last $n-1$ entries of the first column of M. If the first coordinate of w_1 is non-negative, and all other coordinates vanish there is nothing to do. If not, we may choose a Householder reflection (in $\mathbb{R}^{n-1}$) that maps w_1 to $|w_1|e_1^{(n-1)}$ (here the superscript $n-1$ denotes that we consider the basis vector $e_1 \in \mathbb{R}^{n-1}$). Geometrically, such a reflection is obtained by choosing v_1 to be the unit vector that lies in between w_1 and $|w_1|e_1^{(n-1)}$. Explicitly, we set

$$\tilde{v}_1 = |w_1|e_1^{(n-1)} - w_1, \quad v_1 = \frac{\tilde{v}_1}{|\tilde{v}_1|}, \quad P^{(1)} = P_{v_1}. \tag{3.2.5}$$

By Lemma 6, $P^{(1)} \in \mathsf{O}(n-1)$ is a Householder reflection that maps w_1 to $|w|e_1^{(n-1)}$. It may be extended to a Householder reflection in $\mathsf{O}(n)$, by defining

$$Q^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & P^{(1)} \end{pmatrix}. \tag{3.2.6}$$

Then the matrix

$$M^{(1)} := Q^{(1)} M \left(Q^{(1)} \right)^T = Q^{(1)} M Q^{(1)}, \tag{3.2.7}$$

is similar to M. By construction, the first row of $M^{(1)}$ is $(M_{11}, |w_1|, 0, \ldots, 0)$, and the first column is $(M_{11}, |w_1|, 0, \ldots, 0)^T$. Thus, we may write

$$M^{(1)} = \begin{pmatrix} T^{(1)} & |w_1|(e_1^{(n-1)})^T \\ |w_1|e_1^{(n-1)} & N^{(1)} \end{pmatrix}, \tag{3.2.8}$$

where $T^{(1)}$ is a (trivial) 1×1 tridiagonal matrix and $N^{(1)} \in \text{Symm}(n-1)$.

   2. The proof is completed by induction. Assume that $M^{(k)} \in \mathsf{Symm}(n)$ has the form

$$M^{(k)} = \begin{pmatrix} T^{(k)} & |w_k|(e_1^{(n-k)})^T \\ |w_k|e_1^{(n-k)} & N^{(k)}) \end{pmatrix}, \tag{3.2.9}$$

where $T^{(k)} \in \mathsf{Tridiag}(k)$ and $N^{(k)} \in \mathrm{Symm}(n-k)$, $1 \le k \le n-1$. We apply the procedure of step 1 to $N^{(k)}$ to obtain a vector v_k, a Householder reflection $P^{(k)} = P_{v_k}$, and an orthogonal transformation of $M^{(k)}$,

$$Q^{(k)} = \begin{pmatrix} I_k & 0 \\ 0 & P^{(k)} \end{pmatrix} \in \mathsf{O}(n), \quad M^{(k+1)} = Q^{(k)} M^{(k)} Q^{(k)}. \qquad (3.2.10)$$

Note that $Q^{(k)}$ leaves the first k rows and columns of $M^{(k)}$ unchanged, thus it does not destroy the tridiagonal structure of the first k rows and columns. Thus, $M^{(k+1)}$ has the form (3.2.9) with the index k replaced by $k+1$.

The procedure terminates when $k = n-2$, and yields

$$M = QTQ^T, \quad Q = Q^{(n-2)}Q^{(n-3)} \ldots Q^{(1)}. \qquad (3.2.11)$$

3. It is simplest to prove (3.3.3) probabilistically. Informally, the k-th step of the procedure above is a change to polar coordinates in $\mathbb{R}^{n-k}$, with $b_k \ge 0$ playing the role of the radius, and the factor $b_k^{n-k-1} db_k d\omega_{n-1-k}$ being the pushforward of Lebesgue measure in $\mathbb{R}^{n-k}$ to polar coordinates. More precisely, assume that $M \sim$ GOE. We note that the first step of the above procedure leaves M_{11} alone. Thus, $a_1 = M_{11} \sim \mathcal{N}(0,1)$. Moreover, the term b_1 is the length of the first column of M, not including the diagonal term M_{11}. Since a χ_m^2 random variable has the same law as the length of a vector in $\mathbb{R}^m$ whose entries are iid $\mathcal{N}(0,1)$ random variables, we see that $b_1 \sim \chi_{n-1}$. Further, the vector w_1 is uniformly distributed on S^{n-2}. We next observe that by the independence and invariance of the Gaussian ensembles, the matrix $N^{(1)}$ in (3.2.8) $\sim$ GOE$(n-1)$. Indeed, $\tilde{M}_1$, the lower-right $(n-1) \times (n-1)$ block of M, is a GOE$(n-1)$ matrix, and the reflection $P^{(1)}$ is independent of $\tilde{M}_1$. Thus, $N^{(1)} = P^{(1)} \tilde{M}_1 P^{(1)}$ has law GOE$(n-1)$. Thus, $a_2 \sim \mathcal{N}(0,1)$ and $b_2 \sim \chi_{n-2}$. An obvious induction now shows that if $M \sim$ GOE then $T \sim$ Hermite(1), and the vectors w_k, are uniformly distributed on S^{n-1-k}, $1 \le k \le n-2$. Comparing the two laws, we find (with $\beta = 1$)

$$e^{-\frac{\beta \mathrm{Tr}(M^2)}{2}} DM = C_n e^{-\frac{\beta \mathrm{Tr}(T^2)}{2}} da_j \prod_{k=1}^{n-1} b_k^{n-k-1} db_k \prod_{l=1}^{n-2} d\omega_l \qquad (3.2.12)$$

The exponential weights cancel, and yield the Jacobian formula (3.3.3). $\qquad \square$

3.3 Tridiagonalization on $\mathsf{Her}(n)$ and $\mathsf{Quart}(n)$

Theorem 20 admits a natural extension to $\mathsf{Her}(n)$ and $\mathsf{Quart}(n)$.

Theorem 24. *For every $M \in \mathsf{Her}(n)$ (resp. $\mathsf{Quart}(n)$) there exists a tridiagonal matrix $T \in \mathsf{Jac}(n)$ and $Q \in \mathsf{U}(n)$ (resp. $\mathsf{USp}(n)$) such that*

$$M = QTQ^*. \qquad (3.3.1)$$

The transformation (3.3.1) is given by a change of variables

$$\mathsf{Her}(n) \to \overline{\mathsf{Jac}(n)} \times \left(S_{\mathbb{F}}^{n-2} \times S_{\mathbb{F}}^{n-3} \times \ldots S_{\mathbb{F}}^1\right), \qquad (3.3.2)$$

where $S_{\mathbb{F}}^l$ denotes the unit sphere in $\mathbb{F}^l$, with $\mathbb{F} = \mathbb{C}$ (resp. $\mathbb{H}$). The volume form DM on $\mathsf{Her}(n)$ (resp. $\mathsf{Quart}(n)$) transforms as follows:

$$DM = C_n \prod_{j=1}^{n} da_j \prod_{k=1}^{n-1} b_k^{n-k-1} db_k \prod_{l=1}^{n-2} d\omega_l \qquad (3.3.3)$$

where $d\omega_l$ denotes uniform measure on the sphere $S_{\mathbb{F}}^l$, and C_n is a universal constant.

Remark 25. Note that the matrix T is always real, whereas the entries of M and Q are in $\mathbb{C}$ or $\mathbb{H}$.

The proof of Theorem 24 is in the same vein as that of Theorem 20. It is only necessary to replace the Householder projections in $\mathsf{O}(n)$ with projections in $\mathsf{U}(n)$ and $\mathsf{USp}(n)$. For example, given $v \in \mathbb{C}^n$ with $|v| = 1$, the associated Householder projection in $\mathsf{U}(n)$ is $P_v = I - vv^*$. Step 3 in the proof of Theorem 24 also explains the role of the parameter β in the definition of the Hermite-β ensembles. The k-th step of the Householder transformation maps a standard Gaussian vector in $\mathbb{C}^{n-k}$ to its magnitude and direction. The law of the magnitude is now $\chi_{2(n-k)}$ (or $\chi_{\beta(n-k)}$ with $\beta = 2$). Similarly, the direction of the Gaussian vector is uniformly distributed on the unit sphere in $\mathbb{C}^{n-k-1}$.

3.4 Inverse spectral theory for Jacobi matrices

Bounded Jacobi operators admit a complete and beautiful spectral theory that is intimately tied to orthogonal polynomials and continued fractions. We first introduce this theory for finite Jacobi matrices, since it underlies Theorem 19. As usual, write

$$T = U\Lambda U^T, \quad U \in \mathsf{O}(n), \qquad (3.4.1)$$

for the diagonalization of T. We also define the *Weyl chamber*

$$\mathcal{W}^n = \{\Lambda \in \mathbb{R}^n \,|\, \lambda_1 < \lambda_2 < \ldots < \lambda_n\}. \qquad (3.4.2)$$

For each $\Lambda \in \mathcal{W}^n$, its *isospectral manifold* is the set

$$\mathcal{M}_\Lambda = \{T \in \overline{\mathsf{Jac}(n)} \,\big|\, T = U\Lambda U^T, \text{ for some } U \in \mathsf{O}(n)\}. \qquad (3.4.3)$$

The following theorem shows that the interior of the isospectral manifold is diffeomorphic to the positive orthant of the unit sphere.

Theorem 26. *The spectral mapping*

$$\mathcal{S} : \mathsf{Jac}(n) \to \mathcal{W}^n \times S_+^{n-1}, \quad T \mapsto (\Lambda, U^T e_1), \qquad (3.4.4)$$

is an analytic diffeomorphism.

The isospectral manifold admits several distinct parametrizations. First, it is clear that we may use the simplex Σ_n instead of the orthant S_+^{n-1}. Indeed, let $u = U^T e_1$ denote the first row of the matrix of eigenvectors and define $c_j = u_k^2$, $1 \leq k \leq n$. Since $U \in \mathsf{O}(n)$, $\sum_{k=1}^n u_k^2 = 1$. Thus, $u \in S^{n-1}$ and $c \in \Sigma_n$. Lemma 7 below shows that u_k can be chosen to be strictly positive, which allows us to restrict attention to the positive orthant S_+^{n-1}

Theorem 26 may also be viewed as a mapping to the *spectral measure*

$$T \mapsto \mu = \sum_{j=1}^n u_j^2 \delta_{\lambda_j} = \sum_{j=1}^n c_j \delta_{\lambda_j}. \qquad (3.4.5)$$

It is often more convenient to work with the Cauchy transform of the spectral measure, μ. Define the τ-*function*,

$$\mu \mapsto \tau(z) = \int_{\mathbb{R}} \frac{1}{x - z} \mu(dx) = \sum_{j=1}^n \frac{u_j^2}{\lambda_j - z}, \quad z \in \mathbb{C} \backslash \{\lambda_1, \ldots, \lambda_n\}. \qquad (3.4.6)$$

The inverse $\tau \mapsto \mu$ is obtained by computing the poles and residues of τ.

The τ-function may also be written as a ratio of polynomials of degree $n - 1$ and n respectively. Let $T_k \in \mathsf{Jac}(k)$ denote the lower-right $k \times k$ submatrix of T, $1 \leq k \leq n$. It follows from Cramer's rule that

$$\tau(z) = e_1^T (T - z)^{-1} e_1 = \frac{\det(T_{n-1} - zI)}{\det(T - zI)} = \frac{\prod_{j=1}^{n-1} (\lambda_j^{(n-1)} - z)}{\prod_{j=1}^n (\lambda_j^{(n)} - z)}, \qquad (3.4.7)$$

where $\Lambda^{(k)}$ denotes the diagonal matrix of eigenvalues of T_k and $\Lambda^{(n)} = \Lambda$. We will show that the ordered eigenvalues of T_{n-1} and T_n interlace, i.e.

$$\lambda_1^{(n)} < \lambda_1^{(n-1)} < \lambda_2^{(n)} < \ldots < \lambda_{n-1}^{(n-1)} < \lambda_n^{(n)}. \qquad (3.4.8)$$

Thus, *interlacing sequences* provide another parametrization of $\mathsf{Jac}(n)$. A convenient visal description of interlacing sequences, called *diagrams*, was introduced by Kerov and Vershik [13]. The importance of these alternate parametrizations (spectral measures, τ-function, diagrams) is that they provide a transparent framework for the analysis of the limit $n \to \infty$.

The surprising aspect of Theorem 3.4.4 is that the spectral data (Λ, u) provides enough information to reconstruct the matrix T. There are two reconstruction procedures. The first involves orthogonal polynomials, the second uses the theory of continued fractions. We explain the use of orthogonal polynomials below, and outline the theory of continued fractions in the exercises. In order to develop these procedures, it is first necessary to establish basic properties of the eigenvalues of Jacobi matrices.

Lemma 7. *Assume $T \in \mathsf{Jac}(n)$. Then*

    1. The first entry of each non-zero eigenvector is non-zero. In particular, we may normalize the eigenvectors to ensure $u_k > 0$ for $1 \leq k \leq n$.

2. The eigenvalues of T are distinct.

Proof. We write the eigenvalue equation $Tv = zv$ in coordinates.

$$b_{k-1}v_{k-1} + (a_k - z)\, v_k + b_k v_{k+1} = 0, \quad 1 \le k \le n, \qquad (3.4.9)$$

with the convention $b_0 = b_n = 0$. Since the off-diagonal terms b_k are strictly positive, we may solve this linear system recursively. Given v_1, we find

$$v_2 = \frac{v_1(z - a_1)}{b_1}, \quad v_3 = \frac{v_1}{b_1 b_2}\left((a_2 - z)(a_1 - z) - b_1^2\right), \text{ etc.} \qquad (3.4.10)$$

Thus, $v \equiv 0 \in \mathbb{R}^n$ if $v_1 = 0$. Further, the solution space to the eigenvalue equation $Tv = \lambda v$ has dimension at most 1. $\qquad \square$

Lemma 8. *The characteristic polynomials $d_k(z) = \det(zI - T_k)$ satisfy the recurrence relations*

$$d_{k+1}(z) = (z - a_{n-k})d_k(z) - b_{n-k}^2\, d_{k-1}(z), \quad 1 \le k \le n-1, \qquad (3.4.11)$$

with the initial condition $d_0(z) \equiv 1$ and the convention $b_n = 0$.

Proof. Expand the determinant $\det(zI - T_k)$ about the k-th row, and compute the minors associated to $z - a_{n-k}$ and b_{n-k}. $\qquad \square$

Lemma 9. *The eigenvalues of T_k and T_{k+1} interlace, $1 \le k \le n-1$.*

Proof. We consider the τ-functions for the minors T_k,

$$\tau_k(z) = \frac{\det(T_k - zI)}{\det(T_{k+1} - zI)} = -\frac{d_k(z)}{d_{k+1}(z)}. \qquad (3.4.12)$$

By the recurrence relation (3.4.11), we have

$$-\frac{1}{\tau_k(z)} = z - a_{n-k} + b_{n-k}^2 \tau_{k-1}(z). \qquad (3.4.13)$$

We claim that on the real line, $\tau_k(x)$ is strictly increasing between the zeros of π_k. Indeed, it is clear that $\tau_1(x) = (a_n - x)^{-1}$ has this property, and upon differentiating (3.4.13) we find that

$$\frac{1}{\tau_k^2}\tau_k' = 1 + b_{n-k}^2 \tau_{k-1}' > 0,$$

except at poles. The claim follows by induction.

Since τ_k is strictly increasing between poles, by the intermediate value theorem, it has exactly one zero between any two poles. By (3.4.12), the zeros of τ_k are the eigenvalues of T_k, and the poles of τ_k are the eigenvalues of T_{k+1}. Thus, they interlace. $\qquad \square$

A remarkable feature of the spectral theory of Jacobi matrices is that the orthogonal polynomials associated to the spectral measure $\mu(T)$ may be used to reconstruct T. In order to state this assertion precisely, let us recall some basic facts about orthogonal polynomials. Assume given a probability measure μ on $\mathbb{R}$ that has finite-moments of all orders, i.e.,

$$\int_{\mathbb{R}} |x|^\alpha \, \mu(dx) < \infty, \quad \alpha > 0. \tag{3.4.14}$$

We may apply the Gram-Schmidt procedure to the monomials $\{x^k\}_{k=0}^\infty$ to construct a sequence of polynomials that are orthogonal in $L^2(\mathbb{R}, \mu)$. There are two standard normalizations that one may adopt.

Orthonormal polynomials, denoted $\{p_k\}_{k=0}^\infty$, have the property that p_k is of degree k, $k = 0, 1, 2, \ldots$, and

$$\int_{\mathbb{R}} p_k(x) p_l(x) \, \mu(dx) = \delta_{kl}. \tag{3.4.15}$$

Monic polynomials, denoted $\{\pi_k\}_{k=0}^\infty$ have the property that $\pi_k(x)$ is of degree k and the coefficient of x^k is 1. Further,

$$\int_{\mathbb{R}} \pi_k(x) \pi_l(x) \, \mu(dx) = 0, \quad k \neq l. \tag{3.4.16}$$

Lemma 10 (Three-term recurrence for orthogonal polynomials)**.** *Given* $(\Lambda, u) \in \mathcal{W}^n \times S_+^{n-1}$, *let* $\mu(\Lambda, u) = \sum_{k=1}^n u_k^2 \delta_{\Lambda_k}$. *Then the associated monic orthogonal polynomials* $\{\pi_k\}_{k=0}^n$, *satisfy the three-term recurrence (3.4.17)*

$$\pi_k(z) = (z - a_k)\pi_{k-1}(z) - b_{k-1}^2 \, \pi_{k-2}(z), \quad 2 \leq k \leq n, \tag{3.4.17}$$

where the coefficients a_k *and* b_k *are given by*

$$a_k = \frac{\int_{\mathbb{R}} x\pi_k^2 \, \mu(dx)}{\int_{\mathbb{R}} \pi_k^2(x) \, \mu(dx)}, \quad b_{k-1}^2 = \frac{\int_{\mathbb{R}} x\pi_k \pi_{k-1}(x) \, \mu(dx)}{\int_{\mathbb{R}} \pi_{k-1}^2(x) \, \mu(dx)}, \quad k = 1, \ldots, K, \tag{3.4.18}$$

and $b_0 = 0$. *The recurrence (3.4.18) defines a Jacobi matrix* $T(\mu)$.

Remark 27. If μ is not a discrete measure of the form (3.4.5), but has bounded support, the recurrence (3.4.17) defines a bounded Jacobi operator on $l^2(\mathbb{C})$.

Proof. Given any μ as in (3.4.14), we obtain the sequence $\{\pi_k\}$ using the Gram-Schmidt procedure. When μ is of the form (3.4.5), the vector space $L^2(\mathbb{R}, \mu)$ has dimension n and the Gram-Schmidt procedure yields an orthogonal basis $\{\pi_0, \pi_1, \ldots, \pi_{n-1}\}$ for $L^2(\mathbb{R}, \mu)$.

The three-term recurrence for the orthogonal polynomials is obtained as follows. Since $x\pi_k(x)$ is a polynomial of degree $k + 1$ it can be expressed as a

linear combination $x\pi_k(x) = \sum_{j=0}^{k+1} c_{j,k}\pi_j(x)$. Since the π_j are monic, we must have $c_{k+1,k} = 1$. Moreover, for $j = 0, \ldots, k-2$

$$\int_{\mathbb{R}} x\pi_k(x)\pi_j(x)\,\mu(dx) = \int_{\mathbb{R}} \pi_k(x)x\pi_j(x)\,\mu(dx) = 0,$$

since $x\pi_j$ lies in the span of $\{\pi_0, \ldots, \pi_{k-1}\}$. Thus, $c_{j,k} = 0$ for $j = 0, \ldots, k-2$ and we find

$$x\pi_k(x) = \pi_{k+1}(x) + c_{k,k}\pi_k(x) + c_{k-1,k}\pi_{k-1}(x). \tag{3.4.19}$$

It remains to show that $c_{k-1,k} > 0$. By orthogonality, $\int_{\mathbb{R}} x\pi_k(x)\pi_{k+1}(x)\,\mu(dx) = \int_{\mathbb{R}} \pi_{k+1}^2(x)\,\mu(dx)$. Thus, $c_{k,k-1} > 0$ for all k such that $\pi_{k-1}(x)$ does not vanish in $L^2(\mathbb{R}, \mu)$. This is (3.4.17) aside from a change in notation.   □

Proof of Theorem 26. We have defined a forward map $T \mapsto \mu(T)$ as follows. The matrix T defines a τ-function $\tau(z) = e_1^T(T - zI)^{-1}e_1$, which is expressed as a ratio of characteristic polynomials in (3.4.7). The poles of $\tau(z)$ are the eigenvalues of T. The norming constants are the residues at the poles, and are given by

$$u_k^2 = -\frac{d_{n-1}(\lambda_k)}{d_n'(\lambda_k)}, \quad 1 \le k \le n. \tag{3.4.20}$$

The inverse map $\mu \to T(\mu)$ is given by Lemma 10. The orthogonal polynomials defined by μ satisfy a three-term recurrence whose coefficients determine T.

We only need to show that the map $\mu \mapsto T(\mu) \mapsto \mu(T(\mu))$ is the identity. Let $\mu \cong (\Lambda, u)$ be given and define $T(\mu)$ by the recurrence relations. We will show that

$$e_1^T(T - zI)^{-1}e_1 = \int_{\mathbb{R}} \frac{1}{x-z}\mu(dx) = \sum_{k=1}^{n} \frac{u_k^2}{\lambda_k - z}. \tag{3.4.21}$$

We expand both sides of this equation for large z, to see that all we have to establish is the relation

$$e_1^T T^k e_1 = \int_{\mathbb{R}} x^k \mu(dx), \quad 0 \le k \le n-1. \tag{3.4.22}$$

The remainder of the proof is left as an exercise to the reader.   □

Remark 28. Observe that the recurrence relation (3.4.17) may be rewritten as the matrix equation,

$$\begin{pmatrix} a_1 - z & 1 & 0 & \ldots & 0 \\ b_1^2 & a_2 - z & 1 & \ldots & 0 \\ 0 & b_2^2 & a_3 - z & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & \ldots & b_{n-1}^2 & a_n - z \end{pmatrix} \cdot \begin{pmatrix} \pi_0(z) \\ \pi_1(z) \\ \vdots \\ \pi_{k-1}(z) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -\pi_k(z) \end{pmatrix}.$$

$$\tag{3.4.23}$$

Since $\pi_0(z) = 1$, each zero of $\pi_k(z)$ is an eigenvalue of the matrix above. Thus, $\pi_k(z) = \det(zI - \tilde{T}_k)$ where $\tilde{T}_k$ denotes the *upper-left* $k \times k$ submatrix of T (compare with T_k and $d_k(z) = \det(zI - T_k)$).

Thus, given μ, the entries of T are obtained from "top to bottom". However, given T, the τ-function is the limit of τ-functions $-d_k(z)/d_{k+1}(z)$ computed 'bottom-to-top'.

3.5 Jacobians for tridiagonal ensembles

We can now combine Theorem 26 with the definition of Hermite-β ensembles to state a refined version of Theorem 19.

Theorem 29. *For each $\beta > 0$, the law of the Hermite-β ensembles in spectral variables $(\Lambda, u) \in \mathcal{W}^n \times S_+^{n-1}$ is given by the density*

$$p_{\text{Hermite}}(\Lambda, u) = \frac{1}{Z_{n,\beta}} \left(e^{-\frac{\beta}{4}\text{Tr}(\Lambda^2)} |\triangle(\Lambda)|^\beta D\Lambda \right) \left(\prod_{j=1}^n u_j^\beta \, du_j \right). \qquad (3.5.1)$$

In particular, Λ and u are independent.

Theorem 29 follows from a computation of the Jacobian of the spectral map $\mathcal{S} : \text{Jac}(n) \to \mathcal{W}^n \times S_+^{n-1}$.

Theorem 30. *The volume forms on $\text{Jac}(n)$ and $\mathcal{W}^n \times S_+^{n-1}$ are related by*

$$DT = \prod_{j=1}^n da_j \prod_{k=1}^{n-1} b_k^{n-k-1} \, db_k = C_n \triangle(\Lambda) D\Lambda \prod_{k=1}^n u_j \, du_j. \qquad (3.5.2)$$

where C_n is a universal constant.

Remark 31. We have suppressed the explicit form of the universal constants in the statement of the lemma to focus on the marginals on $\mathcal{W}^n$ and S_+^{n-1} respectively. The computation of the constants is an interesting exercise (see [4]).

While Theorem 30 is an analytic/geometric assertion, the simplest proof uses probabilistic reasoning, as in step 3 of the proof of Theorem 20. Since we have computed the Jacobian for the diagonalizing map $\text{Symm}(n) \to \mathbb{R}^n \times \text{O}(n)$ (Weyl's formula) and the tridiagonalizing map $\text{Symm}(n) \to \text{Jac}(n)$ (Theorem 20), the ratio of these Jacobians may be used to compute the Jacobian of the spectral map $\text{Jac}(n) \to \mathcal{W}^n \times S_+^{n-1}$. The main point is that by the $\text{O}(n)$ invariance of GOE, the top row of the eigenvector matrix must be uniformly distributed on S^{n-1} and is independent of Λ. This gives the term $\prod_{k=1}^n u_j \, du_j$ in equation (3.5.2). As Dumitriu and Edelman remark, this is a 'true random matrix theory' calculation. Another approach to (3.5.2) uses symplectic geometry.

Lemma 11 (Vandermonde determinant in (a, b) coordinates).

$$\triangle(\Lambda) = \prod_{j<k}(\lambda_j - \lambda_k) = \frac{\prod_{k=1}^{n-1} b_k^k}{\prod_{j=1}^{n} u_j}. \tag{3.5.3}$$

Proof. 1. Recall that $\Lambda^{(l)}$ denotes the diagonal matrix of eigenvalues of T_l and that $\pi_l(x) = \prod_{j=1}^{l}(x - \lambda_j^{(l)})$. Therefore, we have the identity

$$\prod_{j=1}^{l}\prod_{k=1}^{l-1}\left|\lambda_j^{(l)} - \lambda_k^{(l-1)}\right| = \left|\prod_{j=1}^{l}\pi_{l-1}\left(\lambda_j^{(l)}\right)\right| = \left|\prod_{k=1}^{l-1}\pi_l\left(\lambda_k^{(l-1)}\right)\right|. \tag{3.5.4}$$

Since π_{l-1} and π_l are related through the three-term recurrence

$$\pi_l(x) = (x - a_l)\pi_{l-1}(x) - b_{l-1}^2\pi_{l-2}(x),$$

we have

$$\left|\prod_{k=1}^{l-1}\pi_l(\lambda_k^{(l-1)})\right| = b_{l-1}^{2(l-1)}\left|\prod_{k=1}^{l-1}\pi_{l-2}(\lambda_k^{(l-1)})\right| = b_{l-1}^{2(l-1)}\left|\prod_{j=1}^{l-2}\pi_{l-1}(\lambda_j^{(l-2)})\right|.$$

We apply this identity repeatedly, starting with $l = n$ to obtain

$$\left|\prod_{k=1}^{n-1}\pi_n(\lambda_k^{(n-1)})\right| = b_{n-1}^{2(n-1)}\left|\prod_{j=1}^{n-2}\pi_{n-1}(\lambda_j^{(n-2)})\right|$$

$$= b_{n-1}^{2(n-1)}b_{n-2}^{2(n-2)}\left|\prod_{k=1}^{n-3}\pi_{n-2}(\lambda_k^{(n-3)})\right| = \ldots = \prod_{k=1}^{n-1}b_k^{2(n-k-1)}.$$

2. The coefficients u_j^2 are the residue of $\tau_n(z)$ at the poles λ_j, i.e.

$$u_k^2 = \left|\frac{\pi_{n-1}(\lambda_k)}{\pi_n'(\lambda_k)}\right|, \quad 1 \le k \le n. \tag{3.5.5}$$

Observe also that

$$\pi_n'(\lambda_k) = \prod_{j\neq k}(\lambda_j - \lambda_k), \text{ and } \prod_{k=1}^{n}\pi_n'(\lambda_k) = \triangle(\Lambda)^2. \tag{3.5.6}$$

Therefore,

$$\prod_{j=1}^{n}u_j^2 = \frac{1}{\triangle(\lambda)^2}\prod_{k=1}^{n}|\pi_{n-1}(\lambda_k)| = \frac{\prod_{k=1}^{n-1}b_k^{2k}}{\triangle(\lambda)^2}. \tag{3.5.7}$$

$\square$

Proof of Theorem 30. 1. Our goal is to compute the Jacobian of the spectral mapping $\mathcal{S}$,

$$DT = \frac{\partial(T(a,b))}{\partial(\Lambda, u)} D\Lambda Du, \tag{3.5.8}$$

where $Du = \prod_{j=1}^{n} du_j$. Rather than compute the change of variables directly, we will compute the push-forward of GOE onto $\mathsf{Jac}(n)$ and $\mathcal{W}^n \times S_+^{n-1}$ separately, and obtain the Jacobian above.

2. Consider the push-forward of GOE under the map $M \mapsto (\Lambda, u)$, where $M = U\Lambda U^T$ is the diagonalization of M. Since Λ and the matrix of eigenvalues U are independent, Λ and $u = U^T e_1$ are independent. Since U is distributed according to Haar measure on $\mathsf{O}(n)$, the vector u is uniformly distributed on S^{n1} and the push-forward of GOE is the measure

$$p(\Lambda, u) D\Lambda Du = c_n e^{-\frac{1}{2}\mathrm{Tr}(\Lambda)^2} \triangle(\Lambda) D\Lambda Du. \tag{3.5.9}$$

3. Next consider the push-forward of GOE under the map $M \mapsto T$, where $M = QTQ^T$ denotes the tridiagonalization of M. As we have seen in the proof of Theorem 20, T and U are independent, and the marginal distribution of T is given by

$$\tilde{p}(T) \, DT = C_n e^{-\frac{1}{2}\mathrm{Tr}(T^2)} \prod_{j=1}^{n} da_j \prod_{k=1}^{n-1} b_k^{n-k-1} \, db_k. \tag{3.5.10}$$

4. Since $T \in \mathsf{Jac}(n)$ and $(\Lambda, u) \in \mathcal{W}^n \times S_+^{n-1}$ are in bijection, we have

$$p(\Lambda, u) = \tilde{p}(T(\Lambda, u)) \frac{DT}{D(\Lambda, u)}. \tag{3.5.11}$$

We compare the expressions in (3.5.9) and (3.5.10) and use Lemma 11 to obtain

$$\frac{DT}{D(\Lambda, u)} = \frac{C_n}{c_n} \frac{\prod_{k=1}^{n-1} b_k}{\prod_{j=1}^{n} u_j}. \tag{3.5.12}$$

The constants are computed in [4]. □

Proof of Theorem 29. The law of We change variables using the spectral mapping and Theorem 30 to obtain the following identity for the law of the Hermite$-$ β ensembles

$$C_{n,\beta} e^{-\frac{\beta}{4}\mathrm{Tr}(T^2)} \prod_{k=1}^{n-1} b_k^{\beta(n-k)-1} \, DT \tag{3.5.13}$$

$$= C_{n,\beta} \left(e^{-\frac{\beta}{4}\mathrm{Tr}(\Lambda^2)} \triangle(\Lambda)^\beta \, D\Lambda \right) \left(Du \prod_{j=1}^{n} u_j^{\beta-1} \right). \tag{3.5.14}$$

Since the distribution factors, Λ and u are independent with the laws stated in Theorem 29. □

3.6 Exercises

1. Write a numerical code to sample matrices from both GOE and the Hermite$-$1 ensemble. Verify numerically that a suitably normalized density of eigenvalues for the GOE matrix approaches the semicircle law as n increases ($n = 100$ should be ample). Is the same true for the Hermite $-$ 1 ensemble? Why or why not?

2. Consider the tridiagonal matrix $T \in \mathsf{Jac}(n)$ that has entries $a_j = 0$, $1 \leq j \leq n$, $b_k = 1$, $1 \leq k \leq n - 1$.

   (a) Compute explicitly the spectral measure using Chebyshev polynomials (compare T with the recurrence relations for the Chebyshev polynomials).

   (b) Plot histograms of two distributions related to T for $n = 100$: (i) the empirical distribution of eigenvalues ($\frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k}$); (ii) the spectral density $\sum_{k=1}^{n} u_k^2 \delta_{\lambda_k}$. Can you identify the limit in (i)?

(This exercise will be relevant for an enumeration problem relating Brownian excursion to the Riemann-ζ function).

3. Establish uniqueness and smoothness in the proof of Theorem 26.

4. Use equation (3.4.12) to recursively expand τ_n as a continued fraction. Combine this with the uniqueness step in Q.2 to deduce an alternative approach to Theorem 26 that avoids the theory of orthogonal polynomials.

5. The following property of the function $-z^{-1}$ is relevant in the continued fraction scheme. Symmetric matrices have a partial order: Given A, $B \in \mathsf{Symm}(n)$ we say that $A \geq B$ if $u^T A u \geq u^T B u$ for every $u \in \mathbb{R}^n$. Suppose $A \geq B \geq 0$. Show that $-A^{-1} \geq -B^{-1}$.

6. This problem is a follow-up to exercise 5 in HW 1. Given a map f as in that exercise, compute an (explicit) expression for its derivative Df.

7. Compute the following normalization constants:

   (a) The normalization constants $Z_{n,\beta}$ in the standard definitions of GOE, GUE and GSE with exponential weight $e^{-\frac{\beta}{4} \mathrm{Tr}(M^2)}$.

   (b) The constant $C_{n,\beta}$ in (3.5.13).

   (c) The constant C_n in the Jacobian for ensembles (3.3.3) (compare with your calculation of the volume of the unit sphere in HW1).

8. The proofs of Dumitriu and Edelman finesse the following issue: given $T \in \mathsf{Jac}(n)$ it requires some care to find a decomposition for the tangent space $T_T \mathsf{Jac}(n)$, especially the isospectral manifold, $\mathcal{M}_T$, that is analogous to Lemma 2. As in that lemma, we may split $T_T \mathsf{Jac}(n)$ into orthogonal subspaces that correspond to diagonal matrices $\dot{\Lambda}$ and $Q^T \dot{Q} \in \mathfrak{o}(n)$. However, while each $Q^T \dot{Q} \in \mathfrak{o}(n)$ generates a curve in $T_T \mathsf{Symm}(n)$, not all $Q^T \dot{Q}$ give rise to curves in $T_T \mathsf{Jac}(n)$. Verify this. Explore this issue further by trying to find a basis for the isospectral manifold $\mathcal{M}_T$ (see equation (3.4.3)).

3.7 Exercises

3.8 Notes

To include in improved version.

1. Tridiagonal matrices as weights in enumerations problems.

2. Example: Chebyshev polynomials, Brownian excursion as a scaling limit of Dyck paths and relation with ζ-function.

Chapter 4

Determinantal formulas: From Vandermonde to Fredholm

Our purpose in this section is to present the elegant determinantal formulas of Dyson, Gaudin and Mehta for invariant matrix ensembles on $\mathsf{Her}(n)$. These formulas combine three distinct elements: (i) the Weyl formula on $\mathsf{Her}(n)$; (ii) the theory of orthogonal polynomials; (iii) Fredholm determinants. We first introduce these formulas for GUE. We then use the asymptotic properties of Hermite polynomials to establish their scaling limits (Theorem 2, Theorem 5 and Theorem 8). While the eigenvalues of GOE and GSE do not have a determinantal structure, they have a related Pfaffian structure, which is described in a later chapter.

4.1 Probabilities as determinants

In what follows we will adopt the following notation. In order to avoid confusion, we let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ denote the *unordered* eigenvalues of M, and $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathcal{W}^n$ denote the *ordered* eigenvalues of M. The probability density of x, denoted $P_n(x_1, \ldots, x_n)$, is obtained from the Weyl formula.

$$P^{(n)}(x_1, \ldots, x_n) = \frac{1}{Z_n} \triangle(x)^2 e^{-\frac{1}{2} \sum_{k=1}^{n} x_k^2}. \tag{4.1.1}$$

Observe that $P^{(n)}$ is invariant under permutations $(x_1, \ldots, x_n) \mapsto (x_{\sigma_1}, \ldots, x_{\sigma_n})$, $\sigma \in S_n$. In practice, our interest lies not in the joint density of all n eigenvalues, but statistics such as the law of the largest eigenvalue. Thus, what is required is an analytical technique to extract such information from (4.1.1) by integrating out degrees of freedom to obtain information on the joint distribution of m-eigenvalues, $1 \leq m \leq n$. More precisely, given m and a Borel function

$f : \mathbb{R}^m \to \mathbb{R}$ that is symmetric under permutations,

$$f(x_1, \ldots, x_m) = f(x_{\sigma_1}, \ldots, x_{\sigma_m}), \quad \sigma \in S_m, \qquad (4.1.2)$$

we consider random variables of the type

$$N_f = \sum_{(j_1, \ldots, j_m) \in [1,n]^m, \, j_k \text{distinct}} f(x_{j_1}, \ldots, x_{j_m}). \qquad (4.1.3)$$

Expectations of random variables of the form (4.1.3) are given by

$$\mathbb{E}(N_f) = \int_{\mathbb{R}^m} f(x_1, \ldots, x_m) \, R_m^{(n)}(x_1, \ldots, x_m) \, dx_1 \ldots dx_m, \qquad (4.1.4)$$

where we have defined the *m-point correlation function*

$$R_m^{(n)}(x_1, \ldots, x_m) \qquad\qquad\qquad (4.1.5)$$
$$= \frac{n!}{(n-m)!} \int_{\mathbb{R}^{n-m}} P^{(n)}(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) \, dx_{m+1} \ldots dx_m.$$

The combinatorial factor in (4.1.3) arises as follows. There are $\binom{n}{m}$ ways of picking subsets of m distinct indices from $[1, n]$. On the other hand,

$$R_m^{(n)}(x_1, \ldots, x_m) = R_m^{(n)}(x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_m}), \quad \sigma \in S_m. \qquad (4.1.6)$$

and the integral on the right hand side of (4.1.6) appears $m!$ times when integrating over the complementary $n - m$ variables for each choice of indices $\{j_1, \ldots, j_m\} \in [1, n]^m$.

Theorem 32. *The joint density and m-point functions for* $\mathrm{GUE}(n)$ *are*

$$P^{(n)}(x_1, \ldots, x_n) = \frac{1}{n!} \det\left(K^{(n)}(x_j, x_k)_{1 \leq j,k \leq n} \right), \qquad (4.1.7)$$

$$R_m^{(n)}(x_1, \ldots, x_m) = \det\left(K^{(n)}(x_j, x_k)_{1 \leq j,k \leq m} \right), \qquad (4.1.8)$$

where the integral kernel $K^{(n)}$ *is defined by the Hermite wave functions*

$$K_n(x, y) = \sum_{k=0}^{n-1} \psi_k(x)\psi_k(y). \qquad (4.1.9)$$

Remark 33. The kernel K_n may be simplified using identities for the Hermite polynomials. The Christoffel-Darboux formula (B.2.6) allows us to write

$$K_n(x, y) = \sqrt{n} \frac{\psi_n(x)\psi_{n-1}(y) - \psi_n(x)\psi_{n-1}(y)}{x - y}. \qquad (4.1.10)$$

Further, eliminating ψ_{n-1} with the identity (B.2.4) yields

$$K_n(x, y) = \frac{\psi_n(x)\psi_n'(y) - \psi_n(x)\psi_{n'}(y)}{x - y} - \frac{1}{2}\psi_n(x)\psi_n(y). \qquad (4.1.11)$$

A particular consequence of Theorem 32 is the following fundamental formula. Assume S is a bounded Borel set, let $\mathbf{1}$ denote its indicator function, and let $A_m(S)$ denote the probability that the set S contains precisely m eigenvalues for $M \in \mathrm{GUE}(n)$.

Theorem 34. *The generating function of $\{A_m(S)\}_{m=0}^{\infty}$ is given by the formula*

$$\det\left(I - zK^{(n)}\mathbf{1}_S\right) = \sum_{m=0}^{\infty} A_m(S)(1-z)^m, \quad z \in \mathbb{C}, \tag{4.1.12}$$

where $\det\left(I - zK^{(n)}\mathbf{1}_S\right)$ *denotes the Fredholm determinant of the kernel*

$$K^{(n)}\mathbf{1}_S(x,y) = \sum_{k=0}^{n-1} \mathbf{1}_S(x)\psi_k(x)\psi_k(y)\mathbf{1}_S(y). \tag{4.1.13}$$

Theorem 32 and Theorem 34 illustrate the general spirit of determinantal formulas in random matrix theory. The density of a joint distribution is expressed as a determinant of an integral operator with finite rank. One may then use the theory of orthogonal polynomials, in particular, results on the asymptotics of orthogonal polynomials, to establish the basic limit theorems outlined in Chapter 1 (see Theorems 35 and Theorem 36 below).

Appendices B and C provide brief introductions to Hermite polynomials and Fredholm determinants respectively.

4.2 The m-point correlation function

Proof of Theorem 32. We form linear combinations of the rows of the Vandermonde matrix to obtain

$$\triangle(x) = \det \begin{pmatrix} \mathfrak{h}_0(x_1) & \mathfrak{h}_0(x_2) & \dots & \mathfrak{h}_0(x_n)\dots \\ \mathfrak{h}_1(x_1) & \mathfrak{h}_1(x_2) & \dots & \mathfrak{h}_1(x_n)\dots \\ \vdots & \vdots & & \vdots \\ \mathfrak{h}_{n-1}(x_1) & \mathfrak{h}_{n-1}(x_2) & \dots & \mathfrak{h}_{n-1}(x_n)\dots \end{pmatrix}. \tag{4.2.1}$$

The calculations above would apply to any set of monic polynomials of degree $0, 1, 2, \dots, n-1$. The Hermite polynomials and wave functions are relevant because they satisfy the orthogonality relations

$$\int_{\mathbb{R}} \mathfrak{h}_j(x)\mathfrak{h}_k(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \delta_{jk}k!, \tag{4.2.2}$$

and allow the inclusion of an exponential weight. Precisely, the Hermite wave-functions

$$\psi_k(x) = \frac{1}{\sqrt{k!}} P_k(x)\frac{e^{-x^2/4}}{(2\pi)^{1/4}}, \tag{4.2.3}$$

satisfy the orthogonality relation

$$\int_{\mathbb{R}} \psi_j(x)\psi_k(x)\,dx = \delta_{jk}, \tag{4.2.4}$$

and form a basis for $\mathrm{L}^2(\mathbb{R})$. Let H denote the matrix with entries $H_{jk} = \psi_{j-1}(x_k)$. It follows from (4.2.1) and (4.2.3) that

$$e^{-\frac{|x|^2}{2}}\triangle(x)^2 = \det H^2 = \det H^T H = \det\left(K^{(n)}(x_j, x_k)_{1\leq j,k\leq n}\right), \tag{4.2.5}$$

using the identity

$$\left(H^T H\right)_{jk} = \sum_{l=1}^{n} H_{lj}H_{lk} = \sum_{l=0}^{n-1} \psi_l(x_j)\psi_l(x_k) = K^{(n)}(x_j, x_k). \tag{4.2.6}$$

Therefore, the joint density $P^{(n)}(x)$ is proportional to $\det K^{(n)}$. To determine the constant of proportionality we evaluate the integral

$$\int_{\mathbb{R}^n} \det(H)^2 dx_1 \dots dx_n = \int_{\mathbb{R}^n} \left(\det\left[\psi_{j-1}(x_k)\right]_{1\leq j,k\leq n}\right)^2 dx_1 \dots dx_n$$

$$= \sum_{\sigma,\tau\in S_n} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \int_{\mathbb{R}^n} \prod_{j=1}^{n} \psi_{\sigma_j-1}(x_j)\psi_{\tau_j-1}(x_j)\,dx_1 \dots dx_n$$

$$\tag{4.2.7}$$

$$= \sum_{\sigma,\tau\in S_n} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \prod_{j=1}^{n} \delta_{\sigma_j,\tau_j} = \sum_{\sigma,\tau\in S_n} \mathbf{1}_{\{\sigma=\tau\}} = n!\,.$$

We combine (4.2.7) and (4.2.6) to obtain the first assertion in Theorem 32:

$$P^{(n)}(x_1,\dots,x_n) = \frac{1}{n!}\det\left(K^{(n)}(x_j, x_k)_{1\leq j,k\leq n}\right).$$

The formulas for the correlation functions may be obtained by induction, beginning with

$$R^{(n)}(x_1,\dots,x_n) = \det\left[K^{(n)}(x_j, x_k)\right]_{1\leq j,k\leq n}. \tag{4.2.8}$$

First, the orthonormality relations (4.2.4) imply

$$\int_{\mathbb{R}} K^{(n)}(x, x) = n, \quad \int_{\mathbb{R}} K^{(n)}(x, z)K^{(n)}(z, y)\,dz = K^{(n)}(x, y). \tag{4.2.9}$$

Assume (4.1.8) holds for an index $m+1\leq n$. We then have

$$
\begin{aligned}
R_m^{(n)}(x_1,\dots,x_m) &= \frac{1}{n-m}\int_{\mathbb{R}} R_{m+1}^{(n)}(x_1,\dots,x_m,x_{m+1})\,dx_{m+1}\\
&= \frac{1}{n-m}\int_{\mathbb{R}} \det\left[K^{(n)}(x_j, x_k)\right]_{1\leq j,k\leq m+1} dx_{m+1}\\
&= \frac{1}{n-m}\sum_{\sigma\in S_{m+1}} \mathrm{sgn}(\sigma)\int_{\mathbb{R}} K^{(n)}(x_1, x_{\sigma_1})\dots K^{(n)}(x_{m+1}, x_{\sigma_{m+1}})\,dx_{m+1}.
\end{aligned}
$$

If $\sigma_{m+1} = m + 1$ in this sum, then the first equality in (4.2.9) implies

$$\int_{\mathbb{R}} K^{(n)}(x_1, x_{\sigma_1}) \cdots K^{(n)}(x_{m+1}, x_{\sigma_{m+1}}) \, dx_{m+1} \qquad (4.2.10)$$

$$= n \, K^{(n)}(x_1, x_{\sigma_1}) \cdots K^{(N)}(x_m, x_{\sigma_m}).$$

If $\sigma_{m+1} \neq m + 1$, there exists $j \leq m$ and $k \leq m$ such that $\sigma_j = m + 1$ and $\sigma_{m+1} = k$. We then use the second equality in (4.2.9) to find

$$\int_{\mathbb{R}} K^{(n)}(x_1, x_{\sigma_1}) \cdots K^{(n)}(x_{m+1}, x_{\sigma_{m+1}}) \, dx_{m+1} \qquad (4.2.11)$$

$$= K^{(n)}(x_1, x_{\sigma'_1}) \cdots K^{(n)}(x_m, x_{\sigma'_m}).$$

where σ' is a permutation of $\{1, \ldots, m\}$ such that $\sigma'_j = k$ and $\sigma'_l = \sigma_l$ if $l \neq j$. Each permutation $\sigma' \in S_m$ may come from m permutations $\sigma \in S_{m+1}$. Further, $\mathrm{sgn}(\sigma') = -\mathrm{sgn}(\sigma)$ since these permutations differ by a single swap. Therefore, using equations (4.2.10) and (4.2.11) we have

$$\int_{\mathbb{R}} \det \left[K^{(n)}(x_j, x_k) \right]_{1 \leq j,k \leq m+1} dx_{m+1} = (n - m) \det \left[K^{(n)}(x_j, x_k) \right]_{1 \leq j,k \leq m}.$$

$$\square$$

4.3 Determinants as generating functions

Proof of Theorem 34. The Fredholm determinant $\det \left(I - z K^{(n)} \mathbf{1} \right)$ is an entire function of z. Thus, equation (4.1.12) is equivalent to the statement

$$A_m(S) = \frac{1}{m!} \left(-\frac{d}{dz} \right)^m \det \left(I - z K^{(n)} \mathbf{1} \right) \Big|_{z=1}. \qquad (4.3.1)$$

We first prove formula (4.3.1) in the case $m = 0$. Let $\mathbf{1}$ denote the characteristic function of the set S. The probability that all eigenvalues lie outside S is given by

$$\int_{\mathbb{R}^n} \left(\prod_{j=1}^{n} (1 - \mathbf{1}(x_j)) \right) P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \qquad (4.3.2)$$

$$= \sum_{j=0}^{n} (-1)^j \int_{\mathbb{R}^n} \rho_j(\mathbf{1}(x_1), \ldots, \mathbf{1}(x_n)) P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n,$$

where $\rho_j^n(x_1, \ldots, x_n)$ is the j-th symmetric function in n variables. For example,

$$\rho_0^n(x) = 1, \quad \rho_1^n(x) = \sum_{j=1}^{n} x_j, \quad \rho_2^n(x) = \sum_{j<k}^{n} x_j x_k, \quad \rho_n^n(x) = \prod_{j=1}^{n} x_j.$$

Since each term in ρ_j^n consists of j products of the form $\mathbf{1}(x_{\sigma_j})$ for some permutation $\sigma \in S_n$, we integrate over the remaining $n - j$ variables, and use the permutation invariance of $R_j^{(n)}$ to obtain

$$\int_{\mathbb{R}^n} \rho_j^n(\mathbf{1}(x_1), \ldots, \mathbf{1}(x_n)) P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \qquad (4.3.3)$$

$$= \frac{(n-j)!}{n!} \binom{n}{j} \int_{\mathbb{R}^j} \mathbf{1}(x_1) \ldots, \mathbf{1}(x_j) R_j^{(n)}(x_1, \ldots, x_j) \, dx_1 \ldots dx_j,$$

$$= \frac{1}{j!} \int_{\mathbb{R}^j} \det \left(K^{(n)} \mathbf{1}(x_k, x_l)_{1 \leq k, l \leq j} \right) \, dx_1 \ldots dx_j. \qquad (4.3.4)$$

In the last equality, we have used (4.1.8) and multiplied the kernel on the left and right by the diagonal matrix $D = \operatorname{diag}(\mathbf{1}(x_1), \ldots, \mathbf{1}(x_j))$, so that

$$\mathbf{1}(x_1) \ldots \mathbf{1}(x_j) R_j^{(n)}(x_1, \ldots, x_j) = \mathbf{1}^2(x_1) \ldots \mathbf{1}^2(x_j) R_j^{(n)}(x_1, \ldots, x_j)$$

$$= \det \left(D(x_k) K(x_k, x_l) D(x_l)_{1 \leq k, l \leq j} \right) = \det \left(K^{(n)} \mathbf{1}(x_k, x_l)_{1 \leq k, l \leq j} \right),$$

where $K^{(n)} \mathbf{1}$ is defined in (4.1.13). We now combine (4.3.2) and (4.3.4) to obtain

$$\sum_{j=0}^{n} (-1)^j \int_{\mathbb{R}^n} \rho_j(\mathbf{1}(x_1), \ldots, \mathbf{1}(x_n)) P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n$$

$$= \det(I - K^{(n)} \mathbf{1}), \qquad (4.3.5)$$

using the infinite series (C.1.8) for the Fredholm determinant (only n terms are non-zero, since $K^{(n)}$ has rank n).

We now turn to the case $m \geq 1$. Equation (4.3.2) must now be modified to allow exactly m eigenvalues within S and $n - m$ eigenvalues ouside S. Since there are $\binom{n}{m}$ ways of selecting m distinct eigenvalues, we find

$$A_m(S) = \qquad (4.3.6)$$

$$\binom{n}{m} \int_{\mathbb{R}^n} \prod_{j=1}^{m} \mathbf{1}(x_j) \prod_{k=1}^{n-m} (1 - \mathbf{1}(x_{m+k})) \, P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n$$

$$\binom{n}{m} \sum_{k=0}^{n-m} (-1)^k \int_{\mathbb{R}^n} \rho_k^{(n-m)} (\mathbf{1}(x_{m+1}), \ldots, \mathbf{1}(x_n)) \, P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n.$$

As in (4.3.3) and (4.3.4) we use the fact that ρ_k^{n-m} consists of a product of k

terms all of which integrate to the same value. Thus, the sum above is

$$\binom{n}{m} \sum_{k=0}^{n-m} (-1)^k \binom{n-m}{k} \int_{\mathbb{R}^n} \prod_{j=1}^{m+k} \mathbf{1}(x_j) \, P^{(n)}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n$$

$$= \frac{1}{m!} \sum_{k=0}^{n-m} \frac{(-1)^k}{k!} \int_{\mathbb{R}^{m+k}} \prod_{j=1}^{m+k} \mathbf{1}(x_j) R_{m+1}^{(n)}(x_1, \ldots, x_{m+k}) \, dx_1 \ldots dx_{m+k}$$

$$= \frac{1}{m!} \sum_{k=0}^{n-m} \frac{(-1)^k}{k!} \int_{\mathbb{R}^{m+k}} \det\left(K\mathbf{1}(x_p, x_q)_{1 \le p,q \le m+k}\right) \, dx_1 \ldots dx_{m+k}$$

$$= \frac{1}{m!} \left(-\frac{d}{dz}\right)^m \det(I - zK\mathbf{1})\big|_{z=1}. \tag{4.3.7}$$

In the second equality, we have simplified the combinatorial factors as follows:

$$\frac{1}{m!} = \binom{n}{m}\binom{n-m}{k}\frac{n!}{(n-m-k)!}.$$

In the last line, we have used formula (C.1.11) for the derivative of a Fredholm determinant. $\qquad \square$

4.4 Scaling limits I: the semicircle law

The empirical measure of the eigenvalues

$$L_n(dx) = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k}(dx) \tag{4.4.1}$$

has the expected density

$$\mathbb{E}L_n(dx) = \frac{1}{n} K_n(x, x) \, dx. \tag{4.4.2}$$

The above expression is somewhat more transparent in its weak form. For every $f \in C_0(\mathbb{R})$, we have

$$\mathbb{E}\langle L_n, f \rangle = \frac{1}{n} \int_{\mathbb{R}} f(x) R_1^{(n)}(x) \, dx = \frac{1}{n} \int_{\mathbb{R}} f(x) K_n(x, x) \, dx, \tag{4.4.3}$$

by Theorem 32 and equation (4.1.8). The value of the kernel K_n on the diagonal is determined by the Christoffel-Darboux relation (4.1.10) and L'Hospital's lemma:

$$K_n(x, x) = \sqrt{n}\left(\psi_n'(x)\psi_{n-1}(x) - \psi_n(x)\psi_{n-1}'(x)\right). \tag{4.4.4}$$

The scaling limit of $\mathbb{E}L_n$ is the semicircle law defined in (1.2.1).

Lemma 12.

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} K_n\left(x\sqrt{n}, x\sqrt{n}\right) = p_{\mathrm{sc}}(x), \quad x \in \mathbb{R}. \tag{4.4.5}$$

Further, for any $\varepsilon > 0$, the convergence is uniform on the set $\{x \,||x - 2| \ge \varepsilon\}$.

Proof. The lemma follows from the Plancherel-Rotach asymptotics for the Hermite wave functions (see Cases 1 and 2 and equations (B.4.1)–(B.4.4)) in Appendix B). Define the rescaled wave functions

$$\Psi_{n+p}(x) = n^{\frac{1}{4}} \psi_{n+p}(x\sqrt{n}), \quad p = -2, -1, 0. \tag{4.4.6}$$

We use the identity (B.2.4) to eliminate ψ'_n and ψ'_{n-1} from (4.4.4) and find after a few computations that

$$\frac{1}{\sqrt{n}} K_n\left(x\sqrt{n}, x\sqrt{n}\right) = \Psi_{n-1}^2(x) - \sqrt{\frac{n-1}{n}} \Psi_{n-2}(x)\Psi_n(x). \tag{4.4.7}$$

We now use the asymptotic relations (B.4.2) and (B.4.4) depending on whether $|x| < 2$ or $|x| > 2$. Since the region $|x| > 2$ corresponds to exponential decay with a rate proportional to n, we focus on the region $|x| < 2$. In order to simplify notation, let

$$\theta = n\left(\varphi - \frac{1}{2}\sin 2\varphi\right) - \frac{1}{2}\varphi - \frac{\pi}{4}. \tag{4.4.8}$$

(This is the argument of the cosine in (B.4.16) when $p = -1$.) Then (4.4.7) and (B.4.2) yield

$$\frac{1}{\sqrt{n}} K_n\left(x\sqrt{n}, x\sqrt{n}\right)$$

$$\sim \frac{1}{\pi \sin \varphi}\left(\cos^2 \theta - \cos(\theta + \varphi)\cos(\theta - \varphi)\right) = \frac{1}{2\pi}\sqrt{4 - x^2},$$

using $x = 2\cos\varphi$ and the trigonometric formula $\cos 2\alpha = 2\cos^2\alpha - 1$. A similar calculation with (B.4.4) shows that the limit vanishes outside the set $|x| > 2$. The assertion of uniformity in the convergence follows from the assertion of uniform convergence in the Plancherel-Rotach asymptotics. $\qquad\square$

4.5 Scaling limits II: the sine kernel

Recall from Definition 4 that K_{sine} is the integral kernel on $\mathbb{R} \times \mathbb{R}$ given by

$$K_{\mathsf{sine}}(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)}, \quad x \neq y, \tag{4.5.1}$$

and $K_{\mathsf{sine}}(x, x,) = 1$. It defines an integral operator on $L^2(S)$ for every bounded, measurable set S. We can now prove a stronger version of Theorem 5.

Theorem 35. *For each integer $m = 0, 1, 2, \ldots$ and bounded, measurable set S*

$$\lim_{n\to\infty} \mathbb{P}\left(M \in \mathrm{GUE}(n) \text{ has } m \text{ eigenvalues in } \sqrt{n}S\right)$$

$$= \frac{1}{m!}\left(-\frac{d}{dz}\right)^m \det\left(I - zK_{\mathsf{sine}}\mathbf{1}_S\right)|_{z=1}. \tag{4.5.2}$$

The proof of Theorem 35 is a consequence of the following

Lemma 13. *Let S be a bounded measurable set. Then*

$$\lim_{n\to\infty} \sup_{x,y\in S} \left| K_n\left(\frac{x}{\sqrt{n}}, \frac{y}{\sqrt{n}}\right) - K_{\mathsf{sine}}(x,y) \right| = 0. \tag{4.5.3}$$

Proof. Let us define the rescaled wave function

$$\Psi_n(x) = n^{\frac{1}{4}} \psi_n\left(\frac{x}{\sqrt{n}}\right). \tag{4.5.4}$$

The identity (B.2.4) now takes the form

$$\Psi_{n-1}(x) = \Psi_n'(x) + \frac{x}{2n}\Psi_n(x), \tag{4.5.5}$$

and the rescaled kernel takes the form

$$K_n\left(\frac{x}{\sqrt{n}}, \frac{y}{\sqrt{n}}\right) = \frac{\Psi_n(x)\Psi'(y) - \Psi_n'(x)\Psi_n(y)}{x-y} - \frac{1}{2n}\Psi_n(x)\Psi_n(y). \tag{4.5.6}$$

We now use (B.3.4) (when n is even) and (B.3.4) (when n is odd) to obtain (4.5.1). $\qquad\square$

Proof of Theorem 35. Let $\tilde{K}_n(x,y)$ denote the rescaled kernel $K_n(\frac{x}{\sqrt{n}}, \frac{y}{\sqrt{n}})$. It follows from Lemma 13 and the definition of the Fredholm determinant that

$$\lim_{n\to\infty} \det\left(I - z\tilde{K}_n\mathbf{1}_S\right) = \det\left(I - zK_{\mathsf{sine}}\mathbf{1}_S\right), \quad z \in \mathbb{C}, \tag{4.5.7}$$

and that the convergence is uniform in n for z in a bounded set. In particular, the derivatives at $z = 1$ converge for all m, that is

$$\lim_{n\to\infty} \left(-\frac{d}{dz}\right)^m \det\left(I - z\tilde{K}_n\mathbf{1}_S\right)\Big|_{z=1} = \left(-\frac{d}{dz}\right)^m \det\left(I - zK_{\mathsf{sine}}\mathbf{1}_S\right)\big|_{z=1}. \tag{4.5.8}$$

By Theorem 34, this is equivalent to (4.5.2). $\qquad\square$

4.6 Scaling limits III: the Airy kernel

Recall from Definition 7 that K_{Airy} is the continuous integral kernel on $\mathbb{R} \times \mathbb{R}$ given by

$$K_{\mathsf{Airy}}(x,y) = \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y}, \quad x \neq y. \tag{4.6.1}$$

The fluctuations at the edge of the spectrum are described as follows. Let $(x_1, \ldots, x_n)$ denote the unordered eigenvalues of a matrix $M \in \mathrm{GUE}(n)$ and let us consider the shifted and rescaled points

$$s_k = n^{\frac{1}{6}}\left(x - 2\sqrt{n}\right), \quad k = 1, \ldots, n. \tag{4.6.2}$$

For each nonnegative integer m and bounded, measurable set S, let $B_m^{(n)}(S)$ denote the probability that exactly m of the points $s_1, \ldots, s_n$ lie in S when $M \in \mathrm{GUE}(n)$.

Theorem 36.

$$\lim_{n\to\infty} B_m^{(n)}(S) = \frac{1}{m!}\left(-\frac{d}{dz}\right)^m \det\left(I - zK_{\mathsf{Airy}}\mathbf{1}_S\right)\big|_{z=1}. \qquad (4.6.3)$$

Remark 37. The assumption that S is bounded is necessary for K_{sine}. The sine-kernel has a (weak) rate of decay $|x|^{-1}$ as $|x| \to \infty$ and the Fredholm determinant $\det(I - zK_{\mathsf{sine}}\mathbf{1}_S)$ is not finite unless S is bounded. However, the Airy function, and the thus the Airy kernel, has strong decay as x and $y \to \infty$. The Fredholm determinant $\det(I - zK_{\mathsf{Airy}}\mathbf{1}_S)$ is well-defined in $L^2(S)$ for sets S that are bounded below, but not above, such as $S = (a,\infty)$ for any $a \in \mathbb{R}$. Such sets will be considered when we compute the Tracy-Widom distribution. See Exercise 5.

The proof of Theorem 36 follows from the Plancherel-Rotach asymptotics for the Hermite polynomials, in particular the Airy asymptotics in the transition zone (see Case 3 and (B.4.5)–(B.4.6) in Appendix B). The following lemma plays a role analogous to that of Lemma 13 in the proof of Theorem 35.

Lemma 14. *Let S be a bounded measurable set. Then*

$$\lim_{n\to\infty} \sup_{x,y\in S}\left|\frac{1}{n^{\frac{1}{6}}}K_n\left(2\sqrt{n} + \frac{x}{n^{\frac{1}{6}}}, 2\sqrt{n} + \frac{y}{n^{\frac{1}{6}}}\right) - K_{\mathsf{Airy}}(x,y)\right| = 0. \qquad (4.6.4)$$

Proof. Let us define the rescaled wave-functions

$$\Psi_n(x) = n^{\frac{1}{12}}\psi_n\left(2\sqrt{n} + \frac{x}{n^{\frac{1}{6}}}\right). \qquad (4.6.5)$$

We then use identity (B.2.4) to find

$$n^{\frac{1}{6}}K_n\left(2\sqrt{n} + \frac{x}{n^{\frac{1}{6}}}, 2\sqrt{n} + \frac{y}{n^{\frac{1}{6}}}\right)$$
$$= \frac{\Psi_n(x)\Psi_n'(y) - \Psi_n'(x)\Psi_n(y)}{x - y} - \frac{1}{2n^{\frac{1}{3}}}\Psi_n(x)\Psi_n(y). \qquad (4.6.6)$$

As noted in Appendix B, as $n \to \infty$, $\Psi_n(x)$ converges to $\mathrm{Ai}(x)$ and the convergence is uniform for x in compact subsets of $\mathbb{C}$. Thus, in addition, $\Psi_n'(x) \to \mathrm{Ai}'(x)$ uniformly in compact sets, and (4.6.4) follows. $\qquad \square$

4.7 Exercises

1. Prove the Christoffel-Darboux identity (B.1.7) for Hermite polynomials. (This is a standard relation and it is easy to find a proof in many texts, but try to do it on your own.)

2. Use the method of steepest descent to establish the asymptotic formula (A.3.1) for the Airy function. This is an easy application of the method of steepest descent.

3. In order to appreciate the power of the Plancherel-Rotach asymptotics, some numerical calculations will help.

(a) Develop a numerical scheme to compute all the roots of the n-th Hermite polynomial $\mathfrak{h}_n$. Plot the empirical distribution of roots for $n = 100$. Can you determine the limiting density of suitably rescaled roots?

(b) Numerically compute the Hermite wave functions for large n, say $n = 100$, and compare the rescaled wave function with the Plancherel-Rotach asymptotic formulas in all three regions (oscillatory, decaying and transition).

4. Use the method of steepest descent to establish the Plancherel-Rotach asymptotics in the region of exponential decay (equation (B.4.4)). This requires more care than Q.2.

5. Establish the following a priori bound on the Airy kernel. For any $a \in \mathbb{R}$,

$$\sup_{x,y} e^{x+y} |K_{\mathsf{Airy}}(x,y)| < \infty. \tag{4.7.1}$$

Let S be the semi-infinite interval (a, ∞). Use the above estimate to establish that the Fredholm determinant $\det(I - zK_{\mathsf{Airy}}\mathbf{1}_S)$ is an entire function.

4.8 Notes

To include in improved version.

1. Moment estimates to strengthen convergence to semicircle law.

2. Definition of determinantal processes.

3. Pair correlation function for the sine kernel.

Chapter 5

The equilibrium measure

In this section we establish properties of the equilibrium measure for general invariant ensembles. We also relate the equilibrium measure to the classical theory of orthogonal polynomials and Fekete points.

5.1 The log-gas

Let $V : \mathbb{R} \to \mathbb{R}$ denote a potential such that $V(x) \to \infty$ sufficiently rapidly as $|x| \to \infty$. The log-gas with size n and potential nV is a system of n identical charged particles constrained to the line interacting via pairwise Coulomb repulsion and the potential nV (we have scaled the potential V by n in order to ensure a scaling limit). The total energy of the system in any configuration $x \in \mathbb{R}^n$ is given by

$$E(x) = n \sum_{j=1}^{n} V(x_j) + \frac{1}{2} \sum_{j \neq k} \log \frac{1}{|x_j - x_k|}. \tag{5.1.1}$$

A fundamental postulate of equilibrium statistical mechanics is that the probability density of finding the system in a state x at inverse temperature $\beta > 0$ is

$$\frac{1}{Z_{n,V}(\beta)} e^{-\beta E(x)}, \tag{5.1.2}$$

where $Z_{n,V}$ is the partition function

$$Z_{n,V}(\beta) = \int_{\mathbb{R}^n} e^{-\beta E(x)} \, Dx. \tag{5.1.3}$$

The log-gas provides us with a physical caricature of eigenvalue repulsion. On one hand, we see that the energy $E(x)$ has two complementary terms: the logarithmic potential drives charges apart, but the potential V confines them in space. On the other hand, let V define an invariant probability measure of the form (1.1.3) on $\mathsf{Symm}(n)$, $\mathsf{Her}(n)$ or $\mathsf{Quart}(n)$. As a consequence of Weyl's

formula (Theorem 13), the equilibrium density (5.1.2) is precisely the joint law of the eigenvalues for these ensembles at $\beta = 1$, 2 and 4 respectively. It is in this sense that the 'eigenvalues repel'.

We have scaled the energy V with n in (5.1.1) in order to obtain a simple description of the scaling limit when $n \to \infty$. In order to study this limit, we view the energy function as a functional of the empirical measure, L_n, rather than a configuration $x \in \mathbb{R}^n$. For $(r, s) \in \mathbb{R}^2$ let

$$e(r, s) = \frac{1}{2}V(r) + \frac{1}{2}V(s) + \log \frac{1}{|r - s|}, \tag{5.1.4}$$

and given a probability measure μ on the line, define the functional

$$I[\mu] = \int_{\mathbb{R}} \int_{\mathbb{R}} e(r, s)\, \mu(dr)\, \mu(ds). \tag{5.1.5}$$

Observe that if L_n is the empirical measure associated to $x \in \mathbb{R}^n$, then

$$E(x) = n^2 \left(\frac{1}{n} \sum_{j=1}^{n} V(x_j) + \frac{1}{n^2} \sum_{j \neq k} \log \frac{1}{|x_j - x_k|} \right) = n^2 \tilde{I}[L_n], \tag{5.1.6}$$

and we may rewrite the partition function in the form

$$Z_{n,V}(\beta) = \int_{\mathbb{R}^n} e^{-n^2 \beta \tilde{I}[L_n]}\, Dx. \tag{5.1.7}$$

Here $\tilde{I}[L_n]$ denotes the renormalized functional

$$\tilde{I}[\mu] = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{r \neq s} e(r, s)\, \mu(dr)\, \mu(ds), \tag{5.1.8}$$

that takes into account all interaction terms in $I[\mu]$, except the singular self-interaction term from $I[\mu]$. The logarithmic singularity in $e(r, s)$ is integrable if $\mu(ds)$ has an absolutely continuous density. Thus, if the particles in the log-gas spread out sufficiently as $n \to \infty$, we expect that μ has a smooth density, and

$$\lim_{n \to \infty} \frac{1}{n^2} \log Z_{n,V}(\beta) = \min_{\mu} I[\mu]. \tag{5.1.9}$$

In order to establish this relation, it is first necessary to obtain a precise analytical understanding of this minimization problem. We first prove such results under the formal assumption that there exists an $R > 0$ such that $V(x) = +\infty$ for $|x| > R$. This simply means that we first restrict attention to measures with support within the interval $[-R, R]$. Once the ideas are clear in this setting, we turn to measures with support on the line.

5.2 Energy minimization for the log-gas

5.2.1 Case 1: bounded support

Let $\mathcal{P}_R$ denote the set of probability measures on the interval $[-R, R]$. Recall that the natural topology on $\mathcal{P}_R$ is the weak topology (we adopt the probabilists convention for what is conventionally termed the weak-$*$ topology). A sequence of measures $\{\mu_k\}_{k=1}^{\infty} \in \mathcal{P}_R$ converges weakly to $\mu \in \mathcal{P}_R$ if

$$\lim_{n \to \infty} \langle \mu_n, f \rangle = \langle \mu, f \rangle, \tag{5.2.1}$$

for every function $f \in C(\mathbb{R})$. This topology is natural, because it yields compactness by Helly's theorem: Each sequence $\{\mu_k\}_{k=1}^{\infty} \in \mathcal{P}_R$ has a subsequence that converges weakly to a measure in $\mathcal{P}_R$.

Theorem 38. *Assume V is a continuous function on $[-R, R]$. There exists a unique probability measure $\mu_* \in \mathcal{P}_R$ such that*

$$I[\mu_*] = \min_{\mu \in \mathcal{P}_R} I[\mu]. \tag{5.2.2}$$

The proof of Theorem 38 is a demonstration of the classical method of the calculus of variations. There are two distinct ideas at work: existence follows from the fact that the functional $I[\mu]$ is weakly lower semicontinuous; uniqueness follows from the fact that $I[\mu]$ is a strictly convex function on $\mathcal{P}_R$.

Lemma 15. *Suppose the sequence $\{\mu_n\}_{n=1}^{\infty} \in \mathcal{P}_R$ converges weakly to $\mu \in \mathcal{P}_R$. Then*

$$I[\mu] \le \liminf_{n \to \infty} I[\mu_n]. \tag{5.2.3}$$

Lemma 16. *Let $\mu_0 \ne \mu_1$ be two measures in $\mathcal{P}_R$ and let $\mu_\theta = (1 - \theta)\mu_0 + \theta\mu_1$ denote their convex combination for each $\theta \in (0, 1)$. Then*

$$I[\mu_\theta] < (1 - \theta)I[\mu_0] + \theta I[\mu_1]. \tag{5.2.4}$$

Proof of Theorem 38. Existence. Since V is bounded, the function $e(x, y)$ is bounded below on $[-R, R]$. Therefore, $\inf_{\mu \in \mathcal{P}_R} I[\mu] > -\infty$. Further, since the logarithmic singularity is integrable, $I[\mu] < \infty$ for any measure that is absolutely continuous. Thus, we may assume that there is a sequence of measures $\{\mu_k\}_{k=1}^{\infty}$ such that

$$\lim_{k \to \infty} I[\mu_k] = \inf_{\mu \in \mathcal{P}_R} I[\mu] < infty. \tag{5.2.5}$$

Since $\mathcal{P}_R$ is compact in the weak topology, we may extract a convergent subsequence, also labeled $\{\mu_k\}_{k=1}^{\infty}$ for simplicity. Let μ_* denote the weak limit of this subsequence. We then use Lemma 15 to obtain the chain of inequalities

$$\inf_{\mu \in \mathcal{P}_R} I[\mu] \le I[\mu_*] \le \liminf_{k \to \infty} I[\mu_k] = \inf_{\mu \in \mathcal{P}_R} I[\mu]. \tag{5.2.6}$$

Thus, μ_* is a minimizer.

Uniqueness. Assume μ_* and ν_* are two distinct minimizers. We apply Lemma 16 to their convex combination with $\theta = 1/2$ to obtain the contradiction

$$\inf_{\mu \in \mathcal{P}_R} I[\mu] \leq I[\tfrac{1}{2}\mu_* + \tfrac{1}{2}\nu_*] < \tfrac{1}{2}\left(I[\mu_*] + I[\nu_*]\right) = \inf_{\mu \in \mathcal{P}_R} I[\mu]. \qquad (5.2.7)$$

$$\square$$

5.2.2   Weak lower semicontinuity

We now turn to the proof of Lemma 15. We first observe that for each monomial $r^j s^k$ in the variables r and s, the quadratic functional

$$\mu \mapsto \int_{-R}^{R}\int_{-R}^{R} r^j s^k \, \mu(dr)\,\mu(ds) = \left(\int_{-R}^{R} r^j \mu(dr)\right)\left(\int_{-R}^{R} s^k \mu(ds)\right)$$

is weakly continuous since it is the product of two bounded linear functionals on $\mathcal{P}_R$. Since each polynomial $p(r,s)$ in the variables (r,s) is a finite sum of monomials, the functional

$$\mu \mapsto \int_{-R}^{R}\int_{-R}^{R} p(r,s)\mu(dr)\,\mu(ds)$$

is also weakly continuous. Finally, since each continuous function $f \in C([-R,R]^2)$ may be uniformly approximated by polynomials, the quadratic functional

$$\mu \mapsto \int_{-R}^{R}\int_{-R}^{R} f(r,s)\mu(dr)\,\mu(ds)$$

is weakly continuous.

The function $e(s,t)$ defined in (5.1.4) is *not* continuous on $[-R,R]^2$ since the logarithmic term is unbounded on the diagonal $s = t$. However, for any $M > 0$, the truncated function $e_M(r,s) = \min(e(r,s),M)$ is continuous. Thus, given a weakly convergent sequence of measures $\{\mu_k\}_{k=1}^{\infty}$ with limit $\mu \in \mathcal{P}_R$ we find

$$\int_{-R}^{R}\int_{-R}^{R} e_M(r,s)\mu(dr)\,\mu(ds) = \lim_{k\to\infty}\int_{-R}^{R}\int_{-R}^{R} e_M(r,s)\mu_k(dr)\,\mu_k(ds)$$

$$\leq \liminf_{k\to\infty}\int_{-R}^{R}\int_{-R}^{R} e(r,s)\mu_k(ds)\mu_k(ds) = \liminf_{k\to\infty} I[\mu_k].$$

We let $M \to \infty$ on the left hand side and use the monotone convergence theorem to obtain (5.2.3).

5.2.3   Strict convexity

Lemma 16 is a particular consequence of a general fact in potential theory. The essential idea is to recognize that the function $z \mapsto -\log|z|$ is the fundamental

solution to Laplace's equation in $\mathbb{C} \cong \mathbb{R}^2$. More precisely, given a signed measure μ with a smooth density $\rho(z)$, supported in the ball $B_R \subset \mathbb{C}$ the unique solution to Poisson's equation with Dirichlet boundary condition

$$-\triangle \psi = \mu, \quad z \in \mathbb{C} \backslash \Omega, \quad \psi(z) = 0, \quad |z| = R, \tag{5.2.8}$$

is given by the integral formula

$$\psi(z) = \int_{B_R} G(z, w) \rho(w) \, Dw, \quad z \in B_R, \tag{5.2.9}$$

where Dw denotes the two-dimensional area element in $\mathbb{C}$ and $G(z, w)$ is the Green's function for Poisson's equation in the ball B_R with Dirichlet boundary conditions,

$$G(z, w) = \frac{1}{2\pi} \log \left(\frac{|w|}{R} \frac{|z - w^R|}{|z - w|} \right), \quad w^R = \frac{R^2 w}{|w|^2}, \quad z, w \in B_R. \tag{5.2.10}$$

The function $G(z, w)$ is obtained by the method of images: the image point w^R is the reflection of the point $w \in B_R$ in the circle ∂B_R [12, §4.1]. What matters here is that the dominant term in the Green's function is the logarithmic term $-\log|z - w|$, just as in equation (5.1.5), and the positivity of

$$\int_{B_R} \int_{B_R} G(z, w) \, \mu(dz) \, \mu(dw) = -\int_{B_R} \psi(w) \triangle \psi(w) \, ds = \int_{B_R} |\nabla \psi(w)|^2 \, Dw > 0. \tag{5.2.11}$$

However, in contrast with (5.1.5) here we have assumed that $\mu(dw)$ has a smooth density $\rho(w)$, whereas the measures of interest in (5.1.5) are concentrated on an interval, and may have no regularity. Thus, some care is needed in formulating and proving a theorem on positivity analogous to (5.2.11).

Recall that a signed Borel measure μ on the line may be uniquely decomposed into two positive measures $\mu_\pm$ respectively such that $\mu = \mu_+ - \mu_-$. The Fourier transform of a measure is defined by

$$\hat{\mu}(u) = \int_{\mathbb{R}} e^{-ius} \mu(ds), \quad u \in \mathbb{R}. \tag{5.2.12}$$

The Fourier transform is a well-defined distribution. If $\mu_\pm$ are finite measures on $[-R, R]$, the Fourier transform is a continuous function of u that decays to zero as $|u| \to \infty$ by the Riemann-Lebesgue lemma.

Lemma 17. *Assume $\mu = \mu_+ - \mu_-$ is a signed measure on $[-R, R]$ such that*

$$\int_{-R}^R \mu_+(dr) = \int_{-R}^R \mu_-(dr) < \infty. \tag{5.2.13}$$

Then we have the identity

$$\int_{-R}^R \int_{-R}^R \log \frac{1}{|r - s|} \left(\mu_+(dr)\mu_+(ds) + \mu_-(dr)\mu_-(ds) \right) \tag{5.2.14}$$

$$= \int_{-R}^R \int_{-R}^R \log \frac{1}{|r - s|} \left(\mu_+(dr)\mu_-(ds) + \mu_-(dr)\mu_+(ds) \right) + \int_0^\infty \frac{|\hat{\mu}(u)|^2}{u} \, du.$$

In particular, $I[\mu] > 0$ if μ is non-zero and satisfies (5.2.13).

Remark 39. Equation (5.2.14) simply says that

$$\int_{-R}^{R} \int_{-R}^{R} \log \frac{1}{|r-s|} \, \mu(dr) \, \mu(ds) = \int_{0}^{\infty} \frac{|\hat{\mu}(u)|^2}{u} \, du. \tag{5.2.15}$$

for a signed measure μ with $\int_{-R}^{R} \mu(ds) = 0$. This identity has been written in the form (5.2.14) in order to ensure that there are no ill-defined terms of the form $\infty - \infty$. It is now clear from (5.1.4) and (5.1.5) that $I[\mu] > 0$ for such measures.

Proof. This proof is from [3, p.142]. We first regularize the logarithm at 0 and use the following integral representation. For any real s and $\varepsilon > 0$

$$\log(s^2 + \varepsilon^2) = \log \varepsilon^2 + 2 \, \mathsf{Im} \int_{0}^{\infty} e^{-\varepsilon u} \frac{e^{isu} - 1}{iu} \, du. \tag{5.2.16}$$

We apply this integral representation to the following regularization of $I[\mu]$, and use the fact that $\int_{-R}^{R} \mu(dr) = 0$, to obtain

$$\int_{-R}^{R} \int_{-R}^{R} \log \left((r-s)^2 + \varepsilon^2 \right) \mu(dr) \mu(ds)$$

$$= 2 \, \mathsf{Im} \int_{0}^{\infty} e^{-\varepsilon u} \int_{-R}^{R} \int_{-R}^{R} \frac{e^{i(r-s)u} - 1}{iu} \, \mu(dr) \mu(ds) \, du$$

$$= 2 \, \mathsf{Im} \int_{0}^{\infty} e^{-\varepsilon u} \frac{|\hat{\mu}(u)|^2}{iu} \, du = -2 \int_{0}^{\infty} e^{-\varepsilon u} \frac{|\hat{\mu}(u)|^2}{u} \, du.$$

We may rewrite this identity in terms of $\mu_{\pm}$ as follows:

$$\int_{-R}^{R} \int_{-R}^{R} \log \frac{1}{\sqrt{(r-s)^2 + \varepsilon^2}} \left(\mu_+(dr)\mu_+(ds) + \mu_-(dr)\mu_-(ds) \right) \tag{5.2.17}$$

$$= \int_{-R}^{R} \int_{-R}^{R} \log \frac{1}{\sqrt{(r-s)^2 + \varepsilon^2}} \left(\mu_+(dr)\mu_-(ds) + \mu_-(dr)\mu_+(ds) \right) + \int_{0}^{\infty} e^{-\varepsilon u} \frac{|\hat{\mu}(u)|^2}{u} \, du.$$

We now let $\varepsilon \downarrow 0$ and use the monotone convergence theorem to obtain (5.2.14) $\qquad \square$

Finally, let us prove Lemma 16. Suppose μ_0 and μ_1 be two measures in $\mathcal{P}_R$ as in (5.2.4). The difference

$$(1-\theta)I[\mu_0]+\theta I[\mu_1]-I[\mu_\theta] = \theta(1-\theta) \int \int \log \frac{1}{|r-s|} \, (\mu_0 - \mu_1) \, (dx) \, (\mu_0 - \mu_1) \, (dx)$$

in the sense of signed measures. Thus, it is strictly positive when $\mu_0 \neq \mu_1$ by Lemma 17.

5.2.4 Case 2: Measures on the line

Having explained the main ideas behind Theorem 38 for finite measures, let us turn to the measures on the line. The proof of uniqueness requires no change, since it is easily verified that Lemma 17 holds for measures in $\mathcal{P}_{\mathbb{R}}$. However, it is necessary to modify the proof of existence to account for a possible loss of compactness: a sequence of measures in $\mathcal{P}_{\mathbb{R}}$ may drift off to infinity (e.g. $\mu_k = \delta_k$, $k \in \mathbb{Z}$). The appropriate condition required for compactness here is the following.

Definition 40. A sequence of measures $\{\mu_k\}_{k=1}^{\infty} \in \mathcal{P}_{\mathbb{R}}$ is *tight* if for every $\varepsilon > 0$ there exists $M_{\varepsilon} > 0$ such that

$$\sup_{k \geq 1} \mu_k \left(\mathbb{R} \backslash [-M_{\varepsilon}, M_{\varepsilon}] \right) < \varepsilon. \tag{5.2.18}$$

Compactness of measures in $\mathcal{P}_{\mathbb{R}}$ is provided by the Prokhorov-Varadarajan criterion: the sequence $\{\mu_k\}_{k=1}^{\infty} \in \mathcal{P}_{\mathbb{R}}$ has a subsequence that converges to a measure $\mu \in \mathcal{P}_{\mathbb{R}}$ if and only if the sequence $\{\mu_k\}_{k=1}^{\infty}$ is tight [22]. In practice, application of this criterion requires a uniform estimate on the tails of the measures $\{\mu_k\}_{k=1}^{\infty}$. Such a bound is possible only if the growth of the confining potential $V(x)$ as $|x| \to \infty$ is faster than the divergence of $\log |x|$ as $|x| \to \infty$. We formalize this requirement as follows. For any $\varepsilon > 0$, observe that

$$|r - s| = |r - 1 - (s - 1)| \leq \sqrt{r^2 + 1} \sqrt{s^2 + 1}. \tag{5.2.19}$$

Therefore, we have the lower bound

$$\log \frac{1}{r - s} \geq \frac{1}{2} \left(\log \frac{1}{r^2 + 1} + \log \frac{1}{s^2 + 1} \right). \tag{5.2.20}$$

Let us define the function

$$l(s) = \frac{1}{2} \log \frac{1}{s^2 + 1} + \frac{1}{2} V(s). \tag{5.2.21}$$

If $l(s)$ is bounded below, then by adding a constant to V if necessary, we can ensure that $l(s) \geq 0$ for all s. Clearly, this does not change the nature of the minimization problem.

Theorem 41. *Assume $V(s)$ is a continuous function such that $l(s)$ is bounded below and $l(s) \to \infty$ as $|s| \to \infty$.*

(a) *There exists a unique probability measure $\mu_* \in \mathcal{P}_{\mathbb{R}}$ such that*

$$I[\mu_*] \leq \min_{\mu \in \mathcal{P}_{\mathbb{R}}} I[\mu]. \tag{5.2.22}$$

(b) *The support of the measure μ_* is contained within a finite interval.*

Proof. (a) Since V is bounded below and the addition of a constant to V does not change the minimization problem, we may assume that $l(s) \geq 0$. Then

$$e(r,s) = \log \frac{1}{|r-s|} + \frac{1}{2}V(r) + \frac{1}{2}V(s) \geq l(r) + l(s) \geq 0, \qquad (5.2.23)$$

and $c := \inf_{\mu \in \mathcal{P}_\mathbb{R}} I[\mu] \geq 0$. Suppose $\mu_k{}_{k=1}^\infty$ is an infimizing sequence: i.e. $\lim_{k \to \infty} I[\mu_k] = c$. Without loss of generality, we may assume that $I[\mu_k] \leq c+1$ for all k. Tightness of the sequence $\{\mu_k\}_{k=1}^\infty$ follows from the following (Chebyshev) inequality. For any $M > 0$,

$$c + 1 \geq I[\mu_k] = \int_\mathbb{R} \int_\mathbb{R} e(r,s) \mu_k(dr) \mu_k(ds) \qquad (5.2.24)$$

$$\geq 2 \int_\mathbb{R} l(s) \mu_k(ds) \geq 2l_M \int_{|s|>M} \mu_k(ds) = 2l_M \mu_k(\mathbb{R} \setminus [-M, M]),$$

where $l_M = \inf_{|s| \geq M} l(s)$. Since $\lim_{|s| \to \infty} l(s) = \infty$, $l_M \to \infty$ as $M \to \infty$. Thus, for any $\varepsilon > 0$, we may choose $M = M_\varepsilon$ large enough so that (5.2.18) holds. The rest of the proof of part (a) follows that of Theorem 38.

(b) For any $M > 0$, let S_M denote the set $(-\infty, M) \cup (M, \infty)$. We will show that $\mu_*(S_M) = 0$ if M is large enough. The proof relies on varying the measure μ_* by adding more mass proportional to μ_* in the set S_M. More precisely, let ν denote the restriction of μ_* to the set S_M, and for any $t \in (-1, 1)$, define the measures

$$\mu_t = \frac{\mu_* + t\nu}{1 + t\nu(S_M)}. \qquad (5.2.25)$$

We then find that $I[\mu_t]$ is a differentiable function of t, with

$$0 = \left. \frac{dI[\mu_t]}{dt} \right|_{t=0} = 2 \int_{S_M} \nu(ds) \int_\mathbb{R} \mu_*(dr) e(r,s) - 2\nu(S_M) I[\mu_*]. \qquad (5.2.26)$$

The estimate (5.2.23) and positivity of l yields the lower bound

$$2 \int_{S_M} \nu(ds) \int_\mathbb{R} \mu_*(dr) e(r,s) \qquad (5.2.27)$$

$$\geq \int_{S_M} l(s) \nu(ds) + \int_\mathbb{R} l(r) \mu_*(dr) \geq \int_{S_M} l(s) \nu(ds) \geq l_M \nu(S_M).$$

As in part (a), $l_M \to \infty$ as $M \to \infty$. Thus, for M sufficiently large, we have $l_M - I[\mu_*]) > 0$ and since ν is a positive measure, we have the (trivial) estimate

$$2(l_M - I[\mu_*]) \nu(S_M) \geq 0. \qquad (5.2.28)$$

On the other hand, the inequalities (5.2.26) and (5.2.27) yield the opposite inequality

$$2(l_M - I[\mu_*]) \nu(S_M) \leq 0. \qquad (5.2.29)$$

Thus, $\nu(S_M) = 0$ for all M such that $l_M > I[\mu_*]$. $\square$

5.3   Fekete points

A second approach to the energy minimization problem relies on a study of the minimizers of the function $E(x)$ defined in (5.1.1) for $x \in \mathbb{R}^n$, and a potential V that satisfies the assumptions of Theorem 41. For any such potential, $0 \le E(x) < \infty$ for any $x \in \mathbb{R}^n$ such that $x_j \ne x_k$, $j \ne k$. Thus, for each n, there exists a set of points $F_n \subset \mathbb{R}^n$, such that

$$E(x_*) = \min_{x \in \mathbb{R}^n} E(x), \quad x_* \in F_n. \tag{5.3.1}$$

The set F_n is called the set of n-Fekete points. The Fekete points are naturally connected to the minimization problem for the functional $I[\mu]$ through the modified functional $H[L_n]$, where $L_n(x)$ is the empirical measure associated to a point $x \in \mathbb{R}^n$. Let δ_n denote the rescaled energy of Fekete points

$$\delta_n = \frac{1}{n(n-1)} E(x^{(n)}). \tag{5.3.2}$$

The main result is then the following

Theorem 42. *Assume V satisfies the assumptions of Theorem 41. Let $\{x^{(n)}\}_{n=1}^\infty$ be a sequence of points $x^{(n)} \in F_n$ and Then*

  (a) *The rescaled energy of Fekete points increases monotonically to $I[\mu_*]$.*

$$0 \le \delta_n \le \delta_{n+1} \le I[\mu_*]. \tag{5.3.3}$$

  (b) *The empirical measures $L(x^{(n)})$ converge weakly to μ_*.*

Proof of (a). We first prove the estimates (5.3.3). The uniform upper bound on $E(x^{(n)})$ is obtained as follows. Fix a positive integer n and a point $x^{(n)} \in F_n$. By definition, for any $s = (s_1, \ldots, s_n) \in \mathbb{R}^n$,

$$E(x^{(n)}) \le E(s) = \frac{1}{2} \sum_{j,k=1}^n \left(V(s_j) + V(s_k) \right) + \sum_{j \ne k=1}^n \log \frac{1}{|s_j - s_k|}. \tag{5.3.4}$$

Let $\mu(ds)$ be any probability measure on the line. We integrate (5.3.4) with respect to the n-fold tensorized probability measure $\mu \otimes \mu \cdots \otimes \mu$ on $\mathbb{R}^n$ to obtain

$$E(x^{(n)}) \tag{5.3.5}$$

$$\le \int_{\mathbb{R}^n} \left[\frac{1}{2} \sum_{j,k=1}^n \left(V(s_j) + V(s_k) \right) + \sum_{j \ne k=1}^n \log \frac{1}{|s_j - s_k|} \right] \mu(ds_1)\mu(ds_2) \cdots \mu(ds_n)$$

$$= n(n-1) \int_{\mathbb{R}} \int_R e(r,s)\mu(ds)\mu(dr) = I[\mu],$$

since for each value of the indices j and k only the integrals over $\mu(ds_j)$ and $\mu(ds_k)$ give contributions that are not unity and there are $n(n-1)$ possible unordered pairings of j and k. In particular, $E(x^{(n)}) \le n(n-1)I[\mu_*]$.

The monotonicity of δ_n follows from the following argument. Suppose $x^{(n+1)} = (x_1, \ldots, x_{n+1})$ is point in the Fekete set F_{n+1}. We fix an index m, $1 \le m \le n+1$ and use the definition of E in (5.1.1) to obtain

$$
e^{-\frac{1}{n(n+1)}E(x^{(n+1)})} = \left(\prod_{1 \le j \ne k \le n+1} |x_j - x_k| e^{-\frac{V(x_j)}{2}} e^{-\frac{V(x_k)}{2}} \right)^{\frac{1}{n(n+1)}} \tag{5.3.6}
$$

$$
= \left(\prod_{j \ne m} |x_j - x_m| e^{-\frac{V(x_j)}{2} - \frac{V(x_m)}{2}} \right)^{\frac{2}{n(n+1)}} \left(\prod_{j,k \ne m} |x_j - x_k| e^{-\frac{V(x_j)}{2}} e^{-\frac{V(x_k)}{2}} \right)^{\frac{1}{n(n+1)}}
$$

$$
\le \left(\prod_{j \ne m} |x_j - x_m| e^{-\frac{V(x_j)}{2}} e^{-\frac{V(x_m)}{2}} \right)^{\frac{2}{n(n+1)}} e^{-\delta_n \frac{n-1}{n+1}}
$$

since the second term is the energy $E(\hat{x})$ of the point $\hat{x} \in \mathbb{R}^n$ obtained from $x^{(n)}$ by projecting out the coordinate x_m.

Since m is arbitrary, we take the product over $1 \le m \le n+1$ to obtain

$$
e^{-\frac{1}{n}E(x^{(n+1)})} \le e^{-(n-1)\delta_n} \left(\prod_{1 \le m \le n+1} \prod_{1 \le j \le n+1, j \ne m} |x_j - x_m| e^{-\frac{V(x_j)}{2} - \frac{V(x_m)}{2}} \right)^{\frac{2}{n(n+1)}}
$$

$$
= e^{-(n-1)\delta_n} e^{-\frac{2}{n(n+1)}E(x^{(n+1)})}. \tag{5.3.7}
$$

This inequality simplifies to $\delta_n \le \delta_{n+1}$. $\square$

Proof of (b). While the self-energy of all the Fekete points is infinite, inequality (5.3.3) shows that a suitably renormalized energy is finite, and bounded above by $I[\mu_*]$. This inequality, in combination with an easy modification of the Chebyshev inequality (5.2.24) also shows that the empirical measures $L(x^{(n)})$ are tight. Thus, there exists a convergent subsequence and a limiting probability measure $\nu \in \mathcal{P}_{\mathbb{R}}$ such that the empirical measures $L^{(n)}$ defined by the Fekete points $x^{(n)}$ converge weakly to ν as $n \to \infty$.

For any $M > 0$, we introduce the cut-off energy $e_M(r,s) = \min(M, e(r,s))$ and observe that

$$
\delta_n = \frac{1}{n(n-1)} E(x^{(n)}) = \frac{n^2}{n(n-1)} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{r \ne s} e(r,s) L^{(n)}(dr) L^{(n)}(ds)
$$

$$
\ge \frac{n^2}{n(n-1)} \int_{\mathbb{R}} \int_{\mathbb{R}} e_M(r,s) L^{(n)}(dr) L^{(n)}(ds) - \frac{M}{n-1}.
$$

Since the function $e_M(r,s)$ is continuous and $0 \le e_M(r,s) \le M$, we may interchange limits as $n \to \infty$, and use Theorem 42(a) to obtain

$$
I[\mu_*] \ge \liminf_{n \to \infty} \delta_n \ge \int_{\mathbb{R}} \int_{\mathbb{R}} e_M(r,s) \nu(dr) \nu(ds). \tag{5.3.8}
$$

We now let $M \to \infty$ and use the monotone convergence theorem and the fact that μ_* is a minimizer to obtain

$$I[\mu_*] \geq I[\mu] \geq I[\mu_*]. \tag{5.3.9}$$

Since μ_* is unique, it follows that $\mu_* = \nu$.

This argument proves that every subsequential limit of $L^{(n)}$ is μ_*. Thus, the entire sequence converges to μ_*. $\qquad\square$

5.4 Exercises

The first three questions are related. The goal is to formulate and analyze the equation for the equilibrium measure μ_* associated to the potential $V(x)$. In order to simplify your calculations, assume that μ_* has a continuous density ψ, in all the problems below. The last two questions discuss enumeration problems related to the Catalan numbers.

1. Basics of the Hilbert transform. Let $G(z)$ denote the Stieltjes transform

$$G(z) = \int_{-\infty}^{\infty} \frac{1}{s-z} \mu_*(ds) = \int_{-\infty}^{\infty} \frac{1}{s-z} \psi(s)(ds), \quad z \in \mathbb{C}\backslash\mathrm{supp}(\mu_*). \tag{5.4.1}$$

The Hilbert transform of ψ is the limit of the Stieltjes transform as $z \to x \in \mathbb{R}$. The Hilbert transform also differs from the Stieltjes transform by the inclusion of a factor of π (since this makes the Fourier transform of the operator $\mathcal{H}$ particularly simple). That is, given μ_* as above, we set

$$\mathcal{H}\psi(x) = \frac{1}{\pi}\mathrm{p.v.} \int_{-\infty}^{\infty} \frac{\psi(s)}{x-s}\, ds := \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \frac{x-s}{(x-s)^2 + \varepsilon^2}\, \psi(s)\, ds. \tag{5.4.2}$$

(a) Show that $\mathcal{H}\psi$ is a bounded function when $\psi(x)$ is continuous.

(b) Show that μ_* may be recovered from G by evaluating the jump in the imaginary part of G across the support of μ_*:

$$\lim_{\varepsilon \to 0} \frac{1}{2\pi i} \left(G(x + i\varepsilon) - G(x - i\varepsilon) \right) = \psi(x). \tag{5.4.3}$$

(c) Compute the Hilbert transform of the following functions to obtain a feel for it (answers are on wikipedia):

$$e^{ix}, \quad \delta_0(x), \quad \mathbf{1}_{[a,b]}(x).$$

2. Integral equation for ψ. Assume V is differentiable and satisfies the assumptions of Theorem 41 so that μ_* has compact support. Show that if μ_* has a density ψ as above, then it satisfies the integral equation

$$\mathcal{H}\psi(x) = \frac{1}{2\pi}V'(x) \quad \text{on} \quad \mathrm{supp}(\mu_*). \tag{5.4.4}$$

3. Fixed point equation for the resolvent. One solution to (5.4.4) uses the Stieltjes transform $G(z)$. Assume that $V(x)$ is a polynomial of degree $d \geq 2$.

(a) Show that G satisfies the quadratic equation

$$G^2(z) + V'(z)G(z) + P(z) = 0, \tag{5.4.5}$$

where $P(z)$ is a polynomial of degree $d - 2$ whose coefficients are determined by the moments of μ_* of degree lower than d. The solution branch is determined by the requirement that $G(z) \sim -1/z$ as $z \to \infty$ which is immediate from (5.4.1).

(b) Equation (5.4.5) may be solved by making further assumptions on the form of μ_*. In particular, assume that $V(z)$ is even, that the support of μ_* is a single interval $[-2a, 2a]$, and show that (5.4.5) simplifies to

$$G(z) = Q(z)\sqrt{z^2 - 4a^2} - \frac{1}{2}V'(z) \tag{5.4.6}$$

where $Q(z)$ is a polynomial of degree $d - 2$ whose coefficients are determined by the condition that $G(z) \sim -1/z$ as $z \to \infty$.

(c) Apply these ideas to compute the equilibrium measure for the quartic potential

$$V(x) = \frac{1}{2}x^2 + \frac{g}{4}x^4. \tag{5.4.7}$$

Show that

$$G(z) = \left(\frac{1}{2} + \frac{g}{2}x^2 + ga^2\right)\sqrt{x^2 - 4a^2} - \frac{1}{2}\left(x + gx^3\right), \tag{5.4.8}$$

where a^2 solves the quadratic equation

$$3ga^4 + a^2 - 1 = 0. \tag{5.4.9}$$

(d) Compute the associated density $\psi(x)$ and plot it as g varies.

4. Establish the identity (1.3.11).

5. Show that the Catalan numbers enumerate the number of Dyck paths as discussed below equation (1.3.12).

5.5 Notes

To include in improved version.

1. Fixed point equation for equilibrium measure.

2. Properties of Hilbert transform.

3. Convergence of k-point distribution to tensor product of equilibrium measure.

Appendix A

The Airy function

A.1 Integral representation

There are several different conventions for the definition of the Airy function. The standardization adopted here follows [1]. The Airy function, $\mathrm{Ai}(x)$ is defined as the oscillatory integral

$$\mathrm{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{t^3}{3} + xt\right) dt = \frac{1}{\pi} \lim_{b \to \infty} \int_0^b \cos\left(\frac{t^3}{3} + xt\right) dt. \quad \text{(A.1.1)}$$

This is an improper integral, that is, the integral converges conditionally, not absolutely. In order to obtain an absolutely convergent integral, it is necessary to work in the complex plane. Let C denote a contour in the complex plane that starts and ends at the point at infinity, and is asymptotically tangent to the rays $e^{-i\pi/3}$ and $e^{+i\pi/3}$ respectively. Then first setting $t = -iz$ and then deforming the contour, we have

$$\mathrm{Ai}(x) = \frac{1}{2\pi i} \int_{-\infty}^\infty e^{\left(\frac{z^3}{3} - xz\right)} dz = \frac{1}{2\pi i} \int_C e^{\left(\frac{z^3}{3} - xz\right)} dz. \quad \text{(A.1.2)}$$

The integral is absolutely convergent for every $x \in \mathbb{C}$ on the contour C. Indeed, with $z = re^{i\theta}$,

$$\left| e^{\left(\frac{z^3}{3} - xz\right)} \right| \le e^{|x|r} e^{-r^3 \cos(3\theta)/3} \sim e^{-r^3/3} e^{r|x|} \quad \text{(A.1.3)}$$

as $z \to \infty$ along the rays $\theta = \pm\pi/3$. Thus, $\mathrm{Ai}(x)$ is an entire function.

A.2   Differential equation

We differentiate under the integral sign (justified by (A.1.3)) and integrate by parts to obtain

$$\text{Ai}''(x) = \frac{1}{2\pi i} \int_C z^2 e^{\left(\frac{z^3}{3} - xz\right)} dz \tag{A.2.1}$$

$$= \frac{1}{2\pi i} \int_C \frac{d}{dz} e^{\frac{z^3}{3}} e^{-xz} \, dz = -\frac{1}{2\pi i} \int_C e^{\frac{z^3}{3}} \frac{d}{dz} e^{-xz} \, dz = x\text{Ai}(x).$$

Thus, $\text{Ai}(x)$ satisfies the *Airy differential equation*

$$y'' = xy, \quad x \in \mathbb{C}. \tag{A.2.2}$$

This differential equation has a scaling invariance: if $y(x)$ is a solution, so are $y(\omega x)$ and $y(\omega^2 x)$ where $\omega = e^{2\pi i/3}$ is a cube root of unity. Thus, both $\text{Ai}(\omega x)$ and $\text{Ai}(\omega^2 x)$ solve (A.2.2). Each of these solutions is linearly independent of $\text{Ai}(x)$. A solution to (A.2.2) that is real when x is real, and is linearly independent from $\text{Ai}(x)$, is obtained from the linear combination

$$\text{Bi}(x) = e^{\pi i/6}\text{Ai}(\omega x) + e^{\pi - i/6}\text{Ai}(\omega^2 x). \tag{A.2.3}$$

A.3   Asymptotics

The functions $\text{Ai}(x)$ and $\text{Bi}(x)$ have the following asymptotic properties.

Asymptotics as $x \to \infty$.

$$\zeta = \frac{2}{3}x^{\frac{3}{2}}, \quad \text{Ai}(x) \sim \frac{e^{-\zeta}}{2x^{\frac{1}{4}}\sqrt{\pi}}, \quad \text{Bi}(x) \sim \frac{x^{\frac{1}{4}}}{\sqrt{\pi}}e^{\zeta}. \tag{A.3.1}$$

Asymptotics as $x \to -\infty$.

$$\zeta = \frac{2}{3}(-x)^{\frac{3}{2}}, \quad \text{Ai}(x) \sim \frac{1}{x^{\frac{1}{4}}\sqrt{\pi}} \sin\left(\zeta + \frac{\pi}{4}\right), \quad \text{Bi}(x) \sim \frac{1}{x^{\frac{1}{4}}\sqrt{\pi}} \cos\left(\zeta + \frac{\pi}{4}\right).$$
$$\tag{A.3.2}$$

Appendix B

Hermite polynomials

In this chapter, μ denotes the weight function

$$\mu(dx) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx. \tag{B.0.1}$$

The Hermite polynomials $\{\mathfrak{h}_k\}_{k=0}^{\infty}$ are the monic family of polynomials of degree k orthogonal with respect to the weight μ.

B.1   Basic formulas

$$\mathfrak{h}_k(x) = e^{\frac{x^2}{2}}\left(-\frac{d}{dx}\right)^k e^{\frac{-x^2}{2}}. \tag{B.1.1}$$

$$\mathfrak{h}_k(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}(-i\xi)^k e^{-\frac{1}{2}(\xi-ix)^2}\,d\xi. \tag{B.1.2}$$

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathfrak{h}_k(x)\mathfrak{h}_l(x)e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}k!\delta_{kl}. \tag{B.1.3}$$

$$x\mathfrak{h}_k(x) = \mathfrak{h}_{k+1}(x) + k\mathfrak{h}_{k-1}(x), \quad k \geq 1. \tag{B.1.4}$$

$$\mathfrak{h}_k'(x) = k\mathfrak{h}_{k-1}(x). \tag{B.1.5}$$

$$\mathfrak{h}_k''(x) - x\mathfrak{h}_k'(x) + k\mathfrak{h}_k(x) = 0. \tag{B.1.6}$$

$$\sum_{j=0}^{k-1}\frac{1}{j!}\mathfrak{h}_j(x)\mathfrak{h}_j(y) = \frac{(\mathfrak{h}_k(x)\mathfrak{h}_{k-1}(y) - \mathfrak{h}_{k-1}(x)\mathfrak{h}_k(y))}{(k-1)!(x-y)}, \quad x \neq y. \tag{B.1.7}$$

Relation (B.1.1) may be treated as an alternate definition of the Hermite polynomials. On the other hand, since we have defined the Hermite polynomials as the monic orthogonal polynomials obtained by applying the Gram-Schmidt procedure to the set $\{1, x, x^2, \ldots\}$ in $L^2(\mathbb{R}, \mu)$, equation (B.1.1) may be verified as follows. First, it is clear from (B.1.1) that $\mathfrak{h}_k(x)$ is a monic polynomial of degree k and that $\mathfrak{h}_0(x) = 1$, $\mathfrak{h}_1(x) = x$. By induction, if it has been established that property (B.1.1) defines the Hermite polynomials for $j \leq k - 1$, then it is only necessary to show that the monic polynomial

$$P_k(x) = e^{\frac{x^2}{2}} \left(-\frac{d}{dx} \right)^k e^{\frac{-x^2}{2}},$$

is the same as $\mathfrak{h}_k$. The polynomial P_k is orthogonal to $\mathfrak{h}_j$, $0 \leq j \leq k-1$ because

$$\int_{\mathbb{R}} P_k(x) \mathfrak{h}_j(x) \mu(dx) = \int_{\mathbb{R}} \left(\frac{d}{dx} \right)^k \mathfrak{h}_j(x) \mu(dx) = 0,$$

since H_j has degree less than k. Since P_k is monic, it must be $\mathfrak{h}_k$. The same calculation serves to establish (B.1.3).

The integral representation (B.1.2) follows from the formula for the Fourier transform of a Gaussian

$$e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x} e^{-\frac{\xi^2}{2}} \, d\xi, \tag{B.1.8}$$

and the identity (B.1.1).

The two-term recurrence relation follows from (3.4.18) and (B.1.3) (see also Remark 27). The coefficient a_k vanishes because equation (B.1.1) shows that $\mathfrak{h}_k^2$ is an even polynomial for all k. The coefficient b_k^2 may be rewritten

$$b_k^2 = \frac{\int x \mathfrak{h}_{k-1}(x) \mathfrak{h}_k(x) \mu(dx)}{\int \mathfrak{h}_{k-1}^2 \mu(dx)} = \frac{\int x \mathfrak{h}_{k-1}(x) \mathfrak{h}_k(x) \mu(dx)}{\int \mathfrak{h}_k^2 \mu(dx)} \frac{\int \mathfrak{h}_k^2(x) \mu(dx)}{\int \mathfrak{h}_{k-1}^2 \mu(dx)} = 1 \cdot k^2,$$
$$\tag{B.1.9}$$

by (B.1.3).

The differential equation (B.1.5) is obtained by rewriting (B.1.1) in the form

$$e^{-\frac{x^2}{2}} \mathfrak{h}_k(x) = (-1)^k \left(\frac{d}{dx} \right)^k e^{-\frac{x^2}{2}},$$

differentiating both sides, and then multiplying by $e^{\frac{x^2}{2}}$. Equation (B.1.6) is obtained by differentiating (B.1.5) and using (B.1.4). The proof of the Christoffel-Darboux identity is left as an exercise to the reader.

B.2   Hermite wave functions

The Hermite wave functions $\{\psi\}_{k=0}^{\infty}$ are defined by

$$\psi_k(x) = \frac{1}{\sqrt{k!}} \frac{e^{-x^2/4}}{(2\pi)^{1/4}} \mathfrak{h}_k(x), \quad k = 0, 1, 2, \ldots \tag{B.2.1}$$

The following properties of the Hermite wave-functions follow immediately from the corresponding properties of the Hermite polynomials.

$$\int_{\mathbb{R}} \psi_k(x)\psi_l(x)\, dx = \delta_{kl}. \tag{B.2.2}$$

$$x\psi_k(x) = \sqrt{k+1}\,\psi_{k+1}(x) + \sqrt{k}\,\psi_{k-1}(x). \tag{B.2.3}$$

$$\psi'_k(x) = -\frac{x}{2}\psi_k(x) + \sqrt{k}\,\psi_{k-1}(x). \tag{B.2.4}$$

$$\psi''_k(x) + \left(k + \frac{1}{2} - \frac{x^2}{4}\right)\psi_k(x) = 0. \tag{B.2.5}$$

$$\sum_{k=0}^{n-1} \psi_k(x)\psi_k(y) = \sqrt{n}\,\frac{(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(x))}{x - y}. \tag{B.2.6}$$

B.3 Small x asymptotics

The following classical formulas capture the asymptotics of the Hermite polynomials near the origin [1, §22.15].

$$\lim_{n\to\infty} \frac{(-1)^n}{2^n}\frac{\sqrt{n}}{n!}\, \mathfrak{h}_{2n}\left(\frac{x}{\sqrt{2n}}\right) = \frac{1}{\sqrt{\pi}}\cos x. \tag{B.3.1}$$

$$\lim_{n\to\infty} \frac{(-1)^n}{2^n\, n!}\, \mathfrak{h}_{2n+1}\left(\frac{x}{\sqrt{2n}}\right) = \sqrt{\frac{2}{\pi}}\,\sin x. \tag{B.3.2}$$

Further, the convergence to the limit is uniform over x in a bounded interval.

In comparing equations (B.3.1) and (B.3.2) with a standard reference such as [1], the reader should note that there are two conventions in the definition of Hermite polynomials. The exponential weight in earlier sources was chosen to be e^{x^2}, which differs from our choice (B.0.1). The relation between the Hermite polynomials, $\{H_n(x)\}$ in [1], and those used here are:

$$H_n(x) = 2^{\frac{n}{2}}\,\mathfrak{h}_n(x\sqrt{2}), \quad \mathfrak{h}_n(x) = 2^{-\frac{n}{2}}\,H_n\left(\frac{x}{\sqrt{2}}\right). \tag{B.3.3}$$

These formulas may be immediately translated into asymptotic formulas for the Hermite wave functions, using Stirling's approximation for the factorial.

$$\lim_{n\to\infty} (2n)^{1/4}(-1)^n\,\psi_{2n}\left(\frac{x}{\sqrt{2n}}\right) = \frac{1}{\sqrt{\pi}}\cos x. \tag{B.3.4}$$

$$\lim_{n\to\infty} (2n)^{1/4}(-1)^n \psi_{2n+1}\left(\frac{x}{\sqrt{2n}}\right) = \frac{1}{\sqrt{\pi}} \sin x. \tag{B.3.5}$$

The asymptotic formulas (B.3.1) and (B.3.2) are proved by applying Laplace's method to the integral formula (B.1.2). We only explain how to prove (B.3.1) since equation (B.3.2) is similar. Since $(i)^{2n} = (-1)^n$, we take the real part of (B.1.2) to find

$$(-1)^{2n}\mathfrak{h}_{2n}\left(\frac{x}{\sqrt{2n}}\right) = \sqrt{\frac{2}{\pi}}\, e^{\frac{x^2}{4n}} \int_0^\infty \xi^{2n} e^{-\frac{\xi^2}{2}} \cos\left(\frac{x\xi}{\sqrt{2n}}\right) d\xi$$

$$= \frac{2^{n+1} n^{n+\frac{1}{2}}}{\sqrt{\pi}} \int_0^\infty e^{-n\left(t^2 - 2\log t\right)} \cos xt\, dt, \tag{B.3.6}$$

by rescaling $\xi = \sqrt{n}\, t$. We now apply Laplace's method to the integral above. The function $g(t) = t^2 - 2\log t$ has a single minimum on the interval $(0, \infty)$ at $t = 1$. At this point

$$g(1) = 1, \quad g'(1) = 0, \quad g''(1) = 4. \tag{B.3.7}$$

Laplace's approximation now yields

$$\int_0^\infty e^{-ng(t)} \cos xt\, dx \sim e^{-n} \sqrt{\frac{\pi}{2n}} \cos x, \tag{B.3.8}$$

which when combined with (B.3.6) implies

$$(-1)^{2n}\mathfrak{h}_{2n}\left(\frac{x}{\sqrt{2n}}\right) \sim 2^{n+\frac{1}{2}} n^n e^{-n} \cos x. \tag{B.3.9}$$

Equation (B.3.9) is equivalent to (B.3.1) by Stirling's approximation. Further, it is easy to check that the error is uniformly small for x in a bounded set.

B.4 Plancherel-Rotach asymptotics

Another asymptotic regime is obtained when we consider $x = O(\sqrt{n})$ and let $n \to \infty$. The limit is oscillatory or exponential depending on the range of x. This is to be expected: for each n, the polynomial $\mathfrak{h}_n(x)$, and thus the wave function $\psi_n(x)$, has n zeros. The largest and smallest of the zeros are approximately $\pm\sqrt{(n+1/2)}$. The oscillatory regime is obtained when $x(n+1/2)^{-1/2}$ lies well within the interval $(-1, 1)$. Outside this interval, the Hermite wave function decays exponentially fast. A more delicate calculation, using the Airy function, is required to understand the transition from oscillatory to exponential behavior.

We will prove a weaker version of the Plancherel-Rotach formulas, that suffices for our needs. These formula are as follows.

Case 1. Oscillatory behavior.

$$x = 2\cos\varphi, \qquad 0 < \varphi < \pi. \tag{B.4.1}$$

$$n^{\frac{1}{4}}\psi_{n+p}\left(x\sqrt{n}\right) \sim \frac{1}{\sqrt{\pi\sin\varphi}}\cos\left[n\left(\varphi - \frac{1}{2}\sin 2\varphi\right) + \left(p + \frac{1}{2}\right)\varphi - \frac{\pi}{4}\right]. \tag{B.4.2}$$

The convergence is uniform for φ in a compact subset of $(0,\pi)$.

Case 2. Exponential decay.

$$|x| = 2\cosh\varphi, \qquad 0 < \varphi. \tag{B.4.3}$$

$$n^{\frac{1}{4}}\psi_{n+p}\left(x\sqrt{n}\right) \sim \frac{e^{(p+\frac{1}{2})\varphi}}{\sqrt{2\pi\sinh\varphi}}e^{-\frac{n}{2}\left(e^{2\varphi}+1-2\varphi\right)}. \tag{B.4.4}$$

The convergence is uniform for φ in a compact subset of $(0,\infty)$. Observe that $e^{2\varphi} + 1 - 2\varphi > 0$ when $\varphi > 0$, ensuring exponential decay.

Case 3. The transition region.

$$x = 2\sqrt{n} + \frac{s}{n^{\frac{1}{6}}} \quad s \in \mathbb{C}, \tag{B.4.5}$$

$$n^{\frac{1}{12}}\,\psi_n(x) \sim \mathrm{Ai}(s) + O\left(n^{-\frac{2}{3}}\right). \tag{B.4.6}$$

The convergence is uniform for s in a compact subset of $\mathbb{C}$.

All three asymptotic relations are obtained by the method of steepest descent. Assume $x \in \mathbb{R}$. We fix an integer p, use the integral identity (B.1.2) with $k = n + p$, and rescale $\xi = nt$ to obtain

$$\mathfrak{h}_{n+p}\left(x\sqrt{n}\right) = \left(-i\sqrt{n}\right)^{n+p}\sqrt{\frac{n}{2\pi}}\int_{-\infty}^{\infty} t^{n+p}e^{-\frac{n}{2}(t-ix)^2}\,dt \tag{B.4.7}$$

$$= \left(-i\sqrt{n}\right)^{n+p}\sqrt{\frac{n}{2\pi}}\left(\int_0^{\infty} e^{-\frac{n}{2}(t-ix)^2}\,dt + (-1)^{n+p}\int_0^{\infty} e^{-\frac{n}{2}(t+ix)^2}\,dt\right)$$

$$:= \left(-i\sqrt{n}\right)^{n+p}\sqrt{\frac{n}{2\pi}}\left(I_{n+p}(x) + (-1)^{n+p}I_{n+p}(-x)\right). \tag{B.4.8}$$

The integral $I_{n+p}(x)$ may be rewritten in the form

$$I_{n+p}(x) = \int_0^{\infty} t^p e^{-ng(t)}\,dt, \quad g(t) = \frac{1}{2}(t - ix)^2 - \log t. \tag{B.4.9}$$

As is usual, the first step is to determine the critical points where $g'(t) = 0$. This reduces to the quadratic equation $t^2 - ixt - 1 = 0$. The three distinct asymptotic limits arise from the three distinct possibilities for the roots.

(a) $|x| < 2$. The function g has two critical points on the unit circle, given by

$$t_{\pm} = \frac{ix \pm \sqrt{4 - x^2}}{2} = ie^{\mp i\varphi}, \qquad\qquad \text{(B.4.10)}$$

where x and φ are related through (B.4.1).

(b) $|x| > 2$. The two critical points lie on the imaginary axis, and may be written in the form

$$t_{\pm} = i\left(\frac{x \pm \sqrt{x^2 - 4}}{2}\right) = i\,\mathrm{sgn}(x)e^{\pm\varphi}, \qquad\qquad \text{(B.4.11)}$$

where each branch of φ is defined through the relation (B.4.3).

(c) $|x| = 2$. The two critical points coalesce into a single value $t = 1$. A further blow-up is necessary to obtain the Airy asymptotics (B.4.6).

Let us first consider the integral $I_{n+p}(x)$ in case (a), and let us assume that $x > 0$ to be concrete. We deform the integral over $(0, \infty)$ into an integral through the steepest descent contour, Γ, through the critical point t_+ as shown in Figure ??. The existence of such a contour may be deduced by continuity, beginning with the observation that when $x = 0$, Γ is simply the segment $(0, \infty)$ along the real line. While in general, Γ is given by the equation $\mathsf{Im}(g(t)) = 0$, it is not important for us to solve for the contour explicitly: all that is required is to understand the phase of $g''(t_+)$ in order to check that the integral over $(0, \infty)$ can be deformed to an integral over Γ as x varies. It is easy to check that when $|x| < 2$

$$g''(t_+) = 1 + \frac{1}{t_+^2} = 1 - e^{2i\varphi} = \left(-ie^{i\varphi}\right)(2\sin\varphi). \qquad\qquad \text{(B.4.12)}$$

Thus, we have

$$I_{n+p}(x) = \int_0^{\infty} t^p e^{-ng(t)}\, dt = e^{-ng(t_+)} \int_{\Gamma} t^p e^{-n(g(t)-g(t_+))}\, dt$$

$$\approx e^{-ng(t_+)} t_+^p \left.\frac{dt}{ds}\right|_{t_+} \int_{-\infty}^{\infty} e^{-\frac{n}{2}|g''(t_+)|s^2}\, ds. \qquad\qquad \text{(B.4.13)}$$

In the second line, we have used the fact that $\mathsf{Im}(g(t)) = 0$ on Γ, and we have further approximated the integral over Γ by an integral over the tangent to Γ at t_+. More precisely, the approximation here is

$$g''(t_+)(t - t_+)^2 = |g''(t_+)|s^2,$$

which implies

$$\left.\frac{dt}{ds}\right|_{t_+} = e^{i\left(\frac{\pi}{4} - \frac{\varphi}{2}\right)}. \qquad\qquad \text{(B.4.14)}$$

We now combine the values

$$t_+ = ie^{-i\varphi}, \quad g(t_+) = -\frac{e^{2i\varphi}}{2} + i\left(\varphi - \frac{\pi}{2}\right),$$

with (B.4.13) and (B.4.14) to obtain

$$I_{n+p}(x) \sim e^{\frac{n}{2}\cos 2\varphi}\sqrt{\frac{\pi}{n\sin\varphi}}e^{i\left(\frac{n}{2}\sin 2\varphi + (n+p+\frac{1}{2})\left(\frac{\pi}{2}-\varphi\right)\right)}. \tag{B.4.15}$$

Finally, since x is real, we have $\overline{I_{n+p}(x)} = I_{n+p}(-x)$. We combine (B.4.8) with (B.4.15) to obtain

$$\mathfrak{h}_{n+p}(x\sqrt{n}) \sim n^{\frac{n+p}{2}}\sqrt{\frac{2}{\sin\varphi}}e^{\frac{n}{2}\cos 2\varphi}\cos\left[n\left(\varphi - \frac{1}{2}\sin 2\varphi\right) + \left(p + \frac{1}{2}\right)\varphi - \frac{\pi}{4}\right], \tag{B.4.16}$$

where x and φ are related via (B.4.1). We now use (B.2.1) and Stirling's approximation to obtain (B.4.2).

The asymptotics in case (b) are obtained as follows. We make the change of variables (B.4.3), and deform the domain of integration for I_{n+p} to the contour consisting of two straight lines shown in Figure **??**. The remaining calculations are left to the reader. The final asymptotic relations are

$$\mathfrak{h}_{n+p}(x\sqrt{n}) \sim n^{\frac{n+p}{2}}\frac{e^{-\frac{n}{2}}}{\sqrt{\sinh\varphi}}e^{\left(p+\frac{1}{2}\right)\varphi - \frac{n}{2}(\sinh(2\varphi)-2\varphi)}, \tag{B.4.17}$$

which combines with (B.2.1) and Stirling's approximation to yield (B.4.4).

We now turn to case (c). We only present the main change of variables that underly the result. We begin with the integral representation(B.4.7) and substitute

$$t = i + \frac{r}{n^{\frac{1}{3}}}, \quad x = 2\sqrt{n} + \frac{s}{n^{\frac{1}{6}}}, \tag{B.4.18}$$

moving the integral over $\mathbb{R}$ to an integral over the line $i + \mathbb{R}$, to obtain

$$\mathfrak{h}_n(x) = (-i\sqrt{n})^n\frac{n^{\frac{1}{6}}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{nh(r)}\,dr, \tag{B.4.19}$$

where

$$h(r) = \log\left(i + \frac{r}{n^{\frac{1}{3}}}\right) - \frac{1}{2}\left(\left(i + \frac{r}{n^{\frac{1}{3}}}\right) - i\left(2 + \frac{s}{n^{\frac{2}{3}}}\right)\right)^2$$

$$\approx \frac{1}{2} + \log i + \frac{s}{n^{\frac{2}{3}}} + \frac{1}{n}\left(isr + \frac{i}{3}r^3\right) + \frac{s^2}{2n^{\frac{4}{3}}}, \tag{B.4.20}$$

using the Taylor series for the logarithm. The terms that depend on s may be pulled out of the integral and we are left with

$$\mathfrak{h}_n(x) \sim \frac{n^{\frac{n}{2}+\frac{1}{6}}}{\sqrt{2\pi}}e^{\frac{n}{2}}e^{sn^{\frac{1}{3}}}\int_{-\infty}^{\infty}e^{isr+\frac{i}{3}r^3}\,dr \tag{B.4.21}$$

$$= \sqrt{2\pi}n^{\frac{n}{2}+\frac{1}{6}}e^{\frac{n}{2}}e^{sn^{\frac{1}{3}}}\operatorname{Ai}(s).$$

We now use the definition (B.2.1) and Stirling's approximation to obtain (B.4.6).

Appendix C

Fredholm determinants

C.1 Definitions

Our purpose in this section is to explain the notion of a Fredholm determinant and resolvent in a simple and concrete setting. The ideas presented here originated in Fredholm's attempt to find a solution formula akin to Cramer's rule for linear integral equations. The notion of a determinant for an infinite-dimensional linear operator is, of course, of independent interest and has attracted the interest of many mathematicians. Simon's book provides an excellent overview of current knowledge [20].

Assume given a bounded, measurable kernel $K : [0,1] \times [0,1] \to \mathbb{R}$ and a continuous function $h : [0,1] \to \mathbb{R}$. Fix a spectral parameter $z \in \mathbb{C}$ and consider the linear integral equation

$$\varphi(x) - z \int_0^1 K(x,y)\varphi(y)\,dy = h(x), \quad x \in [0,1]. \tag{C.1.1}$$

The integral equation (C.1.1) may be written in the more compact form

$$(I - zK)\varphi = h, \tag{C.1.2}$$

where $I - zK$ denotes the bounded linear operator on $L^2([a,b])$ defined by

$$\varphi \mapsto (I - zK)\varphi, \quad (I - zK)\varphi(x) = \varphi(x) - z \int_a^b K(x,y)\varphi(y)\,dy \quad x \in [a,b]. \tag{C.1.3}$$

Integral equations such as (C.1.1) may naturally be viewed as continuum limits of linear equations. More precisely, we fix a positive integer n, consider a uniform grid $x_j = j/n$, define the vector $h_j^{(n)} = h(x_j)$, matrix $K_{j,k}^{(n)} = n^{-1}K(x_j, x_k)$, $1 \le j, k \le n$ and discretize (C.1.1) by the linear equation

$$\varphi_j^{(n)} - z \sum_{k=1}^n K_{j,k}^{(n)} \varphi_k^{(n)} = h_j^{(n)}, \quad 1 \le j \le n. \tag{C.1.4}$$

Equation (C.1.4) has a unique solution if and only if $\det(I_n - zK^{(n)}) \neq 0$. By linearity, the solution for arbitrary $h^{(n)}$ is determined by the resolvent $R^{(n)} = (I_n - zK^{(n)})^{-1}$, which is given by Cramer's rule.

$$R_{j,k}^{(n)} = (-1)^{j+k} \frac{\det(M_{jk})}{\det(I_n - zK^{(n)})}, \tag{C.1.5}$$

where M_{jk} denotes the matrix obtained from $I_n - zK^{(n)}$ by removing the j-th row and k-th column. Further, if z_j, $j = 1, \ldots n$, denote the zeros of the polynomial $\det(I_n - zK^{(n)})$, the eigenvalues of $K^{(n)}$ are given by $1/z_j$. Both these notions may be extended to (C.1.1) via the Fredholm determinant. The basic observation that allows passage to the limit is the identity

$$\det(I_n - zK^{(n)}) = \tag{C.1.6}$$

$$1 - \frac{z}{n} \sum_{j_1=1}^{N} K(x_{j_1}, x_{j_2}) + \frac{z^2}{2!} \frac{1}{n^2} \sum_{j_1,j_2=1}^{n} \begin{vmatrix} K(x_{j_1}, x_{j_1}) & K(x_{j_1}, x_{j_2}) \\ K(x_{j_2}, x_{j_1}) & K(x_{j_2}, x_{j_2}) \end{vmatrix} + \ldots$$

The coefficient of z^k in the expansion above may be computed by differentiating the left hand side k times with respect to z, and setting $z = 0$. Since K is continuous, as $n \to \infty$, the k-th term in the sum above converges to the integral

$$\frac{(-z)^k}{k!} \int_{[0,1]^k} \det\left(K(x_p, x_q)_{1 \leq p,q \leq k}\right) dx_1 \ldots dx_k. \tag{C.1.7}$$

Definition-Theorem 43. The Fredholm determinant of the operator $I - zK$ is the entire function of z defined by the convergent series

$$D(z) := \det(I - zK) = \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} \int_{[0,1]^k} \left(\det(K(x_p, x_q)_{1 \leq p,q \leq k})\right) dx_1 \cdots dx_k. \tag{C.1.8}$$

Proof. It is only necessary to show that the series(C.1.7) is convergent for all $z \in \mathbb{C}$. The determinant of a $k \times k$ matrix A with columns $a_1, \ldots, a_k$ is the (signed) volume of the parallelopiped spanned by the vectors $a_1, \ldots, a_k$. Therefore,

$$|\det(A)| \leq |a_1||a_2| \cdots |a_k| \leq \left(\max_{1 \leq j \leq k} |a_j|\right)^k. \tag{C.1.9}$$

We have assumed that K is bounded on $[0,1] \times [0,1]$, say $\max |K| \leq M < \infty$. By the inequality above,

$$|(\det(K(x_p, x_q)_{1 \leq p,q \leq k}))| \leq k^{k/2} M^k. \tag{C.1.10}$$

Thus, the k-term in the series (C.1.8) is dominated by

$$\left| \frac{(-z)^k}{k!} \int_{[0,1]^k} \left(\det(K(x_p, x_q)_{1 \leq p,q \leq k})\right) dx_1 \cdots dx_k \right|$$

$$\leq (|z|M)^k \frac{k^{k/2}}{k!} \sim \frac{1}{\sqrt{2\pi}} (|z|Me)^k \frac{1}{k^{\frac{k+1}{2}}},$$

where we have used Stirling's approximation in the last step. □

Since $D(z)$ is entire, we may differentiate term-by-term to obtain

$$\left(-\frac{d}{dz}\right)^k \det(I - zK) \tag{C.1.11}$$
$$= \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} \int_{[0,1]^{m+k}} \left(\det(K(x_p, x_q)_{1 \le p, q \le m+k})\right) dx_1 \cdots dx_{m+k}.$$

Recall that the zeros of a non-zero entire function form a discrete, countable set. The entire function $\det(I - \lambda^{-1}K)$ is an infinite-dimensional generalization of the characteristic polynomial of the matrix $K^{(n)}$ in the following sense:

Theorem 44 (Eigenvalues of K). *Assume that K is a continuous kernel. The complex number λ is an eigenvalue of K if and only if $D(\lambda^{-1}) = 0$.*

For more on Fredholm determinants, see [14, Ch.24].

C.2 Some context

C.3 Separable kernels

Appendix D

Notation

D.1 Indices

The integers m and n are reserved for the number of rows and columns of a matrix. For square matrices, we use n. The letters j and k are used to denote indices. The letter i is reserved for $\sqrt{-1}$.

D.2 Fields

$\mathbb{R}$ real numbers.

$\mathbb{C}$ complex numbers with imaginary unit $i = \sqrt{-1}$.

$\mathbb{H}$ quaternions with imaginary units $\mathfrak{e}_1$, $\mathfrak{e}_2$, $\mathfrak{e}_3$.

$\mathbb{F}$ general notation for one of the above fields.

$\mathbb{T}^n$ the n-dimensional real torus.

Σ_n the n-dimensional simplex, $x \in \mathbb{R}^n$, $\sum_{k=1}^n x_k = 1$, $x_k \geq 0$, $1 \leq k \leq n$.

For $x \in \mathbb{C}$ and $x \in \mathbb{H}$ we use $\bar{x}$ to denote the complex and quaternion conjugate respectively. The absolute value of a number $x \in \mathbb{F}$ is always denoted $|x|$. The same notation is used for the Euclidean length of a vector in $\mathbb{F}^n$, but the distinction between the two uses of $|\cdot|$ will be clear from the context. For example, for $x \in \mathbb{F}^n$, $x = (x_1, \ldots, x_n)$, $x_j \in \mathbb{F}$, $1 \leq j \leq n$, we write

$$|x|^2 = \sum_{j=1}^n |x_j|^2. \qquad \text{(D.2.1)}$$

D.3   Matrices

The fundamental spaces of matrices are denoted as follows:

$\mathbb{M}^{\mathbb{F}}_{m \times n}$ $m \times n$ matrix with entries from a field $\mathbb{F}$.

$\mathbb{M}^{\mathbb{F}}_{n \times n}$ $n \times n$ matrix with entries from a field $\mathbb{F}$.

Symm(n) real, symmetric $n \times n$ matrices.

Her(n) complex, Hermitian $n \times n$ matrices.

Quart(n) real, self-dual quaternion $n \times n$ matrices.

Jac(n)

We write $M^{\dagger}$ for the adjoint of a matrix $M \in \mathbb{M}^{\mathbb{F}}_{m \times n}$. Matrices in Symm$(n)$, Her$(n)$ and Quart(n) are self-adjoint: $M = M^{\dagger}$, but the notion of duality is distint in each setting, since the underlying field is different. For $M \in$ Symm(n), $M^{\dagger} = M^T$; if $M \in$ Her(n), then $M^{\dagger} = \bar{M}^T = M^*$; and if $M \in$ Quart(n), then $M^{\dagger} = \bar{M}^T$. All these matrices have real eigenvalues, and the matrices are said to be positive definite if all eigenvalues are *strictly positive*. We denote the subset of positive definite matrices by Symm$_+(n)$, Her$_+(n)$ and Quart$_+(n)$ respectively. The Hilbert-Schmidt norm of a matrix $M \in \mathbb{M}^{\mathbb{F}}_{n \times n}$ is denoted

$$\|M\|^2 = \text{Tr}\left(M^{\dagger}M\right) \sum_{j,k=1}^{n} |M_{j,k}|^2. \tag{D.3.1}$$

$B_r(M)$ denotes the ball of radius r centered at M in the norm $\|\cdot\|$.

D.4   Lie groups

The following notation is adopted for the classical groups and their Lie algebras.

O(n), $\mathfrak{o}(n)$ the real, orthogonal group.

SO(n), $\mathfrak{so}(n)$ the special orthogonal group.

U(n), $\mathfrak{u}(n)$ the unitary group.

USp(n), $\mathfrak{usp}(n)$ the group of unitary symplectic matrices, or the compact symplectic group.

D.5   Banach spaces

The following notation is used for standard Banach spaces.

$C(J)$ The space of continuous functions on an interval J equipped with the supremum norm.

$\mathcal{P}_J$ The space of probability measures on an interval J equipped with the weak topology.

$\langle \mu, f \rangle$ The duality pairing between measures and continuous functions on the interval J given by

$$\langle \mu, f \rangle = \int_J f(x)\, \mu(dx). \tag{D.5.1}$$

$C_0(\mathbb{R})$ The space of continuous function on $\mathbb{R}$ that vanish at infinity, equipped with the supremum norm.

$C_b(\mathbb{R})$ The space of bounded continuous function on $\mathbb{R}$ that vanish at infinity, equipped with the supremum norm.

Bibliography

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Courier Dover Publications, 1972.

[2] P. BOURGADE, *On random matrices and L-functions*, PhD thesis, Ph. D. thesis, 2009.

[3] P. DEIFT, C. D. LEVERMORE, AND C. E. WAYNE, eds., *Dynamical Systems and Probabilistic Methods in Partial Differential Equations*, vol. 31 of Lecture Notes in Applied Mathematics, American Mathematical Society, Providence, First ed., 1996, ch. Integrable Hamiltonian Systems, pp. 103–138.

[4] I. DUMITRIU AND A. EDELMAN, *Matrix models for beta ensembles*, Journal of Mathematical Physics, 43 (2002), pp. 5830–5847.

[5] F. J. DYSON, *Statistical theory of the energy levels of complex systems. I*, Journal of Mathematical Physics, 3 (1962), pp. 140–156.

[6] A. EDELMAN AND B. SUTTON, *From random matrices to stochastic operators*, Journal of Statistical Physics, 127 (2007), pp. 1121–1165.

[7] H. M. EDWARDS, *Riemann's zeta function*, vol. 58, Academic Press, 1974.

[8] J. HARER AND D. ZAGIER, *The euler characteristic of the moduli space of curves*, Inventiones Mathematicae, 85 (1986), pp. 457–485.

[9] S. HASTINGS AND J. MCLEOD, *A boundary value problem associated with the second Painlevé transcendent and the Korteweg-de Vries equation*, Archive for Rational Mechanics and Analysis, 73 (1980), pp. 31–51.

[10] E. HILLE, *Ordinary differential equations in the complex domain*, Courier Dover Publications, 1997.

[11] M. JIMBO, T. MIWA, Y. MÔRI, AND M. SATO, *Density Matrix of an impenetrable Bose Gas and the fifth Painleve Transcendent*, Physica 1D, 1 (1980), pp. 80–158.

[12] F. John, *Partial differential equations*, vol. 1 of Applied Mathematical Sciences, Springer-Verlag, New York, fourth ed., 1991.

[13] S. V. Kerov, *Asymptotic representation theory of the symmetric group and its applications in analysis*, American Mathematical Society, Providence, RI, 2003.

[14] P. Lax, *Functional Analysis*, John Wiley and Sons, 2002.

[15] B. M. McCoy and T. T. Wu, *The two-dimensional Ising model*, vol. 22, Harvard University Press, 1973.

[16] M. L. Mehta, *Random Matrices*, Elsevier, San Diego and London, third ed., 2004.

[17] M. L. Mehta and M. Gaudin, *On the density of eigenvalues of a random matrix*, Nuclear Physics, 18 (1960), pp. 420–427.

[18] J. Ramírez, B. Rider, and B. Virág, *Beta ensembles, stochastic Airy spectrum, and a diffusion*, Journal of the American Mathematical Society, 24 (2011), pp. 919–944.

[19] B. Riemann, *Collected Works of Bernhard Riemann*, Dover Publications, 1953.

[20] B. Simon, *Trace ideals and their applications*, vol. 120 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, second ed., 2005.

[21] R. P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge University Press, 2011.

[22] D. W. Stroock, *Probability theory: an analytic view*, Cambridge university press, 2010.

[23] G. 't Hooft, *A planar diagram theory for strong interactions*, Nuclear Physics B, 72 (1974), pp. 461–473.

[24] C. A. Tracy and H. Widom, *Level–Spacing Distributions and the Airy Kernel*, Communications in Mathematical Physics, 159 (1994), pp. 151–174.

[25] J.-B. Zuber, *Introduction to random matrices.* Lectures at ICTS, Bangalore, 2012.