Author:
Bailey, Emma C

Title:
Generalized Moments of Characteristic Polynomials of Random Matrices

University of Bristol

Generalized Moments of Characteristic Polynomials of Random Matrices

Emma Bailey

May 2020

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science, School of Mathematics.

ABSTRACT

A central theme of this thesis is random matrix theory and its connection to log-correlated fields. We present results on the statistical properties of characteristic polynomials of matrices from one of the three 'classical' compact matrix groups: unitary, $U(N)$; symplectic $Sp(2N)$; or orthogonal $O(2N)$ (in particular $SO(2N) \subset O(2N)$). Take $A \in G(N)$, a matrix from $G(N) \in \{U(N), Sp(2N), SO(2N)\}$, then denote by $P_N(A,\theta) = \det(I - A\exp(-i\theta))$ its characteristic polynomial. The logarithm of $|P_N(A,\theta)|$ displays logarithmic correlations. We focus too on connections; random matrix theory and its influence on number theory, probability, statistical physics, and combinatorics.

Our main results concern the study of *moments of moments* of characteristic polynomials. We give kth moment, defined with respect to the matrix group average, of the random variable corresponding to the 2βth moment of $P_N(A,\theta)$ with respect to the uniform measure $\frac{d\theta}{2\pi}$, for all $k, \beta \in \mathbb{N}$. We show that these moments of moments are polynomials in N and give the respective degrees in each unitary, symplectic, and orthogonal case.

For unitary matrices, this resolves a conjecture of Fyodorov and Keating [82] regarding the scaling of the moments with N as $N \to \infty$, for $k, \beta \in \mathbb{N}$. In the symplectic and orthogonal cases, we show that the leading order *differs* from the unitary case. Unifying all the moments of moments is the underlying integrable system. We here emphasise a connection with representation theory, giving a formulation of the moments in each case in terms of combinatorial counts. Additionally, we develop a branching model of the moments of moments and demonstrate that the Fyodorov and Keating conjecture extends to this setting.

We also analyse mixed moments of unitary characteristic polynomials asymptotically, and relate the solution to a particular Painlevé differential equation. Additionally, the asymptotic behaviour of moments of logarithmic derivatives of unitary characteristic polynomials near the unit circle are determined.

ACKNOWLEDGEMENTS

Thanks to my supervisor, Jon Keating, for all of his generous advice, support, and for sharing such interesting problems with me. Thanks too to my family (both academic and genetic), friends, and particularly Dan, for supporting me when the problems got interesting.

DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Tuesday 19th May, 2020
Emma Bailey

Contents

List of Figures

Notation

The following notation will be consistent throughout the document, unless otherwise stated.

$\mathrm{GL}(N, \mathbb{F})$	General linear group of matrices of size $N \times N$ with entries in the field $\mathbb{F}$
A	Matrix $A = (a_{ij})$
A^T	Transpose of A
$\overline{A}$	Complex conjugate of A
A^*	Matrix adjoint; $A^* = \overline{A^T} = (\overline{a_{ji}})$
I_N	$N \times N$ identity matrix, written I when the dimension is clear from context
$\Delta(x_1, \ldots, x_N)$	The $N \times N$ Vandermonde determinant
$f(x) = O(g(x))$ as $x \to \infty$	There exists some positive constant c such that for all large enough x, $\lvert f(x)\rvert \leq c\lvert g(x)\rvert$. If we write $O_{\alpha,\beta}(g(x))$ then we mean that the constant c depends on α, β
$f(x) = o(g(x))$ as $x \to \infty$	Provided $\lim_{x\to\infty} \left\lvert \frac{f(x)}{g(x)} \right\rvert = 0$
$f(x) = \Omega(g(x))$ as $x \to \infty$	Provided there exists some positive constant c such that $\limsup_{x\to\infty} \left\lvert \frac{f(x)}{g(x)} \right\rvert \geq c$
$f(x) \sim g(x)$ as $x \to \infty$	Provided $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$

Chapter 1

Introduction

1.1 Background

The study of random matrices is expansive. From modelling buses in Mexico (see Krbálek and Šeba [118] and Baik et al. [12]) to quantum systems (see for example Wigner [158, 159]), random matrix theory is applicable to a broad range of mathematics. For this reason, it is prudent to direct our gaze here to the necessary players.

The original results covered in chapters 3, 4, 6 and 7, as well as the computations featuring in chapter 5 all, naturally, concern random matrices. In particular, we study various moments of characteristic polynomials. However, the motivation for the problems presented, and the proof techniques used to resolve them, have a much broader ancestry.

Thus, this section is devoted to covering the groundwork necessary in order to present our results fully. In section 1.1.1, we focus on the elements of random matrix theory that we will require: namely the classical compact groups and their characteristic polynomials. Section 1.1.2 introduces the relevant number theoretic functions to which (perhaps initially surprisingly) our results are connected. The topic of extremal value theory is presented in section 1.1.3 since our work is related to various conjectures of Fyodorov and Keating [81, 82] on extreme values of characteristic polynomials. Since the logarithm of the real part of the characteristic polynomial of a unitary matrix displays logarithmic correlations, section 1.1.4 outlines the foundations of log-correlated fields. Additionally, suitably renormalized, the characteristic polynomial of a unitary matrix converges to the Gaussian multiplicative chaos measure; therefore we introduce the theory in section 1.1.5. Finally, in section 1.1.6, we cover symmetric function theory. It transpires that moments of characteristic polynomials can be rephrased in terms of symmetric functions, and we will regularly exploit such a connection.

1.1.1 Random matrix theory

Within this thesis we will be concerned with the study of matrices from the classical compact matrix groups, namely *unitary*, *symplectic*, and *orthogonal*. The canonical example will always be unitary matrices, with symplectic and orthogonal matrices being unitary matrices with further restrictions. For a more general account of these topics, and many more, see [1, 124].

Definition 1.1.1 (Unitary Group)**.** *Denoted by $U(N)$, the unitary group of $N \times N$ matrices is*

$$U(N) := \{A \in \mathrm{GL}(N, \mathbb{C}) : A^*A = AA^* = I\}.$$

Recall that we write $A^* = \overline{A^T}$ for the conjugate transpose of a matrix A. A key property of unitary matrices (in fact any of the matrices considered henceforth), is that their eigenvalues lie on the unit circle in the complex plane, see figure 1.1. The remaining two matrix groups that we require are the following.

Definition 1.1.2 (Symplectic Group)**.** *Denoted by $Sp(2N)$, the (unitary) symplectic group*[1] *of $2N \times 2N$ matrices is*

$$Sp(2N) := \{A \in U(2N) : A\Omega A^T = \Omega\},$$

where Ω is the block matrix

$$\Omega := \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix}. \tag{1.1.1}$$

The eigenvalues of symplectic matrices come in complex conjugate pairs.

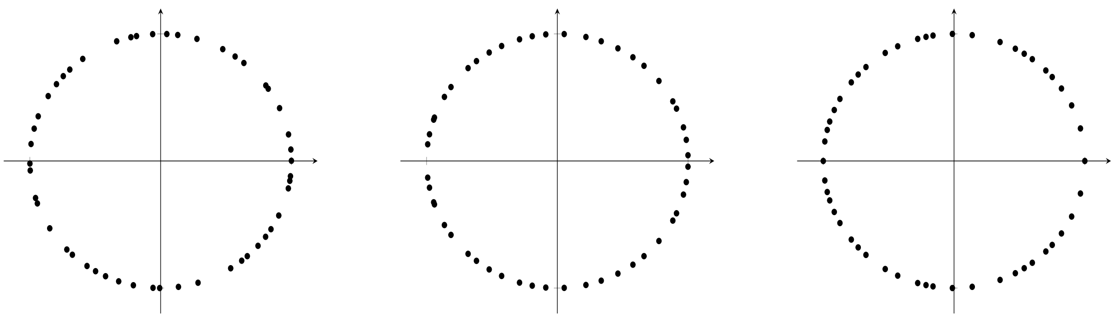
Definition 1.1.3 (Orthogonal Group)**.** *Denoted by $O(N)$, the orthogonal group of $N \times N$ matrices is*

$$O(N) := \{A \in \mathrm{GL}(N, \mathbb{R}) : AA^T = A^T A = I\}.$$

The special orthogonal group, $SO(N)$ is the subgroup of orthogonal matrices with determinant $+1$. When we specialise to even *dimensional special orthogonal matrices, the eigenvalues come in complex conjugate pairs, and we will commonly write*

$$SO(2N) := \{A \in O(2N) : \det(A) = +1\}$$

for the group of such matrices.



(a) Unitary. (b) Symplectic. (c) Special Orthogonal.

Figure 1.1: Comparing the eigenvalues of random 50×50 unitary (1.1a), symplectic (1.1b), and special orthogonal (1.1c) matrices.

Let $A \in G(N)$ be a matrix from one of groups $G(N) \in \{U(N), Sp(2N), SO(2N)\}$. We write for its characteristic polynomial

$$P_{G(N)}(A, \theta) := \det(I - Ae^{-i\theta}), \tag{1.1.2}$$

Since A will have its eigenvalues on the unit circle, we write the polynomial variable as $e^{-i\theta}$ and consider real θ. Furthermore, if $A \in Sp(2N)$ or $A \in SO(2N)$ then, whilst the matrix size is $2N$, the eigenvalues of A come in N complex conjugate pairs, hence the subscript N rather than $2N$. Whenever we focus on one particular compact group then for notational simplicity we will often, temporarily, write $P_N \equiv P_{G(N)}$.

Recall that a matrix A from any of the classical compact groups $G(N) \in \{U(N), Sp(2N), O(N)\}$ will have its eigenvalues on the unit circle. For $G(N) = U(N)$, we may express them as $e^{i\theta_1}, \ldots, e^{i\theta_N}$,

[1] Sometimes written in the literature as $USp(2N)$.

with $\theta_j \in [0, 2\pi)$. However, for $G(N) \in \{Sp(2N), SO(2N)\}$, one has to take in to account additional symmetries. For example, take $S \in Sp(2N)$. Its eigenvalues come in conjugate pairs $e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_N}$ for $\theta_j \in [0, \pi)$. Thus its characteristic polynomial factors as

$$P_{Sp(2N)}(S, \theta) = \det(I - Se^{-i\theta}) = \prod_{j=1}^{N}(1 - e^{i\theta_j}e^{-i\theta})(1 - e^{-i\theta_j}e^{-i\theta}). \tag{1.1.3}$$

Similar notation is used, for example, for $O \in SO(2N)$.

The benefit of working with $U(N)$, $Sp(2N)$, and $O(N)$ (and its subgroup $SO(N)$) is that they are compact Lie groups and hence one can endow each group with a uniform measure, namely a *Haar* measure. When we take $U(N)$ together with its Haar measure, they form the *Circular Unitary Ensemble*, or CUE. This ensemble was first introduced by Dyson [70], together with the Circular Orthogonal and Circular Symplectic Ensembles, COE and CSE respectively[2].

One way to express the Haar measure μ_{Haar} on a particular compact group is via the explicit formulae of Weyl [156]. We use the unitary group as an example. Firstly, we define a *class function* on $A \in U(N)$ by the property that $f(A) \coloneqq f(\theta_1, \ldots, \theta_N)$ (where the θ_j are the eigenphases of A) is symmetric in all of its variables. Then,

$$\int_{U(N)} f(A) d\mu_{\text{Haar}}(A) = \frac{1}{(2\pi)^N N!} \int_0^{2\pi} \cdots \int_0^{2\pi} f(\theta_1, \ldots, \theta_N) \prod_{1 \le j < k \le N} \left| e^{i\theta_j} - e^{i\theta_k} \right|^2 d\theta_1 \cdots d\theta_N. \tag{1.1.4}$$

Often we rewrite terms such as the product in integrand in the right hand side of (1.1.4) using the *Vandermonde determinant* notation. The square $N \times N$ Vandermonde matrix V is

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{N-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^{N-1} \end{pmatrix}, \tag{1.1.5}$$

so $V_{i,j} = x_i^{j-1}$. The determinant of V is the *Vandermonde determinant*, which we write as

$$\Delta(x_1, \ldots, x_N) \coloneqq \det(V) = \prod_{1 \le i < j \le N} (x_j - x_i). \tag{1.1.6}$$

Hence (1.1.4) can be rewritten in this notation as

$$\int_{U(N)} f(A) d\mu_{\text{Haar}}(A) = \frac{1}{(2\pi)^N N!} \int_0^{2\pi} \cdots \int_0^{2\pi} f(\theta_1, \ldots, \theta_N) \left| \Delta(e^{i\theta_1}, \ldots, e^{i\theta_N}) \right|^2 d\theta_1 \cdots d\theta_N. \tag{1.1.7}$$

A key fact of the Haar measure is that it is invariant under unitary transformations, i.e.

$$d\mu_{\text{Haar}}(A) = d\mu_{\text{Haar}}(UAU^*) \tag{1.1.8}$$

for $U \in U(N)$. This gives an insight as to why we would expect a formula such as (1.1.4) to hold.

[2]It is important to emphasise that, unlike with the CUE, the COE and CSE are not simply $O(N)$ or $Sp(2N)$ together with their respective Haar measures. See Mehta [124] for further details.

Notice that A may be diagonalised so that

$$A = U \begin{pmatrix} e^{i\theta_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{i\theta_N} \end{pmatrix} U^* \tag{1.1.9}$$

for some $U \in U(N)$. Thus, using the invariance of the measure, integrals over A can essentially be seen as integrals over the eigenphases of A. Similar formulae hold for symplectic and orthogonal matrices, see Weyl [156]. Henceforth, for ease of notation, we may write

$$dA := d\mu_{\text{Haar}}(A) \tag{1.1.10}$$

whenever which Haar measure we are referring to is clear from context.

1.1.2 Number theory

One of the major developments in the recent history of random matrix theory is the connection to number theory. We here review the relevent number theoretic concepts. We begin with the definition of one of the central functions of number theory.

Definition 1.1.4 (Riemann zeta function)**.** *Let* $s \in \mathbb{C}$ *with* $\text{Re}(s) > 1$. *Then the Riemann zeta function is defined by*

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

In general, summations of the type

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} \tag{1.1.11}$$

for $s, a_n \in \mathbb{C}$ are known as *Dirichlet series*. Hence, $\zeta(s)$ can be defined by a Dirichlet series with $a_n = (1, 1, 1, \dots)$. Equivalently, $\zeta(s)$ can also be expressed as a product over primes, known as an Euler product,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \tag{1.1.12}$$

also for $\text{Re}(s) > 1$. Here and for the remainder of the thesis, products of the form appearing in (1.1.12) are over primes p, unless otherwise explicitly stated. The connection between the summation presentation and the Euler product formulation follows by the fundamental theorem of arithmetic.

One can analytically extend $\zeta(s)$ to all of the complex plane, with the exception of a simple pole at $s = 1$ (many proofs of this fact can be found, for example, in [71]). A consequence of the meromorphic continuation is the *functional equation*,

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\tfrac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s), \tag{1.1.13}$$

where $\Gamma(z)$ is the usual analytic extension of the factorial function, taking the value $(z-1)!$ whenever $z \in \mathbb{N}$.

Using (1.1.13), it is easy to see that there are zeros at the negative even integers coming from the sine function. These are known as *trivial* zeros. The other zeros at positive even integers due to $\sin(\frac{\pi s}{2})$ are cancelled by the simple poles attributed to $\Gamma(1-s)$. The functional equation also leads to many symmetries for the remaining *non-trivial* zeros. If one denotes by ρ_n any zero of $\zeta(s)$ other

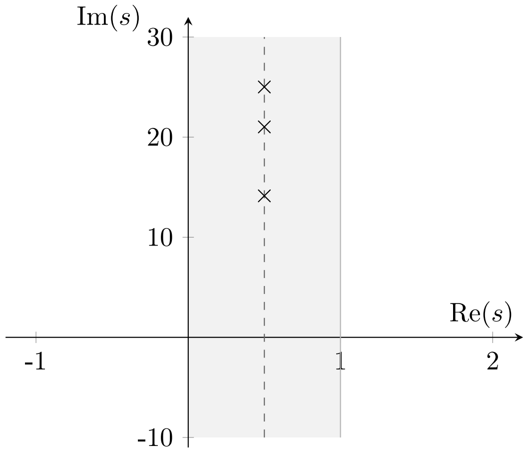


Figure 1.2: Plot showing the critical strip for $\zeta(s)$ and the first three (positive) non-trivial zeros. Provably all non-trivial zeros of $\zeta(s)$ lie within the shaded region (i.e. $0 < \mathrm{Re}(s) < 1$), and the Riemann hypothesis states that all should lie on the dashed line (i.e. $\mathrm{Re}(s) = 1/2$).

than those at the negative even integers (i.e. the non-trivial ones), then a consequence of the prime number theorem[3] of Hadamard and de la Vallée Poussin (proved independently, see for example [154]) is that ρ_n must lie in the *critical strip*, $0 < \mathrm{Re}(s) < 1$. The infamous Riemann hypothesis, introduced by Riemann [140], says that the non-trivial zeros in fact lie in the centre of the critical strip, see also figure 1.2.

Conjecture 1.1.5 (Riemann hypothesis). $\mathrm{Re}(\rho_n) = \frac{1}{2}$ *for all* n*.*

One often refers to the line $\mathrm{Re}(s) = 1/2$ as the 'critical line' or the 'half line'.

At time of writing, the record for the proportion of non-trivial Riemann zeros that provably lie on the half line is 0.41729 ('more than 5/12'), due to Pratt et al. [134]. We note that there has only been incremental, though nevertheless impressive, progress since the result of Conrey [43] which gave 'more than 2/5' of zeros on the critical line - for instance, interim improvements to 'more than 41%' due to Bui et al. [31]. Interestingly in the context of this thesis, the methods used by Pratt et al. share their origins with many of the results presented in the subsequent chapters via the 'Ratios Theorem' (see in particular chapter 7, section 7.1.4).

Moving beyond the non-trivial zeros of $\zeta(s)$, a related number theoretic question is the size of $\zeta(1/2 + it)$ for $t \in \mathbb{R}$, either on average or the exceptionally large values. Selberg proved the following central limit theorem for $\zeta(s)$ [144].

Theorem 1.1.6 (Selberg's Central Limit theorem [144]). *For any rectangle* $B \subset \mathbb{C}$*,*

$$\lim_{T\to\infty} \mathbb{P}\left\{ T \le t \le 2T : \frac{\log \zeta(\frac{1}{2} + it)}{\sqrt{\frac{1}{2}\log\log\frac{t}{2\pi}}} \in B \right\} = \frac{1}{2\pi}\int\int_B e^{-\frac{1}{2}(x^2+y^2)} dxdy.$$

Thus, both the real part and the imaginary part of the logarithm of the zeta function independently tend to a Gaussian random variable. Note that this means that the typical size of $\log|\zeta(1/2 + it)|$ is $O(\sqrt{\log\log t})$.

More generally, one is interested in determining both the size of the moments of $\zeta(s)$ over stretches of the critical line,

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2} + it)|^{2\beta} dt, \tag{1.1.14}$$

[3] In broad terms, the prime number theorem states that the number of primes up to some given positive number x grows asymptotically like $x/\log(x)$.

and the extreme values of $\zeta(1/2+it)$. Such questions have motivated much of the work described in chapter 2, as well as our original work, which is the subject of the remaining chapters. There are also natural extensions to the Riemann zeta function via (1.1.11), known as L-functions. These, and their associated moments, are also covered in chapter 2 (see in particular section 2.1.3).

1.1.3 Extreme value theory

As mentioned in section 1.1.2, we will be interested in the extreme values of various functions (for example, characteristic polynomials, or $\zeta(s)$). The study of extreme values is the study of rare events, and a general introduction to the topic can be found in Leadbetter, Lindgren, and Rootz [121], and de Haan and Ferreira [57]. A central result within the field is the Fisher-Tippett-Gnedenko Theorem [88].

Theorem 1.1.7 (Fisher-Tippett-Gnedenko Theorem)**.** *Let $X_1, X_2, \dots, X_n$ be independent and identically distributed (iid) random variables. Define*

$$M_n \coloneqq \max\{X_1, \dots, X_n\}.$$

If there exists $a_n > 0$ for all n, and b_n such that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = F(x),$$

where F is non-degenerate, then F is an extreme value cumulative distribution function and belongs to one of three classes:

$$\begin{aligned}
&\text{(I)} \quad \Lambda(x) = e^{-e^{-x}} \quad \text{for } x \in \mathbb{R},\\
&\text{(II)} \quad \Phi_\alpha(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-x^{-\alpha}} & \text{if } x > 0, \end{cases} \quad \text{for some } \alpha > 0,\\
&\text{(III)} \quad \Psi_\alpha(x) = \begin{cases} e^{-(-x)^{\alpha}} & \text{if } x < 0 \\ 1 & \text{if } x \geq 0, \end{cases} \quad \text{for some } \alpha > 0.
\end{aligned}$$

Type (I), Type (II), and Type (III) are known as Gumbel, Fréchet, and Weibull distributions respectively. Of particular interest within the context of this exposition is the following result, see for example Leadbetter et al. [121], Theorem 1.5.3.

Theorem 1.1.8. *Let $\{Z_1, Z_2, \dots\}$ be independent and identically distributed standard Gaussian random variables. As in theorem 1.1.7, let $M_n \coloneqq \max\{Z_1, \dots, Z_n\}$. Then define*

$$a_n \coloneqq \frac{1}{\sqrt{2\log n}}, \tag{1.1.15}$$

and

$$b_n \coloneqq \sqrt{2\log n} - \frac{\log\log n + \log 4\pi}{2\sqrt{2\log n}}. \tag{1.1.16}$$

Then

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = e^{-e^{-x}}. \tag{1.1.17}$$

Thus, after rescaling, the maximum of a collection of standard Gaussian random variables has a Gumbel (Type (I)) distribution.

Since the proof highlights some key techniques, and the correct shape of the maximum M_n, we sketch it here. For full details see for example [121]. The statement of theorem 1.1.8 is an important example against which we will compare many processes, see for example section 1.1.4 and the extended discussion of the literature in chapter 2.

We may trivially rewrite

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n \leq u_n) = (1-(1-\Phi(u_n)))^n \tag{1.1.18}$$

where $u_n = u_n(x) = a_n x + b_n$ and Φ is the usual standard Gaussian distribution function. Now, if one can find a_n, b_n defining u_n such that

$$1 - \Phi(u_n) = \frac{1}{n}\tau(x) + o\left(\frac{1}{n}\right) \tag{1.1.19}$$

for $\tau(x)$ some function of x to be determined, then

$$\lim_{n\to\infty} \mathbb{P}(M_n \leq u_n) = e^{-\tau(x)}. \tag{1.1.20}$$

Using classical Gaussian tail asymptotics, we have that $1 - \Phi(u_n) \sim \phi(u_n)/u_n$, where $\phi(x)$ is the standard normal density function. Hence one would need

$$\frac{1}{n}\tau(x) = \frac{1}{u_n\sqrt{2\pi}} e^{-\frac{1}{2}u_n^2} \tag{1.1.21}$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{a_n x + b_n} e^{-\frac{1}{2}(a_n^2 x^2 + b_n^2 + 2a_n b_n x)}. \tag{1.1.22}$$

Analysing (1.1.22) at $x = 0$ implies that

$$b_n = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi}{2\sqrt{2\log n}}, \tag{1.1.23}$$

which in turn implies

$$a_n = \frac{1}{\sqrt{2\log n}}. \tag{1.1.24}$$

We thus conclude that, with this choice of a_n, b_n,

$$\lim_{n\to\infty} \mathbb{P}\left(M_n \leq a_n x + b_n\right) = \lim_{n\to\infty} (1-(1-\Phi(u_n)))^n \tag{1.1.25}$$

$$= \lim_{n\to\infty}\left(1 - \frac{1}{n}e^{-x}\right)^n \tag{1.1.26}$$

$$= e^{-e^{-x}} \tag{1.1.27}$$

as required.

Thus, we have shown that for standard Gaussian random variables the approximate size of the maximum is

$$M_n \approx b_n + a_n \mathcal{M} \tag{1.1.28}$$

where $\mathcal{M}$ is a Gumbel random variable. It will soon be useful to consider a dyadic number of variables. Take $Y_1, \ldots, Y_{2^n}$ Gaussian random variables, centred and with variance $\sigma^2 n$. Set

$$M_{2^n} := \max\{Y_1, \ldots, Y_{2^n}\}. \tag{1.1.29}$$

Using the above technique, one finds that

$$a_n = 1, \tag{1.1.30}$$
$$b_n = cn - \frac{1}{2}\frac{\sigma^2}{c}\log n, \tag{1.1.31}$$

where $c = \sqrt{2\sigma^2 \log 2}$. Hence,

$$M_{2^n} \approx cn - \frac{\sigma^2}{c}\log n + \mathcal{M} \tag{1.1.32}$$

where $\mathcal{M}$ has a Gumbel distribution. The fact that the leading order is linear in n, and the subleading term is logarithmic in n with a coefficient of $1/2$ will be an informative comparison with statements in the next section.

Further, we will also often make the choice $\sigma^2 = (1/2)\log 2$ (i.e. each of the 2^n random variables are distributed as $Y_j \sim \mathcal{N}(0, (1/2)\log 2^n)$). In this particular case

$$M_{2^n} \approx \log 2^n - \frac{1}{4}\log\log n + \mathcal{M} \tag{1.1.33}$$
$$= n\log 2 - \frac{1}{4}\log\log n + \mathcal{M}. \tag{1.1.34}$$

1.1.4 Log-correlated fields

We give a brief recap of the key elements needed hereinafter; for a more thorough review of the wider research area see Duplantier et al. [69], and Arguin [4] for a survey with similar aims to this exposition.

Most generally, one can define a stochastic process $X_n = \{X_n(v) : v \in V_n\}$ on a metric space V_n with a distance $|\cdot|$, so that the dimension of the space $\dim V_n$ depends on n. Then, the defining feature of log-correlated fields is the form of the covariance[4],

$$\mathbb{E}[X_n(v)X_n(w)] \approx -\log|v - w|, \tag{1.1.35}$$

for $v, w \in V_n$.

The choice of discrete field V_n with 2^n points will prove to be a particularly instructive example. Take a log-correlated field $X_n = \{X_n(v), v \in V_n\}$ for this choice of V_n, so that X_n has covariance as defined by the right hand side of (1.1.35), and such that the $X_n(v)$ are centred with variance $\mathbb{E}[X_n(v)]^2 = \sigma^2 n$.

Then one expects that the maximum $M_{2^n} = \max_{v\in V_n} X_n(v)$ behaves as

$$M_{2^n} \approx cn - \frac{3}{2}\frac{\sigma^2}{c}\log n + \mathcal{M}, \tag{1.1.36}$$

where $c = \sqrt{2\sigma^2 \log 2}$ and $\nu = \dim V_n$. The distribution of $\mathcal{M}$ is expected to be *no longer Gumbel*[5].

We here highlight the similarities and differences between the maximum of log-correlated processes (1.1.36) and the maximum of independent Gaussian random variables, see (1.1.32). For both, the leading order of the maximum is linear in n, and the subleading term is of the order $\log n$, but the subleading coefficient differs between the cases: $-1/2$ versus $-3/2$. If the random variables are all

[4] Here we write '≈' to encompass any covariance structure which has a logarithmic singularity when the two points meet.

[5] For the log-correlated fields that we will henceforth be considering, in fact it is believed that the density of $\mathcal{M}$ is

$$2e^{-x}K_0(2e^{-\frac{x}{2}}), \tag{1.1.37}$$

where $K_0(z)$ is the modified Bessel function of the second kind. As observed by Kundu et al. [119], the density in (1.1.37) matches that for the sum of two independent Gumbel random variables. Further discussion of this can be found in chapter 2.

independent, then the maximum isn't 'pulled down' as much as when the variables are log-correlated. Such behaviour is expected to be universal within each case.

Important examples of log-correlated processes in the context of this document are branching random walks and the logarithm of the characteristic polynomial of a random unitary matrix. Further, fundamental models within the physics literature involve log-correlated fields. We now discuss two rich models exhibiting such logarithmic correlations: the generalized Random Energy Model (GREM) and branching random walks. The models will serve as informative examples for chapter 2, and lay the groundwork for the branching model featuring in chapter 6.

The Random Energy Model

The 'Random Energy Model' (REM), first introduced by Derrida [62], is a spin glass model from condensed matter physics. The REM is a stochastic process on the hypercube $\{-1,1\}^n$. For each point in the state space, one associates the independent random variable $X_n(v) \in \mathcal{N}(0,n)$.

Translating in to the language of log-correlated fields, we have the process $X_n = \{X_n(v) : v \in \{-1,1\}^n\}$ (so the dimension of the space is 2^n). From the point of view of statistical physics, the partition function is central to investigations, and for the REM is defined as

$$Z_n(\beta) := \sum_{v\in\{-1,1\}^{2^n}} \exp(-\beta X_n(v)), \tag{1.1.38}$$

where β represents the inverse temperature of the system (so $\beta > 0$), with energy $X_n(v)$. Note that the maximum of the REM will follow (1.1.32) with $\sigma^2 = 1$. Additionally, the *free energy* associated to $Z_n(\beta)$ is

$$f_n(\beta) = -\frac{1}{\beta}\log Z_n(\beta). \tag{1.1.39}$$

The free energy is related to the extreme values of the energy, $M_{2^n} := \max_{v\in\{-1,1\}^n} X_n(v)$, via

$$\lim_{\beta\to\infty} f_n(\beta) = -\lim_{\beta\to\infty}\frac{1}{\beta}\log Z_n(\beta) = -M_{2^n}, \tag{1.1.40}$$

and hence whenever the following limits are well-defined,

$$\lim_{\beta\to\infty}\lim_{n\to\infty}\frac{f_n(\beta)}{n} = -\lim_{n\to\infty}\frac{M_{2^n}}{n}. \tag{1.1.41}$$

This shows that one can recover the correct order of the maximum of the system together with its coefficient by studying the free energy. Such an approach will prove critical in later chapters.

One can generalize the REM so that the random variables $X_n(v)$, $v \in \{-1,1\}^{2^n}$ are no longer independent, but instead depend on the distance $|v-w|$ for $v,w \in \{-1,1\}^{2^n}$. This is a *generalized* random energy model (GREM). Adaptations in particular which introduce *logarithmic* correlations of the form (1.1.35) have inspired great research interest, for example [34, 63, 78, 82], and are particularly relevent for this thesis. Once again, a similar process to that described above yields a connection between the free energy for the GREM and its maximum, which should instead now follow (1.1.36).

Branching Random Walks

Perhaps the simplest log-correlated case where the '3/2' coefficient can be proven to appear is the case of a Gaussian random walk on a binary tree. This is also sometimes referred to as the hierarchical Gaussian field. We first define such a process. Take a rooted binary tree of depth n and, in the language introduced at the start of section 1.1.4, we let V_n be the leaves of such a tree. For a fixed leaf

$v \in \{1, \ldots, 2^n\}$, the random variable $X_n(v)$ is given by

$$X_n(v) = \sum_{m=1}^{n} Y_m(v), \tag{1.1.42}$$

where $Y_m(v) \sim \mathcal{N}(0, \sigma^2)$. Thus, $X_n(v)$ is a random walk from root to leaf v, where at each level $m \in \{1, \ldots, n\}$, an independent and identically distributed Gaussian random variable is collected. Clearly, the distribution of $Y_m(v)$ does not depend on m, v, but we retain the notation to make the connection with the level and leaf within the binary tree clear. Note also for a comparison with the (G)REM, that $X_n(v) \sim \mathcal{N}(0, n\sigma^2)$. Figure 1.3 illustrates the process.

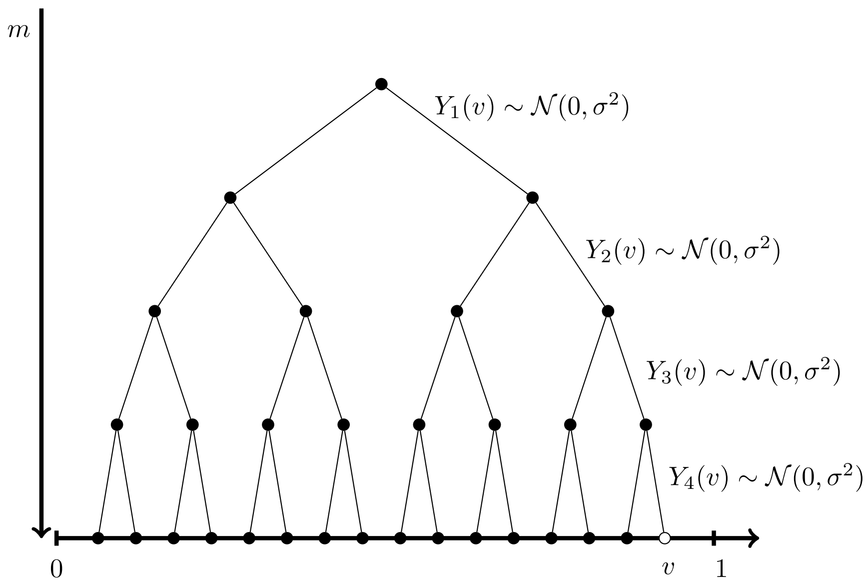


Figure 1.3: An example of a random walk $X_n(v)$ on a binary tree of depth $n = 4$, from root to leaf v. The weightings are Gaussian random variables with mean 0 and variance σ^2.

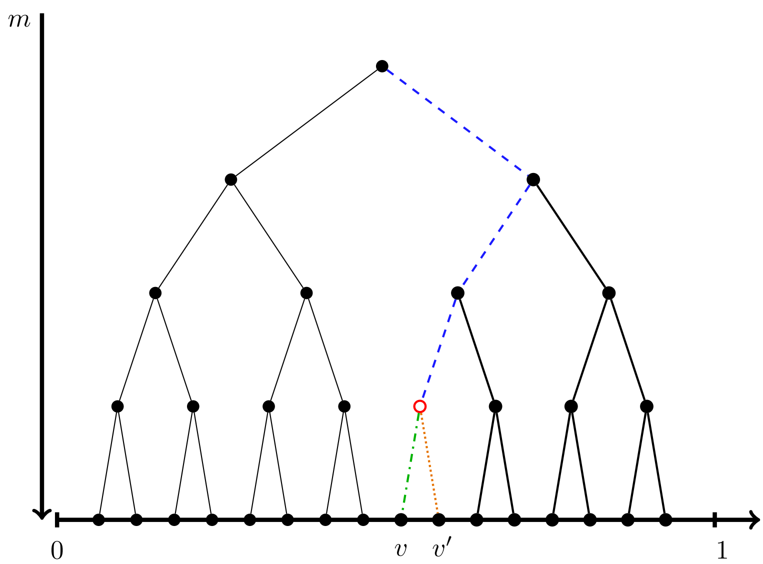


Figure 1.4: A tree structure of depth 4 with an example of splitting. The last common ancestor of leaves v, v' is illustrated by the 'hollow' (red) node and occurs at level 3.

We remark here that the binary tree random walk is crucial not only for determining maxima of other log-correlated processes (see detailed discussions on the recent literature in chapter 2), but also for the novel work presented in chapter 6.

To see that X_n is log-correlated, one determines the covariance structure. Firstly, one clearly has

$$\mathbb{E}[X_n(v)] = 0 \tag{1.1.43}$$

$$\mathbb{E}[X_n(v)^2] = \sigma^2 n. \tag{1.1.44}$$

To determine the covariance, we need to define the *last common ancestor* of two leaves v, v'.

Definition 1.1.9 (Last common ancestor)**.** *The last common ancestor of two leaves v and v' of a binary tree, written* $\mathrm{lca}(v, v')$*, is the level of the first point at which the paths from root to v and from root to v' diverge. Figure 1.4 gives an example of last common ancestors.*

Then, one calculates that

$$\mathbb{E}[Y_j(v)Y_j(v')] = \begin{cases} \sigma^2, & \text{if } |v - v'| \le 2^{-j} \\ 0, & \text{if } |v - v'| > 2^{-j}, \end{cases} \tag{1.1.45}$$

since the leaves are equally spaced within the interval $[0, 1]$. Thus, the covariance of $X_n(v)$ and $X_n(v')$ depends on the last common ancestor of v, v'. Specifically,

$$\mathbb{E}[X_n(v)X_n(v')] = \sigma^2 \cdot \mathrm{lca}(v, v'). \tag{1.1.46}$$

This shows that $X_n(v)$ displays the symptoms of a log-correlated process. More precisely, take $0 \le r \le 1$ such that $rn \in \mathbb{N}$. Then consider, for a fixed leaf $v \in \{1, \dots, 2^n\}$, the proportion of neighbours w whose covariance with v is at least $\sigma^2 nr$. It is precisely those w whose last common ancestor is at level rn:

$$\frac{1}{2^n} \# \left\{ w \in \{1, \dots, 2^n\} : \frac{\mathbb{E}[X_n(v)X_n(w)]}{\mathbb{E}[X_n(v)^2]} \ge r \right\} = 2^{-rn}. \tag{1.1.47}$$

Such a property is indicative of a log-correlated system, see for example [4].

Given that the branching random walk has a log-correlated structure, a natural question is to investigate its maximum. Bramson [29] was the first to identify to subleading order (and, notably, the subleading constant 3/2) the maximum of the above field X_n.

Theorem 1.1.10 (Bramson [29])**.** *Let $X_n(v)$ be a branching random walk from root to leaf v on a binary tree of depth n, with centred Gaussian increments of variance σ^2. Set $M_{2^n} := \max_{v \in \{1,\dots,2^n\}} X_n(v)$. Then*

$$M_{2^n} = cn - \frac{3}{2} \frac{\sigma^2}{c} \log n + x \tag{1.1.48}$$

where $c = \sqrt{2\sigma^2 \log 2}$ and x is a bounded fluctuating term.

One sees the 'log-correlated' constant 3/2 appearing; the importance of the canonical example of branching random walks on binary trees is evident. Such processes will make frequent reappearances hereafter. In particular, we will often set $\sigma^2 = (1/2) \log 2$. Given this choice, (1.1.48) becomes

$$M_{2^n} = n \log 2 - \frac{3}{4} \log n + x. \tag{1.1.49}$$

Given the regularity with which we will set $\sigma^2 = (1/2) \log 2$, we may equally refer to the 'log-correlated' constant as being 3/4, (making the 'independent' constant 1/4).

1.1.5 Gaussian multiplicative chaos

Within this section, we give a top-level account of Gaussian multiplicative chaos (GMC) and relevant tools from the perspective of this thesis. For an excellent review of the topic, we direct the reader to the paper of Rhodes and Vargas [138].

The origins of GMC trace back to the work of Kahane [108], who introduced the theory for understanding the exponential of a Gaussian field whose covariance has a logarithmic singularity. To this end, take $D \subset \mathbb{R}^d$ a subdomain, and $X = \{X(v) : v \in D\}$ a Gaussian field so that

$$\mathbb{E}[X(v)] = 0, \tag{1.1.50}$$

and

$$\begin{aligned} \mathbb{E}[X(v)X(w)] &= \max\{-\log|v-w|, 0\} + g(v,w) && (1.1.51)\\ &\sim -\log|v-w|, && (1.1.52) \end{aligned}$$

as $v \to w$, and for g some bounded function over $D \times D$. Clearly the covariance (1.1.52) implies a connection to the log-correlated fields discussed previously.

The log-singularity present in (1.1.52) is precisely the cause of the difficulty when constructing the measure associated with the exponential of the field,

$$e^{\gamma X(v) - \frac{\gamma^2}{2}\mathbb{E}[X(v)^2]}dx, \tag{1.1.53}$$

for some $\gamma \in \mathbb{R}$. A natural solution is to 'regularize' the field X: introduce a smooth cut-off $X_n(v)$ so that in the large n limit, $X_n(v) \to X(v)$. For such a cut-off, if one is able evaluate

$$e^{\gamma X_n(v) - \frac{\gamma^2}{2}\mathbb{E}[X_n(v)^2]}dx \overset{n\to\infty}{\longrightarrow} \mu_\gamma(dx), \tag{1.1.54}$$

for some limiting measure $\mu_\gamma(dx)$, then one may define (1.1.53) to be said limit. Kahane's theorem [108], stated below, defines such an X_n and shows that the limiting measure μ_γ is only non-trivial for a certain range of γ.

Theorem 1.1.11 (Kahane [108]). *Let $v, w \in D$ and assume that there exists a continuous and bounded function $g : D \times D \to \mathbb{R}$ such that*

$$\mathbb{E}[X(v)X(w)] = \max\{-\log|v-w|, 0\} + g(v,w), \tag{1.1.55}$$

and further that the covariance has a decomposition

$$\mathbb{E}[X(v)X(w)] = \sum_{l=1}^{\infty} K_l(v,w), \tag{1.1.56}$$

for K_l continuous and positive definite covariance kernels. Let Y_l be the Gaussian field with mean 0 and covariance given by K_l, such that Y_l is independent from $Y_{l'}$ for $l \neq l'$. Set $X_n = Y_1 + \cdots + Y_n$. Then for $\gamma \in \mathbb{R}$, the measures

$$\mu_{\gamma,n}(dx) := e^{\gamma X_n(x) - \frac{\gamma^2}{2}\sum_{l=1}^{n} K_l(x,x)}dx, \tag{1.1.57}$$

converge almost surely in the space of Radon measure (for topology given by weak convergence) to some random measure $\mu_\gamma(dx)$. This convergence is independent of the regularization of X. The measure

μ_γ is the zero measure for $\gamma^2 \geq 2d$ (where recall our field is defined with respect to $D \subset \mathbb{R}^d$), and non-trivial for $\gamma^2 < 2d$.

For a final comment on GMC measures, we examine the range for which such measures are non-trivial: $\gamma^2 < 2d$. Frequently, the range $[0, 2d)$ is broken up in to two sections, known as the L^1- and L^2-phase:

$$L^2\text{−phase}: \quad 0 \leq \gamma^2 < d, \tag{1.1.58}$$

$$L^1\text{−phase}: \quad d \leq \gamma^2 < 2d. \tag{1.1.59}$$

When demonstrating convergence to the GMC measure for a particular field X, it is often the case that different techniques are required between the ranges, as shown in chapter 2, section 2.2.3. Additionally we remark that, as formulated above, the measure $\mu_\gamma(dx)$ is trivial for $\gamma^2 \geq 2d$. It is possible to construct a GMC measure for $\gamma^2 \geq 2d$. The case of $\gamma^2 = 2d$ yields a phase transition and is known as *critical chaos*. Conjecturally, see for example [138], all constructions for the critical chaos measure are the same. In the *super-critical* regime of $\gamma^2 > 2d$, one can define a new class of chaos known as *atomic multiplicative chaos*, and we refer the reader to [138] and the references within for further details.

1.1.6 Symmetric function theory

A function f in n variables is *symmetric* if it is invariant under permutations of its arguments. That is, if $\sigma \in S_n$ (the group of permutations on n symbols) and if

$$f(x_1, x_2, \ldots, x_n) = f(\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_n)), \tag{1.1.60}$$

then f is a symmetric function.

It will transpire that the study of symmetric functions is intimately connected with moments of characteristic polynomials. For this review, we only require a small, though mathematically rich, portion of the field. For a comprehensive account of the wider theory, see the books of Stanley and Macdonald [122, 147, 148].

We begin with a fundamental definition.

Definition 1.1.12 (Partition)**.** *Let $l \in \mathbb{N}$. A* partition *λ of length $l = l(\lambda)$ is a sequence of non-increasing non-negative integers with l non-zero elements. Thus, if $\lambda = (\lambda_i)_{i=1}^l$ is an l-long partition, then*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l \tag{1.1.61}$$

with $\lambda_i \in \mathbb{N}$ for $i \in \{1, \ldots, l\}$. It is sometimes useful to extend partitions with finitely many zeros. In this case, we identify all partitions which share the same non-zero portion,

$$\lambda = (\lambda_1, \ldots, \lambda_l) = (\lambda_1, \ldots, \lambda_l, 0, \ldots, 0). \tag{1.1.62}$$

The weight *of a partition λ is written $|\lambda|$ and is the sum of its elements,*

$$|\lambda| = \sum_{i=1}^{l} \lambda_i. \tag{1.1.63}$$

If $|\lambda| = n$ for $n \in \mathbb{N}$ then we say λ partitions *n and write $\lambda \vdash n$.*

Finally, if $\lambda = (\lambda_i)_{i=1}^l$ has $m_i = m_i(\lambda)$ elements equal to i, then we can write λ in multiplicative

(a) The Young diagram $\lambda = (5, 4, 2, 1, 1)$.

(b) The Young diagram for $\lambda' = (5, 3, 2, 2, 1)$, the conjugate partition to $\lambda = (5, 4, 2, 1, 1)$.

Figure 1.5: Examples of Young diagrams and conjugate partitions.

notation,

$$\lambda = \langle 1^{m_1} 2^{m_2} \cdots \lambda_1^{m_{\lambda_1}} \rangle = (\underbrace{\lambda_1, \ldots, \lambda_1}_{m_{\lambda_1}}, \ldots, \underbrace{2, \ldots, 2}_{m_2}, \underbrace{1, \ldots, 1}_{m_1}). \tag{1.1.64}$$

Given a partition λ, one can pictorially represent λ using a Young diagram (sometimes called a Ferrers diagram).

Definition 1.1.13 (Young diagram)**.** *Let $\lambda = (\lambda_i)_{i=1}^{l(\lambda)}$ be a partition. The Young diagram of λ is a collection of $|\lambda|$ boxes arranged in $l = l(\lambda)$ left-justified rows. The first row has λ_1 boxes, the second has λ_2, and so on until the lth row which has λ_l boxes. Figure 1.5a shows an example of a Young diagram.*

Given a partition λ (or equivalently the Young diagram of λ), one can define the conjugate *of λ, written λ'. We have that $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ is defined by the condition that the Young diagram of λ' is the transpose (reflection in the main diagonal) of the diagram of λ. Notice that this means that $m_i(\lambda') = \lambda_i - \lambda_{i+1}$, $\lambda'_1 = l(\lambda)$, and $\lambda_1 = l(\lambda')$. Figure 1.5b gives an example of the transpose operation on Young diagrams.*

Definition 1.1.14 (Hook–length)**.** *Take a partition λ. For a given cell (i, j) in λ (the top left box corresponds to the co-ordinates $(1, 1)$, the box below it to $(2, 1)$ etc.), the* arm length *of that cell is $a(i, j)$, the number of boxes strictly to the right. Similarly, the* leg length*, $g(i, j)$, is the number of boxes strictly below the cell. The* hook–length *of a cell $h(i, j) \coloneqq a(i, j) + g(i, j) + 1$ (i.e. the boxes to the right of the cell, those below, and including the cell itself). Thus, for the top left cell $(1, 1)$, the arm length is $\lambda_1 - 1$, the leg length is $l(\lambda) - 1$ and the hook–length is $\lambda_1 + l(\lambda) - 1$.*

The hook–length *for the partition λ is then defined as*

$$h_\lambda \coloneqq \prod_{(i,j) \in \lambda} h(i, j). \tag{1.1.65}$$

Definition 1.1.15 (Young tableau)**.** *Given a totally ordered finite alphabet of symbols and a partition λ, a* Young tableau *is a Young diagram of shape λ with each cell of the diagram filled with a symbol from the alphabet.*

Henceforth, we will always use the alphabet of n numbers $\{1, \ldots, n\}$. Figure 1.6a shows an example of a Young tableau with entries from $\{1, \ldots, n\}$.

Definition 1.1.16 ((Semi)standard Young tableau)**.** *Given λ, a* standard Young tableau *(SYT) is a Young tableau of shape λ with the additional requirement that entries must strictly increase down columns and across rows. This means that the alphabet must have at least* $\max\{\lambda_1, l(\lambda)\}$ *symbols.*

A semistandard Young tableau *(SSYT) has instead that the rows may* weakly increase*, whilst the columns must still strictly increase (and hence at least $l(\lambda)$ symbols are required). Thus, a standard Young tableau is a semistandard Young tableau. Figures 1.6b and 1.6c demonstrate the standard and semistandard properties.*

4	1	7	3	3	8
2	5	1	10		
5	8				
9	2				

(a) A Young tableau for $\lambda = (6,4,2,2)$ with entries in $\{1,\ldots,10\}$.

1	2	3	5	7	8
3	5	6	8		
4	7				
6	8				

(b) A standard Young tableau for $\lambda = (6,4,2,2)$ with entries in $\{1,\ldots,8\}$.

1	1	2	3	3	5
2	3	4	5		
3	4				
5	5				

(c) A semistandard Young tableau for $\lambda = (6,4,2,2)$ with entries in $\{1,\ldots,5\}$.

Figure 1.6: Examples of Young tableaux, including those with the standard and semistandard restrictions.

The set of all standard Young tableaux of shape λ *and entries in* $\{1,\ldots,n\}$ *is denoted* $\mathrm{SYT}_n(\lambda)$*, and similarly the set of all such semistandard Young tableaux is* $\mathrm{SSYT}_n(\lambda)$*.*

For the subsequent definition, we introduce the following multivariate encoding of a given Young tableau. Take a tableau T with entries in $\{1,\ldots,n\}$. We define the *type* of T to be

$$t = t(T) \coloneqq (t_1,\ldots,t_n) \tag{1.1.66}$$

where t_j is the number of times j appears in T. For example, the type of the tableau T appearing in figure 1.6c is $t = (2,2,4,2,4)$. For a tableau T with type t, the multivariate notation

$$x^T = x_1^{t_1}\cdots x_n^{t_n} \tag{1.1.67}$$

can also be used to encode the filling.

We are now ready to define the central symmetric function for our purposes.

Definition 1.1.17 (Schur function)**.** *Let* λ *be a partition and take* $n \in \mathbb{N}$*, with* $n \geq l(\lambda)$*. Then the* Schur function *(or Schur polynomial) is*

$$\begin{aligned} s_\lambda(x_1,\ldots,x_n) &\coloneqq \sum_{T\in\mathrm{SSYT}_n(\lambda)} x^T && (1.1.68)\\ &= \sum_{T\in\mathrm{SSYT}_n(\lambda)} x_1^{t_1}\cdots x_n^{t_n}. && (1.1.69)\end{aligned}$$

Note that if $n > l(\lambda)$ *then we extend* λ *with zeros until it has length* n*.*

Before giving an example of constructing a Schur function for a given λ and n, we give an alternative (and equivalent) definition. One can also express a Schur function as the following quotient of determinants,

$$s_\lambda(x_1,\ldots,x_n) = \frac{\det\begin{pmatrix} x_1^{\lambda_1+n-1} & x_2^{\lambda_1+n-1} & \cdots & x_n^{\lambda_1+n-1}\\ x_1^{\lambda_2+n-2} & x_2^{\lambda_2+n-2} & \cdots & x_n^{\lambda_2+n-2}\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{\lambda_n} & x_2^{\lambda_n} & \cdots & x_n^{\lambda_n}\end{pmatrix}}{\Delta(x_1,\ldots,x_n)}, \tag{1.1.70}$$

where $\Delta(x_1,\ldots,x_n)$ is the usual Vandermonde determinant, defined in (1.1.6), and λ has been implicitly extended with zeros if necessary. It is from (1.1.70) that one can most easily see that $s_\lambda(x_1,\ldots,x_n)$ is a symmetric polynomial: both numerator and denominator in (1.1.70) are clearly alternating polynomials[6] in $x_1,\ldots,x_n$. The fact that the numerator is alternating also immediately implies that it is

[6] A function $g(x_1,\ldots,x_n)$ is alternating if $g(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = -g(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n)$ for any $i,j \in \{1,\ldots,n\}$, $i \neq j$.

divisible by $\Delta(x_1, \ldots, x_n)$: the numerator is clearly zero whenever two variables x_i, x_j both equal some x, and the denominator shares the same property. This justifies why $s_\lambda(x_1, \ldots, x_n)$ is a polynomial (which is immediate from the definition 1.1.17).

Example 1.1.18. *Let $\lambda = (2, 1)$ and $n = 3$. Then to construct $s_\lambda(x_1, x_2, x_3)$, we first identify all the semistandard Young tableaux of shape λ with entries*[7] *in $\{1, 2, 3\}$. All such possibilities are given below.*

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}$$

Each tableau corresponds to a summand in definition 1.1.17. Hence,

$$s_{(2,1)}(x_1, x_2, x_3) = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2 + 2x_1 x_2 x_3. \tag{1.1.71}$$

We also verify that (1.1.70) *yields the same polynomial.*

$$s_{(2,1)}(x_1, x_2, x_3) = \frac{\det\begin{pmatrix} x_1^4 & x_2^4 & x_3^4 \\ x_1^2 & x_2^2 & x_3^2 \\ 1 & 1 & 1 \end{pmatrix}}{(x_2 - x_3)(x_1 - x_3)(x_1 - x_2)} \tag{1.1.72}$$

$$= \frac{\det\begin{pmatrix} x_1^4 & (x_2 - x_1)(x_2^3 + x_1 x_2^2 + x_1^2 x_2 + x_1^3) & (x_3 - x_2)(x_3^3 + x_2 x_3^2 + x_2^2 x_3 + x_2^3) \\ x_1^2 & (x_2 - x_1)(x_2 + x_1) & (x_3 - x_2)(x_3 + x_2) \\ 1 & 0 & 0 \end{pmatrix}}{(x_2 - x_3)(x_1 - x_3)(x_1 - x_2)} \tag{1.1.73}$$

$$= \frac{\det\begin{pmatrix} x_2^3 + x_1 x_2^2 + x_1^2 x_2 + x_1^3 & x_3^3 + x_2 x_3^2 + x_2^2 x_3 + x_2^3 \\ x_2 + x_1 & x_3 + x_2 \end{pmatrix}}{x_3 - x_1} \tag{1.1.74}$$

$$= \frac{\det\begin{pmatrix} x_2^3 + x_1 x_2^2 + x_1^2 x_2 + x_1^3 & (x_3 - x_1)(x_1^2 + x_2^2 + x_3^2 + x_1 x_3 + x_1 x_2 + x_2 x_3) \\ x_2 + x_1 & x_3 - x_1 \end{pmatrix}}{x_1 - x_3} \tag{1.1.75}$$

$$= (x_2 + x_1)(x_1^2 + x_2^2 + x_3^2 + x_1 x_3 + x_1 x_2 + x_2 x_3) - (x_2^3 + x_1 x_2^2 + x_1^2 x_2 + x_1^3) \tag{1.1.76}$$

$$= x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_2^2 x_3 + x_2 x_3^2 + x_1 x_3^2 + 2x_1 x_2 x_3, \tag{1.1.77}$$

as expected.

1.2 Overview

The remainder of this thesis will heavily draw on the ideas and tools developed within this chapter. In chapter 2, we discuss various pertinent parts of the literature, in order to provide context for our original research. In particular, we will focus on various random matrix moments, and their connection to number theory. We also draw our attention to a series of conjectures due to Fyodorov and Keating [82], one of which we address within this thesis.

In chapter 3, we prove a theorem concerning 'moments of moments' of characteristic polynomials of unitary matrices. These moments of moments (c.f. definition 2.2.8) are an average both over the

[7] The fact here that $|\lambda| = n = 3$ is purely coincidental, this is not a requirement for a Schur function.

matrix group and over the unit circle (where the characteristic polynomial naturally acts). This result resolves a conjecture of Fyodorov and Keating.

We develop the ideas of chapter 3 in chapter 4 to investigate moments of moments of symplectic and orthogonal characteristic polynomials. The proof techniques we employ here differ from those of chapter 3; we focus much more on a combinatorial approach and extensively use the theory introduced in section 1.1.6.

A consequence of our results of chapters 3 and 4 is that the moments of moments are polynomials in the matrix size. In chapter 5, we present and discuss computer code that we have written to calculate the full polynomials for small moments parameters. We also investigate the highly structured nature of these functions.

Building upon the ideas established in section 1.1.4, within chapter 6 we develop a model of the moments of moments. This is a branching model exhibiting logarithmic correlations. We are able not only to recreate the results of chapters 3 and 4 in this 'idealized' setting, but also we see the full phase change predicted by Fyodorov and Keating.

Finally, in chapter 7, we handle more general moments of unitary characteristic polynomials. In particular, we prove results for mixed moments (averages involving both the characteristic polynomial and its derivative), as well as for moments of the logarithmic derivative. These have natural number theoretic consequences, which we also describe. Further, our results exhibit a connection to the solution of a particular non-linear second-order differential equation.

Chapters 3, 4, and 7 are all based on published or submitted work with co-authors. The next section, as well as the introductions to the respective chapters, explains the present author's contributions.

1.3 Authorship

Original research can be found within all subsequent chapters, and here we emphasise where such results can be found. Where the results have appeared in papers (either published or submitted), we give the relevant reference. Additional details are given within the introduction to the respective chapters regarding authorship and, for those based on existing papers, how the chapters differ from the respective manuscripts.

(i) In chapter 2, section 2.2.4, we sketch an argument suggesting that the leading order of the maximum of $|P_{Sp(2N)}(A, \theta)|$ or $|P_{SO(2N)}(A, \theta)|$ should match the Fyodorov and Keating conjecture for the unitary case. This (unpublished) work is joint with Prof. Paul Bourgade and was conducted during a research visit to CIMS.

(ii) Theorem 3.1.4 and theorem 3.1.5 are the key results of chapter 3. They are theorems 1.2 and 1.3 of our paper with J. P. Keating [15], published in *Communications in Mathematical Physics*.

(iii) In chapter 4, we prove theorem 4.1.1 and theorem 4.1.2. These are theorems 1.1 and 1.2 in our paper with T. Assiotis and J. P. Keating [9], *submitted*.

(iv) The polynomials given in chapter 5 originally featured as examples within [9, 15]. The computer code, given within the chapter, was prompted by conversations with Dr. Chris Hughes. Additional examples and plots have been calculated for this document.

(v) We rephrase the theorems of chapters 3 and 4 in terms of a branching random walk model within chapter 6. All results presented within this chapter are original to this thesis. The project was conducted under the supervision of Prof. Jon Keating, and the work involving symplectic and orthogonal models was furthered through discussions with Prof. Paul Bourgade.

(vi) Finally, in chapter 7 we prove theorems 7.1.1 and 7.1.2 which appear as theorems 1.1 and 1.2 in [14], published in *Journal of Mathematical Physics*. This project originated at an American Institute of Mathematics workshop, and hence is highly collaborative. Further particulars of the present author's contribution can be found in the introduction to this chapter.

Chapter 2

Literature review

The chapter is largely dedicated to a review of aspects of the literature that are relevant to the forthcoming chapters. In section 2.1, we cover the general theory of, and the key results for, moments of characteristic polynomials. We also emphasise the various connections throughout to number theory. In particular, our results in chapters 3 and 4 generalize the moments which feature sections 2.1.1, 2.1.2, and 2.1.3. Furthermore, section 2.1.4 relates to the results presented in chapter 7.

In section 2.2, we review various conjectures of Fyodorov and Keating. These are connected to both the results in section 2.1, and also to our original work, see in particular chapter 3. We also highlight some of the recent progress towards proving the Fyodorov and Keating conjectures, in order to provide context for our research. Within section 2.2.4, we argue that the Fyodorov and Keating conjecture for the maximum of unitary characteristic polynomials should, at least to leading order, also hold for symplectic and orthogonal characteristic polynomials. We outline calculations in support of this assertion. The research for this section was conducted as part of a research visit to the Courant Institute for Mathematical Research (CIMS) with Prof. Paul Bourgade.

2.1 Moments of characteristic polynomials

The study of moments of characteristic polynomials of matrices from the compact groups has been a central theme of random matrix theory in recent years. In this part of the review, we present and discuss some of the landmarks within the theory. The focus for section 2.1.1 will be the calculation of the moments and value distribution of unitary characteristic polynomials. This is work of Keating and Snaith [115, 116]. In section 2.1.2, number theoretic moments come to the fore, in particular those of the Riemann zeta function. The conjectural connection between the two seemingly disparate fields of random matrix theory and number theory is explored. At the close of this section, we move to more general moments, including mixed moments and moments of log-derivatives of characteristic polynomials. We also discuss a generalization of the moments of the Riemann zeta function considered in section 2.1.2 to moments of more general L-functions.

2.1.1 Unitary moments

The central function of this thesis is the characteristic polynomial of a matrix A from one of the compact groups discussed in the introduction. Namely we consider matrices $A \in G(N)$ for the group $G(N) \in \{U(N), Sp(2N), SO(2N)\}$. Recall that the characteristic polynomial of A is written

$$P_{G(N)}(A, \theta) = \det(I - Ae^{-i\theta}) \tag{2.1.1}$$

where we emphasise that $P_{G(N)}$ naturally acts on the unit circle, so $\theta \in [0, 2\pi)$.

The canonical example will always be $A \in U(N)$, and therefore temporarily we focus on this unitary case. We henceforth assume that $P_N \equiv P_{U(N)}$ is the characteristic polynomial of such a unitary matrix. For a fixed point $\theta \in [0, 2\pi)$, the 2βth moments of P_N are

$$M_N(\beta) := \int_{U(N)} |P_N(A, \theta)|^{2\beta} dA, \tag{2.1.2}$$

where as usual (see (1.1.8)) dA is the Haar measure on $U(N)$, and to ensure integrability $\mathrm{Re}(2\beta) > -1$. Due to the rotational invariance of the Haar measure on $U(N)$, $M_N(\beta)$ will be independent[1] of θ. By the Weyl integration formula (1.1.4), the 2βth moment is equivalent to

$$M_N(\beta) = \frac{1}{(2\pi)^N N!} \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{j=1}^{N} \left| 1 - e^{i(\theta_j - \theta)} \right|^{2\beta} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_N})|^2 d\theta_1 \cdots d\theta_N. \tag{2.1.3}$$

In [116], Keating and Snaith computed $M_N(\beta)$ for $\mathrm{Re}(2\beta) > -1$. We emphasise that their calculation is exact, i.e. for finite N.

Theorem 2.1.1 (Keating-Snaith [116]). *Let $A \in U(N)$ and $\mathrm{Re}(2\beta) > -1$. Then*

$$M_N(\beta) = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j + 2\beta)}{(\Gamma(j + \beta))^2}, \tag{2.1.4}$$

where $\Gamma(z)$ is the usual extension of the factorial function.

Before discussing some pertinent parts of the proof of theorem 2.1.1, we make a few comments. Firstly, note that (2.1.4) clearly has an analytic continuation in β to the rest of the complex plane. Further, since theorem 2.1.1 gives a finite N formula, we may analyse it as $N \to \infty$. This yields

$$M_N(\beta) \sim c_U(\beta) N^{\beta^2} \tag{2.1.5}$$

where

$$c_U(\beta) = \frac{\mathcal{G}^2(1 + \beta)}{\mathcal{G}(1 + 2\beta)} \tag{2.1.6}$$

and $\mathcal{G}(z)$ is the Barnes $\mathcal{G}$-function (so $\mathcal{G}(z + 1) = \Gamma(z)\mathcal{G}(z)$, and $\mathcal{G}(1) = 1$). When β is an integer, the statement of theorem 2.1.1 reads

$$M_N(\beta) = \prod_{0 \le i < j \le \beta - 1} \left(\frac{N}{i + j + 1} + 1 \right) \sim c_U(\beta) N^{\beta^2}, \tag{2.1.7}$$

where the leading order coefficient simplifies to

$$c_U(\beta) := \prod_{j=0}^{\beta - 1} \frac{j!}{(j + \beta)!}. \tag{2.1.8}$$

By (2.1.7), we also learn that for $\beta \in \mathbb{N}$, $M_N(\beta)$ is a polynomial in N of degree β^2.

The key tool in the proof of theorem 2.1.1 is the celebrated Selberg integral, see for example chapter 17 of [124]. Such an integral has many equivalent forms (and extensions); the one we give here is most relevant for the proof of theorem 2.1.1.

[1] Such a statement is no longer true of averages over $Sp(2N)$ or $SO(2N)$, see for example figure 1.1.

Theorem 2.1.2 (Selberg's Integral [124]). *Let $n \in \mathbb{N}$ and let as usual $\Delta(x_1, \ldots, x_n)$ denote the Vandermonde determinant (taking $\Delta(x_1) = 1$ for $n = 1$). Take $a, b, \alpha, \beta, \gamma \in \mathbb{C}$ with $\mathrm{Re}(a), \mathrm{Re}(b), \mathrm{Re}(\alpha), \mathrm{Re}(\beta)$ all strictly positive, $\mathrm{Re}(\alpha + \beta) > 1$, and*

$$-\frac{1}{n} < \mathrm{Re}(\gamma) < \min\left(\frac{\mathrm{Re}(\alpha)}{n-1}, \frac{\mathrm{Re}(\beta)}{n-1}, \frac{\mathrm{Re}(\alpha+\beta+1)}{2(n-1)}\right). \tag{2.1.9}$$

Then

$$J(a, b, \alpha, \beta, \gamma, n) := \int_{-\infty}^{\infty} \cdots \int_{\infty}^{\infty} |\Delta(x_1, \ldots, x_n)|^{2\gamma} \prod_{j=1}^{n} (a + ix_j)^{-\alpha} (b - ix_j)^{-\beta} dx_1 \cdots dx_n \tag{2.1.10}$$

$$= \frac{(2\pi)^n}{(a+b)^{(\alpha+\beta)n - \gamma n(n-1) - n}} \prod_{j=0}^{n-1} \frac{\Gamma(1 + \gamma + j\gamma)\Gamma(\alpha + \beta - (n-1+j)\gamma - 1)}{\Gamma(1+\gamma)\Gamma(\alpha - j\gamma)\Gamma(\beta - j\gamma)}. \tag{2.1.11}$$

To prove theorem 2.1.1, one coerces (2.1.3) in to a form amenable to an application of Selberg's formula (2.1.11). This is achieved by writing the integrand of (2.1.3) in terms of absolute values of sine-functions, setting $\theta = 0$ (which one can do due to the rotational invariance of the Haar measure on $U(N)$), and making a simple change of variables. One then sees that

$$M_N(\beta) = \frac{2^{N^2 + 2\beta N}}{(2\pi)^N N!} J(1, 1, N + \beta, N + \beta, 1, N), \tag{2.1.12}$$

delivering the statement of theorem 2.1.1.

The motivation for computing $M_N(\beta)$ was to establish the value distribution of the real and imaginary parts of the logarithm of $P_N(A, \theta)$. Notice that $M_N(\beta)$ exactly is the generating function for the moments of real part[2]. Keating and Snaith show that, as $N \to \infty$, that the real and imaginary parts tend independently to Gaussian random variables.

Theorem 2.1.3 (Keating-Snaith Central Limit Theorem [116]). *Take any rectangle $B \subset \mathbb{C}$. Then for fixed $\theta \in [0, 2\pi)$,*

$$\lim_{N\to\infty} \mathrm{meas} \left\{ A \in U(N) : \frac{\log P_N(A, \theta)}{\sqrt{\frac{1}{2}\log N}} \in B \right\} = \frac{1}{2\pi} \int\int_B e^{-\frac{1}{2}(x^2+y^2)} dx dy, \tag{2.1.13}$$

where the measure of the set is taken to be the usual Haar measure on $U(N)$.

Hence, on average $\log |P_N(A, \theta)| \sim \sqrt{(1/2)\log N}$. It is interesting to observe the similarities between (2.1.13) and the Selberg central limit theorem 1.1.6. Indeed, both the Gaussian nature of $\log |P_N(A, \theta)|$ and the scaling in theorem 2.1.3 will be important over the course of the subsequent section.

2.1.2 Random matrix theory and number theory

The origins of the connection between random matrix theory and number theory can be traced back to a conversation in 1971 between Hugh Montgomery and Freeman Dyson, introduced at tea at the Institute for Advanced Study by Sarvadaman Chowla. The conversation turned to the study of pair correlations of eigenvalues and of Riemann zeta zeros, which turn out to have an identical formulation.

We begin with Dyson's theorem, and a sketch of its derivation, (see Dyson [70] or Mehta [124]) on the two-point correlation function of eigenphases for $A \in U(N)$. Let the eigenvalues of A be $e^{i\theta_1}, \ldots, e^{i\theta_N}$,

[2]Thus, they also consider $\exp(it \,\mathrm{Im} \log P_N(A, \theta))$, the moments generating function for the imaginary part. The techniques used for the imaginary part mirror those for the real part.

and rescale the eigenphases θ_j so that on average they have unit spacing,

$$\phi_j := \frac{\theta_j N}{2\pi}. \tag{2.1.14}$$

The two-point correlation function for A is then

$$R_2(A,x) := \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{N}\sum_{k=-\infty}^{\infty}\delta(x+kN-(\phi_n-\phi_m)). \tag{2.1.15}$$

Dyson established the following result.

Theorem 2.1.4 (Dyson's Pair Correlation [70])**.** *Take $A \in U(N)$ and let $\phi_1,\ldots,\phi_N$ be the normalized eigenphases of A, see* (2.1.14)*. Then for test functions f such that $f(x) \to 0$ as $|x| \to \infty$,*

$$\lim_{N\to\infty}\int_{U(N)}\int_{-\infty}^{\infty} f(x)R_2(A,x)\,dx\,dA = \int_{-\infty}^{\infty} f(x)\left(\delta(x)+1-\left(\frac{\sin(\pi x)}{\pi x}\right)^2\right)dx. \tag{2.1.16}$$

Hence, taking the test function to be $f(x)=1$ for $x\in[\alpha,\beta]$ and 0 otherwise one finds

$$\lim_{N\to\infty}\int_{U(N)}\frac{1}{N}\left|\{\phi_n,\phi_m : \alpha\le\phi_n-\phi_m\le\beta\}\right|\,dA = \int_{\alpha}^{\beta}\left(\delta(x)+1-\left(\frac{\sin(\pi x)}{\pi x}\right)^2\right)dx. \tag{2.1.17}$$

We now sketch Dyson's proof. We first notice that $R_2(A,x)$ is periodic in x, so its Fourier series can be calculated to be

$$R_2(A,x) = \frac{1}{N^2}\sum_{k=-\infty}^{\infty}|\operatorname{Tr}A^k|^2 e^{\frac{2\pi i k x}{N}}. \tag{2.1.18}$$

The following additional result of Dyson will be useful.

Theorem 2.1.5 (Dyson [70])**.** *Let $k\in\mathbb{Z}$. Then,*

$$\int_{U(N)}|\operatorname{Tr}A^k|^2 dA = \begin{cases} N^2 & \text{if } k=0\\ |k| & \text{if } |k|\le N\\ N & \text{if } |k|>N.\end{cases} \tag{2.1.19}$$

Hence

$$\int_{U(N)}R_2(A,x)dA = \frac{1}{N^2}\sum_{k=-\infty}^{\infty}e^{\frac{2\pi i k x}{N}}\times\begin{cases} N^2 & \text{if } k=0\\ |k| & \text{if } |k|\le N\\ N & \text{if } |k|>N\end{cases} \tag{2.1.20}$$

$$= \sum_{k=-\infty}^{\infty}\delta(x-kN)+1-\left(\frac{\sin(\pi x)}{\pi x}\right)^2, \tag{2.1.21}$$

which gives the result.

The equivalent calculation for the non-trivial zeros of the Riemann zeta function was done by Montgomery [126], though completely independently of the work of Dyson. Recall from chapter 1, section 1.1.2, that the non-trivial zeros of the Riemann zeta function $\zeta(s)$ are conjecturally of the form $s=\sigma+it$ for $\sigma=1/2$ and $t\in\mathbb{R}$, and provably $\sigma\in(0,1)$, $t\in\mathbb{R}$.

Now denote the non-trivial zeros of $\zeta(s)$ by $\rho_n = 1/2+it_n$, with $\mathrm{Re}(t_n)>0$ and where the ordering on the zeros is by height[3]. Define

$$N(T) := |\{n : 0 \leq \mathrm{Re}(t_n) \leq T\}| \tag{2.1.22}$$

to be the number of non-trivial zeros up to height T. Then it can be shown, see for example [154], that

$$N(T) \sim \frac{T}{2\pi} \log \frac{T}{2\pi e}, \tag{2.1.23}$$

as $T \to \infty$. This proves that there are infinitely many non-trivial zeros. Further, the mean density increases logarithmically with height T.

Mimicking the calculation for the pair correlation of eigenvalues, we rescale the zeros so that they have unit mean spacing,

$$w_n := \frac{t_n}{2\pi} \log \frac{t_n}{2\pi}. \tag{2.1.24}$$

Montgomery's pair correlation conjecture is that,

$$\lim_{T\to\infty} \frac{1}{T} |\{w_n, w_m \in [0, T] : \alpha \leq w_n - w_m \leq \beta\}| = \int_\alpha^\beta \left(\delta(x) + 1 - \left(\frac{\sin(\pi x)}{\pi x} \right)^2 \right) dx. \tag{2.1.25}$$

Such a conjecture is motivated by a theorem of Montgomery [126]. This result can be stated as

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n,m\leq N} f(w_n - w_m) = \int_{-\infty}^{\infty} f(x) \left(\delta(x) + 1 - \left(\frac{\sin(\pi x)}{\pi x} \right)^2 \right) dx. \tag{2.1.26}$$

where f is a test function with Fourier transform supported on $(-1, 1)$ (and vanishes elsewhere), and such that the left and right sides of (2.1.26) converge. The pair correlation conjecture corresponds to choosing f as the indicator function on $[\alpha, \beta)$ (whose Fourier transform does not vanish outside $(-1, 1)$).

The similarity between (2.1.25) and (2.1.17) is clearly apparent. Further numerical evidence for the conjecture has been supplied by the work of Odlyzko [129], and heuristics and computations have been completed on general k-point correlation functions [23, 24, 97, 141]. A visual comparison can be found in figure 2.1. There we have compared plots of 50 points taken from a uniform distribution on the unit circle, with the eigenvalues of a random matrix drawn from $U(50)$, and 50 consecutive[4] non-trivial zeros of $\zeta(s)$, scaled to wrap around the unit circle. The figure shows that the zeta zeros display similar 'repulsion' as the eigenvalues are known to show (see (1.1.4)), and also that their distribution seems far from uniform.

Further evidence for a connection can be found by comparing theorem 2.1.3 with Selberg's central limit theorem 1.1.6 for the Riemann zeta function. For convenience, we reproduce both statements here. For any rectangle $B \in \mathbb{C}$,

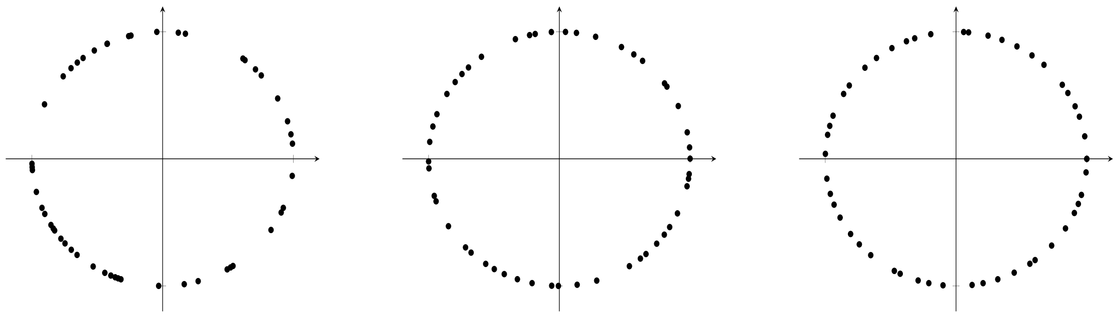
$$\lim_{T\to\infty} \frac{1}{T} \left| \left\{ t \in [T, 2T] : \frac{\log \zeta(\frac{1}{2} + it)}{\sqrt{\frac{1}{2} \log \log \frac{T}{2\pi}}} \in B \right\} \right| = \frac{1}{2\pi} \int\int_B e^{-\frac{1}{2}(x^2+y^2)} dx dy, \tag{2.1.27}$$

and

$$\lim_{N\to\infty} \mathrm{meas} \left\{ A \in U(N) : \frac{\log P_N(A, \theta)}{\sqrt{\frac{1}{2} \log N}} \in B \right\} = \frac{1}{2\pi} \int\int_B e^{-\frac{1}{2}(x^2+y^2)} dx dy. \tag{2.1.28}$$

[3] Recall from section 1.1.2 that if ρ_n is a non-trivial zeros of $\zeta(s)$, then so is $\overline{\rho_n}$. Thus, it suffices to only consider zeros with positive imaginary part.

[4] We do not use the first 50 zeros since it takes some time for the zeros to 'warm up' and demonstrate the pairwise repulsion.



(a) 50 points drawn from the uniform distribution on the unit circle.

(b) Eigenvalues of a random unitary 50×50 matrix.

(c) 50 consecutive non-trivial zeros of $\zeta(s)$, from 201^{st} to 250^{th}, data from [153], scaled to lie on the unit circle.

Figure 2.1: Comparing 50 points distributed uniformly on the unit circle (2.1a), with 50 eigenvalues of a random unitary matrix (2.1b), and with 50 scaled consecutive non-trivial zeros of $\zeta(s)$ (2.1c).

In both cases, the real and imaginary parts of the respective logarithms tend independently to Gaussian random variables. Additionally, if one sets

$$N = \log \tfrac{T}{2\pi} \tag{2.1.29}$$

in the scaling for (2.1.27), then it matches the scaling in (2.1.28). The same identification in the unit mean scaling (2.1.24) means that the average density of zeros matches the average density of eigenvalues in (2.1.14).

A natural question, therefore, is if one can exploit this apparent connection? Can one model, in a statistical sense, the Riemann zeta function by unitary characteristic polynomials? It turns out that one can do so, very successfully. To this end, we focus on a long-standing number theoretic conjecture.

Conjecture 2.1.6. *It is widely believed that*

$$\mathscr{M}_T(\beta) \coloneqq \lim_{T\to\infty} \frac{1}{T} \int_0^T |\zeta(\tfrac{1}{2} + it)|^{2\beta} dt \tag{2.1.30}$$

$$= c_\zeta(\beta) a_\zeta(\beta) \left(\log \tfrac{T}{2\pi}\right)^{\beta^2}, \tag{2.1.31}$$

with

$$a_\zeta(\beta) \coloneqq \prod_p \left[\left(1 - \frac{1}{p}\right)^{\beta^2} \left(\sum_{m=0}^{\infty} \left(\frac{\Gamma(\beta+m)}{m!\Gamma(\beta)} \right)^2 p^{-m} \right) \right], \tag{2.1.32}$$

where the product is over primes p, and $c_\zeta(\beta)$ is some function depending on the moment parameter β.

The only known values of $\mathscr{M}_T$ are $\beta = 0, 1, 2$. The case of $\beta = 0$ is trivial, $\beta = 1$ was computed by Hardy and Littlewood [91], and Ingham [102] determined $\beta = 2$. In each case, conjecture 2.1.6 is satisfied and additionally one learns that

$$c_\zeta(0) = 1 \tag{2.1.33}$$

$$c_\zeta(1) = 1 \tag{2.1.34}$$

$$c_\zeta(2) = \frac{2}{4!} \tag{2.1.35}$$

Using number theoretic arguments, Conrey and Ghosh [52] and Conrey and Gonek [53] conjectured the precise forms of the 6^{th} and 8^{th} moments, giving,

$$c_\zeta(3) = \frac{42}{9!} \tag{2.1.36}$$
$$c_\zeta(4) = \frac{24024}{16!}. \tag{2.1.37}$$

Hence it seems that $c_\zeta(\beta) \cdot (\beta^2)!$ is integral for $\beta \in \mathbb{N}$. Additionally, Ramachandra [136] and Heath-Brown [96] have established a lower bound $\mathscr{M}_T(\beta) \gg (\log \frac{T}{2\pi})^{\beta^2}$ for positive, rational β. Upper bounds of the correct size are known, conditional on the Riemann hypothesis, due to arguments of Soundararajan [146] and Harper [93].

Given the above discussion, we are led to compare $\zeta(1/2 + it)$ with $P_N(A, \theta)$ for $A \in U(N)$, and identify $N = \log \frac{T}{2\pi}$. Hence, $\mathscr{M}_T(\beta)$ should be modelled by $M_N(\beta)$, where recall

$$M_N(\beta) = \int_{U(N)} |P_N(A, \theta)|^{2\beta} dA \sim \frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)} N^{\beta^2}. \tag{2.1.38}$$

This is precisely the conclusion of Keating and Snaith [116] who used their moment computation, theorem 2.1.1, to conjecture that

$$c_\zeta(\beta) = \frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}. \tag{2.1.39}$$

Observe that (2.1.39) matches with the all known and conjectural cases (2.1.33)–(2.1.37).

Most recently, a series of papers due to Conrey and Keating [37–41] have unified the number theoretic and random matrix methods. As commented above, Keating and Snaith's work implies, for integer β, that the 2βth moment of $\zeta(1/2+it)$ is of the order of $(\log \frac{T}{2\pi})^{\beta^2}$. Conrey et al. [46] extended this conjecture to

$$\frac{1}{T} \int_0^T |\zeta(\tfrac{1}{2} + it)|^{2\beta} dt = Q_\beta(\log \tfrac{T}{2\pi}) + O\left(T^{-\frac{1}{2}+\varepsilon}\right) \tag{2.1.40}$$

for integer β, where $Q_\beta(x)$ is a polynomial in x of degree β^2 which has a multiple contour integral representation. Such complex analytic integrals prove essential for our analysis in chapter 3. The conjectures of Conrey et at. extend too to other L-functions.

The classical number theoretic approach to understanding moments of the Riemann zeta function uses Dirichlet polynomial approximations of $\zeta(s)$ and higher powers. Analysis in this direction, as discussed previously, gives the second and fourth moments and strong conjectural forms for the sixth and the eight moments. However, this method fails for $\beta \geq 5$ since it predicts *negative values.*

Over the course of five successive works, Conrey and Keating demonstrate why the Dirichlet polynomial method in its traditional form fails: for higher β it is essential that one uses *much longer* Dirichlet polynomials than is conventional to avoid missing crucial terms. Whilst their work is still conjectural for higher moments, it provides convincing evidence (in particular, the moments are no longer negative) for proceeding in such a fashion.

The recipe that they derive is based on divisor sums and can be viewed as a type of 'multi-dimensional Hardy-Littlewood circle method', or 'Manin-type stratification'. Consider

$$\int_0^\infty \left(\prod_{\alpha \in A} \zeta(\tfrac{1}{2} + it + \alpha)\right) \left(\prod_{\beta \in B} \zeta(\tfrac{1}{2} - it + \beta)\right) \psi\left(\frac{t}{T}\right) dt, \tag{2.1.41}$$

where A, B are sets of size β of 'shifts' and ψ is a smooth function of compact support. Note that by letting all $\alpha, \beta \to 0$ we recover the desired 2βth moment. Conrey and Keating demonstrate that, conjecturally, (2.1.41) can be evaluated either by multiple contour integrals of the type appearing in

(2.1.40), or by examining the Dirichlet series

$$\prod_{\alpha \in A} \zeta(\tfrac{1}{2} + it + \alpha) = \sum_{n=1}^{\infty} \frac{\tau_A(n)}{n^s}, \tag{2.1.42}$$

where the arithmetic divisor function $\tau_A(n)$ is exactly defined by the Euler product expansion of the left hand side of (2.1.42). Their conclusion, which for the first time unites the random matrix and number theoretic approaches, is that the result is the same regardless of the method chosen.

2.1.3 Symplectic and orthogonal moments and families of L-functions

The discussion of section 2.1.2, comparing the Riemann zeta function with unitary characteristic polynomials, has been significantly extended by Katz and Sarnak [109, 110] and Keating and Snaith [115]. This extension connects more general L-functions with the other compact matrix groups. Rather than averaging along the critical line for a fixed function (e.g. $\zeta(s)$), the conjecture is that if one looks at a fixed symmetry point (e.g. $s = 1/2$) and averages over a 'family' of number theoretic functions, then instead of unitary statistics one should comparatively recover statistics for the other compact matrix groups. Within this section, we make such notions clearer by means of various examples.

Most generally, an L-function is a Dirichlet series with an Euler product and a functional equation (recall (1.1.11), (1.1.12), and (1.1.13)). However, in order to keep this review self-contained we focus on just three canonical examples. For a comprehensive account, we refer to [103].

Following the work of Katz and Sarnak [109, 110] and Keating and Snaith [116], we use the concept of a 'family of L-functions'. These families are split by 'symmetry type'. This indicates which random matrix ensemble (unitary, symplectic, or orthogonal) corresponds to the family. Additional information on families and examples can be found in the introductions of [46] and [44].

Let $G(N)$ be either $U(N), Sp(2N)$, or $SO(2N)$. A generalisation of the unitary moments given in (2.1.2) is the following.

$$M_{G(N)}(\beta, \theta) := \int_{G(N)} |P_{G(N)}(A, \theta)|^{2\beta} dA, \tag{2.1.43}$$

where the average is over the relevant Haar measure for the chosen matrix group. If $G(N) = Sp(2N)$ or $G(N) = SO(2N)$ then the measure is no longer invariant under rotations (hence why we now emphasise θ in the notation). Since the eigenvalues come in complex conjugate pairs, one often fixes $\theta = 0$ as the symmetry point (just as $s = 1/2$ is a symmetry point of the zeros $\zeta(s)$) and computes with $M_{G(N)}(\beta, 0)$.

We first observe that for $A \in Sp(2N)$ with characteristic polynomial $P_{Sp(2N)}(A, \theta)$, then

$$P_{Sp(2N)}(A, 0) = \prod_{j=1}^{N} \left(1 - e^{i\theta_j}\right)\left(1 - e^{-i\theta_j}\right) = 2^{2N} \prod_{j=1}^{N} \sin^2\left(\tfrac{\theta_j}{2}\right). \tag{2.1.44}$$

This shows that the value of the characteristic polynomial at the symmetry point is always real and positive (similarly for $A \in SO(2N)$), so trivially $|P_{Sp(2N)}(A, 0)| = P_{Sp(2N)}(A, 0)$. However, for consistency with the unitary case we write $|P_{Sp(2N)}(A, 0)|$.

Theorem 2.1.7 (Keating and Snaith [115])**.** *Let $A \in Sp(2N)$ with characteristic polynomial $P_{Sp(2N)}(A, \theta)$. Then*

$$\int_{Sp(2N)} |P_{Sp(2N)}(A, 0)|^{2\beta} dA = 2^{\beta(2\beta+1)} \left(\prod_{j=1}^{2\beta} \frac{j!}{(2j)!} \right) \prod_{j=1}^{2\beta} (N + j) \prod_{1 \le i < j \le 2\beta} \left(N + \frac{i+j}{2} \right) \tag{2.1.45}$$

$$\sim c_{Sp}(\beta) N^{\frac{2\beta(2\beta+1)}{2}}, \tag{2.1.46}$$

as $N \to \infty$, where

$$c_{Sp}(\beta) := 2^{4\beta^2} \frac{\mathcal{G}(1+2\beta)\sqrt{\Gamma(1+2\beta)}}{\sqrt{\mathcal{G}(1+4\beta)\Gamma(1+4\beta)}}. \tag{2.1.47}$$

As usual $\mathcal{G}$ is the Barnes $\mathcal{G}$-function, see (2.1.6)*. Thus for integer β,*

$$\int_{Sp(2N)} |P_{Sp(2N)}(A,0)|^{2\beta} dA \sim c_{Sp}(\beta) N^{\frac{2\beta(2\beta+1)}{2}} \tag{2.1.48}$$

where the leading coefficient specialises to

$$c_{Sp}(\beta) = \frac{1}{\prod_{j=1}^{2\beta}(2j-1)!!} \tag{2.1.49}$$

and $k!! = k(k-2)(k-4)\cdots(k-(2\lceil \frac{k}{2} \rceil - 2))$.

Theorem 2.1.8 (Keating and Snaith [115])**.** *Let $A \in SO(2N)$ with characteristic polynomial $P_{SO(2N)}(A,\theta)$. Then*

$$\int_{SO(2N)} |P_{SO(2N)}(A,0)|^{2\beta} dA = 2^{\frac{2\beta(2\beta+1)}{2}} \left(\prod_{j=1}^{2\beta-1} \frac{j!}{(2j)!} \right) \prod_{1 \le i < j \le 2\beta} \left(N + \frac{i+j}{2} - 1 \right) \tag{2.1.50}$$

$$\sim c_{SO}(\beta) N^{\frac{2\beta(2\beta-1)}{2}}, \tag{2.1.51}$$

as $N \to \infty$, where

$$c_{SO}(\beta) := 2^{2\beta^2} \frac{\mathcal{G}(1+2\beta)\sqrt{\Gamma(1+4\beta)}}{\sqrt{\mathcal{G}(1+4\beta)\Gamma(1+2\beta)}}. \tag{2.1.52}$$

For integer β,

$$\int_{SO(2N)} |P_{SO(2N)}(A,0)|^{2\beta} dA \sim c_{SO}(\beta) N^{\frac{2\beta(2\beta-1)}{2}}, \tag{2.1.53}$$

where the leading coefficient specialises to

$$c_{SO}(\beta) = \frac{2^{2\beta}}{\prod_{j=1}^{2\beta-1}(2j-1)!!}. \tag{2.1.54}$$

From theorems 2.1.7 and 2.1.8, one learns that the behaviour at the symmetry point $\theta = 0$ is *different* between the three ensembles. Both results are proved in a similar way to theorem 2.1.1, i.e. using the Selberg integral.

It will also be useful to highlight the work of Keating and Odgers [112], who computed low moments for the symplectic and orthogonal characteristic polynomials away from the symmetry point.

Theorem 2.1.9 (Keating and Odgers [112])**.** *Let $\theta \in \mathbb{R}$ and $G(N) = SO(2N)$ or $Sp(2N)$. Then*

$$\int_{G(N)} |P_{G(N)}(A,\theta)|^2 dA = \begin{cases} \dfrac{1}{|1-e^{-2i\theta}|^2} \left[(2N+3) - \displaystyle\sum_{j=0}^{2N+3} e^{(N+1-j)2i\theta} \right] & \text{if } G(N) = Sp(2N), \\ (2N+1) + \displaystyle\sum_{j=0}^{2N+1} e^{(N-j)2i\theta} & \text{if } G(N) = SO(2N). \end{cases} \tag{2.1.55}$$

Hence, for fixed $\theta \neq n\pi$, $n \in \mathbb{Z}$, and as $N \to \infty$, their result gives

$$\int_{G(N)} |P_{G(N)}(A,\theta)|^2 dA \sim \begin{cases} \dfrac{2}{|1-e^{-2i\theta}|^2}\, N & \text{if } G(N) = Sp(2N), \\ 2N & \text{if } G(N) = SO(2N). \end{cases} \tag{2.1.56}$$

Notice that the second moment in the symplectic case is still dependent on θ.

Keating and Odgers also computed the asymptotic behaviour of the symplectic and orthogonal moments for fixed $\theta \neq n\pi$.

Theorem 2.1.10 (Keating and Odgers [112])**.** *Fix $\theta \neq n\pi$ and let $G(N) = SO(2N)$ or $Sp(2N)$. Then as $N \to \infty$*

$$\int_{G(N)} |P_{G(N)}(A,\theta)|^{2\beta} dA \sim \begin{cases} \dfrac{c_G(\beta)}{|1-e^{-2i\theta}|^{\beta(\beta+1)}} N^{\beta^2} & \text{if } G(N) = Sp(2N), \\ \dfrac{c_G(\beta)}{|1-e^{-2i\theta}|^{\beta(\beta-1)}} N^{\beta^2} & \text{if } G(N) = SO(2N), \end{cases} \tag{2.1.57}$$

where

$$c_G(\beta) := 2^{\beta^2} \prod_{j=0}^{\beta-1} \frac{j!}{(\beta+j)!}. \tag{2.1.58}$$

Thus, on the scale of mean eigenvalue spacing, there is a transition from orthogonal or symplectic statistics back to the unitary form, as $N \to \infty$. Just as the moments of unitary characteristic polynomials were used to deduce information about moments of the Riemann zeta function, one can now use these results on the moments of symplectic and orthogonal characteristic polynomials to learn about more general L-functions.

By way of example, we first rephrase the results given in section 2.1.2 (comparing $P_{U(N)}(A,\theta)$ with $\zeta(1/2+it)$) in terms of a *family* with *unitary symmetry.* This will serve as a guide for the subsequent two examples of families with different symmetry types.

Family with unitary symmetry

Recall that an L-function is a Dirichlet series that also has an Euler product and a functional equation. Thus, the Riemann zeta function is an L-function, and in fact is the simplest L-function. Additionally, as has been evidenced in section 2.1.2, the statistics of $\zeta(1/2+it)$ mirror those of eigenvalues of unitary matrices. Therefore, we say that $\zeta(1/2+it)$ belongs to a 'unitary' family and we write the family as

$$\{\zeta(\tfrac{1}{2}+it) |\, t \geq 0\}. \tag{2.1.59}$$

This is the archetypal example of a unitary family. To any symmetry family, there is an associated 'height' by which the elements are ordered. For the Riemann zeta function, the family is ordered by the height up the critical line t. Finally, the unitary symmetry is demonstrated, for example, through a comparison between the moments

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2}+it)|^{2\beta} dt \sim c_\zeta(\beta) a(\beta) \left(\log \tfrac{T}{2\pi}\right)^{\beta^2} \tag{2.1.60}$$

and

$$\int_{U(N)} |P_{U(N)}(A,\theta)|^{2\beta} dA \sim c_U(\beta) N^{\beta^2}, \tag{2.1.61}$$

where, as given in section 2.1.2, the first result is conjectural beyond $\beta = 2$, and the latter is theorem 2.1.1, proved by Keating and Snaith.

We now introduce two different types of families of L-functions, and similarly identify the compact random matrix group corresponding to the symmetry type. Just as above, we are able to compare conjectural forms for the number theoretic moments with known results for the random matrix moments.

Family with symplectic symmetry

To begin, we once more revisit the Riemann zeta function. Recall definition 1.1.4,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{2.1.62}$$

convergent to the right of $\mathrm{Re}(s) = 1$. The simplest extension to $\zeta(s)$ is the Dirichlet L-function for the non-trivial character of conductor 3, defined as follows.

Definition 2.1.11. *Define* $\chi_{-3} : \mathbb{N} \to \mathbb{C}$ *by*

$$\chi_{-3}(n) := \begin{cases} 0 & \text{if } n \equiv 0 \mod 3 \\ 1 & \text{if } n \equiv 1 \mod 3 \\ -1 & \text{if } n \equiv -1 \mod 3. \end{cases} \tag{2.1.63}$$

This is a periodic function with period 3. *The Dirichlet L-function corresponding to* χ_{-3} *is*[5]

$$L(s, \chi_{-3}) := \sum_{n=1}^{\infty} \frac{\chi_{-3}(n)}{n^s} = 1 - \frac{1}{2^s} + \frac{1}{4^s} - \frac{1}{5^s} + \cdots \tag{2.1.64}$$

convergent for $\mathrm{Re}(s) > 1$. *Like* $\zeta(s)$ *(see* (1.1.12) *and* (1.1.13)*),* $L(s, \chi_{-3})$ *has an Euler product and a functional equation (see for example [103]).*

Of course, one can generalise definition 2.1.11 to more general moduli than just 3. Take an integer d such that

$$d = \begin{cases} k & \text{if } k \equiv 1 \mod 4, k \text{ square-free, or,} \\ 4m & \text{if } m \equiv 2 \text{ or } 3 \mod 4, \text{ and } m \text{ is square-free.} \end{cases} \tag{2.1.65}$$

If d satisfies either of the above conditions then d is called a *fundamental discriminant*[6]. The first few positive fundamental discriminants are $d = 1, 5, 8, 12, 13, 17, 21, \ldots$, and the first negative values are $d = -3, -4, -7, -8, \ldots$.

Define for such d

$$\chi_d(n) := \left(\frac{d}{n}\right), \tag{2.1.66}$$

where $\left(\frac{d}{n}\right)$ is the Kronecker symbol, the generalisation of the Legendre symbol. Explicitly, for an integer n with prime decomposition

$$n = u \cdot p_1^{e_1} \cdots p_k^{e_k}, \tag{2.1.67}$$

with $u = \pm 1$ and p_j prime, then the Kronecker symbol is

$$\left(\frac{d}{n}\right) := \left(\frac{d}{u}\right) \prod_{j=1}^{k} \left(\frac{d}{p_j}\right)^{e_j}. \tag{2.1.68}$$

[5] The reason for using -3 rather than 3 in the notation will become apparent shortly.

[6] The name is due to the fact that such d are the discriminants of quadratic number fields, with $d = 1$ being the 'degenerate' quadratic field $\mathbb{Q}$.

In the right hand side of (2.1.68) $\left(\frac{a}{p}\right)$ is the Legendre symbol, which take the values for $p \neq 2$,

$$\left(\frac{a}{p}\right) := \begin{cases} 0 & \text{if } a \equiv 0 \mod p, \\ 1 & \text{if } a \equiv m^2 \mod p \text{ and } a \not\equiv 0 \mod p, \\ -1 & \text{if } a \not\equiv m^2 \mod p \text{ and } a \not\equiv 0 \mod p. \end{cases} \tag{2.1.69}$$

and

$$\left(\frac{a}{2}\right) := \begin{cases} 0 & \text{if } a \text{ is even}, \\ 1 & \text{if } a \equiv \pm 1 \mod 8, \\ -1 & \text{if } a \equiv \pm 3 \mod 8. \end{cases} \tag{2.1.70}$$

Finally, $\left(\frac{a}{1}\right) := 1$, $\left(\frac{a}{-1}\right) := -1$ if $a < 0$ and 1 otherwise, and $\left(\frac{a}{0}\right) := 1$ if $a = \pm 1$ and 0 otherwise. Then χ_d is called a *real Dirichlet character*. When $d = 1$, χ_d is the trivial character (taking the value 1 for all n), and for other fundamental discriminants d, χ_d is a real, primitive, *quadratic* Dirichlet character of modulus d. Notice that for $d = -3$ (the first negative fundamental discriminant), $\chi_{-3}(n) = \left(\frac{-3}{n}\right)$ which matches (2.1.63).

Given χ_d, a real, quadratic Dirichlet character modulo a fundamental discriminant d, the associated *Dirichlet L-function* is

$$L(s, \chi_d) := \sum_{n=1}^{\infty} \frac{\chi_d(n)}{n^s}. \tag{2.1.71}$$

Such L-functions again have an Euler product, a functional equation, and a meromorphic continuation to the full complex plane (see [103]). Further, they have an associated Riemann hypothesis which conjectures that the non-trivial zeros of $L(s, \chi_d)$ are also on the critical line $\mathrm{Re}(s) = 1/2$. For a fixed d, the statistics of the non-trivial zeros of $L(s, \chi_d)$ high up the critical line again display *unitary* symmetries and hence form a unitary family,

$$\{L(\tfrac{1}{2} + it, \chi_d) |\ d \text{ a fixed, fundamental discriminant}, t \geq 0\}, \tag{2.1.72}$$

see Rudnick and Sarnak [141]. However, one can instead take the value of $L(s, \chi_d)$ at a fixed symmetry point $s = 1/2$ and average over d. One then should recover *symplectic* symmetries (see for example [44, 115]). To this end, define the family

$$\{L\left(\tfrac{1}{2}, \chi_d\right) \mid d \text{ a fundamental discriminant}, \chi_d(n) = \left(\tfrac{d}{n}\right)\}, \tag{2.1.73}$$

ordered by $|d|$. This is an example of a *symplectic family*. To demonstrate this, we again look at moments. It is conjectured that (see for example [44])

$$\frac{1}{D^*} \sum_{|d| \leq D}^{*} L(\tfrac{1}{2}, \chi_d)^{2\beta} \sim c_{L_D}(\beta) a_{L_D}(\beta) \left(\log D\right)^{\frac{2\beta(2\beta+1)}{2}}, \tag{2.1.74}$$

as $D \to \infty$. The sum is only over fundamental discriminants d; D^* is the length of the sum; and $a_{L_D}(\beta)$ has a similar form to (2.1.32) (see Conrey et al. [46] for example for the full formulation). The conjecture is based on work of Jutila [107] and Soundararajan [145], and so the values of $c_{L_D}(\beta)$ are known for $\beta = 1, 2, 3$ and conjectured for $\beta = 4$.

Recall theorem 2.1.7, the result of Keating and Snaith which gives the asymptotic behaviour of the

2βth moments of symplectic characteristic polynomials at the symmetry point,

$$\int_{Sp(2N)} |P_{Sp(2N)}(A,0)|^{2\beta} dA \sim c_{Sp}(\beta) N^{\frac{2\beta(2\beta+1)}{2}}, \qquad (2.1.75)$$

where $c_{Sp}(\beta)$ is the leading order moment coefficient depending on β, see (2.1.49). As with the comparison between $\zeta(s)$ and unitary polynomials, we here associate matrix size N with the logarithm of the 'height' of the family: $N \sim \log D$. With this dictionary in place, the symplectic form of (2.1.74) is evident. Further, just as in the unitary case, theorem 2.1.7, provides the conjecture that $c_{L_D}(\beta) = c_{Sp}(\beta)$. All known values of $c_{L_D}(\beta)$ indeed satisfy such a relation.

Family with orthogonal symmetry

Finally we give an example of an orthogonal family. In this case, one considers two categories: even and odd[7]. The even families are related to the matrices from $SO(2N)$, and the odd to those from $SO(2N+1)$. However, since the 'odd' L-functions take the value 0 at their symmetry point[8], we only consider 'even' families. In general, the L-functions displaying orthogonal symmetries are more complicated than the Dirichlet L-functions, see (2.1.71). Arguably the simplest example is derived from L-functions attached to elliptic curves.

Consider an elliptic curve E defined over $\mathbb{Q}$,

$$E : y^2 = x^3 + ax + b, \qquad (2.1.76)$$

for $a, b \in \mathbb{Z}$ such that the discriminant $\Delta = -16(4a^3 + 27b^2) \neq 0$ (which ensures that E has distinct roots, or equivalently, is non-singular). If the pair $x, y \in \mathbb{C}$ form a solution to the equation defining the curve E, then we say they lie on E and sometimes write $(x, y) \in E$. Given a prime $p \nmid \Delta$, one can consider the number of points on E modulo p. This leads to the following definition

$$a_p := p + 1 - |\{(x, y) \in E : x, y \in \mathbb{Z}/p\mathbb{Z}\}|. \qquad (2.1.77)$$

These coefficients a_p are used when constructing the L-function for the curve E, which is defined by its Euler product[9]

$$L(s, E) := \prod_p \left(1 - a_p p^{-s} + \mathbb{1}_{p\nmid\Delta} p^{-2s+1}\right)^{-1}, \qquad (2.1.78)$$

where $\mathbb{1}_{p\nmid\Delta}$ is 1 for the 'good primes' not dividing the discriminant, and 0 otherwise. Just as with $\zeta(s)$, from the Euler product one can derive the appropriate Dirichlet series, and in turn the meromorphic continuation and the functional equation [104].

The orthogonal family is built from 'twisting' $L(s, E)$ by the Dirichlet character $\chi_d(n)$, see (2.1.66). If

$$L(s, E) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \qquad (2.1.79)$$

is the Dirichlet series corresponding to (2.1.78), and the sequence $(a_n)_{n\in\mathbb{N}}$ is the appropriate sequence formed by expanding the Euler product, then for d a fundamental discriminant,

$$L_E(s, \chi_d) := \sum_{n=1}^{\infty} \frac{a_n \chi_d(n)}{n^s} \qquad (2.1.80)$$

[7]The parity comes from the sign of the functional equation for the L-function, see for example [46].

[8]This is a consequence of their functional equation, see for example [46].

[9]This is formulation is the algebraic convention, though it will mean that the symmetry point is at $s = 1$, rather than $s = 1/2$. However, one can simply renormalize each local factor to shift the critical line for $L(s, E)$ to the traditional $\mathrm{Re}(s) = 1/2$, see [104].

is the twisted L-function for E. Hence, fix E an elliptic curve over $\mathbb{Q}$ such that the sign of its functional equation is $+1$ (i.e. provided the L-function is even). Then, one can define the family

$$\{L_E(1,\chi_d)|\ d \text{ a fundamental discriminant, } \chi_d(n) = \left(\tfrac{d}{n}\right)\}, \tag{2.1.81}$$

which is ordered by $|d|$.

This is an example of an *orthogonal family.* The corresponding Riemann hypothesis for $L_E(s,\chi_d)$ places the critical line at $\mathrm{Re}(s) = 1$, rather than the conventional $1/2$. This is why the L-function is evaluated at 1 in (2.1.81). However, this difference is merely due to the conventional normalization one chooses when defining elliptic curve L-functions; it is simple to redefine L_E so that its critical line is shifted to $\mathrm{Re}(s) = 1/2$.

It is conjectured that (see for example [44])

$$\frac{1}{D^*}\sum_{|d|\le D}^{*} L_E(1,\chi_d)^{2\beta} \sim c_{L_E}(\beta)a_{L_E}(\beta)\,(\log D)^{\frac{2\beta(2\beta-1)}{2}}, \tag{2.1.82}$$

as $D \to \infty$. The sum is only over fundamental discriminants d; D^* is the length of the sum; and $a_{L_E}(\beta)$ again has a similar form to (2.1.32) (see Conrey et al. [48]). Once again, small moments, and hence small values of c_{L_E}, have been computed, see [44].

Keating and Snaith showed that the asymptotic behaviour of the 2βth moments of special orthogonal characteristic polynomials at the symmetry point is asymptotically

$$\int_{SO(2N)} |P_{SO(2N)}(A,0)|^{2\beta} dA \sim c_{SO}(\beta)N^{\frac{2\beta(2\beta-1)}{2}}, \tag{2.1.83}$$

where $c_{SO}(\beta)$ is the leading order moment coefficient depending on β, see (2.1.54). Like in the case of Dirichlet L-functions, by comparing N with $\log D$, the orthogonal symmetry is evident. Additionally, this furnishes the conjecture that $c_{L_E}(\beta)$ is equal to $c_{SO}(\beta)$ [115].

2.1.4 Mixed moments and log-derivative moments

To conclude the review of random matrix moments, we explore 'mixed' averages. These will be moments of characteristic polynomials and their derivatives. Once more, there will be a connection with the averages of certain number theoretical functions. Our results in chapter 7 build on the literature presented here. There, additionally, we show that the joint moments of characteristic polynomials and their derivatives are related to solutions of certain non-linear differential equations.

For the entirety of this section, we will focus just on the unitary polynomial case. For ease of notation we will write $P_N(A,\theta) \equiv P_{U(N)}(A,\theta)$. Then, one is interested in determining the following 'mixed moments',

$$F_N(h,k) := \int_{U(N)} |P_N(A,0)|^{2k-2h}|P_N'(A,0)|^{2h} dA, \tag{2.1.84}$$

which are 'mixed' in the sense that they concern products of the characteristic polynomial and its derivative.

However, rather than working directly with $P_N(A,\theta)$ it turns out to be profitable to consider

$$Z_A(\theta) := e^{\frac{iN}{2}(\theta+\pi)}e^{-\frac{i}{2}\sum_{j=1}^{N}\theta_j}P_N(A,\theta) \tag{2.1.85}$$

where the eigenvalues of A are $e^{i\theta_1},\dots,e^{i\theta_N}$. Then $Z_A(\theta)$ is real for real θ, and $|Z_A(\theta)| = |P_N(A,\theta)|$.

The mixed moments for Z_A were considered by Hughes[10] [98]

$$\tilde{F}_N(h,k) \coloneqq \int_{U(N)} |Z_A(0)|^{2k-2h} |Z_A'(0)|^{2h} dA \tag{2.1.86}$$

for $\mathrm{Re}(h) > -1/2$ and $\mathrm{Re}(k) > \mathrm{Re}(h) - 1/2$. Hughes proves the following result.

Theorem 2.1.12 (Hughes [98])**.** *For integer $h \geq 1$, and $k \geq h$ also integer,*

$$\tilde{F}_N(h,k) = \int_{U(N)} Z_A(0)^{2k-2h} Z_A'(0)^{2h} dA \tag{2.1.87}$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha^{2k}} \sum_{n=0}^{2h} (-1)^n \binom{2h}{n} e^{-\frac{iNn\alpha}{2}} T_{k,n}(\alpha) \tag{2.1.88}$$

where[11]

$$T_{k,n}(\alpha) = \det \begin{pmatrix} \binom{N+k}{k} & \cdots & \binom{N+2k-1}{k} \\ \vdots & & \vdots \\ \binom{N+k}{2k-n-1} & \cdots & \binom{N+2k-1}{2k-n-1} \\ z_{k-n+1,1}(\alpha) & \cdots & z_{k-n+1,k}(\alpha) \\ \vdots & & \vdots \\ z_{k,1}(\alpha) & \cdots & z_{k,k}(\alpha) \end{pmatrix} \tag{2.1.89}$$

and where for $i = k-n+1, \ldots, k$ and $j = 1, \ldots, k$,

$$z_{i,j}(\alpha) = \sum_{m=0}^{N+j-i} \binom{N+k+j-1}{k+i+m-1} \binom{m+i-k+n-1}{i-k+n-1} (e^{i\alpha} - 1)^m. \tag{2.1.90}$$

Hughes analyses (2.1.88) as $N \to \infty$, and shows that

$$\tilde{F}(h,k) \coloneqq \lim_{N \to \infty} \frac{1}{N^{k^2+2h}} \mathbb{E}\left[Z_A(0)^{2k-2h} Z_A'(0)^{2h}\right] \tag{2.1.91}$$

is analytic in k in the range $\mathrm{Re}(k) > h - 1/2$. However, the constraint of integrality for h remains. Additionally, Hughes explicitly computes $\tilde{F}(h,k)$ for small integer values of h. Using symmetric function theory, Dehaye [58] extended the result of Hughes by giving $\tilde{F}(h,k)$ in terms of a much simpler, rational, function of k (which again may be analytically extended), together with a ratio of Barnes $\mathcal{G}$-functions, though still for integer h. Precisely his result is as follows.

Theorem 2.1.13 (Dehaye [58])**.** *For $h, k \in \mathbb{N}$ with $2k - 2h > -1$,*

$$\int_{U(N)} |Z_A(0)|^{2k-2h} |Z_A'(0)|^{2h} dA = \frac{(-1)^h}{2^{2h}} \tilde{F}_N(0,k) \sum_{n=0}^{2h} \frac{(2h)!(-N)^{2h-n}}{(2h-n)!} C_N(n,k) \tag{2.1.92}$$

where C_N is expressed as a sum over partitions

$$C_N(n,k) \coloneqq (-2)^n \sum_{\substack{\lambda \vdash n \\ l(\lambda) \leq k}} \frac{[k]_\lambda [-N]_\lambda}{[2k]_\lambda h_\lambda^2} \tag{2.1.93}$$

[10] Hughes also explains how to recover the mixed moments $F_N(h,k)$ for $P_N(A,\theta)$ (see (2.1.84)) from those for $Z_A(\theta)$, denoted by $\tilde{F}_N(h,k)$.

[11] Since Z_A is real for real θ, one may freely drop the absolute values on Z_A, Z_A'.

with the generalized Pochhammer symbol $[c]_\lambda$ *defined as*

$$[c]_\lambda := \prod_{i=1}^{l(\lambda)} (c-i+1)_{\lambda_i} \tag{2.1.94}$$

$$= \prod_{i=1}^{l(\lambda)} \prod_{j=1}^{\lambda_i} (c-i+j) \tag{2.1.95}$$

and h_λ *is the 'hook-length' of the partition* λ *(see* (1.1.14)*).*

Additionally, the right hand side of (2.1.92) *extends meromorphically in* k *to* $\mathrm{Re}(k) > h - 1/2$ *for fixed* $h \in \mathbb{N}$*.*

Returning to examining $\tilde{F}_N(h,k)$, we make some connections to previously studied results. Note that $\tilde{F}_N(0,k)$ defined by (2.1.86) is exactly $M_N(k)$, as computed by Keating and Snaith, see theorem 2.1.1. Thus, the Hughes and Dehaye results recover the usual asymptotic for $F(0,k)$. Similarly, $\tilde{F}_N(k,k)$, i.e. moments of the derivative, can also be analysed by standard techniques [54]. In general, there is substantial interest in extending the results of Dehaye and Hughes to beyond integer h. The only result to date in this direction is the work of Winn [160], which allows $h = (2m-1)/2$ for $m \in \mathbb{N}$ (note that this means that the power on the derivative is still integral, though for the first time, odd).

Theorem 2.1.14 (Winn [160])**.** *Take* $m, k \in \mathbb{N}$ *with* $h := (2m-1)/2$ *and* $2k - 2h > -1$*. Then*

$$\int_{U(N)} |Z_A(0)|^{2k-2h} |Z_A'(0)|^{2h} dA$$

$$= \frac{(-1)^{h+\frac{1}{2}}}{2^{2h-1}\pi} \tilde{F}_N(0,k) \left(\sum_{n=1}^{2h} \sum_{l=1}^{n} \binom{2h}{n-l} \frac{(-1)^l}{l} (-N)^{2h-n} n! C_N(n,k) \right. \tag{2.1.96}$$

$$\left. + \sum_{n=2h+1}^{kN} \frac{(2h)!(n-2h-1)!}{N^{n-2h}} C_N(n,k) \right). \tag{2.1.97}$$

One motivation for studying mixed moments such as (2.1.86) again comes from number theory. In particular, $\tilde{F}(h,k)$ is conjecturally related to

$$\lim_{T\to\infty} \frac{1}{T} \frac{1}{\left(\log \frac{T}{2\pi}\right)^{k^2+2h}} \int_0^T |\zeta(\tfrac{1}{2}+it)|^{2k-2h} |\zeta'(\tfrac{1}{2}+it)|^{2h} dt, \tag{2.1.98}$$

just as the moments of $P_N(A,\theta)$ were related to the moments of $\zeta(1/2+it)$. For example, take Hardy's function

$$Z(t) := e^{i\phi(t)} \zeta(\tfrac{1}{2}+it), \tag{2.1.99}$$

where

$$\phi(t) = \mathrm{Im}\left(\log \frac{\Gamma\left(\frac{1+2it}{4}\right)}{\pi^{\frac{it}{2}}} \right), \tag{2.1.100}$$

(note that this mimics the definition (2.1.85) of $Z_A(\theta)$). Conrey and Ghosh [51] proved (under the assumption of the Riemann hypothesis) that

$$\frac{1}{T} \frac{1}{(\log T)^2} \int_1^T |Z(t)||Z'(t)| dt \sim \frac{e^2-5}{4\pi}. \tag{2.1.101}$$

This corresponds to $k=1$ and $h=1/2$ in (2.1.98). By allowing odd integer powers, Winn's result [160] shows that

$$\tilde{F}\left(\tfrac{1}{2},1\right) = \frac{e^2-5}{4\pi}. \tag{2.1.102}$$

Once again, one is able to find a connection between random matrix moments and (conjectural) number theoretic moments.

Chapter 7 presents our published results concerning mixed moments of the form (2.1.86). There, we also give the equivalent, complementary, number theoretic conjectures, following examples such as that of Winn and Conrey and Ghosh given here.

In chapter 7 we also handle certain moments of logarithmic derivatives of unitary characteristic polynomials,

$$\int_{U(N)} \left| \frac{P_N'(A,\theta)}{P_N(A,\theta)} \right|^{2k} dA. \tag{2.1.103}$$

Once more, such averages lead to conjectures for the equivalent number theoretic averages.

2.2 The Fyodorov-Keating conjectures

The second half of our review will concern a series of conjectures due to Fyodorov, Hiary, and Keating [81], and, in more detail, Fyodorov and Keating [82]. These conjectures are influenced by statistical mechanics and log-correlated fields (discussed in chapter 1 section 1.1.4) and concern both unitary characteristic polynomials as well as the Riemann zeta function.

We first state the conjectures in section 2.2.1, and then we explore why one should expect such behaviour in section 2.2.2. The remaining sections of this chapter are dedicated to reviewing the recent progress towards proving such conjectures. At the relevant points we highlight where our results (primarily those of chapter 3) contribute to the research landscape.

Besides section 2.2.4, for the remainder of this chapter we will be concerned with unitary characteristic polynomials. Thus, until we explicitly state otherwise, $P_N(A,\theta) \equiv P_{U(N)}(A,\theta)$ will represent the characteristic polynomial of $A \in U(N)$.

2.2.1 The conjectures

The following conjectures will be intrinsic to much of the latter part of this thesis. In particular, proving conjecture 2.2.4 is entirely the motivation behind the work presented in chapter 3, which, in turn, has implications for conjectures 2.2.2 and 2.2.3.

All of the following are, directly or indirectly, related to maxima of log-correlated fields. For context, we recall (1.1.36) here specialised with $\sigma^2 = \frac{1}{2}\log 2$. Specifically, we take V_n to be a metric space and $X_n = \{X_n(v), v \in V_n\}$ to be a log-correlated field with mean zero and $\mathbb{E}[X_n(v)^2] = \sigma^2 n = \frac{n}{2}\log 2$. We further write $N = 2^n$.

$$\max_{v\in V_n} X_n(v) \approx \log N - \frac{3}{4}\log\log N + \mathcal{M}, \tag{2.2.1}$$

where $\mathcal{M}$ is an $O(1)$ random variable. With these choices of σ^2 and N, the 'log-correlated constant' is renormalized to 3/4, versus the 'independent constant' of 1/4 (compared to the 3/2 vs. 1/2 between (1.1.36) and (1.1.32)).

The first, most general, conjecture of Fyodorov and Keating [82] is the following.

Conjecture 2.2.1. *For $\theta \in [0,L)$, $L \in (0,2\pi]$, and a matrix A sampled uniformly from $U(N)$,*

$$\max_{\theta\in[0,L)} \log|P_N(A,\theta)| \sim a_{N_L} + b_{N_L} x_{A,N_L}, \tag{2.2.2}$$

as $N_L \to \infty$, where,

$$a_{N_L} = \log N_L - \frac{3}{4}\log\log N_L + o(1) \quad \text{and} \quad b_{N_L} = 1 + O\left(\frac{1}{\log N_L}\right), \tag{2.2.3}$$

and where $N_L := NL/2\pi$ is the average number of eigenvalues of the associated $N \times N$ unitary matrix A in the interval $[0, L)$. The random variable x_{A,N_L} has probability density $p(x_{A,N_L})$.

The similarity between conjecture 2.2.1 and the maximum of the log-correlated field, see (2.2.1), is striking, and suggests that $\log|P_N(A,\theta)|$ has a log-correlated structure. Once more, we emphasise that if the constant on the subleading term for a_{N_L} in (2.2.2) (i.e. $-3/4$) was instead conjectured to be $-1/4$, and the limiting distribution of x_{A,N_L} were a single Gumbel random variable, then one would expect $\log|P_N(A,\theta)|$ to behave like *independent* Gaussian random variables. Instead, as shown in section 2.2.2, in fact $\log|P_N(A,\theta)|$ does exhibit logarithmic correlations[12] for different θ.

The 'full circle' case (i.e. $L = 2\pi$) is arguably the most interesting, and is where the most progress can be made. Given this, conjecture 2.2.1 becomes the following.

Conjecture 2.2.2. *For $A \in U(N)$ sampled uniformly, one expects*

$$\max_{\theta\in[0,2\pi)} \log|P_N(A,\theta)| = \log N - \frac{3}{4}\log\log N + x_{A,N}, \tag{2.2.4}$$

where $(x_{A,N}, N \in \mathbb{N})$ is a sequence of random variables which converge in distribution.

Fyodorov and Keating further conjecture that the random variable $x_{A,N}$ should converge in distribution to $x = \mathscr{G}_1 + \mathscr{G}_2$, a sum of two independent Gumbel random variables. Thus $p(x)$, the probability density for x, decays like $x\exp(-x)$ as $x \to \infty$:

$$p(x) = 2e^{-x}K_0\left(2e^{-\frac{x}{2}}\right) \overset{x\to\infty}{\sim} xe^{-x}, \tag{2.2.5}$$

where K_0 is a Bessel function.

Given the connection between random matrix polynomials and number theory presented in section 2.1.2, a link between conjecture 2.2.2 and maximum of the Riemann zeta function is to be expected. Recall that the correct choice is to identify $N = \log T/2\pi$, see the discussion around (2.1.29). Then, the number theoretic version of conjecture 2.2.2 due to Fyodorov and Keating is the following.

Conjecture 2.2.3. *Let $t \sim U[T, 2T]$ (i.e. t is taken uniformly from the interval $[T, 2T]$). Then*

$$\max_{h\in[0,2\pi)} \log|\zeta(\tfrac{1}{2} + i(t+h))| = \log\log T - \frac{3}{4}\log\log\log T + O_{\mathbb{P}}(1), \tag{2.2.6}$$

where $O_{\mathbb{P}}(1)$ encompasses a term bounded in probability as $T \to \infty$.

In section 2.2.2, we review the heuristic calculation that is used to justify conjecture 2.2.2 (and hence also conjecture 2.2.3). The final conjecture of Fyodorov and Keating relevant to this thesis is also based on said heuristic calculation (and thus, if proven in full generality, would provide a method for proving conjectures 2.2.2 and 2.2.3). A central function is the following moment,

$$g_N(\beta; A) := \frac{1}{2\pi}\int_0^{2\pi} |P_N(A,\theta)|^{2\beta} d\theta. \tag{2.2.7}$$

Note, these moments differ from those calculated by Keating and Snaith, see (2.1.2). The random variable $g_N(\beta; A)$ is 2βth moment of a fixed matrix $A \in U(N)$, averaged around the unit circle with

[12] There is also interest in the imaginary part of $\log P_N(A,\theta)$, see for example the paper of Fyodorov and Le Doussal [79].

respect to the uniform measure on it. This is random with respect to the matrix A. Thus, we may additionally define the moments of g_N with respect to the Haar measure on $U(N)$. Performing this combination is called the *moments of moments*, written

$$\mathrm{MoM}_{U(N)}(k,\beta) := \int_{U(N)} g_N(\beta;A)^k dA. \tag{2.2.8}$$

The final conjecture of Fyodorov and Keating [82] that we consider gives the asymptotic behaviour of $\mathrm{MoM}_{U(N)}(k,\beta)$.

Conjecture 2.2.4. *For $k \in \mathbb{N}$*

$$\mathrm{MoM}_{U(N)}(k,\beta) = \begin{cases} \left(\dfrac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k \dfrac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)} N^{k\beta^2} & \text{if } k < \frac{1}{\beta^2}, \\ \gamma_{k,\beta} N^{k^2\beta^2-k+1} & \text{if } k > \frac{1}{\beta^2}. \end{cases} \tag{2.2.9}$$

At the transition point $k = \beta^{-2}$, we will see that (c.f. chapter 3) the moments of moments will grow like $N \log N$.

Conjectures 2.2.2 and 2.2.3 will be the focus of the rest of this chapter. Justification for them is given in section 2.2.2. Significant progress has been made on both fronts and is covered in section 2.2.3. At the end of this chapter, see section 2.2.4, we sketch an argument which implies that, at least at leading order, conjecture 2.2.2 should still hold for symplectic and orthogonal characteristic polynomials. Conjecture 2.2.4 is the main motivation behind our work given in chapter 3 where, with an additional assumption on β, we establish the conjecture. There we also discuss further why one should expect conjecture 2.2.4 to hold. Chapter 4 explores the moments of moments of symplectic and orthogonal polynomials, and shows that the leading order behaviour differs from unitary, just as Keating and Snaith show for moments at the symmetry point, see theorems 2.1.1, 2.1.7, and 2.1.8.

2.2.2 Justification for conjectures 2.2.2, 2.2.3, and 2.2.4

Since conjecture 2.2.2 concerns the maximum of the real part of the logarithm of characteristic polynomials, it will be convenient to define

$$V_N(A,\theta) := -2\log|P_N(A,\theta)|. \tag{2.2.10}$$

The reasons for the -2 coefficient will become clear in what follows. The theorem of Keating and Snaith (theorem 2.1.3) reveals that $V_N(A,\theta)$ satisfies a central limit theorem,

$$V_N(A,\theta) \sim \mathcal{N}(0, 2\log N). \tag{2.2.11}$$

Furthermore, Hughes et al. [100] show that

$$\log|P_N(A,\theta)| = -\,\mathrm{Re}\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \frac{\mathrm{Tr}\, A^n}{\sqrt{n}} e^{-in\theta}. \tag{2.2.12}$$

The following result[13] of Diaconis and Shahshahani [64] concerning powers of traces of A is also useful.

Theorem 2.2.5 (Diaconis and Shashahani [64])**.** *Take a sequence $(X_j)_{j=1}^{\infty}$ of independent and identically distributed complex random variables whose real and imaginary parts are centred Gaussians with*

[13] Their result can also be deduced from the Strong Szegö theorem [150, 152].

variance 1/2. Then, for any fixed k and $A \in U(N)$, as $N \to \infty$,

$$\left(\operatorname{Tr} A, \tfrac{1}{\sqrt{2}} \operatorname{Tr} A^2, \ldots, \tfrac{1}{\sqrt{k}} \operatorname{Tr} A^k\right) \xrightarrow{d} (X_1, \ldots, X_k). \tag{2.2.13}$$

Hence, the coefficients $\operatorname{Tr} A^n/\sqrt{n}$ in (2.2.12) tend to independent and identically distributed complex Gaussian variables. By (2.2.12) together with theorem 2.2.5, see for example [82], one can show that

$$\mathbb{E}[V_N(A,\theta_1)V_N(A,\theta_2)] \sim -2\log 2\left|\sin\left(\frac{\theta_1-\theta_2}{2}\right)\right|, \tag{2.2.14}$$

where the average is over A with respect to the Haar measure. For large but finite N, one therefore has that $\mathbb{E}[V_N(A,\theta_1)V_N(A,\theta_2)] \approx -2\log|\theta_1-\theta_2|$ for $|\theta_1-\theta_2| \sim 1/N$ (the scale of mean eigenvalue spacing). If instead θ_1, θ_2 were much closer together than one would expect, then $\mathbb{E}[V_N(A,\theta_1)V_N(A,\theta_2)]$ would be well-approximated by the variance $\mathbb{E}[V_N(A,\theta_1)^2] = 2\log N$ (c.f. (2.2.11)). Hence, overall, as $N \to \infty$,

$$\mathbb{E}[V_N(A,\theta_1)V_N(A,\theta_2)] \approx \begin{cases} -2\log|\theta_1-\theta_2| & \text{for } \dfrac{1}{N} \ll |\theta_1-\theta_2| \ll 1 \\ 2\log N & \text{for } |\theta_1-\theta_2| \ll \dfrac{1}{N}. \end{cases} \tag{2.2.15}$$

Thus, the log-correlated behaviour of $V_N(A,\theta)$ (and hence $P_N(A,\theta)$) is apparent, providing justification for the appearance of $c = 3/4$ in conjecture 2.2.2 rather than $c = 1/4$ (which one would expect if the variables were independent).

Similarly, for the Riemann zeta function, one defines

$$V_\zeta(t,h) := -2\log\left|\zeta\left(\tfrac{1}{2} + i(t+h)\right)\right|, \tag{2.2.16}$$

for fixed $t \in \mathbb{R}$. Again, the presence of the coefficient -2 will be explained shortly. By Selberg's central limit theorem (c.f. theorem 1.1.6) we have that $V_\zeta(t,h)$ converges to a Gaussian distribution with a dependence on the shift h, as $t \to \infty$

$$V_\zeta(t,h) \sim \mathcal{N}(0, 2\log\log(t+h)). \tag{2.2.17}$$

Furthermore, the correlation for two points h_1, h_2 can be calculated in the following way (see [28] or the appendix of [82] for all the details). Firstly, we use the Euler product for zeta to expand each $V_\zeta(t,h_j)$,

$$\begin{aligned}&\mathbb{E}[V_\zeta(t,h_1)V_\zeta(t,h_2)] \\ &\quad= 4\sum_{p_1,p_2}\sum_{n_1,n_2=1}^{\infty}\frac{1}{n_1}\frac{1}{n_2}\frac{1}{p_1^{n_1/2}}\frac{1}{p_2^{n_2/2}}\mathbb{E}[\cos(n_1(t+h_1)\log p_1)\cos(n_2(t+h_2)\log p_2)]\end{aligned} \tag{2.2.18}$$

where the expectation is the average over $[t-h/2, t+h/2]$ for some h satisfying $1/\log t \ll h \ll t$. Note that in the large t limit this interval will contain an increasing number of zeros. Then by expanding the product of cosines in (2.2.18) and using that the main term comes from the diagonal contribution ($p_1 = p_2$, $n_1 = n_2$) with standard prime estimates, one finds

$$\mathbb{E}[V_\zeta(t,h_1)V_\zeta(t,h_2)] \approx \begin{cases} -2\log|h_1-h_2|, & \text{for } \dfrac{1}{\log t} \ll |h_1-h_2| \ll 1 \\ 2\log\log t, & \text{for } |h_1-h_2| \ll \dfrac{1}{\log t}. \end{cases} \tag{2.2.19}$$

The similarity to the covariance of $V_N(A,\theta)$ is evident. Once again there is a dependence on the distance between the points h_1, h_2. If h_1 is very close to h_2, then essentially $V_\zeta(A, h_1)$ and $V_\zeta(t, h_2)$ are perfectly correlated. However, if they are separated on the same scale as θ_1 and θ_2 must be in (2.2.15) (i.e. making the usual identification $N \sim \log t$), then instead one sees the logarithmic correlation.

Such structure has important ramifications. Recall that Selberg's central limit theorem reveals that on average $\log |\zeta(1/2 + it)|$ is on the order of $\sqrt{(1/2)\log\log t}$. The Lindelöf Hypothesis states that

$$|\zeta(\tfrac{1}{2} + it)| = o(t^\varepsilon) \tag{2.2.20}$$

for any $\varepsilon > 0$. Under the Riemann hypothesis, one has that

$$|\zeta(\tfrac{1}{2} + it)| = O\left(\exp\left(\frac{c_1 \log t}{\log\log t}\right)\right), \tag{2.2.21}$$

for some constant c_1 (see for example [154]). However, it is also known (without any assumptions) that

$$|\zeta(\tfrac{1}{2} + it)| = \Omega\left(\exp\left(\sqrt{\frac{\log t}{\log\log t}}\right)\right). \tag{2.2.22}$$

Above, Ω is as defined by Hardy and Littlewood [91], so $f(x) = \Omega(g(x))$ should be interpreted saying that $f(x)$ takes the value $g(x)$ infinitely often. Hence, maxima of zeta must lie between (2.2.21) and (2.2.22).

If it happened that $\log |\zeta(1/2 + i(t + h_1))|$ and $\log |\zeta(1/2 + i(t + h_2))|$ could be considered independent of each other, then the Fisher-Tippett-Gnedenko theorem (see chapter 1, section 1.1.3, and theorem 1.1.8) could be applied. Such an assumption was made by Montgomery when attempting to estimate the typical size of $\log |\zeta(1/2+it)|$. If one sets $X_1, \ldots, X_n$ to be n local maxima of $\log |\zeta(1/2+it)|$, assumed to be independent, then theorem 1.1.8 would imply that the typical size is

$$\log |\zeta(\tfrac{1}{2} + it)| = O\left(\exp\left(c_2 \sqrt{\log t \log\log t}\right)\right) \tag{2.2.23}$$

for some constant c_2. Montgomery's conjecture is considerably larger than $\sqrt{(1/2)\log\log t}$, the typical value of $\log |\zeta(1/2 + it)|$, and is closer to (2.2.22) than (2.2.21). However, as now established, Montgomery's assumption of independence is incorrect.

In random matrix theory, the range of theta in conjecture 2.2.2 is natural. Number theoretically, however, there is no such periodicity and so different ranges for the maximum are just as valid. Indeed, 'long range' maxima (for ranges of length $O(T)$ rather than $O(1)$) have been the subject of much recent study.

Bondarenko and Seip [25, 26] show that

$$\max_{t \in [1,T]} |\zeta(\tfrac{1}{2} + it)| \geq \exp\left((1 + o(1))\sqrt{\frac{\log T \log\log\log T}{\log\log T}}\right). \tag{2.2.24}$$

This has recently been improved to

$$\max_{t \in [1,T]} |\zeta(\tfrac{1}{2} + it)| \geq \exp\left((\sqrt{2} + o(1))\sqrt{\frac{\log T \log\log\log T}{\log\log T}}\right). \tag{2.2.25}$$

by de la Bretèche and Tenenbaum [30]. Both results are unconditional, and more generally cover intervals of length $[T^\alpha, T]$ for $\alpha \in [0, 1)$.

Using techniques from random matrix theory, Farmer, Gonek and Hughes [73] conjecture

$$\max_{t\in[1,T]} |\zeta(\tfrac{1}{2}+it)| = \exp\left(\left(\tfrac{1}{\sqrt{2}}+o(1)\right)\sqrt{\log T\log\log T}\right). \tag{2.2.26}$$

A different, and perhaps more tractable, question therefore is to determine the correct size of the maximum in *short* intervals. Ranges of length $O(1)$ (rather than $O(T)$ considered in (2.2.25) and (2.2.25)) are precisely those covered by conjecture 2.2.3. In particular, since these ranges are so much shorter, numerical calculations now become feasible.

We now focus our attention on the techniques used by Fyodorov and Keating to construct the precise form of conjectures 2.2.2, 2.2.3, and 2.2.4. This technique is inspired by a class of problems in statistical mechanics.

Recall the definition of $g_N(\beta;A)$ from (2.2.7),

$$g_N(\beta;A) := \frac{1}{2\pi}\int_0^{2\pi} |P_N(A,\theta)|^{2\beta}d\theta \equiv \frac{1}{2\pi}\int_0^{2\pi} e^{-\beta V_N(A,\theta)}d\theta, \tag{2.2.27}$$

where $\beta > 0$. Compare the definition of g_N to (1.1.38): both can be viewed as partition functions in the language of statistical mechanics. Then, $V_N(A,\theta)$ is the 'energy' for the system and β is the inverse temperature. The reason for including -2 in the definition of $V_N(A,\theta)$ and $V_\zeta(t,h)$ is now also apparent.

Other, similar, problems have been addressed by re-expressing the question in terms of the language of statistical mechanics (see in particular the paper of Fyodorov and Bouchaud [78]). Recall from section 1.1.4 that a related, important function is the corresponding *free energy* for the system. Fyodorov and Keating exactly work with the free energy,

$$\mathcal{F}(\beta) := -\frac{1}{\beta}\log g_N(\beta;A). \tag{2.2.28}$$

and show that the maximum of $\log|P_N(A,\theta)|$ can be recovered as the large β limit

$$\lim_{\beta\to\infty}\mathcal{F}(\beta) = -2\max_{\theta\in[0,2\pi)}\log|P_N(A,\theta)|. \tag{2.2.29}$$

A similar construction can be made for $\log|\zeta(1/2+it)|$.

Fyodorov and Keating also point out that the free energy (2.2.28) demonstrates the 'freezing' property (related to be parameter β, which in the statistical mechanics language represents inverse temperature). What they mean by this will be clear shortly. Firstly, they define the *normalized* free energy to be

$$\mathscr{F}(\beta) := -\frac{1}{\beta\log N}\log\left(Ng_N(\beta;A)\right). \tag{2.2.30}$$

By considering the average of $\mathscr{F}(\beta)$ with respect to the Haar measure, they argue (see [82]) that for β small (i.e. high temperature) the average of $\mathscr{F}(\beta)$ is governed by the typical values of $P_N(A,\theta)$. Recall that the Keating and Snaith result gives that the moments behave like N^{β^2} (c.f. theorem 2.1.1).

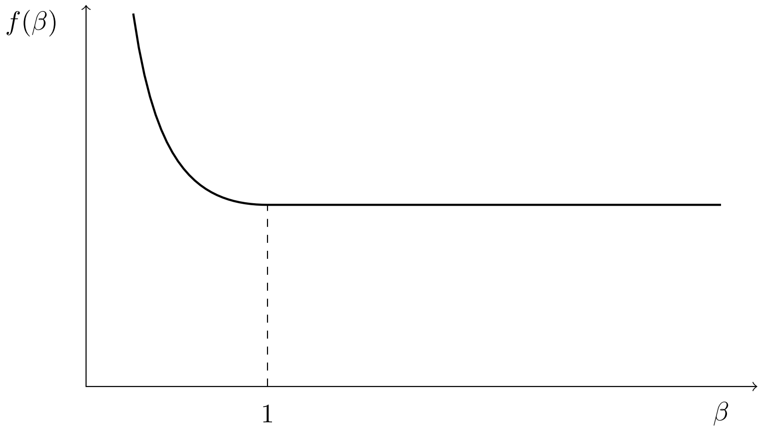


Figure 2.2: Graph showing freezing of the free energy as β (inverse temperature) changes, where $f(\beta) = -\lim_{N\to\infty} \mathbb{E}[\mathscr{F}(\beta)]$.

Hence, for small β one expects

$$-\mathbb{E}[\mathscr{F}(\beta)] = \frac{1}{\beta \log N}\mathbb{E}[\log(N g_N(\beta; A))] \tag{2.2.31}$$
$$= \frac{1}{\beta \log N}\mathbb{E}\left[\log\left(\frac{N}{2\pi}\int_0^{2\pi} |P_N(A,\theta)|^{2\beta} d\theta\right)\right] \tag{2.2.32}$$
$$\sim \frac{\log\left(N \cdot N^{\beta^2}\right)}{\beta \log N} \tag{2.2.33}$$
$$= \beta + \frac{1}{\beta}. \tag{2.2.34}$$

However, as β grows large (i.e. as temperature decreases), the free energy will instead be governed by the extreme values, see (2.2.29). The Fyodorov-Keating conjecture 2.2.2 gives that extreme values of $\log|P_N(A,\theta)|$ should scale as $\log N$ to leading order. Hence in the large N limit for large β, instead

$$-\mathbb{E}[\mathscr{F}(\beta)] = \frac{1}{\beta \log N}\mathbb{E}\left[\log\left(\frac{N}{2\pi}\int_0^{2\pi} |P_N(A,\theta)|^{2\beta} d\theta\right)\right] \tag{2.2.35}$$
$$\sim \frac{\log\left(N \cdot N^{2\beta}\right)}{\beta \log N} \tag{2.2.36}$$
$$\to 2, \tag{2.2.37}$$

as β grows large. This is the meaning of *freezing* for this system: as the temperature moves from small to large β (i.e. temperature decreases), the free energy reaches a critical temperature and thereafter remains constant. Here, the critical temperature is $\beta = 1$, see figure 2.2. Hence, as N grows large,

$$-\mathbb{E}[\mathscr{F}(\beta)] \sim \begin{cases} \beta + \frac{1}{\beta} & \text{if } \beta \leq 1 \\ 2 & \text{if } \beta > 1. \end{cases} \tag{2.2.38}$$

Returning to the question of the maximum of $P_N(A,\theta)$, notice that the free energy also provides insight. If one has sufficiently fine information about the moments of the random variable $g_N(\beta; A)$ with respect to the Haar measure on $U(N)$, then by (2.2.29) this would reveal information about the maximum of $\log|P_N(A,\theta)|$ (precisely by performing the calculation outlined in (2.2.29)). The exact quantity of interest is

$$\mathbb{E}[g_N(\beta;A)^k] = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi}\mathbb{E}\left[\prod_{j=1}^{k}|P_N(A,\theta_j)|^{2\beta}\right]d\theta_1\cdots d\theta_k, \tag{2.2.39}$$

where $\mathbb{E}[\cdot]$ is the Haar measure on $U(N)$, and provided that $k\in\mathbb{N}$.

As is further explored in chapter 3, section 3.1.1 and chapter 5, the integrand of (2.2.39) can be written as a Toeplitz determinant $D_N(f)$ with symbol $f(z)=\prod_{j=1}^{k}|z-e^{i\theta_j}|^{2\beta}$. The values $z=e^{i\theta_j}$ give rise to so-called 'Fisher-Hartwig' singularities. Using Widom's result [157] on the Fisher-Hartwig asymptotic formula, the integrand of (2.2.39) can be seen to be

$$\mathbb{E}\left[\prod_{j=1}^{k}|P_N(A,\theta_j)|^{2\beta}\right]\propto|\Delta(e^{i\theta_1},\ldots,e^{i\theta_k})|^{-\beta^2} \tag{2.2.40}$$

as $N\to\infty$. Thus, provided that the Fisher-Hartwig singularities at $e^{i\theta_1},\ldots,e^{i\theta_k}$ remain fixed and distinct, one can apply Selberg's integral and find that

$$\mathbb{E}[g_N(\beta;A)^k]\sim\left(\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k\frac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)}N^{k\beta^2}. \tag{2.2.41}$$

In order to ensure that the Fisher-Hartwig singularities remain fixed and distinct, one requires the restriction $k<1/\beta^2$, see for example [82]. This is the justification (and indeed, the outline of the proof) for this regime of conjecture 2.2.4. Outside of this range, i.e. $k>1/\beta^2$, the Fisher-Hartwig singularities coalesce, and so computing $\mathbb{E}[g_N(\beta;A)^k]$ proves to be much more difficult. In fact, one either requires a uniform Fisher-Hartwig asymptotic formula valid for when the singularities coalesce, or an alternative approach. Indeed, it is precisely the contributions from the coalescing singularities that leads to the different leading order behaviour in conjecture 2.2.4. For further discussion and results on this regime, see chapter 3.

This concludes the review of the heuristics behind conjectures 2.2.2, 2.2.3, and 2.2.4. We now summarize the recent advances towards proving the two 'maxima' Fyodorov-Keating conjectures. The final conjecture (conjecture 2.2.4) concerning the leading order of the *moments of moments* is the subject of chapter 3.

2.2.3 Progress towards conjectures 2.2.2 and 2.2.3

There has been much interest in resolving the conjectures of Fyodorov and Keating in recent years. This has been almost entirely successful in the random matrix case, and significant progress has been made towards settling the number theoretic case.

As has now been evidenced multiple times, typically problems in random matrix theory are more tractable than their equivalent formulations in number theory. Thus, one often uses the random matrix results to inform the number theoretic calculation. However, in this case, a model of the Riemann zeta function was the lynchpin to almost all the subsequent successful improvements. Thus, we first present this number theoretic model, and state the result of Arguin, Belius, and Harper [7]. Their work shows that, to subleading order and including the predicted subleading coefficient, the model of the Riemann zeta function follows the Fyodorov and Keating conjecture.

We subsequently give an outline of their proof, and quickly focus on the identification of the approximate branching structure found within. It is highly instructive to first map out the calculation for an *exact* branching random walk. These techniques are fundamental not only to the result of Arguin, Belius, and Harper [7], but also to all the proofs of the ensuing random matrix and number theoretic results. A discussion of these forms the remaining part of this section.

Identification of a branching structure in a model of $\zeta(1/2+it)$

It is well known that the primes display a behaviour akin to 'deterministic chaos'. To try to understand the Riemann zeta function, and in particular the behaviour of its maximum in short intervals, one might try to use the pseudo-random structure and try to model it. Using work of Soundararajan [146], Harper [92] constructed, and demonstrated the validity of, a randomised model of the zeta function. Subsequently, Arguin et al. [7] examined this model and established to subleading order an adaptation of the conjecture of Fyodorov and Keating in this case. A crucial part of this work is the discovery of an approximate *tree structure* in the model of $\zeta(s)$. As will be discussed, it is also possible to demonstrate such structure within the true zeta function. Recall that branching random walks were introduced in chapter 1, section 1.1.4.

We here restate conjecture 2.2.3, though the maximum is taken over a (trivially) different length of interval. For $\tau \in [T, 2T]$ chosen uniformly at random,

$$\max_{h\in[0,1]} \log|\zeta(\tfrac{1}{2}+i(\tau+h))| = \log\log T - \frac{3}{4}\log\log\log T + O_{\mathbb{P}}(1), \tag{2.2.42}$$

where $O_{\mathbb{P}}(1)$ is a term bounded in probability as $T\to\infty$.

Harper's idea is to construct a random model of the zeta function by modelling the primes in the Euler product representation as random variables on the unit circle. He shows that [92], under the Riemann hypothesis, there exists a set H of measure at least 0.99 with $H \subseteq [T, T+1]$ such that

$$\log|\zeta(\tfrac{1}{2}+i\eta)| = \mathrm{Re}\left\{\sum_{p\le T} p^{-(\frac{1}{2}+i\eta)}\frac{\log\frac{T}{p}}{\log T}\right\} + O(1), \tag{2.2.43}$$

for all $\eta \in H$. The set of small measure $[T, T+1]\backslash H$ where the result fails essentially covers those points close to the zeros of zeta. In order to capture the 'quasi-randomness' of the primes, Harper introduces the following random variables. Let $(U_p, p \text{ prime})$ be a sequence of independent random variables distributed uniformly on the unit circle[14]. Heuristically speaking[15], U_p models $p^{-i\tau}$, see also figure 2.3. Thus, when $\eta = \tau + h$, we make this substitution (and drop the term $\log(T/p)/\log T$ since it only plays a small 'smoothing' role in (2.2.43)). It is then easy to justify, see for example [7], that the random variable

$$\sum_{p\le T}\frac{\mathrm{Re}(p^{-ih}U_p)}{\sqrt{p}} \tag{2.2.44}$$

for $h\in[0,1]$ is a good model for $\log|\zeta(1/2+i(\tau+h))|$.

Arguin et al. [7] prove the following.

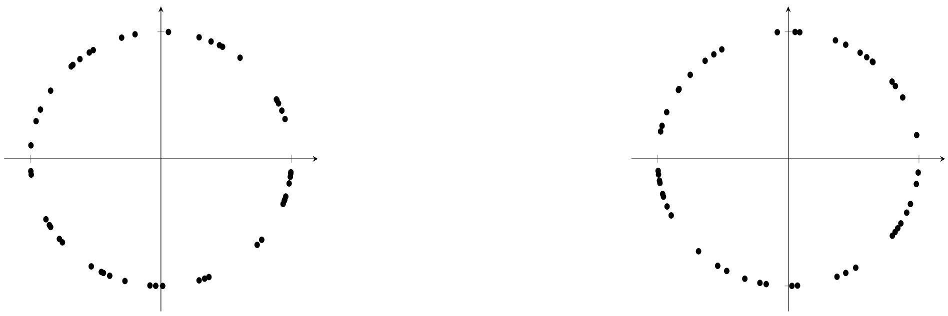
Theorem 2.2.6 (Arguin et a. [7])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(U_p, p$ prime$)$ be independent random variables distributed uniformly on the unit circle. Then*

$$\max_{h\in[0,1]}\sum_{p\le T}\frac{\mathrm{Re}(p^{-ih}U_p)}{\sqrt{p}} = \log\log T - \frac{3}{4}\log\log\log T + o_{\mathbb{P}}(\log\log\log T).$$

The error term $o_{\mathbb{P}}(\log\log\log T)$ converges to zero in probability when divided by $\log\log\log T$.

[14] In the literature, these are sometimes referred to as *Steinhaus* random variables.

[15] One can make the argument more rigorous, see [92]



(a) Plotting 50 evaluations of a Steinhaus random variable.

(b) Plotting p^{-100i} for the first fifty primes p.

Figure 2.3: Comparing 50 evaluations of a Steinhaus random variable with $p^{-i\tau}$, for the first 50 primes and $\tau = 100$.

In other words, the maximum over $h \in [0,1]$ of the model for $\log|\zeta(1/2+i(\tau+h))|$ where τ is sampled uniformly from $[T, 2T]$ matches both the leading and subleading order terms of conjecture 2.2.3.

We now outline the key technique that allowed Arguin et al. to prove the above result. We particularly concentrate on the identification of the approximate branching structure. Once this structure is apparent, we then discuss the general method of proof that would be employed if one instead had an *exact* branching structure. It is this recipe that proves key to the success of the proof of theorem 2.2.6, and to the progress towards conjectures 2.2.2 and 2.2.3 detailed later within this section.

In order to make the comparison with the branching random walk clear, we set without loss of generality $T = e^{2^n}$ for some large $n \in \mathbb{N}$, where T is the height at which the interval is situated up the critical line. Now set

$$M_{2^n} := \log 2^n - \frac{3}{4}(\log\log 2^n) \tag{2.2.45}$$
$$= n\log 2 - \frac{3}{4}\log n + O(1), \tag{2.2.46}$$

i.e. the right hand side of (2.2.42) given the above assumption that $\log T = 2^n$.

Define the random process

$$\left(X_n(h) := \sum_{p \le e^{2^n}} \frac{\mathrm{Re}(p^{-ih}U_p)}{\sqrt{p}}, h \in [0,1] \right). \tag{2.2.47}$$

Thus, theorem 2.2.6 follows if one can show

$$\lim_{n\to\infty} \mathbb{P}\left(M_{2^n}(-\varepsilon) \le \max_{h\in[0,1]} X_n(h) \le M_{2^n}(\varepsilon) \right) = 1, \quad \forall \varepsilon > 0, \tag{2.2.48}$$

where

$$M_{2^n}(\varepsilon) = n\log 2 - \frac{3}{4}\log n + \varepsilon \log n \tag{2.2.49}$$

and $X_n(h)$ is defined by (2.2.47).

Clearly, $M_{2^n}(\varepsilon)$ is reminiscent of the maximum of the branching random walk, see (1.1.49). This observation implies that there may be branching structure evident within $X_n(h)$. In order to pursue this further, denote the summands in (2.2.47) as

$$W_p(h) := \frac{\mathrm{Re}(U_p p^{-ih})}{\sqrt{p}}. \tag{2.2.50}$$

Straightforwardly one shows

$$\mathbb{E}[X_n(h_1)] = 0 \tag{2.2.51}$$

$$\mathbb{E}\left[W_p(h_1)W_p(h_2)\right] = \frac{1}{2p}\cos(|h_1 - h_2|\log p) \tag{2.2.52}$$

$$\mathbb{E}\left[X_n(h_1)X_n(h_2)\right] = \frac{1}{2}\sum_{p\le e^{2^n}} \frac{\cos(|h_1 - h_2|\log p)}{p} \tag{2.2.53}$$

for $h_1, h_2 \in [0,1]$. The first calculation (2.2.51) follows by symmetry. For the second, (2.2.52), one first rewrites $\mathrm{Re}(U_{p_j}p_j^{-ih_j})$ for $j = 1,2$ using U_p and $\overline{U_p}$. Then one uses the fact that all terms are zero in expectation unless $p_1 = p_2$. The final equality follows immediately from (2.2.52) and (2.2.47).

Thus the covariance of $X_n(h)$ depends on distance between the points h_1 and h_2 in a *logarithmic* way:

$$\mathbb{E}\left[X_n(h_1)X_n(h_2)\right] \approx \begin{cases} \frac{1}{2}\log|h_1 - h_2|^{-1} & \text{if } |h_1 - h_2| \ge 2^{-n} \\ \frac{1}{2}\log 2^n & \text{if } |h_1 - h_2| < 2^{-n}. \end{cases} \tag{2.2.54}$$

That is, if $|h_1 - h_2| \ge 2^{-n}$ (the scale of separation on average) then one can estimate $\frac{1}{2}\log|h_1 - h_2|^{-1}$ from (2.2.53). However, if the two points lie closer together than 2^{-n}, i.e. they lie 'unusually' close, then the covariance is approximately

$$\mathbb{E}\left[X_n(h_1)X_n(h_2)\right] = \frac{1}{2}\sum_{p\le e^{2^n}} \frac{1}{p} \approx \frac{1}{2}\log 2^n. \tag{2.2.55}$$

An interpretation of this observation is that if the points h_1, h_2 are far enough apart, then the $X_n(h_1)$ and $X_n(h_2)$ are almost exactly uncorrelated. However, when h_1 and h_2 lie close together on the scale of mean separation, there is almost perfect correlation.

The above calculation should be compared to those for the exact branching random walk (which we do shortly). Indeed, Arguin et al. comment that these approximate results are exact in that setting. To clearly exhibit the source of the branching structure, they break the sum in (2.2.47) in to dyadic-like partitions, defining

$$Y_m(h) := \sum_{2^{m-1}<\log p\le 2^m} W_p(h) = \sum_{2^{m-1}<\log p\le 2^m} \frac{\mathrm{Re}(U_p p^{-ih})}{\sqrt{p}}. \tag{2.2.56}$$

Hence, by (2.2.52) we have

$$\mathbb{E}[Y_m(h)^2] = \sum_{2^{m-1}<\log p\le 2^m} \frac{1}{2p} \tag{2.2.57}$$

$$\mathbb{E}[Y_m(h_1)Y_m(h_2)] = \sum_{2^{m-1}<\log p\le 2^m} \frac{1}{2p}\cos(|h_1 - h_2|\log p). \tag{2.2.58}$$

Thus the model of the Riemann zeta function can be written as

$$X_n(h) = \sum_{m=0}^{n} Y_m(h), \tag{2.2.59}$$

resembling a branching random walk as described in section 1.1.4.

The following lemma of Arguin et al. gives the distance at which two walks $X_n(h_1)$ and $X_n(h_2)$

become essentially uncorrelated. For ease of notation write

$$h_1 \wedge h_2 := \lfloor \log_2 |h_1 - h_2|^{-1} \rfloor. \tag{2.2.60}$$

This 'wedge' notation should be compared to the last common ancestor of two leaves h_1, h_2 of a binary tree, recall definition 1.1.9.

Lemma 2.2.7 (Arguin et al. [7]). *For $h_1, h_2 \in \mathbb{R}$, $m \geq 1$,*

$$\mathbb{E}[Y_m(h_1)^2] = \frac{1}{2}\log 2 + O\left(e^{-c\sqrt{2^m}}\right), \tag{2.2.61}$$

$$\mathbb{E}[Y_m(h_1)Y_m(h_2)] = \begin{cases} \frac{1}{2}\log 2 + O\left(2^{-2(h_1\wedge h_2 - m)}\right) + O\left(e^{-c\sqrt{2^m}}\right) & \text{if } m \leq h_1 \wedge h_2, \\ O\left(2^{-(m-h_1\wedge h_2)}\right) & \text{if } m > h_1 \wedge h_2, \end{cases} \tag{2.2.62}$$

for some constant c.

The proof follows from a strong form of the prime number theorem and integration by parts. Often, one has to handle the case of $m = 0$ (i.e. the contribution from small primes) separately, which is the cause of the requirement $m \geq 1$ in the statement of lemma 2.2.7.

Lemma 2.2.7 shows that the $Y_m(h)$, which act as the increments of the branching random walk, are essentially perfectly correlated when h_1 and h_2 lie close, relative to m the index of the increment (which will correspond to depth in the binary tree). Otherwise, effectively, they are perfectly uncorrelated.

The proof of theorem 2.2.6 is inspired by the techniques that one would use if $X_n(h)$ were an exact branching random walk. Thus, we now outline the method in this precise setting, and comment how Arguin et al. are able to adapt this proof to the approximate situation for the model of the Riemann zeta function. As mentioned above, this method is also used in the results towards conjectures 2.2.2 and 2.2.3.

General method of proof

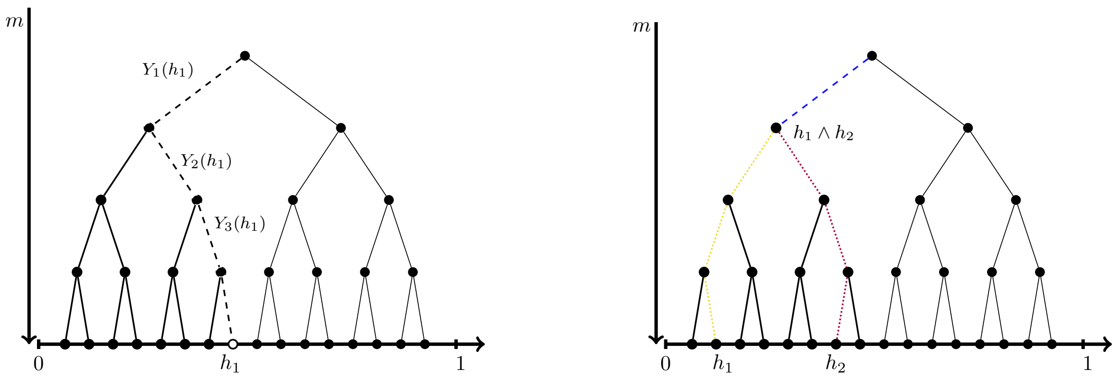
Harper's model of Riemann zeta function, (2.2.47), has now been shown by Arguin et al. [7] to demonstrate approximate branching. We now wish to compare this with the branching random walk introduced in section 1.1.4. The exact analogy is as follows. Create a binary tree of depth n so that its 2^n leaves are equally spaced within the interval $[0, 1]$ - see figure 2.4a. To each branch attach an independent, centred Gaussian random variable with variance $\sigma^2 = \frac{1}{2}\log 2$. The m levels, $m = 1, \dots, n$ correspond to the 'dyadic' decomposition of the primes for the Riemann zeta model (2.2.56) (again, the contribution from small primes, $m = 0$, is handled separately).

Recall that, given two nodes within the tree h_1, h_2, the last common ancestor is the largest m such that the walks from root to h_1 and h_2 are the same up to level m, and diverge thereafter. Usually we write $m = \text{lca}(h_1, h_2)$, but to emphasise the analogy with the model of $\zeta(s)$, we here write $m = h_1 \wedge h_2$ - c.f. (2.2.60). Figure 2.4b shows two walks from root to leaf nodes h_1 and h_2, following the same (dashed) path up to level $h_1 \wedge h_2$. Thereafter they diverge and hence become independent.

The general method used to prove theorem 2.2.6 is based on the following recipe, applied to the exact branching random walk (see also [4] and the introduction of [7]). Many of the results towards conjectures 2.2.2 and 2.2.3 also use the same method.

The first step is to *identify a branching structure*. For the branching random walk, this is trivially

$$X_n(h) = \sum_{m=1}^{n} Y_m(h), \tag{2.2.63}$$



(a) An exact branching random walk on a binary tree of depth 4, with random variable weightings $Y_m(h_l) \sim \mathcal{N}(0, \frac{1}{2}\log 2)$ for level $m \in \{1, 2, 3, 4\}$, and leaf $l \in \{1, \ldots, 2^4\}$.

(b) A tree structure with an example of the level $h_1 \wedge h_2$, after which the walks $X_4(h_1)$ (i.e. the (red) path from root to h_1) and $X_4(h_2)$ (i.e. the (yellow) path from root to h_2) become uncorrelated.

Figure 2.4: Pictorial representation of the exact branching structure for the model of $\zeta(s)$.

where $Y_m(h) \sim \mathcal{N}(0, \frac{1}{2}\log 2)$. Compare this to the approximate branching structure for the model of $\zeta(s)$, (2.2.59).

One now considers the *number of exceedances*,

$$Z(t) := \left|\{l \leq 2^n : X_n\left(\tfrac{l}{2^n}\right) \geq t\}\right|, \tag{2.2.64}$$

i.e. the number of leaves[16] $\frac{l}{2^n}$ such that the value of the walk associated to that leaf exceeds the value t. The relationship between maxima and $Z(t)$ is clearly the following

$$\max_{l \in \{1,\ldots,2^n\}} X_n\left(\tfrac{l}{2^n}\right) \geq t \quad \iff \quad Z(t) \geq 1. \tag{2.2.65}$$

In order to identify the correct size of the maximum of $X_n(h)$, one finds t so that $\mathbb{P}(Z(t) \geq 1) = o(1)$ (c.f. theorem 1.1.10 and the subsequent sketch of proof).

Thus, one proceeds to bound $\mathbb{P}(Z(t) \geq 1)$. An *upper bound* is attained through a union bound,

$$\mathbb{P}(Z(t) \geq 1) \leq 2^n \mathbb{P}(X_n\left(\tfrac{1}{2^n}\right) \geq 0). \tag{2.2.66}$$

Standard Gaussian tail estimates give

$$\mathbb{P}(Z(t) \geq 1) \leq 2^n \mathbb{P}(X_n\left(\tfrac{1}{2^n}\right) \geq t) \tag{2.2.67}$$

$$\approx e^{\log 2^n} \frac{\sqrt{n}}{t} e^{-\frac{t^2}{\log 2^n}} \tag{2.2.68}$$

$$= \frac{\sqrt{n}}{t} e^{-\frac{t^2}{\log 2^n} + \log 2^n}, \tag{2.2.69}$$

implying that $\mathbb{P}(Z(t) \geq 1) \leq o(1)$ when $t = t(n) = \log 2^n - (\frac{1}{4} - \varepsilon)\log n$. Recall from chapter 1, section 1.1.3, that this is approximately the correct order of the maximum for *independent* Gaussian random variables. However, one instead here has *log-correlated* random variables. This means that we expect the maximum to be lower than this value of $t(n)$. Shortly, we discuss an adaptation to this upper bound that delivers the correct size of the maximum.

[16]One assumes that the leaves are equally spaced at $\frac{1}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1$ in $[0, 1]$.

The Paley-Zygmund inequality delivers a *lower bound*,

$$\mathbb{P}(Z(t) \geq 1) \geq \frac{\mathbb{E}[Z(t)]^2}{\mathbb{E}[Z(t)^2]}, \tag{2.2.70}$$

and similarly one can show [5,7] for the branching random walk that the second moment is exponentially larger than the first moment squared; it is inflated by those exceeding walks 'pulling up' neighbours.

Altogether, this implies that $Z(t)$ is not the right quantity to consider. Instead, one alters the definition of $Z(t)$ to take in to account strong structure underlying the model. It turns out, [7], that with high probability, a walk $X_m(h)$ up to level $1 \leq m \leq n$ lies below a *linear barrier* $\log 2^m + B$ for some B growing slowly with n. If rather one works with

$$\tilde{Z}(t) \coloneqq |\{l \leq 2^n : X_n\left(\tfrac{l}{2^n}\right) \geq t, \text{ and } X_m\left(\tfrac{l}{2^n}\right) \leq \log 2^m + B, \forall m \leq n\}| \tag{2.2.71}$$

then (with some modifications), calculating the above upper and lower bounds (2.2.66) and (2.2.70) with $\tilde{Z}(t)$ replacing $Z(t)$ is precisely the right approach.

Making such a method rigorous for functions with *approximate* branching structure, such as the model of the Riemann zeta function described above, is where the technicalities lie. Generally, however, all proofs of this type follow the method outlined above.

Progress Towards Conjecture 2.2.2

We will now present an overview of the stages of progress made following the paper of Arguin, Belius, and Harper [7], and subsequently highlight the approximate branching structure in each case.

Recall that the Fyodorov-Keating conjecture (conjecture 2.2.2) for the maximum of $\log |P_N(A, \theta)|$ is the following. For $A \in U(N)$ sampled uniformly with respect to the Haar measure, we expect

$$\max_{\theta \in [0, 2\pi)} \log |P_N(A, \theta)| = \log N - \frac{3}{4} \log \log N + x_{A,N}, \tag{2.2.72}$$

where $(x_{A,N}, N \in \mathbb{N})$ is a sequence of random variables which converge in distribution.

The first step towards a proof of this conjecture was made by Arguin, Belius, and Bourgade [5]. The following is their main theorem (theorem 1.2 in [5]), establishing the conjecture to leading order.

Theorem 2.2.8 (Arguin, Belius, and Bourgade [5])**.** *For $N \in \mathbb{N}$, let $A \in U(N)$ be sampled uniformly with respect to the Haar measure. Then*

$$\lim_{N \to \infty} \frac{\max_{\theta \in [0, 2\pi)} \log |P_N(A, \theta)|}{\log N} = 1 \quad \textit{in probability.}$$

Soon after, the work of Paquette and Zeitouni [132] verified that the subleading term in the conjecture is correct. Precisely, their main result (theorem 1.2 in [132]) is as follows.

Theorem 2.2.9 (Paquette and Zeitouni [132])**.** *For $N \in \mathbb{N}$, let $A \in U(N)$ be sampled uniformly. Then*

$$\lim_{N \to \infty} \frac{\max_{\theta \in [0, 2\pi)} \log |P_N(A, \theta)| - \log N}{\log \log N} = -\frac{3}{4} \quad \textit{in probability.}$$

Importantly, their work establishes the constant $-3/4$, the characteristic coefficient in the subleading term for processes with logarithmic correlations. Finally, the best result as of the time of writing, is the work of Chhaibi, Madaule, and Najnudel [35] who answer the conjecture up to tightness (theorem 1.2 in [35]).

Theorem 2.2.10 (Chhaibi, Madaule, and Najnudel [35]). *If $A \in U(N)$ is chosen uniformly with respect to the Haar measure, and $\{\theta_1, \ldots, \theta_N\}$ is the set of eigenphases of A, then the family of random variables*

$$\left(\max_{\theta \in [0,2\pi) \setminus \{\theta_1,\ldots,\theta_N\}} \log |P_N(A,\theta)| - \left(\log N - \frac{3}{4} \log\log N \right) \right)_{N \geq 2} \tag{2.2.73}$$

is tight.

Chhaibi et al. in fact prove a stronger statement concerning the CβE, the *Circular Beta Ensemble*, a probability distribution on n points on the unit circle,

$$\frac{1}{(2\pi)^n Z_{n,\beta}} \prod_{1 \leq j < k \leq n} |e^{i\theta_j} - e^{i\theta_k}|^\beta d\theta_1 \cdots d\theta_n, \tag{2.2.74}$$

for a certain constant $Z_{n,\beta}$. Thus, $\beta = 2$ recovers the CUE (c.f. (1.1.4)) and theorem 2.2.10 is the $\beta = 2$ particular case of their more general result. They also prove that the equivalent family of random variables, where the real part of the logarithm of the characteristic polynomial is replaced by the imaginary part, is also tight.

Hence only the identification of the distribution of the fluctuating term of conjecture 2.2.2 remains. Recall that this is conjectured to be a sum of two independent Gumbel random variables. The only results to date in this direction relate to yet another model, and are due to Remy [137]. We give Remy's results after outlining the branching structure intrinsic to the proofs of theorems 2.2.8, 2.2.9, and 2.2.10. As emphasised previously, identifying an approximate branching structure is inherent to proving the results as it permits a connection with the exact branching random walk model.

Branching structure for $\log |P_N(A,\theta)|$

Following the progress made by identifying the approximate branching structure in a model of the Riemann zeta function by Arguin et al. [7], a crucial part of proving the various results towards conjecture 2.2.2 is the identification of a quasi-tree structure within the logarithm of the characteristic polynomial.

Recall that the first result towards a proof of conjecture 2.2.2 was the result of Arguin, Belius, and Bourgade - theorem 2.2.8. Following a similar procedure to that outlined in (2.2.56), they show that the tree-like structure emerges from a multiscale decomposition. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(\theta) = \log |1 - e^{i\theta}|, \tag{2.2.75}$$

and observe that the Fourier series of f is

$$-\sum_{j=1}^{\infty} \frac{\mathrm{Re}(e^{-ij\theta})}{j}. \tag{2.2.76}$$

This means that

$$\log |P_N(A,\theta)| = -\sum_{j=1}^{\infty} \frac{\mathrm{Re}(\mathrm{Tr}(A^j) e^{-ij\theta})}{j}, \tag{2.2.77}$$

where $\theta \in [0, 2\pi)$. Compare this to (2.2.12), where such a decomposition was used to justify the log-correlated nature of $\log |P_N(A,\theta)|$. Note that (2.2.77) is ill-defined when θ is an eigenphase of A. To resolve this, let the set of eigenphases of A be $\mathcal{E}_A = \{\theta_1, \ldots, \theta_N\}$. Then for $\theta \in \mathcal{E}_A$, define both the left and the right hand side of (2.2.77) to be $-\infty$.

We recall two important properties of the traces of powers of unitary matrices. Firstly, due to orthogonality of characters of the unitary group (equivalently, the rotational invariance of the Haar measure), traces are uncorrelated (see for example [64]):

$$\mathbb{E}_{A\in U(N)}\left[\operatorname{Tr}(A^j)\overline{\operatorname{Tr}(A^k)}\right] = \delta_{j,k}k, \tag{2.2.78}$$

provided that $k \leq N$. Secondly, recall the result of Diaconis and Shahshahani, theorem 2.2.5, concerning the convergence of powers of traces,

$$\left(\frac{\operatorname{Tr}(A^j)}{\sqrt{j}}\right)_{j\geq 1} \overset{N\to\infty}{\longrightarrow} \left(\mathcal{N}_j^{\mathbb{C}}\right)_{j\geq 1}. \tag{2.2.79}$$

Further, the speed of the convergence is superexponential [106].

Together, these mean that Arguin et al. may truncate the sum in (2.2.77) at N, just gaining an error at the level of $O(1)$. Further, one can decompose the now finite sum as

$$X_n(\theta) = \sum_{m=1}^{n} W_m(\theta) = -\sum_{m=1}^{n}\sum_{e^{m-1}\leq j<e^m} \frac{\operatorname{Re}(\operatorname{Tr}(A^j)e^{-ij\theta})}{j}, \quad \text{for } n \in \{0,\dots,\log N\}, \tag{2.2.80}$$

which should be reminiscent of the procedure used by Arguin et al. [7] for the model of the Riemann zeta function[17]. It should therefore not be surprising that (2.2.80) can be interpreted as a branching random walk. The increments $W_m(\theta)$ are, by the above discussion, uncorrelated and have variance approximately $\frac{1}{2}$. If one takes two points θ_1, θ_2, then the covariance of the increments is calculated (see [5]) to be

$$\mathbb{E}\left[W_m(\theta_1)W_m(\theta_2)\right] = \frac{1}{2}\sum_{e^{m-1}\leq j<e^m} \frac{\cos(j|\theta_1-\theta_2|)}{j} = \begin{cases} \frac{1}{2} + O\left(e^{m-\theta_1\wedge\theta_2}\right) & \text{if } m \leq \theta_1\wedge\theta_2 \\ O\left(e^{-2(m-\theta_1\wedge\theta_2)}\right) & \text{if } m > \theta_1\wedge\theta_2, \end{cases} \tag{2.2.81}$$

where

$$\theta_1\wedge\theta_2 := -\log||\theta_1-\theta_2|| \tag{2.2.82}$$
$$= -\log(\min\{|\theta_1-\theta_2|, 2\pi-|\theta_1-\theta_2|\}) \tag{2.2.83}$$

(c.f. the definition (2.2.60)). As is now usual, one sees that the increments are essentially perfectly correlated for levels m before $\theta_1\wedge\theta_2$, and then almost perfectly uncorrelated thereafter.

It turns out that it is first easier to work with a slightly different decomposition, which is more in-keeping with (2.2.59) and (2.2.56).

$$X_{(1-\delta)\log N}(\theta) = \sum_{m=1}^{K-1} Y_m(\theta) \tag{2.2.84}$$

for some large integer K, and $\delta = K^{-1}$, and where

$$Y_m(\theta) := \sum_{\frac{m-1}{K}\log N<j\leq\frac{m}{K}\log N} W_j(\theta). \tag{2.2.85}$$

[17]Recall that we compare N with height $\log T$ up the critical line. In (2.2.59), $\log T = 2^n$ so the number of levels (equivalently, the length of the walk $X_n(h)$) is on the order of $\log\log T$, which corresponds to the $\log N$ levels in (2.2.80).

Therefore

$$X_{(1-\delta)\log N}(\theta) = -\sum_{m=1}^{K-1} \sum_{N^{\frac{m-1}{K}} < j \leq N^{\frac{m}{K}}} \frac{\mathrm{Re}(\mathrm{Tr}(A^j)e^{-ij\theta})}{j}. \tag{2.2.86}$$

Then, by (2.2.81), the $Y_m(\theta)$ are centred with variance approximately $\frac{1}{2K}\log N$. As usual, the upper bound on $X_{(1-\delta)\log N}$ follows from an adapted union bound. The analysis of this approximate branching random walk uses the same recipe[18] as that used by Arguin et al. [7] for the model of $\zeta(s)$, see the general method of proof outlined previously.

The lower bound requires more work and in particular a *truncated second moment* argument is required (a modification of the Paley-Zygmund inequality described previously). This controls the second moment, which as discussed following (2.2.70), would otherwise dominate the square of the first moment. Combining the upper and lower bound delivers the claimed leading order.

Swiftly following the result of Arguin, Belius, and Bourgade, conjecture 2.2.2 was verified to subleading order by Paquette and Zeitouni. The key improvement by Paquette and Zeitouni which permitted them to upgrade the result of Arguin et al. is a careful comparison between the field

$$\log|\det(I - zA)| \tag{2.2.87}$$

and a centred Gaussian field inside the unit circle on optimal scales.

Once more, the result is achieved using upper and lower bounds. For the lower bound, Paquette and Zeitouni study the field $\log|\det(I - zA)|$ *inside* the unit circle, i.e. $|z| < 1$. It turns out that this is particularly convenient since there $\log|\det(I - zA)|$ is harmonic, so almost surely

$$\sup_{|z|<1} \log|\det(I - zA)| = \max_{z\in\{z\in\mathbb{C}:|z|=1\}} \log|\det(I - zA)|. \tag{2.2.88}$$

This means that one can instead analyse the behaviour of the field on the interior of the disk and any value there will provide a lower bound for the maximum on the boundary. Conversely, note that

$$\log|\det(I - zA)| = \sum_{j=1}^{N} \log\left|1 - ze^{i\theta_j}\right|, \tag{2.2.89}$$

where the eigenvalues of A are $e^{i\theta_1}, \ldots, e^{i\theta_N}$. Then, one can construct the following upper bound. For any $M > 0$, Paquette and Zeitouni show that there exists an $\tilde{N}(M)$ sufficiently large such that for all integers $N > \tilde{N}(M)$ and any $\{\theta_1, \ldots, \theta_N\} \in [0, 2\pi)$ one has

$$\max_{|z|=1} \sum_{j=1}^{N} \log\left|1 - e^{i(\theta-\theta_j)}\right| \leq \max_{|z|=1-\frac{M}{N}} \sum_{j=1}^{N} \log\left|1 - e^{i(\theta-\theta_j)}\right| + M. \tag{2.2.90}$$

One is then motivated to understand the behaviour of the field $\log|\det(I - zA)|$ at points equally spaced around the unit circle, shifted to lie just inside:

$$\left\{(1 - \tfrac{M}{N})e^{2\pi i\frac{j}{N}}\right\}_{j=1}^{N}, \tag{2.2.91}$$

and M can be set to 2, see [132]. As has been demonstrated multiple times in this review, any attempt to complete this calculation under the assumption of independence will be doomed to fail. Inspired by

[18] Though with significant additional technicalities, in particular Riemann–Hilbert analysis is required to justify the truncated sum described above.

work of Bramson [29] and Arguin, Belius, and Harper [7], Paquette and Zeitouni instead work with a barrier event that constrains the particles to lie below some level. Another key aspect is making rigorous the comparison between $\log|\det(I-zA)|$ and $G(z)$, a centred, real-valued, log-correlated Gaussian field. After this is established, they then directly appeal to the body of literature concerning the maximum of log-correlated Gaussian fields, and in particular branching random walks and branching Brownian motion.

Finally, the most recent improvement towards fully establishing conjecture 2.2.2 is due to Chhaibi, Madaule, and Najnudel [35]. They proved that the conjectured family of random variables is tight. Once more, they use a branching structure and a truncated second moment method akin to Arguin, Belius, and Bourgade, and Paquette and Zeitouni. The key addition is the use of the theory of *orthogonal polynomials on the unit circle* (OPUC) and Verblunsky coefficients, see Szegö [151].

If $\mathbb{D}$ is the unit circle in the complex plane, and μ is a probability measure on $\mathbb{D}$, then one can create a sequence of monic polynomials $\{\Phi_k(z), k = 0, 1, \dots\}$ which are orthogonal with respect to μ by applying the Gram-Schmidt orthogonalisation procedure to the sequence $\{z^n : n = 0, 1, \dots\}$. These polynomials can be generated using the Szegö recurrence relation,

$$\Phi_{k+1}(z) = z\Phi_k(z) - \overline{\alpha_k}\Phi_k^*(z), \tag{2.2.92}$$

where

$$\Phi_k^*(z) = z^k\overline{\Phi_k\left(\overline{z}^{-1}\right)} \tag{2.2.93}$$

and $\Phi_0(z) = 1$. The $*$ operator reverses the order of the coefficients of the polynomial and the numbers α_k are known as *Verblunsky coefficients*. Chhaibi et al. note that one can use the OPUC to understand the behaviour of the characteristic polynomials through the following lemma (which is a particular case of lemma 2.1 in Chhaibi et al. [35]).

Lemma 2.2.11 (Chhaibi, Madaule, Najnudel [35]). *The following family of random variables is tight,*

$$\left(\sup_{\theta\in[0,2\pi)\setminus\{\theta_1,\dots,\theta_N\}} \log|P_N(A,\theta)| - \sup_{\theta\in[0,2\pi)} \log|\Phi_{N-1}^*(e^{i\theta})|\right)_{N\geq 1}.$$

Hence, one deduces that proving theorem 2.2.10 is equivalent to proving that the family of random variables

$$\left(\max_{\theta\in[0,2\pi)} \log|\Phi_N^*(e^{i\theta})| - \left(\log N - \frac{3}{4}\log\log N\right)\right)_{N\geq 2} \tag{2.2.94}$$

is tight. Further, one can also express $\log\Phi_N^*(e^{i\theta})$ as a sum of logarithms of a function of Verblunsky coefficients and continuous real functions called Prüfer phases. This introduces a martingale structure which is particularly useful when determining the extreme values of polynomials $(\Phi_k^*)_{k\geq 0}$. With some careful construction, one can then define a new field

$$Z_k(\theta) := \sum_{j=0}^{k-1} \frac{X_j^{\mathbb{C}} e^{i\psi_j(\theta)}}{\sqrt{j+1}} \tag{2.2.95}$$

for $\theta \in \mathbb{R}$, where $X_j^{\mathbb{C}}$ is a complex Gaussian of variance 1, and $\psi_j(\theta)$ are so-called Prüfer phases. This reduces the problem once more to now just studying this new field $Z_k(\theta)$ as they further show that the family of random variables

$$\left(\sup_{\theta\in[0,2\pi)} \left|\log\Phi_k^*(e^{i\theta}) - Z_k(\theta)\right|\right)_{k\geq 0} \tag{2.2.96}$$

is tight. Thus the final goal is to show that

$$\sup_{\theta\in[0,2\pi)} \mathrm{Re}(Z_N(\theta)) = \log N - \frac{3}{4}\log\log N + O(1), \tag{2.2.97}$$

where $O(1)$ is a tight family of random variables. Whilst highly technical, in broad strokes the proof, as all previous proofs in the section have done, follows from an upper bound by the first moment, and a lower bound using the second moment.

Given that Chhaibi et al. have shown conjecture 2.2.2 up to tightness, all that remains is to identify the distribution of the fluctuating term (conjectured to be the sum of two independent Gumbel random variables). To date, the only process towards this goal is due to Remy [137] who considers a related model.

Identification of the Gumbel random variables

A recent paper of Remy [137] establishes the sum of two independent Gumbel random variables in an analogous problem. We will discuss this paper shortly, but first we take a slight detour to understand further the connection between random matrix theory and Gaussian multiplicative chaos (GMC). Recall that GMC measures were introduced in chapter 1, section 1.1.4. In particular we focus on the work of Webb [155] and Nikula, Saksman, and Webb [128].

In two successive works, it was determined that for $\alpha \in (-1/2, 2)$ and as $N \to \infty$,

$$\frac{|P_N(A,\theta)|^\alpha}{\mathbb{E}[|P_N(A,\theta)|^\alpha]}\frac{d\theta}{2\pi} \xrightarrow{\text{law}} e^{\alpha X(\theta) - \frac{\alpha^2}{2}\mathbb{E}[X(\theta)^2]}\frac{d\theta}{2\pi}. \tag{2.2.98}$$

Thus, if one considers the left hand side as a sequence of measures on the unit circle, by (2.2.98) they converge in law to the GMC measure found on the right hand side of the statement. Recall that we write $X(\theta)$ for a centred and logarithmically correlated Gaussian field, where explicitly[19]

$$\mathbb{E}[X(\theta_1)X(\theta_2)] = -\frac{1}{2}\log|e^{i\theta_1} - e^{i\theta_2}|. \tag{2.2.99}$$

We here emphasise the regions in which one can select the parameter α (c.f. (1.1.59) and (1.1.58), and the surrounding discussion). It was first shown by Webb [155] that the convergence (2.2.98) is true in the L^2-phase, which corresponds to taking $\alpha \in \left(-\frac{1}{2}, \sqrt{2}\right)$. In fact, Webb proves a more general version allowing for twists of the characteristic polynomial by powers of the exponential of its argument. In a subsequent work due to Nikula, Saksman, and Webb [128], the convergence (2.2.98) was extended to include the L^1-phase which is equivalent to taking $\alpha \in [\sqrt{2}, 2)$. Motivated by the theory of multiplicative chaos, it is conjectured that the limiting object for $\alpha > 2$ will be zero, thus the interesting behaviour has now been categorised.

Remy also works with the field $X(\theta)$, though with a different normalisation in the covariance. For convenience, we have translated Remy's result to be consistent with the normalisation used throughout this chapter. Define

$$Y_\alpha = \frac{1}{2\pi}\int_0^{2\pi} e^{\alpha X(\theta)} d\theta, \tag{2.2.100}$$

for $\alpha \in (0, 2)$. As usual, Y_α is rigorously defined as a limit using an appropriate cut-off $X_\varepsilon(\theta)$ of $X(\theta)$,

$$e^{\alpha X(\theta)} := \lim_{\varepsilon\to 0} e^{\alpha X_\varepsilon(\theta) - \frac{\alpha}{2}\mathbb{E}[X_\varepsilon(\theta)^2]} \tag{2.2.101}$$

[19] Whilst the normalisation constant of 1/2 appearing here is not standard within the log-correlated literature, its role is to mimic the random matrix setting. It does not affect the range of validity of the measure.

(see Remy [137] or section 1.1.4 for details). It is instructive to compare this to the definition of $g_N(\beta; A)$ (2.2.27). The main theorem of Remy proves a conjecture of Fyodorov and Bouchaud [78].

Theorem 2.2.12. *Take $\alpha \in (0, 2)$. For all $\rho \in \mathbb{R}$ such that $\rho\alpha^2 < 4$, one has*

$$\mathbb{E}[Y_\alpha^\rho] = \frac{\Gamma\left(1 - \frac{\rho\alpha^2}{4}\right)}{\Gamma\left(1 - \frac{\alpha^2}{4}\right)^\rho}. \tag{2.2.102}$$

Consider the statement of theorem 2.2.12 alongside the results of Webb and Nikula et al., and set $\alpha = 2\beta$ and $\rho = k$. With this specialisation, (2.2.98) becomes

$$\frac{|P_N(A, \theta)|^{2\beta}}{\mathbb{E}[|P_N(A, \theta)|^{2\beta}]} \frac{d\theta}{2\pi} \to e^{2\beta X(\theta) - 2\beta^2 \mathbb{E}[X(\theta)^2]} \frac{d\theta}{2\pi}, \tag{2.2.103}$$

for $2\beta \in (-1/2, 2)$, as $N \to \infty$. Hence, by the result of Remy, one expects for $k\beta^2 < 1$, and $2\beta \in (0, 2)$,

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} \frac{|P_N(A, \theta)|^{2\beta}}{\mathbb{E}[|P_N(A, \theta)|^{2\beta}]} d\theta\right)^k\right] \to \mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} e^{2\beta X(\theta) - 2\beta^2 \mathbb{E}[X(\theta)^2]} d\theta\right)^k\right] \tag{2.2.104}$$

$$= \frac{\Gamma(1 - k\beta^2)}{\Gamma(1 - \beta^2)^k}. \tag{2.2.105}$$

Given the rotational invariance of the Haar measure and the result of Keating and Snaith 2.1.1, one has that (assuming that the convergence (2.2.105) holds),

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} |P_N(A, \theta)|^{2\beta} d\theta\right)^k\right] \sim \left(\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k \frac{\Gamma(1 - k\beta^2)}{\Gamma^k(1 - \beta^2}) N^{k\beta^2}, \tag{2.2.106}$$

which is precisely the first regime of conjecture 2.2.4.

Additionally, Remy considers the ‘critical’ GMC at $\alpha = 2$. In this case, the measure is denoted $-X(\theta)e^{X(\theta)}d\theta$ and is found via

$$-X(\theta)e^{2X(\theta)}d\theta := -\lim_{\varepsilon \to 0} \left(X_\varepsilon(\theta) - 2\mathbb{E}[X_\varepsilon(\theta)^2]\right) e^{2X_\varepsilon(\theta) - 2\mathbb{E}[X_\varepsilon(\theta)^2]} d\theta, \tag{2.2.107}$$

again for a suitable cut-off X_ε. Such a construction gives a non-trivial random positive measure, see [67, 68, 133]. Now define

$$Y' := -\int_0^{2\pi} X(\theta)e^{2X(\theta)}d\theta. \tag{2.2.108}$$

It was shown by Aru, Powell, and Sepúlveda [8] that Y' is related to the limit as $\alpha \to 2$ of Y_α,

$$Y' = \lim_{\alpha \to 2} \frac{Y_\alpha}{2 - \alpha} \tag{2.2.109}$$

in probability. One can hence deduce (see [137]) the density $f_{Y'}$ of Y',

$$f_{Y'}(y) = \begin{cases} y^{-2}e^{-y^{-1}} & \text{if } y \geq 0 \\ 0 & \text{if } y < 0, \end{cases} \tag{2.2.110}$$

so $\log Y'$ has a standard Gumbel law. Finally, recent results (see [21, 65]), have shown that for suitable cut-offs X_ε,

$$\max_{\theta \in [0, 2\pi)} X_\varepsilon(\theta) - \log\frac{1}{\varepsilon} + \frac{3}{4}\log\log\frac{1}{\varepsilon} \to \mathscr{G} + \log Y' + C, \tag{2.2.111}$$

where $\mathscr{G}$ is a Gumbel random variable independent from Y', and C is a constant depending on the cut-off chosen.

Hence, combining (2.2.111), (2.2.110), with N corresponding to $1/\varepsilon$ and $\log|P_N(A,\theta)|$ to $X_\varepsilon(\theta)$, this analogy would imply that

$$\max_{\theta\in[0,2\pi)} \log|P_N(A,\theta)| - \log N + \frac{3}{4}\log\log N \to \mathscr{G}_1 + \mathscr{G}_2 + C, \tag{2.2.112}$$

thus adding further weight to conjecture 2.2.2 of Fyodorov and Keating. This approach also suggests that it may be easier to instead establish

$$\max_{\theta\in[0,2\pi)} \log|P_N(A,\theta)| - \log N + \frac{3}{4}\log\log N \to \mathscr{G}_1 + \log Y' + C. \tag{2.2.113}$$

Progress towards conjecture 2.2.3

Finally, we give a brief review of the progress to date towards resolving the conjecture of Fyodorov and Keating concerning the maximum of $\log|\zeta(1/2+it)|$ in short intervals. Since this document primarily concerns random matrix results, we here focus just on the key theorems and identifying the source of the branching structure.

Recall that the proof technique of identifying the branching structure originated with Harper's model of $\log|\zeta(1/2+it)|$, see [92]. We have also now shown that there has been significant progress towards a proof of conjecture 2.2.2. For this final part of the section, we give the various results towards a proof of conjecture 2.2.3, which recall has the form

$$\max_{h\in[0,2\pi)} \log|\zeta(\tfrac{1}{2}+i(\tau+h))| = \log\log T - \frac{3}{4}\log\log\log T + O_{\mathbb{P}}(1), \tag{2.2.114}$$

where $\tau \sim U[T,2T]$ and the final term $O_{\mathbb{P}}(1)$ represents a term bounded in probability as T grows large.

The first result of Najnudel [127] proves the conjecture to leading order under the Riemann hypothesis (and also proves an analogous result for the imaginary part of $\log\zeta(1/2+it)$).

Theorem 2.2.13 (Najnudel [127])**.** *Take $\varepsilon > 0$. Then, under the Riemann hypothesis,*

$$\frac{1}{T}meas.\left\{T \le t \le 2T : (1-\varepsilon)\log\log T < \max_{|t-h|\le 1} \log|\zeta(\tfrac{1}{2}+ih)| < (1+\varepsilon)\log\log T\right\} \to 1 \tag{2.2.115}$$

as $T\to\infty$.

For Najnudel, the random walk structure originates in the Euler product. In particular, note that

$$\log\zeta(s+it) = \sum_p \sum_{k=1}^{\infty} \frac{p^{-k(s+it)}}{k} = \sum_{n\ge1} \frac{l(n)}{n^{s+it}} \tag{2.2.116}$$

for $\mathrm{Re}(s) > 1$ and $t\in\mathbb{R}$, and where $l(n) = 1/k$ if $n = p^k$ and 0 otherwise.

Common to all the proofs is that the upper bound for the leading order, i.e. demonstrating that $\max_{|t-h|\le1}\log|\zeta(1/2+ih)| < (1+\varepsilon)\log\log T$ (see for example section 2 of [6]), is straightforward.

For the lower bound, Najnudel first proves the following. For any integrable function $\phi:\mathbb{R}\to\mathbb{R}$ such that its Fourier transform $\hat{\phi}$ is compactly supported, then, for $\mathrm{Re}(s) > 1$,

$$L_\phi(s) := \int_{-\infty}^{\infty} \log\zeta(s+it)\phi(t)dt \tag{2.2.117}$$

is well defined and

$$L_\phi(s) = \sum_{n\geq 1} \frac{l(n)}{n^s}\hat{\phi}(\log n), \tag{2.2.118}$$

which can be continued to the whole of the complex plane. One then wishes to extend such a result inside the critical strip. However, difficulties arise associated with the zeros of $\zeta(s)$ and the pole at $s = 1$. Assuming the Riemann hypothesis clearly gives an exact awareness of the location of the zeros of $\zeta(s)$ inside the strip. Therefore, for $\sigma \in [1/2, 1)$, $\tau \in \mathbb{R}$, $H > 0$, Najnudel is able to prove, under the Riemann hypothesis and provided ϕ satisfies the conditions above and additionally is dominated by any negative power at infinity, that

$$\int_{-\infty}^{\infty} \log\zeta\left(\sigma + i\left(\tau + \tfrac{t}{H}\right)\right)\phi(t)dt = \sum_{n\geq 1}\frac{l(n)}{n^{\sigma+i\tau}}\hat{\phi}\left(\tfrac{\log n}{H}\right) + O_\phi\left(1 + \tfrac{e^{O_\phi(H)}}{1+|\tau|}\right). \tag{2.2.119}$$

If one chooses, as Najnudel does, ϕ satisfying the conditions of (2.2.119), with additionally $\psi := \hat{\phi}$ taking values in $[0, 1]$, even and equal to 1 at 0, then for $H > 1$ and $\tau \in \mathbb{R}$ one can define

$$\Lambda_\psi(\tau, H) := \sum_{n\geq 1}\frac{l(n)}{n^{\frac{1}{2}+i\tau}}\psi\left(\tfrac{\log n}{n}\right). \tag{2.2.120}$$

Najnudel then proves (propositions 5.1, 5.2 in [127]) that, without too big an error term, and by making careful choices for τ, H in (2.2.120) (see the statement of proposition 5.2 in [127]), the maximum of $\mathrm{Re}(\kappa\Lambda_\psi(\tau, H))$ provides a lower bound on the maximum of $\mathrm{Re}(\log\zeta(1/2 + it))$ in the relevant short intervals, where $\kappa \in \{0, 1, \ldots, H - 1\}$. Indeed, one can also show (see section 6 of [127]) that, after applying a multiscale-type decomposition and using the random approximation to terms of the form p^{it} (see Harper's model, discussed at the start of section 2.2.3), that $\Lambda_\psi(\tau, H)$ has an *approximate branching structure*.

Soon after Najnudel's result, Arguin, Belius, Bourgade, Radziwiłł and Soundararajan were able to remove the assumption of the Riemann hypothesis.

Theorem 2.2.14 (Arguin, Belius, Bourgade, Radziwiłł, Soundararajan [6]). *For any $\varepsilon > 0$, as $T \to \infty$, we have*

$$\frac{1}{T}meas\left\{T \leq t \leq 2T : (1-\varepsilon)\log\log T < \max_{|t-h|\leq 1}\log|\zeta(\tfrac{1}{2} + ih)| < (1+\varepsilon)\log\log T\right\} \to 1. \tag{2.2.121}$$

Arguin et al. are able to bypass assuming the Riemann hypothesis by examining $\zeta(\sigma + i(t + h))$ for σ slightly away from 1/2 and proving results for *most*, not *all* t. As mentioned above, the upper bound is relatively simple to prove. The lower bound requires more work. Firstly, Arguin et al. establish that large values just off the critical line imply large values lying on the critical line, thus permitting them to work slightly to the right of the line $1/2 + it$. They then construct a 'mollifier' (i.e. a function $M(s)$ so that, just to the right of the critical line, $M(s)\zeta(s) \approx 1$). This allows them to show that for almost all $t \in [T, 2T]$, one instead may work with a significantly shorter Dirichlet polynomial

$$\sum_{p\leq X}\mathrm{Re}\left(\sum_{p\leq X}\frac{1}{p^{\frac{1}{2}+\varepsilon+ih}}\right) \tag{2.2.122}$$

for an X *much smaller* than T. In particular they show that a large value of the maximum of (2.2.122) (over $|t - h| \leq 1$) implies a large value of $\max_{|t-h|\leq 1}\log|\zeta(1/2 + ih)|$. Similarly to Najnudel [127], it is within (2.2.122) that one may find an approximate branching structure. The techniques for handling branching structures described at the start of this section are again key to proving the required lower bound.

Most recently, Harper has provided a nearly sharp upper bound.

Theorem 2.2.15 (Harper [94]). *For any real function $g(T)$ tending to infinity with T, we have*

$$\max_{|h|\leq 1/2} \log|\zeta(\tfrac{1}{2}+i(t+h))| \leq \log\log T - \frac{3}{4}\log\log\log T + \frac{3}{2}\log\log\log\log T + g(T), \tag{2.2.123}$$

for a set of $t \in [T, 2T]$ with measure $(1+o(1))T$.

The main difference in the proof of Harper's result is how $\zeta(s)$ is approximated by Dirichlet polynomials. Harper approximates the zeta function in mean square by the product of two Dirichlet polynomials: one over smooth numbers and one over rough numbers. Typically, one discards the latter in favour of short Dirichlet series, see for example (2.2.122). However, Harper shows that in order to achieve the desired subleading term, one must work with the full range and use the contribution from the large primes to one's advantage. Doing so provides the claimed bound. This is similar in spirit to the work of Conrey and Keating [37–41] on the techniques for understanding higher moments of $\zeta(1/2+it)$, discussed in section 2.1.2.

This concludes the review of the recent progress towards proving both conjectures 2.2.2 and conjecture 2.2.3. As has been demonstrated, the identification throughout of an approximate branching structure is essential. Conjecture 2.2.4 is also of substantial research interest, and is the topic of chapter 3. In chapter 6 we focus on the approximate branching structure present within $\log|P_N(A,\theta)|$. Just as Arguin et al. [7] used the randomized model of $\zeta(s)$ to prove a version of conjecture 2.2.3, there we prove a regime of conjecture 2.2.4 adapted to this model. Additionally we outline a method for proving such an adapted conjecture in full generality.

2.2.4 Maxima of orthogonal and symplectic polynomials

For the final part of this chapter, we consider the maximum size of orthogonal and symplectic characteristic polynomials. A natural extension to the preceding discussion would be to determine the size of the maximum of $\log|P_{G(N)}(A,\theta)|$ for $A \in SO(2N)$ or $A \in Sp(2N)$, and conclude whether or not it differs from the unitary case. Here we provide evidence towards the conclusion that the maximum in either situation does not differ, at least to leading order, to that of the unitary case.

It is not initially obvious that this should be the case. Recall that the moments of symplectic and orthogonal characteristic polynomials at the symmetry point (c.f. theorem 2.1.7 and theorem 2.1.8) behave differently from unitary. This may have implied that the leading order of the maximum differs from conjecture 2.2.2. However, the work of Keating and Odgers (c.f. theorem 2.1.10) shows that, away from the symmetry point, the statistics revert back to those of unitary matrices. It is this behaviour which seems to win out; our argument described below suggests that, to leading order, the maximum of $\log|P_{G(N)}(A,\theta)|$ is $\log N$ regardless of the choice of $G(N) \in \{U(N), Sp(2N), SO(2N)\}$. The work described here came out of conversations with Prof. Paul Bourgade during a research visit to CIMS.

We begin with the orthogonal case. Firstly, we state central limit theorem for $\log|P_{SO(2N)}(A,0)|$, which is work of Keating and Snaith [115].

$$\frac{\log|P_{SO(2N)}(A,0)| + \frac{1}{2}\log N}{\sqrt{\log N}} \sim \mathcal{N}(0,1). \tag{2.2.124}$$

To find the correct level of the maximum for $SO(2N)$, we need to estimate how (2.2.124) changes as θ varies close to the symmetry point. Keating and Odgers proved that (corollary 3 in [112]) for

$\frac{1}{N} \ll \theta \ll 1$,

$$\mathbb{E}[\log |P_{SO(2N)}(A,\theta)|] \sim \frac{1}{2}\log(1-e^{2i\theta}), \tag{2.2.125}$$

hence for this range of θ, the average value of the logarithm behaves like $\frac{1}{2}\log\theta$. Hence, one proposes that

$$\frac{\log |P_{SO(2N)}(A,\theta)| - \frac{1}{2}\log\theta}{\sqrt{\sigma^2 \log N}} \sim \mathcal{N}(0,1). \tag{2.2.126}$$

To determine the value of σ^2, we use the second moment of the characteristic polynomial. Recall that Keating and Odgers show (see (2.1.56)) for $\frac{1}{N} \ll \theta \ll 1$

$$\mathbb{E}[|P_{SO(2N)}(A,\theta)|^2] \sim N. \tag{2.2.127}$$

Alternatively, we could compute

$$\mathbb{E}[|P_{SO(2N)}(A,\theta)|^2] = \mathbb{E}[\exp(2\log |P_{SO(2N)}(A,\theta)|)] \tag{2.2.128}$$

$$\sim \mathbb{E}[\exp(2Z)] \tag{2.2.129}$$

$$= \exp(2\sigma^2 \log N + \log\theta). \tag{2.2.130}$$

where Z is a Gaussian random variable with mean $\frac{1}{2}\log\theta$ and variance $\sigma^2 \log N$, following (2.2.126). Then by comparing (2.2.130) and (2.2.127), we find that

$$\sigma^2 = \frac{1}{2} - \frac{\log\theta}{2\log N}. \tag{2.2.131}$$

So for $\frac{1}{N} \ll \theta \ll 1$, we would expect

$$\log |P_{SO(2N)}(A,\theta)| \sim \mathcal{N}(\tfrac{1}{2}\log\theta, \sigma^2 \log N), \tag{2.2.132}$$

with σ^2 given by (2.2.131). As a sanity check, we consider $\theta = 1/N$. Then (2.2.132) implies, as $N \to \infty$,

$$\log |P_{SO(2N)}(A,0)| \sim \mathcal{N}(-\tfrac{1}{2}\log N, \log N), \tag{2.2.133}$$

as expected by (2.2.124). Additionally, away from the symmetry point, we should instead recover unitary statistics. One sees this by setting $\theta = 1$ in (2.2.132), whereby $\log |P_{SO(2N)}(A,1)$ would instead be modelled by a normal random variable with mean 0 and variance $(1/2)\log N$.

The next task is to determine the correct value of $\alpha \in [0,1]$ such that

$$\max_{\theta\in[0,2\pi)} \log |P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log |P_{SO(2N)}(A,\theta)|] > \alpha \log N \tag{2.2.134}$$

with high probability. By estimating the Gaussian density function Φ, we have

$$\mathbb{P}\left(\frac{\log |P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log |P_{SO(2N)}(A,\theta)|]}{\sqrt{\sigma^2 \log N}} > x\right) \sim e^{-\frac{x^2}{2}}. \tag{2.2.135}$$

where $x = \frac{\alpha\sqrt{\log N}}{\sigma}$. Hence

$$\mathbb{P}(\log |P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log |P_{SO(2N)}(A,\theta)|] > \alpha \log N) \sim e^{-\frac{\alpha^2 \log N}{2\sigma^2}} \tag{2.2.136}$$

$$= N^{-\frac{\alpha^2}{1-\frac{\log\theta}{\log N}}}. \tag{2.2.137}$$

Taking a union bound and exponential scaling, $e^l \le N\theta \le e^{l+1}$ for $1 \le l \le \log N$,

$$\begin{aligned}
\mathbb{P}&\left(\max_{\theta\in[0,2\pi)} \log|P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log|P_{SO(2N)}(A,\theta)|] > \alpha\log N\right) \\
&\le \sum_{\theta} \mathbb{P}(\log|P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log|P_{SO(2N)}(A,\theta)|] > \alpha\log N) && (2.2.138)\\
&= \sum_{l=1}^{\log N} e^l N^{-\frac{\alpha^2}{1-\frac{\log\theta}{\log N}}} && (2.2.139)\\
&= \sum_{l=1}^{\log N} e^l N^{-\frac{\alpha^2}{2-\frac{l}{\log N}}}. && (2.2.140)
\end{aligned}$$

Set

$$f(l) := l - \frac{\alpha^2}{2-\frac{l}{\log N}}\log N. \tag{2.2.141}$$

We want to determine which l maximises $f(l)$, so differentiating gives

$$f'(l) = 1 - \frac{\alpha^2}{\left(2-\frac{l}{\log N}\right)^2}. \tag{2.2.142}$$

Thus $l = (2-\alpha)\log N$ and plugging back into (2.2.140) implies $\alpha = 1$. Hence, one would expect that

$$\frac{\max_{\theta\in[0,2\pi)} \log|P_{SO(2N)}(A,\theta)| - \mathbb{E}[\log|P_{SO(2N)}(A,\theta)|]}{\log N} \overset{N\to\infty}{\longrightarrow} 1. \tag{2.2.143}$$

This means that, to leading order, one would expect that the maximum of $\log|P_{SO(2N)}(A,\theta)|$ matches that of $\log|P_{U(N)}(A,\theta)|$ (c.f. conjecture 2.2.2).

For completeness, we can repeat the calculation for $A \in Sp(2N)$. The central limit theorem, due to Keating and Snaith [115], at the symmetry point is

$$\frac{\log|P_{Sp(2N)}(A,0)| - \frac{1}{2}\log N}{\sqrt{\log N}} \sim \mathcal{N}(0,1). \tag{2.2.144}$$

Keating and Odgers [112] computed the mean value of the logarithm of the characteristic polynomial in this case,

$$\mathbb{E}[\log|P_{Sp(2N)}(A,\theta)|] = -\frac{1}{2}\log(1-e^{2i\theta}), \tag{2.2.145}$$

hence the average value of the logarithm behaves as $-\frac{1}{2}\log\theta$. As above, we use the second moment of the characteristic polynomial to determine the variance. Keating and Odgers show that for this range of θ, the second moment still depends on θ (which differs from the orthogonal case), see (2.1.56),

$$\begin{aligned}
\mathbb{E}[|P_{Sp(2N)}(A,\theta)|^2] &\sim \frac{2N}{|1-e^{-2i\theta}|^2} && (2.2.146)\\
&\sim \frac{N}{\theta^2}. && (2.2.147)
\end{aligned}$$

Mimicking the calculation for the orthogonal case, we find also that

$$\begin{aligned}
\mathbb{E}[|P_{Sp(2N)}(A,\theta)|^2] &= \mathbb{E}[\exp(2\log|P_{Sp(2N)}(A,\theta)|)] && (2.2.148)\\
&\sim \mathbb{E}[\exp(2Z)] && (2.2.149)\\
&= \exp(2\sigma^2\log N - \log\theta) && (2.2.150)
\end{aligned}$$

where Z is a Gaussian random variable with mean $-\frac{1}{2}\log\theta$ and variance $\sigma^2 \log N$. Comparing (2.2.150) to (2.2.147), we again find that

$$\sigma^2 = \frac{1}{2} - \frac{\log\theta}{2\log N}. \tag{2.2.151}$$

So for $\frac{1}{N} \ll \theta \ll 1$, we would expect

$$\log|P_{Sp(2N)}(A,\theta)| \sim \mathcal{N}(-\tfrac{1}{2}\log\theta, \sigma^2\log N). \tag{2.2.152}$$

This matches the usual central limit theorem at the symmetry point when $\theta \sim 1/N$, see (2.2.144).

The calculation hence now proceeds identically to the orthogonal case, yielding the prediction that

$$\frac{\max_{\theta\in[0,2\pi)}\log|P_{Sp(2N)}(A,\theta)| - \mathbb{E}[\log|P_{Sp(2N)}(A,\theta)|]}{\log N} \overset{N\to\infty}{\longrightarrow} 1. \tag{2.2.153}$$

2.3 Conclusion

In section 2.1, we focussed on the connection between random matrix theory and number theory, in particular using moments of characteristic polynomials to model various moments of L-functions. This narrative provides context for our results in chapters 4 and 7.

Further, this theme continued in section 2.2; the maxima conjectures of Fyodorov and Keating were formulated both for unitary characteristic polynomials and for the Riemann zeta function using the analogy developed in section 2.1. Recall conjecture 2.2.4

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_{U(N)}(A,\theta)|^{2\beta}d\theta\right)^k\right] \sim \begin{cases}\left(\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k \frac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)}N^{k\beta^2} & \text{if } k<\frac{1}{\beta^2},\\ \gamma_{k,\beta}N^{k^2\beta^2-k+1} & \text{if } k>\frac{1}{\beta^2},\end{cases} \tag{2.3.1}$$

which was used to justify conjecture 2.2.2. We refer to the left hand side of (2.3.1) as a 'moments of moments'. One could rephrase this in terms of moments of moments of the Riemann zeta function. Conjecture 2.2.4 is the subject of chapter 3, and this generalization to number theoretic moments of moments is discussed there.

We argued in section 2.2.4 that, at least to leading order, conjecture 2.2.2 should extend to symplectic and special orthogonal characteristic polynomials. However, the question of the behaviour of their moments of moments remains unanswered. We show in chapter 4 that their asymptotic behaviour in fact *differs* from the unitary case, and that each compact group has its own unique asymptotic form. This is more in line with results such as theorem 2.1.7 and theorem 2.1.8, in which Keating and Snaith showed that at the symmetry point the moments over each compact group differ.

Recall also from section 2.2.2 that studying moments of moments of unitary characteristic polynomials could lead to fine information about their maximum. We have discussed that the distribution of the fluctuating term in conjecture 2.2.2 remains open, and conjecture 2.2.4 could provide a method for understanding it. To date, only the result of Remy [137] concerning a model of $\log|P_{U(N)}(A,\theta)|$ has clearly identified the sum of independent Gumbel random variables. However, note that that result holds in the regime $k\beta^2 < 1$, which we showed in section 2.2.2 is the easier regime for understanding the moments of moments of $P_{U(N)}(A,\theta)$ due to properties of the associated Fisher-Hartwig singularities. This provides further reasons for being interested in understanding the moments of moments in the regime $k\beta^2 > 1$. This in particular is the subject of chapter 3.

Finally, we emphasise once more the importance of the underlying approximate branching structure of both unitary characteristic polynomials and the Riemann zeta function. This identification proved crucial to the series of impressive results towards complete proofs of conjectures 2.2.2 and 2.2.3. In chapter 6 we continue this idea to moments of moments calculations for models of $\log|P_{G(N)}(A,\theta)|$. We are able to prove an analogy of (2.3.1) for this model, as well as rederiving the results of chapters 3 and 4. An interesting direction for future research would be to explore how to utilise this approximate branching structure to strengthen our results of chapters 3 and 4.

Chapter 3

Unitary moments of moments

The basis for this chapter is the paper '*On the moments of the moments of the characteristic polynomials of random unitary matrices*', which is work of the present author and J. P. Keating [15], published in *Communications in Mathematical Physics*. The project was carried out by the present author, under the advisement of J. P. Keating. We also thank Edva Roditty-Gershon and Scott Harper for helpful discussions.

The material within this chapter largely only differs from [15] in that arguments and calculations have been expanded for clarity. We have included additional relevant literature for context, in particular within section 3.2. Moreover, elements of the introduction have been updated in light of recent results. Any such addenda are the work of the present author. Furthermore, the appendix of [15] has been incorporated in to the text (with remarks inline signifying where this has taken place), except for the section giving explicit examples of the moments of moments (section 6.1 in [15]), which has been relocated to chapter 5.

3.1 Introduction

This chapter concerns certain moments of unitary characteristic polynomials, so throughout we write

$$P_N(A,\theta) \equiv P_{U(N)}(A,\theta) = \det(I - Ae^{-i\theta}), \tag{3.1.1}$$

to denote the characteristic polynomial of an $N \times N$ unitary matrix A on the unit circle in the complex plane. The typical values taken by P_N when A is chosen at random, uniformly with respect to Haar measure on the unitary group $U(N)$ (i.e. from the Circular Unitary Ensemble), have been the subject of extensive study as demonstrated in chapter 2. The moments of P_N and its logarithm were computed in [116] using the Selberg integral, see theorem 2.1.1, and compared with the corresponding moments of the Riemann zeta function, $\zeta(s)$, on its critical line ($\mathrm{Re}(s) = 1/2$) (as discussed in section 2.1.2). Recall that it follows from these calculations that $\log P_N(A,\theta)/\sqrt{\frac{1}{2}\log N}$ satisfies a central limit theorem when $N \to \infty$, in that the real and imaginary parts independently converge to normal random variables with zero mean and unit variance (c.f. theorem 2.1.3). This is true as well without normalising, in a distributional sense [100].

The correlations of $\log|P_N(A,\theta)|$ can be computed using, for example, formulae due to Diaconis and Shahshahani [64], and can be shown to satisfy (see section 2.2.2, (2.2.15))

$$\mathbb{E}_{A\in U(N)}\left[\log|P_N(A,\theta)|\log|P_N(A,\theta+x)|\right]\sim\begin{cases}\frac{1}{2}\log N & \text{if } |x|<<\frac{1}{N}\\ -\frac{1}{2}\log|x| & \text{if } 1>>|x|>>\frac{1}{N}\end{cases} \tag{3.1.2}$$

when $N\to\infty$. (The imaginary part of $\log P_N(A,\theta)$ exhibits similar behaviour.)

As shown throughout chapter 2, the fact that $\log|P_N(A,\theta)|$ behaves like a log-correlated Gaussian random function has stimulated a good deal of interest recently, as it suggests a connection with other similar random fields such as those associated with the branching random walk, branching Brownian Motion, the 2-dimensional Gaussian Free Field, and Liouville quantum gravity. This observation, together with heuristic calculations and numerical experiments (c.f. [80]), motivated the series of conjectures [81, 82] of section 2.2 concerning the maximum of $|P_N(A,\theta)|$ on the unit circle,

$$P_{\max}(A)=\max_{\theta\in[0,2\pi)}|P_N(A,\theta)|. \tag{3.1.3}$$

Recall from chapter 2 that the heuristic calculations described in [82] are based on an analysis of the random variable

$$g_N(\beta;A):=\frac{1}{2\pi}\int_0^{2\pi}|P_N(A,\theta)|^{2\beta}d\theta \tag{3.1.4}$$

which is the 2βth moment of $|P_N(A,\theta)|$ with respect to the uniform measure on the unit circle $\frac{d\theta}{2\pi}$ for a fixed $A\in U(N)$. Specifically, the calculations centre on computing the moments of this random variable with respect to an average over $A\in U(N)$. We refer to these as the *moments of the moments* of $P_N(A,\theta)$, such averages were first defined in chapter 2, see (2.2.8),

$$\mathrm{MoM}_{U(N)}(k,\beta):=\mathbb{E}_{A\in U(N)}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_N(A,\theta)|^{2\beta}d\theta\right)^k\right]. \tag{3.1.5}$$

They will be the main focus of our attention.

We also emphasise that the integrand of $g_N(\beta;A)$, when appropriately normalised,

$$\frac{|P_N(A,\theta)|^{2\beta}}{\mathbb{E}\left[|P_N(A,\theta)|^{2\beta}\right]}\frac{d\theta}{2\pi} \tag{3.1.6}$$

has been the subject of considerable interest as highlighted in chapter 2, section 2.2.3. There we gave the results of Webb and Nikula et al. [128, 155], showing that (3.1.6) converges to a limiting Gaussian multiplicative chaos measure (see [20, 108, 139] and chapter 1, section 1.1.4) for $\beta\in(-\frac{1}{4},1)$ (c.f. [142] for a corresponding result for the Riemann zeta function on the critical line). Importantly, there is expected to be a freezing transition [82] at $\beta=1$, leading to a different regime of behaviour when $\beta>1$ (recall the discussion around (2.2.38)).

One of the main conjectures of [82] is conjecture 2.2.4. This states that when $N\to\infty$

$$\mathrm{MoM}_{U(N)}(k,\beta)\sim\begin{cases}\left(\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k\frac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)}N^{k\beta^2} & \text{for } k<1/\beta^2\\ \gamma_{k,\beta}N^{k^2\beta^2-k+1} & \text{for } k>1/\beta^2\end{cases} \tag{3.1.7}$$

where $\mathcal{G}(s)$ is the Barnes $\mathcal{G}$-function and $\gamma_{k,\beta}$ is an unspecified function of k and β[1]. At the transition point $k=\beta^2$, for $k\geq 2$, one should expect that the moments of moments grow like $N\log N$. Recall from chapter 2, section 2.2.2, that one justification for this conjecture follows from a heuristic calculation of

[1]By $A(N)\sim B(N)$, we mean that $A(N)/B(N)\to 1$ when $N\to\infty$.

the moments of moments when k is an integer [81, 82, 111]. This is based on the following expansion for $k \in \mathbb{N}$

$$\mathrm{MoM}_{U(N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi}\left[\mathbb{E}\prod_{j=1}^{k}|P_N(A,\theta_j)|^{2\beta}\right]d\theta_1\cdots d\theta_k. \tag{3.1.8}$$

As explained in section 2.2.2, the integrand in (3.1.8) can be computed asymptotically when $N \to \infty$ and the θ_js are fixed and distinct using the appropriate Fisher-Hartwig formula [74]. The resulting integrals over the θ_js can then be computed when $k < 1/\beta^2$ using the Selberg integral, leading to the expression in the conjecture (3.1.7) in this range. This expression diverges as k approaches $1/\beta^2$ from below. The reason for this is that when $k \geq 1/\beta^2$, singularities associated with coalescences of the θ_js become important. Developing a precise asymptotic in the range $k \geq 1/\beta^2$ therefore requires a Fisher-Hartwig formula that is valid uniformly as the Fisher-Hartwig singularities coalesce. From this perspective, the regime $k \geq 1/\beta^2$ is the more challenging one. Such a uniform Fisher-Hartwig asymptotic formula has only recently been achieved by Fahs [72], by extending a calculation of Claeys and Krasovsky [36]. We highlight that our work described subsequently was completed prior to Fahs' result, and additionally it gives insight in to the leading order coefficient $\gamma_{k,\beta}$ in (3.1.7). As described in section 3.1.1, the increasing results of Claeys and Krasovsky, and Fahs, are only fine enough to give such information for $k = 1, 2$.

In chapter 2, section 2.2.3, we discussed a closely analogous problem in which $\log|P_N(A,\theta)|$ (c.f. (3.5.2)) is replaced by a random Fourier series with the same correlation structure. Such series can be considered as one-dimensional models of the two-dimensional Gaussian Free Field − the analogue of conjecture (3.1.7), due to Fyodorov and Bouchaud [78], was there shown to have recently been proved in the regime $k < 1/\beta^2$ for all k and β by Remy [137] using ideas from conformal field theory [120], see theorem 2.2.12.

We note that the conjecture described above extends to the other circular ensembles (i.e. to the CβE) [35,80,116] and to the Gaussian ensembles [83–85]. We note as well that there are extensive mathematics and physics literatures on log-correlated Gaussian fields; see, for example chapter 1 section 1.1.4 as well as [65], [84] and [34], and references contained therein. There has been a particular focus on the freezing transition at $\beta = 1$. In the case of uncorrelated Gaussian fields this is well understood; see for example the Random Energy Model discussed in section 1.1.4, and [62, 117]. For log-correlated fields the freezing transition continues to be a focus of research; see, for example, [82, 149] and references therein.

Our focus here will be on the conjecture for the asymptotics of the moments of moments (3.1.7) when $k \in \mathbb{N}$ and $\beta \in \mathbb{N}$. Note that this immediately places us in the regime where $k\beta^2 \geq 1$, and so in the more difficult regime which is dominated by coalescing Fisher-Hartwig singularities, and where progress, at time of the publication of our result, had been limited to the cases of $k = 1, 2$. Here one can exploit connections with representation theory and integrable systems that have not been incorporated in the probabilistic approaches taken previously. Specifically, we shall use three different, but equivalent, exact (rather than asymptotic) expressions for the integrand in (3.1.8). This allows us to circumvent the problems described above associated with coalescing Fisher-Hartwig singularities. We also note that our results include the freezing transition point at $\beta = 1$, see the discussion after (2.2.29) in section 2.2.2.

The first of these expressions, which takes the form of a combinatorial sum and was proved in [45], enables us to compute $\mathrm{MoM}_{U(N)}(k,\beta)$ exactly and explicitly for small values of k and β, when both take values in $\mathbb{N}$. This suggests a refinement of conjecture (3.1.7) in this case:

$$\mathrm{MoM}_{U(N)}(k,\beta) = \mathrm{Poly}_{k^2\beta^2-k+1}(N), \tag{3.1.9}$$

where $\mathrm{Poly}_{k^2\beta^2-k+1}(N)$ is a polynomial in the variable N of degree $k^2\beta^2-k+1$. This obviously implies (3.1.7) in the range $k \geq 1/\beta^2$ for $k, \beta \in \mathbb{N}$. In chapter 5 we give explicit examples of these polynomials. This method can be used to establish that $\mathrm{MoM}_{U(N)}(k, \beta)$ is in general a polynomial in N, but does not straightforwardly determine the order of the polynomial in question.

We then go on to analyse (3.1.9) using two alternative approaches. The first of these uses a second formula for the integrand in (3.1.8) that is based on the representation theory of the unitary group and involves expressing $\mathrm{MoM}_{U(N)}(k, \beta)$ in terms of a sum of semistandard Young tableaux via the theory of symmetric functions (c.f. chapter 1, section 1.1.6). The application of the theory of symmetric functions in this context was developed by Bump and Gamburd [33], who used it to analyse moments of characteristic polynomials, following [116] and [45]. We explore their result further in section 3.2. It allows us to prove that $\mathrm{MoM}_{U(N)}(k, \beta)$ is bounded by a polynomial function of N of degree less than or equal to $k^2\beta^2$ at integer values of k, β, and N. The other approach involves a third formula for the integrand in (3.1.8), which takes the form of a multiple contour integral and which was also proved in [45]. This allows us to compute the large-N asymptotics of $\mathrm{MoM}_{U(N)}(k, \beta)$, using methods developed in [112, 113]. We show in this way that $\mathrm{MoM}_{U(N)}(k, \beta)$ is an analytic function of N that grows like $N^{k^2\beta^2-k+1}$ as $N \to \infty$. This approach allows us to obtain a formula for the leading coefficient of the polynomial in (3.1.9), which corresponds to evaluating the function $\gamma_{k,\beta}$ in (3.1.7) when k and β are both integers. Combining these various results allows us to deduce that $\mathrm{MoM}_{U(N)}(k, \beta)$ is a polynomial in N of order $k^2\beta^2 - k + 1$, thereby proving (3.1.9).

The fact that $\mathrm{MoM}_{U(N)}(k, \beta)$ is a polynomial in the variable N when k and β both take values in $\mathbb{N}$ means that in this case we have an exact formula. This is a consequence of this problem being integrable, as is clear from the analysis based on symmetric functions. From the perspective of asymptotics, it means that we know the complete structure of the asymptotics of $\mathrm{MoM}_{U(N)}(k, \beta)$; that is, we know the general form of all terms in the asymptotic expansion, not just the leading order term.

We emphasize that our main motivation here is to prove (3.1.7), and in particular its refinement (3.1.9), in the regime $k\beta^2 \geq 1$ where, at time of publication, previous approaches had failed in general (i.e. other than when $k = 2$) because they require a general Fisher-Hartwig formula valid as k singularities coalesce. Our approach circumvents this obstacle.

This chapter is structured as follows. In the next section we state some formulae for $\mathrm{MoM}_{U(N)}(k, \beta)$ that can be obtained straightforwardly from expressions already in the literature and formulate our general results as theorems. We also discuss the recent developments due to Fahs [72], which gives an alternative proof of (3.1.7) (though with a non-explicit leading order coefficient). In section 3.2, we explain the calculation involving symmetric functions, and then in section 3.3 we describe the calculation involving multiple integrals. In section 3.5 we discuss some connections between our main result and approaches to analysing rigorously the value distribution of $P_{\max}(A)$, in the context of conjectures 2.2.2 and 2.2.3 made in [81, 82], as well as setting out some thoughts on potential extensions and applications. This includes moments of the Riemann zeta function and other L-functions in short intervals, as well as other random matrix ensembles. Elsewhere, in chapter 5, we calculate $\mathrm{MoM}_{U(N)}(k, \beta)$ for small values of k and β, motivating (3.1.9).

3.1.1 Results for $\mathrm{MoM}_{U(N)}(1, \beta)$ and $\mathrm{MoM}_{U(N)}(2, \beta)$

We first set out in this section some results concerning $\mathrm{MoM}_{U(N)}(k, \beta)$ that can be obtained straightforwardly from calculations in the literature and that prove (3.1.9) when $k = 1$ and $k = 2$. Such results were the best known, prior to our result, at time of publication. At the end of this section, we review the recent progress due to Fahs [72].

The case $k = 1, \beta \in \mathbb{N}$ follows immediately from the moment formula of Keating and Snaith [116] (c.f. theorem 2.1.1, also [17]), and matches with the conjecture (3.1.9). Specifically,

$$\mathrm{MoM}_{U(N)}(1,\beta) = \mathbb{E}[|P_N(A,\theta)|^{2\beta}] = \prod_{0\le i,j\le \beta-1}\left(1+\frac{N}{i+j+1}\right), \tag{3.1.10}$$

which is clearly a polynomial in N of degree β^2. In this case the leading order coefficient can be calculated [116] to be

$$\prod_{j=0}^{\beta-1}\frac{j!}{(j+\beta)!}. \tag{3.1.11}$$

The calculation of the average in (3.1.10) was carried out in [116] using the Weyl integration formula and Selberg's integral, as emphasised in chapter 2, section 2.1.1. Bump and Gamburd [33] later gave an alternative proof using symmetric function theory. In this second approach, the expression (3.1.11) was obtained by counting certain semistandard Young tableaux. We shall see these parallel stories of symmetric function theory and complex analysis continuing for higher values of k, and such calculations are expanded in section 3.2.

A proof of $\mathrm{MoM}_{U(N)}(2,\beta)$ due to Keating et al. [113]

A proof of (3.1.9) when $k=2, \beta\in\mathbb{N}$ follows directly from formulae given in [113]. It differs from the proof given by Claeys and Krasovsky [36] (which we explore below) demonstrating (3.1.7) for all β, but without identifying the polynomial structure when $\beta\in\mathbb{N}$ - see the discussion following their calculation for more details. We outline the calculation of Keating et al. first here.

For $A\in U(N)$, the *secular coefficients* of A, written $\mathrm{Sc}_n(A)$, are the coefficients of its characteristic polynomial

$$\det(I+xA) = \sum_{n=0}^{N}\mathrm{Sc}_n(A)x^n. \tag{3.1.12}$$

The following theorem is proved in [113] (theorem 1.5 in that paper).

Theorem 3.1.1 (Keating et al. [113])**.** *For $A\in U(N)$, define*

$$I_\eta(m;N) := \int_{U(N)}\Big|\sum_{\substack{j_1+\cdots+j_\eta=m\\ 0\le j_1,\ldots,j_\eta\le N}}\mathrm{Sc}_{j_1}(A)\cdots\mathrm{Sc}_{j_\eta}(A)\Big|^2 dA. \tag{3.1.13}$$

If $c=m/N, c\in[0,\eta]$, then $I_\eta(m;N)$ is a polynomial in N and

$$I_\eta(m;N) = \gamma_\eta(c)N^{\eta^2-1}+O_\eta(N^{\eta^2-2}), \tag{3.1.14}$$

where

$$\gamma_\eta(c) = \sum_{0\le l<c}\binom{\eta}{l}^2(c-l)^{(\eta-l)^2+l^2-1}p_{\eta,l}(c-l), \tag{3.1.15}$$

with $p_{\eta,l}(c-l)$ being polynomials in $(c-l)$.

To see how theorem 3.1.1 proves (3.1.9) when $k=2$, $\beta\in\mathbb{N}$, we first make use of the generating series for $I_\eta(m,N)$ given in [113],

$$\sum_{0\le m\le \eta N} I_\eta(m;N)x^m = \int_{U(N)}\det(I-A)^\eta\det(I-A^*x)^\eta dA. \tag{3.1.16}$$

Recall the definition of $\mathrm{MoM}_{U(N)}(2,\beta)$,

$$\mathrm{MoM}_{U(N)}(2,\beta) = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} \mathbb{E}\left[|P_N(A,\theta_1)|^{2\beta}|P_N(A,\theta_2)|^{2\beta}\right] d\theta_1 d\theta_2. \tag{3.1.17}$$

We now show that the integrand of (3.1.17) can be expressed in terms of the generating series for $I_\eta(m,N)$, (3.1.16). Firstly, we make use of the 'functional equation' of the characteristic polynomial of a unitary matrix U,

$$\det(I - Ue^{-i\theta}) = (-1)^N e^{-i\theta N}\det(U)\det(I - U^* e^{i\theta}). \tag{3.1.18}$$

Then, we apply (3.1.18) to the integrand of the right hand side of (3.1.16), with $\eta = 2\beta$, and $x = e^{i(\theta_1-\theta_2)}$,

$$\begin{aligned}\det(I-A)^{2\beta}\det(I-A^*e^{i(\theta_1-\theta_2)})^{2\beta} &= \det(I-A)^{\beta} \\ &\times\left((-1)^N\det(A)\det(I-A^*)\right)^{\beta} \\ &\times\det(I-A^*e^{i(\theta_1-\theta_2)})^{\beta} \\ &\times\left(\frac{(-1)^N e^{-i(\theta_2-\theta_1)N}}{\det(A)}\det(I-Ae^{-i(\theta_1-\theta_2)N})\right)^{\beta} \qquad (3.1.19) \\ &= e^{-i\beta(\theta_2-\theta_1)N}|\det(I-A)|^{2\beta}|\det(I-Ae^{-i(\theta_1-\theta_2)})|^{\beta}. \qquad (3.1.20)\end{aligned}$$

Finally, using that the Haar measure is invariant under rotations, we have that

$$\begin{aligned}\sum_{0\le m\le 2\beta N} I_{2\beta}(m;N)e^{i(\theta_1-\theta_2)m} &= \int_{U(N)}\det(I-A)^{2\beta}\det(I-A^*e^{i(\theta_1-\theta_2)})^{2\beta}dA \qquad (3.1.21)\\ &= e^{-i\beta(\theta_2-\theta_1)N}\int_{U(N)}|\det(I-A)|^{2\beta}|\det(I-Ae^{-i(\theta_1-\theta_2)})|^{\beta}dA \qquad (3.1.22)\\ &= e^{-i\beta(\theta_2-\theta_1)N}\int_{U(N)}|\det(I-Ae^{-i\theta_1})|^{2\beta}|\det(I-Ae^{-i\theta_2})|^{\beta}dA \qquad (3.1.23)\\ &== e^{-i\beta(\theta_2-\theta_1)N}\mathbb{E}\left[|P_N(A,\theta_1)|^{2\beta}|P_N(A,\theta_2)|^{2\beta}\right]. \qquad (3.1.24)\end{aligned}$$

Hence, we may now use theorem 3.1.1 to calculate $\mathrm{MoM}_{U(N)}(2,\beta)$:

$$\begin{aligned}\mathrm{MoM}_{U(N)}(2,\beta) &= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}\mathbb{E}\left[|P_N(A,\theta_1)|^{2\beta}|P_N(A,\theta_2)|^{2\beta}\right]d\theta_1 d\theta_2 \qquad (3.1.25)\\ &= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}\int_{U(N)}\left|\det(I-Ae^{-i\theta_1})\right|^{2\beta}\left|\det(I-Ae^{-i\theta_2})\right|^{2\beta}dA d\theta_1 d\theta_2 \qquad (3.1.26)\\ &= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} e^{i\beta(\theta_2-\theta_1)N}\sum_{0\le m\le 2\beta N} I_{2\beta}(m;N)e^{i(\theta_1-\theta_2)m}d\theta_1 d\theta_2 \qquad (3.1.27)\\ &= \sum_{0\le m\le 2\beta N} I_{2\beta}(m;N)\delta_{m-\beta N}. \qquad (3.1.28)\end{aligned}$$

Immediately, theorem 3.1.1 gives us that $\mathrm{MoM}_{U(N)}(2,\beta)$ is a polynomial in N, and we have the correct leading order,

$$\mathrm{MoM}_{U(N)}(2,\beta) \sim \gamma_{2\beta}(\beta)N^{4\beta^2-1} + O_\beta(N^{4\beta^2-2}), \tag{3.1.29}$$

provided that $\gamma_{2\beta}(\beta) \neq 0$. Theorem 3.1.1 was proved by two methods: symmetric function theory and complex analysis. The former determines an equivalent structure for $\gamma_\eta(c)$ to that given in (3.1.15) coming from a standard lattice point count, which proves that $I_\eta(m;N)$ is a polynomial in N and makes it clear that $\gamma_{2\beta}(\beta) \neq 0$. By using complex analysis the result regarding the leading order in N can be established and the form for $\gamma_\eta(c)$ given in (3.1.15) is found.

Moments of moments, Toeplitz determinants, and Riemann–Hilbert problems

The conclusion of Keating et al. can also be arrived at using the work of Claeys and Krasovsky [36]. Their motivation was to study the asymptotic behaviour of Toeplitz determinants corresponding to symbols with two Fisher-Hartwig singularities (recall the introduction to Toeplitz determinants given in chapter 2, section 2.2.2). Their results are uniform in that they describe the transition between the two singularities being distinct and when they are permitted to merge. There is also a connection to a solution of a certain non-linear second-order ordinary differential equation, known as Painlevé V.

The following discussion of these techniques did not appear in [15]. We additionally cover the recent extension of the results of Claeys and Krasovsky, due to Fahs [72]. The work of Fahs concerns k distinct Fisher-Hartwig singularities, and gives uniform asymptotics as they are allowed to coalesce.

As stated in chapter 2, section 2.2.2, the $N \times N$ Toeplitz determinant for the symbol f is defined by

$$D_N(f) \coloneqq \det(\hat{f}_{j-k})_{j,k=1}^N \tag{3.1.30}$$

where f is a real-valued, 2π-periodic, integrable function with Fourier coefficients

$$\hat{f}_j \coloneqq \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-ij\theta}d\theta. \tag{3.1.31}$$

As previously seen, the connection between Toeplitz determinants and random matrix averages is the following identity [151], sometimes called the Heine identity,

$$D_N(f) = \frac{1}{(2\pi)^N N!}\int_0^{2\pi}\cdots\int_0^{2\pi}\prod_{j=1}^N f(\theta_j)|\Delta(e^{i\theta_1},\dots,e^{i\theta_N})|^2 d\theta_1\cdots d\theta_N. \tag{3.1.32}$$

Thus,

$$\mathbb{E}\left[\prod_{j=1}^k |P_N(A,\theta_j)|^{2\beta}\right] = D_N(f) \tag{3.1.33}$$

where the symbol $f(z) = \prod_{j=1}^k |z - e^{i\theta_j}|^{2\beta}$, and as usual $\mathbb{E}[\cdot]$ is expectation over $U(N)$. One says that the symbol f has k Fisher-Hartwig (FH) singularities at $e^{i\theta_1},\dots,e^{i\theta_k}$.

In the simplest case, $k=1$ (which recall was determined by Keating and Snaith, see theorem 2.1.1), then we have just a single FH singularity. Without loss of generality, due to the rotational invariance of the Haar measure, we can write the symbol $f^*(z) = |z-1|^{2\beta}$ in this case. Then, one can show (see for example (1.13) in [36]),

$$\log D_N(f^*) = \beta^2 \log N + \log\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)} + O\left(\frac{1}{N}\right), \tag{3.1.34}$$

provided that $\mathrm{Re}(\beta) > -1/2$. One immediately sees that this precisely coincides with the asymptotic form for $\mathrm{MoM}_{U(N)}(1,\beta)$ calculated by Keating and Snaith.

The main result[2] of Claeys and Krasovsky is to handle symbols with two FH singularities.

[2]Adapted here slightly from the statement given in [36] (there, theorem 1.15) for consistency of notation.

Theorem 3.1.2 (Claeys and Krasovsky [36]). *Take $\beta > -1/4$. Let*

$$f(z) = |z - e^{i\theta_1}|^{2\beta}|z - e^{-i\theta_2}|^{2\beta}. \tag{3.1.35}$$

Then for $0 < t_1 < \pi$ and as $N \to \infty$,

$$\int_0^{t_1} D_N(f)dt = \begin{cases} c_1(t_1,\beta)N^{2\beta^2}(1+o(1)) & \text{if } 2\beta^2 < 1 \\ c_2 N\log N(1+o(1)) & \text{if } 2\beta^2 = 1 \\ c_3(\beta)N^{4\beta^2-1}(1+o(1)) & \text{if } 2\beta^2 > 1. \end{cases} \tag{3.1.36}$$

The constants c_1, c_2, c_3 are explicitly given in [36], and are additionally related to a solution to the Painlevé V *differential equation.*

Thus, (3.1.36) exactly matches[3] (3.1.7) with $k = 2$.

After the publication of [15] (the paper that this chapter is based on), Fahs extended the results of Claeys and Krasovsky from 2 merging singularities to $k \in \mathbb{N}$ merging singularities. Fahs' results capture the leading order behaviour in each regime, as well as at the transition point $k = 1/\beta^2$, though are not precise enough to learn information about the leading order coefficients beyond $k = 2$. Specifically, he proves the following result (again, we have slightly adapted the statement given in [72], theorem 1.1 there, for notational consistency).

Theorem 3.1.3 (Fahs [72]). *Let $k \in \mathbb{N}$ and set*

$$f(z) := \prod_{j=1}^{k} |z - e^{i\theta_j}|^{2\beta}, \tag{3.1.37}$$

with $\beta \geq 0$ and $0 \leq \theta_1 < \theta_2 < \cdots < \theta_k \leq 2\pi$. Then as $N \to \infty$,

$$\log D_N(f) = k\beta^2 \log N - 2\beta^2 \sum_{1\leq i<j\leq k} \log\left(\sin\left|\frac{\theta_i - \theta_j}{2}\right| + \frac{1}{N}\right) + O(1), \tag{3.1.38}$$

where the error term is uniform for $0 \leq \theta_1 < \cdots < \theta_k \leq 2\pi$.

To see that this verifies (3.1.7), define for $\varepsilon > 0$,

$$I_\varepsilon(\beta) := \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi} \prod_{1\leq i<j\leq k}\left(\sin\left|\frac{\theta_i-\theta_j}{2}\right| + \varepsilon\right)^{-2\beta^2} d\theta_1\cdots d\theta_k. \tag{3.1.39}$$

If $0 < k\beta^2 < 1$ then (3.1.39) is well-defined at $\varepsilon = 0$ and can be calculated using the Selberg integral (as remarked by Fyodorov and Keating, see section 2.2.2),

$$I_0(\beta) = \frac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)}. \tag{3.1.40}$$

Combining (3.1.40) with theorem 3.1.3 gives (after a little work, see section 2.1 in [72] for the details)

$$\mathrm{MoM}_{U(N)}(k,\beta) = \left(\frac{\mathcal{G}^2(1+\beta)}{\mathcal{G}(1+2\beta)}\right)^k \frac{\Gamma(1-k\beta^2)}{\Gamma^k(1-\beta^2)} N^{k\beta^2}, \tag{3.1.41}$$

provided $0 < k\beta^2 < 1$. In the other regimes, one has to tread more carefully. Such analysis can be

[3] Within their proof it is clear how to deal with the constraint on t_1, see the comments after (1.47) in [36].

found in [72], and recovers the result that

$$\mathrm{MoM}_{U(N)}(k,\beta) = O(1)N^{k^2\beta^2-k+1} \tag{3.1.42}$$

for $k\beta^2 > 1$, and for $k\beta^2 = 1$, one sees that $\log \mathrm{MoM}_{U(N)}(k,\beta) = \log N + \log\log N + O(1)$. This verifies (3.1.7), but obfuscates the leading order coefficient (which recall we are interested in since it is related to determining $\max_{\theta\in[0,2\pi)} \log |P_N(A,\theta)|$).

Within their proofs, both theorem 3.1.2 and theorem 3.1.3 first relate the Toeplitz determinants to a system of polynomials orthogonal on the unit circle. Then, they both characterize these polynomials as a Riemann–Hilbert problem and use the associated techniques in order to determine the respective results. The topic of Riemann–Hilbert problems is outside the scope of this thesis, but the interested reader may seek out [27, 60, 61] for further details, and [22] for an overview of connections to random matrix theory.

3.1.2 Statement of results

Our approach combines the methods and formulae developed in [33, 45, 112, 113, 116]; in particular we make use of the complex analytic techniques employed in [112, 113]. We first reformulate (3.1.9) in terms of symmetric function theory and a lattice point count function. This gives a polynomial bound on $\mathrm{MoM}_{U(N)}(k,\beta)$ at integer values of k, β, and N. We next use a representation in terms of multiple contour integrals; this furnishes an expression for $\mathrm{MoM}_{U(N)}(k,\beta)$ as an entire function of N and allows us to prove the following theorem.

Theorem 3.1.4. *Let $k,\beta \in \mathbb{N}$. Then*

$$\mathrm{MoM}_{U(N)}(k,\beta) = \gamma_{k,\beta} N^{k^2\beta^2-k+1} + O(N^{k^2\beta^2-k}), \tag{3.1.43}$$

where $\gamma_{k,\beta}$ can be written explicitly in the form of an integral.

Using a combinatorial sum equivalent to the multiple contour integrals due to [45], we then deduce the following result.

Theorem 3.1.5. *Let $k,\beta \in \mathbb{N}$. Then $\mathrm{MoM}_{U(N)}(k,\beta)$ is a polynomial in N.*

These theorems together prove (3.1.9) for $k,\beta \in \mathbb{N}$.

3.2 A symmetric function theoretic representation

As with the cases $k = 1, 2$ and $\beta \in \mathbb{N}$, we can rephrase the problem in terms of symmetric function theory. For a thorough introduction to this topic, see [122] and [147]. Much of the relevant theory was introduced in chapter 1, and any additional tools required for the subsequent calculations are given within this section so that this review may be self-contained.

The aim of this section is twofold. Firstly, we highlight the role that symmetric function theory plays in the analysis of the moments of moments as k increases. Secondly, in understanding how the results of Bump and Gamburd [33] and Keating et al. [113] generalize for higher k, we recover an explicit polynomial bound on $\mathrm{MoM}_{U(N)}(k,\beta)$ at integer values of k, β, and N. Much of the theory behind these results is also applicable to the theorems stated and proved within chapter 4.

We begin with a review of a result of Bump and Gamburd [33]. Their work uses the representation theory of $U(N)$ in order to calculate various moments, and provides a beautiful, combinatorial interpretation of many of the random matrix moments already encountered. In particular, they rederive theorem 2.1.1, the finite-N formula for $M_N(\beta) = \mathbb{E}[|P_N(A,\theta)|^{2\beta}]$, provided $\beta \in \mathbb{N}$.

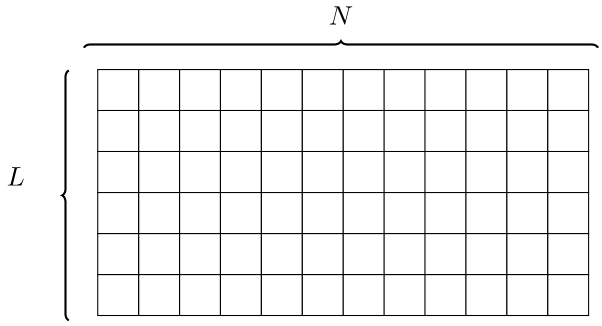


Figure 3.1: Example of a rectangular Young tableau of shape $\lambda = \langle N^L \rangle$.

Theorem 3.2.1 (Bump and Gamburd [33]). *Let $\beta \in \mathbb{N}$. Then*

$$\mathbb{E}[|P_N(A,\theta)|^{2\beta}] = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2\beta)}{\Gamma(j+\beta)^2}. \tag{3.2.1}$$

In order to prove theorem 3.2.1, Bump and Gamburd re-express the moments of characteristic polynomials as an average over Schur polynomials. They derive the following proposition[4].

Proposition 3.2.2 (Bump and Gamburd [33]). *Take $K, L, N \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_{K+L} \in \mathbb{C}$, then*

$$\int_{U(N)} \prod_{l=1}^{L} \det(I + \alpha_l^{-1} A^*) \prod_{k=1}^{K} \det(I + \alpha_{L+k} A) dA = \frac{s_{\langle N^L \rangle}(\alpha_1, \ldots, \alpha_{K+L})}{\prod_{l=1}^{L} \alpha_l^N}. \tag{3.2.2}$$

Before we comment on the proof, we recall the definition of a Schur function from chapter 1, section 1.1.6. Given a partition λ (a non-increasing sequence of integers with finitely many non-zero terms), the Schur function in n variables is

$$s_\lambda(x_1, \ldots, x_n) = \sum_{T \in \mathrm{SSYT}_n(\lambda)} x_1^{t_1} \cdots x_n^{t_n}, \tag{3.2.3}$$

where the sum is over all semistandard Young tableaux of shape λ with entries from $\{1, \ldots, n\}$. The partition λ is implicitly extended with zeros until it has length n, and $t_j = t_j(T)$ is the number of times j appears in T. One may also write partitions in multiplicative notation, so if m_j is the number of times j appears in λ, then $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle$. Hence, the Schur function appearing in the statement of proposition 3.2.2 relates to the partition $\lambda = \langle N^L \rangle$, whose Young diagram is given by figure 3.1.

The proof follows by using the dual Cauchy identity (see for example [122])

$$\prod_{i,j} (1 + x_i y_j) = \sum_{\lambda} s_{\lambda'}(x) s_\lambda(y) \tag{3.2.4}$$

to expand the determinants in the integrand in the left hand side of (3.2.2) in terms of Schur functions. Importantly, Schur functions are related to the representation theory of $U(N)$. In particular, see for example [32], a special case of the Weyl character formula gives the following.

Proposition 3.2.3 (Weyl character formula [156]). *Choose $N \in \mathbb{N}$ and let λ be a partition with length $l(\lambda) \leq N$. Take $A \in \mathrm{GL}(N, \mathbb{C})$ with eigenvalues $t_1, \ldots, t_N$. Define $\chi_\lambda(A) = \chi_\lambda^N(A) \coloneqq s_\lambda(t_1, \ldots, t_N)$. Then, the function χ_λ is the character of an irreducible analytic representation of* $\mathrm{GL}(N, \mathbb{C})$. *The character χ_λ is irreducible and its restriction to $U(N)$ is also irreducible.*

[4] The statement of proposition 3.2.2 is more generally stated in [33], giving a further equality for (3.2.2) as a permutation sum. The representation of the moments as a permutation sum had also already been proved by Conrey et al. [45], though again using a different method.

Using the dual Cauchy identity, integrating over $U(N)$, and employing orthogonality of characters thus gives the final statement of proposition 3.2.2.

Theorem 3.2.1 follows from proposition 3.2.2 by setting $L = K = \beta$ and $\alpha_1 = \cdots = \alpha_{2\beta} = 1$ (using the rotational invariance of the Haar measure). Thus

$$\int_{U(N)} |P_N(A,\theta)|^{2\beta} dA = s_{\langle N^\beta \rangle}(\overbrace{1,\ldots,1}^{2\beta}). \tag{3.2.5}$$

Using (3.2.3), the Schur polynomial $s_\lambda(1^n)$ is equal to the number of semistandard Young tableaux of shape λ with entries in $\{1,\ldots,n\}$. Using the following classical result on counting such tableaux, the 'Hook-content formula' (see [147]), one has the statement of the theorem.

Lemma 3.2.4. *The number of SSYT of shape λ with entries in $1,2,\ldots,n$ can be found by evaluating the Schur polynomial $s_\lambda(1,\ldots,1)$, where the argument is an n-long vector. We implicitly extend λ with zeros until it has length n. Then*

$$s_\lambda(1,1,\ldots,1) = \prod_{1\le i<j\le n} \frac{\lambda_i - \lambda_j + j - i}{j-i}, \tag{3.2.6}$$

which is a polynomial in $\lambda_i - \lambda_j$.

The following result is our extension of proposition 3.2.2 to allow the average to be taken over a product of characteristic polynomials. Our result gives that this is equal to an evaluation of a similar normalized Schur function to the one appearing in proposition 3.2.2. This proposition first appeared in [15] as proposition 2.1, though here we give the full details of the proof.

Proposition 3.2.5. *For $N,k,\beta \in \mathbb{N}$, we have*

$$\mathbb{E}_{A\in U(N)}\left[\prod_{j=1}^{k} |P_N(A,\theta_j)|^{2\beta}\right] = \frac{s_{\langle N^{k\beta}\rangle}(e^{-i\underline{\theta}})}{\prod_{j=1}^{k} e^{-iN\beta\theta_j}}, \tag{3.2.7}$$

where $s_\lambda(x_1,\ldots,x_n)$ is the Schur polynomial in n variables with respect to the partition λ, and we write $\lambda = \langle \alpha^n \rangle = (\underbrace{\alpha,\ldots,\alpha}_{n})$ and

$$e^{i\underline{\theta}} = (\overbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}^{\beta},\overbrace{e^{i\theta_2},\ldots,e^{i\theta_2}}^{\beta},\ldots,\overbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}^{\beta},\overbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}^{\beta},\overbrace{e^{i\theta_2},\ldots,e^{i\theta_2}}^{\beta},\ldots,\overbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}^{\beta}). \tag{3.2.8}$$

Proof. The proof follows naturally from the proof of proposition 3.2.2, and incorporates techniques of Dehaye [59]. Firstly, we manipulate the expectation over the unitary group using the functional equation for the characteristic polynomial,

$$\mathbb{E}\left[\prod_{j=1}^{k} |P_N(A,\theta_j)|^{2\beta}\right] = \int_{U(N)} \prod_{j=1}^{k} P_N(A,\theta_j)^\beta \prod_{j=1}^{k} P_N(A^*,-\theta_j)^\beta dA \tag{3.2.9}$$

$$= (-1)^{Nk\beta} \prod_{j=1}^{k} \left(e^{i\theta_j}\right)^{\beta N} \int_{U(N)} \overline{\det(A)}^\beta \prod_{j=1}^{2k\beta} P_N(A,\alpha_j) dA \tag{3.2.10}$$

where

$$\underline{\alpha} = (\alpha_1, \ldots, \alpha_{2k\beta}) \tag{3.2.11}$$
$$= (\underbrace{\theta_1, \ldots, \theta_1}_{\beta}, \underbrace{\theta_2, \ldots, \theta_2}_{\beta}, \ldots, \underbrace{\theta_k, \ldots, \theta_k}_{\beta}, \underbrace{\theta_1, \ldots, \theta_1}_{\beta}, \ldots, \underbrace{\theta_k, \ldots, \theta_k}_{\beta}). \tag{3.2.12}$$

We now use the dual Cauchy identity (see (3.2.4)) to rewrite the determinants in (3.2.10) in terms of the characters of the unitary group (i.e. Schur polynomials). Thus

$$\mathbb{E}\left[\prod_{j=1}^{k} |P_N(A,\theta_j)|^{2\beta}\right] = (-1)^{Nk\beta} \prod_{j=1}^{k} \left(e^{i\theta_j}\right)^{\beta N}$$
$$\times \int_{U(N)} \sum_{\lambda} (-1)^{|\lambda|} \overline{s_{\langle (k\beta)^N\rangle}(A)} s_{\lambda'}(A) s_{\lambda}(e^{-i\alpha_1}, \ldots, e^{-i\alpha_{2k\beta}}) dA \tag{3.2.13}$$

where $s_\lambda(A)$ is the Schur polynomial $s_\lambda(x_1, \ldots, x_N)$ evaluated at the eigenvalues of A. Using orthogonality of characters, the integral with respect to the Haar measure is zero unless $\lambda = \langle N^{k\beta}\rangle$, and we arrive at the statement of the result.

$$\mathbb{E}\left[\prod_{j=1}^{k} |P_N(A,\theta_j)|^{2\beta}\right] = \frac{s_{\langle N^{k\beta}\rangle}(e^{-i\underline{\alpha}})}{\prod_{j=1}^{k} e^{-iN\beta\theta_j}}. \tag{3.2.14}$$

□

Using proposition 3.2.5, we can rewrite $\mathrm{MoM}_{U(N)}(k,\beta)$ in terms of Schur functions,

$$\mathrm{MoM}_{U(N)}(k,\beta) = \frac{1}{(2\pi)^k} \int_0^{2\pi} \cdots \int_0^{2\pi} \frac{s_{\langle N^{k\beta}\rangle}(e^{-i\underline{\theta}})}{\prod_{j=1}^{k} e^{-iN\beta\theta_j}} \prod_{j=1}^{k} d\theta_j \tag{3.2.15}$$
$$= \frac{1}{(2\pi)^k} \int_0^{2\pi} \cdots \int_0^{2\pi} \sum_{T} e^{-i\theta_1(\tau_1 - N\beta)} \cdots e^{-i\theta_k(\tau_k - N\beta)} \prod_{j=1}^{k} d\theta_j \tag{3.2.16}$$

where

$$\tau_j = t_{2(j-1)\beta+1} + \cdots + t_{2j\beta} \quad \text{for } j = 1, \ldots, k, \tag{3.2.17}$$

and as usual $t_n = t_n(T)$ is the number of entries in the tableau T equal to n. By computing the θ integrals in (3.2.16), one finds that the contribution is zero unless $\tau_j = N\beta$ for all j. Relating this back to the semistandard Young tableaux, this means that $\mathrm{MoM}_{U(N)}(k,\beta)$ is equal to the number of SSYT of rectangular shape $\langle N^{k\beta}\rangle$ with entries in $\{1, \ldots, 2k\beta\}$ with the *additional constraint* that there have to be $N\beta$ entries from each of the sets

$$\{2(j-1)\beta + 1, \ldots, 2j\beta\}, \quad \text{for } j \in \{1, \ldots, k\}. \tag{3.2.18}$$

Thus we define *restricted* SSYT (RSSYT) to be those SSYT, $\widetilde{T}$, satisfying this additional condition. We have therefore found that

$$\mathrm{MoM}_{U(N)}(k,\beta) = \sum_{\widetilde{T}} 1, \tag{3.2.19}$$

where the sum is now over $\widetilde{T}$, a set of restricted SSYT described above. When specialised to the case of $k = 1, \beta \in \mathbb{N}$, this approach matches the proof of theorem 3.2.1 of Bump and Gamburd.

Since the set of RSSYT is a proper subset of all SSYT, we have that the number of RSSYT of rectangular shape $\lambda = \langle N^{k\beta}\rangle$ is bounded by a polynomial in N of degree $k^2\beta^2$. This concludes the proof of the bound on $\mathrm{MoM}_{U(N)}(k,\beta)$ for integer values of k, β, and N.

Following the publication of [15], Assiotis and Keating [10] extended this analysis using *Gelfand-Tsetlin* patterns. They are able to recover theorem 3.1.4 using this combinatorial approach, and get an alternative interpretation of $\gamma_{k,\beta}$, the leading order coefficient. Such an approach is well-adaptable to the case of symplectic and orthogonal moments of moments, and is the subject of chapter 4.

3.3 A multiple contour integral representation

In this section we give the proof of theorem 3.1.4. We rely on a series of lemmas which we state in section 3.3.1 and then prove in section 3.3.2.

3.3.1 Structure of proof of theorem 3.1.4

Recall from the introduction and section 3.1.1 that theorem 3.1.4 is known for $k = 1, 2$, so we will henceforth focus on integers $k > 2$ (though the method we now develop can be adapted for the cases $k = 1, 2$ as well). A key element of the proof is the following result (lemma 2.1) of Conrey et al. [45].

Lemma 3.3.1. *For $\alpha_j \in \mathbb{C}$,*

$$\int_{U(N)} \prod_{j=m+1}^{n} \det(I - Ae^{\alpha_j}) \prod_{j=1}^{m} \det(I - A^* e^{-\alpha_j}) dA$$
$$= \frac{(-1)^{n(n-1)/2}}{(2\pi i)^n m!(n-m)!} \prod_{q=m+1}^{n} e^{N\alpha_q} \oint \cdots \oint \frac{e^{-N\sum_{l=m+1}^{n} z_l} \Delta(z_1, \ldots, z_n)^2 dz_1 \cdots dz_n}{\prod_{1\le l\le m<q\le n}(1-e^{z_q - z_l}) \prod_{l=1}^{n}\prod_{q=1}^{n}(z_l - \alpha_q)},$$

where the contours enclose the poles at $\alpha_1, \ldots, \alpha_n$ and $\Delta(z_1, \ldots, z_n) = \prod_{i<j}(z_j - z_i)$ is the Vandermonde determinant.

This multiple contour integral is nearly identical to the conjectural form for shifted moments of the zeta function, as discussed in section 2.1.2.

Before using lemma 3.3.1, we first define

$$I_{k,\beta}(\theta_1, \ldots, \theta_k) := \mathbb{E}_{A\in U(N)} \left(\prod_{j=1}^{k} |P_N(A, \theta_j)|^{2\beta} \right), \tag{3.3.1}$$

which captures the average over the unitary group and thus

$$\mathrm{MoM}_{U(N)}(k,\beta) = \frac{1}{(2\pi)^k} \int_0^{2\pi} \cdots \int_0^{2\pi} I_{k,\beta}(\theta_1, \ldots, \theta_k) d\theta_1 \cdots d\theta_k. \tag{3.3.2}$$

Our focus now switches to understanding $I_{k,\beta}(\underline{\theta})$. Following Keating et al. [113], we use lemma 3.3.1 to expand the average over the CUE to a multiple contour integral.

$$I_{k,\beta}(\underline{\theta}) = \frac{(-1)^{k\beta} e^{-i\beta N\sum_{j=1}^{k}\theta_j}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \oint \cdots \oint \frac{e^{-N(z_{k\beta+1}+\cdots+z_{2k\beta})} \Delta(z_1, \ldots, z_{2k\beta})^2 dz_1 \cdots dz_{2k\beta}}{\prod_{m\le k\beta<n}(1-e^{z_n - z_m}) \prod_{m=1}^{2k\beta}\prod_{n=1}^{k}(z_m + i\theta_n)^{2\beta}}. \tag{3.3.3}$$

We note that equations (3.3.2) and (3.3.3) define $\mathrm{MoM}_{U(N)}(k,\beta)$ as an analytic function of N.

We deform each of the $2k\beta$ contours so that any one now consists of a sum of k small circles surrounding each of the poles at $-i\theta_1, \ldots, -i\theta_k$, given by $\Gamma_{-i\theta_l}$ for $l \in \{1, \ldots, k\}$, and connecting straight lines whose contributions will cancel (just as in [113], we follow the procedure outlined in [112]).

This means that we will have a sum of $k^{2k\beta}$ multiple integrals,

$$I_{k,\beta}(\underline{\theta}) = \frac{(-1)^{k\beta} e^{-i\beta N \sum_{j=1}^{k} \theta_j}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \sum_{\varepsilon_j \in \{1,\ldots,k\}} J_{k,\beta}(\underline{\theta}; \varepsilon_1, \ldots, \varepsilon_{2k\beta}), \tag{3.3.4}$$

where

$$J_{k,\beta}(\underline{\theta}; \varepsilon_1, \ldots, \varepsilon_{2k\beta}) = \int_{\Gamma_{-i\theta_{\varepsilon_1}}} \cdots \int_{\Gamma_{-i\theta_{\varepsilon_{2k\beta}}}} \frac{e^{-N(z_{k\beta+1}+\cdots+z_{2k\beta})} \Delta(z_1, \ldots, z_{2k\beta})^2 dz_1 \cdots dz_{2k\beta}}{\prod_{m\leq k\beta<n} (1-e^{z_n - z_m}) \prod_{m=1}^{2k\beta} \prod_{n=1}^{k} (z_m + i\theta_n)^{2\beta}} \tag{3.3.5}$$

is the multiple contour integral with $2k\beta$ contours each specialised around one of the k poles determined by the vector $\underline{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2k\beta})$.

In fact, many of the summands do not contribute to the sum due to the highly symmetric nature of the integrand. The following lemma, which is a generalized version of lemma 4.11 in [113], determines exactly which summands make no contribution.

Lemma 3.3.2. *Let a choice of contours in* (3.3.4) *be denoted by* $\underline{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2k\beta})$ *where* $\varepsilon_j \in \{1, \ldots, k\}$. *If any particular pole is overrepresented in* $\underline{\varepsilon}$ *(i.e. some pole* $-i\theta^*$ *features in at least* $2\beta+1$ *contours), then that summand is identically zero.*

Thus we have that

$$I_{k,\beta}(\underline{\theta}) = \frac{(-1)^{k\beta} e^{-i\beta N \sum_{j=1}^{k} \theta_j}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \sum_{l_1=0}^{2\beta} \cdots \sum_{l_{k-1}=0}^{2\beta} c_{\underline{l}}(k,\beta) J_{k,\beta;\underline{l}}(\underline{\theta}), \tag{3.3.6}$$

where $J_{k,\beta;\underline{l}}(\underline{\theta})$ is the integral $J_{k,\beta}(\underline{\theta}; \underline{\varepsilon})$ with contours given by

$$\underline{\varepsilon} = (\overbrace{1,\ldots,1}^{l_1}, \overbrace{2,\ldots,2}^{l_2}, \ldots, \overbrace{k-1,\ldots,k-1}^{l_{k-1}}, \overbrace{k,\ldots,k}^{2\beta}, \overbrace{k-1,\ldots,k-1}^{2\beta-l_{k-1}}, \ldots, \overbrace{1,\ldots,1}^{2\beta-l_1}),$$

and $c_{\underline{l}}(k,\beta)$ is a product of binomial coefficients capturing the symmetry exhibited by the integrand:

$$\begin{aligned} c_{\underline{l}}(k,\beta) = &\binom{k\beta}{l_1}\binom{k\beta-l_1}{l_2}\binom{k\beta-(l_1+l_2)}{l_3}\cdots\binom{k\beta-\sum_{m=1}^{k-2} l_m}{l_{k-1}} \\ &\times \binom{k\beta}{2\beta-l_1}\binom{(k-2)\beta+l_1}{2\beta-l_2}\cdots\binom{k\beta-\sum_{m=1}^{k-2}(2\beta-l_m)}{2\beta-l_{k-1}}. \end{aligned} \tag{3.3.7}$$

So $c_{\underline{l}}(k,\beta)$ counts the number of ways of picking l_1 of the first $k\beta$ contours and $2\beta - l_1$ of the second $k\beta$ contours to surround $-i\theta_1$, and then repeating on the remaining $k\beta - l_1$ contours in the first half and $(k-2)\beta + l_1$ contours in the second half, and so on.

Note also that the coefficient $c_{\underline{l}}(k,\beta)$ will not allow 'overcrowding' of either half. If for example we were to set $l_1 = \cdots = l_{k-1} = 2\beta$ then we would be trying to fit $2(k-1)\beta$ labels on to $k\beta$ contours, which clearly cannot be done since we assume $k > 2$. However, $c_{\underline{l}}(k,\beta)$ contains the binomial coefficient

$$\binom{k\beta-\sum_{m=1}^{k-2} l_m}{l_{k-1}} = \binom{(4-k)\beta}{2\beta} = 0 \quad \text{for } k > 2, \tag{3.3.8}$$

and so is zero for this choice of $\underline{l}$.

Next we perform the change of variables,

$$z_n = \frac{v_n}{N} - i\alpha_n,$$

where

$$\alpha_n = \begin{cases} \theta_1 & \text{if } n \in \{1,\ldots,l_1\} \cup \{2(k-1)\beta+1+l_1,\ldots,2k\beta\} \\ \theta_2 & \text{if } n \in \{l_1+1,\ldots,l_1+l_2\} \cup \{2(k-2)\beta+1+l_1+l_2,\ldots,2(k-1)\beta+l_1\} \\ \vdots & \quad\vdots \\ \theta_{k-1} & \text{if } n \in \{\sum_{m=1}^{k-2} l_m+1,\ldots,\sum_{m=1}^{k-1} l_m\} \cup \{2\beta+1+\sum_{m=1}^{k-1} l_m,\ldots,4\beta+\sum_{m=1}^{k-2} l_m\} \\ \theta_k & \text{if } n \in \{\sum_{m=1}^{k-1} l_m+1,\ldots,\sum_{m=1}^{k-1} l_m+2\beta\}. \end{cases} \tag{3.3.9}$$

which shifts all the contours to be small circles surrounding the origin. We have in particular used here that (3.3.6) defines exactly the relationship between the contours and the poles at the $-i\theta_j$. This variable change also allows us to pull out the dependence on N in the integrand. Then up to terms of order $1/N$ smaller[5], we have that the integrand of $J_{k,\beta;\underline{l}}(\underline{\theta})$ is

$$\begin{aligned} &\frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} e^{iN\sum_{m=k\beta+1}^{2k\beta}\alpha_m} \prod_{\substack{m<n \\ \alpha_m\neq\alpha_n}} (i\alpha_m - i\alpha_n)^2 \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} \left(\frac{v_n-v_m}{N}\right)^2 \prod_{m=1}^{2k\beta} \frac{dv_m}{N}}{\prod_{\substack{m\le k\beta<n \\ \alpha_n\neq\alpha_m}} \left(1-e^{\frac{v_n-v_m}{N}} e^{i(\alpha_m-\alpha_n)}\right) \prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} \left(\frac{v_m-v_n}{N}\right) \prod_{m=1}^{2k\beta}\prod_{n=1}^{k} \left(\frac{v_m}{N}+i(\theta_n-\alpha_m)\right)^{2\beta}} \\ &\sim \frac{e^{iN\sum_{m=k\beta+1}^{2k\beta}\alpha_m}}{N^{2k\beta}} \frac{\prod_{\substack{m<n \\ \alpha_m\neq\alpha_n}} (i\alpha_m - i\alpha_n)^2}{\prod_{\substack{m<n \\ \alpha_m\neq\alpha_n}} (i\alpha_m - i\alpha_n)^2 \prod_{m=1}^{2k\beta}\left(\frac{v_m}{N}\right)^{2\beta}} \\ &\quad\times \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} \left(\frac{v_n-v_m}{N}\right)^2 \prod_{m=1}^{2k\beta} dv_m}{\prod_{\substack{m\le k\beta<n \\ \alpha_n\neq\alpha_m}} \left(1-e^{\frac{v_n-v_m}{N}} e^{i(\alpha_m-\alpha_n)}\right) \prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} \left(\frac{v_m-v_n}{N}\right)} &(3.3.10) \\ &= \frac{e^{iN\sum_{m=k\beta+1}^{2k\beta}\alpha_m} N^{4k\beta^2}}{N^{2k\beta}} \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} \left(\frac{v_n-v_m}{N}\right)^2 \prod_{m=1}^{2k\beta} \frac{dv_m}{v_m^{2\beta}}}{\prod_{\substack{m\le k\beta<n \\ \alpha_n\neq\alpha_m}} \left(1-e^{\frac{v_n-v_m}{N}} e^{i(\alpha_m-\alpha_n)}\right) \prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} \left(\frac{v_m-v_n}{N}\right)}. &(3.3.11) \end{aligned}$$

To determine the power of N coming from the terms originating from the Vandermonde determinant, we count the sizes of the following sets,

$$\#\{(m,n) : 1 \le m < n \le 2k\beta\} = \binom{2k\beta}{2} = k\beta(2k\beta-1) \tag{3.3.12}$$

$$\#\{(m,n) : 1 \le m < n \le 2k\beta, \alpha_m \neq \alpha_n\} = 2k\beta^2(k-1) \tag{3.3.13}$$

$$\#\{(m,n) : 1 \le m < n \le 2k\beta, \alpha_m = \alpha_n\} = k\beta(2\beta-1). \tag{3.3.14}$$

The count (3.3.12) is trivial. One sees that, for example, (3.3.13) results from the following calculation. Firstly, we recall from (3.3.9) the structure of the vector $\underline{\alpha}$,

$$(\overbrace{\theta_1,\ldots,\theta_1}^{l_1},\overbrace{\theta_2,\ldots,\theta_2}^{l_2},\ldots,\overbrace{\theta_{k-1},\ldots,\theta_{k-1}}^{l_{k-1}},\underbrace{\theta_k,\ldots,\theta_k}_{2\beta},\overbrace{\theta_{k-1},\ldots,\theta_{k-1}}^{2\beta-l_{k-1}},\ldots,\overbrace{\theta_1,\ldots,\theta_1}^{2\beta-l_1}). \tag{3.3.15}$$

As we are looking for pairs (m,n) such that $m<n$ and $\alpha_m \neq \alpha_n$, it is clear that any of the first l_1 choices of θ_1 can be paired with any of the following $2k\beta - l_1$ options, except for the final $2\beta - l_1$ as these are also θ_1. Thus, the total number of pairs (m,n) with $m \in \{1,\ldots,l_1\}$, $m<n$, and $\theta_n \neq \theta_1$ is $l_1(2k\beta - l_1 - (2\beta - l_1))$. Continuing in this fashion we see that in general, for $m \in \{1,\ldots,\sum_{j=1}^{k-1} l_j\}$,

[5]Henceforth, whenever we write 'up to terms of order $1/N$' followed by a statement of the form $A(N) \sim B(N)$ we mean that $A(N) = B(N)(1+O(1/N))$ as $N \to \infty$.

the number of such pairs is given by considering

$$\sum_{i=1}^{k-1} l_i\Big(2\beta + \sum_{j=i+1}^{k-1} l_j + \sum_{\substack{j=1\\ j\neq i}}^{k-1}(2\beta - l_j)\Big). \tag{3.3.16}$$

Similarly, for $m \in \{l_1 + \cdots + l_{k-1} + 1, \ldots, l_1 + \cdots + l_{k-1} + 2\beta\}$, the number of pairs satisfying the correct conditions is

$$\sum_{i=1}^{k-1} 2\beta(2\beta - l_i), \tag{3.3.17}$$

and if $m \in \{l_1 + \cdots + l_{k-1} + 2\beta + 1, \ldots, 2k\beta\}$ then we get

$$\sum_{i=1}^{k-1}(2\beta - l_i)\sum_{j=1}^{i-1}(2\beta - l_j). \tag{3.3.18}$$

In total therefore, we have to evaluate

$$\sum_{i=1}^{k-1} l_i\Big(2\beta + \sum_{j=i+1}^{k-1} l_j + \sum_{\substack{j=1\\ j\neq i}}^{k-1}(2\beta - l_j)\Big) + \sum_{i=1}^{k-1} 2\beta(2\beta - l_i) + \sum_{i=1}^{k-1}(2\beta - l_i)\sum_{j=1}^{i-1}(2\beta - l_j). \tag{3.3.19}$$

By collecting like terms we see that (3.3.19) is equal to

$$\begin{aligned}
&2\beta(k-1)\sum_{i=1}^{k-1} l_i - \sum_{i=1}^{k-1}\sum_{j=1}^{i-1} l_i l_j + 2\beta(2\beta(k-1) - \sum_{i=1}^{k-1} l_i) + \sum_{i=1}^{k-1}\sum_{j=1}^{i-1}(4\beta^2 - 2\beta(l_i + l_j) + l_i l_j)\\
&\quad = 2\beta(k-1)\sum_{i=1}^{k-1} l_i + 2\beta(2\beta(k-1) - \sum_{i=1}^{k-1} l_i) + \sum_{i=1}^{k-1}\sum_{j=1}^{i-1}(4\beta^2 - 2\beta(l_i + l_j)) &(3.3.20)\\
&\quad = 2\beta^2 k(k-1) + 2\beta\left((k-2)\sum_{i=1}^{k-1} l_i - \sum_{i=1}^{k-1}\sum_{j=1}^{i-1}(l_i + l_j)\right) &(3.3.21)\\
&\quad = 2\beta^2 k(k-1) + 2\beta\left((k-2)\sum_{i=1}^{k-1} l_i - \sum_{i=1}^{k-1}(i-1)l_i - \sum_{j=1}^{k-2}(k-1-j)l_j\right) &(3.3.22)\\
&\quad = 2\beta^2 k(k-1) + 2\beta\left((k-2)\sum_{i=1}^{k-1} l_i - \sum_{i=1}^{k-1}(i-1+k-1-i)l_i\right) &(3.3.23)\\
&\quad = 2\beta^2 k(k-1) + 2\beta\left((k-2)\sum_{i=1}^{k-1} l_i - (k-2)\sum_{i=1}^{k-1} l_i\right) &(3.3.24)\\
&\quad = 2k\beta^2(k-1) &(3.3.25)
\end{aligned}$$

as claimed in (3.3.13). Then (3.3.14) can be deduced immediately since it is the difference between (3.3.12) and (3.3.13).

To count the remaining power of N that remains in the denominator of the integrand in (3.3.11), we define the following index sets inspired by the product terms in (3.3.11).

$$A_{k,\beta;\underline{l}} := \{(m,n) : 1 \leq m \leq k\beta < n \leq 2k\beta, \alpha_m = \alpha_n\} \tag{3.3.26}$$

$$B_{k,\beta;\underline{l}} := \{(m,n) : 1 \leq m \leq k\beta < n \leq 2k\beta, \alpha_m \neq \alpha_n\}, \tag{3.3.27}$$

so $|A_{k,\beta;\underline{l}}| + |B_{k,\beta;\underline{l}}| = k^2\beta^2$. Hence

$$\prod_{\substack{m<n \\ \alpha_m=\alpha_n}} \left(\frac{v_n - v_m}{N}\right)^2 = \frac{1}{N^{2k\beta(2\beta-1)}} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n - v_m)^2 \tag{3.3.28}$$

$$\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} \left(\frac{v_m - v_n}{N}\right) = \frac{1}{(-N)^{|A_{k,\beta;\underline{l}}|}} \prod_{(m,n)\in A_{k,\beta;\underline{l}}} (v_n - v_m)\,. \tag{3.3.29}$$

Returning once more to the integrand of $J_{k,\beta;\underline{l}}(\underline{\theta})$, we have that up to terms of order $1/N$ smaller it is equal to

$$e^{iN\sum_{m=k\beta+1}^{2k\beta}\alpha_m}(-N)^{|A_{k,\beta;\underline{l}}|}\frac{e^{-\sum_{m=k\beta+1}^{2k\beta}v_m}\prod_{\substack{m<n \\ \alpha_m=\alpha_n}}(v_n-v_m)^2\prod_{m=1}^{2k\beta}\frac{dv_m}{v_m^{2\beta}}}{\prod_{(m,n)\in B_{k,\beta;\underline{l}}}\left(1-e^{\frac{v_n-v_m}{N}}e^{i(\alpha_m-\alpha_n)}\right)\prod_{(m,n)\in A_{k,\beta;\underline{l}}}(v_n-v_m)}. \tag{3.3.30}$$

If we set $l_k = k\beta - (l_1 + \cdots + l_{k-1})$, then the calculation of the size of $A_{k,\beta;\underline{l}}$ follows the method outlined for (3.3.13) and we have

$$(-1)^{|A_{k,\beta;\underline{l}}|} = (-1)^{\sum_{j=1}^{k} l_j(2\beta-l_j)} = (-1)^{k\beta},$$

and so

$$I_{k,\beta}(\underline{\theta}) \sim \sum_{l_1,\ldots,l_{k-1}=0}^{2\beta} \frac{c_{\underline{l}}(k,\beta)N^{|A_{k,\beta;\underline{l}}|}}{(2\pi i)^{2k\beta}((k\beta)!)^2}\int_{\Gamma_0}\cdots\int_{\Gamma_0}\frac{e^{-iN(\beta\sum_{j=1}^{k}\theta_j-\sum_{m=k\beta+1}^{2k\beta}\alpha_m)}f(\underline{v};\underline{l})\prod_{m=1}^{2k\beta}dv_m}{\prod_{(m,n)\in B_{k,\beta;\underline{l}}}\left(1-e^{\frac{v_n-v_m}{N}}e^{i(\alpha_m-\alpha_n)}\right)}, \tag{3.3.31}$$

where we isolate the terms with no θ dependence and denote them by $f(\underline{v};\underline{l})$, so explicitly

$$f(\underline{v};\underline{l}) = \frac{e^{-\sum_{m=k\beta+1}^{2k\beta}v_m}\prod_{\substack{m<n \\ \alpha_m=\alpha_n}}(v_n-v_m)^2}{\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}}(v_n-v_m)\prod_{m=1}^{2k\beta}v_m^{2\beta}}. \tag{3.3.32}$$

We now focus on the denominator of the integrand in (3.3.31), which is the final term involving N. It will prove to be fruitful to use the properties of the α_js to remove the dependence on $(m,n) \in B_{k,\beta;\underline{l}}$. To this end, one can split the set $B_{k,\beta;\underline{l}}$ into $\binom{k}{2}$ disjoint subsets:

$$B_{k,\beta;\underline{l}} = \bigcup_{1\le\sigma<\tau\le k} S_{\sigma,\tau}, \tag{3.3.33}$$

where

$$S_{\sigma,\tau} := \{(m,n) \in B_{k,\beta;\underline{l}} : \alpha_m - \alpha_n = \pm(\theta_\tau - \theta_\sigma)\}, \tag{3.3.34}$$

and further partition $S_{\sigma,\tau}$ into two subsets $S^+_{\sigma,\tau}, S^-_{\sigma,\tau}$, where

$$S^+_{\sigma,\tau} = \{(m,n) \in B_{k,\beta;\underline{l}} : \alpha_m - \alpha_n = \theta_\tau - \theta_\sigma\} \tag{3.3.35}$$

$$S^-_{\sigma,\tau} = \{(m,n) \in B_{k,\beta;\underline{l}} : \alpha_m - \alpha_n = \theta_\sigma - \theta_\tau\}. \tag{3.3.36}$$

The goal of the next lemma is to use the structure of the vector $\underline{\alpha}$, to 'decouple' the pair (m,n) from the term $\exp(i(\alpha_m - \alpha_n))$.

Lemma 3.3.3.

$$\prod_{(m,n)\in B_{k,\beta;\underline{l}}} \left(1-\exp\left(\frac{v_n-v_m}{N}\right)\exp\left(i(\alpha_m-\alpha_n)\right)\right)^{-1}$$

$$= (-1)^{g(k,\beta;\underline{l})} \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\leq\sigma<\tau\leq k}}^{\infty} \exp\left(i\sum_{\gamma<\rho}(\theta_\rho-\theta_\gamma)(t_{\gamma,\rho}+|S^-_{\gamma,\rho}|)\right) \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right),$$

where

$$(v_n-v_m)^{\pm} = \begin{cases} v_n-v_m & \text{if } (m,n)\in S^+_{\sigma,\tau} \\ v_m-v_n & \text{if } (m,n)\in S^-_{\sigma,\tau}, \end{cases}$$

the first sum appearing in the right hand size of the statement of the lemma should be read as

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\leq\sigma<\tau\leq k}}^{\infty} = \sum_{t_{1,2}=-\infty}^{\infty}\sum_{t_{1,3}=-\infty}^{\infty}\cdots\sum_{t_{1,k}=-\infty}^{\infty}\sum_{t_{2,3}=-\infty}^{\infty}\cdots\sum_{t_{k-1,k}=-\infty}^{\infty} \tag{3.3.37}$$

and the second sum is over vectors (of weights) $\underline{x}=(x_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}}$ *subject to constraints given by* $(\star)$*, which are*

$$\sum_{(m,n)\in S_{\sigma,\tau}} x_{m,n} = t_{\sigma,\tau}+|S^-_{\sigma,\tau}|, \quad 1\leq\sigma<\tau\leq k \tag{3.3.38}$$

$$x_{m,n}\in\mathbb{Z}, \qquad \forall(m,n)\in B_{k,\beta;\underline{l}} \tag{3.3.39}$$

$$H(-x_{m,n}\,\mathrm{Re}\{(v_n-v_m)^{\pm}\})=1, \qquad \forall(m,n)\in B_{k,\beta;\underline{l}}, \tag{3.3.40}$$

where $\mathrm{H}(x)$ *is the Heaviside step function (so* $\mathrm{H}(x)=1$ *if* $x\geq 0$ *and* $\mathrm{H}(x)=0$ *if* $x<0$*). Furthermore, the prefactor can be expressed as follows,*

$$(-1)^{g(k,\beta;\underline{l})} = (-1)^{\sum_{\sigma<\tau}|S^-_{\sigma,\tau}|} \prod_{\substack{(m,n)\in S_{\sigma,\tau} \\ 1\leq\sigma<\tau\leq k}} (-\operatorname{sgn}(\mathrm{Re}\{(v_n-v_m)^{\pm}\})). \tag{3.3.41}$$

Using lemma 3.3.3 with (3.3.31) we have

$$\begin{aligned} I_{k,\beta}(\underline{\theta}) \sim & \sum_{l_1,\ldots,l_{k-1}=0}^{2\beta} \frac{\tilde{c}_{\underline{l}}(k,\beta)N^{|A_{k,\beta;\underline{l}}|}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \\ & \times \int_{\Gamma_0}\cdots\int_{\Gamma_0} f(\underline{v};\underline{l})\exp\left(iN\Big(\sum_{m=k\beta+1}^{2k\beta}\alpha_m - \beta\sum_{j=1}^{k}\theta_j\Big)\right) \\ & \qquad \times \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\leq\sigma<\tau\leq k}}^{\infty} \exp\left(i\sum_{1\leq\gamma<\rho\leq k}(\theta_\rho-\theta_\gamma)(t_{\gamma,\rho}+|S^-_{\gamma,\rho}|)\right) \\ & \qquad \times \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right)\prod_{m=1}^{2k\beta} dv_m, \end{aligned} \tag{3.3.42}$$

where

$$\tilde{c}_{\underline{l}}(k,\beta) = (-1)^{g(k,\beta;\underline{l})}c_{\underline{l}}(k,\beta),$$

and the function $g(k,\beta;\underline{l})$ is as described in the statement of lemma 3.3.3. Now we relate $I_{k,\beta}(\underline{\theta})$ back to $\mathrm{MoM}_{U(N)}(k,\beta)$ using (3.3.2), and deduce the following lemma.

Lemma 3.3.4.

$$\mathrm{MoM}_{U(N)}(k,\beta) \sim \sum_{l_1,\dots,l_{k-1}=0}^{2\beta} \frac{\tilde{c}_{\underline{l}}(k,\beta)N^{|A_{k,\beta;\underline{l}}|}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \int_{\Gamma_0}\cdots\int_{\Gamma_0} f(\underline{v};\underline{l}) \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right)$$
$$\times \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)} \prod_{m=1}^{2k\beta} dv_m.$$

In order to take the asymptotic analysis further we require the next lemma.

Lemma 3.3.5.

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right) \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)}$$
$$\sim N^{|B_{k,\beta;\underline{l}}|-k+1}\kappa_k \left((k-1)\beta-\sum_{j=1}^{k-1} l_j\right)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} \Psi_{k,\beta;\underline{l}}(((k-1)\beta-\sum_{j=1}^{k-1} l_j)\underline{v}),$$

where κ_k is a constant depending on k,

$$\Psi_{k,\beta;\underline{l}}(\underline{v}) = \int\cdots\int_{\substack{\underline{y}=(y_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}} \\ (\tilde{\ddagger})}} \exp\left(\sum y_{m,n}(v_n-v_m)^{\pm}\right) \prod dy_{m,n},$$

and $(\tilde{\ddagger})$ denotes normalised constraints related to those previously denoted $(\star)$, see proof for more details.

Using lemma 3.3.4 and lemma 3.3.5 we can prove the following which establishes out the power of N we seek.

Lemma 3.3.6.

$$\mathrm{MoM}_{U(N)}(k,\beta) \sim \gamma_{k,\beta} N^{k^2\beta^2-k+1}$$

where

$$\gamma_{k,\beta} = \sum_{l_1,\dots,l_{k-1}=0}^{2\beta} c_{k,\beta;\underline{l}}((k-1)\beta-\sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} P_{k,\beta}(l_1,\dots,l_{k-1}),$$

$c_{k,\beta;\underline{l}}$ is some constant depending on $k,\beta,\underline{l}$, and

$$P_{k,\beta}(l_1,\dots,l_{k-1}) = \frac{(-1)^{g(k,\beta;\underline{l})}}{(2\pi i)^{2k\beta}((k\beta)!)^2}$$
$$\times \int_{\Gamma_0}\cdots\int_{\Gamma_0} \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n-v_m)^2}{\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} (v_n-v_m) \prod_{m=1}^{2k\beta} v_m^{2\beta}} \Psi_{k,\beta;\underline{l}}(((k-1)\beta-\sum_{j=1}^{k-1} l_j)\underline{v}) \prod_{m=1}^{2k\beta} dv_m,$$

with $\Psi_{k,\beta;\underline{l}}(\underline{v})$ as defined in lemma 3.3.5, and $g(k,\beta;\underline{l})$ given by (3.3.59).

The last step is to prove that we do indeed have the correct asymptotic, which is achieved through the final lemma.

Lemma 3.3.7. *For $k,\beta\in\mathbb{N}$, $\gamma_{k,\beta}\neq 0$ where $\gamma_{k,\beta}$ is as defined in lemma 3.3.6.*

Hence combining lemma 3.3.6 and lemma 3.3.7 gives us theorem 3.1.4,

$$\mathrm{MoM}_{U(N)}(k,\beta) = \gamma_{k,\beta} N^{k^2\beta^2-k+1} + O(N^{k^2\beta^2-k}). \tag{3.3.43}$$

3.3.2 Details of the proof of theorem 3.1.4

Proof of lemma 3.3.2. We recall the statement of the lemma.

Lemma. *Let a choice of contours in*

$$J_{k,\beta}(\underline{\theta};\varepsilon_1,\ldots,\varepsilon_{2k\beta}) = \int_{\Gamma_{-i\theta_{\varepsilon_1}}} \cdots \int_{\Gamma_{-i\theta_{\varepsilon_{2k\beta}}}} \frac{e^{-N(z_{k\beta+1}+\cdots+z_{2k\beta})}\Delta(z_1,\ldots,z_{2k\beta})^2 dz_1\cdots dz_{2k\beta}}{\prod_{m\le k\beta<n}(1-e^{z_n-z_m})\prod_{m=1}^{2k\beta}\prod_{n=1}^{k}(z_m+i\theta_n)^{2\beta}}$$

be denoted by $\underline{\varepsilon} = (\varepsilon_1,\ldots,\varepsilon_{2k\beta})$ *where* $\varepsilon_j \in \{1,\ldots,k\}$*. If any one of the* k *poles is overrepresented in* $\underline{\varepsilon}$ *(i.e. some pole* $-i\theta^*$*,* $\theta^* \in \{\theta_1,\ldots,\theta_k\}$*, features in at least* $2\beta+1$ *contours), then* $J_{k,\beta}(\underline{\theta};\underline{\varepsilon})$ *is identically zero for that choice of* $\underline{\varepsilon}$*.*

The proof closely follows the proof of lemma 4.11 in [113]. We show the case where $-i\theta_1$ is 'over-represented', and without loss of generality assume that the choice of contour is given by

$$\underline{\varepsilon}^* = (\overbrace{1,\ldots,1}^{2\beta+1},\overbrace{2,\ldots,2}^{2\beta-1},\overbrace{3,\ldots,3}^{2\beta},\ldots,\overbrace{k,\ldots,k}^{2\beta}). \tag{3.3.44}$$

The other cases follow similarly.

To $J_{k,\beta}(\underline{\theta},\underline{\varepsilon}^*)$ apply the change of variable $z_j \mapsto z_j - i\theta_1$ and consider the function,

$$G(z_1,\ldots,z_{2\beta+1}) \coloneqq \frac{e^{-N(z_{k\beta+1}+\cdots+z_{2k\beta})}\Delta(z_1,\ldots,z_{2k\beta})}{\prod_{m\le k\beta<n}(1-e^{z_n-z_m})\prod_{m=1}^{2k\beta}\prod_{n=2}^{k}(z_m+i(\theta_n-\theta_1))^{2\beta}\prod_{m=2\beta+2}^{2k\beta} z_m^{2\beta}}, \tag{3.3.45}$$

which is analytic around zero. The integrand of $J_{k,\beta}(\underline{\theta};\underline{\varepsilon}^*)$ is

$$e^{iNk\beta\theta_1}\frac{G(z_1,\ldots,z_{2\beta+1})\Delta(z_1,\ldots,z_{2k\beta})dz_1\cdots dz_{2k\beta}}{\prod_{m=1}^{2\beta+1} z_m^{2\beta}}.$$

We appeal to the residue theorem to compute $J_{k,\beta}(\underline{\theta};\underline{\varepsilon}^*)$, and the proof follows if we can show that the coefficient of $\prod_{m=1}^{2\beta+1} z_m^{2\beta-1}$ in $G(z_1,\ldots,z_{2\beta+1})\Delta(z_1,\ldots,z_{2k\beta})$ is zero. Since $G(z_1,\ldots,z_{2\beta+1})$ is analytic around zero, we focus on the Vandermonde determinant and use the following expansion,

$$\Delta(z_1,\ldots,z_{2k\beta}) = \sum_{\sigma\in S_{2k\beta}} \mathrm{sgn}(\sigma)\prod_{m=1}^{2k\beta} z_m^{\sigma(m)-1}.$$

Thus, we are searching for terms in this expansion of the form $\prod_{m=1}^{2\beta+1} z_m^{\sigma(m)-1}$ with $\sigma(m)-1\le 2\beta-1$ for $m=1,\ldots,2\beta+1$. However, there is no term of this form as σ is a permutation on the set $\{1,\ldots,2k\beta\}$, so for at least one $m\in\{1,\ldots,2\beta+1\}$, $\sigma(m)\ge 2\beta+1$. By the residue theorem we conclude that $J_{k,\beta}(\underline{\theta};\underline{\varepsilon}^*)$ is zero. ☐

Proof of lemma 3.3.3. We recall the statement of the lemma.

Lemma.

$$\prod_{(m,n)\in B_{k,\beta;\underline{l}}} \left(1-\exp\left(\frac{v_n-v_m}{N}\right)\exp\left(i(\alpha_m-\alpha_n)\right)\right)^{-1}$$
$$= (-1)^{g(k,\beta;\underline{l})} \sum_{\substack{t_{\sigma,\tau}=-\infty \\ for\ 1\le\sigma<\tau\le k}}^{\infty} \exp\left(i\sum_{\gamma<\rho}(\theta_\rho-\theta_\gamma)(t_{\gamma,\rho}+|S^-_{\gamma,\rho}|)\right) \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right),$$

where

$$(v_n-v_m)^{\pm} = \begin{cases} v_n-v_m & \text{if } (m,n)\in S^+_{\sigma,\tau} \\ v_m-v_n & \text{if } (m,n)\in S^-_{\sigma,\tau}, \end{cases} \tag{3.3.46}$$

the first sum appearing in the right hand size of the statement of the lemma should be read as

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ for\ 1\le\sigma<\tau\le k}}^{\infty} = \sum_{t_{1,2}=-\infty}^{\infty}\sum_{t_{1,3}=-\infty}^{\infty}\cdots\sum_{t_{1,k}=-\infty}^{\infty}\sum_{t_{2,3}=-\infty}^{\infty}\cdots\sum_{t_{k-1,k}=-\infty}^{\infty} \tag{3.3.47}$$

and the second sum is over vectors (of weights) $\underline{x}=(x_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}}$ *subject to constraints given by* $(\star)$*, which are*

$$\begin{aligned} \sum_{(m,n)\in S_{\sigma,\tau}} x_{m,n} &= t_{\sigma,\tau}+|S^-_{\sigma,\tau}|, && 1\le\sigma<\tau\le k \\ x_{m,n} &\in \mathbb{Z}, && \forall (m,n)\in B_{k,\beta;\underline{l}} \\ H(-x_{m,n}\operatorname{Re}\{(v_n-v_m)^{\pm}\}) &= 1, && \forall (m,n)\in B_{k,\beta;\underline{l}}, \end{aligned}$$

where $\mathrm{H}(x)$ *is the Heaviside step function (so* $\mathrm{H}(x)=1$ *if* $x\ge 0$ *and* $\mathrm{H}(x)=0$ *if* $x<0$*). Furthermore, the prefactor can be expressed as follows,*

$$(-1)^{g(k,\beta;\underline{l})} = (-1)^{\sum_{\sigma<\tau}|S^-_{\sigma,\tau}|} \prod_{\substack{(m,n)\in S_{\sigma,\tau} \\ 1\le\sigma<\tau\le k}} (-\operatorname{sgn}(\operatorname{Re}\{(v_n-v_m)^{\pm}\})).$$

Firstly, recall the definition of the sets $B_{k,\beta;\underline{l}}, S^+_{\sigma,\tau}$, and $S^-_{\sigma,\tau}$,

$$\begin{aligned} B_{k,\beta;\underline{l}} &:= \{(m,n): 1\le m\le k\beta<n\le 2k\beta, \alpha_m\ne\alpha_n\}, \\ S^+_{\sigma,\tau} &:= \{(m,n)\in B_{k,\beta;\underline{l}}: \alpha_m-\alpha_n=\theta_\tau-\theta_\sigma\}, \\ S^-_{\sigma,\tau} &:= \{(m,n)\in B_{k,\beta;\underline{l}}: \alpha_m-\alpha_n=\theta_\sigma-\theta_\tau\}. \end{aligned}$$

We use the partition of $B_{k,\beta;\underline{l}}$ by the sets $S^+_{\sigma,\tau}, S^-_{\sigma,\tau}$ (although not emphasised in the notation, these sets also depend on k,β, and $l_1,\ldots,l_{k-1}$) to break up the product appearing on the left hand side of the statement of the lemma as follows,

$$\prod_{(m,n)\in B_{k,\beta;\underline{l}}} \left(1-e^{\frac{v_n-v_m}{N}}e^{i(\alpha_m-\alpha_n)}\right)^{-1}$$

$$= \prod_{1\le\sigma<\tau\le k} \prod_{(m,n)\in S^+_{\sigma,\tau}} \left(1-e^{\frac{v_n-v_m}{N}}e^{i(\theta_\tau-\theta_\sigma)}\right)^{-1} \prod_{(p,q)\in S^-_{\sigma,\tau}} \left(1-e^{\frac{v_q-v_p}{N}}e^{i(\theta_\sigma-\theta_\tau)}\right)^{-1} \tag{3.3.48}$$

$$= \prod_{1\le\sigma<\tau\le k} \Bigg[e^{i(\theta_\tau-\theta_\sigma)|S^-_{\sigma,\tau}|} \prod_{(m,n)\in S^+_{\sigma,\tau}} \left(1-e^{\frac{v_n-v_m}{N}}e^{i(\theta_\tau-\theta_\sigma)}\right)^{-1}$$

$$\times \prod_{(p,q)\in S^-_{\sigma,\tau}} \left(e^{i(\theta_\tau-\theta_\sigma)}-e^{\frac{v_q-v_p}{N}}\right)^{-1}\Bigg] \tag{3.3.49}$$

$$= \prod_{1\le\sigma<\tau\le k} \Bigg[(-1)^{|S^-_{\sigma,\tau}|} e^{i(\theta_\tau-\theta_\sigma)|S^-_{\sigma,\tau}|} \prod_{(p,q)\in S^-_{\sigma,\tau}} e^{\frac{v_p-v_q}{N}}$$

$$\times \prod_{(m,n)\in S_{\sigma,\tau}} \left(1-e^{\frac{1}{N}(v_n-v_m)^\pm}e^{i(\theta_\tau-\theta_\sigma)}\right)^{-1}\Bigg] \tag{3.3.50}$$

where

$$(v_n-v_m)^\pm = \begin{cases} v_n-v_m & \text{for } (m,n)\in S^+_{\sigma,\tau} \\ v_m-v_n & \text{for } (m,n)\in S^-_{\sigma,\tau}. \end{cases}$$

or equivalently,

$$(v_n-v_m)^\pm = \begin{cases} v_n-v_m & \text{if } m\le k\beta<n \text{ and } \varepsilon_m>\varepsilon_n \\ v_m-v_n & \text{if } m\le k\beta<n \text{ and } \varepsilon_m<\varepsilon_n, \end{cases}$$

which is equivalent to asking if $(m,n)\in S^+_{\sigma,\tau}$ or $S^-_{\sigma,\tau}$ respectively if we recall that

$$\underline{\varepsilon} = (\overbrace{1,\ldots,1}^{l_1},\overbrace{2,\ldots,2}^{l_2},\ldots,\overbrace{k-1,\ldots,k-1}^{l_{k-1}},\overbrace{k,\ldots,k}^{2\beta},\overbrace{k-1,\ldots,k-1}^{2\beta-l_{k-1}},\ldots,\overbrace{1,\ldots,1}^{2\beta-l_1}). \tag{3.3.51}$$

For a fixed choice σ,τ, and a fixed pair (m,n), we use the following expansion

$$\begin{aligned}(1-&\exp(\tfrac{1}{N}(v_n-v_m)^\pm)\exp(i(\theta_\tau-\theta_\sigma)))^{-1}\\ &= -\operatorname{sgn}(\operatorname{Re}\{(v_n-v_m)^\pm\})\\ &\quad\times\sum_{t=-\infty}^{\infty}\exp(\tfrac{1}{N}(v_n-v_m)^\pm t)\exp(i(\theta_\tau-\theta_\sigma)t)\,\mathrm{H}(-t\operatorname{Re}\{(v_n-v_m)^\pm\}),\end{aligned} \tag{3.3.52}$$

where $\mathrm{H}(x)$ is the Heaviside step function (so $\mathrm{H}(x)=1$ if $x\ge0$ and $\mathrm{H}(x)=0$ if $x<0$). The above equality holds since when $\operatorname{Re}\{(v_n-v_m)^\pm\}$ is strictly negative, one can view the left hand side of (3.3.52) as the evaluation of the following geometric series,

$$\left(1-\exp(\tfrac{1}{N}(v_n-v_m)^\pm)\exp(i(\theta_\tau-\theta_\sigma))\right)^{-1} = \sum_{s=0}^{\infty}\exp(\tfrac{1}{N}(v_n-v_m)^\pm s)\exp(i(\theta_\tau-\theta_\sigma)s). \tag{3.3.53}$$

With $\operatorname{Re}\{(v_n-v_m)^\pm\}$ in this range, the Heaviside function in (3.3.52) kills all negative values of t, and the resulting expression is precisely the right hand side of (3.3.53). Otherwise, if $\operatorname{Re}\{(v_n-v_m)^\pm\}$ is strictly positive (by assumption it cannot be zero), then the obvious manipulation of the left hand side of (3.3.52) is

$$\frac{1}{1-\exp(\frac{1}{N}(v_n-v_m)^{\pm})\exp(i(\theta_\tau-\theta_\sigma))} = -\frac{\exp(-\frac{1}{N}(v_n-v_m)^{\pm})\exp(-i(\theta_\tau-\theta_\sigma))}{1-\exp(-\frac{1}{N}(v_n-v_m)^{\pm})\exp(-i(\theta_\tau-\theta_\sigma))} \tag{3.3.54}$$

$$= -\sum_{s=1}^{\infty}\exp(-\tfrac{1}{N}(v_n-v_m)^{\pm}s)\exp(-i(\theta_\tau-\theta_\sigma)s). \tag{3.3.55}$$

When $\mathrm{Re}\{(v_n-v_m)^{\pm}\}$ is strictly positive, then the only range of t which survives in the right hand side of (3.3.52) once more matches the summation range of (3.3.55).

Now, incorporating (3.3.52) into the final product of (3.3.50), we have

$$\prod_{(m,n)\in B_{k,\beta;\underline{l}}} \left(1-\exp\left(\tfrac{1}{N}(v_n-v_m)\right)\exp\left(i(\alpha_m-\alpha_n)\right)\right)^{-1}$$

$$= \prod_{1\le\sigma<\tau\le k}\Bigg[(-1)^{|S^-_{\sigma,\tau}|}\exp(i(\theta_\tau-\theta_\sigma)|S^-_{\sigma,\tau}|)\prod_{(p,q)\in S^-_{\sigma,\tau}}\exp\left(\frac{v_p-v_q}{N}\right)$$
$$\times\prod_{(m,n)\in S_{\sigma,\tau}}\left(-\,\mathrm{sgn}(\mathrm{Re}\{(v_n-v_m)^{\pm}\})\right)$$
$$\times\prod_{(m,n)\in S_{\sigma,\tau}}\sum_{t=-\infty}^{\infty}\exp\left(\tfrac{1}{N}(v_n-v_m)^{\pm}t\right)\exp(i(\theta_\tau-\theta_\sigma)t)\,\mathrm{H}(-t\,\mathrm{Re}\{(v_n-v_m)^{\pm}\})\Bigg] \tag{3.3.56}$$

$$= (-1)^{g(k,\beta;\underline{l})}\prod_{1\le\sigma<\tau\le k}\Bigg[\sum_{t_{\sigma,\tau}=-\infty}^{\infty}\exp\left(i(\theta_\tau-\theta_\sigma)(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)\right)$$
$$\times\sum_{\substack{\underline{x}=(x_{m,n})\\ \sum_{(m,n)\in S_{\sigma,\tau}}x_{m,n}=t_{\sigma,\tau}+|S^-_{\sigma,\tau}|}}\prod_{(m,n)\in S_{\sigma,\tau}}\exp\left(\tfrac{1}{N}(v_n-v_m)^{\pm}x_{m,n}\right)\mathrm{H}(-x_{m,n}\,\mathrm{Re}\{(v_n-v_m)^{\pm}\})\Bigg] \tag{3.3.57}$$

$$= (-1)^{g(k,\beta;\underline{l})}\sum_{\substack{t_{\sigma,\tau}=-\infty\\ \text{for } 1\le\sigma<\tau\le k}}^{\infty}\exp\left(i\sum_{\gamma<\rho}(\theta_\rho-\theta_\gamma)(t_{\gamma,\rho}+|S^-_{\gamma,\rho}|)\right)\sum_{\substack{(x_{m,n})\\ (\star)}}\exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right). \tag{3.3.58}$$

The overall sign in (3.3.58) is

$$(-1)^{g(k,\beta;\underline{l})} = (-1)^{\sum_{\sigma<\tau}|S^-_{\sigma,\tau}|}\prod_{\substack{(m,n)\in S_{\sigma,\tau}\\ 1\le\sigma<\tau\le k}}\left(-\,\mathrm{sgn}(\mathrm{Re}\{(v_n-v_m)^{\pm}\})\right), \tag{3.3.59}$$

the first sum in (3.3.58) should be read as

$$\sum_{\substack{t_{\sigma,\tau}=-\infty\\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} = \sum_{t_{1,2}=-\infty}^{\infty}\sum_{t_{1,3}=-\infty}^{\infty}\cdots\sum_{t_{1,k}=-\infty}^{\infty}\sum_{t_{2,3}=-\infty}^{\infty}\cdots\sum_{t_{k-1,k}=-\infty}^{\infty},$$

and the second sum is now over the ‘full’ vectors $\underline{x} = (x_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}}$ whose elements $x_{m,n}$ are integers subject to constraints given by $(\star)$, which are

$$\sum_{(m,n)\in S_{\sigma,\tau}} x_{m,n} = t_{\sigma,\tau} + |S^-_{\sigma,\tau}|, \quad 1 \le \sigma < \tau \le k \tag{3.3.60}$$

$$x_{m,n} \in \mathbb{Z}, \qquad \forall (m,n) \in B_{k,\beta;\underline{l}} \tag{3.3.61}$$

$$H(-x_{m,n}\operatorname{Re}\{(v_n - v_m)^{\pm}\}) = 1, \qquad \forall (m,n) \in B_{k,\beta;\underline{l}}. \tag{3.3.62}$$

This means that the vector $\underline{x}$ should be thought of as being made up of concatenated subsequences $(x_{m,n})_{(m,n)\in S_{\sigma,\tau}}$ for each $1 \le \sigma < \tau \le k$, and each subsequence must satisfy the constraints (3.3.60)–(3.3.62). This completes the proof of lemma 3.3.3.

□

Proof of lemma 3.3.4. Firstly, recall the statement of the lemma.

Lemma.

$$\begin{aligned}\operatorname{MoM}_{U(N)}(k,\beta) \sim &\sum_{l_1,\ldots,l_{k-1}=0}^{2\beta} \frac{\tilde{c}_{\underline{l}}(k,\beta) N^{|A_{k,\beta;\underline{l}}|}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \\ &\times \int_{\Gamma_0} \cdots \int_{\Gamma_0} f(\underline{v};\underline{l}) \sum_{\substack{t_{\sigma,\tau}=-\infty \\ for\ 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right) \\ &\times \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)} \prod_{m=1}^{2k\beta} dv_m.\end{aligned}$$

We begin with lemma 3.3.42,

$$\begin{aligned}I_{k,\beta}(\underline{\theta}) \sim &\sum_{l_1,\ldots,l_{k-1}=0}^{2\beta} \frac{\tilde{c}_{\underline{l}}(k,\beta) N^{|A_{k,\beta;\underline{l}}|}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \\ &\times \int_{\Gamma_0} \cdots \int_{\Gamma_0} f(\underline{v};\underline{l}) \exp\left(iN\Big(\sum_{m=k\beta+1}^{2k\beta} \alpha_m - \beta\sum_{j=1}^{k}\theta_j\Big)\right) \\ &\times \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \exp\left(i\sum_{1\le\gamma<\rho\le k}(\theta_\rho - \theta_\gamma)(t_{\gamma,\rho} + |S^-_{\gamma,\rho}|)\right) \\ &\times \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right) \prod_{m=1}^{2k\beta} dv_m,\end{aligned}$$

and use the structure of the α_m to deduce that

$$\exp\left(iN\Big(\sum_{m=k\beta+1}^{2k\beta} \alpha_m - \beta\sum_{j=1}^{k}\theta_j\Big)\right) = \exp\left(iN\sum_{j=1}^{k-1}(\beta - l_j)(\theta_j - \theta_k)\right). \tag{3.3.63}$$

Combining (3.3.42), (3.3.2), (3.3.63) and switching the order of integration we have that

$$\mathrm{MoM}_{U(N)}(k,\beta) \sim \sum_{l_1,\dots,l_{k-1}=0}^{2\beta} \frac{\tilde{c}_{\underline{l}}(k,\beta) N^{|A_{k,\beta;\underline{l}}|}}{(2\pi)^k (2\pi i)^{2k\beta}((k\beta)!)^2}$$

$$\times \int_{\Gamma_0} \cdots \int_{\Gamma_0} \Bigg[f(\underline{v};\underline{l}) \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\frac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right)$$

$$\times \int_0^{2\pi} \cdots \int_0^{2\pi} \exp\left(iN \sum_{j=1}^{k-1} (\beta - l_j)(\theta_j - \theta_k)\right)$$

$$\times \exp\left(i \sum_{\gamma<\rho} (\theta_\rho - \theta_\gamma)(t_{\gamma,\rho} + |S^-_{\gamma,\rho}|)\right) \prod_{n=1}^{k} d\theta_n \Bigg] \prod_{m=1}^{2k\beta} dv_m. \quad (3.3.64)$$

By noting that $\theta_\rho - \theta_\gamma = \theta_k - \theta_\gamma - (\theta_k - \theta_\rho)$, we now see, importantly, that the θ integral will just be a function of differences $(\theta_j - \theta_k)$, $j \in \{1, \dots, k-1\}$. Focussing on the inner integral in (3.3.64) we have

$$\sum_{\sigma<\tau} (\theta_\tau - \theta_\sigma)(|S^-_{\sigma,\tau}| + t_{\sigma,\tau})$$

$$= \sum_{\sigma<\tau} (\theta_k - \theta_\sigma)(|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) - \sum_{\sigma<\tau} (\theta_k - \theta_\tau)(|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) \quad (3.3.65)$$

$$= \left(\sum_{\sigma=1}^{k-1} (\theta_k - \theta_\sigma) \sum_{\tau=\sigma+1}^{k} (|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) \right) - \left(\sum_{\tau=2}^{k-1} (\theta_k - \theta_\tau) \sum_{\sigma=1}^{\tau-1} (|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) \right) \quad (3.3.66)$$

$$= (\theta_k - \theta_1) \sum_{\tau=2}^{k} (|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) + \sum_{n=2}^{k-1} (\theta_k - \theta_n) \sum_{\tau=n+1}^{k} (|S^-_{n,\tau}| + t_{n,\tau})$$
$$- \sum_{n=2}^{k-1} (\theta_k - \theta_n) \sum_{\sigma=1}^{n-1} (|S^-_{\sigma,n}| + t_{\sigma,n}) \quad (3.3.67)$$

$$= (\theta_k - \theta_1) \sum_{\tau=2}^{k} (|S^-_{\sigma,\tau}| + t_{\sigma,\tau}) + \sum_{n=2}^{k-1} (\theta_k - \theta_n) \left(\sum_{\tau=n+1}^{k} (|S^-_{n,\tau}| + t_{n,\tau}) - \sum_{\sigma=1}^{n-1} (|S^-_{\sigma,n}| + t_{\sigma,n}) \right) \quad (3.3.68)$$

$$= \sum_{n=1}^{k-1} (\theta_k - \theta_n) \left(\sum_{\tau=n+1}^{k} (|S^-_{n,\tau}| + t_{n,\tau}) - \sum_{\sigma=1}^{n-1} (|S^-_{\sigma,n}| + t_{\sigma,n}) \right). \quad (3.3.69)$$

Thus, inserting (3.3.69) into (3.3.64) we can evaluate the integral over the $\theta_1, \dots, \theta_k$ to find

$$\int_0^{2\pi} \cdots \int_0^{2\pi} \exp\left(i \sum_{j=1}^{k-1} (\theta_k - \theta_j) \left(N(l_j - \beta) + \sum_{\rho=j+1}^{k} (t_{j,\rho} + |S^-_{j,\rho}|) - \sum_{\gamma=1}^{j-1} (t_{\gamma,j} + |S^-_{\gamma,j}|) \right)\right) \prod_{n=1}^{k} d\theta_n$$

$$= (2\pi)^k \prod_{j=1}^{k-1} \delta_{N(l_j-\beta) + \sum_{\rho=j+1}^{k} (t_{j,\rho} + |S^-_{j,\rho}|) - \sum_{\gamma=1}^{j-1} (t_{\gamma,j} + |S^-_{\gamma,j}|)}, \quad (3.3.70)$$

where the δ is a Kronecker δ-function. Considering this in the context of (3.3.64) we have the result. □

Proof of lemma 3.3.5. We restate the claim for context.

Lemma.

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right) \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)}$$

$$\sim N^{|B_{k,\beta;\underline{l}}|-k+1}\kappa_k \left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} \Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)\underline{v}\right),$$

where κ_k is a constant depending on k,

$$\Psi_{k,\beta;\underline{l}}(\underline{v}) = \int\cdots\int_{\substack{\underline{y}=(y_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}} \\ (\tilde{\ddagger})}} \exp\left(\sum y_{m,n}(v_n-v_m)^{\pm}\right)\prod dy_{m,n},$$

and $(\tilde{\ddagger})$ denotes normalised versions of constraints $(\star)$ as described in the proof of lemma 3.3.3, more details given in the proof.

Recall that the first sum in the left hand side of the statement of the lemma should be interpreted as

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} = \sum_{t_{1,2}=-\infty}^{\infty}\sum_{t_{1,3}=-\infty}^{\infty}\cdots\sum_{t_{1,k}=-\infty}^{\infty}\sum_{t_{2,3}=-\infty}^{\infty}\cdots\sum_{t_{k-1,k}=-\infty}^{\infty},$$

and the second sum runs over vectors $\underline{x} = (x_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}} = (x_{m,n})_{(m,n)\in\bigcup_{\sigma<\tau} S_{\sigma,\tau}}$ with integer elements subject to the following constraints

$$\sum_{(m,n)\in S_{1,2}} x_{m,n} = t_{1,2} + |S^-_{1,2}| \tag{3.3.71}$$

$$\sum_{(m,n)\in S_{1,3}} x_{m,n} = t_{1,3} + |S^-_{1,3}| \tag{3.3.72}$$

$$\vdots$$

$$\sum_{(m,n)\in S_{1,k}} x_{m,n} = t_{1,k} + |S^-_{1,k}| \tag{3.3.73}$$

$$\sum_{(m,n)\in S_{2,3}} x_{m,n} = t_{2,3} + |S^-_{2,3}| \tag{3.3.74}$$

$$\vdots$$

$$\sum_{(m,n)\in S_{k-1,k}} x_{m,n} = t_{k-1,k} + |S^-_{k-1,k}| \tag{3.3.75}$$

$$H(-x_{m,n}\operatorname{Re}\{(v_n-v_m)^{\pm}\}) = 1, \quad \forall (m,n)\in B_{k,\beta;\underline{l}}. \tag{3.3.76}$$

We now focus on the product of δ-functions in the left hand side of the statement of the lemma,

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{\underline{x} \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right) \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)}, \tag{3.3.77}$$

which constrain (3.3.77) to be zero unless the following hold,

$$\sum_{j=2}^{k}(t_{1,j}+|S^-_{1,j}|) = N(\beta - l_1) \tag{3.3.78}$$

$$\sum_{j=3}^{k}(t_{2,j}+|S^-_{2,j}|) - (t_{1,2}+|S^-_{1,2}|) = N(\beta - l_2) \tag{3.3.79}$$

$$\vdots$$

$$(t_{k-2,k-1}+|S^-_{k-2,k-1}|+t_{k-2,k}+|S^-_{k-2,k}|) - \sum_{j=1}^{k-3}(t_{j,k-2}+|S^-_{j,k-2}|) = N(\beta - l_{k-2}) \tag{3.3.80}$$

$$(t_{k-1,k}+|S^-_{k-1,k}|) - \sum_{j=1}^{k-2}(t_{j,k-1}+|S^-_{j,k-1}|) = N(\beta - l_{k-1}). \tag{3.3.81}$$

These conditions form an underdetermined system of linear equations. There are $\binom{k}{2}$ variables and $k-1$ equations, hence $\binom{k}{2}-(k-1)$ free parameters. We eliminate the $k-1$ dependent variables from (3.3.77) using the linear equations (3.3.78)-(3.3.81). The following rewriting of the system picks out which we choose to discard.

$$t_{1,k} = N(\beta - l_1) - \sum_{j=2}^{k-1}(t_{1,j}+|S^-_{1,j}|) - |S^-_{1,k}| \tag{3.3.82}$$

$$t_{2,k} = N(\beta - l_2) + t_{1,2} + |S^-_{1,2}| - |S^-_{2,k}| - \sum_{j=3}^{k-1}(t_{2,j}+|S^-_{2,j}|) \tag{3.3.83}$$

$$\vdots$$

$$t_{k-2,k} = N(\beta - l_{k-2}) + \sum_{j=1}^{k-3}(t_{j,k-2}+|S^-_{j,k-2}|) - (t_{k-2,k-1}+|S^-_{k-2,k-1}|+|S^-_{k-2,k}|) \tag{3.3.84}$$

$$t_{k-1,k} = N(\beta - l_{k-1}) + \sum_{j=1}^{k-2}(t_{j,k-1}+|S^-_{j,k-1}|) - |S^-_{k-1,k}|. \tag{3.3.85}$$

Thus, the $k-1$ outer sums over $t_{1,k},\ldots,t_{k-1,k}$ in (3.3.77) collapse and we are left with

$$\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k-1}}^{\infty} \sum_{\substack{x \\ (\ddagger)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n - v_m)^{\pm}\right), \tag{3.3.86}$$

where $(\ddagger)$ reflects constraints on $(x_{m,n})$ as before, with (3.3.82)-(3.3.85) substituted in. Since the Heaviside function $H(-x_{m,n}\,\mathrm{Re}\{(v_n - v_m)^{\pm}\})$ is equal to 1 for all $(m,n)\in B_{k,\beta;\underline{l}}$, this sum converges exponentially quickly.

Summing over all weights $x_{m,n}$ and using (3.3.82)-(3.3.85), we have that

$$\sum_{(m,n)\in B_{k,\beta;\underline{l}}} x_{m,n} = N((k-1)\beta - \sum_{j=1}^{k-1} l_j) + \sum_{1\le\sigma<\tau\le k-1}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|). \tag{3.3.87}$$

Clearly if any of the $t_{\sigma,\tau}$ grow faster than N, then there must be a least one weight $x^*_{m,n}$ having the same growth rate. So for large $t_{\sigma,\tau}$, the summands in (3.3.87) will not contribute to the leading order.

In order to pull out the correct power of N, we employ the following general lemma about geometric sums (see Keating et al. [113], lemma 4.12).

Lemma 3.3.8. *As* $K \to \infty$*,*

$$\sum_{\substack{k_1+\cdots+k_d=K\\k_i\geq 0}} \exp\left(\frac{1}{K}\sum_{i=1}^{d} k_j z_j\right) = K^{d-1} \int\cdots\int_{\substack{x_1+\cdots+x_d=1\\x_j\geq 0}} e^{\sum x_j z_j} dx_1\cdots dx_d + O(K^{d-2}).$$

From this we deduce that the leading power of K in the left hand side is given by the dimension of the space described by the weights k_j subject to any rules placed upon them. Within lemma 3.3.8, one can think of the weights as forming a d-dimensional vector where the sum of the elements must equal K. Thus one has $d-1$ degrees of freedom in choosing the vector elements.

This means that the information we need to extract from the constraints given by (‡) in (3.3.86) is the dimension of the space spanned by the vector $\underline{x} = (x_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}}$ subject to those restrictions. This will give us the claimed power of N in the statement of lemma 3.3.5.

Before we apply lemma 3.3.8, we first incorporate $\sum_{\sigma<\tau}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)$ into one of the weights using (3.3.87). To do this, pick one weight, say $x_{1,2}$, and shift it by $\sum_{1\leq\sigma<\tau\leq k-1}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)$ to get an equivalent form of the inner sum in the right hand side of (3.3.86),

$$\begin{aligned}\sum_{\substack{\underline{x}\\(\ddagger)}} \exp&\left(\tfrac{1}{N}\sum_{\substack{(m,n)\in B_{k,\beta;\underline{l}}\\(m,n)\neq(1,2)}} x_{m,n}(v_n-v_m)^{\pm} + \tfrac{1}{N}\big(x_{1,2}+\sum_{\sigma<\tau}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)\big)(v_2-v_1)^{\pm}\right)\\ &= \exp\left(\tfrac{1}{N}\sum_{\sigma<\tau}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)(v_2-v_1)^{\pm}\right)\sum_{\substack{\underline{x}\\(\ddagger')}}\exp\left(\tfrac{1}{N}\sum_{(m,n)\in B_{k,\beta;\underline{l}}} x_{m,n}(v_n-v_m)^{\pm}\right),\end{aligned} \tag{3.3.88}$$

The constraints (‡′) can be deduced from (‡) by applying the described shift to the weights. In particular, this now means (3.3.87) becomes

$$\sum_{(m,n)\in B_{k,\beta;\underline{l}}} x_{m,n} = N((k-1)\beta - \sum_{j=1}^{k-1} l_j). \tag{3.3.89}$$

Now, we apply lemma 3.3.8 to the sum term on the right hand side of (3.3.88), with $K = N((k-1)\beta-\sum_{j=1}^{k-1} l_j)$. (The case when $(k-1)\beta = \sum_{j=1}^{k-1} l_j$ does not contribute at leading order, for reasons to be discussed at the end of the proof.) To determine the leading power of K, we count the amount of choice we have in choosing the weights $x_{m,n}$, for a fixed choice of $t_{\sigma,\tau}$, $1\leq\sigma<\tau\leq k-1$. From (3.3.82)-(3.3.85), we see that for each $\binom{k}{2}$ equation we lose one degree of freedom in choosing the weights, so in total, we have $|B_{k,\beta;\underline{l}}|-\binom{k}{2}$ degrees of freedom in determining $\underline{x}$. Thus, by lemma 3.3.8 we have

$$\begin{aligned}\sum_{\substack{\underline{x}\\(\ddagger')}}&\exp\left(\tfrac{1}{N}\sum_{(m,n)\in B_{k,\beta;\underline{l}}} x_{m,n}(v_n-v_m)^{\pm}\right)\\ &= \Big(N((k-1)\beta-\sum_{j=1}^{k-1} l_j)\Big)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}}\Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta-\sum_{j=1}^{k-1} l_j\right)\underline{v}\right) + O\big(N^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}-1}\big),\end{aligned} \tag{3.3.90}$$

where

$$\Psi_{k,\beta;\underline{l}}(\underline{v}) = \underset{\substack{\underline{y}=(y_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}} \\ (\tilde{\ddagger})}}{\int\cdots\int} \exp\left(\sum y_{m,n}(v_n - v_m)^{\pm}\right)\prod dy_{m,n},$$

and $(\tilde{\ddagger})$ denotes the normalised version of the constraints $(\ddagger')$ (since we need only consider $t_{\sigma,\tau}$ growing at most like N for each $1 \le \sigma < \tau \le k-1$, asymptotically the constraints $(\tilde{\ddagger})$ will be $O(1)$, and in particular will not depend on $t_{\sigma,\tau}$). We write $\Psi_{k,\beta;\underline{l}}(c\underline{v})$ for $\Psi_{k,\beta;\underline{l}}(cv_1,\ldots,cv_{2k,\beta})$.

Then, combining (3.3.90) with (3.3.88) and (3.3.86), we have

$$\begin{aligned}
&\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k-1}}^{\infty} \sum_{\substack{\underline{x} \\ (\ddagger)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right) \\
&\sim \left(N((k-1)\beta - \sum_{j=1}^{k-1} l_j)\right)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} \\
&\quad\times \sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k-1}}^{\infty} \Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)\underline{v}\right) \prod_{1\le\sigma<\tau\le k-1} \exp\left(\tfrac{1}{N}(t_{\sigma,\tau}+|S^-_{\sigma,\tau}|)(v_2-v_1)^{\pm}\right)
\end{aligned} \tag{3.3.91}$$

$$\sim \kappa_k N^{\binom{k}{2}-(k-1)}\left(N((k-1)\beta - \sum_{j=1}^{k-1} l_j)\right)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} \Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)\underline{v}\right), \tag{3.3.92}$$

for some constant κ_k depending on k. Note that the case where $\sum_{j=1}^{k-1} l_j = (k-1)\beta$ falls in to the subleading order terms.

Thus at leading order,

$$\begin{aligned}
&\sum_{\substack{t_{\sigma,\tau}=-\infty \\ \text{for } 1\le\sigma<\tau\le k}}^{\infty} \sum_{\substack{(x_{m,n}) \\ (\star)}} \exp\left(\tfrac{1}{N}\sum x_{m,n}(v_n-v_m)^{\pm}\right) \prod_{j=1}^{k-1} \delta_{N(l_j-\beta)+\sum_{\rho=j+1}^{k}(t_{j,\rho}+|S^-_{j,\rho}|)-\sum_{\gamma=1}^{j-1}(t_{\gamma,j}+|S^-_{\gamma,j}|)} \\
&\qquad \sim N^{|B_{k,\beta;\underline{l}}|-k+1}\kappa_k\left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} \Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta - \sum_{j=1}^{k-1} l_j\right)\underline{v}\right).
\end{aligned} \tag{3.3.93}$$

□

Proof of lemma 3.3.6. We restate the lemma.

Lemma.

$$\mathrm{MoM}_{U(N)}(k,\beta) \sim \gamma_{k,\beta} N^{k^2\beta^2-k+1}$$

where

$$\gamma_{k,\beta} = \sum_{l_1,\ldots,l_{k-1}=0}^{2\beta} c_{k,\beta;\underline{l}}((k-1)\beta - \sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} P_{k,\beta}(l_1,\ldots,l_{k-1}),$$

$c_{k,\beta;\underline{l}}$ *is a constant depending on* $k,\beta,l_1,\ldots,l_{k-1}$ *(more details given below), and*

$$\begin{aligned}
P_{k,\beta}(l_1,\ldots,l_{k-1}) &= \frac{(-1)^{g(k,\beta;\underline{l})}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \\
&\quad\times \int_{\Gamma_0}\cdots\int_{\Gamma_0} \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m}\prod_{\substack{m<n \\ \alpha_m=\alpha_n}}(v_n-v_m)^2}{\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}}(v_n-v_m)\prod_{m=1}^{2k\beta} v_m^{2\beta}} \Psi_{k,\beta;\underline{l}}(((k-1)\beta - \sum_{j=1}^{k-1} l_j)\underline{v}) \prod_{m=1}^{2k\beta} dv_m,
\end{aligned}$$

with $\Psi_{k,\beta;\underline{l}}(\underline{v})$ as defined in lemma 3.3.5, and $g(k,\beta;\underline{l})$ given by (3.3.59).

Recall the definition of the sets $A_{k,\beta;\underline{l}}$ and $B_{k,\beta;\underline{l}}$,

$$A_{k,\beta;\underline{l}} \coloneqq \{(m,n) : 1 \leq m \leq k\beta < n \leq 2k\beta, \alpha_m = \alpha_n\} \tag{3.3.94}$$

$$B_{k,\beta;\underline{l}} \coloneqq \{(m,n) : 1 \leq m \leq k\beta < n \leq 2k\beta, \alpha_m \neq \alpha_n\}, \tag{3.3.95}$$

so $|A_{k,\beta;\underline{l}}| + |B_{k,\beta;\underline{l}}| = k^2\beta^2$. Using this fact with lemma 3.3.4 and lemma 3.3.5 we have

$$\begin{aligned}\mathrm{MoM}_{U(N)}(k,\beta) \sim N^{k^2\beta^2-k+1} &\sum_{l_1,\dots,l_{k-1}=0}^{2\beta} \frac{(-1)^{g(k,\beta;\underline{l})} c_{k,\beta;\underline{l}}((k-1)\beta - \sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}}}{(2\pi i)^{2k\beta}((k\beta)!)^2} \\ &\times \int_{\Gamma_0} \cdots \int_{\Gamma_0} f(\underline{v};\underline{l}) \Psi_{k,\beta;\underline{l}}\Big(\Big((k-1)\beta - \sum_{j=1}^{k-1} l_j\Big)\underline{v}\Big) \prod_{m=1}^{2k\beta} dv_m,\end{aligned} \tag{3.3.96}$$

where $c_{k,\beta;\underline{l}}$ is a constant encompassing the two constants given in (3.3.7) and the statement of lemma 3.3.5,

Lemma 3.3.6 then follows recalling the definition of $f(\underline{v};\underline{l})$,

$$f(\underline{v};\underline{l}) = \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n - v_m)^2}{\prod_{\substack{m\leq k\beta<n \\ \alpha_m=\alpha_n}} (v_n - v_m) \prod_{m=1}^{2k\beta} v_m^{2\beta}},$$

and by setting $\gamma_{k,\beta}$ and $P_{k,\beta}(l_1,\dots,l_{k-1})$ as claimed.

□

Proof of lemma 3.3.7. Finally, recall the statement of lemma 3.3.7.

Lemma. *For $k,\beta \in \mathbb{N}$, $\gamma_{k,\beta} \neq 0$ where $\gamma_{k,\beta}$ is as defined in lemma 3.3.6.*

Thus, we have to show that

$$\sum_{l_1,\dots,l_{k-1}=0}^{2\beta} c_{k,\beta;\underline{l}}((k-1)\beta - \sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} P_{k,\beta}(l_1,\dots,l_{k-1}) \neq 0, \tag{3.3.97}$$

for $c_{k,\beta;\underline{l}}$ some constant depending on $k,\beta,\underline{l}$ and

$$\begin{aligned}P_{k,\beta}(l_1,\dots,l_{k-1}) = &\frac{(-1)^{g(k,\beta;\underline{l})}}{(2\pi i)^{2k\eta}((k\beta)!)^2} \\ &\times \int_{\Gamma_0} \cdots \int_{\Gamma_0} \frac{e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n - v_m)^2}{\prod_{\substack{m\leq k\beta<n \\ \alpha_m=\alpha_n}} (v_n - v_m) \prod_{m=1}^{2k\beta} v_m^{2\beta}} \Psi_{k,\beta;\underline{l}}(((k-1)\beta - \sum_{j=1}^{k-1} l_j)\underline{v}) \prod_{m=1}^{2k\beta} dv_m,\end{aligned} \tag{3.3.98}$$

Since $c_{k,\beta;\underline{l}}$ is a constant encompassing both (3.3.7) and the constant appearing in the statement of lemma 3.3.5, it is clearly non-zero and its sign is independent of the sum parameters $l_1,\dots,l_{k-1}$. Further, at leading order we need only consider parameters l_j such that $l_1 + \cdots + l_{k-1} \neq (k-1)\beta$. To prove the required result, we show that $((k-1)\beta - \sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}} P_{k,\beta}(\underline{l}) \neq 0$ (in fact, it is strictly positive) for such $l_1,\dots,l_{k-1}$. To do this, we will appeal to the residue theorem.

Fix a choice of $l_1,\dots,l_{k-1}$ in agreement with the various constraints. To show that $P_{k,\beta}(\underline{l})$ is non-zero, firstly denote the integrand in (3.3.98) by $q_{k,\beta}(l_1,\dots,l_{k-1})$. Then by the residue theorem we have to show that there is a term of the form $(v_1 \cdots v_{2k\beta})^{2\beta-1}$ with non-zero coefficient in the expansion of

$$q_{k,\beta}(l_1,\dots,l_{k-1}) \prod_{m=1}^{2k\beta} v_m^{2\beta} = \Psi_{k,\beta;\underline{l}}\Big(\big((k-1)\beta - \sum_{j=1}^{k-1} l_j\big)\underline{v}\Big) e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} \frac{\prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n - v_m)^2}{\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} (v_n - v_m)}. \quad (3.3.99)$$

Now, simplifying the product terms of the right hand side of (3.3.99),

$$\frac{\prod_{\substack{m<n \\ \alpha_m=\alpha_n}} (v_n - v_m)^2}{\prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} (v_n - v_m)} = \prod_{\substack{m\le k\beta<n \\ \alpha_m=\alpha_n}} (v_n - v_m) \prod_{n=1}^{k} \Delta(v_{\sum_{j=1}^{n-1} l_j+1},\dots,v_{\sum_{j=1}^{n} l_j})^2$$
$$\times \prod_{n=1}^{k} \Delta(v_{\sum_{j=1}^{n} l_j+2(k-n)\beta+1},\dots,v_{\sum_{j=1}^{n-1} l_j+2(k-(n-1))\beta})^2, \quad (3.3.100)$$

where $l_k = k\beta - (l_1 + \cdots + l_{k-1})$. We use the following expansion of the Vandermonde determinant,

$$\Delta(x_1,\dots,x_n)^2 = \sum_{\sigma,\tau\in S_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \prod_{i=1}^{n} x_i^{\sigma(i)+\tau(i)-2}. \quad (3.3.101)$$

From any of the terms appearing in the first product of Vandermonde determinants in the right hand side of (3.3.100), we find a term of the form

$$l_n! \prod_{i=\sum_{j=1}^{n-1} l_j+1}^{\sum_{j=1}^{n} l_j} v_i^{l_n-1}, \quad n\in\{1,\dots,k\}, \quad (3.3.102)$$

and similarly for any of the terms in the second product. Thus, the Vandermonde determinants collectively contribute a term of the form

$$\prod_{i=1}^{l_1} v_i^{l_1-1} \prod_{i=l_1+1}^{l_1+l_2} v_i^{l_2-1} \cdots \prod_{i=\sum_{j=1}^{k-1} l_j+1}^{k\beta} v_i^{l_k-1} \prod_{i=k\beta+1}^{\sum_{j=1}^{k-1} l_j+2\beta} v_i^{2\beta-l_k-1} \cdots \prod_{i=2(k-1)\beta+1+l_1}^{2k\beta} v_i^{2\beta-l_1-1}, \quad (3.3.103)$$

and this term has a strictly positive coefficient

$$\prod_{j=1}^{k} l_j!(2\beta-l_j)!. \quad (3.3.104)$$

One sees (3.3.104) as follows[6]. We are interested in determining the coefficient of terms of the form $(x_1\cdots x_n)^{n-1}$ in the square of the Vandermonde determinant,

$$\Delta(x_1,\dots,x_n)^2 = \sum_{\sigma,\tau\in S_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \prod_{i=1}^{n} x_i^{\sigma(i)+\tau(i)-2}. \quad (3.3.105)$$

Thus, we require that $\sigma(i)+\tau(i) = n+1$ for all $i\in\{1,\dots,n\}$, and in particular we want to show that this coefficient is strictly positive.

Immediately, we see that there will be $n!$ terms of the required form since fixing $\sigma(i)$ completely determines $\tau(i)$. Consider the bijection

$$\phi : \{1,\dots,n\} \to \{1,\dots,n\}, \quad i \mapsto n+1-i. \quad (3.3.106)$$

[6]The following calculation appeared in the appendix of [15].

The order of ϕ is 2 and if n is even, it has no fixed point, whereas if n is odd there is a unique fixed point $(n+1)/2$. Thus, $\phi \in S_n$ and it consists of $n/2$ transpositions if n is even, and $(n-1)/2$ transpositions if n is odd. Now set $\tau = \phi \circ \sigma$, so $\tau \in S_n$, and $\tau(i) = n+1-\sigma(i)$. Given σ, we have found our unique permutation. To determine the sign of τ, note that $\operatorname{sgn}(\tau) = \operatorname{sgn}(\phi)\operatorname{sgn}(\sigma)$, and

$$\operatorname{sgn}(\phi) = (-1)^{\lfloor \frac{n}{2} \rfloor} = \begin{cases} +1 & \text{if } n \equiv 0,1 \mod 4 \\ -1 & \text{if } n \equiv 2,3 \mod 4. \end{cases} \tag{3.3.107}$$

Thus, the coefficient of $(x_1 \cdots x_n)^{n-1}$ in $\Delta(x_1, \dots, x_n)^2$ is $\operatorname{sgn}(\phi)n!$. It now follows that the coefficient of

$$\prod_{i=1}^{l_1} v_i^{l_1-1} \prod_{i=l_1+1}^{l_1+l_2} v_i^{l_2-1} \cdots \prod_{i=\sum_{j=1}^{k-1} l_j+1}^{k\beta} v_i^{l_k-1} \prod_{i=k\beta+1}^{\sum_{j=1}^{k-1} l_j+2\beta} v_i^{2\beta-l_k-1} \cdots \prod_{i=2(k-1)\beta+1+l_1}^{2k\beta} v_i^{2\beta-l_1-1} \tag{3.3.108}$$

in

$$\prod_{n=1}^{k} \Delta(v_{\sum_{j=1}^{n-1} l_j+1}, \dots, v_{\sum_{j=1}^{n} l_j})^2 \prod_{n=1}^{k} \Delta(v_{\sum_{j=1}^{n} l_j+2(k-n)\beta+1}, \dots, v_{\sum_{j=1}^{n-1} l_j+2(k-(n-1))\beta})^2 \tag{3.3.109}$$

is given by

$$(-1)^{\sum_{j=1}^{k}\left(\lfloor \frac{l_j}{2} \rfloor + \lfloor \frac{2\beta-l_j}{2} \rfloor\right)} \prod_{j=1}^{k} l_j!(2\beta-l_j)! = (-1)^{k\beta+\sum_{j=1}^{k}\left(\lfloor \frac{l_j}{2} \rfloor + \lfloor \frac{-l_j}{2} \rfloor\right)} \prod_{j=1}^{k} l_j!(2\beta-l_j)! \tag{3.3.110}$$

$$= (-1)^{k\beta}(-1)^{\sum_{j=1}^{k} \delta_{\{l_j \text{ is odd}\}}} \prod_{j=1}^{k} l_j!(2\beta-l_j)! \tag{3.3.111}$$

$$= (-1)^{k\beta}(-1)^{\#\{j: l_j \text{ is odd}\}} \prod_{j=1}^{k} l_j!(2\beta-l_j)!. \tag{3.3.112}$$

This proves the result since the parity of $\#\{j : l_j \text{ is odd}\}$ is the same as the parity of $k\beta$ as $\sum_{j=1}^{k} l_j = k\beta$.

We now expand the remaining product in (3.3.100) as

$$\prod_{\substack{m \le k\beta < n \\ \alpha_m = \alpha_n}} (v_n - v_m) = \prod_{\substack{m \in \{1,\dots,l_1\} \\ n \in \{l_1+2(k-1)\beta+1,\dots,2k\beta\}}} (v_n - v_m) \cdots \prod_{\substack{m \in \{\sum_{j=1}^{k-1} l_j+1,\dots,k\beta\} \\ n \in \{k\beta+1,\dots,\sum_{j=1}^{k-1} l_j+2\beta\}}} (v_n - v_m). \tag{3.3.113}$$

From the first product in the right hand side of (3.3.113) we take the term $\prod_{i=1}^{l_1}(-v_i)^{2\beta-l_1}$. The second gives $\prod_{i=l_1+1}^{l_1+l_2}(-v_i)^{2\beta-l_2}$, and so on. Hence, in total from (3.3.100) we have a term of the form

$$(-1)^{k\beta} \prod_{j=1}^{k} \left(\prod_{i=\sum_{n=1}^{j-1} l_n+1}^{\sum_{n=1}^{j} l_n} v_i^{2\beta-1} \right) \prod_{j=1}^{k} \left(\prod_{i=\sum_{n=1}^{j} l_n+2(k-j)\beta+1}^{\sum_{n=1}^{j-1} l_n+2(k-(j-1))\beta} v_i^{2\beta-l_j-1} \right). \tag{3.3.114}$$

We now use the exponential function in (3.3.99) to give us the remaining contribution,

$$e^{-\sum_{m=k\beta+1}^{2k\beta} v_m} = \sum_{t=0}^{\infty} \frac{\left(-\sum_{m=k\beta+1}^{2k\beta} v_m\right)^t}{t!} \tag{3.3.115}$$

$$= \sum_{t=0}^{\infty} \frac{(-1)^t}{t!} \sum_{a_{k\beta+1}+\cdots+a_{2k\beta}=t} \binom{t}{a_{k\beta+1}, \dots, a_{2k\beta}} \prod_{i=k\beta+1}^{2k\beta} v_i^{a_i}, \tag{3.3.116}$$

where the multinomial coefficient is

$$\binom{n}{c_1, \ldots, c_m} = \frac{n!}{c_1! \cdots c_m!}. \tag{3.3.117}$$

To complete the construction of the term of the form $(v_1 \cdots v_{2k\beta})^{2\beta-1}$, we need

$$a_i = \begin{cases} l_k & \text{for } i \in \{k\beta+1, \ldots, \sum_{j=1}^{k-1} l_j + 2\beta\} \\ l_{k-1} & \text{for } i \in \{\sum_{j=1}^{k-1} l_j + 2\beta + 1, \ldots, \sum_{j=1}^{k-2} l_j + 4\beta\} \\ \vdots & \quad \vdots \\ l_1 & \text{for } i \in \{l_1 + 2(k-1)\beta + 1, \ldots, 2k\beta\}. \end{cases} \tag{3.3.118}$$

Hence the required coefficient comes from looking at the term for which $t = \sum_i a_i = \sum_i l_i(2\beta - l_i)$, which has coefficient

$$\frac{(-1)^{\sum_{i=1}^k l_i(2\beta-l_i)} \binom{\sum l_i(2\beta-l_i)}{l_k, \ldots, l_k, \ldots, l_1, \ldots, l_1}}{\left(\sum_{i=1}^k l_i(2\beta - l_i)\right)!} = \frac{(-1)^{k\beta}}{(l_1!)^{2\beta-l_1} \cdots (l_k!)^{2\beta-l_k}}. \tag{3.3.119}$$

Thus, we have constructed a term of the form $(v_1 \cdots v_{2k\beta})^{2\beta-1}$ which has strictly positive coefficient (the prefactors of $(-1)^{k\beta}$ in (3.3.114) and (3.3.119) cancel each other) given by

$$\frac{\prod_{j=1}^k l_j!(2\beta - l_j!)}{(l_1!)^{2\beta-l_1} \cdots (l_k!)^{2\beta-l_k}}. \tag{3.3.120}$$

In fact this is the only way to construct a term of this form from the integrand. One sees this as follows[7]. When trying to construct the term of the form $(v_1 \cdots v_{2k\beta})^{2\beta-1}$ in

$$\prod_{q=1}^k \Delta(v_{\sum_{j=1}^{q-1} l_j+1}, \ldots, v_{\sum_{j=1}^q l_j})^2 \prod_{q=1}^k \Delta(v_{\sum_{j=1}^q l_j + 2(k-q)\beta+1}, \ldots, v_{\sum_{j=1}^{q-1} l_j + 2(k-(q-1))\beta})^2 \prod_{\substack{m \le k\beta < n \\ \alpha_m = \alpha_n}} (v_n - v_m), \tag{3.3.121}$$

first note that the variables v_m, for $m \in \{1, \ldots, k\beta\}$ only appear in the Vandermonde determinants and the products

$$\prod_{\substack{m \le k\beta < n \\ \alpha_m = \alpha_n}} (v_n - v_m) = \prod_{\substack{m \in \{1, \ldots, l_1\} \\ n \in \{2(k-1)\beta+1+l_1, \ldots, 2k\beta\}}} (v_n - v_m) \cdots \prod_{\substack{m \in \{\sum_{j=1}^{k-1} l_j + 1, \ldots, k\beta\} \\ n \in \{k\beta+1, \ldots, \sum_{j=1}^{k-1} l_j + 2\beta\}}} (v_n - v_m). \tag{3.3.122}$$

In particular, after fixing $q \in \{1, \ldots, k\}$ take v_j with $j \in \{\sum_{i=1}^{q-1} l_i + 1, \ldots, \sum_{i=1}^q l_i\}$. Then v_j only appears in the following two terms:

$$\Delta(v_{\sum_{i=1}^{q-1} l_i+1}, \ldots, v_{\sum_{i=1}^q l_i})^2 \text{ and } \prod_{\substack{m \in \{\sum_{i=1}^{q-1} l_i+1, \ldots, \sum_{i=1}^q l_i\} \\ n \in \{2k\beta - \sum_{i=1}^q (2\beta - l_i)+1, \ldots, 2k\beta - \sum_{i=1}^{q-1}(2\beta - l_i)\}}} (v_n - v_m). \tag{3.3.123}$$

In particular these are both homogeneous polynomials: the former of degree $l_q(l_q - 1)$ in l_q variables and the latter is of degree $l_q(2\beta - l_q)$ in 2β variables. We will show that the only way to construct a term of the form $(v_1 \cdots v_{2k\beta})^{2\beta-1}$ is as described following (3.3.100). Without loss of generality, we will set $q = 1$ and assume $l_1 \ge 2$. From the above discussion, the square of the Vandermonde determinant

[7] The following argument was originally part of the appendix of [15].

consists of terms of the form

$$v_1^{a_1}\cdots v_{l_1}^{a_{l_1}}, \text{ with } \sum_{i=1}^{l_1} a_i = l_1(l_1-1). \tag{3.3.124}$$

Similarly, the product term is built of elements of the form

$$v_1^{b_1}\cdots v_{l_1}^{b_{l_1}} v_{2(k-1)\beta+1+l_1}^{b_{l_1+1}}\cdots v_{2k\beta}^{b_{2\beta}}, \text{ with } \sum_{i=1}^{2\beta} b_i = l_1(2\beta-l_1),\ 0\le b_i\le 2\beta-l_1. \tag{3.3.125}$$

Hence, each term of

$$\Delta(v_1,\dots,v_{l_1})^2 \prod_{\substack{m\in\{1,\dots,l_1\}\\ n\in\{2(k-1)\beta+1+l_1,\dots,2k\beta\}}} (v_n-v_m) \tag{3.3.126}$$

is of the form

$$v_1^{a_1+b_1}\cdots v_{l_1}^{a_{l_1}+b_{l_1}} v_{2(k-1)\beta+1+l_1}^{b_{l_1+1}}\cdots v_{2k\beta}^{b_{2\beta}}, \tag{3.3.127}$$

with a_i, b_i satisfying the homogenous conditions. To reach our goal, we need to find all possibilities for $a_i, 1\le i\le l_1$ and $b_i, 1\le i\le 2\beta$ that $a_i+b_i = 2\beta-1$ for $i\in\{1,\dots,l_1\}$. This implies that we need $\sum_{i=1}^{l_1}(a_i+b_i) = l_1(2\beta-1)$. Now note that the 'homogeneous conditions' in (3.3.124) and (3.3.125) together mean that

$$\sum_{i=1}^{l_1}(a_i+b_i)+\sum_{l_1+1}^{2\beta} b_i = \sum_{i=1}^{l_1} a_i + \sum_{i=1}^{2\beta} b_i = l_1(2\beta-1). \tag{3.3.128}$$

Thus, we must set $b_{l_1+1},\dots,b_{2\beta}=0$ if we want to construct the required term. This leaves us with finding all a_i, b_i $1\le i\le l_1$ such that all the following are satisfied,

$$0\le b_i\le 2\beta-l_1, \tag{3.3.129}$$

and

$$a_i+b_i = 2\beta-1, \tag{3.3.130}$$

$$\sum_{i=1}^{l_1} a_i = l_1(l_1-1), \tag{3.3.131}$$

$$\sum_{i=1}^{l_1} b_i = l_1(2\beta-l_1). \tag{3.3.132}$$

However, the latter two conditions imply that we must have $b_i = 2\beta-l_1$ for all $1\le i\le l_1$ which in turn gives us that $a_i = l_1-1$ for all $1\le i\le l_1$, and these are the only possible choices. This exactly matches the construction described following (3.3.100). The case for $q\in\{2,\dots,k\}$ follows similarly.

All that is left in order to show the statement of lemma 3.3.7 is that the term

$$(-1)^{g(k,\beta;\underline{l})}((k-1)\beta-\sum_{j=1}^{k-1} l_j)^{|B_{k,\beta;\underline{l}}|-\binom{k}{2}}\Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta-\sum_{j=1}^{k-1} l_j\right)\underline{v}\right)$$

only contributes a positive coefficient, where recall

$$\Psi_{k,\beta;\underline{l}}\left(\left((k-1)\beta-\sum_{j=1}^{k-1}l_j\right)\underline{v}\right)=$$

$$\underset{\substack{\underline{y}=(y_{m,n})_{(m,n)\in B_{k,\beta;\underline{l}}}\\(\tilde{\ddagger})}}{\int\cdots\int} \exp\left(((k-1)\beta-\sum_{j=1}^{k-1}l_j)\sum y_{m,n}(v_n-v_m)^{\pm}\right)\prod dy_{m,n}. \tag{3.3.133}$$

Calculating $\gamma_{k,\beta}$ involves computing derivatives of (3.3.133) and evaluating it at $\underline{v}=0$ by the residue theorem. We consider the case where $(k-1)\beta > l_1+\cdots+l_{k-1}$, the other case follows similarly. Incorporating in the sign $(-1)^{g(k,\beta;\underline{l})}$ in to the integrand of the right hand side of (3.3.133) (where for simplicity we ignore the positive prefactor in the exponent since it doesn't contribute to the overall sign) we have

$$\begin{aligned}(-1)^{g(k,\beta;\underline{l})}\prod_{(m,n)\in B_{k,\beta;\underline{l}}}&\exp\left(y_{m,n}(v_n-v_m)^{\pm}\right)\\&=\prod_{\sigma<\tau}\Bigg[\prod_{(m,n)\in S^{-}_{\sigma,\tau}}(\mathrm{sgn}(\mathrm{Re}\{(v_m-v_n)\})\exp\left(-y_{m,n}(v_n-v_m)\right))\\&\qquad\times\prod_{(m,n)\in S^{+}_{\sigma,\tau}}(-\mathrm{sgn}(\mathrm{Re}\{(v_n-v_m)\})\exp\left(y_{m,n}(v_n-v_m)\right))\Bigg].\end{aligned} \tag{3.3.134}$$

Thus, in order to show that $\gamma_{k,\beta}$ is strictly positive we need to establish that differentiating the right hand side of (3.3.134) contributes an overall positive sign. To see that this is true, first note that since each of the v_m, for $m\in\{1,\ldots,2k\beta\}$, in $P_{k,\beta}(\underline{l})$ has a pole of even order at 0, and by the residue theorem we are required to differentiate the exponential term in (3.3.133) an odd number of times. Then, by the requirements of the conditions on the Riemann integral in the right hand side of (3.3.133), for each (m,n), one has that the Heaviside function ensures that the product $y_{m,n}\,\mathrm{Re}\{(v_n-v_m)^{\pm}\}$ is negative. It is easy to check that in each case, after differentiating an odd number of times, that the term on the right hand of (3.3.134) is positive. This concludes the proof of lemma 3.3.7. □

3.4 Polynomial structure

In this section we prove theorem 3.1.5. The technique we use relies on a formula for $I_{k,\beta}(\theta_1,\ldots,\theta_k)$ (c.f. (3.3.1) and (3.3.2)) that follows from an expression obtained by Conrey et al. [45]. This takes the form of a combinatorial sum and is a special case of a more general expression explored further in chapter 5.

Theorem 3.4.1. *Let $\Xi_{k\beta}$ be the set of $\binom{2k\beta}{k\beta}$ permutations $\sigma\in S_{2k\beta}$ such that*

$$\sigma(1)<\sigma(2)<\cdots<\sigma(k\beta), \tag{3.4.1}$$

$$\sigma(k\beta+1)<\cdots<\sigma(2k\beta), \tag{3.4.2}$$

and

$$\underline{\omega}=(\underbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}_{\beta},\ldots,\underbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}_{\beta},\underbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}_{\beta},\ldots,\underbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}_{\beta}). \tag{3.4.3}$$

(Notice the similarity between the structure of $\underline{\omega}$ and (3.2.8).)

Then,

$$\left(\prod_{j=k\beta+1}^{2k\beta}\omega_j^N\right) I_{k,\beta}(\theta_1,\ldots,\theta_k) = \sum_{\sigma\in\Xi_{k\beta}} \frac{\left(\omega_{\sigma(k\beta+1)}\omega_{\sigma(k\beta+2)}\cdots\omega_{\sigma(2k\beta)}\right)^N}{\prod_{l\le k\beta<q}(1-\omega_{\sigma(l)}\omega_{\sigma(q)}^{-1})}. \tag{3.4.4}$$

Therefore,

$$\mathrm{MoM}_{U(N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi}\prod_{j=k\beta+1}^{2k\beta}\omega_j^{-N}\sum_{\sigma\in\Xi_{k\beta}}\frac{\left(\omega_{\sigma(k\beta+1)}\omega_{\sigma(k\beta+2)}\cdots\omega_{\sigma(2k\beta)}\right)^N}{\prod_{l\le k\beta<q}(1-\omega_{\sigma(l)}\omega_{\sigma(q)}^{-1})}d\theta_1\cdots d\theta_k. \tag{3.4.5}$$

The individual summands in the integrand in this expression have poles of finite order (when $\omega_{\sigma(q)} = \omega_{\sigma(l)}$). These cancel with zeros in the numerator in the complete sum, as they must because $I_{k,\beta}(\theta_1,\ldots,\theta_k)$ is bounded, being the average of a product of polynomials [45]. The function remaining after this cancellation may be computed by applying l'Hôpital's rule a finite number of times. This function is therefore a polynomial in the variables $e^{i\theta_1},\ldots,e^{i\theta_k}$ with coefficients that are each polynomial functions of N (coming from the derivatives associated with applying l'Hôpital's rule). Upon integrating only the coefficient of the constant term remains, which is polynomial in N. This concludes the proof of theorem 3.1.5. In principle one could compute the order of the polynomial this way, but in general we found the approach based on the asymptotic evaluation of the integral representation, set out in the previous section, to be more straightforward. In specific cases the calculation is feasible, as demonstrated in chapter 5.

3.5 Summary and outlook

Our main result within this chapter is a proof that the moments of the moments of the characteristic polynomials of random unitary matrices, $\mathrm{MoM}_{U(N)}(k,\beta)$, are polynomial functions of N, of order $k^2\beta^2 - k + 1$, when k and β both take values in $\mathbb{N}$. This proves the conjecture for the leading order asymptotics made in [81, 82] when k and β both take values in $\mathbb{N}$. Moreover, it goes further in establishing that an exact formula exists when k and β both take values in $\mathbb{N}$, and, in passing, establishes the general structure of the (finite) asymptotic expansion for $\mathrm{MoM}_{U(N)}(k,\beta)$ in this case.

It is clear from the calculation set out in section 3.2 that we have an exact formula when k and β both take values in $\mathbb{N}$ because of an underlying integrable structure: the approach based on symmetric function theory, and hence on representation theory, yields an exact formula in terms of a count of certain restricted semistandard Young tableaux. The symmetric functions used in section 3.2 may be related to certain generalized hypergeometric functions (c.f. [122] and, for example [75]), and it would be interesting to explore this calculation in that context, especially if doing so extends the results to non-integer values of k and β. We see our calculation as a first step in that direction and anticipate pursuing this further.

We note in passing that the formula we establish using the multiple integral approach provides as a byproduct an asymptotic expression for the count of semistandard Young tableaux that arises in the calculation. Additionally, given the recent result of Fahs [72], understanding the role that Toeplitz determinants, and hence Fisher-Hartwig asymptotics and solutions to Painlevé equations, play in such asymptotic combinatorial counts is a very interesting problem.

As suggested in chapter 2, section 2.2.2, the moments of the moments we study here play a central role in the heuristic analysis in [81,82,111] of the value distribution of $\log P_{\max}(A)$. Recall the statement of the Fyodorov-Keating conjecture (c.f. conjecture 2.2.2) in this case. As $N\to\infty$,

$$\log P_{\max}(A) = \log N - \frac{3}{4}\log\log N + x_{A,N}, \tag{3.5.1}$$

where $x_{A,N}$ is a random variable that is $O_{\mathbb{P}}(1)$ and which has a limiting value distribution that is a sum of two Gumbel distributions.

We covered in chapter 2, section 2.2.3 the fact that several components of the Fyodorov-Keating conjectures have recently been proved: the first term on the righthand side of (3.5.1) was established in [5], the second term in [132], and the tightness of $x_{A,N}$ in [35]. All of these calculations have utilised a hierarchical branching structure in the Fourier expansion of $\log |P_N(A,\theta)|$,

$$\log |P_N(A,\theta)| = -\mathrm{Re} \sum_{k=1}^{\infty} \frac{\mathrm{Tr} A^k}{k} \exp(ik\theta), \tag{3.5.2}$$

which we emphasise is similar to that found in other log-correlated Gaussian fields such as the branching random walk and the two-dimensional Gaussian Free Field; that is, they have utilised general probabilistic methods. When $\log |P_N(A,\theta)|$ (c.f. (3.5.2)) is replaced by a random Fourier series with the same correlation structure – such series can be considered as one-dimensional models of the two-dimensional Gaussian Free Field – the analogue of conjecture (3.1.7), due to Fyodorov and Bouchaud [78], has recently been proved for all k and β in the regime $k < 1/\beta^2$ by Remy [137] using ideas from conformal field theory [120], see section 2.2.3.

Formally, the $\beta \to \infty$ asymptotics of $g_N(\beta; A)$ determines $P_{\max}(A)$, and so it is natural to seek to understand the value distribution of $P_{\max}(A)$ by calculating the moments of $g_N(\beta; A)$ and then taking the large-β limit. However, doing this requires the moments for all k and β, not just the integer moments. Moreover, the controlling range is when freezing dominates and $k\beta^2$ is close to 1. Our results therefore cannot be applied as they stand. This is one reason why the possibility of using the integrable structure to extend them to non-integer values of k and β is attractive. When $k = 1$ the Selberg integral makes this possible. (And in the somewhat similar problem of the joint moments of the characteristic polynomial and its derivative, Painlevé theory provides a route (c.f. [14, 18], and chapter 7).) The result of Fahs [72], whilst verifying the asymptotic formula for real $\beta > 0$, obscures the leading order coefficient in the regimes $k\beta^2 > 1$, and so also does not provide fine enough information.

The association between characteristic polynomials of random matrices and the theory of the Riemann zeta function, see section 2.1.2, motivated the analogous conjecture 2.2.3,

$$\zeta_{\max}(T) = \max_{0 \le x < 2\pi} |\zeta(\tfrac{1}{2} + iT + ix)|, \tag{3.5.3}$$

where T is random [81, 82, 111]. These correspond to replacing N in (3.5.1) by $\log T$ (c.f [116]). Recall that in this case too there has been recent progress in proving the leading order term in the resulting formula when $T \to \infty$ [6, 127], based on calculations that mirror those for the extremes of characteristic polynomials.

The multiple-integral approach we have developed here also applies to the Riemann zeta function, using the representation established in [46], giving explicit (conjectural) formulae for the integer moments of moments over short intervals of the critical line in that case. These take the form of polynomials in $\log T$ up to an error that is a power of T smaller. This is important because in numerical computations of the moments one is necessarily restricted to finite intervals, and it is a key question how moments computed in different intervals fluctuate. Our formula for the moments of moments gives an answer to this question; this calculation is joint work of the present author and J. P. Keating [16].

The methods of calculation we have developed here to understand the unitary moments of moments extend also to the other classical compact groups. This applies to both the representations in terms of symmetric functions as well as the multiple contour integrals. Such a theory was originally developed in [33, 45]. The extension of our results to the other compact random matrix groups therefore follows the method outlined in this chapter. Such a calculation is the subject of chapter 4, though we focus on the combinatorial approach of section 3.2 rather than complex analytic techniques of section 3.3.

Once we understand the moments of moments for all the compact matrix groups, one then may apply the results to the corresponding moments of moments for each of the three symmetry classes of L-functions, see section 2.1.3. For further details see our subsequent paper [16]. Our results within this chapter in particular also apply immediately (and unconditionally) to the moments of the moments of function field L-functions defined over $\mathbb{F}_q$ in the limit $q \to \infty$. This follows from equidistribution results in that case (c.f. [111]).

Finally, our formulae have already been applied to analysing the results of numerical computations using randomly generated unitary matrices, where they explain the fluctuations in the moments of the characteristic polynomials evaluated by averaging over the unit circle [80]. We anticipate further similar applications and extensions to other numerical computations of the moments of spectral determinants.

Chapter 4

Symplectic and orthogonal moments of moments

This chapter is based on the paper '*On the moments of the moments of the characteristic polynomials of Haar distributed symplectic and orthogonal matrices*'. This paper was co-authored by the present author, T. Assiotis, and J. P. Keating [9]. The project came out of conversations with the co-authors.

The responsibility for writing the initial manuscript was split between T. Assiotis and the present author, and the diagrams were produced by the present author. Together, we proved the two theorems given within this chapter, theorem 4.1.1 and 4.1.2, under the advisement of J. P. Keating. Section 4.4.3 was lead by T. Assiotis. The proof of the polynomial part of theorems 4.1.1 and 4.1.2 (see section 4.3) adapts the technique presented in chapter 3, section 3.4. The idea for the proof of lemma 4.4.4 (and also lemma 4.5.5) originates from the paper of T. Assiotis and J. P. Keating [10].

Originally, the paper additionally contained examples of the polynomials whose existence is the subject of the two main theorems 4.1.1 and 4.1.2. Like in chapter 3, the code used to produce these polynomials was work of the present author, and they now form part of chapter 5 (and hence do not feature here). Otherwise this chapter closely follows the manuscript [9], except for where arguments have been expanded for clarity, and additional comments and examples have been inserted. Notation has also been changed, so to be consistent with the rest of this thesis. All such changes and inclusions are due to the present author.

4.1 Introduction

In comparison to the previous chapter, we now move away from working with unitary characteristic polynomials to focus on those related to symplectic and orthogonal matrices. It will be convenient to emphasise which group in particular is being considered, and so we write

$$P_{G(N)}(A,\theta) = \det\left(I - Ae^{-i\theta}\right)$$

for the characteristic polynomial on the unit circle of a matrix $A \in G(N)$, for $G(N) \in \{Sp(2N), SO(2N)\}$. Recall that $Sp(2N)$ denotes the group of $2N \times 2N$ (unitary) symplectic matrices, and $SO(2N)$ denotes the group of $2N \times 2N$ orthogonal matrices with determinant $+1$. We emphasise that the eigenvalues of matrices from $Sp(2N)$ and $SO(2N)$ lie on the unit circle and come in complex conjugate pairs, namely they are of the form: $e^{i\phi_1}, e^{-i\phi_1}, e^{i\phi_2}, e^{-i\phi_2}, \ldots, e^{i\phi_N}, e^{-i\phi_N}$. In particular, we have that:

$$\overline{P_{G(N)}(A,\theta)} = P_{G(N)}(A,-\theta). \tag{4.1.1}$$

Endowing the groups $Sp(2N)$ and $SO(2N)$ with the normalized Haar measure, we denote by $\mathbb{E}_{A\in G(N)}$ the mathematical expectation with respect to the corresponding measure on $G(N)$. Recall that we defined the *moments of moments* for a unitary characteristic polynomial in chapter 2 (c.f. (2.2.8)), which were the focus of chapter 3. Their generalization to $A \in G(N)$ is the subject of the present chapter, and so we define

$$\mathrm{MoM}_{G(N)}\,(k,\beta) := \mathbb{E}_{A\in G(N)}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_{G(N)}(A,\theta)|^{2\beta}d\theta\right)^k\right]. \tag{4.1.2}$$

Our focus will be on the asymptotics of $\mathrm{MoM}_{G(N)}\,(k,\beta)$ in the limit as $N\to\infty$ when k and β are fixed integers.

When $G(N)$ is the unitary group $U(N)$, we have already emphasised the great deal of interest in its moments of moments, see chapter 2, section 2.2.2 and chapter 3. The conjectured asymptotic formula for $\mathrm{MoM}_{U(N)}\,(2,\beta)$ was proved by Claeys and Krasovsky using a Riemann–Hilbert analysis [36], and for all non-negative integer values of k and β by Bailey and Keating [15] (c.f. theorem 3.1.4) using an approach based on exact formulae for finite N. An alternative approach when k and β are non-negative integers was developed by Assiotis and Keating [10]. They used the connection between unitary moments and representation theory (see section 3.2), and additionally drew upon the link with constrained Gelfand-Tsetlin patterns. This yields the same results as found in [15], but leads to an alternative explicit formula for the coefficient appearing in the leading-order contribution to the asymptotics in terms of the volume of the associated Gelfand-Tsetlin polytopes; i.e. it provides a geometrical interpretation for this constant. Recently, Fahs has extended the approach developed in [36] to give a proof of the asymptotic formula for $\mathrm{MoM}_{U(N)}\,(k,\beta)$ for non-negative integer values of k and general non-negative real β, but without an explicit expression for the coefficient of the leading order term. There is considerable interest in removing the assumption that k is a non-negative integer though this is likely to require new ideas. Finally, there has also been a good deal of progress in proving the associated conjectures for the extreme value statistics of the characteristic polynomials; see, for example, [5, 35, 132].

Our purpose here is to extend the approach developed in [10] to give formulae for $\mathrm{MoM}_{G(N)}\,(k,\beta)$, when k and β are non-negative integers and when $G(N)$ is either of the groups $Sp(2N)$ and $SO(2N)$, in terms of the associated constrained Gelfand-Tsetlin patterns (which are different to those that appear in the unitary case in [10]). We then establish asymptotic formulae in which the volumes of the related Gelfand-Tsetlin polytopes appear. Importantly, we find that the leading order asymptotic dependence on N is conditional on the group in question.

We now have a well developed understanding of how to use results for random matrices to make conjectures about the corresponding questions in number theory, see for example chapter 2, section 2.1.2. For example, formulae for the moments of the moments of the characteristic polynomials of random unitary matrices, and for the extreme value statistics of the characteristic polynomials, can be used to motivate conjectures for the moments of the moments and for the extreme value statistics of the Riemann zeta-function on short intervals of its critical line [81, 82]. There has recently been progress in proving these conjectures; see, for example, [6, 94, 95, 127].

Our results here provide a similar basis for conjecturing formulae for the moments of the moments of L-functions from orthogonal and symplectic families. For example, one could consider L-functions associated with quadratic twists of elliptic curves, and quadratic Dirichlet L-functions. The two averages are, first, over a short section of the critical line (e.g. a section of length 2π) centred on the symmetry point of the functional equation, and, second, over members of the family (i.e. in the two examples given, over twists). This application is explored further by Bailey and Keating [16].

It would be interesting to extend the approach developed by Claeys and Krasovsky in [36] and Fahs in [72] to the orthogonal and symplectic groups. This would require uniform asymptotics for determinants of the form Toeplitz + Hankel as the singularities merge; as far as we are aware this theory remains to be developed. It would also be interesting to explore the implications of our results for orthogonal and symplectic analogues of Gaussian Multiplicative Chaos, along the lines of the corresponding theory in the unitary case (see, for example, [128, 155] and section 1.1.5).

4.1.1 Statement of results

Our results are as follows.

Theorem 4.1.1. *Let $G(N) = Sp(2N)$. Let $k, \beta \in \mathbb{N}$. Then,* $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ *is a* polynomial function *in N. Moreover,*

$$\mathrm{MoM}_{Sp(2N)}(k,\beta) = \mathfrak{c}_{Sp}(k,\beta) N^{k\beta(2k\beta+1)-k} + O\left(N^{k\beta(2k\beta+1)-k-1}\right), \tag{4.1.3}$$

where the leading order term coefficient $\mathfrak{c}_{Sp}(k,\beta)$ is the volume of a convex region defined in section 4.4.2 and is strictly positive.

Theorem 4.1.2. *Let $G(N) = SO(2N)$. Let $k, \beta \in \mathbb{N}$. Then,* $\mathrm{MoM}_{SO(2N)}(k,\beta)$ *is a* polynomial function *in N. Moreover,*

$$\mathrm{MoM}_{SO(2N)}(1,1) = 2(N+1) \tag{4.1.4}$$

otherwise,

$$\mathrm{MoM}_{SO(2N)}(k,\beta) = \mathfrak{c}_{SO}(k,\beta) N^{k\beta(2k\beta-1)-k} + O\left(N^{k\beta(2k\beta-1)-k-1}\right), \tag{4.1.5}$$

where the leading order term coefficient $\mathfrak{c}_{SO}(k,\beta)$ is given as a sum of volumes of convex regions described in section 4.5.2 and is strictly positive.

Recall that in chapter 3, we recovered the equivalent result for the unitary group, showing that for $k, \beta \in \mathbb{N}$

$$\mathrm{MoM}_{U(N)}(k,\beta) = \mathfrak{c}_{U}(k,\beta) N^{k^2\beta^2-k+1} + O\left(N^{k^2\beta^2-k}\right), \tag{4.1.6}$$

which too is a polynomial in N. Thus, the leading power of N for each of the matrix groups differs.

4.1.2 Strategy of proof

In order to prove our main results we combine the approaches that were developed in [15] and [10] (see chapter 3, and also [113]) for treating the simpler case of the unitary group. We first adapt an argument presented in [15] to prove that $\mathrm{MoM}_{U(N)}(k,\beta)$ is a polynomial in N. Then, in order to obtain the leading order term and an expression for its coefficient, we develop the combinatorial approach of [10] to this setting.

Recall from chapter 3 that the equivalent theorem for the unitary group, (4.1.6), was proved using a multiple contour integral representation due to Conrey et al. [45] for the average of a product of characteristic polynomials over $U(N)$. There are equivalent multiple contour integral formulae for symplectic and orthogonal polynomials [45], and so one could equally pursue that method of proof for theorems 4.1.1 and 4.1.2. This would also result in a different interpretation of the leading order coefficients $\mathfrak{c}_{Sp}(k,\beta)$ and $\mathfrak{c}_{SO}(k,\beta)$. However, due to the natural integrable structure of the moments of moments, and the beautiful combinatorics that one encounters, we have chosen to instead use the representation theoretic method of [10, 33].

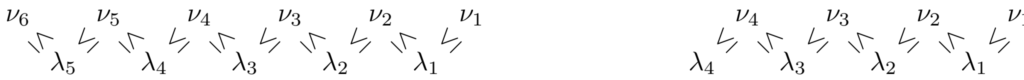
The outline of the proof is as follows. We first obtain an expression for $\mathrm{MoM}_{G(N)}(k, \beta)$ in terms of certain combinatorial objects, namely Gelfand-Tsetlin patterns, satisfying some (quite involved) constraints. We do this by making use of formulae due to Bump and Gamburd [33] that express averages of products of characteristic polynomials over the classical compact groups in terms of certain associated characters. The next step can be seen as taking a discrete to continuous limit, which gives the leading order coefficient as the volume of an explicit polytope, see sections 4.2.3, 4.4.2, and 4.5.2 for more precise statements.

There are certain important, not entirely technical, differences to the unitary group setting explored in [10]. In particular, the combinatorial objects we work with, namely the symplectic and orthogonal Gelfand-Tsetlin patterns, are more complicated than their unitary counterparts. For example, in order to apply the results required for the discrete to continuous limit in the orthogonal case, we first need to perform a decomposition of the corresponding patterns. The most significant difference however is the complexity of the constraints involved in the orthogonal and symplectic settings. For the case of the unitary group, the constraints only depend on a single level of the pattern, whereas for the cases considered in this paper they involve several levels. We review the theory and the calculation for the unitary case in section 4.2.

This complication has the following consequences. Firstly, from the discrete to continuous limit argument it is not immediately clear that the leading order coefficient is actually strictly positive (which, as shown in section 4.2, is straightforward in the unitary case). We manage to overcome this problem by a careful analysis of the different types of constraints. This is one of the more challenging parts of the paper, and the argument is supplemented by a number of diagrams. Secondly, the intricacies of the constraints prevents us, at least at present, from obtaining a more explicit expression for the leading order coefficient as was done in [10] (such an expression has been used to connect this coefficient to Painlevé equations for $k = 2$, see [113] and [19]). However we do not believe that this is an intrinsic limitation of our approach, since, as we show in section 4.4.3 for example, whenever such a leading order coefficient in an allied problem has been computed explicitly by different methods, it can fact also be reproduced by calculating volumes of Gelfand-Tsetlin polytopes.

However, there are some overarching themes common to the analysis of each case. For example, the moments of the moments are polynomials in N for each compact group. Although the leading order power of N (i.e. the degree of the polynomial) differs, they share the same construction. By this we mean the following: the degree of the polynomial can be seen as having a 'compact group average' part, and a 'unit circle average' part. For the unitary case, recall that degree of the moments of moments is $k^2\beta^2 - k + 1$. The $k^2\beta^2$ contribution is akin to the β^2 leading order of the Keating and Snaith 2βth moment calculation (except our moments consist of $2k\beta$ characteristic polynomials rather than 2β), which is an average over $U(N)$. The remaining $-(k-1)$ contribution relates to the fact that in the unitary case (partly due to the rotational invariance of the Haar measure) we have $k-1$ unique constraints coming from subsequently averaging around the unit circle. A similar statement is true of the degrees of the symplectic and orthogonal moments of moments polynomials (c.f. theorems 2.1.7 and 2.1.8 for the 'compact group average' calculations at the symmetry point).

Finally we emphasise that because of the combinatorial nature of the method of proof, many of the arguments are more suited to a visual explanation. Indeed, an unfortunate consequence of translating such arguments to text is the notation that they necessitate. Thus, wherever possible, we have provided accompanying diagrams to the proofs, and sign-post them throughout to aid understanding.



(a) Interlacing with $\lambda \in \mathrm{S}_5$, $\nu \in \mathrm{S}_6$

(b) Interlacing with $\lambda \in \mathrm{S}_4$, $\nu \in \mathrm{S}_4$

Figure 4.1: Examples of interlacing of signatures. Figure 4.1a shows signatures whose length differs by 1, and figure 4.1b shows two signatures of the same length. Note that in both cases, the numbering is from right to left, as to be in-keeping with future definitions. The inequalities are explicitly shown, though it is not standard to do so.

4.2 Preliminaries

4.2.1 Symplectic and orthogonal Gelfand-Tsetlin patterns and Schur polynomials

We will now give some background on symplectic and orthogonal Schur polynomials. These are related to the Schur polynomials already encountered in this thesis, see sections 1.1.6 and 3.2. The extension to the other compact groups can be defined via the characters of irreducible representations of the corresponding classical compact groups, c.f. the equivalent unitary construction, proposition 3.2.3. From this perspective, making use of the Weyl character formula, one obtains well-known explicit expressions in terms of ratios of determinants (which we also record below). For our purposes however, we shall need some equivalent (see [135]) combinatorial definitions in terms of sums over objects called Gelfand-Tsetlin patterns. We mainly follow the recent exposition in section 2 of [11].

Firstly, we generalize the definition of a partition, to allow for general integer entries.

Definition 4.2.1 (Signature)**.** *A signature λ of length n is a sequence of n non-increasing integers $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n)$. We denote the set of all such signatures by S_n. We also denote the set of the signatures with non-negative entries by S_n^+. Note that this is* distinct *from the definition of a partition since we keep track of* trailing zeros*, whereas partitions ν are identified with any other partition which has the same non-zero entries.*

For $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathrm{S}_n^+$ we define $\lambda^- := (\lambda_1, \ldots, \lambda_{n-1}, -\lambda_n)$. If $\lambda_1 = \cdots = \lambda_n = L$ then we also write $\lambda = \langle L^n \rangle$.

We record the next definition to cover the case of two given signatures interacting in a certain way.

Definition 4.2.2 (Interlacing)**.** *We say that signatures $\lambda \in \mathrm{S}_n$ and $\nu \in \mathrm{S}_{n+1}$ interlace, and write $\lambda \prec \nu$, if:*

$$\nu_1 \geq \lambda_1 \geq \nu_2 \geq \cdots \geq \nu_n \geq \lambda_n \geq \nu_{n+1}. \tag{4.2.1}$$

Similarly, we say that $\lambda \in \mathrm{S}_n$ and $\nu \in \mathrm{S}_n$ interlace, and still write $\lambda \prec \nu$ if:

$$\nu_1 \geq \lambda_1 \geq \nu_2 \geq \cdots \geq \nu_n \geq \lambda_n. \tag{4.2.2}$$

It is common to draw interlacing signatures, so to emphasise the their interaction. A pictorial representation of interlacing is given in figure 4.1. Note that there, as with future definitions and to be in keeping with the literature, we draw signatures from right to left.

The next definition introduces a 'full' *Gelfand-Tsetlin* pattern. We do not make use of full patterns in this chapter, beyond the next discussion, but their introduction facilitates the subsequent definition of 'half' *Gelfand-Tsetlin* patterns.

$$\begin{array}{ccccccc} 1 & & 1 & & 6 & & 9 \\ & 1 & & 3 & & 7 & \\ & & 2 & & 5 & & \\ & & & 4 & & & \end{array}$$

(a) Example of a Gelfand-Tsetlin pattern of length 4.

$$\begin{array}{ccccccc} 1 & & 2 & & 3 & & 4 \\ & 1 & & 2 & & 4 & \\ & & 2 & & 4 & & \\ & & & 3 & & & \end{array}$$

(b) Example of a Gelfand-Tsetlin pattern of length 4 with fixed top row $\nu = (4, 3, 2, 1)$.

Figure 4.2: Examples of Gelfand-Tsetlin Patterns.

Definition 4.2.3 (Gelfand-Tsetlin pattern)**.** *A non-negative Gelfand-Tsetlin pattern of length/depth n is a sequence of signatures $\left(\lambda^{(i)}\right)_{i=1}^{n}$ such that $\lambda^{(i)} \in \mathrm{S}_i^+$ and*

$$\lambda^{(1)} \prec \lambda^{(2)} \prec \cdots \prec \lambda^{(n-1)} \prec \lambda^{(n)}. \tag{4.2.3}$$

One writes GT_n^+ for the set of all such patterns. Given a signature $\nu \in S_n^+$, it is often useful to additionally consider those Gelfand-Tsetlin pattern with top row ν. The set of such patterns is written $GT_n^+(\nu)$.

Finally, it is common to draw Gelfand-Tsetlin patterns as a triangular array, essentially a generalization of the representation of interlacing signatures shown in figure 4.1. An example of this pictorial representation is given by figure 4.2.

The connection with semistandard Young tableaux can now be explored. For a signature $\nu \in \mathrm{S}_n^+$, there is a well-known bijection between semistandard Young tableaux of shape ν with entries in $\{1, \ldots, n\}$ and non-negative Gelfand-Tsetlin patterns of length n with fixed top row ν. The correspondence is as follows, see for example [10, 89].

Given $\nu \in \mathrm{S}_n^+$ and a Gelfand-Tsetlin pattern $P = \left(\lambda^{(i)}\right)_{i=1}^{n}$ where $\lambda^{(n)} = \nu$ (so $P \in GT_n^+(\nu)$), we retrieve the corresponding semistandard Young tableau by inserting the digit $1 \leq m \leq n$ into the cells of the tableau of shape $\mu^{(m)} = \lambda^{(m)} \backslash \lambda^{(m-1)}$ with $\mu^{(1)} = \lambda^{(1)} \backslash \varnothing$. If $\lambda^{(m)} \backslash \lambda^{(m-1)}$ is empty then do not insert m. Such a procedure is outlined in example 4.2.4.

In the other direction, we proceed in the following way. Given a signature $\nu \in \mathrm{S}_n^+$ and a semistandard Young tableaux Y of shape ν with a filling using entries from $\{1, \ldots, n\}$, one obtains the relevant Gelfand-Tsetlin pattern $P = \left(\lambda^{(i)}\right)_{i=1}^{n} \in GT_n^+(\nu)$ by setting $\lambda^{(m)}$ to be the Young diagram consisting of cells of Y with entries less than or equal to m, and removing trailing zeros to ensure $\lambda^{(m)} \in S_m^+$. An example of this algorithm is shown in example 4.2.5.

Example 4.2.4 (Gelfand-Tsetlin pattern to semistandard Young tableau.)**.** *Take $P \in GT_4^+(\nu)$ for $\nu = (4, 3, 2, 1)$, as shown in figure 4.2b. To find the associated semistandard Young tableau Y of shape ν with entries in $\{1, 2, 3, 4\}$ (note that the similarity between the signature ν and the set of possible entries is just coincidental here) we apply the described procedure.*

Since the shape of the tableau is known, we draw it first.

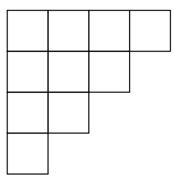


Figure 4.3: Young tableau Y of shape $\nu = (4, 3, 2, 1)$.

To insert the entries $\{1, 2, 3, 4\}$ *in to* Y*, we begin with the entry* 1*. We take the first signature in the definition of* P*, which by figure 4.2b is* $\lambda^{(1)} = (3)$*. We begin with the subtableau of* Y *corresponding to* $\lambda^{(1)}$*, shown shaded below, and insert the value* 1 *there.*

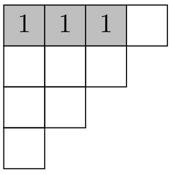
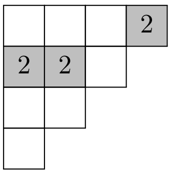
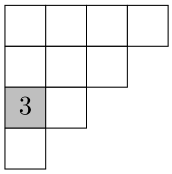


Figure 4.4: The Young tableau Y of shape $\nu = (4, 3, 2, 1)$, with 1 inserted.

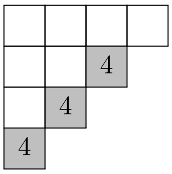
By repeating the procedure, we find the relevant semistandard Young tableau. Figure 4.5 shows the process, where at each stage the available boxes for shading are highlighted. The notation $\lambda^{(m)} \backslash \lambda^{(m-1)}$ *means remove the subshape* $\lambda^{(m-1)}$ *from* $\lambda^{(m)}$*.*



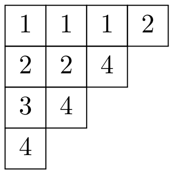
(a) Inserting 2 in to Y.



(b) Inserting 3 in to Y.



(c) Inserting 4 in to Y.



(d) The semistandard Young tableau corresponding to the Gelfand-Tsetlin pattern P.

Figure 4.5: Example of applying the bijection between semistandard Young tableaux and Gelfand-Tsetlin patterns, from Gelfand-Tsetlin pattern to Young tableau.

Example 4.2.5 (Semistandard Young tableau to Gelfand-Tsetlin pattern.)**.** *Take* Y *to be the following semistandard Young tableau of shape* $\nu = (9, 6, 1, 1)$*. We will show that its corresponding Gelfand-Tsetlin pattern is shown in figure 4.2a.*

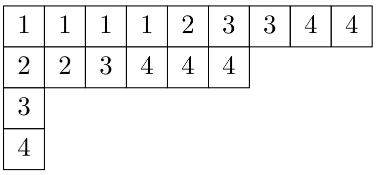


Figure 4.6: A semistandard Young tableau Y of shape $\nu = (9, 6, 1, 1)$.

To construct $P = (\lambda^{(1)}, \ldots, \lambda^{(4)}) \in GT_4^+(\nu)$ *from* Y *we apply the described procedure. First set* $\lambda^{(1)}$ *to be partition for the Young diagram consisting of the cells of* Y *with entries less than or equal to* 1*. Such cells are shaded in the diagram below.*

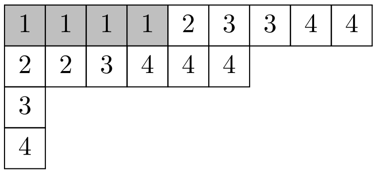
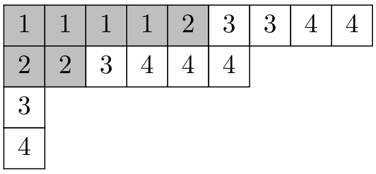
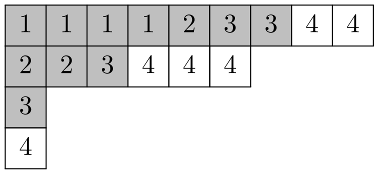


Figure 4.7: Y with cells with entries less than or equal to 1 shaded.

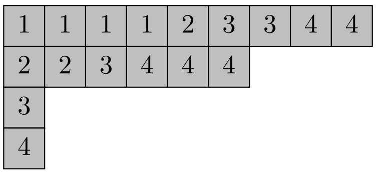
Hence $\lambda^{(1)} = (4)$. We continuing in this way, setting $\lambda^{(m)}$ equal to the partition corresponding to the subset of Y with entries less than or equal to m. The procedure is shown in figure 4.8, with the subset of Y shaded at each stage.



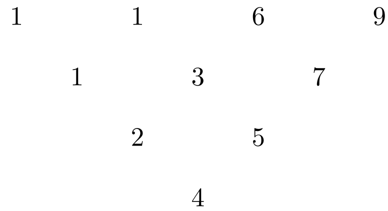
(a) Setting $\lambda^{(2)} = (5, 2)$.



(b) Setting $\lambda^{(3)} = (7, 3, 1)$.



(c) Setting $\lambda^{(4)} = (9, 6, 1, 1)$.



(d) The Gelfand-Tsetlin pattern corresponding to the semistandard Young tableau Y.

Figure 4.8: Example of applying the bijection between semistandard Young tableaux and Gelfand-Tsetlin patterns, from Young tableau to Gelfand-Tsetlin pattern.

Recall that in chapter 3, section 3.2, we showed that the unitary moments of moments are equal to a count of *restricted* semistandard Young tableaux of rectangular shape. However, only a polynomial bound on $\mathrm{MoM}_{U(N)}(k, \beta)$ was justified there. Assiotis and Keating took the analysis further, exploiting the connection described above with Gelfand-Tsetlin patterns. In particular, they show the following.

Proposition 4.2.6 (Assiotis and Keating [10]). *Let $k, \beta \in \mathbb{N}$. Then $\mathrm{MoM}_{U(N)}(k, \beta)$ is equal to the number of Gelfand-Tsetlin patterns of length $2k\beta$ with top row*

$$\lambda^{(2k\beta)} = (\overbrace{N, \ldots, N}^{k\beta}, \overbrace{0, \ldots, 0}^{k\beta}), \tag{4.2.4}$$

which satisfy the following constraints

$$\lambda_1^{(2j\beta)} + \cdots + \lambda_{2j\beta}^{(2j\beta)} = jN\beta \tag{4.2.5}$$

for $j \in \{1, \ldots, k\}$.

The proof of proposition 4.2.6 follows directly from our observation (3.2.15) and the bijection between semistandard Young tableaux and Gelfand-Tsetlin patterns described above.

Note that the structure of the top row (4.2.4) means that two sections of the resulting Gelfand-Tsetlin pattern are 'frozen'. This is due to the non-negativity and the interlacing condition of Gelfand-Tsetlin patterns. Finally, in order to rederive theorem 3.1.4 using this combinatorial approach, Assiotis and Keating require the following two definitions. Ideas along these lines are also necessary for the proofs of theorems 4.1.1 and 4.1.2.

Definition 4.2.7. *Let $\mathcal{I}_n \subset \mathbb{R}^n \times \mathbb{R}^n$ denote the collection of real $n \times n$ matrices so that if $(x_i^{(j)})_{i,j=1}^n \in \mathcal{I}_n$, then*

$$x_1^{(j)} \leq x_2^{(j)} \leq \cdots \leq x_n^{(j)}, \tag{4.2.6}$$

$$x_i^{(1)} \geq x_i^{(2)} \geq \cdots \geq x_i^{(n)}, \tag{4.2.7}$$

$$\tag{4.2.8}$$

*for $i, j \in \{1, \ldots, n\}$. Thus, I_n is the set of $n \times n$ real matrices $x_i^{(j)}$ whose entries are non-*decreasing *along the rows (indexed by the subscript) and non-*increasing *down columns (indexed by the superscript).*

Definition 4.2.8. *Let $\mathcal{M}_N(k, \beta)$ be the set of integer arrays $x = (x_i^{(j)})_{i,j=1}^{k\beta} \in \mathbb{Z}^{k\beta} \times \mathbb{Z}^{k\beta}$ satisfying*

(i) $0 \leq x_i^{(j)} \leq N$, for $1 \leq i, j \leq k\beta$,

(ii) for $l = 1, \ldots, \lfloor \frac{k}{2} \rfloor$,

$$\sum_{j=1}^{2\beta l} x_j^{(2\beta l-(j-1))} = l\beta N, \tag{4.2.9}$$

$$\sum_{j=1}^{2\beta l} x_{k\beta-2\beta l+j}^{(k\beta-(j-1))} = l\beta N, \tag{4.2.10}$$

$$\tag{4.2.11}$$

(iii) the matrix $x \in \mathcal{I}_{k\beta}$.

Assiotis and Keating then define a further bijection between Gelfand-Tsetlin patterns and $\mathcal{M}_N(k, \beta)$, see [10] for the full description. This allows them to reduce the problem once again to

$$\mathrm{MoM}_{U(N)}(k, \beta) = \#\mathcal{M}_N(k, \beta), \tag{4.2.12}$$

for $k, \beta \in \mathbb{N}$. Finally, moving from a discrete to a continuous setting, and using theorem 4.2.16, they are able to re-establish theorem 3.1.4. Moreover, the leading order coefficient $\mathfrak{c}_U(k, \beta)$ is hence related to the volume of certain a convex region (which in turn can be related to the Painlevé V equation when $k = 2$, c.f. also chapter 7).

Moving away from the unitary case, we now define the related notion of a half Gelfand-Tsetlin pattern, see figure 4.9 for an example. We often drop the 'Gelfand-Tsetlin' from 'half Gelfand-Tsetlin pattern' henceforth for brevity. Symplectic and orthogonal Gelfand-Tsetlin patterns will be half patterns with additional properties. Such patterns are intimately related to moments of symplectic and orthogonal characteristic polynomials, just as (unitary) Gelfand-Tsetlin patterns are via the bijection to semistandard Young tableaux described above.

Definition 4.2.9 (Half (Gelfand-Tsetlin) pattern)**.** *Let n be a positive integer. A half (Gelfand-Tsetlin) pattern of length n is given by a sequence of interlacing signatures $\left(\lambda^{(i)}\right)_{i=1}^n$ such that $\lambda^{(2i-1)}, \lambda^{(2i)} \in \mathrm{S}_i$ and the interlacing is as follows:*

$$\lambda^{(1)} \prec \lambda^{(2)} \prec \cdots \prec \lambda^{(n-1)} \prec \lambda^{(n)}.$$

We call the first entries on the odd rows, namely $\lambda_i^{(2i-1)}$, the odd starters.

We arrive to the definition of a symplectic Gelfand-Tsetlin pattern, see figure 4.10a for an illustration. Again, we will usually refer to these as just 'symplectic patterns'.

Definition 4.2.10 (Symplectic patterns)**.** *Let n be a positive integer. A $(2n)$–symplectic Gelfand-Tsetlin pattern $P = \left(\lambda^{(i)}\right)_{i=1}^{2n}$ is a half pattern of length $2n$ all of whose entries are non-negative integers. For fixed complex numbers $(x_1, \ldots, x_n)$ we associate to the pattern P a weight $w_{sp}(P)$ (dependence on $x_1, \ldots, x_n$ is suppressed from the notation and will be clear from context in what follows) given by:*

$$w_{sp}(P) = \prod_{i=1}^{n} x_i^{\sum_{j=1}^{i} \lambda_j^{(2i)} - 2\sum_{j=1}^{i} \lambda_j^{(2i-1)} + \sum_{j=1}^{i-1} \lambda_j^{(2i-2)}},$$

with $\lambda^{(0)} \equiv 0$. For $\nu \in \mathrm{S}_n^+$, we write SP_ν for the set of all $(2n)$–symplectic Gelfand-Tsetlin patterns with top row $\lambda^{(2n)} = \nu$.

We now give the combinatorial definition of the symplectic Schur polynomial as a sum of weights over symplectic patterns. This should be seen in the context of the definition of a (unitary) Schur polynomial, see definition 1.1.17.

Definition 4.2.11 (Symplectic Schur polynomial)**.** *Let $\nu \in \mathrm{S}_n^+$. We define the symplectic Schur polynomial by:*

$$sp_\nu^{(2n)}(x_1, \ldots, x_n) = \sum_{P \in SP_\nu} w_{sp}(P). \tag{4.2.13}$$

It can be shown (see [135]) that this combinatorial definition coincides with the following determinantal form given by the Weyl character formula:

$$sp_\nu^{(2n)}(x_1, \ldots, x_n) = \frac{\det\left(x_i^{\nu_j+n-j+1} - x_i^{-(\nu_j+n-j+1)}\right)_{i,j=1}^{n}}{\det\left(x_i^{n-j+1} - x_i^{-(n-j+1)}\right)_{i,j=1}^{n}}. \tag{4.2.14}$$

We move on to the definition of orthogonal patterns. This is slightly more involved than the symplectic case since some of the elements are now permitted to be negative. We will use the notation

$$\mathrm{sgn}(x) = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0. \end{cases}$$

Definition 4.2.12 (Orthogonal patterns)**.** *Let n be a positive integer. A $(2n-1)$–orthogonal Gelfand-Tsetlin pattern $P = \left(\lambda^{(i)}\right)_{i=1}^{2n-1}$ is a half pattern of length $2n-1$ all of whose entries are either all integers or all half-integers*[1] *and which moreover satisfy:*

(i) All entries except odd starters are non-negative.

(ii) The odd starters satisfy $|\lambda_i^{(2i-1)}| \leq \min\{\lambda_{i-1}^{(2i-2)}, \lambda_i^{(2i)}\}$ for $i = 2, \ldots, n-1$ and moreover $|\lambda_1^{(1)}| \leq \lambda_1^{(2)}$ and $|\lambda_n^{(2n-1)}| \leq \lambda_{n-1}^{(2n-2)}$.

For fixed complex numbers $(x_1, \ldots, x_n)$ we associate to the pattern P a weight $w_o(P)$ given by:

$$w_o(P) = \prod_{i=1}^{n} x_i^{\mathrm{sgn}(\lambda_i^{(2i-1)})\mathrm{sgn}(\lambda_{i-1}^{(2i-3)})\left[\sum_{j=1}^{i} |\lambda_j^{(2i-1)}| - 2\sum_{j=1}^{i-1} |\lambda_j^{(2i-2)}| + \sum_{j=1}^{i-1} |\lambda_j^{(2i-3)}|\right]},$$

[1]It transpires that for our application the entries of $(2n-1)$–orthogonal Gelfand-Tsetlin patterns are always all integers.

with $\lambda^{(0)}, \lambda^{(-1)} \equiv 0$*. For* $\nu \in \mathrm{S}_n$*, we write* OP_ν *for the set of all* $(2n-1)$*–orthogonal Gelfand-Tsetlin patterns with top row* $\lambda^{(2n-1)} = \nu$*.*

See figure 4.10b for an example of an orthogonal Gelfand-Tsetlin pattern.

As in the symplectic case, we have the following combinatorial definition of the orthogonal Schur polynomial as a sum of weights over orthogonal patterns.

Definition 4.2.13 (Orthogonal Schur polynomial)**.** *Let* $\nu \in \mathrm{S}_n^+$*. We define the orthogonal Schur polynomial by:*

$$o_\nu^{(2n)}(x_1, \ldots, x_n) = \sum_{P \in OP_\nu \cup OP_{\nu^-}} w_o(P). \tag{4.2.15}$$

Again, it can be shown (see [135]) that this combinatorial definition coincides with the following determinantal expression given by the Weyl character formula:

$$o_\nu^{(2n)}(x_1, \ldots, x_n) = \frac{2\det\left(x_i^{\nu_j+n-j} + x_i^{-(\nu_j+n-j)}\right)_{i,j=1}^n}{\det\left(x_i^{n-j} + x_i^{-(n-j)}\right)_{i,j=1}^n}. \tag{4.2.16}$$

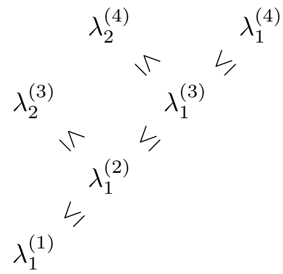
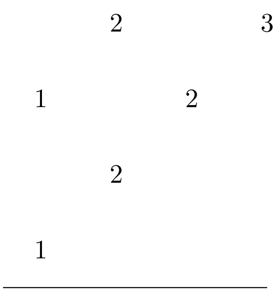
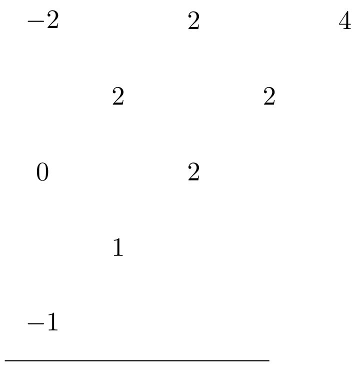


Figure 4.9: A half pattern of length 4, $(\lambda^{(i)})_{i=1}^4$, with the interlacing explicitly shown.



$w_{sp}(P) = x_2$

(a) An example of a (4)-symplectic Gelfand-Tsetlin pattern P, with its corresponding weight $w_{sp}(P)$ below for some complex numbers x_1, x_2 as appearing in definition 4.2.10.



$w_o(P) = (x_1 x_2 x_3^2)^{-1}$

(b) An example of a (5)-orthogonal Gelfand-Tsetlin pattern P, with its corresponding weight $w_o(P)$ below for some complex numbers x_1, x_2, x_3 as appearing in definition 4.2.12.

Figure 4.10: Figures giving examples of symplectic and orthogonal Gelfand-Tsetlin patterns.

4.2.2 Averages of products of characteristic polynomials as Schur polynomials

We have the following results due to Bump and Gamburd, see sections 5 and 6 in [33] (note that [33] uses the equivalent definition of Schur polynomials in terms of determinants, see (4.2.14) and (4.2.16)). These relate products of characteristic polynomials averaged (with respect to Haar measure) over the classical compact groups with Schur polynomials.

Proposition 4.2.14. *Let n be a positive integer and $x_1, \ldots, x_n$ be complex numbers. Then,*

$$\mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{n}\det\left(I-Ax_j\right)\right]=(x_1\cdots x_n)^N sp^{(2n)}_{\langle N^n\rangle}\left(x_1,\ldots,x_n\right). \tag{4.2.17}$$

Proposition 4.2.15. *Let n be a positive integer and $x_1, \ldots, x_n$ be complex numbers. Then,*

$$\mathbb{E}_{A\in SO(2N)}\left[\prod_{j=1}^{n}\det\left(I-Ax_j\right)\right]=(x_1\cdots x_n)^N o^{(2n)}_{\langle N^n\rangle}\left(x_1,\ldots,x_n\right). \tag{4.2.18}$$

Bump and Gamburd proved propositions 4.2.14 and 4.2.15 expanding the averages of the products of characteristic polynomials over the relevant Weyl integration and character formulae [156]. In our applications below we will be taking particular choices of the complex numbers $x_1, \ldots, x_n$ lying on the unit circle in the complex plane for some even integer n.

4.2.3 Asymptotics of the number of lattice points in convex sets

We have the following theorem on the number of lattice points in convex regions of Euclidean space, see for example section 2 in [143]. Recall that in a convex region, the line segment joining any two points also lies completely within the region.

Theorem 4.2.16. *Assume $\mathcal{S}\subset\mathbb{R}^L$ is a convex region contained in a closed ball of radius ρ. Then,*

$$\#\left(\mathcal{S}\cap\mathbb{Z}^L\right)=\mathrm{vol}_L\left(\mathcal{S}\right)+O_L\left(\rho^{L-1}\right), \tag{4.2.19}$$

where the implicit constant in the error term depends only on L.

We will prove our main results on the asymptotics of the moments of the moments by applying the theorem above with some judicious choices (different for each group) of the convex set $\mathcal{S}$.

4.2.4 Averages of products of characteristic polynomials as combinatorial sums

Instead of expressing the averages of products of characteristic polynomials over the various matrix groups in terms of their Schur polynomials, one can instead view them as combinatorial sums. These descriptions follow from work of Conrey et al. [45] and will be used when determining the polynomial structure of the moments of moments. Such an approach was also used in chapter 3 to prove the analogous statement for unitary moments of moments.

Proposition 4.2.17. *Let n be a positive integer and $x_1, \ldots, x_n$ be complex numbers. Then,*

$$\mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{n}\det\left(I-Ax_j\right)\right]=(x_1\cdots x_n)^N\sum_{\varepsilon_j\in\{-1,1\}}\frac{\prod_{j=1}^{n}x_j^{\varepsilon_j N}}{\prod_{1\le i\le j\le n}(1-x_i^{-\varepsilon_i}x_j^{-\varepsilon_j})}$$

Proposition 4.2.18. *Let n be a positive integer and $x_1, \ldots, x_n$ be complex numbers. Then,*

$$\mathbb{E}_{A\in SO(2N)}\left[\prod_{j=1}^{n} \det(I - Ax_j)\right] = (x_1 \cdots x_n)^N \sum_{\varepsilon_j \in \{-1,1\}} \frac{\prod_{j=1}^{n} x_j^{\varepsilon_j N}}{\prod_{1\le i<j\le n}(1 - x_i^{-\varepsilon_i} x_j^{-\varepsilon_j})}$$

The proofs of propositions 4.2.17 and 4.2.17 by Conrey et al. [45] require careful manipulations of multiple contour integral representations of the products of characteristic polynomials (much like those used in chapter 3 to prove theorem 3.1.4). However, the statements of propositions 4.2.17 and 4.2.17 were rederived by Bump and Gamburd [33] using their representation approach, yielding a simplified proof, as well as the identification between the combinatorial sums and the respective characters.

Once more, we will be needing n to be an even integer, and we will be picking the complex numbers $x_1, \ldots, x_n$ in a particular way, always lying on the unit circle in the complex plane.

4.3 Polynomial structure

In this section we prove the following proposition. This, together with results stated in sections 4.4 and 4.5 will prove theorem 4.1.1 and 4.1.2.

Proposition 4.3.1. *Let $G(N) = Sp(2N)$, or $G(N) = SO(2N)$, and $k, \beta \in \mathbb{N}$. Then* $\mathrm{MoM}_{G(N)}(k, \beta)$ *is a polynomial function of N.*

Proof. We make use of the expressions for averages through the different matrix groups due to Conrey et al. [45] that were introduced in section 4.2.4. The argument follows that for the moments of the moments of the characteristic polynomials of unitary matrices, presented in [15].

We begin with the symplectic case. We apply Fubini's Theorem to obtain:

$$\mathrm{MoM}_{Sp(2N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi} \mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{2k\beta} \det(I - Ax_j)\right] d\theta_1 \cdots d\theta_k, \tag{4.3.1}$$

where, by recalling observation (4.1.1):

$$\underline{x} = (\underbrace{e^{-i\theta_1},\ldots,e^{-i\theta_1}}_{\beta},\underbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}_{\beta},\underbrace{e^{-i\theta_2},\ldots,e^{-i\theta_2}}_{\beta},\underbrace{e^{i\theta_2},\ldots,e^{i\theta_2}}_{\beta},\ldots,\underbrace{e^{-i\theta_k},\ldots,e^{-i\theta_k}}_{\beta},\underbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}_{\beta}).$$

Then, by proposition 4.2.17, we can write the moments of moments in the following form.

$$\mathrm{MoM}_{Sp(2N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi} \sum_{\varepsilon_j\in\{-1,1\}} \frac{\prod_{j=1}^{2k\beta} x_j^{\varepsilon_j N}}{\prod_{1\le i\le j\le 2k\beta}(1 - x_i^{-\varepsilon_i} x_j^{-\varepsilon_j})} d\theta_1 \cdots d\theta_k.$$

Above, each summand appears to have a pole of finite order (when $x_i^{\varepsilon_i} = x_j^{-\varepsilon_j}$), but these cancel with zeros in the numerator when the sum is considered as a whole. This is clearly the case since the average of a product of polynomials is bounded [45]. Following this calculation, one may compute the resulting function by applying l'Hôpital's rule a finite number of times, which results in a polynomial function in the variables $e^{i\theta_1}, \ldots, e^{i\theta_k}$, and whose coefficients are themselves polynomials in N. Finally, after performing the integration over the $\theta_1, \ldots, \theta_k$, only the constant term of said polynomial survives, which as noted is a polynomial in N. This concludes the proof of proposition 4.3.1. The argument for the orthogonal case is completely analogous via proposition 4.2.18.

□

4.4 Results for the symplectic group $Sp(2N)$

Here we give the proof of the leading order behaviour and coefficient of $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ as described in theorem 4.1.1. The argument is split in to stages. Firstly, we give an expression for the symplectic moments of moments using symplectic Gelfand-Tsetlin patterns with constraints. Secondly, we observe that part of the pattern is *determined*, and hence only the *free* part plays a role. Finally, by essentially passing from a discrete to a continuous setting and using the lattice point counting results presented in section 4.2.3, we complete the proof.

4.4.1 A combinatorial representation

We begin with a combinatorial representation for $\mathrm{MoM}_{Sp(2N)}(k,\beta)$. Recall that in chapter 3, we showed that for $k,\beta \in \mathbb{N}$, $\mathrm{MoM}_{U(N)}(k,\beta)$ is equal to a count of *restricted* semistandard rectangular Young tableaux. Assiotis and Keating then used this interpretation to show that $\mathrm{MoM}_{U(N)}(k,\beta)$ is equal to the number of (full) non-negative Gelfand-Tsetlin patterns with top row $(N,\dots,N,0,\dots,0)$, satisfying some additional constraints, see proposition 4.2.6. We are able to recover a similar statement here, linking $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ to a count of *symplectic* patterns, with a certain top row, along with some additional constraints.

Proposition 4.4.1. *Let $k,\beta \in \mathbb{N}$. Then, $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ is equal to the number of $(4k\beta)$–symplectic Gelfand-Tsetlin patterns $P = \left(\lambda^{(i)}\right)_{i=1}^{4k\beta}$ with top row $\lambda^{(4k\beta)} = \langle N^{2k\beta}\rangle$, which moreover satisfy the following k constraints for $i = 1,\dots,k$:*

$$\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta}\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right] = \sum_{j=(2i-1)\beta+1}^{2i\beta}\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]. \tag{4.4.1}$$

We denote the set of such patterns by $GT_{Sp}(N;k;\beta)$.

Proof. As in proposition 4.3.1, by an application of Fubini's Theorem we have:

$$\mathrm{MoM}_{Sp(2N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi}\mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{2k\beta}\det\left(I - Ax_j\right)\right]d\theta_1\cdots d\theta_k, \tag{4.4.2}$$

with (using (4.1.1))

$$\underline{x} = (\underbrace{e^{-i\theta_1},\dots,e^{-i\theta_1}}_{\beta},\underbrace{e^{i\theta_1},\dots,e^{i\theta_1}}_{\beta},\underbrace{e^{-i\theta_2},\dots,e^{-i\theta_2}}_{\beta},\underbrace{e^{i\theta_2},\dots,e^{i\theta_2}}_{\beta},\dots,\underbrace{e^{-i\theta_k},\dots,e^{-i\theta_k}}_{\beta},\underbrace{e^{i\theta_k},\dots,e^{i\theta_k}}_{\beta}). \tag{4.4.3}$$

Now, we make use of proposition 4.2.14 along with definition 4.2.11 to rewrite the integrand in (4.4.2) as follows, where the signature determining the set SP_ν is $\nu = \langle N^{2k\beta}\rangle \in \mathrm{S}^+_{2k\beta}$.

$$\begin{aligned}
&\mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{2k\beta}\det\left(I - Ax_j\right)\right] \\
&= \sum_{P\in SP_{\langle N^{2k\beta}\rangle}} \prod_{i=1}^{n} x_i^{\sum_{j=1}^{i}\lambda_j^{(2i)} - 2\sum_{j=1}^{i}\lambda_j^{(2i-1)} + \sum_{j=1}^{i-1}\lambda_j^{(2i-2)}} \qquad (4.4.4)\\
&= \sum_{P\in SP_{\langle N^{2k\beta}\rangle}} \prod_{j=1}^{\beta} e^{-i\theta_1\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}\\
&\quad\times \prod_{j=\beta+1}^{2\beta} e^{i\theta_1\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}\\
&\quad\times \prod_{j=2\beta+1}^{3\beta} e^{-i\theta_2\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}\\
&\quad\times \prod_{j=3\beta+1}^{4\beta} e^{i\theta_2\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}\\
&\quad\times\cdots\\
&\quad\times \prod_{j=(2k-2)\beta+1}^{(2k-1)\beta} e^{-i\theta_k\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}\\
&\quad\times \prod_{j=(2k-1)\beta+1}^{2k\beta} e^{i\theta_k\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}. \qquad (4.4.5)
\end{aligned}$$

Hence, when we perform the integration over the θ, we will make constant use of the fact that

$$\frac{1}{2\pi}\int_0^{2\pi} e^{is\theta}d\theta = \delta_{s=0}. \qquad (4.4.6)$$

The k Kronecker-δ functions that emerge exactly correspond to the constraints found in (4.4.1). One can read this off from the decomposition in (4.4.5); for example the integral over θ_1 will lead to

$$\delta_{\sum_{j=1}^{\beta}\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right] = \sum_{j=\beta+1}^{2\beta}\left[\sum_{l=1}^{j}\lambda_l^{(2j)} - 2\sum_{l=1}^{j}\lambda_l^{(2j-1)} + \sum_{l=1}^{j-1}\lambda_l^{(2j-2)}\right]}. \qquad (4.4.7)$$

Thus, the statement of the proposition readily follows. □

We now make the simple observation that the form of the top signature $\langle N^{2k\beta}\rangle$ essentially fixes the top right triangle of a pattern in $GT_{Sp}(N;k;\beta)$, see figure 4.11. This is similar to the unitary case: there the top row froze the top left and right triangles in the full Gelfand-Tsetlin pattern. Henceforth, it is sufficient to just work with the 'free' (unfrozen) region. In order to formalize this argument, it is convenient to have the following definition, which in essence defines a relabelling of the 'free' coordinates (figure 4.12 may be a useful reference for the definition).

Definition 4.4.2. *Consider the following set of integer arrays* $\left(y^{(i)}\right)_{i=1}^{4k\beta-1} \in \mathbb{Z}^{k\beta(2k\beta+1)}$*, which we denote by* $\mathfrak{I}_{Sp}(N;k;\beta)$*, and which additionally satisfy the following conditions,*

(i) for all $1 \le i \le 2k\beta$, $y^{(i)}, y^{(4k\beta-i)} \in \mathrm{S}^+_{\lfloor\frac{i+1}{2}\rfloor}$,

(ii) both $\left(y^{(i)}\right)_{i=1}^{2k\beta}$ *and* $\left(y^{(4k\beta-i)}\right)_{i=1}^{2k\beta}$ *form* $(2k\beta)$*–symplectic Gelfand-Tsetlin patterns,*

(iii) $0 \le y_j^{(i)} \le N$ *for any valid* i, j,

(iv) the rows $\left(y^{(i)}\right)_{i=1}^{4k\beta-1}$ *fulfil the following constraints:*

In the case k is even, let $i = 1, \ldots, \frac{k}{2}$ *(with* $y^{(0)}, y^{(4k\beta)} \equiv 0$*). Then,*

$$\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta} \left[\sum_{l=1}^{j} y_l^{(2j)} - 2\sum_{l=1}^{j} y_l^{(2j-1)} + \sum_{l=1}^{j-1} y_l^{(2j-2)} \right] = \sum_{j=(2i-1)\beta+1}^{2i\beta} \left[\sum_{l=1}^{j} y_l^{(2j)} - 2\sum_{l=1}^{j} y_l^{(2j-1)} + \sum_{l=1}^{j-1} y_l^{(2j-2)} \right], \tag{4.4.8}$$

and

$$\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta} \left[\sum_{l=1}^{j} y_l^{(4k\beta-2j)} - 2\sum_{l=1}^{j} y_l^{(4k\beta-2j+1)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+2)} \right] = \sum_{j=(2i-1)\beta+1}^{2i\beta} \left[\sum_{l=1}^{j} y_l^{(4k\beta-2j)} - 2\sum_{l=1}^{j} y_l^{(4k\beta-2j+1)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+2)} \right]. \tag{4.4.9}$$

While, when k is odd we have the same constraints as above for $i = 1, \ldots, \frac{k-1}{2}$ *along with:*

$$\sum_{j=(k-1)\beta+1}^{k\beta} \left[\sum_{l=1}^{j} y_l^{(2j)} - 2\sum_{l=1}^{j} y_l^{(2j-1)} + \sum_{l=1}^{j-1} y_l^{(2j-2)} \right] = \sum_{j=(k-1)\beta+1}^{k\beta} \left[\sum_{l=1}^{j} y_l^{(4k\beta-2j)} - 2\sum_{l=1}^{j} y_l^{(4k\beta-2j+1)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+2)} \right]. \tag{4.4.10}$$

Observe that, for both k odd and even there are a total of k constraints.

Condition (i) within definition 4.4.2 follows directly from proposition 4.4.1, as does (iii). Note that since the upper right triangle is frozen, see figure 4.11, the 'free' triangle can essentially be partitioned in to two copies of a sub-symplectic Gelfand-Tsetlin patterns, where one has been reflected in the x-axis and 'glued' on to the top row of the other. This is exactly the content of (ii). The constraints given in (iv) (i.e. (4.4.8)–(4.4.10)) are precisely the same as (4.4.1), just translated in to this new co-ordinate system.

We formalize this relabelling by the following natural bijection between $GT_{Sp}(N;k;\beta)$ and $\mathfrak{I}_{Sp}(N;k;\beta)$:

$$\mathfrak{B}_{Sp} : GT_{Sp}(N;k;\beta) \longrightarrow \mathfrak{I}_{Sp}(N;k;\beta). \tag{4.4.11}$$

This can be seen as follows, and again we recommend using figure 4.12 to accompany this description. Let $(\lambda^{(i)})_{i=1}^{4k\beta} \in GT_{Sp}(N;k;\beta)$. Observe that, by the interlacing $\lambda^{(4k\beta-1)} \prec \langle N^{2k\beta} \rangle = \lambda^{(4k\beta)}$, we have a single free coordinate (the left most element of the second row from the top):

$$\lambda_1^{(4k\beta-1)}, \ldots, \lambda_{2k\beta-1}^{(4k\beta-1)} \equiv N,$$
$$0 \le \lambda_{2k\beta}^{(4k\beta-1)} \le N.$$

We thus relabel $y_1^{(4k\beta-1)} = \lambda_{2k\beta}^{(4k\beta-1)}$. Secondly, again due to the interlacing $\lambda^{(4k\beta-2)} \prec \lambda^{(4k\beta-1)}$, we have:

$$\lambda_1^{(4k\beta-2)}, \ldots, \lambda_{2k\beta-2}^{(4k\beta-2)} \equiv N$$

and moreover,

$$y_1^{(4k\beta-1)} = \lambda_{2k\beta}^{(4k\beta-1)} \leq \lambda_{2k\beta-1}^{(4k\beta-2)} \leq N.$$

We write $y_1^{(4k\beta-2)} = \lambda_{2k\beta-1}^{(4k\beta-2)}$. We continue relabelling in this fashion up to (and including) $\lambda^{(2k\beta+1)}$ (after which no coordinates are necessarily fixed to equal N) and finally, we put $(y^{(i)})_{i=1}^{2k\beta} \equiv (\lambda^{(i)})_{i=1}^{2k\beta}$. Clearly, the map $\mathfrak{B}_{Sp}$ described above is invertible. Thus, by making use of proposition 4.4.1 we obtain the following.

Proposition 4.4.3. *Let $k, \beta \in \mathbb{N}$. Then,*

$$\mathrm{MoM}_{Sp(2N)}(k, \beta) = \#\mathfrak{I}_{Sp}(N; k; \beta).$$

Such a proposition should be reminiscent of the result of Assiotis and Keating, see (4.2.12).

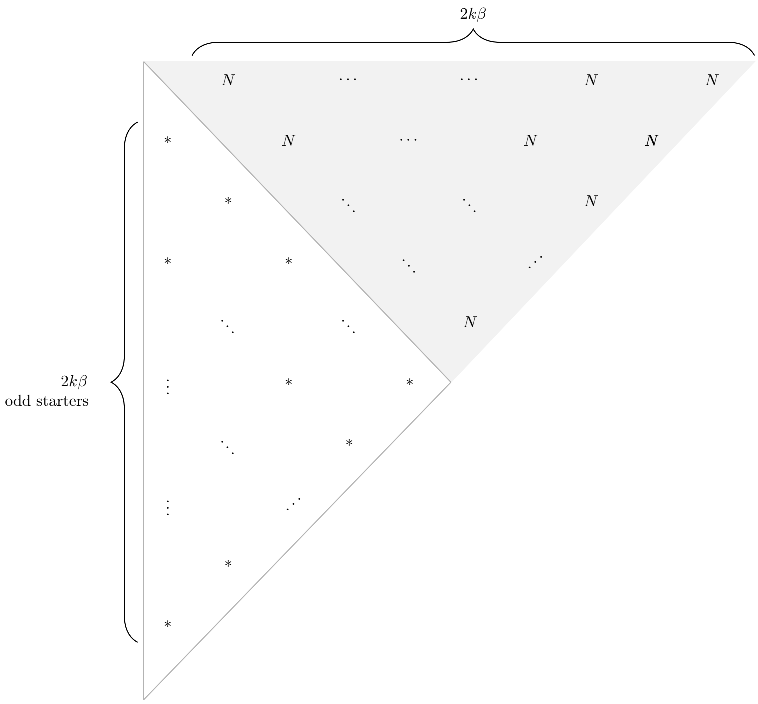


Figure 4.11: Figure depicting the fixed region of $P \in SP_{\langle N^{2k\beta}\rangle}$, a $(4k\beta)$–symplectic Gelfand-Tsetlin pattern. The shaded area represents the fixed region, whilst the unshaded region shows which elements have some freedom in the values that they can take.

4.4.2 Asymptotics and the leading order coefficient

To conclude the proof, we require some final definitions and notation, which will also be useful for the orthogonal case in section 4.5. The overall goal is to determine a continuous analogous to the discrete setting of the previous section. This will eventually permit us to use theorem 4.2.16 to prove theorem 4.1.1. We consider the continuous Weyl chamber:

$$\mathrm{W}_N = \{x = (x_1, \ldots, x_N) \in \mathbb{R}^N : x_1 \geq \cdots \geq x_N\} \tag{4.4.12}$$

and also let $\mathrm{W}_N^+ := \mathrm{W}_N \cap \mathbb{R}_+^N$. Thus W_N and W_N^+ act like continuous versions of the sets of signatures S_N and S_N^+.

We say that $y \in \mathrm{W}_N$ and $x \in \mathrm{W}_{N+1}$ interlace if exactly the inequalities (4.2.1) (from the discrete

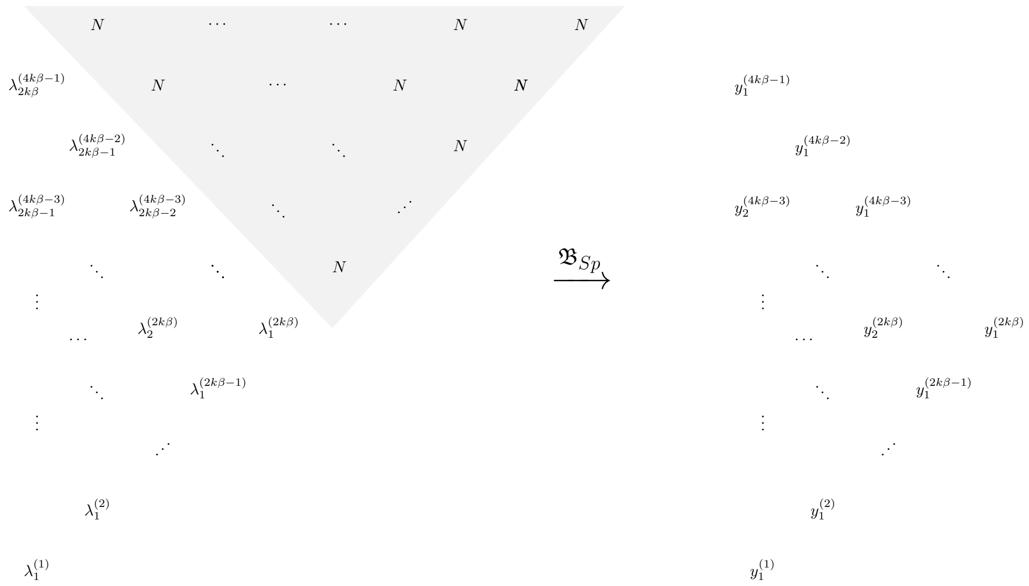


Figure 4.12: Representation of the bijection $\mathfrak{B}_{Sp} : GT_{Sp}(N; k; \beta) \longrightarrow \mathfrak{I}_{Sp}(N; k; \beta)$, giving the relabelling of the coordinates.

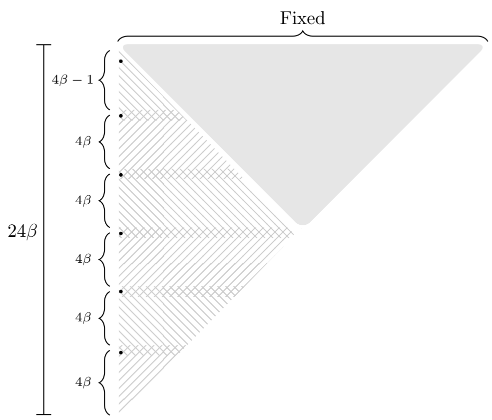


Figure 4.13: Pictorial representations of how the index set $\mathcal{S}_{(k,\beta)}^{Sp}$, and hence the diagram given by $\mathcal{V}_{(k,\beta)}^{Sp}$, for general integer β and $k = 6$ are constructed. A pair (i, j) in index set $\mathcal{S}_{(k,\beta)}^{Sp}$ represents any non-fixed element i in row j of the continuous pattern $\mathcal{V}_{(k,\beta)}^{Sp}$ above, except for the elements depicted by $\bullet$. These are not included in $\mathcal{S}_{(k,\beta)}^{Sp}$, since these are chosen to be fixed by the linear equations. The overlap in the pattern shows the 5 rows $x^{(4\beta)}, \dots, x^{(20\beta)}$ where the constraints overlap.

setting) are satisfied and we also write $y \prec x$ (similarly for $y \in \mathrm{W}_N$ and $x \in \mathrm{W}_N$). The definitions of continuous half-patterns and continuous symplectic and orthogonal Gelfand-Tsetlin patterns are completely analogous to the discrete setting (we simply replace S_i by W_i).

We consider the following index set, which encodes a subset of the elements in the patterns in $\mathfrak{I}_{Sp}(N; k; \beta)$ resulting from applying the relabelling. In particular, the $\mathcal{S}_{(k,\beta)}^{Sp}$ will reference every element of $\mathfrak{I}_{Sp}(N; k; \beta)$, *except* for k purposely chosen elements. Figure 4.13 may be useful to elucidate the definition.

$$
\begin{aligned}
\mathcal{S}^{Sp}_{(k,\beta)} \coloneqq \Bigg\{ &(m,n) : 1 \le m \le \left\lfloor \frac{n+1}{2} \right\rfloor \text{ and } 1 \le n \le 2k\beta; \\
&\text{or } 1 \le m \le \left\lfloor \frac{4k\beta - n + 1}{2} \right\rfloor \text{ and } 2k\beta + 1 \le n < 4k\beta - 1; \\
&n \neq 4\beta, 8\beta, \ldots, 4(k-1)\beta \Bigg\} \\
\cup \Bigg\{ &(m, 4n\beta) : 1 \le m \le 2n\beta - 1 \text{ and } 1 \le n \le \left\lfloor \tfrac{k}{2} \right\rfloor; \\
&\text{or } 1 \le m \le 2(k-n)\beta - 1 \text{ and } \left\lfloor \tfrac{k}{2} \right\rfloor + 1 \le n < k \Bigg\}. \qquad (4.4.13)
\end{aligned}
$$

Thus, the pair (m,n) appears in $\mathcal{S}^{Sp}_{(k,\beta)}$ if and only if $y^{(n)}_m \in \mathfrak{I}_{Sp}(N;k;\beta)$, except for some particular choices of pairs (m,n), which we remove. The k missing pairs are precisely the encodings of $y^{(4\beta)}_{2\beta}$, $y^{(8\beta)}_{4\beta}$, …, $y^{(4(k-2)\beta)}_{4\beta}$, $y^{(4(k-1)\beta)}_{2\beta}$, and $y^{(4k\beta-1)}_{1}$; see figure 4.13 for a visual representation.

Observe that $\mathcal{S}^{Sp}_{(k,\beta)}$ has exactly $k\beta(2k\beta+1) - k$ elements. It is not a coincidence that this is the degree of the polynomial in the statement of theorem 4.1.1. Now define

$$
\mathcal{V}^{Sp}_{(k,\beta)} \coloneqq \{x^{(n)}_m \in \mathbb{R} : (m,n) \in \mathcal{S}^{Sp}_{(k,\beta)}, 0 \le x^{(n)}_m \le 1\} \subset \mathbb{R}^{k\beta(2k\beta+1)-k}, \qquad (4.4.14)
$$

alongside k elements defined as follows,

$$
\begin{aligned}
&x^{(n)}_{\frac{n}{2}} && \text{for } n = 4\beta, 8\beta, \ldots, 4\left\lfloor \tfrac{k}{2} \right\rfloor \beta, && (4.4.15) \\
&x^{(n)}_{\frac{4k\beta-n}{2}} && \text{for } n = 4(\left\lfloor \tfrac{k}{2} \right\rfloor + 1)\beta, \ldots, 4(k-1)\beta, && (4.4.16) \\
&x^{(4k\beta-1)}_{1} && \text{for } n = 4k\beta - 1, && (4.4.17)
\end{aligned}
$$

which are determined by the linear equations (4.4.8)–(4.4.10) (we simply solve for the relevant term) so that:

(i) $0 \le x^{(n)}_m \le 1$, for all $x^{(n)}_m$ described by (4.4.14)–(4.4.17),

(ii) $x^{(n)}, x^{(4k\beta-n)} \in \mathrm{W}^{+}_{\lfloor \frac{n+1}{2} \rfloor}$, for all $n = 1, \ldots, 2k\beta$,

(iii) both $(x^{(n)})^{2k\beta}_{n=1}$ and $(x^{(4k\beta-n)})^{2k\beta}_{n=1}$ form continuous $(2k\beta)$–symplectic Gelfand-Tsetlin patterns.

We call the index set corresponding to the 'determined' elements

$$
\mathcal{T}^{Sp}_{(k,\beta)} \coloneqq \{(m,n) : y^{(n)}_m \in \mathfrak{I}_{Sp}(N;k;\beta)\} \backslash \mathcal{S}^{Sp}_{(k,\beta)}.
$$

Observe that, $\mathcal{V}^{Sp}_{(k,\beta)}$ is convex as an intersection of hyperplanes. Moreover, $\mathcal{V}^{Sp}_{(k,\beta)}$ is contained in the cube $[0,1]^{k\beta(2k\beta+1)-k}$ and hence in a closed ball of radius $\sqrt{k\beta(2k\beta+1)-k}$.

Proof of theorem 4.1.1. The proof of the aspect of the theorem pertaining to the polynomial structure of the moments of moments was given in proposition 4.3.1. For the leading order coefficient term we observe that:

$$
\#\mathfrak{I}_{Sp}(N;k;\beta) = \#\left(\mathbb{Z}^{k\beta(2k\beta+1)-k} \cap \left(N\mathcal{V}^{Sp}_{(k,\beta)}\right)\right),
$$

where for a set $\mathcal{A}$, we write $N\mathcal{A} = \{Nx : x \in \mathcal{A}\}$ for is its dilate by a factor of N. Thus, from proposition 4.4.3 and theorem 4.2.16 with $\mathcal{S} = N\mathcal{V}^{Sp}_{(k,\beta)}$, we obtain:

$$\begin{aligned}\mathrm{MoM}_{Sp(2N)}(k,\beta) = \#\mathfrak{I}_{Sp}(N;k;\beta) &= \#\left(\mathbb{Z}^{k\beta(2k\beta+1)-k} \cap \left(N\mathcal{V}^{Sp}_{(k,\beta)}\right)\right) \\ &= \mathrm{vol}\left(N\mathcal{V}^{Sp}_{(k,\beta)}\right) + O_{k,\beta}\left(N^{k\beta(2k\beta+1)-k-1}\right).\end{aligned}$$

Since,

$$\mathrm{vol}\left(N\mathcal{V}^{Sp}_{(k,\beta)}\right) = N^{k\beta(2k\beta+1)-k}\,\mathrm{vol}\left(\mathcal{V}^{Sp}_{(k,\beta)}\right)$$

we have $\mathfrak{c}_{Sp}(k,\beta) = \mathrm{vol}\left(\mathcal{V}^{Sp}_{(k,\beta)}\right)$. It then suffices to prove that $\mathrm{vol}\left(\mathcal{V}^{Sp}_{(k,\beta)}\right) > 0$ which is the content of lemma 4.4.4 below. ☐

Proving the strict positivity of the constant $\mathfrak{c}_{Sp}(k,\beta)$ is important, because otherwise we simply have a bound for $\mathrm{MoM}_{Sp(2N)}(k,\beta)$. This task is also one of the more complicated parts of this paper. A crucial role is played by a number of figures which elucidate the argument.

Lemma 4.4.4. *Let $k,\beta \in \mathbb{N}$. Then,*

$$\mathfrak{c}_{Sp}(k,\beta) = \mathrm{vol}\left(\mathcal{V}^{Sp}_{(k,\beta)}\right) > 0. \tag{4.4.18}$$

Proof. We consider the following subset $\tilde{\mathcal{V}}^{Sp}_{(k,\beta)} \subset \mathcal{V}^{Sp}_{(k,\beta)}$ defined as for $\mathcal{V}^{Sp}_{(k,\beta)}$, but additionally we require both that $0 < x^{(n)}_m < 1$ and the interlacing is strict:

$$x^{(n+1)}_m > x^{(n)}_m > x^{(n+1)}_{m+1},$$

the above holding also for $x^{(n)}_m$ for $(m,n) \in \mathcal{T}^{Sp}_{(k,\beta)}$ as given in (4.4.15)–(4.4.17). Now, we claim that if there exists at least one element in $\tilde{\mathcal{V}}^{Sp}_{(k,\beta)}$ then $\mathrm{vol}\left(\tilde{\mathcal{V}}^{Sp}_{(k,\beta)}\right) > 0$ since $\tilde{\mathcal{V}}^{Sp}_{(k,\beta)}$ contains a small cube around this element (this clearly implies the statement of the lemma). This can easily be seen as follows. Take a continuous pattern $P = (z^{(n)}_m)_{(m,n)\in\mathcal{S}^{Sp}_{(k,\beta)}} \in \tilde{\mathcal{V}}^{Sp}_{(k,\beta)}$ and let D be the minimal distance between any two elements $z^{(n)}_m$ of P, or between $z^{(n)}_m$ and 0 or 1 (including those $z^{(n)}_m$ corresponding to the points described in (4.4.15),(4.4.16), and (4.4.17)). We observe that if we change each of the coordinates $(z^{(n)}_m)_{(m,n)\in\mathcal{S}^{Sp}_{(k,\beta)}}$ by at most some positive ϵ, then there exists some constant $C_{k,\beta}$ such that the extra values given by $z^{(n)}_m$ for $(m,n) \in \mathcal{T}^{Sp}_{(k,\beta)}$ change by at most $C_{k,\beta} \times \epsilon$. Thus, if $\epsilon = \epsilon(\mathrm{D})$ is small enough we get that $(z^{(n)}_m)_{(m,n)\in\mathcal{S}^{Sp}_{(k,\beta)}} + [-\epsilon,\epsilon]^{k\beta(2k\beta+1)-k} \subset \tilde{\mathcal{V}}^{Sp}_{(k,\beta)}$.

It then suffices to exhibit such an element. We observe that the constraints described in (4.4.8)–(4.4.10) essentially fall in to four distinct categories, hereafter types 1, 2, 3, and 4. These can be visualised as in figures 4.15a, 4.15b, 4.16a, and 4.16b, see also figure 4.14. In each diagram, the shaded triangular region shows the part of the pattern $P \in GT_{Sp}(N;k;\beta)$ which was fixed to be N, and the numbers shown to the left of the pattern are the 'row coefficient'. One can reconstruct the particular constraint described in each figure by first multiplying each row sum by its row coefficient, and the summing the resulting expressions for the top half of the pattern, and equating it with the sum for the bottom half of the pattern (the 'symmetry line' is given by the row with row coefficient 0). For example, figure 4.15a shows the following constraint, ($k = 1, \beta = 3$ in (4.4.10)),

$$\sum_{j=1}^{3}\left[\sum_{l=1}^{j} y_l^{(2j)} - 2\sum_{l=1}^{j} y_l^{(2j-1)} + \sum_{l=1}^{j-1} y_l^{(2j-2)}\right] = \sum_{j=1}^{3}\left[\sum_{l=1}^{j} y_l^{(12-2j)} - 2\sum_{l=1}^{j} y_l^{(13-2j)} + \sum_{l=1}^{j-1} y_l^{(14-2j)}\right]$$

or, equivalently,

$$2\sum_{j=1}^{5}(-1)^j r^{(j)} = 2\sum_{j=7}^{11}(-1)^j r^{(j)},$$

where $r^{(j)}$ is the sum of the elements in row j.

We will first show that it is possible to exhibit an element with strict interlacing and positive distances from 0 and 1 for each of the four types of constraints. We will then argue that these constructions are compatible and yield an element of $\tilde{\mathcal{V}}^{Sp}_{(k,\beta)}$; this fact is not entirely trivial since two consecutive constraints (e.g. $i = 1, 2$ in (4.4.8)) overlap in a single row, see figures 4.19 and 4.21, and clearly interlacing still plays a role.

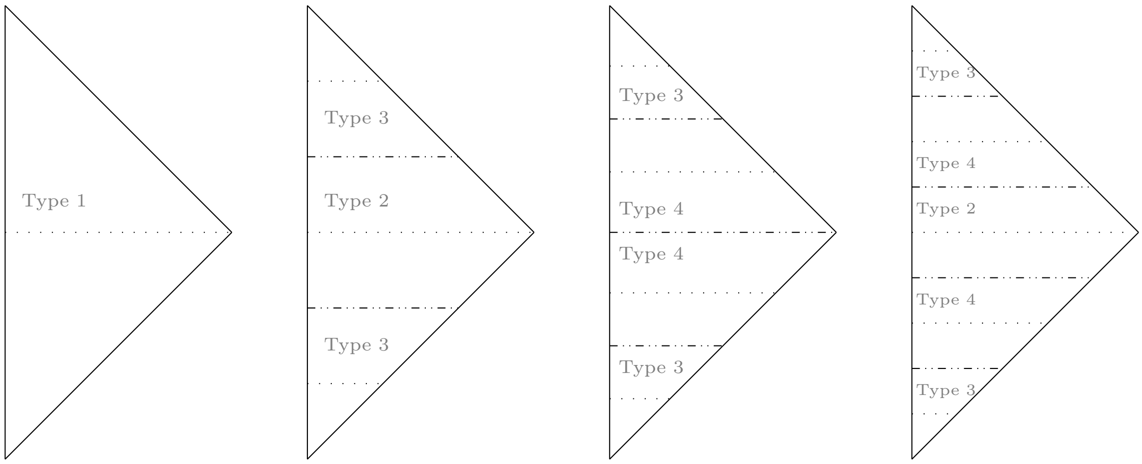


Figure 4.14: Examples of shapes of the constraint types $1, 2, 3, 4$. In each case, the dotted lines show the position of the 'overlaps' within a type; the dashed-dotted lines give the boundaries between different types of constraints (and hence where an 'overlap' will also occur). The left-most diagram shows the constraint of type 1, which only occurs for $k = 1$. The second diagram shows two types of constraint. The middle, pentagonal, shape represents type 2, and the top and bottom triangles are type 3s. This combination of shapes occurs for $k = 3$ (for $k = 2$, the type 2 in the middle disappears, leaving two stacked type 3s). When $k = 4$, type 4 is introduced, and the layout is given in the third triangle. When $k = 5$, we see a mixture of types $2, 3$, and 4 as shown in the rightmost triangle. Thereafter, for higher k, the triangles will resemble either of the two rightmost layouts above, depending on the parity of k, except that the number of 'stacked' type 4 constraints will increase.

The first two types of constraints, types 1 and 2 are shown in figures 4.15a and 4.15b. Type 1 only occurs for $k = 1$ and figure 4.15a shows an example for $k = 1$ and $\beta = 3$. In this case, only (4.4.10) is relevant. The row sum for the $(2k\beta)$th row appears on both sides of (4.4.10), and so this contribution is cancelled out. All the remaining row sums have a coefficient of either $+2$ or -2 in (4.4.10), and precisely which coefficient corresponds to which row can be seen on the left in figure 4.15a. Similarly, type 2 is the generalization of type 1 but for $k > 1$, odd. For these larger values of odd k, the shape of the constraint changes from triangular to pentagonal, but always occurs in the centre portion of the overall pattern. Figure 4.15b shows the type 2 for $k = 3$ and $\beta = 2$. For both said constraints, it is easy to exhibit such an element by symmetry: simply pick the lower half-pattern to have strict interlacing and coordinates a positive distance away from 0 and 1 and reflect in the symmetry line (c.f. the row with factor 0 in either figure).

Constraints of types 3 and 4 are shown in figures 4.16a and 4.16b. Type 3 occurs for $k \geq 2$ and corresponds to (4.4.8), (4.4.9) for $i = 1$ - henceforth we say that a 'lower' type 3 pattern comes from setting $i = 1$ in (4.4.8); whereas an 'upper' type 3 pattern is the analogous object using (4.4.9). The shape of type 3 is always triangular and covers the lowermost and uppermost portion of the overall pattern (c.f. the top and bottom patterned triangles in figure 4.13). figure 4.16a shows type 3 for $k = 2$, $\beta = 2$, and in particular the lower version, corresponding to $i = 1$ in (4.4.8). Note now that all rows have coefficients that are either ± 2, except for the top (resp. for the upper version, bottom) row which gets a coefficient of 1. Type 4 occurs for $k \geq 4$ and represents $i > 1$ in (4.4.8) and (4.4.9); the terms 'lower' and 'upper' are used just as for type 3. Type 4 constraints are trapezoidal, and an example of the lower type is drawn in figure 4.16b for $k = 4, \beta = 2$. Here (as for the general case) the row coefficients are once again symmetrical around the 'overlap' row. For type 3 and type 4 constraints, exhibiting an element is more complicated than type 1 and 2, and we proceed as follows.

In case of a constraint of type 3, we split the configuration as in figure 4.17. This results in a type 1 constraint and a new constraint, hereafter referred to as type 5. In figure 4.17, the top diagram gives an example of this splitting for a general form of a lower type 3, and the particular form of the resulting type 5 constraint is shown in the bottom diagram. For the constraint of type 1 resulting from the splitting, we will again use symmetry. However, the constraint of type 5 requires a separate argument.

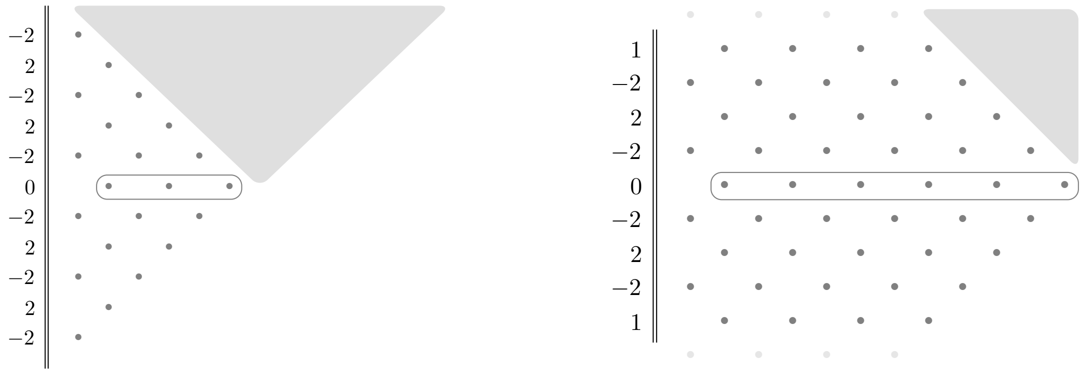
Take $\epsilon > 0$ to be very small according to k and β. We pick the lower half-pattern of constraint type 1, see figure 4.17, so that the distances between any two nearest coordinates, and between the closest coordinate to 0 (and respectively 1), is strictly positive and at most ϵ. We then use reflection through the middle row (the row with 0 as its row coefficient) for the upper half-pattern. We then proceed to the constraint of type 5. We again pick the coordinates, except the largest one (see circled element in figure 4.17) to be at a strictly positive distance of at most ϵ to its neighbour coordinates, and to the edge of the upper half-pattern of the constraint of type 1. Then, the total sum corresponding to constraint type 5 excluding the largest coordinate, which we have yet to pick, is negative and at most $c_{k,\beta} \times \epsilon$ in absolute value, for some constant $c_{k,\beta}$ depending only on k and β. We can then pick the largest coordinate so that this weighted sum over all coordinates is zero as long as $c_{k,\beta} \times \epsilon < 1$.

In order to deal with a constraint of type 4 we split it into a constraint of type 2 and type 5, see figure 4.18. There, the general 'lower' type 4 constraint is shown, along with the method of splitting. One may use exactly the same method described above for type 3 constraints.

Finally, we need to argue that using the procedures above is compatible with putting constraints together. For example, type 3 and type 4 constraints overlap, see figures 4.19 and 4.20, and two type 4 constraints also may overlap, see figures 4.21 and 4.22. With a mixture of type 3 and type 4 (the case for a mixture of two type 4s is analogous), if we use the algorithm above to satisfy the constraint of type 3, then the interlacing forces the coordinates at the edges of the next constraint of type 4 to be 'large', of the order of $c_{k,\beta} \times \epsilon$ for the constant $c_{k,\beta}$ described above. This then forces the largest coordinate of the constraint of type 5 coming from the splitting of the constraint of type 4 to be $\tilde{c}_{k,\beta} \times \epsilon$ for some (possibly much) larger constant $\tilde{c}_{k,\beta}$. However, we note that this does not present any real problems since we only need to apply this procedure a finite number of times and thus as long as we pick ϵ small enough so that $c^*_{k,\beta} \times \epsilon < 1$ for some finite and fixed constant $c^*_{k,\beta}$, the result is as claimed. □

4.4.3 Asymptotics at the symmetry point

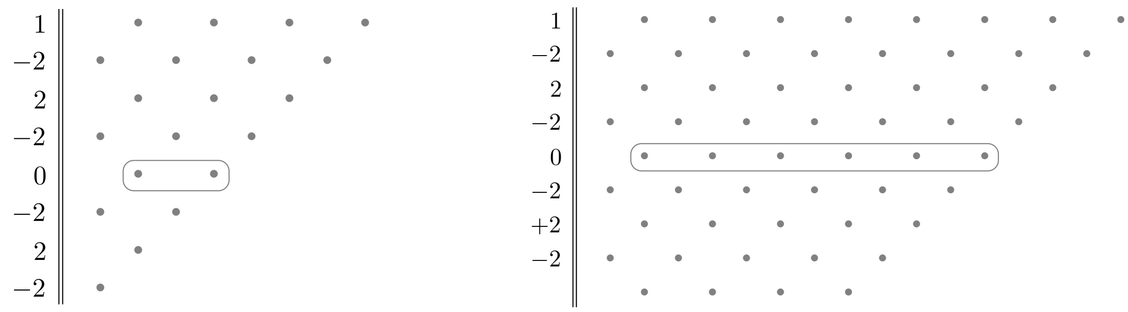
In this subsection we show how the method illustrated above can also be used to recover results of Keating and Snaith on the asymptotics of moments of the characteristic polynomial at the symmetry point, see [115] and theorem 2.1.7. The original proof involved the Selberg integral and asymptotics



(a) Example of constraint type 1. This occurs exclusively for $k = 1$, and is drawn for $k = 1$, $\beta = 3$. The circled coordinates are those which feature in the 'overlap' of the constraint (4.4.10). The grey shaded area shows which elements are fixed to be N. The numbers on the left show the coefficient that appears against any given row sum in (4.4.10).

(b) Example of constraint type 2. This occurs for $k > 1$, k odd, and is drawn for $k = 3$, $\beta = 2$. The circled coordinates are those which feature in the 'overlap' of constraint (4.4.10) (i.e. those in row $2k\beta$). The grey shaded area shows the lower part of the section which is fixed to be N, and the number on the left show the coefficient that appears against any given row sum in (4.4.10).

Figure 4.15: Figures showing constraints of type 1 and 2 for the symplectic case.



(a) Example of constraint type 3. This occurs for $k \geq 2$, and is partly drawn for $k = 2$, $\beta = 2$. The figure depicts the first constraint (i.e. $i = 1$ in (4.4.8)) and the boxed elements are those which appear in the 'overlap' of said constraint. Note that by reflecting this diagram in the x-plane, one gets a figure for the last constraint, i.e. $i = 1$ in (4.4.9). The numbers on the left are the coefficients that appear against the relevant row in (4.4.8), with $i = 1$.

(b) Example of constraint type 4. This occurs for $k \geq 4$ and is drawn for $k = 4, \beta = 2$ and depicts the (lower) constraint for $i = 2$ in (4.4.8). The boxed elements are those which feature in the 'overlap' of the described constraint, and the numbers on the left give the coefficient of a given row sum in (4.4.8). Note that the shape and row coefficients of the upper constraint can be seen by reflecting the diagram in the x-plane.

Figure 4.16: Figures showing constraints of type 3 and 4 for the symplectic case.

for the Barnes $\mathcal{G}$-function. More precisely we recover (c.f. theorem 2.1.7) that for $s \in \mathbb{N}$

$$\mathrm{M}_{Sp}(s) := \mathbb{E}_{A \in Sp(2N)} \left[\det (I - A)^s\right] = \mathrm{c}_{Sp}(s) N^{\frac{s(s+1)}{2}} + O_s\left(N^{\frac{s(s+1)}{2}-1}\right), \tag{4.4.19}$$

where the leading order coefficient is explicit:

$$\mathrm{c}_{Sp}(s) = \frac{1}{\prod_{j=1}^{s} (2j-1)!!}.$$

By applying proposition 4.2.14 with $x_i \equiv 1$ and inserting this into the combinatorial representation of definition 4.2.11 we obtain the following proposition.

Proposition 4.4.5. *Let $s \in \mathbb{N}$. $\mathrm{M}_{Sp}(s)$ is equal to the cardinality of the set $SP_{\langle N^s \rangle}$, namely the number of $(2s)$–symplectic Gelfand-Tsetlin patterns with top row $\langle N^s \rangle$.*

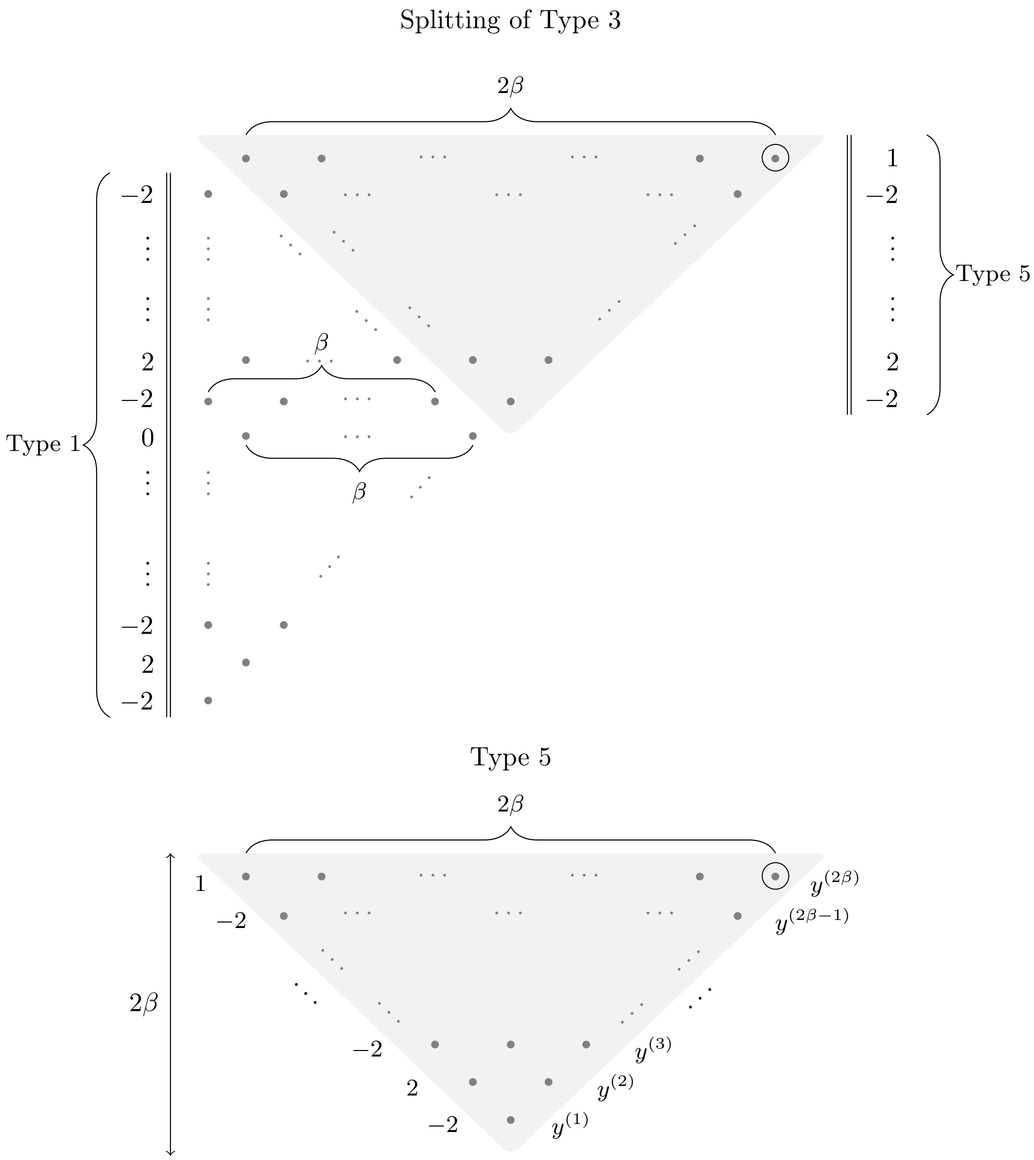


Figure 4.17: Figures giving the construction of a type 5 constraint, which comes from splitting a type 3 constraint (see figure 4.16a). This occurs for $k \geq 2$, and the version for a lower type 3 constraint (i.e. $i = 1$ in (4.4.8)) is drawn in the upper figure to show the situation for general β, and $k \geq 2$. The type 3 constraint is split in to one of type 1 (the unshaded region) and one of a new type, type 5 (the shaded region). The bottom figure shows explicitly the constraint of type 5, which forms a Gelfand-Tsetlin pattern $(y^{(i)})_{i=1}^{2\beta}$, where $y^{(i)} \in \mathrm{W}_i^+$ and $y^{(i)} \prec y^{(i+1)}$. In both diagrams, the circled top right element is the largest, and the numbers on either side show the row sum weightings for $i = 1$ in (4.4.8). The equivalent form for the upper version (i.e. $i = 1$ in (4.4.9)) can be seen by reflecting the top diagram in the x-plane.

As before, the form of the top row fixes the top right triangle of the pattern, see figure 4.23a. An analogous argument to that given in proposition 4.4.3 yields the following.

Proposition 4.4.6. *Let $s \in \mathbb{N}$. Then,*

$$\mathrm{M}_{Sp}(s) = N^{\frac{s(s+1)}{2}} \operatorname{vol}\left(\mathrm{V}_{Sp}(s)\right) + O_s\left(N^{\frac{s(s+1)}{2}-1}\right)$$

where the set $\mathrm{V}_{Sp}(s) \subset [0,1]^{\frac{s(s+1)}{2}}$ consists of joining two continuous half patterns of length s at the top row, as in the figure 4.23b.

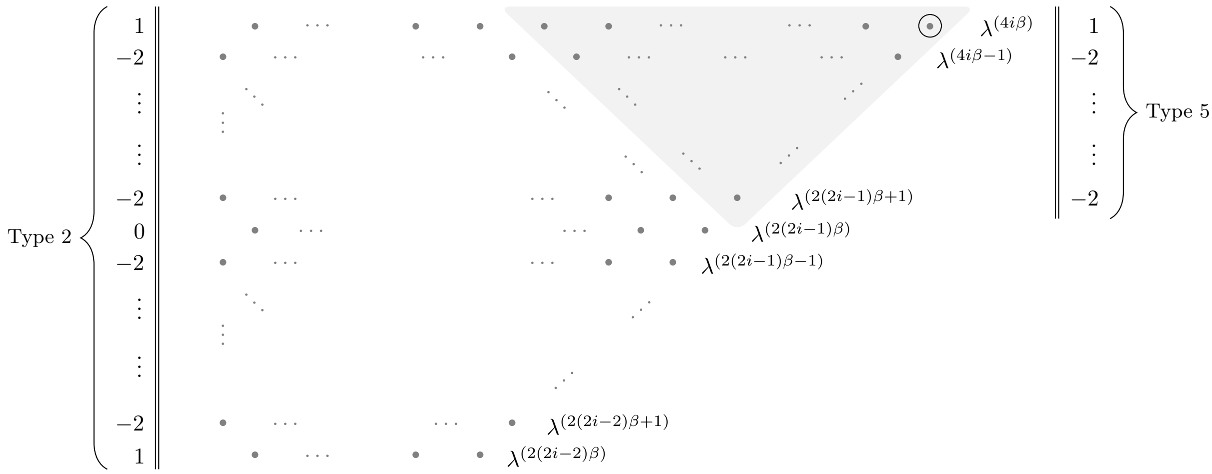


Figure 4.18: Figure showing splitting a type 4 constraint (see figure 4.16b) in to a type 2 and type 5. This occurs for $k \geq 4$, and the lower constraint for some $1 < i \leq \lfloor \frac{k}{2} \rfloor$ in (4.4.8) is drawn in the top figure for general $k \geq 4, \beta$, involving rows $\lambda^{(n)}$ for $n = 2(2i-2)\beta, \dots, 4i\beta$. The type 4 constraint is split in to one of type 2 (the unshaded region) and one of type 5 (the shaded region), see figure 4.17. The circled top right element is the largest, and the numbers on the far left and the far right give the row sum weightings as appearing in (4.4.8). The equivalent form for the upper version (i.e. $1 < i \leq \lfloor \frac{k}{2} \rfloor$ in (4.4.9)) can be seen by reflecting the diagram in the x-plane.

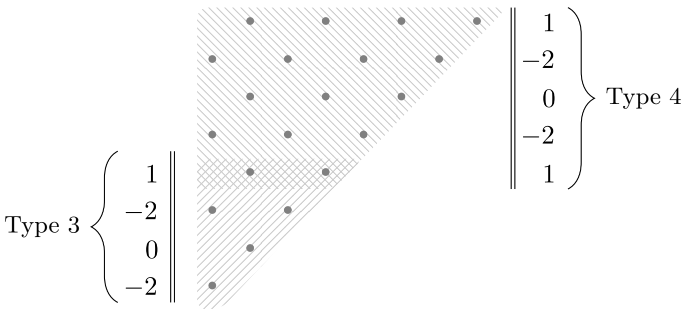


Figure 4.19: Example of a mixture of type 3 and type 4. This example shows $k = 4$, $\beta = 1$, and the interplay between $i = 1$ and $i = 2$ in (4.4.8) is demonstrated through the overlap between the two patterns. The corresponding diagram for $i = 1$ and $i = 2$ in (4.4.9) is simply the reflection of this diagram in the x-plane.

Thus, it suffices to show that the volume of $\mathrm{V}_{Sp}(s)$ can be computed explicitly and equals $\mathrm{c}_{Sp}(s)$. We require the following lemma (which is certainly well-known but we have not located this exact form in the literature).

Lemma 4.4.7. *Let $s \in \mathbb{N}$. The volume of a continuous half pattern of length s with non-negative coordinates and top row $\left(x_1, \dots, x_{\lfloor \frac{s+1}{2} \rfloor}\right) \in \mathrm{W}^+_{\lfloor \frac{s+1}{2} \rfloor}$, that we denote by $\mathrm{vol}_s\left(x_1, x_2, \dots, x_{\lfloor \frac{s+1}{2} \rfloor}\right)$, is given by:*

$$\mathrm{vol}_s\left(x_1, x_2, \dots, x_{\lfloor \frac{s+1}{2} \rfloor}\right) = \prod_{j=1}^{s} \frac{1}{(j-1)!!} \det\left(x_{\lfloor \frac{s+1}{2} \rfloor + 1 - i}^{2(j-1) + \mathbf{1}(s \text{ even})}\right)_{i,j=1}^{\lfloor \frac{s+1}{2} \rfloor}.$$

Proof. Direct computation by induction on s, using multi-linearity of the determinant. □

We finally have:

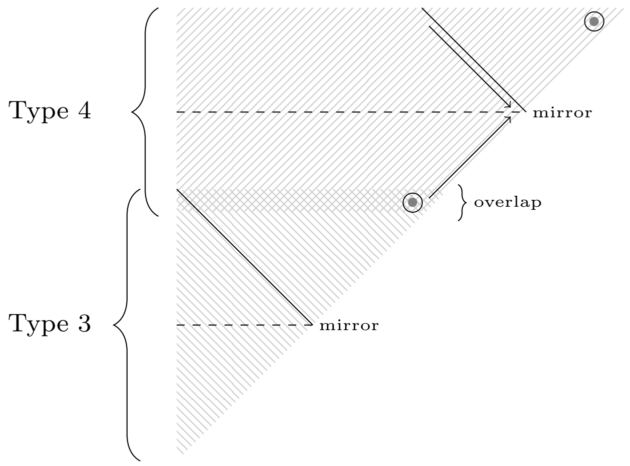


Figure 4.20: Example of combining a split type 3 and a split type 4. The dashed horizontal lines represent the lines of reflection, and the solid diagonal lines show where the splitting of the respective types occurs. The circled elements are the largest element for each section, and the arrows show the location of elements that, due to the interlacing, are forced the be 'large', and also direction of growth.

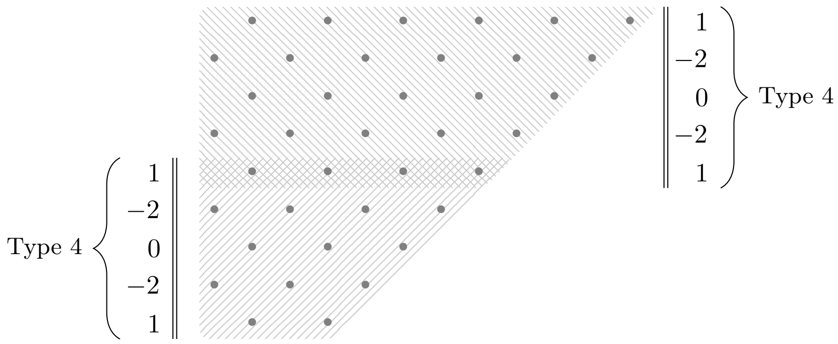


Figure 4.21: Example of a mixture of constraints of type 4. This figure is drawn for $k = 6, \beta = 1$ and depicts the mixture of constraints for $i = 2$ in (4.4.8) and (4.4.9).

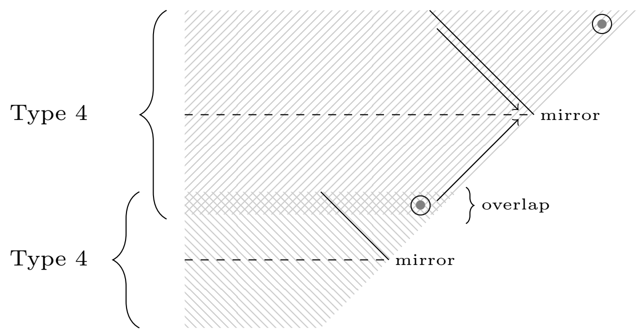


Figure 4.22: Example of combining two split type 4 constraints. The dashed horizontal lines represent the lines of reflection, and the solid diagonal lines show where the splitting of the respective types occurs. The circled elements are the largest element for each section, and the arrows show the location of elements that, due to the interlacing, are forced the be 'large', and also direction of growth.

Proposition 4.4.8. *Let $s \in \mathbb{N}$. Then,*

$$\operatorname{vol}\left(\mathrm{V}_{Sp}(s)\right) = \frac{1}{\prod_{j=1}^{s} (2j-1)!!}.$$

Proof. Recall that, see figure 4.23b, $\mathrm{V}_{Sp}(s)$ is obtained by joining at the top row two continuous half patterns with coordinates in $[0,1]$. We then calculate using lemma 4.4.7 and Andreief's identity:

$$\begin{aligned}
\operatorname{vol}\left(\mathrm{V}_{Sp}(s)\right) &= \int_{1\geq x_1\geq x_2\geq\cdots\geq x_{\lfloor\frac{s+1}{2}\rfloor}\geq 0} \operatorname{vol}_s\left(x_1,x_2,\ldots,x_{\lfloor\frac{s+1}{2}\rfloor}\right)^2 dx_1\cdots dx_{\lfloor\frac{s+1}{2}\rfloor}\\
&= \prod_{j=1}^{s}\left(\frac{1}{(j-1)!!}\right)^2 \det\left(\int_0^1 x^{2(i-1)+2(j-1)+2\mathbf{1}(s\text{ even})}dx\right)_{i,j=1}^{\lfloor\frac{s+1}{2}\rfloor}\\
&= \prod_{j=1}^{s}\left(\frac{1}{(j-1)!!}\right)^2 \det\left(\frac{1}{2\left(i+j-\frac{3}{2}+\mathbf{1}(s\text{ even})\right)}\right)_{i,j=1}^{\lfloor\frac{s+1}{2}\rfloor}.
\end{aligned}$$

In order to evaluate this further one uses the Cauchy determinant formula:

$$\det\left(\frac{1}{x_i-y_j}\right)_{i,j=1}^{n} = \frac{\prod_{i=2}^{n}\prod_{j=1}^{i-1}(x_i-x_j)(y_j-y_i)}{\prod_{i=1}^{n}\prod_{j=1}^{n}(x_i-y_j)}.$$

Applying this with,

$$x_i = 2i-\frac{3}{2}+\mathbf{1}(s\text{ even}),\ y_j = -2j+\frac{3}{2}-\mathbf{1}(s\text{ even})$$

and after some elementary manipulations we readily obtain the statement of the proposition. □

Remark. Similar arguments apply in the setting of $SO(2N)$, see [115] for the original proof.

(a) Figure showing a $(2s)$–symplectic Gelfand-Tsetlin pattern with top row (and hence top right triangle) fixed to be $\langle N^{2s}\rangle$.

(b) Figure showing the two continuous half patterns in $[0,1]$ joined at the top row which give $\mathrm{V}_{Sp}(s)$.

Figure 4.23: Figures showing both the general structure of the (discrete) symplectic half pattern, and the two continuous half patterns formed by the free coordinates joined at the top row.

4.4.4 Computing $\mathrm{MoM}_{Sp(2N)}(1,1)$

Before we move to address the special orthogonal case, we show how one can in fact use proposition 4.4.1 to fully compute $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ in the simplest case.

Let $k=\beta=1$. Then the statement of proposition 4.4.1 gives that

$$\mathrm{MoM}_{Sp(2N)}(1,1) = \#GT_{Sp}(N;1;1), \tag{4.4.20}$$

i.e. the first symplectic moments of moments is equal to the number of (4)–symplectic Gelfand-Tsetlin patterns $P = \left(\lambda^{(i)}\right)_{i=1}^{4}$ with top row (N, N), additionally satisfying the constraint

$$\lambda_1^{(2)} - 2\lambda_1^{(1)} = \lambda_1^{(4)} + \lambda_2^{(4)} - 2\lambda_1^{(3)} - 2\lambda_2^{(3)} + \lambda_1^{(2)}. \quad (4.4.21)$$

Hence $\mathrm{MoM}_{Sp(2N)}(1,1)$ is equal to the number of (4)–symplectic half patterns of the form

$$\begin{array}{ccccc} & N & & N & \\ \lambda_2^{(3)} & & \lambda_1^{(3)} & & \\ & \lambda_1^{(2)} & & & \\ \lambda_1^{(1)} & & & & \end{array}$$

Figure 4.24: The Gelfand-Tsetlin pattern relating to $\mathrm{MoM}_{Sp(2N)}(1,1)$.

with the usual interlacing requirements, and additionally (4.4.21). Using the interlacing, we see that $\lambda_1^{(3)}$ is forced to equal N, and further $0 \le \lambda_1^{(1)}, \lambda_1^{(2)}, \lambda_2^{(3)} \le N$. Hence (4.4.21) now reads

$$\lambda_1^{(1)} = \lambda_2^{(3)}. \quad (4.4.22)$$

Thus, we need to determine the number of (4)–symplectic half patterns of the following form.

$$\begin{array}{ccccc} & N & & N & \\ \lambda_1^{(1)} & & N & & \\ & \lambda_1^{(2)} & & & \\ \lambda_1^{(1)} & & & & \end{array}$$

Figure 4.25: The specialized Gelfand-Tsetlin pattern relating to $\mathrm{MoM}_{Sp(2N)}(1,1)$.

This is clearly equivalent to the following simple combinatorial count,

$$\#GT_{Sp}(N;1;1) = \#\{(\lambda_1^{(1)}, \lambda_1^{(2)}) : 0 \le \lambda_1^{(1)} \le \lambda_1^{(2)} \le N\} \quad (4.4.23)$$

$$= \frac{1}{2}(N+1)(N+2). \quad (4.4.24)$$

Hence

$$\mathrm{MoM}_{Sp(2N)}(1,1) = \frac{1}{2}(N+1)(N+2). \quad (4.4.25)$$

This calculation is independently verified in chapter 5.

4.5 Results for the special orthogonal group $SO(2N)$

We now give the proof of the asymptotic growth of the moments of the moments for $SO(2N)$. The key difference between the argument presented here and that of section 4.4 is that the leading elements in the odd rows of the half-patterns, the 'odd-starters', are now allowed to be positive or negative. This introduces an additional level of complexity due to the fact that now the constraints are not linear (they involve absolute values and signs).

Analogously to the symplectic case outlined in section 4.4, we break the proof down in to steps. Firstly we prove a proposition connecting the moments of moments to a count of restricted orthogonal Gelfand-Tsetlin patterns. Secondly, we note that the constraints on the patterns fix a triangular region, thus the count simplifies down to considering a subregion of the array. This induces a natural bijection between these constrained patterns and certain integer arrays. Finally, by considering the number of fixed parameters and moving to a continuous setting, we may apply theorem 4.2.16 to achieve theorem 4.1.2.

4.5.1 A combinatorial representation

The relevant combinatorial representation for the orthogonal group $SO(2N)$ is the following. Note that the statement of proposition 4.5.1 is almost the same as proposition 4.4.1, except for the extra consideration of the signs of the odd-starters, and now the sums concern absolute values of the elements.

Proposition 4.5.1. *Let $k, \beta \in \mathbb{N}$. Then $\mathrm{MoM}_{SO(2N)}(k,\beta)$ is equal to the number of $(4k\beta-1)$–orthogonal Gelfand-Tsetlin patterns $P = \left(\lambda^{(i)}\right)_{i=1}^{4k\beta-1}$ with top row either $\lambda^{(4k\beta-1)} = \langle N^{2k\beta}\rangle$ or $\lambda^{(4k\beta-1)} = \langle N^{2k\beta}\rangle^-$, which moreover satisfy each of the following k constraints for $i = 1,\ldots,k$:*

$$\begin{aligned}\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta} &\operatorname{sgn}(\lambda_j^{(2j-1)})\operatorname{sgn}(\lambda_{j-1}^{(2j-3)})\left[\sum_{l=1}^{j}|\lambda_l^{(2j-1)}| - 2\sum_{l=1}^{j-1}|\lambda_l^{(2j-2)}| + \sum_{l=1}^{j-1}|\lambda_l^{(2j-3)}|\right] \\ &= \sum_{j=(2i-1)\beta+1}^{2i\beta} \operatorname{sgn}(\lambda_j^{(2j-1)})\operatorname{sgn}(\lambda_{j-1}^{(2j-3)})\left[\sum_{l=1}^{j}|\lambda_l^{(2j-1)}| - 2\sum_{l=1}^{j-1}|\lambda_l^{(2j-2)}| + \sum_{l=1}^{j-1}|\lambda_l^{(2j-3)}|\right],\end{aligned} \tag{4.5.1}$$

where $\lambda^{(0)}, \lambda^{(-1)} \equiv 0$. We let $GT_{SO}(N;k;\beta)$ denote the set of such patterns. Further, we write $GT^+_{SO}(N;k;\beta)$ for the set of such constrained $(4k\beta-1)$–orthogonal patterns with top row $\langle N^{2k\beta}\rangle$, and $GT^-_{SO}(N;k;\beta)$ for the equivalent (but disjoint) set with top row $\langle N^{2k\beta}\rangle^- = (N,\ldots,N,-N)$.

Proof. The proof of proposition 4.5.1 follows entirely the same method as described in the proof of proposition 4.4.1, however we sketch it here for completeness. We focus on the average over the matrix group and, using the natural symmetry of the characteristic polynomial (4.1.1), we have

$$\mathrm{MoM}_{SO(2N)}(k,\beta) = \frac{1}{(2\pi)^k}\int_0^{2\pi}\cdots\int_0^{2\pi}\mathbb{E}_{A\in SO(2N)}\left[\prod_{j=1}^{2k\beta}\det(I-Ax_j)\right]d\theta_1\cdots d\theta_k, \tag{4.5.2}$$

with

$$\underline{x} = (\underbrace{e^{-i\theta_1},\ldots,e^{-i\theta_1}}_{\beta},\underbrace{e^{i\theta_1},\ldots,e^{i\theta_1}}_{\beta},\underbrace{e^{-i\theta_2},\ldots,e^{-i\theta_2}}_{\beta},\underbrace{e^{i\theta_2},\ldots,e^{i\theta_2}}_{\beta},\ldots,\underbrace{e^{-i\theta_k},\ldots,e^{-i\theta_k}}_{\beta},\underbrace{e^{i\theta_k},\ldots,e^{i\theta_k}}_{\beta}). \tag{4.5.3}$$

Using proposition 4.2.15 alongside definition 4.2.13, we rewrite the integrand in (4.5.2) as follows, where the signature determining the set SO_ν is $\nu = \langle N^{2k\beta}\rangle \in \mathrm{S}^+_{2k\beta}$.

$$\begin{aligned}&\mathbb{E}_{A\in Sp(2N)}\left[\prod_{j=1}^{2k\beta}\det\left(I-Ax_j\right)\right] \\ &= \sum_{P\in OP_{\langle N^{2k\beta}\rangle}\cup OP_{\langle N^{2k\beta}\rangle^-}}\prod_{j=1}^{2k\beta} x_j^{\operatorname{sgn}(\lambda_j^{(2j-1)})\operatorname{sgn}(\lambda_{j-1}^{(2j-3)})\left[\sum_{l=1}^{j}|\lambda_l^{(2j-1)}|-2\sum_{l=1}^{j-1}|\lambda_l^{(2j-2)}|+\sum_{l=1}^{j-1}|\lambda_l^{(2j-3)}|\right]}.\end{aligned} \tag{4.5.4}$$

Substituting in (4.5.3) and integrating over the θ, making copious use of the fact that

$$\frac{1}{2\pi}\int_0^{2\pi} e^{is\theta}d\theta = \delta_{s=0}, \tag{4.5.5}$$

we recover the k Kronecker-δ which exactly correspond to the constraints found in (4.5.1). For example the integral over θ_1 will lead to

$$\delta_{\substack{\sum_{j=1}^{\beta} \operatorname{sgn}(\lambda_j^{(2j-1)}) \operatorname{sgn}(\lambda_{j-1}^{(2j-3)}) \left[\sum_{l=1}^{j} |\lambda_l^{(2j-1)} - 2\sum_{l=1}^{j-1} |\lambda_l^{(2j-2)}| + \sum_{l=1}^{j-1} |\lambda_l^{(2j-3)}|\right] = \\ \sum_{j=\beta+1}^{2\beta} \operatorname{sgn}(\lambda_j^{(2j-1)}) \operatorname{sgn}(\lambda_{j-1}^{(2j-3)}) \left[\sum_{l=1}^{j} |\lambda_l^{(2j-1)}| - 2\sum_{l=1}^{j-1} |\lambda_l^{(2j-2)}| + \sum_{l=1}^{j-1} |\lambda_l^{(2j-3)}|\right]}} . \tag{4.5.6}$$

Thus, the statement of the proposition readily follows. □

The case for $k = \beta = 1$ is separate from the general case. This is essentially due to the fact that in this particular situation, the limited number of non-fixed elements in the pattern means that the constraints (4.5.1) behave differently compared to the case for higher k, β. In particular, note that in the case of $GT_{SO}^{+}(N; 1; 1)$ the corresponding constraint does not fix any coordinate, as we see in the proof below. We handle this special case here.

Proposition 4.5.2. *We have that*

$$\mathrm{MoM}_{SO(2N)}(1, 1) = 2(N + 1).$$

Proof. By proposition 4.5.1,

$$\mathrm{MoM}_{SO(2N)}(1, 1) = |GT_{SO}^{+}(N; 1; 1)| + |GT_{SO}^{-}(N; 1; 1)|,$$

where here $GT_{SO}(N; 1; 1)$ is the set of all (3)–orthogonal Gelfand-Tsetlin patterns P with top row either (N, N) or $(N, -N)$, corresponding to the sets $GT_{SO}^{+}(N; 1; 1)$ and $GT_{SO}^{-}(N; 1; 1)$ respectively, and where the constraint (4.5.1) specialises to

$$\operatorname{sgn}(\lambda_1^{(1)})\lambda_1^{(1)} = \operatorname{sgn}(\lambda_2^{(3)}) \operatorname{sgn}(\lambda_1^{(1)})\lambda_1^{(1)}, \tag{4.5.7}$$

see also figure 4.26. The fact that there is only one 'free' parameter here, namely λ_1, is the key difference between this special case and the situation for general k, β. Hence, $|GT_{SO}^{+}(N; 1; 1)| = 2N + 1$ since in this case $\lambda_2^{(3)} = N$ and all values $0 \leq |\lambda_1^{(1)}| \leq N$ are valid. However, in the second case, we have that $\lambda_2^{(3)} = -N$, to (4.5.7) becomes

$$\operatorname{sgn}(\lambda_1^{(1)})\lambda_1^{(1)} = -\operatorname{sgn}(\lambda_1^{(1)})\lambda_1^{(1)}, \tag{4.5.8}$$

hence the only option satisfying constraint (4.5.8) is $\lambda_1^{(1)} \equiv 0$. Thus,

$$\mathrm{MoM}_{SO(2N+1)}(1, 1) = 2(N + 1).$$

$$\begin{array}{ccc} \multicolumn{3}{c}{P \in GT_{SO}^{+}(N; 1; 1)} \\ \hline\hline N & & N \\ & N & \\ \lambda_1^{(1)} & & \end{array} \qquad \begin{array}{ccc} \multicolumn{3}{c}{Q \in GT_{SO}^{-}(N; 1; 1)} \\ \hline\hline -N & & N \\ & N & \\ \lambda_1^{(1)} & & \end{array}$$

Figure 4.26: Cases for determining $\mathrm{MoM}_{SO(2N)}(1, 1)$. The relevant constraint is $\lambda_1^{(1)} = \lambda_1^{(1)} \cdot \operatorname{sgn}(\pm N)$.

□

Henceforth we assume that we are in the general case (i.e. we exclude the case $k = \beta = 1$). Then, we note that by requiring the top row of the pattern P to be either $\langle N^{2k\beta} \rangle$ or $\langle N^{2k\beta} \rangle^{-}$, the top right triangle of $GT_{SO}(N;k,\beta)$ is also determined, as shown in figure 4.27, just as in the symplectic case. We now introduce notation which captures the sign of the odd starters for a given pattern $P \in GT_{SO}(N;k;\beta)$; figure 4.28 may be a useful accompaniment. Note that the ability of the odd starters to be positive or negative is one of the key differences between the orthogonal and the symplectic case.

We consider the following decomposition of $GT_{SO}(N;k;\beta)$ into the disjoint union:

$$GT_{SO}(N;k;\beta) = \bigcup_{\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}} GT_{SO}^{\underline{\varepsilon}}(N;k;\beta), \tag{4.5.9}$$

where $GT_{SO}^{\underline{\varepsilon}}(N;k;\beta)$ is the subset of $GT_{SO}(N;k;\beta)$ where the sign of $\lambda_i^{(2i-1)}$ for $1 \le i \le 2k\beta$ is required to be equal to ε_i. We decompose in this way due to the requirement of convexity in theorem 4.2.16. One then sees that, for instance,

$$GT_{SO}^{+}(N;k;\beta) = \bigcup_{\substack{\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}: \\ \varepsilon_{2k\beta}=1}} GT_{SO}^{\underline{\varepsilon}}(N;k;\beta). \tag{4.5.10}$$

Recall $GT_{SO}^{+}(N;k;\beta)$ was the subset of $GT_{SO}(N;k;\beta)$ with top row fixed to be $\langle N^{2k\beta} \rangle$, hence the sign of the odd starter for this top ($2k\beta$th) row is $+1$. Further examples of this definition are given by figure 4.28.

As in section 4.4, for ease we now concentrate on the 'unfrozen' elements. The following definition formally defines a relabelling of said parts, and figure 4.29 demonstrates the bijection between a given pattern $P \in GT_{SO}^{\underline{\varepsilon}}(N;k;\beta)$ and the renaming. In spirit, this process is the same as that described in definition 4.4.2, though with the added complexity of the signs of the odd starters.

Definition 4.5.3. *We define the set $\mathfrak{I}_{SO}(N;k;\beta)$ by the disjoint union*

$$\mathfrak{I}_{SO}(N;k;\beta) := \bigcup_{\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}} \mathfrak{I}_{SO}^{\underline{\varepsilon}}(N;k;\beta). \tag{4.5.11}$$

For a fixed $\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}$, $\mathfrak{I}_{SO}^{\underline{\varepsilon}}(N;k;\beta)$ is the set of integer arrays[2] $(y^{(i)})_{i=1}^{4k\beta-3} \in \mathbb{Z}^{k\beta(2k\beta-1)}$ *satisfying the following additional requirements:*

(i) $y^{(i)}, y^{(4k\beta-2-i)} \in \mathrm{S}^{+}_{\lfloor \frac{i+1}{2} \rfloor}$ *for* $1 \le i \le 2k\beta - 1$,

(ii) both $(y^{(i)})_{i=1}^{2k\beta-1}$ *and* $(y^{(4k\beta-2-i)})_{i=1}^{2k\beta-1}$ *form* $(2k\beta-1)$*–orthogonal Gelfand-Tsetlin patterns,*

(iii) $0 \le y_j^{(i)} \le N$ *for any valid* i, j,

(iv) the rows $(y^{(i)})_{i=1}^{4k\beta-3}$ *fulfil the following constraints:*

In the case k is even, let $i = 1, \ldots, \frac{k}{2}$ (with $y^{(-1)}, y^{(0)}, y^{(4k\beta-2)}, y^{(4k\beta-1)} \equiv 0$, and $\varepsilon_0 \equiv 1$). Then,

$$\begin{aligned}\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta} \varepsilon_j \varepsilon_{j-1} \left[\sum_{l=1}^{j} y_l^{(2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(2j-2)} + \sum_{l=1}^{j-1} y_l^{(2j-3)} \right] \\ = \sum_{j=(2i-1)\beta+1}^{2i\beta} \varepsilon_j \varepsilon_{j-1} \left[\sum_{l=1}^{j} y_l^{(2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(2j-2)} + \sum_{l=1}^{j-1} y_l^{(2j-3)} \right]\end{aligned} \tag{4.5.12}$$

[2]Note that since the height of orthogonal patterns is always odd, the top *two* rows in the full pattern are frozen, hence why the integer arrays in $\mathfrak{I}_{SO}^{\underline{\varepsilon}}(N;k,\beta)$ are $4k\beta - 3$ long.

and

$$\sum_{j=(2i-2)\beta+1}^{(2i-1)\beta} \varepsilon_{2k\beta-j+1}\varepsilon_{2k\beta-j}\left[\sum_{l=1}^{j} y_l^{(4k\beta-2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(4k\beta-2j)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+1)}\right] \tag{4.5.13}$$

$$= \sum_{j=(2i-1)\beta+1}^{2i\beta} \varepsilon_{2k\beta-j+1}\varepsilon_{2k\beta-j}\left[\sum_{l=1}^{j} y_l^{(4k\beta-2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(4k\beta-2j)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+1)}\right].$$

While, when k is odd we have the same constraints as above for $i = 1, \ldots, \frac{k-1}{2}$ along with:

$$\sum_{j=(k-1)\beta+1}^{k\beta} \varepsilon_j\varepsilon_{j-1}\left[\sum_{l=1}^{j} y_l^{(2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(2j-2)} + \sum_{l=1}^{j-1} y_l^{(2j-3)}\right] \tag{4.5.14}$$

$$= \sum_{j=(k-1)\beta+1}^{k\beta} \varepsilon_{2k\beta-j+1}\varepsilon_{2k\beta-j}\left[\sum_{l=1}^{j} y_l^{(4k\beta-2j-1)} - 2\sum_{l=1}^{j-1} y_l^{(4k\beta-2j)} + \sum_{l=1}^{j-1} y_l^{(4k\beta-2j+1)}\right].$$

Observe that, as in the symplectic case, for both k odd and even there are a total of k constraints.

As stated above, this is essentially the same relabelling as given by definition 4.4.2 in the symplectic case. Nevertheless, we emphasise that (i) and (iii) follow directly from proposition 4.5.1 and the definition of an orthogonal half-pattern. Since the upper right triangle is frozen, see figure 4.27, the 'free' triangle can essentially be partitioned in to two copies of a sub-orthogonal Gelfand-Tsetlin patterns, where one has been reflected in the x-axis and 'glued' on to the top row of the other. This again matches the symplectic case and is exactly the content of (ii). Turning to the constraints given in (iv) (i.e. (4.5.12)–(4.5.14)). These are precisely the same as (4.5.1), just translated in to this new co-ordinate system. Note that since we have 'decoupled' the sign of the odd starter from its absolute value, i.e.

$$\operatorname{sgn}(\lambda_j^{(2j-1)}) = \varepsilon_j \qquad \text{for } 1 \le j \le 2k\beta - 1 \tag{4.5.15}$$

$$|\lambda_j^{(2j-1)}| = y_j^{(2j-1)} \qquad \text{for } 1 \le j \le k\beta \tag{4.5.16}$$

$$|\lambda_j^{(2j-1)}| = y_{2k\beta-j}^{(2j-1)} \qquad \text{for } k\beta + 1 \le j \le 2k\beta - 1 \tag{4.5.17}$$

we therefore have (see also requirement (i)) that the $y_j^{(i)}$ are all positive and so we can remove the absolute values in the translation of (4.5.1).

Then, analogously to how $\mathfrak{B}_{Sp}$ was defined in section 4.4, (see (4.4.11)), one may also define

$$\mathfrak{B}_{SO} : GT_{SO}(N; k; \beta) \longrightarrow \mathfrak{I}_{SO}(N; k; \beta). \tag{4.5.18}$$

The bijection is depicted by figure 4.29, and can be constructed as follows. Take $P \in GT_{SO}(N; k; \beta)$ so $P = (\lambda^{(i)})_{i=1}^{4k\beta-1}$. In particular, there exists $\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}$ such that $P \in GT_{SO}^{\underline{\varepsilon}}(N; k; \beta)$. Due to the interlacing $\lambda^{(4k\beta-3)} \prec \langle N^{2k\beta-1}\rangle = \lambda^{(4k\beta-2)}$, all but one element of $\lambda^{(4k\beta-3)}$ is fixed:

$$\lambda_1^{(4k\beta-3)}, \ldots, \lambda_{2k\beta-2}^{(4k\beta-3)} \equiv N,$$
$$0 \le |\lambda_{2k\beta-1}^{(4k\beta-3)}| \le N.$$

We now set $y_1^{(4k\beta-3)} = |\lambda_{2k\beta-1}^{(4k\beta-3)}|$ and $\varepsilon_{2k\beta-1} = \operatorname{sgn}(\lambda_{2k\beta-1}^{(4k\beta-3)})$. Repeating the same logic, we consider the next pair of interlaced rows $\lambda^{(4k\beta-4)} \prec \lambda^{(4k\beta-3)}$ which once more fixes all but one coordinate:

$$\lambda_1^{(4k\beta-4)}, \dots, \lambda_{2k\beta-3}^{(4k\beta-4)} \equiv N,$$
$$y_1^{(4k\beta-3)} = |\lambda_{2k\beta-1}^{(4k\beta-3)}| \leq \lambda_{2k\beta-2}^{(4k\beta-4)} \leq N.$$

Thus set $y_1^{(4k\beta-4)} = \lambda_{2k\beta-2}^{(4k\beta-4)}$. This process can be repeated up to and including $\lambda^{(2k\beta)}$, after which there are no more coordinates fixed by the interlacing. Thereafter set $y_j^{(i)} = |\lambda_j^{(i)}|$, and throughout use the fact that $\varepsilon_j = \mathrm{sgn}(\lambda_j^{(2j-1)})$. It is apparent that this entire process is invertible, hence the map given by this construction, $\mathfrak{B}_{SO}$ is a bijection. We may then employ proposition 4.5.1 to achieve the following statement.

Proposition 4.5.4. *Let $k, \beta \in \mathbb{N}$. Then*

$$\begin{aligned}\mathrm{MoM}_{SO(2N)}(k,\beta) = \#GT_{SO}(N;k,\beta) &= \sum_{\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}} \#GT_{SO}^{\underline{\varepsilon}}(N;k;\beta) \\ &= \sum_{\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}} \#\mathfrak{I}_{SO}^{\underline{\varepsilon}}(N;k;\beta) \\ &= \#\mathfrak{I}_{SO}(N;k,\beta).\end{aligned}$$

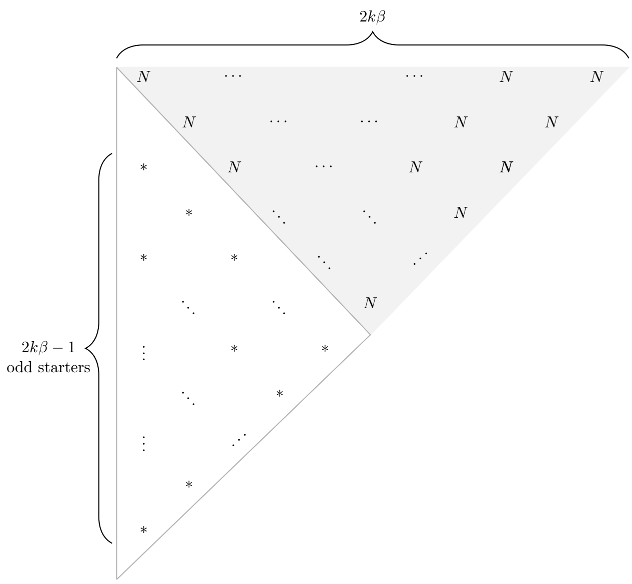


Figure 4.27: Figure depicting the fixed region of a $(4k\beta - 1)$–orthogonal Gelfand-Tsetlin pattern with top row $\langle N^{2k\beta} \rangle$. The shaded area represents the fixed region, whilst the unshaded region shows which elements have some freedom in the values that they can take.

4.5.2 Asymptotics and the leading order coefficient

Recall, from section 4.4.2, that we defined continuous half-patterns and continuous orthogonal Gelfand-Tsetlin patterns using the continuous Weyl chamber,

$$\mathrm{W}_N = \{x = (x_1, \dots, x_N) \in \mathbb{R}^N : x_1 \geq \cdots \geq x_N\}.$$

Once more, the purpose of this process is to define analogous *continuous* forms of the discrete objects considered in the previous section. In section 4.4.2, we defined the index set $\mathcal{S}_{(k,\beta)}^{Sp}$, here we

$\underline{\varepsilon} = (-1, 1, -1, -1)$, $P \in GT_{SO}^{\underline{\varepsilon}}(8; 2; 1)$

-8		8		8		8
	8		8		8	
-2		8		8		
	5		8			
0		6				
	2					
-1						

$\underline{\varepsilon} = (-1, -1, 1, 1)$, $Q \in GT_{SO}^{\underline{\varepsilon}}(5; 2; 1)$

5		5		5		5
	5		5		5	
2		5		5		
	3		5			
-3		4				
	2					
-3						

Figure 4.28: Examples of patterns P, Q in $GT_{SO}^{\underline{\varepsilon}}(N; k; \beta)$ and different, given values of N and $\underline{\varepsilon}$

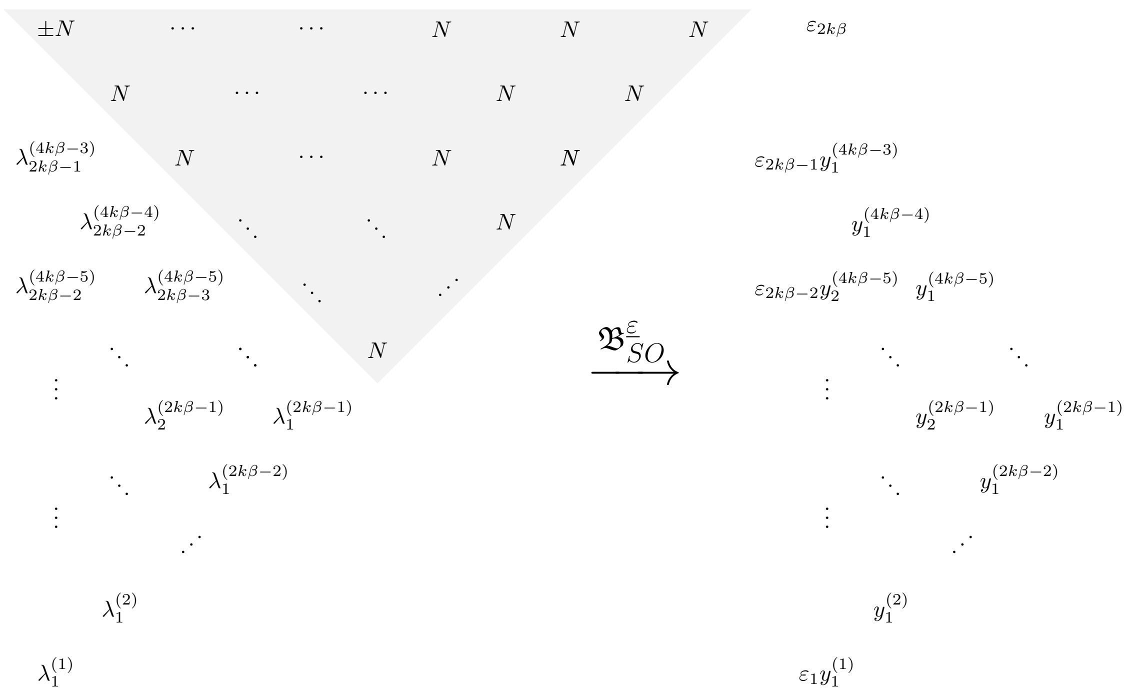


Figure 4.29: Pictorial representation of the relabelling of the coordinates in the Gelfand-Tsetlin pattern on the left, given by the bijection $\mathfrak{B}_{SO}^{\underline{\varepsilon}} : GT_{SO}^{\underline{\varepsilon}}(N; k; \beta) \longrightarrow \mathfrak{I}_{SO}^{\underline{\varepsilon}}(N; k; \beta)$. Above on the right hand side (the image of the bijection), $\varepsilon_j = \mathrm{sgn}(\lambda_j^{(2j-1)})$ for $j = 1, \ldots, 2k\beta - 1$ and the top sign is $\varepsilon_{2k\beta} = \mathrm{sgn}(\lambda_{2k\beta}^{(4k\beta-1)}) = \mathrm{sgn}(\pm N)$.

give the equivalent definition for the orthogonal case. For more explanation of the construction of this set, see the section 4.4.2.

$$\mathcal{S}^{SO}_{(k,\beta)} := \Big\{(m,n) : 1 \le m \le \left\lfloor \frac{n+1}{2} \right\rfloor \text{ and } 1 \le n \le 2k\beta - 1;$$
$$\text{or } 1 \le i \le \left\lfloor \frac{4k\beta - n - 1}{2} \right\rfloor \text{ and } 2k\beta \le n < 4k\beta - 3;$$
$$n \ne 4\beta - 1, 8\beta - 1, \ldots, 4(k-1)\beta - 1\Big\}$$
$$\cup \Big\{(m, 4n\beta - 1) : \ 1 \le m \le 2n\beta - 1 \text{ and } 1 \le n \le \left\lfloor \frac{k}{2} \right\rfloor;$$
$$\text{or } 1 \le m \le 2(k-n)\beta - 1 \text{ and } \left\lfloor \frac{k}{2} \right\rfloor + 1 \le n < k\}\Big\}.$$

Note that the size of the set $\mathcal{S}^{SO}_{(k,\beta)}$ is $k\beta(2k\beta-1) - k$. Again, it is not a coincidence that this is also the degree of the polynomial in the statement of theorem 4.1.2. The set corresponding to the indices 'missing' from $\mathcal{S}^{SO}_{(k,\beta)}$ is the following

$$\mathcal{T}^{SO}_{(k,\beta)} := \{(m, nj) : y^{(n)}_m \in \mathfrak{I}_{SO}(N; k; \beta)\} \backslash \mathcal{S}^{SO}_{(k,\beta)}.$$

Now define the following set $\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})} \subset \mathbb{R}^{k\beta(2k\beta-1)-k}$, which is the continuous version of $\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N; k; \beta)$, except that a particular choice of k of the coordinates from $\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N; k; \beta)$ are determined by the linear equations, (4.5.12)–(4.5.14). Then, $\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}$ comprises the following elements. Firstly, we take coordinates $x^{(n)}_m$ indexed by $(m,n) \in \mathcal{S}^{SO}_{(k,\beta)}$ which moreover satisfy the following:

$$0 \le x^{(n)}_m \le 1, \quad \text{for } (m,n) \in \mathcal{S}^{SO}_{(k,\beta)},$$

and take $\underline{\varepsilon}$ just as in the definition of $\mathcal{I}^{\underline{\varepsilon}}_{SO}(N; k, \beta)$, i.e. a fixed set of signs for the odd-starters. Additionally, $V^{SO}_{(k,\beta;\underline{\varepsilon})}$ contains the following k elements, determined by the linear equations (4.5.12)–(4.5.14) in the definition of $\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N; k; \beta)$,

$$\begin{aligned} x^{(n)}_{\lfloor \frac{n+1}{2} \rfloor} &\quad \text{for } n = 4\beta - 1, 8\beta - 1, \ldots, 4\lfloor \tfrac{k}{2} \rfloor \beta - 1, \\ x^{(n)}_{\lfloor \frac{4k\beta - n - 1}{2} \rfloor} &\quad \text{for } n = 4(\lfloor \tfrac{k}{2} \rfloor + 1)\beta - 1, \ldots, 4(k-1)\beta - 1, 4k\beta - 3. \end{aligned}$$

Thus,

(i) $0 \le x^{(n)}_m \le 1$, for all $x^{(n)}_m \in V^{SO}_{(k,\beta;\underline{\varepsilon})}$,

(ii) $x^{(n)}, x^{(4k\beta - n)} \in \mathrm{W}^+_{\lfloor \frac{n+1}{2} \rfloor}$, for all $n = 1, \ldots, 2k\beta - 1$,

(iii) both $(x^{(n)})^{2k\beta-1}_{n=1}$ and $(x^{(4k\beta-n)})^{2k\beta-1}_{n=1}$ form continuous $(2k\beta - 1)$–orthogonal Gelfand-Tsetlin patterns.

Observe that, just as in the symplectic case, $\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}$ is convex as an intersection of hyperplanes. Moreover, $\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}$ is contained in the cube $[0,1]^{k\beta(2k\beta-1)-k}$ and hence in a closed ball of radius $\sqrt{k\beta(2k\beta-1)-k}$.

Proof of theorem 4.1.2. The fact that the moments of moments are polynomials in N was proven in proposition 4.3.1, and the case of $k = \beta = 1$ was handled above in proposition 4.5.2.

What remains to be shown is the statement concerning the leading order for general k, β. Firstly note that for a given $\underline{\varepsilon} \in \{\pm 1\}^{2k\beta}$:

$$\#\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N;k;\beta) = \#\left(\mathbb{Z}^{k\beta(2k\beta-1)-k} \cap \left(N\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right)\right),$$

where for a set $\mathcal{A}$, we write $N\mathcal{A} = \{Nx : x \in \mathcal{A}\}$ for its dilate by a factor of N. Making use of theorem 4.2.16 with $\mathcal{S} = N\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}$ we get:

$$\begin{aligned}\#\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N;k;\beta) &= \mathrm{vol}\left(N\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right) + O_{k,\beta}\left(N^{k\beta(2k\beta-1)-k-1}\right) \\ &= N^{k\beta(2k\beta-1)-k}\,\mathrm{vol}\left(\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right) + O_{k,\beta}\left(N^{k\beta(2k\beta-1)-k-1}\right).\end{aligned}$$

Thus, by proposition 4.5.4 we obtain:

$$\begin{aligned}\mathrm{MoM}_{SO(2N)}(k,\beta) &= \sum_{\underline{\varepsilon}\in\{\pm1\}^{2k\beta}} \#\mathfrak{I}^{\underline{\varepsilon}}_{SO}(N;k;\beta) \\ &= \sum_{\underline{\varepsilon}\in\{\pm1\}^{2k\beta}} \left[N^{k\beta(2k\beta-1)-k}\,\mathrm{vol}\left(\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right) + O_{k,\beta}\left(N^{k\beta(2k\beta-1)-k-1}\right)\right] \\ &= \mathfrak{c}_{SO}(k,\beta)N^{k\beta(2k\beta-1)-k} + O_{k,\beta}\left(N^{k\beta(2k\beta-1)-k-1}\right)\end{aligned}$$

where

$$\mathfrak{c}_{SO}(k,\beta) = \sum_{\underline{\varepsilon}\in\{\pm1\}^{2k\beta}} \mathrm{vol}\left(\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right). \tag{4.5.19}$$

It then once more suffices to prove that $\mathfrak{c}_{SO}(k,\beta) > 0$, which is the content of lemma 4.5.5 below. □

It is again important that the constant $\mathfrak{c}_{SO}(k,\beta)$ is strictly positive, since this ensures that the claimed order of the polynomial is correct.

Lemma 4.5.5. *Let $k,\beta \in \mathbb{N}$. Then,*

$$\mathfrak{c}_{SO}(k,\beta) > 0. \tag{4.5.20}$$

Proof. Recall that

$$\mathfrak{c}_{SO}(k,\beta) = \sum_{\underline{\varepsilon}\in\{\pm1\}^{2k\beta}} \mathrm{vol}\left(\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right).$$

Thus, the proof of the strict positivity of the leading order coefficient $\mathfrak{c}_{SO}(k,\beta)$ can be deduced from showing that, for at least one choice of $\underline{\varepsilon} \in \{\pm1\}^{2k\beta}$, the volume $\mathrm{vol}\left(\mathcal{V}^{SO}_{(k,\beta;\underline{\varepsilon})}\right)$ is strictly positive. Henceforth, we choose $\underline{\varepsilon} = (1,1,\ldots,1)$. This choice of $\underline{\varepsilon}$ means that all the odd-starters are non-negative, essentially reducing the problem to the already-considered symplectic case. Thus, the argument is near identical to the one given in the symplectic case, see the proof of lemma 4.4.4, aside from trivial differences in the shapes considered. □

4.6 Outlook

The main results of this chapter were the asymptotic formulae for the moments of moments of symplectic and special orthogonal characteristic polynomials for integer moment parameters. We additionally showed that, as was the case with unitary characteristic polynomials, both $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ and $\mathrm{MoM}_{SO(2N)}(k,\beta)$ are polynomials in the matrix size in both cases. This is a result of the problem being integrable; this fact is evident from the connection to Gelfand-Tsetlin patterns used throughout.

Additionally of note is that the leading order behaviour differs across each of the compact groups. Recall that we argued in chapter 2, section 2.2.4 that the maximum around the circle of $\log|P_{G(N)}(A,\theta)|$ for $G(N) \in \{Sp(2N), SO(2N)\}$ should follow, to leading order, the maximum of $\log|P_{U(N)}(A,\theta)|$ as predicted by Fyodorov and Keating [82]. This highlights an interesting comparison between local and global behaviours.

Within this chapter, we also demonstrated how to use propositions 4.4.1 and 4.5.1 to calculate, for small k and β at least, the full moments of moments polynomials. This was done using the representation involving Gelfand-Tsetlin patterns. In chapter 5, we exploit a different connection between averages over the symplectic and orthogonal groups and Toeplitz $\pm$ Hankel determinants to compute (via Mathematica) more examples of these polynomials. This computation reveals further interesting structure which could be the subject of future study.

Finally, we emphasise that our results naturally lead to conjectural formulae for the moments of the moments of L-functions from orthogonal and symplectic families (see chapter 2, section 2.1.3). The first average is taken over a short section of the critical line around the symmetry point for the family, and the second average corresponds to averaging over the family. We direct the interested reader to our subsequent paper [16].

Chapter 5

Computing moments of moments

Computing the full moments of moments for integer moment parameters is an interesting task. The combinatorial (Young tableau or Gelfand-Tsetlin pattern) approach, see chapter 3 section 3.2 or chapter 4, often permits one to compute the first couple of moments by hand, but as soon as k, β increase (and in particular, the k parameter) this technique becomes immediately infeasible[1]. A similar story is true of the multiple contour integral technique used to prove the main theorem of chapter 3.

Further progress is possible using the 'ratios conjecture' of Conrey and Zirnbauer [48, 49] with the aid of computing software (see also the Ratios Theorems in chapter 7, section 7.1.4). Keating and Scott [114] produced initial code based on the ratios conjecture written in Maple, which outputs the first four novel[2] unitary moments of moments (see section 5.2 for the exact polynomials and further discussion of this approach). However, this technique also proves to be inefficient, especially as k increases.

Within this chapter we describe a third method for computing the moments of moments polynomials. We present the code used and discuss some future research directions implied by the structure of the polynomials.

5.1 Introduction

The most efficient technique for computing moments of moments that we have found to date[3] is to use sums of Toeplitz and Hankel determinants. In order to explain the code, we recap the relevant background here (though some of the introduction can also be found in chapter 3, section 3.1.1).

One defines a Toeplitz determinant in the following way. Take $f(z)$ integrable over the unit circle with Fourier coefficients

$$\hat{f}_j = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-ij\theta} d\theta, \tag{5.1.1}$$

for $j \in \mathbb{Z}$. The $N \times N$ *Toeplitz* matrix with symbol $f(z)$ is

$$T = (T_{i,j})_{i,j=1}^{N} = (\hat{f}_{i-j})_{i,j=0}^{N-1}. \tag{5.1.2}$$

The determinant of this matrix, denoted by $D_N(f) = \det(T)$, is the *Toeplitz determinant*. Notice that

[1] In chapter 4 we are easily able to hand-compute the full polynomial for $\mathrm{MoM}_{Sp(2N)}(1,1)$ and $\mathrm{MoM}_{SO(2N)}(1,1)$, see section 4.4.4 and proposition 4.5.2. However, it is clear that as soon as k increases, the combinatorics become much more involved.

[2] Recall that the Keating-Snaith result,theorem 2.1.1, gives a full description of $\mathrm{MoM}_{U(N)}(1,\beta)$.

[3] We thank Dr. Chris Hughes for useful conversations in this direction.

one may rewrite $D_N(f)$ in the following way (the Heine identity),

$$D_N(f) = D_N^T(f) = \frac{1}{(2\pi)^N N!}\int_0^{2\pi}\cdots\int_0^{2\pi}\prod_{j=1}^N f(e^{i\theta_j})\prod_{1\le j<k\le N}|e^{i\theta_j}-e^{i\theta_k}|^2 d\theta_1\cdots d\theta_k. \tag{5.1.3}$$

Comparing (5.1.3) to the Weyl integration formula, (1.1.4), makes the connection to averages of characteristic polynomials explicit. Thus, with a judicious choice of symbol f, one may rewrite the moments of moments of unitary matrices as Toeplitz determinants. This is exactly the technique employed by Claeys and Krasovsky [36] and Fahs [72], see section 3.1.1.

A result of Baik and Rains [13] shows that the same is true for the symplectic and orthogonal cases as well, except that the determinants are now of 'Toeplitz $\pm$ Hankel'–type. Toeplitz $\pm$ Hankel determinants $D_N^{T\pm H}(f)$ are similar to Toeplitz determinants, except they also involve a 'Hankel'–type symbol:

$$D_N^{T\pm H}(f) = \det(f_{i-j} \pm f_{i+j+1})_{i,j=0}^{N-1}. \tag{5.1.4}$$

The exact choices of f and the precise Toeplitz, and Toeplitz $\pm$ Hankel determinants considered, as well as their relationship to the relevant moments of moments are given within section 5.2 and 5.3.

The majority of section 5.2 appeared as appendix 6.1 of [15]. Some additions have been made for additional clarity, and the code used to produce the polynomials has been included. Similarly, the polynomials given in section 5.3 appeared in [9], but the section has been supplemented with the code and various clarifying comments. Section 5.4 focusses on the properties of the polynomials presented in sections 5.2 and 5.3, and plots of the roots of such polynomials are presented since they reveal interesting structure. Such plots are original to this thesis.

5.2 Unitary moments of moments

Here we give explicit examples of the full polynomials $\mathrm{MoM}_{U(N)}(k,\beta)$ for small values of k, β. The formulae we record extend the results of preliminary calculations due to Keating and Scott [114] (c.f. [111]), which formed the basis for some of the numerical computations in [80]. We should remark that the moment formula of Keating and Snaith [116] gives the full polynomials for the case $k = 1$, $\beta \in \mathbb{N}$; see (3.1.10).

The technique employed by Keating and Scott is first presented and discussed. We then give an alternative, and anecdotally more computationally efficient, method for determining the polynomials. Beyond calculation speed and algorithmic simplicity, this second method also easily generalizes to the symplectic and orthogonal case using the results of Baik and Rains [13], see section 5.3.

The approach of Keating and Scott uses the 'Ratios Theorem' first derived by Conrey, Farmer and Zirnbauer [48, 49], and later rederived by many authors including Bump and Gamburd, who used symmetric function theory [33]. Note that we used a special case of this result to prove theorem 3.1.5 in chapter 3, section 3.4, as well as the polynomial structure of the symplectic and orthogonal moments of moments chapter 4, section 4.3. We also make liberal use of the Ratios Theorem in chapter 7. The statement of the theorem can get notationally complex, so here we use the presentation of Conrey and Snaith [55].

First, define for finite sets A, B, C, D,

$$R(A,B;C,D) := \int_{U(N)} \frac{\prod_{\alpha\in A}\det(I - X^* e^{-\alpha})\prod_{\beta\in B}\det(I - Xe^{-\beta})}{\prod_{\gamma\in C}\det(I - X^* e^{-\gamma})\prod_{\delta\in D}\det(I - Xe^{-\delta})} dX. \tag{5.2.1}$$

Further if

$$Z(A,B) := \prod_{\substack{\alpha\in A,\\ \beta\in B}} \frac{1}{(1-e^{-(\alpha+\beta)})}, \tag{5.2.2}$$

then define

$$Z(A,B;C,D) := \frac{Z(A,B)Z(C,D)}{Z(A,D)Z(B,C)}. \tag{5.2.3}$$

Finally, if $S \subset A$ and $T \subset B$ then $\overline{S} = A - S$, $\overline{T} = B - T$, $S^- = \{-\hat{\alpha} : \hat{\alpha} \in S\}$ and similarly for T. Note that here we are using the notation $U + V$, $U - V$ (to be interpreted as $U \cup V$ and $U \backslash V$ respectively for sets U, V) to be consistent with the statement of the theorem in [55].

Theorem 5.2.1 (Ratios Theorem [48, 49])**.** *With $N \geq 0$ and $\mathrm{Re}(\gamma) > 0, \mathrm{Re}(\delta) > 0$ for $\gamma \in C$, $\delta \in D$, $|C| \leq |A| + N$, $|D| \leq |B| + N$, we have*

$$R(A,B;C,D) = \sum_{\substack{S\subset A, T\subset B\\ |S|=|T|}} e^{-N(\sum_{\hat{\alpha}\in S}\hat{\alpha}+\sum_{\hat{\beta}\in T}\hat{\beta})} Z(\overline{S}+T^-, \overline{T}+S^-;C,D),$$

where $A = S + \overline{S}$ and $B = T + \overline{T}$.

To see how this can be used to give the full polynomials for $\mathrm{MoM}_{U(N)}(k,\beta)$, we outline the simplest case with $k = \beta = 1$.

$$\begin{aligned} \mathrm{MoM}_{U(N)}(1,1) &= \frac{1}{2\pi}\int_0^{2\pi} \mathbb{E}x_{A\in U(N)}\left[|P_N(A,\theta)|^2\right] d\theta & (5.2.4)\\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_{U(N)} P_N(A,\theta)P_N(A^*,-\theta)dAd\theta. & (5.2.5)\end{aligned}$$

Thus, we apply the Ratios Theorem with $A = \{i\alpha\}, B = \{i\beta\}$, $C, D = \varnothing$ to find

$$\begin{aligned} \mathrm{MoM}_{U(N)}(1,1) &= \frac{1}{2\pi}\int_0^{2\pi} \lim_{\beta\to-\alpha} Z(A,B) + e^{-iN(\alpha+\beta)}Z(B^-,A^-)d\alpha & (5.2.6)\\ &= \frac{1}{2\pi}\int_0^{2\pi} \lim_{\beta\to-\alpha}\sum_{m=0}^{N} e^{-im(\alpha+\beta)}d\alpha & (5.2.7)\\ &= N+1. & (5.2.8)\end{aligned}$$

Since Keating and Snaith computed a closed form for $\mathrm{MoM}_{U(N)}(1,\beta)$, we can compare (5.2.8) to their formula. For $\beta \in \mathbb{N}$, they show (theorem 2.1.1) that

$$\mathrm{MoM}_{U(N)}(1,\beta) = \prod_{0\leq i,j\leq \beta-1}\left(\frac{N}{i+j+1}+1\right), \tag{5.2.9}$$

which for $\beta = 1$ gives

$$\mathrm{MoM}_{U(N)}(1,1) = N+1, \tag{5.2.10}$$

corroborating (5.2.8).

Higher values of k, β clearly necessitate bigger sets A, B in the application of the Ratios Theorem, and hence many more choices for S, T. This is a limitation of this approach (though to some extent, all the techniques that we describe suffer from the same constraint). Nevertheless, Keating and Scott were able to compute $\mathrm{MoM}_{U(N)}(k,\beta)$ completely for $\beta = 1$, $k \in \{2,3,4\}$, and $k = \beta = 2$. Their polynomials have been recomputed using an alternative approach that we describe below, and are given at the end of this section.

To find (even just slightly) higher moments of moments within a reasonable time-frame, we instead moved to using Toeplitz determinants. The connection between unitary averages and Toeplitz determinants was outlined at the beginning of this chapter. Hence, all that remains is to fix a choice of Toeplitz symbol f. Using (5.1.3) as a guide, we choose,

$$f(z) = f_{\theta_1,\ldots,\theta_k}(z) := \prod_{j=1}^{k}(1 - ze^{-i\theta_j})^{\beta}(1 - z^{-1}e^{i\theta_j})^{\beta}, \tag{5.2.11}$$

where we emphasise that the $e^{i\theta_1}, \ldots, e^{i\theta_k}$ are not related to the eigenvalues of the unitary matrix, but represent the points at which the characteristic polynomial is to be evaluated. We proved in chapter 3 that $\mathrm{MoM}_{U(N)}(k,\beta)$ is a polynomial in N of degree $k^2\beta^2 - k + 1$ for integer k, β. This is the other ingredient essential to the success of this method: since $\mathrm{MoM}_{U(N)}(k,\beta)$ is a polynomial in N, if one has $k^2\beta^2 - k + 2$ evaluations then one can reconstruct the full polynomial. Therefore, we calculate the exact value of

$$\frac{1}{2\pi}\int_0^{2\pi}\cdots\int_0^{2\pi} D_N(f_{\theta_1,\ldots,\theta_k})d\theta_1\cdots d\theta_k, \tag{5.2.12}$$

for $N \in \{1, \ldots, k^2\beta^2 - k + 2\}$. Finally, armed with the $k^2\beta^2 - k + 2$ evaluations, all that remains is to solve a similar linear algebra problem to reconstruct the polynomial.

The code written to do the calculation described above is given here, written in Mathematica.

```
mom_unitary[k_,beta_]:=Module[{t,z},
  (* Given integers k and beta, this function outputs the polynomial in n for the unitary
      moments of moments with these parameters
     Local parameters t1,...,tk represent e^(i theta_1),...,e^(i theta_k)
     Local parameter z represents an eigenvalue *)
  deg = k^2 beta^2 - k + 1;
  vars = Array[t,k];
  f[z_] := Product[(1 - z/vars[[i]])^beta (1 - vars[[i]]/z)^beta, {i, 1, k}];
  poly_evals = Table[Null, {deg + 1}];
  For[m = 1, m <= deg + 1, m++,
    (* Compute the m x m Toeplitz determinant *)
    (* int_theta is the resulting function, and depends on t (theta) *)
    int_theta = Det[Table[Coefficient[f[z], z, i - j], {i, 1, m}, {j, 1, m}]];
    For[i = 1, i <= k, i++,
      (* This is the only term which will survive the integration over the t (theta)
          variables *)
      int_theta = Coefficient[int_theta, vars[[i]], 0]
    ];
    (* This vector stores the evaluations of the polynomial at each n *)
    poly_evals[[m]] = int_theta;
  ];
  (* Using linear algebra, we recover the coefficients of the polynomial *)
  coeffs = Inverse[Table[i^(j-1), {i, 1, deg + 1}, {j, 1, deg + 1}]].poly_evals;
  Return[Factor[Sum[coeffs[[i]]*n^(i - 1), {i, 1, deg + 1}]]]
]
```

The penultimate line of the code performs following linear algebra. Given the first $n+1$ integral evaluations of a polynomial

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \tag{5.2.13}$$

of degree n we can construct the following matrix equation,

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 4 & \cdots & 2^n \\ 1 & 3 & 9 & \cdots & 3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n+1 & (n+1)^2 & \cdots & (n+1)^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} p(1) \\ p(2) \\ p(3) \\ \vdots \\ p(n+1) \end{pmatrix}. \tag{5.2.14}$$

To recover the coefficient a_j of x^j in $p(x)$, we just apply the inverse of the leftmost $(n+1) \times (n+1)$ matrix to both sides.

5.2.1 Examples of unitary polynomials

The code above produces the following polynomials. The first six match those computed by Keating and Scott [114] using the Ratios Theorem method, but the run-time for our method is much quicker. $\mathrm{MoM}_{U(N)}(2,3)$ is novel to this thesis.

$$\begin{aligned}
\mathrm{MoM}_{U(N)}(1,1) &= N+1 \\
\mathrm{MoM}_{U(N)}(2,1) &= \frac{1}{6}(N+3)(N+2)(N+1) \\
\mathrm{MoM}_{U(N)}(3,1) &= \frac{1}{2520}(N+5)(N+4)(N+3)(N+2)(N+1)(N^2+6N+21) \\
\mathrm{MoM}_{U(N)}(4,1) &= \frac{1}{778377600}(N+7)(N+6)(N+5)(N+4)(N+3)(N+2)(N+1) \\
&\quad \times (7N^6+168N^5+1804N^4+10944N^3+41893N^2+99624N+154440) \\
\mathrm{MoM}_{U(N)}(1,2) &= \frac{1}{12}(N+1)(N+2)^2(N+3) \\
\mathrm{MoM}_{U(N)}(2,2) &= \frac{1}{163459296000}(N+7)(N+6)(N+5)(N+4)(N+3)(N+2)(N+1) \\
&\quad \times (298N^8+9536N^7+134071N^6+1081640N^5+5494237N^4+18102224N^3 \\
&\qquad +38466354N^2+50225040N+32432400) \\
\mathrm{MoM}_{U(N)}(2,3) &= \frac{1}{172219132773102415494444188958720000000}(N+1)(N+2)(N+3)(N+4) \\
&\quad \times (N+5)(N+6)(N+7)(N+8)(N+9)(N+10)(N+11) \\
&\quad \times \big(12308743625763N^{24}+1772459082109872N^{23}+121902830804059138N^{22} \\
&\qquad +5328802119564663432N^{21}+166214570195622478453N^{20} \\
&\qquad +3937056259812505643352N^{19}+73583663800226157619008N^{18} \\
&\qquad +1113109355823972261429312N^{17}+13869840005250869763713293N^{16} \\
&\qquad +144126954435929329947378912N^{15}+1259786144898207172443272698N^{14} \\
&\qquad +9315726913410827893883025672N^{13}+58475127984013141340467825323N^{12} \\
&\qquad +311978271286536355427593012632N^{11}+1413794106539529439589778645028N^{10} \\
&\qquad +5427439874579682729570383266992N^{9}+17564370687865211818995713096848N^{8} \\
&\qquad +47561382824003032731805262975232N^{7}+106610927256886475209611301000128N^{6} \\
&\qquad +194861499503272627170466392014592N^{5}+284303877221735683573377603640320N^{4} \\
&\qquad +320989495108428049992898521600000N^{3}+266974288159876385845370793984000N^{2} \\
&\qquad +148918006780282798012340305920000N+43144523802785397500411904000000\big).
\end{aligned}$$

5.3 Symplectic and orthogonal moments of moments

We now give various explicit examples of the polynomials $\mathrm{MoM}_{G(N)}(k,\beta)$ for $G(N) \in \{Sp(2N), SO(2N)\}$ and small, integer values of k, β. These polynomials first appeared[4] in [9] and the computations are the work of the present author.

One could, as commented in section 5.2, use a version of the Ratios Theorem adapted to the symplectic or orthogonal case in order to compute these polynomials.

These examples were calculated using expressions for averages over $Sp(2N), SO(2N)$ involving sums of Toeplitz and Hankel determinants. Such expressions were first derived by Baik and Rains [13].

Theorem 5.3.1 (Baik and Rains [13]). *Let f be a function on the unit circle such that*

$$f_j = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})f(e^{-i\theta})e^{ij\theta}d\theta \tag{5.3.1}$$

is well defined. For $X \in \{Sp(2N), SO(2N)\}$ with eigenvalues[5] $e^{i\phi_1}, e^{-i\phi_1}, \dots, e^{i\phi_N}, e^{-i\phi_N}$ set

$$f(X) := \prod_{j=1}^{N} f(e^{i\phi_j})f(e^{-i\phi_j}). \tag{5.3.2}$$

Then

$$\int_{Sp(2N)} f(X)dX = \det(f_{j-k} - f_{j+k})_{j,k=1}^{N} \tag{5.3.3}$$

$$\int_{SO(2N)} f(X)dX = \frac{1}{2}\det(f_{j-k} + f_{j+k})_{j,k=0}^{N-1}, \tag{5.3.4}$$

except for

$$\int_{SO(0)} f(X)dX = 2. \tag{5.3.5}$$

Thus, in both cases, the matrix averages are equivalent to certain Toeplitz ± Hankel determinants.

5.3.1 The symplectic case

From theorem 5.3.1, it is clear once more that the goal is simply to choose the correct symbol f. Using the symmetry of the eigenvalues of both symplectic and even special orthogonal matrices, the correct choice evidently is

$$f(z) = \prod_{j=1}^{k}(1 - ze^{i\theta_j})^\beta(1 - ze^{-i\theta_j})^\beta(1 - z^{-1}e^{i\theta_j})^\beta(1 - z^{-1}e^{-i\theta_j})^\beta. \tag{5.3.6}$$

Within the construction of both the symplectic and orthogonal moments of moments polynomials, we again crucially use that $\mathrm{MoM}_{G(N)}(k,\beta)$ is a polynomial in N (after theorem 4.1.1 and theorem 4.1.2). Hence, by computing enough evaluations using the Toeplitz ± Hankel determinantal presentations, one can reconstruct the polynomial. As usual, for small k, β this is a computationally feasible task, but the complexity grows swiftly with k, β. The symplectic code is as follows.

[4] With the exception of $\mathrm{MoM}_{Sp(2N)}(2,2)$ and $\mathrm{MoM}_{SO(2N)}(2,2)$, which are original to this chapter.

[5] We here deviate from the traditional notation of $e^{i\theta_j}$ for eigenvalues, since we use $e^{i\theta_j}$ for the points at which we evaluate the characteristic polynomial.

```
mom_sympl[k_,beta_]:=Module[{t,z},
  (* Given integers k and beta, this function outputs the polynomial in n for the
      symplectic moments of moments with these parameters.
     Local parameters t1,...,tk represent e^(i theta_1),...,e^(i theta_k)
     Local parameter z represents an eigenvalue *)
  deg = k beta (2 k beta + 1) - k;
  vars = Array[t,k];
  f[z_] := Product[(1 - z vars[[i]])^beta (1 - z / vars[[i]])^beta
          (1 - vars[[i]] / z)^beta (1 - 1 / (z vars[[i]]))^beta, {i, 1, k}];
  poly_evals = Table[Null, {deg + 1}];
  For[m = 1, m <= deg + 1, m++,
    (* Compute the m x m Toeplitz - Hankel determinant *)
    (* int_theta is the resulting function, and depends on t (theta) *)
    int_theta = Det[Table[Coefficient[f[z], z, i - j]-Coefficient[f[z], z, i + j],
              {i, 1, m}, {j, 1, m}]];
    For[i = 1, i <= k, i++,
      (* This is the only term which will survive the integration over the t (theta)
          variables *)
      int_theta = Coefficient[int_theta, vars[[i]], 0]
    ];
    (* This vector stores the evaluations of the polynomial at each n *)
    poly_evals[[m]] = int_theta;
  ];
  (* Using linear algebra, we recover the coefficients of the polynomial *)
  coeffs = Inverse[Table[i^(j-1), {i, 1, deg + 1}, {j, 1, deg + 1}]].poly_evals;
  Return[Factor[Sum[coeffs[[i]]*n^(i - 1), {i, 1, deg + 1}]]]
]
```

5.3.2 Examples of symplectic polynomials

$$
\begin{aligned}
\mathrm{MoM}_{Sp(2N)}(1,1) &= \frac{1}{2}(N+1)(N+2) \\
\mathrm{MoM}_{Sp(2N)}(1,2) &= \frac{1}{181440}(N+1)(N+2)(N+3)(N+4)(2N+5) \\
&\quad \times \left(23N^4+230N^3+905N^2+1650N+1512\right) \\
\mathrm{MoM}_{Sp(2N)}(1,3) &= \frac{1}{405483668029440000}(N+1)(N+2)(N+3)(N+4)(N+5)(N+6) \\
&\quad \times \big(10253349N^{14}+502414101N^{13}+11401640999N^{12}+158831139621N^{11} \\
&\qquad +1517607151837N^{10}+10524657547803N^{9}+54662663279397N^{8} \\
&\qquad +216189375784263N^{7}+655178814761674N^{6}+1517469287314596N^{5} \\
&\qquad +2654161159219304N^{4}+3424171976788416N^{3}+3125457664755840N^{2} \\
&\qquad +1856618315596800N+563171761152000) \\
\mathrm{MoM}_{Sp(2N)}(2,1) &= \frac{1}{10080}(N+1)(N+2)(N+3)(N+4)\left(3N^4+30N^3+127N^2+260N+420\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{MoM}_{Sp(2N)}(2,2) = {} & \frac{1}{922602496998762940148808155136000000} \\
& \times (N+1)(N+2)(N+3)(N+4)(N+5)(N+6)(N+7)(N+8) \\
& \times (973768123863070N^{26} + 113930870491979190N^{25} + 6397115719464033835N^{24} \\
& \quad + 229468472209599934680N^{23} + 5905547966177597457148N^{22} \\
& \quad + 116095701233631059348562N^{21} + 1812360160544789853842185N^{20} \\
& \quad + 23058330120078491422398060N^{19} + 243489625307024066610062698N^{18} \\
& \quad + 2162212652001456929946764682N^{17} + 16300969840895475542904389485N^{16} \\
& \quad + 105047622582427150960862391840N^{15} + 581366239522098824777561920768N^{14} \\
& \quad + 2771149190931697097022336838782N^{13} + 11391521487556582357601041085935N^{12} \\
& \quad + 40376504846279028038258166555420N^{11} + 123185513678594029275619833254428N^{10} \\
& \quad + 322428740278217975133636090246192N^{9} + 720306825342284199391162987439760N^{8} \\
& \quad + 1363422036976648848604657681704000N^{7} + 2164768121293367194339285639341888N^{6} \\
& \quad + 2844421032168416266996544628302592N^{5} + 3038012656193009832745998965068800N^{4} \\
& \quad + 2576236167351129662498194684416000N^{3} + 1680524836539415222816615034880000N^{2} \\
& \quad + 798336341064739908062588436480000N + 228820063739772554600398848000000) \\
\mathrm{MoM}_{Sp(2N)}(3,1) = {} & \frac{1}{133382785536000}(N+1)(N+2)(N+3)(N+4)(N+5)(N+6) \\
& \times (5810N^{12} + 244020N^{11} + 4746259N^{10} + 56513415N^{9} \\
& \quad + 459233580N^{8} + 2688408450N^{7} + 11665223647N^{6} + 38004428175N^{5} \\
& \quad + 93222284960N^{4} + 171600705780N^{3} + 236485094544N^{2} \\
& \quad + 239758263360N + 185253868800)
\end{aligned}
$$

5.3.3 The orthogonal case

Again, using theorem 5.3.1, we choose the symbol

$$
f(z) = \prod_{j=1}^{k} (1 - ze^{i\theta_j})^{\beta} (1 - ze^{-i\theta_j})^{\beta} (1 - z^{-1}e^{i\theta_j})^{\beta} (1 - z^{-1}e^{-i\theta_j})^{\beta}, \tag{5.3.7}
$$

and compute $k\beta(2k\beta - 1) - k + 1$ evaluations using the Toeplitz + Hankel determinantal presentation to reconstruct $\mathrm{MoM}_{SO(2N)}(k, \beta)$ (though note that special care needs to be taken in the case $k = \beta = 1$ where the degree is instead 1).

The orthogonal code is identical to the symplectic code given in section 5.3.1, except for the required adjustments to the definition of degree and the precise Toeplitz + Hankel determinant to be calculated.

```
mom_ortho[k_,beta_]:=Module[{t,z},
  (* Given integers k and beta, this function outputs the polynomial in n for the
      orthogonal moments of moments with these parameters.
     Local parameters t1,...,tk represent e^(i theta_1),...,e^(i theta_k)
     Local parameter z represents an eigenvalue *)
  deg = If[k == 1 && beta == 1, 1, k beta (2 k beta - 1) - k];
  vars = Array[t,k];
  f[z_] := Product[(1 - z vars[[i]])^beta (1 - z / vars[[i]])^beta
          (1 - vars[[i]] / z)^beta (1 - 1 / (z vars[[i]]))^beta, {i, 1, k}];
  poly_evals = Table[Null, {deg + 1}];
  For[m = 1, m <= deg + 1, m++,
    (* Compute the m x m Toeplitz - Hankel determinant *)
    (* int_theta is the resulting function, and depends on t (theta) *)
    int_theta = (1/2)*Det[Table[Coefficient[f[z], z, i - j]+Coefficient[f[z], z, i + j],
              {i, 0, n-1}, {j, 0, n-1}]];
    For[i = 1, i <= k, i++,
      (* This is the only term which will survive the integration over the t (theta)
          variables *)
      int_theta = Coefficient[int_theta, vars[[i]], 0]
    ];
    (* This vector stores the evaluations of the polynomial at each n *)
    poly_evals[[m]] = int_theta;
  ];
  (* Using linear algebra, we recover the coefficients of the polynomial *)
  coeffs = Inverse[Table[i^(j-1), {i, 1, deg + 1}, {j, 1, deg + 1}]].poly_evals;
  Return[Factor[Sum[coeffs[[i]]*n^(i - 1), {i, 1, deg + 1}]]]
]
```

5.3.4 Examples of orthogonal polynomials

$$
\begin{aligned}
\mathrm{MoM}_{SO(2N)}(1,1) &= 2(N+1)\\
\mathrm{MoM}_{SO(2N)}(1,2) &= \frac{1}{60}(N+1)(N+2)(2N+3)\left(13N^2+39N+20\right)\\
\mathrm{MoM}_{SO(2N)}(1,3) &= \frac{1}{43589145600}(N+1)(N+2)(N+3)(N+4)\\
&\quad\times\big(677127N^{10}+16928175N^9+188303800N^8+1226849750N^7+5186281891N^6\\
&\qquad+14881334615N^5+29392642150N^4+39443286500N^3\\
&\qquad+34230199032N^2+17098220160N+3632428800\big)\\
\mathrm{MoM}_{SO(2N)}(2,1) &= \frac{1}{2}(N+1)^2(N+2)^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{MoM}_{SO(2N)}(2,2) = &\frac{1}{1260285816020642611200000 0}(N+1)(N+2)(N+3)(N+4)(N+5)(N+6)\\
&\times (609707861586N^{20} + 42679550311020N^{19} + 1412430303852455N^{18}\\
&\quad + 29381117133365235N^{17} + 430840519326988896N^{16}\\
&\quad + 4734030811194989220N^{15} + 40443864319087732710N^{14}\\
&\quad + 275114191393973346870N^{13} + 1513575162103740726146N^{12}\\
&\quad + 6802724518095393107220N^{11} + 25122677297620976396115N^{10}\\
&\quad + 76408091504568456233055N^{9} + 191200438955492015678676N^{8}\\
&\quad + 391999359061097734388820N^{7} + 653303373477066454323320N^{6}\\
&\quad + 874117748056919711331840N^{5} + 921264972736828061409696N^{4}\\
&\quad + 742476352946061736638720N^{3} + 435941253732258780710400N^{2}\\
&\quad + 170547181909185489408000N + 35007939333906739200000)\\
\mathrm{MoM}_{SO(2N)}(3,1) = &\frac{1}{1360800}(N+1)(N+2)^2(N+3)^2(N+4)\\
&\times (N^2+5N+9)(31N^4+310N^3+1163N^2+1940N+2100)
\end{aligned}
$$

5.4 Structure of the moments of moments polynomials

We here record some interesting observations about the examples of the moments of moments polynomials recorded in sections 5.2.1, 5.3.2, and 5.3.4. Firstly, the structure of the leading order coefficient in all cases, across each matrix type, is initially striking. This is especially true when one compares them to the general expression for the leading order coefficient that we recover in, for example, theorem 3.1.4. Additionally, one can observe that the polynomials all evaluate to integers at positive, integer N. That this should be the case is not necessarily obvious from the perspective of the multiple contour integral form, see chapter 3, section 3.3. However, once one recalls that the moments of moments in each case are counting various Young tableaux (or equivalently Gelfand-Tsetlin patterns), this phenomenon is therefore explained as a result of this combinatorial representation.

A peculiarity that is not so easily addressed is the highly structured nature of the polynomials. Notice that for $G(N) \in \{U(N), Sp(2N), SO(2N)\}$, the moments of moments factor in the following way,

$$
\mathrm{MoM}_{G(N)}(k,\beta) = c_{k,\beta} f_{g(k,\beta)}(N) \prod_{j=1}^{g(k,\beta)} (N+j), \tag{5.4.1}
$$

where $c_{k,\beta}$ is the relevant leading order coefficient, $f_{g(k,\beta)}(N)$ is a polynomial of degree $d - g(k,\beta)$ in N which may not factor in to linear terms over $\mathbb{Q}$, d is the degree of the full polynomial, and

$$
g(k,\beta) = g(k,\beta;G(N)) \coloneqq \begin{cases} 2k\beta & \text{if } G(N) = Sp(2N) \\ 2k\beta - 1 & \text{if } G(N) = U(N) \\ 2k\beta - 2 & \text{if } G(N) = SO(2N), \end{cases} \tag{5.4.2}
$$

apart from the usual exception of $k = \beta = 1$ for $SO(2N)$.

Recall that for $k = 1$ and $G(N) = U(N)$, Keating and Snaith calculated the moments of moments for finite N, and integer β

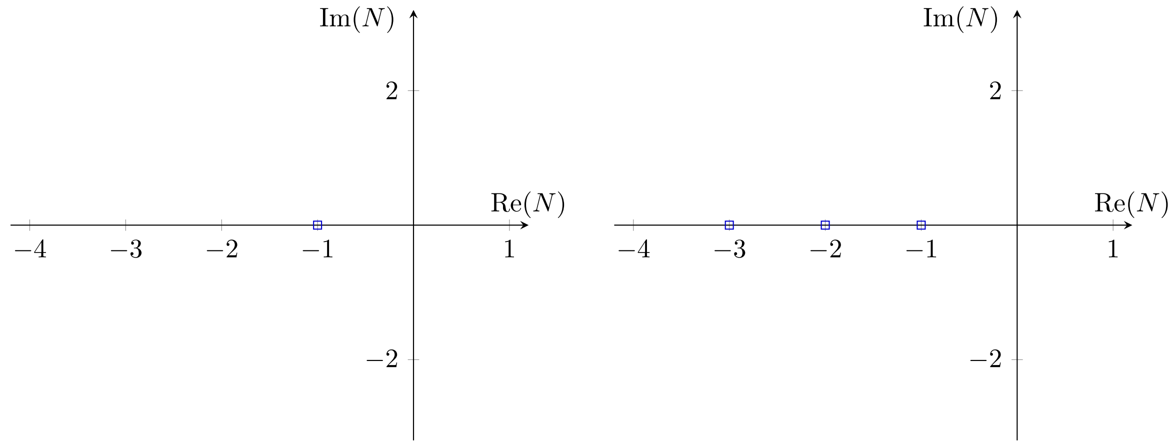
$$\mathrm{MoM}_{U(N)}(1,\beta) = \prod_{0\leq i,j\leq \beta-1}\left(\frac{N}{i+j+1}+1\right), \tag{5.4.3}$$

see theorem 2.1.1. The fact that $\mathrm{MoM}_{U(N)}(1,\beta)$ has such a succinct closed form, alongside the additional examples given for more general moments of moments, leads one to postulate that there might exist a similarly concise closed form for $\mathrm{MoM}_{G(N)}(k,\beta)$.

5.4.1 Roots of the moments of moments polynomials

In order to further investigate the curiously structured factorisation of the various moments of moments polynomials, we have calculated and plotted the roots of the examples given in sections 5.2.1, 5.3.2 and 5.3.4. The charts reveal further symmetries amongst the roots: the symmetry across $\mathrm{Re}(N) = 0$ is to be expected, but the vertical symmetries are yet to be explained.

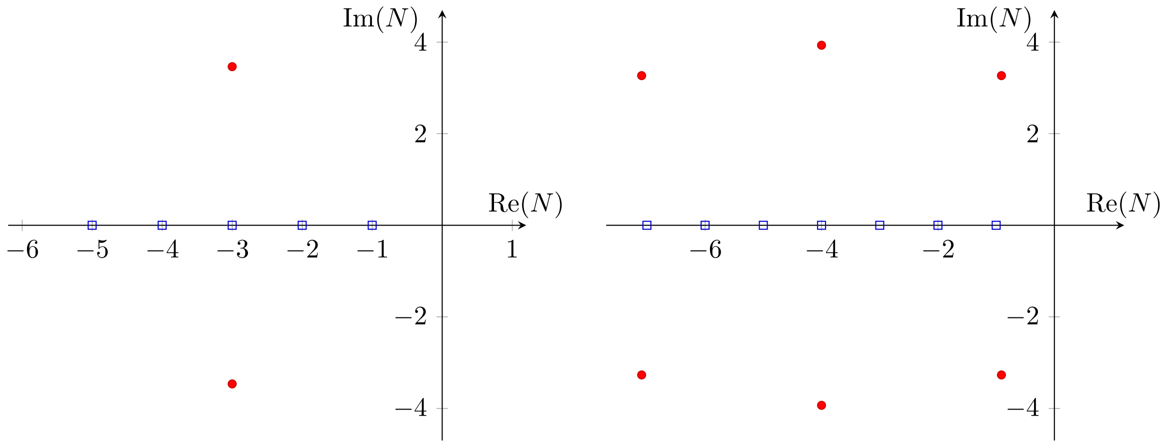
In each of the subsequent plots, the roots associated with the product term in (5.4.1) (i.e. those at $N = -1, -2, \ldots, -g(k,\beta;G(N)))$ are plotted as (blue) squares. Those corresponding to the roots of the remaining polynomial $f_{g(k,\beta)}(N)$ are plotted as (red) circles. In some situations, for example $\mathrm{MoM}_{U(N)}(1,2)$, there are double roots; these are made explicit for each plot within the accompanying caption.



(a) Roots of $\mathrm{MoM}_{U(N)}(1,1)$.

(b) Roots of $\mathrm{MoM}_{U(N)}(2,1)$.

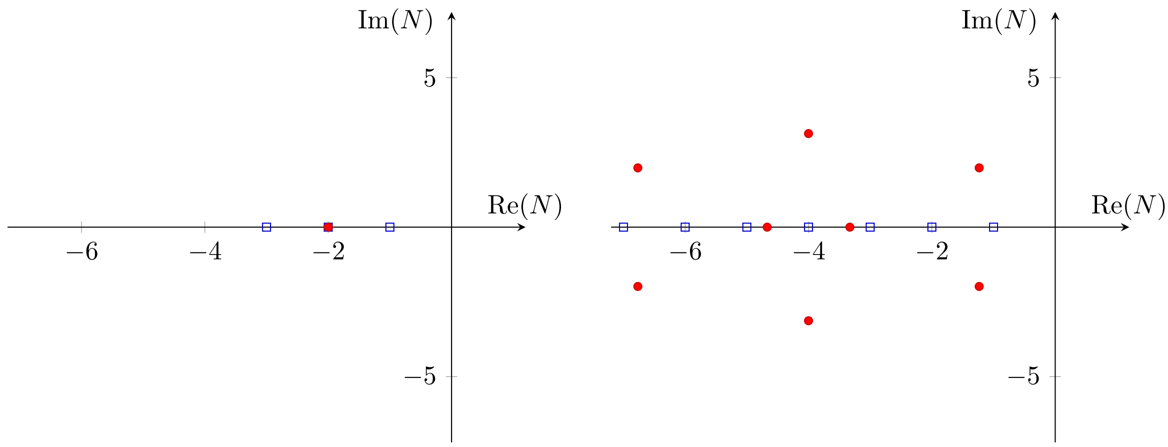
Figure 5.1: Plots of the roots of the unitary moments of moments $\mathrm{MoM}_{U(N)}(1,1)$ and $\mathrm{MoM}_{U(N)}(2,1)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)).



(a) Roots of $\mathrm{MoM}_{U(N)}(3,1)$.

(b) Roots of $\mathrm{MoM}_{U(N)}(4,1)$.

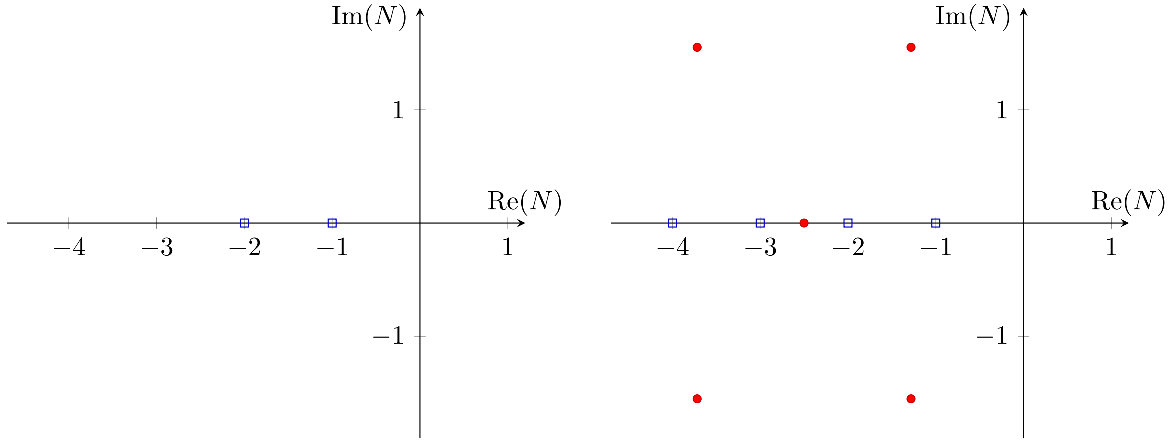
Figure 5.2: Plots of the roots of the unitary moments of moments $\mathrm{MoM}_{U(N)}(3,1)$ and $\mathrm{MoM}_{U(N)}(4,1)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots.



(a) Roots of $\mathrm{MoM}_{U(N)}(1,2)$.

(b) Roots of $\mathrm{MoM}_{U(N)}(2,2)$.

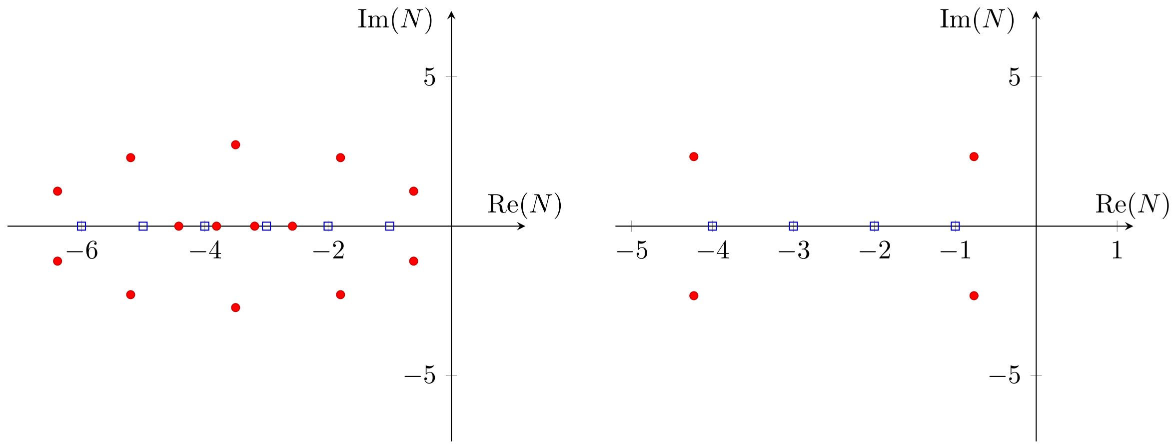
Figure 5.3: Plots of the roots of the unitary moments of moments $\mathrm{MoM}_{U(N)}(1,2)$ and $\mathrm{MoM}_{U(N)}(2,2)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots. Note that $\mathrm{MoM}_{U(N)}(1,2)$ has a repeated root at $N = -2$.



(a) Roots of $\mathrm{MoM}_{Sp(2N)}(1,1)$.

(b) Roots of $\mathrm{MoM}_{Sp(2N)}(1,2)$.

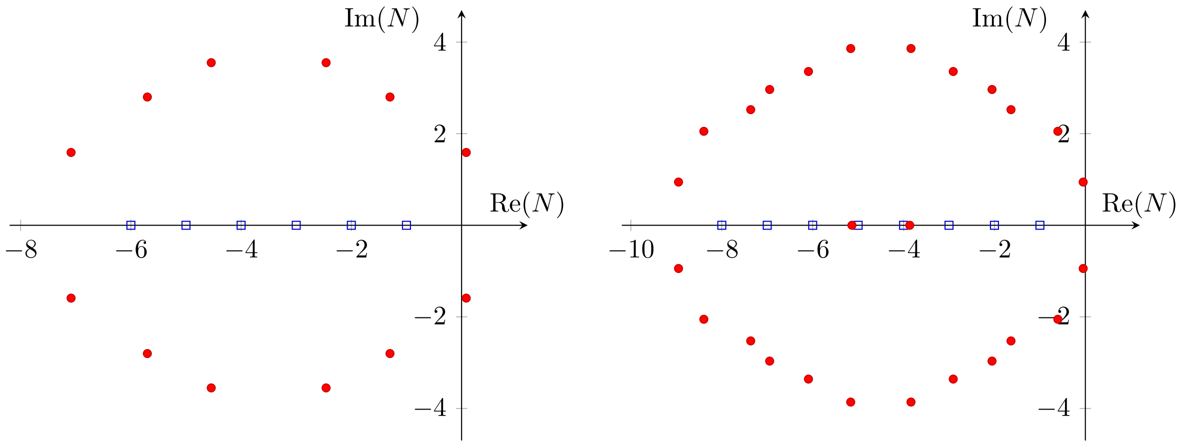
Figure 5.4: Plots of the roots of the symplectic moments of moments $\mathrm{MoM}_{Sp(2N)}(1,1)$ and $\mathrm{MoM}_{Sp(2N)}(1,2)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots.



(a) Roots of $\mathrm{MoM}_{Sp(2N)}(1,3)$.

(b) Roots of $\mathrm{MoM}_{Sp(2N)}(2,1)$.

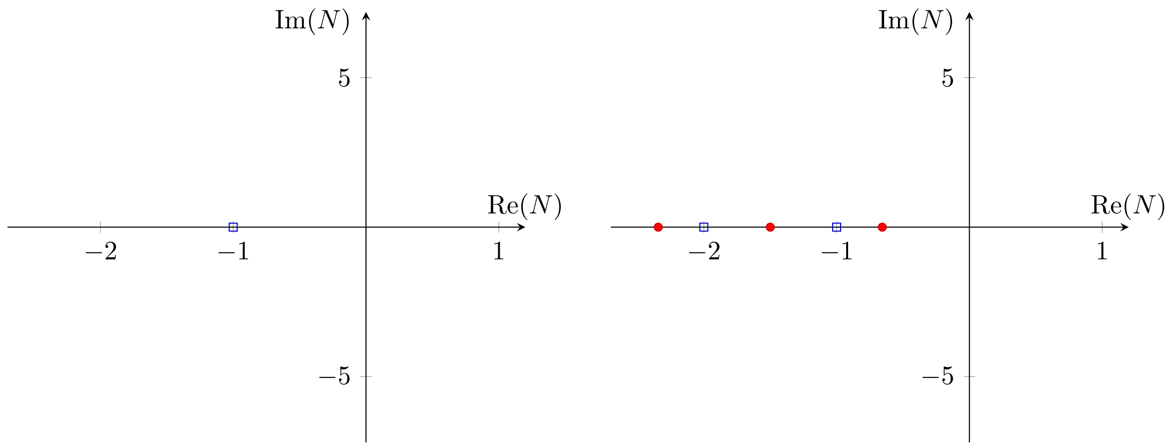
Figure 5.5: Plots of the roots of the symplectic moments of moments $\mathrm{MoM}_{Sp(2N)}(1,3)$ and $\mathrm{MoM}_{Sp(2N)}(2,1)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots.



(a) Roots of $\mathrm{MoM}_{Sp(2N)}(3,1)$.

(b) Roots of $\mathrm{MoM}_{Sp(2N)}(2,2)$.

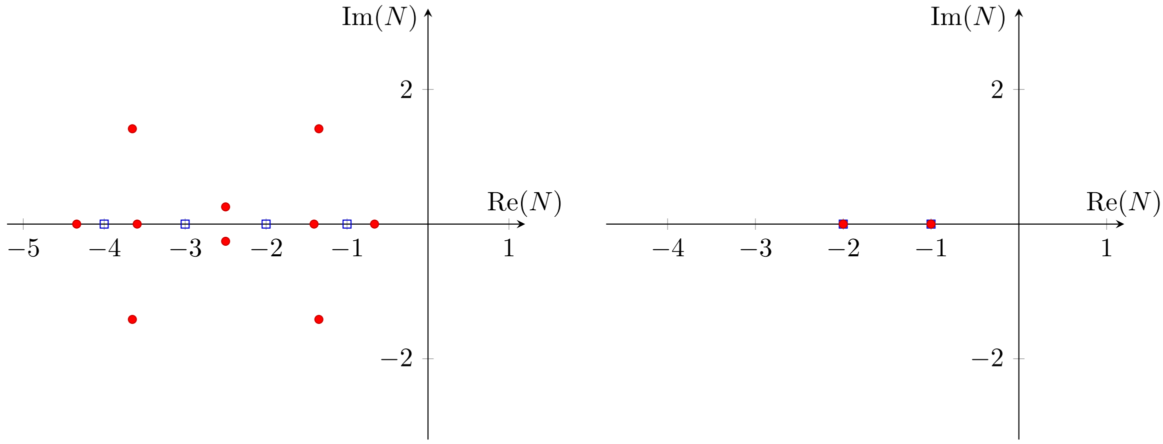
Figure 5.6: Plots of the roots of the symplectic moments of moments $\mathrm{MoM}_{Sp(2N)}(3,1)$ and $\mathrm{MoM}_{Sp(2N)}(2,2)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots.



(a) Roots of $\mathrm{MoM}_{SO(2N)}(1,1)$.

(b) Roots of $\mathrm{MoM}_{SO(2N)}(1,2)$.

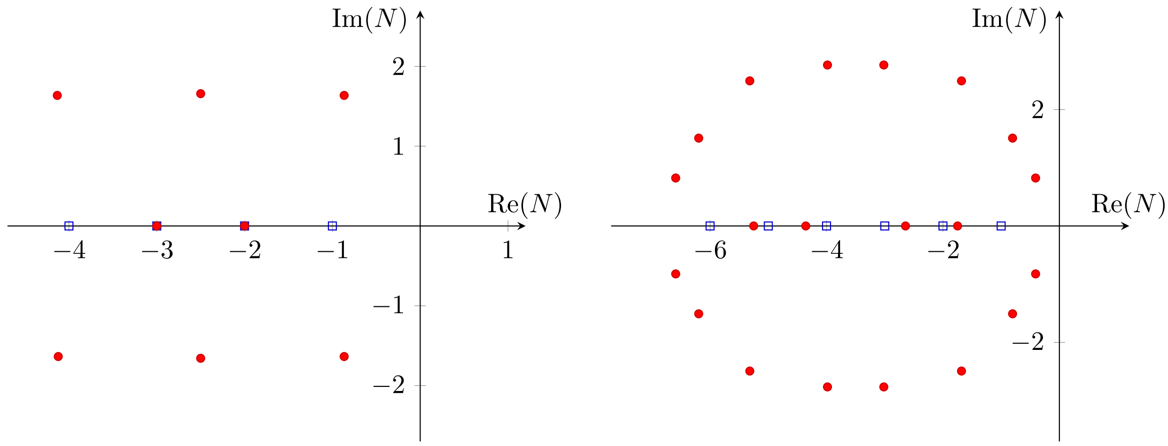
Figure 5.7: Plots of the roots of the orthogonal moments of moments $\mathrm{MoM}_{SO(2N)}(1,1)$ and $\mathrm{MoM}_{SO(2N)}(2,1)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots.



(a) Roots of $\mathrm{MoM}_{SO(2N)}(1,3)$.

(b) Roots of $\mathrm{MoM}_{SO(2N)}(2,1)$.

Figure 5.8: Plots of the roots of the orthogonal moments of moments $\mathrm{MoM}_{SO(2N)}(1,3)$ and $\mathrm{MoM}_{SO(2N)}(2,1)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots. Note that $\mathrm{MoM}_{SO(2N)}(2,1)$ has repeated roots at $N = -1, -2$.



(a) Roots of $\mathrm{MoM}_{SO(2N)}(3,1)$.

(b) Roots of $\mathrm{MoM}_{SO(2N)}(2,2)$.

Figure 5.9: Plots of the roots of the orthogonal moments of moments $\mathrm{MoM}_{SO(2N)}(3,1)$ and $\mathrm{MoM}_{SO(2N)}(2,2)$. The (blue) square points are the roots at $N = -1, -2, \ldots, -g(k,\beta)$ (see (5.4.1)), and the (red) circle points are the remaining roots. Note that $\mathrm{MoM}_{SO(2N)}(3,1)$ has repeated roots at $N = -2, -3$.

Chapter 6

Branching model of moments of moments

In this chapter we outline a model for the moments discussed in chapters 3 and 4. The work presented in this chapter is original to this thesis. Recall that the moments of moments are defined by

$$\mathrm{MoM}_{G(N)}(k,\beta) := \mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} |P_{G(N)}(A,\theta)|^{2\beta} d\theta\right)^k\right],$$

where $G(N) \in \{U(N), SO(2N), Sp(2N)\}$. In chapter 3 we showed that for integer k,β, $\mathrm{MoM}_{U(N)}(k,\beta)$ is a polynomial in N of degree $k^2\beta^2 - k + 1$, in line with a conjecture of Fyodorov and Keating (conjecture 2.2.4). Similarly, in chapter 4, we showed that, again for integer k,β, $\mathrm{MoM}_{Sp(2N)}(k,\beta)$ and $\mathrm{MoM}_{SO(2N)}(k,\beta)$ are also polynomials but of degree $k\beta(2k\beta+1)-k$ and $k\beta(2k\beta-1)-k$ respectively[1]. Recall also that precise information about, for example, the unitary moments of moments (in particular, for non-integer β) would lead to progress towards conjecture 2.2.2.

A key benefit of the model that we develop in this chapter is that it permits the moment parameter β to be non-integer. Whilst results in this direction are known in the unitary case (see section 3.1.1 for a discussion of the work of Claeys and Krasovsky [36] and the extension by Fahs [72]) the benefit of this approach is that it utilises the 'approximate branching structure' which was discussed at length in chapter 2. The identification of such structure was intrinsic to the progress towards proving conjectures 2.2.2 and 2.2.3.

By examining the model in the unitary case for $k = 1,2,3$ and $\beta > 0$ real, we are able to show that the moments of moments for the model follow the Fyodorov-Keating conjecture 2.2.4 including in the 'high temperature' (small β) range and at the critical point $k\beta^2 = 1$. More generally, for $k \in \mathbb{N}$, $k \geq 2$, we also recover the 'low temperature' asymptotic behaviour and the result is comparable to our theorem 3.1.4.

Additionally, this model permits us to investigate the symplectic and orthogonal moments of moments. The only results at time of writing are those stated in chapter 4 and hold for integer moments parameters. We are able to show that, just as in the unitary case, the model captures the 'low temperature' asymptotic behaviour, recovering statements akin to those given in chapter 4, though for non-integer β. This calculation is given in section 6.2.

[1] Except for $\mathrm{MoM}_{SO(2N)}(1,1) = 2(N+1)$.

6.1 The unitary model

Here we introduce the model for the unitary moments of moments. The notation we set out here, will follow through for the symplectic and orthogonal cases in section 6.2. We also note that model involves Gaussian random walks on binary trees, and such processes were defined in chapter 1, section 1.1.4 and explored significantly in chapter 2. Furthermore, since we focus on unitary polynomials, throughout this section we write $P_N(A,\theta) \equiv P_{U(N)}(A,\theta)$ for $A \in U(N)$. As in chapter 2, we write $X_n(l)$ for a particular path through a binary tree of depth n from root to leaf l, where each branch at depth m is weighted by $Y_m(l) \sim \mathcal{N}\left(0, \frac{1}{2}\log 2\right)$

$$X_n(l) = \sum_{m=1}^{n} Y_m(l). \tag{6.1.1}$$

Note that

$$X_n(l) \sim \mathcal{N}(0, \frac{n}{2}\log 2) \tag{6.1.2}$$

and the distribution of $X_n(l)$ does not depend on l (nor does the distribution of $Y_m(l)$ depend on m nor l), but such notation will become useful later. The model is the following. Firstly, we estimate, for a fixed matrix $A \in U(N)$, the average of the characteristic polynomial of A over the circle by a discrete sum,

$$\frac{1}{2\pi}\int_0^{2\pi} |P_N(A,\theta)|^{2\beta} d\theta \sim \frac{1}{N}\sum_{l=1}^{N} |P_N(A,\phi_l)|^{2\beta}, \tag{6.1.3}$$

where $\phi_1, \ldots, \phi_N$ are, for example, taken to be the mid-points between eigenvalues of A. Figure 6.1 shows a plot of $|P_{20}(A,\theta)|$, for some A drawn from $U(20)$. Then, the right hand side of (6.1.3) is just the Riemann sum approximation to the integral over θ.

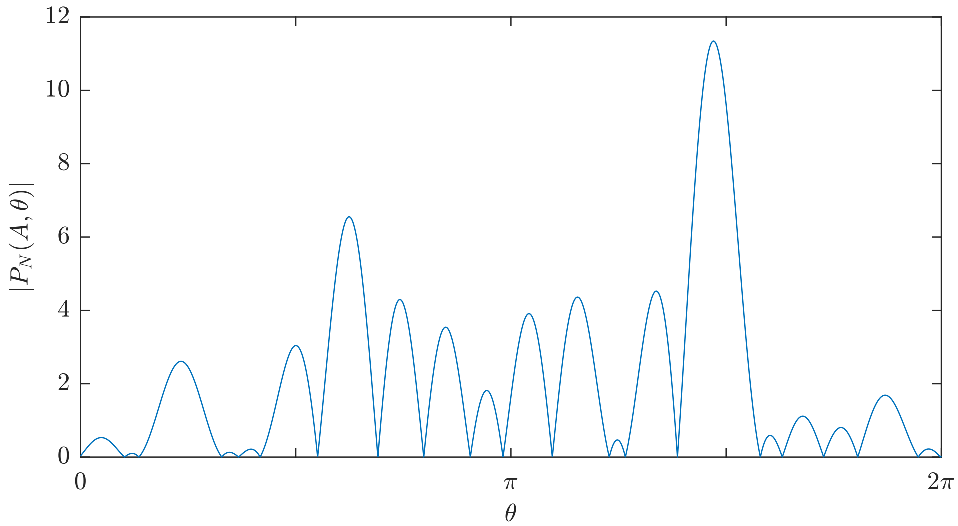


Figure 6.1: Plot of $|P_N(A,\theta)|$, for $\theta \in [0, 2\pi)$ and $A \in U(20)$.

We trivially manipulate the right hand side of (6.1.3) in the following way,

$$\frac{1}{N}\sum_{l=1}^{N} |P_N(A,\phi_l)|^{2\beta} = \frac{1}{N}\sum_{l=1}^{N} e^{2\beta \log|P_N(A,\phi_l)|} \tag{6.1.4}$$

where (6.1.4) now takes the form of a partition function (see chapter 1, section 1.1.4 and chapter 2, section 2.2.2). Ignoring issues of convergence, we replace $\log|P_N(A,\phi_l)|$ with the random variable $X_n(l) \sim \mathcal{N}(0, \frac{1}{2}\log N)$ (i.e. the Keating-Snaith random matrix central limit theorem, see theorem 2.1.3). We identify the random variable $X_n(l)$ with the branching random walk in (6.1.3). Furthermore, in order to emphasise the branching structure, we set $N = 2^n$ for some $n \in \mathbb{N}$. Thus, we compare

$$\frac{1}{N}\sum_{l=1}^{N}|P_N(A,\phi_l)|^{2\beta} \leftrightsquigarrow \frac{1}{2^n}\sum_{l=1}^{2^n} e^{2\beta X_n(l)} \tag{6.1.5}$$

$$= \frac{1}{2^n}\sum_{l=1}^{2^n} e^{2\beta\sum_{m=1}^{n} Y_m(l)} \tag{6.1.6}$$

where, as in (6.1.1), $Y_m(l) \sim \mathcal{N}(0, \frac{1}{2}\log 2)$ (which has no dependence on m nor l, but this emphasis will play a role later).

Under this model, the moments of moments become:

$$\mathrm{MoM}_{U(N)}(k,\beta) = \mathbb{E}_{U(N)}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_N(A,\theta)|^{2\beta}d\theta\right)^k\right] \tag{6.1.7}$$

$$\leftrightsquigarrow \mathbb{E}\left[\left(\frac{1}{2^n}\sum_{l=1}^{2^n} e^{2\beta X_n(l)}\right)^k\right] \tag{6.1.8}$$

$$= \frac{1}{2^{nk}}\sum_{l_1=1}^{2^n}\cdots\sum_{l_k=1}^{2^n}\mathbb{E}\left[e^{2\beta(X_n(l_1)+\cdots+X_n(l_k))}\right], \tag{6.1.9}$$

where the expectation in (6.1.8) and (6.1.9) is with respect to the Gaussian random variables $X_n(l_j)$.

Recall that we showed in chapter 3 that, for $k,\beta \in \mathbb{N}$,

$$\lim_{N\to\infty}\frac{\mathbb{E}_{U(N)}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_N(A,\theta)|^{2\beta}d\theta\right)^k\right]}{N^{k^2\beta^2-k+1}} = \gamma_{k,\beta} \tag{6.1.10}$$

for some positive constant $\gamma_{k,\beta}$ depending only on k,β.

Using the model of the moments of moments, we are able to recover an asymptotic result of the form (6.1.10), albeit with a different constant to $\gamma_{k,\beta}$. However, the benefit of using the model, as stated above, is that it holds for non-integer β, and exploits a branching structure. Furthermore, for $k = 2$ and $k = 3$, we are able demonstrate the phase transition through $k\beta^2 = 1$ as predicted by Fyodorov and Keating [82]. We anticipate that the calculations outlined below can be extended to all $k \in \mathbb{N}$.

6.1.1 The single particle case

In order to introduce the key tools required to analyse the model, we begin with the simplest case of $k = 1$ in (6.1.9), and $\beta > 0$. We find that

$$\frac{1}{2^n}\sum_{l=1}^{2^n}\mathbb{E}\left[e^{2\beta X_n(l)}\right] = \frac{1}{2^n}\sum_{l=1}^{2^n}\mathbb{E}\left[\prod_{m=1}^{n} e^{2\beta Y_m(l)}\right] \tag{6.1.11}$$

$$= \frac{1}{2^n}\sum_{l=1}^{2^n}\prod_{m=1}^{n}\mathbb{E}\left[e^{2\beta Y_m(l)}\right], \tag{6.1.12}$$

since the $Y_m(l)$ are independent. Then, the expectation in (6.1.12) is simply the moment generating function for a centred normal random variable with variance $\frac{1}{2}\log 2$. Hence

$$\frac{1}{2^n}\sum_{l=1}^{2^n}\mathbb{E}\left[e^{2\beta X_n(l)}\right] = \frac{1}{2^n}\sum_{l=1}^{2^n} e^{n\beta^2\log 2} \tag{6.1.13}$$

$$= \frac{1}{2^n}\sum_{l=1}^{2^n}(2^n)^{\beta^2} \tag{6.1.14}$$

$$= N^{\beta^2}, \tag{6.1.15}$$

where recall $N = 2^n$. Recall that the leading order of $\mathrm{MoM}_{U(N)}(1,\beta)$ is N^{β^2} (see theorem 2.1.1 or theorem 3.1.4). Hence, the model captures the correct leading order (though with a different leading order coefficient).

6.1.2 Generalizing to higher values of k

If one takes $k \in \mathbb{N}$, $k > 1$, then the problem no longer simplifies to a simple moment generating function calculation. Instead, as k increases, we think of this as 'loading' the binary tree with k particles at the top (root), each of which fall through a path corresponding to a particular choice of leaf.

Figure 6.2 gives a simple example of this concept. It shows a particular summand in (6.1.9) for $k = 3$ and $N = 2^4 = 16$. In this example, we imagine 3 particles being loaded to the root, and so each summand corresponds to a particular choice of leaf for each particle. Figure 6.2 shows the summand for $\circ$ finishing on the 4th leaf, $\diamond$ falling to the 7th , and $\star$ landing on the 16th leaf (i.e. $l_1 = 4, l_2 = 7, l_3 = 16$ in (6.1.9)). The full 'moments of moments' would correspond to calculating the contribution over all leaves, for all of the particles.

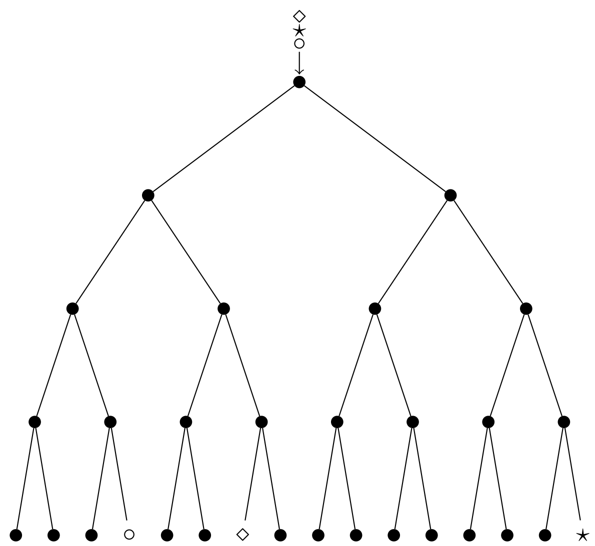


Figure 6.2: A binary tree of depth 4 with three particles $\diamond, \star, \circ$ loaded to the root. Each particle's path through the tree is dictated by a choice of leaf (indicated in the figure by respective symbol).

By employing Jensen's inequality we can prove the following bound. Informally, it shows that when particles 'stick together' they produce a bigger contribution than a path where they diverge. Firstly, write λ for the length of a path from root to depth $\lambda \in \{1, \ldots, n\}$. Then, for some $\kappa \in \{1, \ldots, k\}$, and $Y_m(l) \sim \mathcal{N}(0, \frac{1}{2}\log 2)$,

$$\mathbb{E}\left[\prod_{m=1}^{\lambda} e^{2\beta\kappa Y_m(l)}\right] \geq \left(\mathbb{E}\left[\prod_{m=1}^{\lambda} e^{2\beta Y_m(l)}\right]\right)^{\kappa}. \tag{6.1.16}$$

We interpret the left hand side of (6.1.16) as the contribution to the sum (6.1.9) of κ particles following the same path of length λ. Comparatively, the right hand side (which provides a lower bound) is the contribution of κ independent particles each following their own path of length λ.

Before we introduce the next lemma (which deals with counting paths) the following definition, a generalization of definition 1.1.9, is useful.

Definition 6.1.1. *The last common ancestor* $\mathrm{lca}(l_1, l_2)$ *of two leaves* l_1 *and* l_2 *of a binary tree is the level of the lowest (i.e. furthest from the root) node that has both* l_1 *and* l_2 *as descendants. The last common ancestor of a collection of leaves* $l_1, \ldots, l_k$*, written* $\mathrm{lca}(l_1, \ldots, l_k)$ *is the level of the lowest node that has all* $l_1, \ldots, l_k$ *as descendants. Figure 6.3 gives an example.*

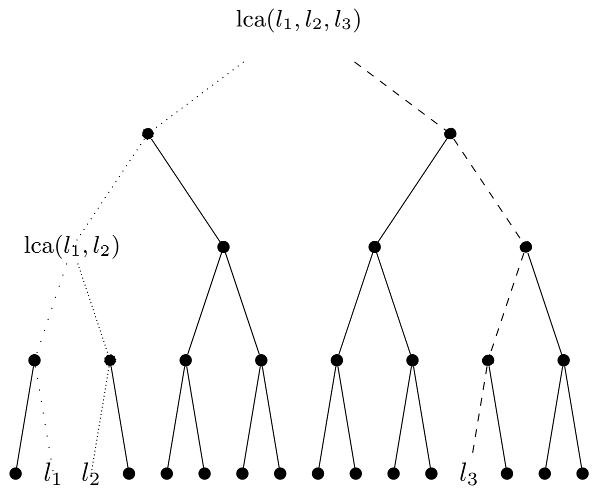


Figure 6.3: A binary tree of depth 4 with three leaves l_1, l_2, l_3 highlighted. The last common ancestor of l_1, l_2 is $\mathrm{lca}(l_1, l_2)$. The last common ancestor of all three (and also $\mathrm{lca}(l_2, l_3)$ and $\mathrm{lca}(l_1, l_3)$) is the root node. The paths are differentiated by dashed and dotted lines.

We now give a lemma which calculates the number of paths sharing a given common ancestor.

Lemma 6.1.2. *Fix a level* λ *in a binary tree of depth* n*, so* $\lambda \in \{0, \ldots, n\}$*. Load the tree with* k *particles whose paths through the tree are determined by leaves* $l_1, \ldots, l_k$*. Let* $c_\lambda(k, n)$ *be the number of paths through a tree of depth* n *traced by the* k *particles such that the last common ancestor* $\mathrm{lca}(l_1, \ldots, l_k)$ *is on level* λ*. Then*

$$c_\lambda(k, n) = 2^\lambda \left(2^{n-\lambda}\right)^k - 2^{\lambda+1} \left(2^{n-(\lambda+1)}\right)^k, \quad \textit{for } \lambda \in \{0, \ldots, n-1\}, \tag{6.1.17}$$

$$c_n(k, n) = 2^n. \tag{6.1.18}$$

Proof. The total number of possible paths that one particle can take is equal to the number of leaves, 2^n. Thus, the total number of paths for k particles is 2^{nk}. Equivalently, the total number of paths is given by summing $c_\lambda(k, n)$ over the levels,

$$2^{nk} = \sum_{\lambda=0}^{n} c_\lambda(k, n) = c_0(k, n) + \sum_{\lambda=1}^{n} c_\lambda(k, n). \tag{6.1.19}$$

The term $c_0(k, n)$ can be calculated easily since it is the number of collections of paths which have at least one path in each 'half' of the tree. Equally, this is the total number of paths less those which stick to one side of the tree:

$$c_0(k,n) = 2^{nk} - 2\left(2^{k(n-1)}\right). \tag{6.1.20}$$

Hence

$$2^{nk} - 2^{k(n-1)+1} + \sum_{\lambda=1}^{n} c_\lambda(k,n) = 2^{nk}, \tag{6.1.21}$$

or equivalently

$$\sum_{\lambda=1}^{n} c_\lambda(k,n) = 2^{k(n-1)+1}. \tag{6.1.22}$$

The same process can be repeated for $\lambda = 1$,

$$\begin{aligned} c_1(k,n) &= 2^{nk} - c_0(k,n) - \sum_{\lambda=2}^{n} c_\lambda(k,n) && (6.1.23)\\ &= 2^{nk} - 2^{nk} + 2\left(2^{k(n-1)}\right) - 2^2\left(2^{k(n-2)}\right) && (6.1.24)\\ &= 2^{nk+1}\left(2^{-k} - 2^{-2k+1}\right). && (6.1.25) \end{aligned}$$

Inductively, we find that for $\lambda = 0, \ldots, n-1$,

$$c_\lambda(k,n) = 2^\lambda\left(2^{k(n-\lambda)} - 2^{\lambda+1}\left(2^{k(n-(\lambda+1))}\right)\right). \tag{6.1.26}$$

One can calculate $c_n(k,n)$ by noting that this is the number of paths with last common ancestor at level n. Hence, the only choice is for all k particles to stick together and $c_n(k,n) = 2^n$, the number of leaves. Alternatively, we can arrive at the same conclusion using the (telescopic) sum (6.1.19):

$$\begin{aligned} 2^{nk} &= \sum_{\lambda=0}^{n-1}\left(2^\lambda\left(2^{k(n-\lambda)}\right) - 2^{\lambda+1}\left(2^{k(n-(\lambda+1))}\right)\right) + c_n(k,n) && (6.1.27)\\ &= 2^{nk} - 2\cdot 2^{k(n-1)} + 2\cdot 2^{k(n-1)} - 2^2\cdot 2^{k(n-2)} + \cdots + 2^{n-1}\cdot 2^k - 2^n + c_n(k,n) && (6.1.28)\\ &= 2^{nk} - 2^n + c_n(k,n), && (6.1.29) \end{aligned}$$

hence $c_n(k,n) = 2^n$. □

The following proposition shows that we can recover the full Fyodorov-Keating conjectured asymptotic behaviour (see conjecture 2.2.4) for the branching model, for $k = 2$.

Proposition 6.1.3. *Take $\beta > 0$ and $k = 2$ and let $N = 2^n$. Then,*

$$\frac{1}{N^2}\sum_{l_1,l_2=1}^{N} \mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2))}\right] \overset{N\to\infty}{\sim} \begin{cases} c_1 N^{4\beta^2-1} & \text{if } 2\beta^2 > 1\\ c_2 N\log N & \text{if } 2\beta^2 = 1\\ c_3 N^{2\beta^2} & \text{if } 2\beta^2 < 1, \end{cases} \tag{6.1.30}$$

for some positive constants c_1, c_2, c_3.

Proof. We begin with the left hand side of (6.1.30), and split in to 'diagonal' and 'off-diagonal' cases.

$$\frac{1}{N^2}\sum_{l_1,l_2=1}^{N} \mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2))}\right] = \frac{1}{N^2}\left(\sum_{l_1=1}^{N}\mathbb{E}\left[e^{4\beta X_n(l_1)}\right] + \sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2))}\right]\right). \tag{6.1.31}$$

Then the diagonal term is calculated just as in section 6.1.1:

$$\frac{1}{N^2}\sum_{l_1=1}^{N}\mathbb{E}\left[e^{4\beta X_n(l_1)}\right] = \frac{1}{N^2}\sum_{l_1=1}^{N} e^{4\beta^2 n\log 2} \tag{6.1.32}$$

$$= N^{4\beta^2-1}. \tag{6.1.33}$$

To handle the off-diagonal case, we use last common ancestors. Since $l_1 \neq l_2$, the two paths traced by the particles must diverge at level $\lambda = \mathrm{lca}(l_1, l_2)$, for some $\lambda \in 0, \ldots, n-1$. We use this fact to break the path in to two sections: the initial section (i.e. up to λ) where the two paths are the same, and the latter section, after λ, where the paths diverge.

$$\frac{1}{N^2}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N}\mathbb{E}\left[\ e^{2\beta(X_n(l_1)+X_n(l_2))}\right]$$

$$= \frac{1}{N^2}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N} 2^{4\beta^2\lambda}\mathbb{E}\left[e^{2\beta(Y_{\lambda+1}(l_1)+\cdots+Y_n(l_1))}e^{2\beta(Y_{\lambda+1}(l_2)+\cdots+Y_n(l_2))}\right]. \tag{6.1.34}$$

As remarked above, two paths from level $\lambda+1$ to n, i.e. $Y_{\lambda+1}(l_1)+\cdots+Y_n(l_1)$ and $Y_{\lambda+1}(l_2)+\cdots+Y_n(l_2)$, must too be independent since their last common ancestor lives at level λ. Hence

$$\frac{1}{N^2}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2))}\right] = \frac{1}{N^2}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N} 2^{4\beta^2\lambda}2^{2\beta^2(n-\lambda)}. \tag{6.1.35}$$

Now we shift from summing over leaves to summing over levels λ, and employ lemma 6.1.2.

$$\frac{1}{N^2}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2}}^{N} 2^{4\beta^2\lambda}2^{2\beta^2(n-\lambda)} = \frac{N^{2\beta^2}}{N^2}\sum_{\lambda=0}^{n-1} c_\lambda(2,n)2^{2\beta^2\lambda} \tag{6.1.36}$$

$$= \frac{N^{2\beta^2}}{N^2}\sum_{\lambda=0}^{n-1}\left(2^{2n-\lambda} - 2^{2n-(\lambda+1)}\right)2^{2\beta^2\lambda} \tag{6.1.37}$$

$$= \frac{N^{2\beta^2}}{N^2}\frac{N(N^{2\beta^2}-N)}{2^{2\beta^2}-2}. \tag{6.1.38}$$

One can conclude from (6.1.38) that there are three different leading order behaviours, depending on the value of β. When $2\beta^2 > 1$, the leading order grows like $N^{4\beta^2-1}$. Instead, if $2\beta^2 < 1$ then $N > N^{2\beta}$ and the overall leading order is $N^{2\beta^2}$. Finally, when $2\beta^2 = 1$, after an application of L'Hôpital's rule, the asymptotic behaviour is $N\log N$. Thus, overall for the branching model of the moments of moments we have, for $k = 2$,

$$\frac{1}{N^2}\sum_{l_1=1}^{N}\sum_{l_2=1}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2))}\right] \overset{N\to\infty}{\sim} \begin{cases} c_1 N^{4\beta^2-1} & \text{if } 2\beta^2 > 1,\\ c_2 N\log N & \text{if } 2\beta^2 = 1,\\ c_3 N^{2\beta^2} & \text{if } 2\beta^2 < 1,\end{cases} \tag{6.1.39}$$

for positive constants c_1, c_2, c_3 deduced from (6.1.33) and (6.1.38). □

In order to prove the similar result for $k = 3$, we first need to extend lemma 6.1.2.

Lemma 6.1.4. *Consider* 3 *particles loaded to a binary tree of depth* n*. Let* $\lambda = \text{lca}(l_1, l_2, l_3)$ *for* $l_j \in \{1, \dots, 2^n\}$ *(so* $\lambda \in \{0, \dots, n\}$*). Then let* $\lambda_{i,j} := \text{lca}(l_i, l_j)$ *for* $i, j \in \{1, 2, 3\}$*,* $i \neq j$*. Define*

$$\Lambda := \max_{\substack{i,j \in \{1,2,3\} \\ i \neq j}} \lambda_{i,j}.$$

Note $\Lambda \in \{\lambda + 1, \dots, n\}$ *unless* $\lambda = n$*, in which case we define* $\Lambda = n$*. In figure 6.3 for example,* $\lambda = 0$ *since immediately the paths split, and* $\Lambda = \text{lca}(l_1, l_2) = 2$ *since* $\text{lca}(l_1, l_3) = \text{lca}(l_2, l_3) = 0$*. Then, similarly to how* c_λ *was defined in lemma 6.1.2, define* $c_{\lambda,\Lambda}(3, n)$ *to be the number of paths through a tree of depth* n *traced by* 3 *particles with a given* λ, Λ*. We have*

$$c_{\lambda,\Lambda}(3, n) = \begin{cases} 3 \cdot 2^{3n-\lambda-\Lambda-2} & \text{if } \lambda \in \{0, \dots, n-1\}, \Lambda \in \{\lambda+1, \dots, n-1\}, \\ 3 \cdot 2^{2n-\lambda-1} & \text{if } \lambda \in \{0, \dots, n-1\}, \Lambda = n, \\ 2^n & \text{if } \lambda = n, \Lambda = n. \end{cases} \tag{6.1.40}$$

Proof. As usual, if $\lambda = n$ (note this doesn't allow Λ to be defined but notationally we write $\Lambda = n$), then $c_{n,n}(3, n)$ is the number of choices of paths that all follow the same route from root to leaf. Since the tree has depth n, we have $c_{n,n}(3, n) = 2^n$.

For $\lambda \in \{0, \dots, n-1\}$ and $\Lambda \in \{\lambda+1, \dots, n-1\}$ we break the calculation down in to stages. There are 2^λ choices for paths from root to level λ. After λ, the three paths must diverge. There are $2 \cdot 3$ choices for the next step: any one of the three paths goes one direction (hereafter referred to as the 'solo' path), and the remaining two go the other. Thereafter, the solo path has $2^{n-\lambda-1}$ choices of leaves to land on. For the remaining paths, they have a choice of $2^{n-\lambda-1}2^{n-\Lambda-1}$ options. Thus, overall for $\lambda \in \{0, \dots, n-1\}$ and $\Lambda \in \{\lambda+1, \dots, n-1\}$,

$$c_{\lambda,\Lambda}(3, n) = 2^\lambda \times (2 \cdot 3) \times 2^{n-\lambda-1}2^{n-\lambda-1}2^{n-\Lambda-1} \tag{6.1.41}$$

$$= 3 \cdot 2^{3n-\lambda-\Lambda-2}. \tag{6.1.42}$$

Finally, for $c_{\lambda,n}(3, n)$, $\lambda \in \{0, \dots, n-1\}$, similarly we find 2^λ choices for the initial path, then $2 \cdot 3$ for the option of subsequent step, then $2^{n-\lambda-1}$ to select from for the solo path, and $2^{n-\lambda-1}$ alternatives for the paired path giving

$$c_{\lambda,n}(3, n) = 3 \cdot 2^{2n-\lambda-1}.$$

As verification, by summing through λ, Λ, we should recover the total number of paths, 2^{3n}.

$$2^n + \sum_{\lambda=0}^{n-1} \left(\sum_{\Lambda=\lambda+1}^{n-1} c_{\lambda,\Lambda}(3, n) + c_{\lambda,n}(3, n) \right) = 2^n + 3 \sum_{\lambda=0}^{n-1} \left(2^{3n-(\lambda+1)} \left(2^{-(\lambda+1)} - 2^{-n} \right) \right)$$

$$+ 3 \sum_{\lambda=0}^{n-1} 2^{2n-(\lambda+1)} \tag{6.1.43}$$

$$= 2^n + 3 \cdot 2^{2n} \cdot \frac{1}{3} 2^{-n} (2^{2n} - 1) \tag{6.1.44}$$

$$= 2^{3n}, \tag{6.1.45}$$

thus completing the proof. □

Now we are ready to handle the case of $k = 3$.

Proposition 6.1.5. *Take $\beta > 0$ and $k = 3$, and let $N = 2^n$. Then,*

$$\frac{1}{N^3}\sum_{l_1,l_2,l_3=1}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right] \overset{N\to\infty}{\sim} \begin{cases} c_1 N^{9\beta^2-2} & \text{if } 3\beta^2 > 1 \\ c_2 N\log N & \text{if } 3\beta^2 = 1 \\ c_3 N^{3\beta^2} & \text{if } 3\beta^2 < 1, \end{cases} \tag{6.1.46}$$

for some positive constants c_1, c_2, c_3.

Proof. Just as with $k = 2$, we separate in to the diagonal and off-diagonal cases,

$$\frac{1}{N^3}\sum_{l_1,l_2,l_3=1}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right]$$
$$= \frac{1}{N^3}\left(\sum_{l_1=1}^{N}\mathbb{E}\left[e^{6\beta X_n(l_1)}\right] + \sum_{\substack{l_1,l_2,l_3=1\\ l_1\neq l_2\neq l_3}}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right]\right). \tag{6.1.47}$$

Then the diagonal term is

$$\frac{1}{N^3}\sum_{l_1=1}^{N}\mathbb{E}\left[e^{6\beta X_n(l_1)}\right] = \frac{1}{N^3}\sum_{l_1=1}^{N} e^{9\beta^2 n\log 2} \tag{6.1.48}$$
$$= N^{9\beta^2-2}. \tag{6.1.49}$$

To handle the off-diagonal case, once more we use last common ancestors. We let λ be the level up to which all three routes follow the same path. Thereafter, one path breaks off in one direction and two go in the other. Λ will be the level beneath λ such that thereafter all paths follow different routes (i.e. the level at which the paired path diverge). For a given λ, Λ, lemma 6.1.4 gives the number of paths displaying this behaviour. Thus,

$$\frac{1}{N^3}\sum_{\substack{l_1,l_2,l_3=1\\ l_1\neq l_2\neq l_3}}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right]$$
$$= \frac{1}{N^3}\sum_{\substack{l_1,l_2,l_3=1\\ l_1\neq l_2\neq l_3}}^{N} 2^{9\beta^2\lambda}\mathbb{E}\left[e^{2\beta(Y_{\lambda+1}(l_1)+\cdots+Y_n(l_1))}e^{2\beta(Y_{\lambda+1}(l_2)+\cdots+Y_n(l_2))}e^{2\beta(Y_{\lambda+1}(l_3)+\cdots+Y_n(l_3))}\right] \tag{6.1.50}$$

where $\lambda = \mathrm{lca}(l_1, l_2, l_3)$. We now use that one path will be completely independent and of length $n-\lambda$. The other two will be dependent up to Λ and thereafter independent. Hence,

$$\frac{1}{N^3}\sum_{\substack{l_1,l_2,l_3=1\\ l_1\neq l_2\neq l_3}}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right]$$
$$= \frac{1}{N^3}\sum_{\lambda=0}^{n-1} 2^{9\beta^2\lambda}2^{\beta^2(n-\lambda)}\left(\sum_{\Lambda=\lambda+1}^{n-1} c_{\lambda,\Lambda}(3,n)2^{4\beta^2(\Lambda-\lambda)}2^{2\beta^2(n-\Lambda)} + c_{\lambda,n}(3,n)2^{4\beta^2(n-\lambda)}\right) \tag{6.1.51}$$
$$= \frac{3N^{\beta^2}}{N^3}\sum_{\lambda=0}^{n-1} 2^{8\beta^2\lambda}\left[N^{2\beta^2+3}2^{-(\lambda+1)}2^{-4\beta^2\lambda}\sum_{\Lambda=\lambda+1}^{n-1}2^{-(\Lambda+1)}2^{2\beta^2\Lambda} + N^{4\beta^2+2}2^{-(\lambda+1)}2^{-4\beta^2\lambda}\right] \tag{6.1.52}$$
$$= \frac{3N^{\beta^2}}{N^3}\sum_{\lambda=0}^{n-1} 2^{4\beta^2\lambda}2^{-(\lambda+1)}\Big(2^{-(\lambda+1)}N^{2\beta^2+3}\frac{N^{2\beta^2-1}2^{\lambda+1} - 2^{2\beta^2(\lambda+1)}}{2^{2\beta^2}-2} + N^{4\beta^2+2}\Big). \tag{6.1.53}$$

Evaluating the final sum and simplifying, we overall find that

$$\frac{1}{N^3}\sum_{l_1,l_2,l_3=1}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+X_n(l_2)+X_n(l_3))}\right] \tag{6.1.54}$$

$$= N^{9\beta^2-2}+3N^{\beta^2-2}\left(\frac{N^{8\beta^2}(2^{6\beta^2}-2)(2^{2\beta^2}-2)}{(2^{6\beta^2}-2^2)(2^{4\beta^2}-2)(2^{2\beta^2}-2)}\right.$$
$$\left.+\frac{N^{2\beta^2+2}2^{2\beta^2}(2^{4\beta^2}-2)}{(2^{6\beta^2}-2^2)(2^{4\beta^2}-2)(2^{2\beta^2}-2)}-\frac{N^{4\beta^2+1}(2^{6\beta^2}-2^2)(2^{2\beta^2}-1)}{(2^{6\beta^2}-2^2)(2^{4\beta^2}-2)(2^{2\beta^2}-2)}\right) \tag{6.1.55}$$

$$= N^{9\beta^2-2}+\frac{3N^{9\beta^2-2}(2^{6\beta^2}-2)}{(2^{6\beta^2}-2^2)(2^{4\beta^2}-2)}$$
$$+\frac{3N^{3\beta^2}2^{2\beta^2}}{(2^{6\beta^2}-2^2)(2^{2\beta^2}-2)}-\frac{3N^{5\beta^2-1}(2^{2\beta^2}-1)}{(2^{4\beta^2}-2)(2^{2\beta^2}-2)} \tag{6.1.56}$$

In order to determine the behaviour of (6.1.56) as $N\to\infty$, we investigate which ranges of β result in which power of N dominating. To do this, we compare the different powers of N appearing in (6.1.56) against each other,

$$5\beta^2-1=3\beta^2 \iff 2\beta^2=1, \tag{6.1.57}$$
$$9\beta^2-2=3\beta^2 \iff 3\beta^2=1, \tag{6.1.58}$$
$$9\beta^2-2=5\beta^2-1 \iff 4\beta^2=1. \tag{6.1.59}$$

Hence, provided that $3\beta^2>1$, the dominant power of N is $9\beta^2-2$. When $3\beta^2<1$, instead the leading order is seen to be $N^{3\beta^2}$. Taking limits, we find that at the critical point $3\beta^2=1$, then the order is $N\log N$ (dominating the term from the diagonal contribution at $3\beta^2=1$, which is N). This concludes the proof. □

Remark. In order to extend the calculations for $k=2$ and $k=3$ to higher k, one would have to formulate a general expression of the form (6.1.17) or (6.1.40) for $k\in\mathbb{N}$. Given this, the recipe described in the proofs of propositions 6.1.3 and 6.1.5 would deliver the result.

The next proposition deals the general case of $k\in\mathbb{N}$, and should be compared to the proven asymptotic behaviour of the true unitary moments of moments, see (6.1.10).

Proposition 6.1.6. *Let $k\in\mathbb{N}$ and $\beta>0$, and set $N=2^n$. If $2\beta^2>1$ then*

$$\lim_{N\to\infty}\frac{\frac{1}{N^k}\sum_{l_1=1}^{N}\cdots\sum_{l_k=1}^{N}\mathbb{E}\left[e^{2\beta(X_n(l_1)+\cdots+X_n(l_k))}\right]}{N^{k^2\beta^2-k+1}}=\gamma_{k,\beta}, \tag{6.1.60}$$

where $\gamma_{k,\beta}$ is some non-zero constant depending on k,β.

Remark. Before we give the proof of proposition 6.1.6, we re-emphasise that the techniques used to prove propositions 6.1.3 and 6.1.5 are fully expected to extend to general $k\in\mathbb{N}$. If this were the case, the constraint $2\beta^2>1$ appearing in the statement of proposition 6.1.6 could be strengthened to $k\beta^2>1$, which is more in-keeping with conjecture 2.2.4. Within the proof, given below, we employ a bound (c.f. (6.1.64)) on the contribution from k paths. If we had a proof for general $k\in\mathbb{N}$ along the lines of propositions 6.1.3 and 6.1.5, then we could replace the bound with a (different) equality coming from such a proof, and thus strengthen the condition from $2\beta^2>1$ to $k\beta^2>1$.

However, note that $2\beta^2>1$ implies that $k\beta^2>1$ for $k\in\mathbb{N}$, $k\geq 2$, so proposition 6.1.6 falls in to the low temperature (high β) range of conjecture 2.2.4 and the statement still matches the conjecture.

Proof. As with the cases of $k = 1, 2, 3$, we first split in to diagonal and off-diagonal terms.

$$\frac{1}{N^k} \sum_{l_1,\dots,l_k=1}^{N} \mathbb{E}\left[e^{2\beta(X_n(l_1)+\cdots+X_n(l_k))}\right]$$
$$= \frac{1}{N^k}\left(\sum_{l_1=1}^{N} \mathbb{E}\left[e^{2k\beta X_n(l_1)}\right] + \sum_{\substack{l_1,\dots,l_k=1 \\ l_1\neq\cdots\neq l_k}}^{N} \mathbb{E}\left[e^{2\beta(X_n(l_1)+\cdots+X_n(l_k))}\right]\right). \tag{6.1.61}$$

Then as usual, the diagonal term is easy to compute:

$$\mathbb{E}\left[e^{2k\beta X_n(l_1)}\right] = e^{n\frac{4k^2\beta^2\log 2}{4}} = N^{k^2\beta^2}. \tag{6.1.62}$$

For the off-diagonal case, we let $\lambda = \mathrm{lca}(l_1, \dots, l_k)$, and split off the contribution up to level λ.

$$\frac{1}{N^k} \sum_{\substack{l_1,\dots,l_k=1 \\ l_1\neq\cdots\neq l_k}}^{N} \mathbb{E}\left[e^{2\beta(X_n(l_1)+\cdots+X_n(l_k))}\right] = \frac{1}{N^k} \sum_{\substack{l_1,\dots,l_k=1 \\ l_1\neq\cdots\neq l_k}}^{N} 2^{k^2\beta^2\lambda}\mathbb{E}\left[\prod_{j=1}^{k} e^{2\beta(Y_{\lambda+1}(l_j)+\cdots+Y_n(l_j))}\right]. \tag{6.1.63}$$

Thus, the result follows if we can show that (6.1.63) doesn't grow faster than $N^{k^2\beta^2-k+1}$, provided that $2\beta^2 > 1$.

Focussing on the expectation in the right hand side of (6.1.63), we use (6.1.16) to find an upper bound,

$$\mathbb{E}\left[\prod_{j=1}^{k} e^{2\beta Y_{\lambda+1}(l_j)+\cdots+Y_n(l_j)}\right] \leq 2^{\beta^2(k-1)^2(n-\lambda)}2^{\beta^2(n-\lambda)}. \tag{6.1.64}$$

By (6.1.16) (essentially Jensen's lemma) we have that the more dependent a path is, the bigger its contribution to the sum. Since the paths must split after level λ, the most dependent a path can be is if $k-1$ particles remain joined, and a single particle becomes independent. The contribution from such a situation is exactly the upper bound given in (6.1.64).

Hence,

$$\frac{1}{N^k} \sum_{\substack{l_1,\dots,l_k=1 \\ l_1\neq\cdots\neq l_k}}^{N} 2^{k^2\beta^2\lambda}\mathbb{E}\left[\prod_{j=1}^{k} e^{2\beta(Y_{\lambda+1}(l_j)+\cdots+Y_n(l_j))}\right] \leq \frac{1}{N^k}\sum_{\lambda=0}^{n-1} c_\lambda(k,n) 2^{k^2\beta^2\lambda}2^{\beta^2(k-1)^2(n-\lambda)}2^{\beta^2(n-\lambda)}, \tag{6.1.65}$$

where $c_\lambda(k, n)$ is the usual count of paths having last common ancestor at level λ. Using lemma 6.1.2 and computing the sum on the right hand side of (6.1.65) we find,

$$\frac{1}{N^k} \sum_{\substack{l_1,\dots,l_k=1 \\ l_1\neq\cdots\neq l_k}}^{N} 2^{k^2\beta^2\lambda}\mathbb{E}\left[\prod_{j=1}^{k} e^{2\beta(Y_{\lambda+1}(l_j)+\cdots+Y_n(l_j))}\right]$$
$$\leq \frac{1}{N^k}\sum_{\lambda=0}^{n-1}(2^{kn-(k-1)\lambda} - 2^{kn-(k-1)(\lambda+1)})2^{k^2\beta^2\lambda}2^{\beta^2(k-1)^2(n-\lambda)}2^{\beta^2(n-\lambda)} \tag{6.1.66}$$
$$= \frac{2^{2\beta^2+k} - 2^{2\beta^2+1}}{2^{2\beta^2k+1} - 2^{2\beta^2+k}}(N^{k^2\beta^2-k+1} - N^{k^2\beta^2-2\beta^2(k-1)}) \tag{6.1.67}$$
$$\ll N^{k^2\beta^2-k+1} \tag{6.1.68}$$

since $2\beta^2 > 1$. This concludes the proof. □

6.2 The symplectic and orthogonal model

We now adapt the branching model developed in section 6.1 to the symplectic and orthogonal cases. We sketch an argument that demonstrates that such a model captures the behaviour of the moments of moments in the 'low temperature' range (large β). The work described in sections 6.2.1 and 6.2.2 came out of conversations with Prof. Paul Bourgade during a research visit to CIMS.

6.2.1 The symplectic case

Recall that, due to the rotational symmetry in the unitary case, the moments

$$\mathbb{E}\left[|P_{U(N)}(A,\theta)|^{2\beta}\right] \tag{6.2.1}$$

are independent of θ. This no longer holds for the symplectic (nor the orthogonal) case. As discussed in chapter 2, Keating and Snaith used the calculation of such moments in [116] to show that

$$\mathbb{E}\left[\log|P_{U(N)}(A,\theta)|\right] = 0 \tag{6.2.2}$$

$$\mathbb{E}\left[(\log|P_{U(N)}(A,\theta)|)^2\right] = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} + O\left(\frac{1}{N^4}\right), \tag{6.2.3}$$

where γ is the Euler-Mascheroni constant. Further, the value distribution of $\log|P_{U(N)}(A,\theta)|$ in the limit $N\to\infty$, is Gaussian with mean and variance given by (6.2.2) and the large N limit of (6.2.3) respectively,

$$\frac{\log|P_{U(N)}(A,\theta)|}{\sqrt{\frac{1}{2}\log N}} \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0,1). \tag{6.2.4}$$

As remarked above, for $A\in Sp(2N)$, the moments of $|P_{Sp(2N)}(A,\theta)|$ are no longer independent of θ. At the symmetry point, $\theta=0$, Keating and Snaith calculated [115] using the Selberg integral

$$\mathbb{E}\left[|P_{Sp(2N)}(A,0)|^{2\beta}\right] = 2^{4\beta N}\prod_{j=1}^{N}\frac{\Gamma(N+j+1)\Gamma(2\beta+j+1/2)}{\Gamma(j+1/2)\Gamma(N+2\beta+1+j)} \tag{6.2.5}$$

$$\sim 2^{2\beta^2}\frac{G(2\beta+1)\sqrt{\Gamma(2\beta+1)}}{\sqrt{G(4\beta+1)\Gamma(4\beta+1)}}N^{\beta(2\beta+1)} \quad \text{as } N\to\infty. \tag{6.2.6}$$

Further, they use (6.2.5) to show that $\log|P_{Sp(2N)}(A,0)|$ tends to a Gaussian random variable as N grows with the following mean and variance

$$\mathbb{E}\left[\log|P_{Sp(2N)}(A,0)|\right] \sim \frac{1}{2}\log N, \tag{6.2.7}$$

$$\mathbb{E}\left[(\log|P_{Sp(2N)}(A,0)|)^2\right] \sim \log N, \tag{6.2.8}$$

as $N\to\infty$. So

$$\frac{\log|P_{Sp(2N)}(A,0)| - \frac{1}{2}\log N}{\sqrt{\log N}} \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0,1). \tag{6.2.9}$$

We now wish to use the model developed in section 6.1, specialized to the symplectic case, to recover the analogue of theorem 4.1.1. Mimicking the procedure outlined in section 6.1, with the usual identification $N=2^n$

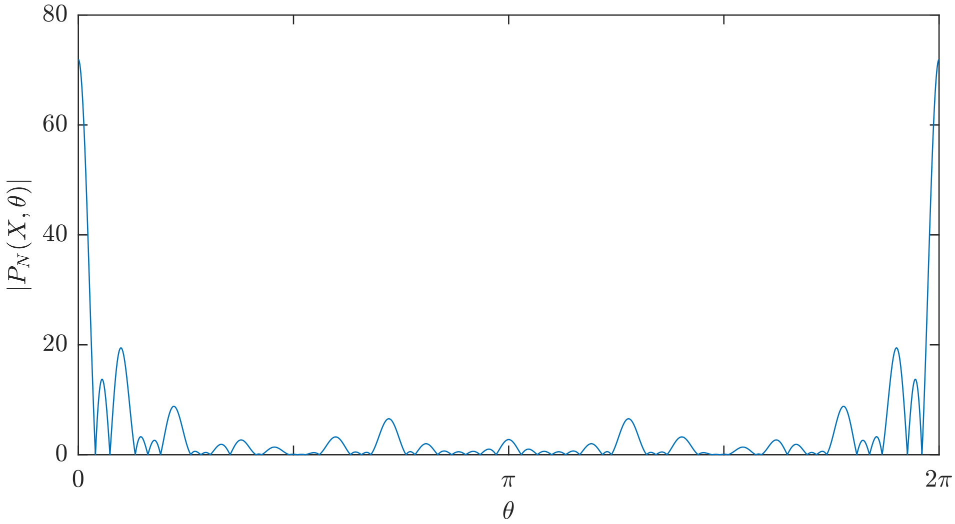


Figure 6.4: Plot of $|P_N(X,\theta)| \equiv |P_{Sp(2N)}(X,\theta)|$, for $X \in Sp(50)$ and $\theta \in [0, 2\pi)$.

$$\mathrm{MoM}_{Sp(2N)}(k,\beta) := \mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} |P_{Sp(2N)}(A,\theta)|^{2\beta} d\theta\right)^k\right] \tag{6.2.10}$$

$$\leftrightsquigarrow \frac{1}{N^k}\mathbb{E}\left[\left(\sum_{j=1}^{N} e^{2\beta \log|P_{Sp(2N)}(A,\phi_j)|}\right)^k\right]. \tag{6.2.11}$$

The exponent in (6.2.11) will be dominated by the value at the symmetry point, since away from $\theta = n\pi$ for $n \in \mathbb{N}$, the statistics revert to the unitary case (see [112] and figure 6.4), which from what will follow is sub-leading. Taking ϕ^* to be the value of $\phi_1, \dots, \phi_N$ closest to this point, we thus compare

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} |P_{Sp(2N)}(A,\theta)|^{2\beta} d\theta\right)^k\right] \leftrightsquigarrow \frac{1}{N^k}\mathbb{E}\left[e^{2k\beta \log|P_{Sp(2N)}(A,\phi^*)|}\right]. \tag{6.2.12}$$

We now use (6.2.9) (ignoring any issues of speed of convergence for the sake of this calculation) to find

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi} |P_{Sp(2N)}(A,\theta)|^{2\beta} d\theta\right)^k\right] \sim \frac{1}{N^k}\mathbb{E}\left[e^{2k\beta Z}\right], \tag{6.2.13}$$

where $(\log N)^{-1/2}(Z - (1/2)\log N)$ is a standard Gaussian random variable. Hence, the model behaves like

$$\frac{1}{N^k}\mathbb{E}\left[e^{2k\beta Z}\right] = \frac{1}{N^k}\mathbb{E}\left[e^{k\beta \log N + 2k^2\beta^2 \log N}\right] \tag{6.2.14}$$

$$= N^{k\beta(2k\beta+1)-k}. \tag{6.2.15}$$

Note that approach is more crude than that described in section 6.1 since the goal here is just to demonstrate the size of the leading order. Regardless, this argument shows that the leading order of the branching model for the symplectic moments of moments matches theorem 4.1.1, where recall we showed in chapter 4 that for $k, \beta \in \mathbb{N}$

$$\lim_{N\to\infty} \frac{\mathrm{MoM}_{Sp(2N)}(k,\beta)}{N^{k\beta(2k\beta+1)-k}} = c_{k,\beta} \tag{6.2.16}$$

for some constant $c_{k,\beta}$ depending on the moment parameters k, β.

6.2.2 The orthogonal case

For completeness, within this section, we repeat the calculation of section 6.2.1 for $A \in SO(2N)$. As was noted in section 6.2.1, the moments of $P_{SO(2N)}(A,\theta)$ for $A \in SO(2N)$ are dependent on θ. At the symmetry point $\theta = 0$, Keating and Snaith [115] calculated (see also theorem 2.1.8).

$$\mathbb{E}\left[|P_{SO(2N)}(A,0)|^{2\beta}\right] = 2^{4\beta N}\prod_{j=1}^{N}\frac{\Gamma(N+j-1)\Gamma(2\beta+j-1/2)}{\Gamma(j-1/2)\Gamma(N+2\beta+j-1)} \tag{6.2.17}$$

$$\sim 2^{2\beta^2}\frac{G(2\beta+1)\sqrt{\Gamma(4\beta+1)}}{\sqrt{G(4\beta+1)\Gamma(2\beta+1)}}N^{\beta(2\beta-1)} \quad \text{as } N\to\infty. \tag{6.2.18}$$

Additionally, Keating and Snaith establish a central limit theorem for $\log|P_{SO(2N)}(A,0)|$ using these moments (see (2.2.124)). For large N, the mean and variance are

$$\mathbb{E}\left[\log|P_{SO(2N)}(A,0)|\right] \sim -\frac{1}{2}\log N, \tag{6.2.19}$$

$$\mathbb{E}\left[(\log|P_{SO(2N)}(A,0)|)^2\right] \sim \log N. \tag{6.2.20}$$

Hence, the limiting distribution of $\log|P_{SO(2N)}(A,0)|$ is

$$\frac{\log|P_{SO(2N)}(A,0)| + \frac{1}{2}\log N}{\sqrt{\log N}} \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0,1). \tag{6.2.21}$$

Recreating the sketch of the model given in section 6.2.2 (see (6.2.11)–(6.2.15)), we have

$$\mathbb{E}\left[\left(\frac{1}{2\pi}\int_0^{2\pi}|P_{SO(2N)}(A,\theta)|^{2\beta}d\theta\right)^k\right] \leftrightsquigarrow \frac{1}{N^k}\mathbb{E}\left[\left(\sum_{j=1}^{N}e^{2\beta\log|P_{SO(2N)}(A,\phi_j)|}\right)^k\right] \tag{6.2.22}$$

$$\sim \frac{1}{N^k}\mathbb{E}\left[e^{2k\beta\log|P_{SO(2N)}(A,\phi^*)|}\right] \tag{6.2.23}$$

$$\sim \frac{1}{N^k}\mathbb{E}\left[e^{2k\beta Z}\right], \tag{6.2.24}$$

where again ϕ^* is the choice of $\phi_1,\dots,\phi_N$ which lies closest to the symmetry point, and in (6.2.24) $(\log N)^{-1/2}(Z + (1/2)\log N)$ is a standard Gaussian random variable. Thus the branching model for the orthogonal moments of moments behaves like

$$\frac{1}{N^k}\mathbb{E}\left[e^{2k\beta Z}\right] = \frac{1}{N^k}\mathbb{E}\left[e^{-k\beta\log N + 2k^2\beta^2\log N}\right] \tag{6.2.25}$$

$$= N^{k\beta(2k\beta-1)-k}. \tag{6.2.26}$$

Recall for comparison, theorem 4.1.2, proved for $k,\beta \in \mathbb{N}$ in chapter 4,

$$\lim_{N\to\infty}\frac{\mathrm{MoM}_{SO(2N)}(k,\beta)}{N^{k\beta(2k\beta-1)-k}} = c_{k,\beta} \tag{6.2.27}$$

for some coefficient $c_{k,\beta}$ depending only on k,β (and different to the leading order coefficient in the symplectic case). Thus, the branching model matches[2] theorem 4.1.2.

[2]Except for the case of $k = \beta = 1$ which had to be dealt with separately too in theorem 4.1.2, so it is unsurprising that the rough approach described here would need to be adapted in this single case.

Chapter 7

Mixed moments and moments of logarithmic derivatives

This chapter is based on the paper '*Mixed moments of characteristic polynomials of random unitary matrices*', published in *Journal of Mathematical Physics*. This paper was co-authored by the present author, S. Bettin, G. Blower, J. B. Conrey, A. Prokhorov, M. O. Rubinstein, and N. C. Snaith [14]. The project originated at the American Institute of Mathematics workshop on *Painlevé Equations and Their Applications*. Whilst the mathematical contribution is shared between all co-authors, the majority of the manuscript was written by the present author, J. B. Conrey, N. C. Snaith, and M. O. Rubinstein, with the specific exception of the paper section 7.7 (which we have not included in this chapter). This latter section, pertaining to a Riemann–Hilbert interpretation of our theorem 7.1.2, was lead by G. Blower and A. Prokhorov.

The present author's mathematical contribution to the project was (in collaboration) theorem 7.1.1 and its proof, as well as (again, in collaboration with co-authors) theorem 7.1.2, in particular the statements and proof of lemma 7.3.1 and lemma 7.3.3. Within section 7.7 of [14], an interpretation of the proof of theorem 7.1.2 in terms of Riemann–Hilbert analysis was presented. Whilst this section contained no mathematical contribution from the author (and so does not feature here) we do include and discuss the interesting general connection between moments involving both characteristic polynomials and their derivatives, and Painlevé transcendents that it facilitated.

The majority of the rest of the chapter is verbatim from [14]. However, we have changed notation to be in-keeping with that of the present document. In particular, the results of [14] are stated in terms of $\Lambda_X(s) = \det(I - X^* s)$ for some $X \in U(N)$, and an associated 'Z' function (see (7.1.6)). Clearly $\Lambda_X(s)$ is very closely related to $P_N(X, \theta)$, the notation for characteristic polynomial preferred throughout this thesis. To avoid confusion, we have translated the statements so that they are in terms of P_N, and its, different, associated 'Z' function. Furthermore, the arguments and explanations have been expanded for clarity, various illuminating examples have been added, and the introduction has been updated to reflect some recent developments. All such changes and inclusions are due to the present author.

7.1 Introduction

Throughout this chapter, A will denote a unitary matrix in $U(N)$. As usual, we write $P_N \equiv P_{U(N)}$ and

$$P_N(A, \theta) = \det(I - Ae^{-i\theta}) = \prod_{j=1}^{N} (1 - e^{i(\theta_j - \theta)}) \tag{7.1.1}$$

for its characteristic polynomial. We remark that in [14] (the paper that this chapter is based on), the results are started in terms of $\Lambda_A(s)$, which is the characteristic polynomial for A^* (the conjugate transpose of A), evaluated at s, so explicitly, $\Lambda_A(s) = \det(I - A^* s)$ and

$$\Lambda_A(e^{-i\theta}) = P_N(A^*, \theta). \tag{7.1.2}$$

We thus have translated the results within this chapter so to be consistent with the notation used throughout this thesis.

As A is unitary, we have the functional equation

$$P_N(A, \theta) = (-1)^N e^{-iN\theta} \det(A) P_N(A^*, -\theta), \tag{7.1.3}$$

see also chapter 3, (3.1.18). Hence, we can define

$$Z_A(\theta) := e^{-i\pi \frac{N}{2}} e^{\frac{i\theta N}{2}} \sqrt{\det(A^*)} P_N(A, \theta) \tag{7.1.4}$$

$$= e^{\frac{iN}{2}(\theta+\pi)} e^{-\frac{i}{2}\sum_{j=1}^N \theta_j} P_N(A, \theta), \tag{7.1.5}$$

so that

$$Z_A(\theta) = (-1)^N Z_{A^*}(-\theta). \tag{7.1.6}$$

This definition makes $Z_A(\theta)$ real for $\theta \in \mathbb{R}$. Recall from chapter 2, section 2.1.4 that this is a useful definition for evaluating mixed moments.

We present our results for this chapter in section 7.1.1. Our theorems naturally imply conjectures for mixed moments and log-derivative moments of $\zeta(1/2 + it)$, these are given in section 7.1.2. We consequently place them in the wider research context in section 7.1.3 and state various theorems that will be useful when proving our results in section 7.1.4.

7.1.1 Results for mixed moments and log-derivative moments

In this paper we will prove the following theorems. The first theorem that we prove expresses the mixed moments of Z_A and Z_A' in terms of derivatives of a determinant involving the modified Bessel function of the first kind I_α. Recall that the modified Bessel's equation is

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - (x^2 + \alpha^2)y = 0 \tag{7.1.7}$$

for some $\alpha \in \mathbb{C}$, and hence $I_\alpha(x)$ is one of the two solutions to (7.1.7)

$$y(x) = \gamma I_\alpha(x) + \delta K_\alpha(x) \tag{7.1.8}$$

for $x \geq 0$ and constants γ, δ. $K_\alpha(x)$ is the modified Bessel function of the second kind. For integer n, I_n can be expressed as the following integral,

$$I_n(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x\cos\theta + in\theta} d\theta. \tag{7.1.9}$$

Our theorems are then as follows (and are proved in section 7.2 and section 7.3 respectively). Note the subtle difference in the moments compared to Hughes' result, given in section 7.1.3 (see also chapter 2, section 2.1.4).

Theorem 7.1.1. *For k, h integers with $2k \geq 2h \geq 0$, we have*

$$\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA$$
$$= (-1)^{\frac{k(k-1)}{2}+k-h} N^{k^2+2k-2h}$$
$$\times \left(\frac{d}{dx}\right)^{2k-2h} \left(e^{-\frac{x}{2}} x^{-\frac{k^2}{2}} \det_{k\times k}\left(I_{i+j-1}(2\sqrt{x})\right)\right) \left(1 + O\left(\frac{1}{N}\right)\right)\Bigg|_{x=0}. \quad (7.1.10)$$

This can further be written in terms of a solution to a Painlevé equation, as expressed in (7.2.40)*. A brief overview of Painlevé equations can be found in section 7.1.3.*

Our second theorem gives the leading asymptotics of the moments of the logarithmic derivative of P_N at a point approaching the unit circle, so we write

$$P_N(A, i\alpha) = \det(I - Ae^{\alpha}) \quad (7.1.11)$$

and we will consider α decreasing on the scale of mean eigenvalue spacing. Additionally,

$$\frac{f'}{f}(x) \quad (7.1.12)$$

will denote the logarithmic derivative of f evaluated at x.

Theorem 7.1.2. *Let* $\mathrm{Re}(\alpha) > 0$ *and* $k \in \mathbb{N}$*,*

$$\int_{U(N)} \left|\frac{P_N'}{P_N}(A, i\alpha)\right|^{2k} dA = \binom{2k-2}{k-1} \frac{N^{2k}}{(2a)^{2k-1}}(1 + O(a)), \quad (7.1.13)$$

where $\alpha = a/N$ and $a = o(1)$ as $N \to \infty$ (so that α depends on N).

These two theorems lead us immediately to conjectures about mixed moments for the Riemann zeta-function, which we give in the following section.

7.1.2 Conjectures for mixed moments and log-derivative moments of $\zeta(1/2 + it)$

Recall that the Riemann zeta-function can be defined by the Dirichlet series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (7.1.14)$$

for $\mathrm{Re}(s) > 1$, and by its meromorphic continuation otherwise. Additionally, by defining

$$\xi(s) := \frac{s(s-1)}{2} \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) \quad (7.1.15)$$

one has the functional equation

$$\xi(s) = \xi(1-s) \quad (7.1.16)$$

where $\xi(s)$ is entire. Therefore Hardy's function

$$Z(t) := \frac{\pi^{-\frac{it}{2}} \Gamma\left(\frac{1}{4} + \frac{it}{2}\right)}{|\Gamma\left(\frac{1}{4} + \frac{it}{2}\right)|} \zeta(\tfrac{1}{2} + it) \quad (7.1.17)$$

is real for real t and satisfies $|Z(t)| = |\zeta(1/2+it)|$. Our theorem 7.1.1 involves Z_A which is the random matrix analogue of Hardy's $Z(t)$ function, and theorem 7.1.2 involves P_N which is the random matrix analogue of ζ. Recall that the conjecture of Keating and Snaith [116], see also theorem 2.1.1, about moments of the Riemann zeta-function may be written as

$$\frac{1}{T}\int_0^T |Z(t)|^{2k} dt \sim \prod_{j=0}^{k-1} \frac{j!}{(j+k)!} a_k (\log T)^{k^2} \tag{7.1.18}$$

for $k \in \mathbb{N}$ and for a certain arithmetic constant a_k, see (2.1.32).

After the work of Hughes [98] and Conrey, Rubinstein and Snaith [54] we expect that the $2k$th moment of $|Z'(t)|$ involves the same arithmetic constant a_k multiplied by a (rational number) geometric factor and $(\log T)^{k^2+2k}$. These ideas translate to a conjecture for the mixed moments we are considering here. We can express our conjecture as follows.

Conjecture 7.1.3. *For non-negative integers k and h with $h \leq k$ we conjecture that as $T \to \infty$,*

$$\frac{\int_0^T \left|\frac{Z'}{Z}(t)\right|^{2k-2h} |Z(t)|^{2k} dt}{\int_0^T |Z(t)|^{2k} dt} \sim (i \log T)^{2k-2k} \left(\frac{d}{dx}\right)^{2k-2h} \exp\left(\frac{x}{2} - \int_0^{4x} (\sigma_{\mathrm{III}'}(s) + k^2)\frac{ds}{s}\right)\Bigg|_{x=0} \tag{7.1.19}$$

where $\sigma_{\mathrm{III}'}$ is defined in (7.2.38).

In this formulation our conjecture appears as an average of $|Z'/Z|^{2k-2h}$ measured against $|Z|^{2k}$. Notice that the arithmetic factors cancel out, as well as the ratio of the product of factorials.

The analogue of theorem 7.1.2 is best expressed in terms of moments of the logarithmic derivative of ζ. In the work of Conrey and Snaith [55] on the n-correlation of the zeros of the Riemann zeta function, the authors use the "recipe"[1] to find a conjectural formula for the average over t of the product of any number of factors of the form

$$\frac{\zeta'}{\zeta}(\tfrac{1}{2} \pm it + \alpha) \tag{7.1.20}$$

with different values of α. In this work we are focussed on the behaviour of the absolute value of such a product when all of the α are the same and $\alpha \to 0^+$.

Conjecture 7.1.4. *For any positive integer k and $a > 0$ we conjecture that*

$$\lim_{T\to\infty} \frac{1}{T(\log T)^{2k}} \int_0^T \left|\frac{\zeta'}{\zeta}(\tfrac{1}{2} \pm it + \alpha)\right|^{2k} dt = c(a) \tag{7.1.21}$$

where $\alpha = a/\log T$ and $a = o(1)$ as $T \to \infty$ (so that α depends on T) and

$$\lim_{a\to 0^+} c(a)(2a)^{2k-1} = \binom{2k-2}{k-1}. \tag{7.1.22}$$

7.1.3 Moments involving P_N and P_N', and Painlevé equations

Moments of characteristic polynomials and their derivatives have been investigated in several recent papers on random matrix theory, (see chapter 2, section 2.1). Part of the interest in these calculations precisely is the subject of the preceding section: they permit an apparent connection with the corresponding averages of L-functions. As evidenced in chapter 2, section 2.1.3 and [2, 18, 42, 54, 55, 58, 66, 99, 123, 125, 160], exploiting this link has been extremely profitable.

[1]The "recipe" is a method for conjecturing moments of L-functions first introduced in [46], see section 2.1 in that paper.

Additionally, it was shown by Forrester and Witte [77] that the leading order coefficient for moments of derivatives of unitary characteristic polynomials, derived by Conrey, Rubinstein and Snaith [54] is related to the solution of a version of the Painlevé III$'$ differential equation. Since the connection between random matrix moments and Painlevé equations features prominently within this chapter, we here give some background to the theory. For a thorough overview, see [90, 105].

Take, for example, the first order differential equation

$$\frac{dy}{dt} = y^2. \tag{7.1.23}$$

The general solution to (7.1.23) is

$$y = \frac{1}{c-t} \tag{7.1.24}$$

for some constant c. A key property to note is that the singularity in the solution is dependent on the initial conditions (since they determine c). One calls such a singularity *movable*. Compare this to the solution to

$$\frac{dy}{dt} = \frac{1}{3y^2}, \tag{7.1.25}$$

which is

$$y = (t-c)^{\frac{1}{3}}, \tag{7.1.26}$$

where the singularity is now a *moveable essential* singularity: a branch point. A natural question would be to classify all first order differential equations of the form

$$P\left(\frac{dy}{dt}, y, t\right) = 0, \tag{7.1.27}$$

(where P is a polynomial in the arguments y, $\frac{dy}{dt}$ with coefficients meromorphic in t) such that their solutions are free of moveable essential singularities. Such a classification is due to Poincaré and Fuchs, see for example [105]. Up to changes of variables and simple transformations, there are only two families of first order differential equations free of essential moveable singularities.

Around the turn of the 20^{th} century, Painlevé and Gambier studied the extension to second order differential equations, i.e. equations of the form

$$\frac{d^2y}{dt^2} = R\left(\frac{dy}{dt}, y, t\right) \tag{7.1.28}$$

where R is a rational function in all its arguments. Their classification, see [86, 87, 101, 130, 131], shows that either one reduces to a simpler case (to the first order case detailed above, or to a linear differential equation) or the equation takes a special form. This latter case yielded 6 new non-linear differential equations: the Painlevé equations. They are as follows.

$$\mathrm{P_I}: \quad \frac{d^2y}{dt^2} = 6y^2 + t \tag{7.1.29}$$

$$\mathrm{P_{II}}: \quad \frac{d^2y}{dt^2} = 2y^3 + ty + \alpha \tag{7.1.30}$$

$$\mathrm{P_{III}}: \quad \frac{d^2y}{dt^2} = \frac{1}{y}\left(\frac{dy}{dt}\right)^2 - \frac{1}{t}\frac{dy}{dt} + \frac{\alpha y^2 + \beta}{t} + \gamma y^3 + \frac{\delta}{y} \tag{7.1.31}$$

$$\mathrm{P_{IV}}: \quad \frac{d^2y}{dt^2} = \frac{1}{2y}\left(\frac{dy}{dt}\right)^2 + \frac{3}{2}y^3 + 4ty^2 + 2(t^2-\alpha)y + \frac{\beta}{y} \tag{7.1.32}$$

$$\mathrm{P_V}: \quad \frac{d^2y}{dt^2} = \left(\frac{1}{2y} + \frac{1}{y-1}\right)\left(\frac{dy}{dt}\right)^2 - \frac{1}{t}\frac{dy}{dt} + \frac{(y-1)^2}{t^2}\left(\alpha y + \frac{\beta}{y}\right) + \frac{\gamma y}{t} + \frac{\delta y(y+1)}{y-1} \tag{7.1.33}$$

$$\mathrm{P_{VI}}: \quad \frac{d^2y}{dt^2} = \frac{1}{2}\left(\frac{1}{y} + \frac{1}{y-1} + \frac{1}{y-t}\right)\left(\frac{dy}{dt}\right)^2 - \left(\frac{1}{t} + \frac{1}{t-1} + \frac{1}{y-t}\right)\frac{dy}{dt} \tag{7.1.34}$$

$$+ \frac{y(y-1)(y-t)}{t^2(t-1)^2}\left(\alpha + \frac{\beta t}{y^2} + \frac{\gamma(t-1)}{(y-1)^2} + \frac{\delta t(t-1)}{(y-t)^2}\right), \tag{7.1.35}$$

where α, β, γ, and δ are arbitrary constants. It turns out that by rescaling, $\mathrm{P_{III}}$ may be written with just 2 independent parameters, and $\mathrm{P_V}$ with 3 independent parameters.

One connection between Painlevé theory and the study of random matrix moments is the aforementioned work of Forrester and Witte [77] which connects the moments of the derivative of unitary characteristic polynomials to $\mathrm{P_{III}}$. Conrey, Rubinstein, Snaith [54] determined that the moments of the derivative are asymptotically

$$\int_{U(N)} |P_N'(A^*, 0)|^{2\beta} dA \sim c_\beta N^{\beta^2+2\beta}, \tag{7.1.36}$$

where

$$c_\beta = (-1)^{\frac{\beta(\beta+1)}{2}} \sum_{m=0}^{\beta} \binom{\beta}{m} \left(\frac{d}{dx}\right)^{k+h} \left(e^{-x} x^{-\frac{\beta^2}{2}} \det_{\beta\times\beta}(I_{i+j-1}(2\sqrt{x}))\right)\Bigg|_{x=0}, \tag{7.1.37}$$

and $I_\nu(z)$ denotes the modified Bessel function of the first kind, see (7.1.8). Forrester and Witte [76] show that the determinant present in (7.1.37) can expressed as a function involving a solution to a particular Painlevé equation. Explicitly, they show [77] that

$$c_\beta = \frac{(-1)^\beta}{A(\beta,\beta)} \sum_{m=0}^{\beta} \binom{\beta}{m} \left(\frac{d}{dx}\right)^{k+h} \exp\left(-\int_0^{4x} (\sigma_{\mathrm{III}'}(s) + \beta^2)\frac{ds}{s}\right)\Bigg|_{x=0}, \tag{7.1.38}$$

where

$$A(n,m) := \prod_{j=1}^{n} \frac{\Gamma(n+1)\Gamma(m+j)}{\Gamma(j+1)} \tag{7.1.39}$$

and $\sigma_{\mathrm{III}'}(s)$ satisfies the particular Painlevé III$'$ equation (the Jimbo–Miwa–Okamoto 'σ-form')

$$(s\,\sigma_{\mathrm{III}'}'')^2 + \sigma_{\mathrm{III}'}'(4\,\sigma_{\mathrm{III}'}' - 1)(\sigma_{\mathrm{III}'} - s\,\sigma_{\mathrm{III}'}') - \frac{\beta^2}{16} = 0, \tag{7.1.40}$$

satisfying the boundary condition

$$\sigma_{\mathrm{III}'}(s) \overset{s\to 0}{\sim} -\beta^2 + \frac{s}{8} + O(s^2), \tag{7.1.41}$$

for $\beta \in \mathbb{N}$.

Using the techniques of [54] we determined a similar relation, theorem 7.1.1. Instead of just averages of moments of $|P_N|$ or $|P_N'|$ separately, we consider mixed moments featuring both the characteristic polynomial and its derivative. This theorem is proved in section 7.2. Subsequently this result also featured in work by a group[2] [18] (see their equation (5)–(79)), although they use different methods, allowing them to extend the result to finite N. In particular, they express the mixed moment in (7.1.10) in terms of a determinant of generalized Laguerre polynomials, and hence use a connection between this determinant and a special case of $\mathrm{P_V}$. By examining the large N limit of the Laguerre determinant, they arrive at the same statement as theorem 7.1.1 including the connection to Painlevé III'.

In section 7.3 we turn to the moments of the logarithmic derivative of the characteristic polynomial.

[2]The work of this group also began at the American Institute of Mathematics workshop.

An exact formula for these moments averaged over $U(N)$ is presented in [55], see theorem 7.1.8, but the asymptotics when N is large and the characteristic polynomials are evaluated close to the unit circle are not easy to extract from that result. However, adapting the method of section 7.2 allows us to work out the leading order term. In section 7.3.3 we compare this with the exact result in a couple of simple cases.

We now summarize some related work on moments involving both the characteristic polynomial and its derivative. We start with a result from the thesis of Chris Hughes [98], recall also chapter 2, section 2.1.4.

Hughes considers the quantity

$$\tilde{F}_N(h,k) := \int_{U(N)} |Z_A(0)|^{2k-2h} |Z'_A(0)|^{2h} dA \tag{7.1.42}$$

where the average is over Haar measure on the unitary group, and Z_A is as defined by (7.1.5). A minor notational warning: Hughes uses $Z_A(\theta)$ for his characteristic polynomial of A evaluated at $e^{-i\theta}$ (i.e. our $P_N(A,\theta)$), and $V_A(\theta)$ for the 'normalized' version (i.e. our $Z_A(\theta)$). His results have been translated in to our notation.

Hughes shows that for integer h,k

$$\lim_{N\to\infty} \frac{1}{N^{k^2+2h}} \tilde{F}_N(h,k) =: \tilde{F}(h,k), \tag{7.1.43}$$

where $\tilde{F}(h,k)$ is given as an expression that can be analytically continued k for $\text{Re}(k) > h - 1/2$. However, the method forces h to be an integer. Note that (7.1.43) is consistent with the statement of theorem 7.1.1 if one manipulates the moments parameters accordingly.

By computing some specific examples, Hughes suggests that for a given integer h, $\tilde{F}(h,k)$ has the form of a rational function of k multiplied by a ratio of Barnes $\mathcal{G}$-functions. As shown in chapter 2, section 2.1.4, Dehaye [58] proved this form for $\tilde{F}(h,k)$, and gave further information about the structure of the rational function of k, but still always for integer h. Winn [160] has given the only example we know of where the exponent on the derivative is not an even integer, by writing down an explicit formula $\tilde{F}_N(h,k)$ when $h = (2m-1)/2$ for $m \in \mathbb{N}$ (i.e. a complementary formula to that of Dehaye for half-integer h), see theorem 2.1.14.

The asymptotics of a similar mixed moment, with just a first power on the derivative of the characteristic polynomial, but for non-integer powers on the characteristic polynomial itself, has been studied in the thesis of Ian Cooper [56] when the average is over the classical compact groups $SO(2N)$ and $Sp(2N)$.

Note, there is interest in allowing the power on the derivatives of the characteristic polynomial to be non-integer, but this appears to be a difficult problem.

Finally, we emphasise once more that Conrey and Snaith obtained a formula (theorem 7.1.8) for the logarithmic derivative of P_N evaluated close to the unit circle. However, our alternative approach allows us to extract the leading order asymptotic in N. The comparative difficulty of both approaches is the subject of section 7.3.3. Recently, Alvarez and Snaith calculated the asymptotics for the log-derivatives of characteristic polynomials of symplectic and orthogonal matrices [3]. They extended the method of section 7.3 and proved the following.

Theorem 7.1.5 (Alvarez and Snaith [3]). *Let $k \in \mathbb{N}$ and $\alpha = a/N$ where $a = o(1)$ as $N \to \infty$ and $\text{Re}(a) > 0$. As $N \to \infty$, for $k \geq 4$,*

$$\int_{Sp(2N)} \left(\frac{P'_N}{P_N}(A, i\alpha) \right)^k dA = (-1)^k \frac{2}{3} \frac{N^k}{a^{k-3}} \frac{(2k-5)!!}{(k-1)!} (1 + O(a)), \tag{7.1.44}$$

and as $N \to \infty$, for $k \geq 2$

$$\int_{SO(2N)} \left(\frac{P_N'}{P_N}(A, i\alpha)\right)^k dA = (-1)^k \frac{2N^k}{a^{k-1}} \frac{(2k-3)!!}{(k-1)!}(1 + O(a)), \tag{7.1.45}$$

where as usual $n!! = n(n-2)(n-4)\cdots(n-(2\lceil \frac{n}{2} \rceil - 2))$.

Alvarez and Snaith also compute the low moments excluded from theorem 7.1.5, as well as log-derivative moments for $SO(2N+1)$.

7.1.4 Mixed moments, log-derivative moments, and the Ratios Formulae

One starting point for averages of characteristic polynomials and derivatives are the Ratios Formulae of Conrey, Farmer and Zirnbauer [48, 49]. Recall that we have already encountered one form of the Ratio Theorem within chapters 3, 4, and 5. The form we give in theorem 7.1.7 will be that of Conrey, Farmer and Zirnbauer. The other representation we give first, also due to Conrey, Farmer and Zirnbauer, and takes the shape of a multiple integral contour integral. This could be compared to, for example, the multiple contour integral representation of the average over $U(N)$ used in the proof of theorem 3.1.4, see lemma 3.3.1 in section 3.3.

Before we state the first form of the Ratios Theorem, we remark that throughout we will write

$$\Delta(w_1, \dots, w_n) = \prod_{1 \leq j < k \leq n} (w_k - w_j) = \det\left(w_j^{k-1}\right)_{j,k=1,\dots,n} \tag{7.1.46}$$

as usual for the Vandermonde determinant, and define

$$z(x) := \frac{1}{1 - e^{-x}}. \tag{7.1.47}$$

Then, the first version of the Ratios Theorem is the following.

Theorem 7.1.6 (Conrey, Farmer, Zirnbauer [49])**.** *Let $K, L, N, Q, R \in \mathbb{N}$. Suppose that the matrix size $N \geq \max\{Q - K, R - L\}$, and $\mathrm{Re}(\gamma_q), \mathrm{Re}(\delta_r) > 0$. Then*

$$\begin{aligned}
&\int_{U(N)} \frac{\prod_{j=1}^{K} P_N(A^*, -i\alpha_j) \prod_{l=K+1}^{K+L} P_N(A, i\alpha_l)}{\prod_{q=1}^{Q} P_N(A^*, -i\gamma_q) \prod_{r=1}^{R} P_N(A, -i\delta_r)} dA \\
&\quad = e^{\frac{N}{2}\left(\sum_{l=1}^{L} \alpha_{K+l} - \sum_{k=1}^{K} \alpha_k\right)} \frac{(-1)^{(K+L)(K+L-1)/2}}{K!L!(2\pi i)^{K+L}} \\
&\qquad \times \oint \cdots \oint e^{\frac{N}{2}\left(\sum_{k=1}^{K} w_k - \sum_{l=1}^{L} w_{K+l}\right)} \frac{\prod_{j=1}^{K} \prod_{l=1}^{L} z(w_j - w_{K+l}) \prod_{q=1}^{Q} \prod_{r=1}^{R} z(\gamma_q + \delta_r)}{\prod_{j=1}^{K} \prod_{r=1}^{R} z(w_j + \delta_r) \prod_{l=1}^{L} \prod_{q=1}^{Q} z(-w_{K+l} + \delta_q)} \\
&\qquad\qquad \times \frac{\Delta(w_1, \dots, w_{K+L})^2 \prod_{j=1}^{K+L} dw_j}{\prod_{j=1}^{K+L} \prod_{k=1}^{K+L} (w_k - \alpha_j)},
\end{aligned} \tag{7.1.48}$$

where the w contours enclose the poles at $\alpha_1, \dots, \alpha_{K+L}$. The function $z(x)$ is as defined in (7.1.47)*.*

The quantity in theorem 7.1.6 can also be written as a permutation sum. This will be the second form of the Ratios Theorem.

Theorem 7.1.7 (Conrey, Farmer, Zirnbauer [49], see also [50]). *Let $K, L, N, Q, R \in \mathbb{N}$. Suppose $N \geq \max\{Q-K, R-L\}$ and $\mathrm{Re}(\gamma_q), \mathrm{Re}(\delta_r) > 0$. We have*

$$\int_{U(N)} \frac{\prod_{j=1}^{K} P_N(A^*, -i\alpha_j) \prod_{l=K+1}^{K+L} P_N(A, i\alpha_l)}{\prod_{q=1}^{Q} P_N(A^*, -i\gamma_q) \prod_{r=1}^{R} P_N(A, -i\delta_r)} dA$$

$$= \sum_{\sigma \in \Xi_{K,L}} e^{N\sum_{k=1}^{K}(\alpha_{\sigma(k)} - \alpha_k)} \frac{\prod_{k=1}^{K}\prod_{l=K+1}^{K+L} z(\alpha_{\sigma(k)} - \alpha_{\sigma(l)}) \prod_{q=1}^{Q}\prod_{r=1}^{R} z(\gamma_q + \delta_r)}{\prod_{r=1}^{R}\prod_{k=1}^{K} z(\alpha_{\sigma(k)} + \delta_r) \prod_{q=1}^{Q}\prod_{l=K+1}^{K+L} z(\gamma_q - \alpha_{\sigma(l)})}. \quad (7.1.49)$$

Above, $\Xi_{K,L}$ denotes the set of permutations σ of $\{1, 2, \dots, K+L\}$ such that

$$1 \leq \sigma(1) < \cdots < \sigma(K) \leq K+L \quad (7.1.50)$$

and

$$1 \leq \sigma(K+1) < \cdots < \sigma(K+L) \leq K+L. \quad (7.1.51)$$

We are interested in the moments of logarithmic derivatives of the characteristic polynomial, which can be derived by differentiation of the Ratios Theorem in the form of theorem 7.1.7. Within this chapter, we will focus on the leading order contribution to averages of the logarithmic derivatives when N, the matrix size, is large and the characteristic polynomial is evaluated close to the unit circle.

The exact formula for these moments of the logarithmic derivative was determined by Conrey and Snaith. However, as already noted, it is complicated to work with and extracting the leading order behaviour is difficult (see section 7.3.3). Note that the theorem below uses *set* rather than *permutation* notation for the arguments of the characteristic polynomials. The result is merely a differentiation of theorem 7.1.7 but given the difficulty of keeping track of all the terms, the discussion of the proof in [55] may be useful.

Theorem 7.1.8 (Conrey and Snaith [55]). *Suppose that $A := \{\alpha_j\}$ and $B := \{\beta_j\}$ for some complex numbers α_j, β_j such that $\mathrm{Re}(\alpha_j) > 0$ and $\mathrm{Re}(\beta_j) > 0$. Then*

$$J(A; B) = J^*(A; B) \quad (7.1.52)$$

where

$$J(A; B) := \int_{U(N)} \prod_{\alpha \in A} \left[(-e^{-\alpha}) \frac{P_N'}{P_N}(A^*, -i\alpha)\right] \prod_{\beta \in B} \left[(-e^{-\beta}) \frac{P_N'}{P_N}(A, -i\beta)\right] dA, \quad (7.1.53)$$

$$J^*(A; B) := \sum_{\substack{S \subset A, T \subset B \\ |S| = |T|}} \Bigg(e^{-N(\sum_{\hat{\alpha} \in S} \hat{\alpha} + \sum_{\hat{\beta} \in T} \hat{\beta})}$$

$$\times \frac{Z(S,T) Z(S^-, T^-)}{Z^\dagger(S, S^-) Z^\dagger(T, T^-)} \sum_{\substack{(A-S)+(B-T) \\ = U_1 + \cdots + U_R \\ |U_r| \leq 2}} \prod_{r=1}^{R} H_{S,T}(U_r) \Bigg), \quad (7.1.54)$$

and

$$H_{S,T}(W)=\begin{cases}\displaystyle\sum_{\hat\alpha\in S}\frac{z'(\alpha-\hat\alpha)}{z(\alpha-\hat\alpha)}-\sum_{\hat\beta\in T}\frac{z'(\alpha+\hat\beta)}{z(\alpha+\hat\beta)} & \text{if } W=\{\alpha\}\subset A-S\\ \displaystyle\sum_{\hat\beta\in T}\frac{z'(\beta-\hat\beta)}{z(\beta-\hat\beta)}-\sum_{\hat\alpha\in S}\frac{z'(\beta+\hat\alpha)}{z(\beta+\hat\alpha)} & \text{if } W=\{\beta\}\subset B-T\\ \displaystyle\left(\frac{z'(\alpha+\beta)}{z(\alpha+\beta)}\right)' & \text{if } W=\{\alpha,\beta\} \text{ with } \alpha\in A-S,\beta\in B-T\\ 0 & \text{otherwise.}\end{cases} \tag{7.1.55}$$

Here $z(x)$ is as defined by (7.1.47), $S^-=\{-s:s\in S\}$, $T^-=\{-t:t\in T\}$ and

$$Z(A,B)=\prod_{\substack{\alpha\in A\\ \beta\in B}}z(\alpha+\beta), \tag{7.1.56}$$

with the dagger on $Z^\dagger(S,S^-)$ imposing the additional restriction that a factor $z(x)$ is omitted if its argument is zero.

We will use both theorem 7.1.6 and 7.1.7 in the proofs of theorem 7.1.1 and 7.1.2. In section 7.3.3 we compute the leading order asymptotic behaviour of the logarithmic-derivative moments (i.e. the subject of theorem 7.1.2) using theorem 7.1.8 in order to demonstrate the difficulty in generally computing the leading order using this method.

7.2 Proof of theorem 7.1.1

There are various ways to write moments of the function $Z_A(\theta)$, defined in (7.1.5). For example, there is an expression as a permutation sum:

$$\begin{aligned}\int_{U(N)}\prod_{j=1}^k Z_{A^*}(-i\alpha_j)Z_A(i\alpha_{j+k})dA\\ =(-1)^{Nk}e^{-\frac{N}{2}\sum_{j=1}^{2k}\alpha_j}\sum_{\sigma\in\Xi}e^{N\sum_{j=1}^k\alpha_{\sigma(j)}}\prod_{\substack{1\le i\le k\\ 1\le j\le k}}z(\alpha_{\sigma(i)}-\alpha_{\sigma(j+k)}),\end{aligned} \tag{7.2.1}$$

where

$$z(x)=\frac{1}{1-e^{-x}}=\frac{1}{x}+O(1), \tag{7.2.2}$$

for small x. Recall also that Ξ denotes the subset of permutations $\sigma\in S_{2k}$, the group of permutations of $\{1,\dots,2k\}$ for which

$$\sigma(1)<\sigma(2)<\cdots<\sigma(k) \tag{7.2.3}$$

and

$$\sigma(k+1)<\sigma(k+2)<\cdots<\sigma(2k). \tag{7.2.4}$$

Note that (7.2.1) is just a simple case of theorem 7.1.7, with a different prefactor because we are using Z_A instead of P_N. This can equivalently be written as

$$\begin{aligned}
&\int_{U(N)} \prod_{j=1}^{k} Z_{A^*}(-i\alpha_j) Z_A(i\alpha_{j+K}) dA \\
&= (-1)^{Nk+\frac{k(k-1)}{2}} \frac{e^{-\frac{N}{2}\sum_{j=1}^{2k}\alpha_j}}{k!(2\pi i)^k} \\
&\times \oint \cdots \oint e^{N\sum_{i=1}^{k} w_i} \prod_{\substack{1\le i\le k \\ 1\le j\le 2k}} z(w_i - \alpha_j) \Delta(w_1, \ldots, w_k)^2 dw_1 \cdots dw_k, \qquad (7.2.5)
\end{aligned}$$

where the contours enclose the αs. This follows from an evaluation of residues (7.2.5), which yields the sum (7.2.1). For more explanation of these expressions, see section 3 of [54], which draws on section 2 of [45]. Additionally, (7.2.5) is exactly the type of multiple contour integral considered in chapter 3, section 3.3. The expression (7.2.5) is similar in spirit to theorem 7.1.6 except that in this simpler case the average can be written as a k-fold integral rather than a $k + L$ dimensional integral as in the theorem.

Now, we are interested in the average

$$\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA. \qquad (7.2.6)$$

We set h and k to be integers, with $2h \ge 0$ and $2k \ge 2h$, and follow closely the method of Conrey, Rubinstein, and Snaith [54]. In order to recover an expression for the mixed moment, we will take $2k - 2h$ derivatives of (7.2.5) by applying

$$\prod_{j=1}^{k-h} \frac{d}{d\alpha_j} \frac{d}{d\alpha_{j+k}}, \qquad (7.2.7)$$

and hence evaluate the result at $\alpha_j = 0$. The following equalities will be useful.

$$\frac{d}{d\alpha} Z_A(i\alpha)\Big|_{\alpha=0} = i\frac{d}{dt} Z_A(t)\Big|_{t=0} = iZ_A'(0). \qquad (7.2.8)$$

Additionally,

$$\frac{d}{d\alpha} Z_{A^*}(-i\alpha)\Big|_{\alpha=0} = -i\frac{d}{dt} Z_{A^*}(t)\Big|_{t=0}. \qquad (7.2.9)$$

Now, we may express the derivative in t of Z_{A^*} at 0 in terms of the derivative in t of Z_A at 0 in the following way. Firstly, we have

$$\frac{d}{dt} Z_{A^*}(t)\Big|_{t=0} = e^{-\frac{i\pi N}{2}} \sqrt{\det A} \left(\frac{iN}{2} P_N(A^*, 0) + i \sum_{j=1}^{N} e^{-i\theta_j} \prod_{\substack{l=1 \\ l\ne j}}^{N} (1 - e^{-i\theta_l}) \right). \qquad (7.2.10)$$

Now observe that

$$\overline{\frac{d}{dt}Z_A(t)\Big|_{t=0}} = \overline{e^{-\frac{i\pi N}{2}}\sqrt{\det A^*}\left(\frac{iN}{2}P_N(A,0) + i\sum_{j=1} e^{i\theta_j}\prod_{\substack{l=1\\ l\neq j}}^{N}(1-e^{i\theta_l})\right)} \tag{7.2.11}$$

$$= -e^{\frac{i\pi N}{2}}\sqrt{\det A}\left(\frac{iN}{2}P_N(A^*,0) + i\sum_{j=1} e^{-i\theta_j}\prod_{\substack{l=1\\ l\neq j}}^{N}(1-e^{-i\theta_l})\right). \tag{7.2.12}$$

Hence

$$\frac{d}{d\alpha}Z_{A^*}(-i\alpha)\Big|_{\alpha=0} = i(-1)^N\overline{Z_A'(0)}. \tag{7.2.13}$$

Differentiating the left hand side of (7.2.5) and using (7.2.8) and (7.2.13), we find

$$\int_{U(N)}\prod_{j=1}^{k}\prod_{l=1}^{k-h}\frac{d}{d\alpha_l}\frac{d}{d\alpha_{k+l}}Z_{A^*}(-i\alpha_j)Z_A(i\alpha_{k+j})\Big|_{\alpha_l,\alpha_{k+l}=0}dA$$

$$= \int_{U(N)}\left(\prod_{j=k+1-h}^{k}Z_{A^*}(0)Z_A(0)\right)$$

$$\times\left(\prod_{j=1}^{k-h}\prod_{l=1}^{k-h}\frac{d}{d\alpha_l}\frac{d}{d\alpha_{k+l}}Z_{A^*}(-i\alpha_j)Z_A(i\alpha_{k+j})\Bigg|_{\alpha_l,\alpha_{k+l}=0}\right)dA \tag{7.2.14}$$

$$= \int_{U(N)}(-1)^{Nh}|Z_A(0)|^{2h}\prod_{j=1}^{k-h}\frac{d}{d\alpha_j}\frac{d}{d\alpha_{k+j}}Z_{A^*}(-i\alpha_j)Z_A(i\alpha_{k+j})\Bigg|_{\alpha_j,\alpha_{k+j}=0}dA \tag{7.2.15}$$

$$= \int_{U(N)}(-1)^{Nh}|Z_A(0)|^{2h}\prod_{j=1}^{k-h}\left(-(-1)^N Z_A'(0)\overline{Z_A'(0)}\right)dA \tag{7.2.16}$$

$$= (-1)^{Nk+k-h}\int_{U(N)}|Z_A(0)|^{2h}|Z_A'(0)|^{2k-2h}dA. \tag{7.2.17}$$

Secondly, differentiating the right hand side of (7.2.5), we (trivially) have

$$(-1)^{Nk+\frac{k(k-1)}{2}}\prod_{j=1}^{k-h}\frac{d}{d\alpha_j}\frac{d}{d\alpha_{k+j}}\Bigg[\frac{e^{-\frac{N}{2}\sum_{j=1}^{2k}\alpha_j}}{k!(2\pi i)^k}$$

$$\times\oint\cdots\oint e^{N\sum_{i=1}^{k}w_i}\prod_{\substack{1\le i\le k\\ 1\le j\le 2k}}z(w_i-\alpha_j)\Delta(w_1,\ldots,w_k)^2dw_1\cdots dw_k\Bigg]\Bigg|_{\alpha_1=\cdots=\alpha_{2k}=0}. \tag{7.2.18}$$

Equating (7.2.17) and (7.2.18), we have

$$\int_{U(N)}|Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h}dA$$

$$= (-1)^{\frac{k(k+1)}{2}-h}\prod_{j=1}^{k-h}\frac{d}{d\alpha_j}\frac{d}{d\alpha_{j+k}}\Bigg[\frac{e^{-\frac{N}{2}\sum_{j=1}^{2k}\alpha_j}}{k!(2\pi i)^k}$$

$$\times\oint\cdots\oint e^{N\sum_{i=1}^{k}w_i}\prod_{\substack{1\le i\le k\\ 1\le j\le 2k}}z(w_i-\alpha_j)\Delta^2(w_1,\ldots,w_k)dw_1\cdots dw_k\Bigg]\Bigg|_{\alpha_1=\cdots=\alpha_{2k}=0}. \tag{7.2.19}$$

Let $\alpha_i = a_i/N$ and $w_i \to w_i/N$, then with the use of (7.2.2) one has

$$\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k+1)}{2}-h} N^{2k-2h} N^{k^2} \prod_{j=1}^{k-h} \frac{d}{da_j} \frac{d}{da_{j+k}} \left[\frac{e^{-\frac{1}{2}\sum_{j=1}^{2k} a_j}}{k!(2\pi i)^k} \right. \\
&\qquad \times \left.\left. \oint \cdots \oint e^{\sum_{i=1}^{k} w_i} \frac{\Delta^2(w_1,\ldots,w_k)}{\prod_{\substack{1\le i\le k\\ 1\le j\le 2k}} (w_i - a_j)} \left(1 + O(\tfrac{1}{N})\right) dw_1 \cdots dw_k \right] \right|_{a_1=\cdots=a_{2k}=0}, \qquad (7.2.20)
\end{aligned}$$

where the contours enclose the as.

The aim now is to separate these integrals. We do this by using a series of results from [54]. To start, the following explicit derivative will be useful.

$$\left. \frac{d}{da} \frac{e^{-\frac{a}{2}}}{\prod_{1\le i\le k}(w_i - a)} \right|_{a=0} = \frac{1}{\prod_{i=1}^{k} w_i} \left(\sum_{j=1}^{k} \frac{1}{w_j} - \frac{1}{2} \right). \qquad (7.2.21)$$

Hence, (7.2.20) becomes

$$\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k+1)}{2}-h} N^{k^2+2k-2h} \frac{1}{k!(2\pi i)^k} \\
&\qquad \times \left. \oint \cdots \oint \frac{\Delta^2(w_1,\ldots,w_k) \left(\sum_{j=1}^{k} \frac{1}{w_j} - \frac{1}{2} \right)^{2k-2h} \left(1 + O(\frac{1}{N})\right)}{e^{-\sum_{i=1}^{k} w_i}} \prod_{j=1}^{k} \frac{dw_j}{w_j^{2k}} \right|_{a_1=\cdots=a_{2k}=0}. \qquad (7.2.22)
\end{aligned}$$

Next we allow $\Delta\left(\frac{d}{dL}\right)$ to have the meaning

$$\Delta\left(\frac{d}{dL}\right) \prod_{i=1}^{k} f(L_i) = \prod_{1\le i<j\le k} \left(\frac{d}{dL_j} - \frac{d}{dL_i} \right) \prod_{i=1}^{k} f(L_i). \qquad (7.2.23)$$

Below, we will use lemma 5 of [54]:

$$\left. \Delta^2\left(\frac{d}{dL}\right) \left(\prod_{i=1}^{k} f(L_i) \right) \right|_{L_i=1} = k! \det_{k\times k} \left(f^{(i+j-2)}(1) \right), \qquad (7.2.24)$$

for any sufficiently differentiable function f.

Borrowing a technique from [54], we replace the factor $\exp(w_1 + \cdots + w_k)$ appearing in (7.2.20) with $\exp(\sum L_i w_i)$, and then pull out the Vandermonde determinant squared from the integrand as a differential operator. Differentiating under the integral sign and substituting $L_i = 1$ then recovers the original integral. The advantage in doing so is that it allows us to separate the resulting multidimensional integral.

Thus, we have

$$
\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k+1)}{2}-h} N^{k^2+2k-2h} \frac{\Delta^2\left(\frac{d}{dL}\right)}{k!(2\pi i)^k} \\
&\quad\quad \times \oint \cdots \oint \frac{e^{\sum_{i=1}^k L_i w_i}\left(\sum_{j=1}^k \frac{1}{w_j} - \frac{1}{2}\right)^{2k-2h}}{\prod_{i=1}^k w_i^{2k}} \left(1+O\left(\tfrac{1}{N}\right)\right) dw_1 \cdots dw_k \Bigg|_{L_i=1}.
\end{aligned} \tag{7.2.25}
$$

We now apply the same technique again; we introduce an extra parameter x and the differential operator in the x variable in order to simplify the $(\sum 1/w_j - 1/2)^{2k-2h}$ appearing in the integrand. This fully separates the integral,

$$
\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k+1)}{2}-h} N^{k^2+2k-2h} \frac{\Delta^2\left(\frac{d}{dL}\right)}{k!} \\
&\quad\quad \times \left(\frac{d}{dx}\right)^{2k-2h} \left[e^{-\frac{x}{2}} \prod_{j=1}^k \left(\frac{1}{2\pi i} \oint \frac{e^{L_j w + \frac{x}{w}}}{w^{2k}} dw\right) \left(1+O\left(\tfrac{1}{N}\right)\right)\right]\Bigg|_{L_j=1, x=0}.
\end{aligned} \tag{7.2.26}
$$

Still following [54] we have, from equation (2.11) of that paper,

$$
\frac{1}{2\pi i} \int_{|z|=1} \frac{e^{Lz+\frac{t}{z}}}{z^{2k}} dz = \frac{L^{2k-1} I_{2k-1}(2\sqrt{Lt})}{(Lt)^{k-\frac{1}{2}}} =: f_t(L), \tag{7.2.27}
$$

where I_ν is the I-Bessel function, see (7.1.8). We now use (7.2.24) and write

$$
\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k-1)}{2}+k-h} N^{k^2+2k-2h} \left(\frac{d}{dx}\right)^{2k-2h} \left[e^{-\frac{x}{2}} \det_{k\times k}\left(f_x^{(i+j-2)}(1)\right) \left(1+O\left(\tfrac{1}{N}\right)\right)\right]\Bigg|_{x=0},
\end{aligned} \tag{7.2.28}
$$

where $f_t(L)$ is as defined in (7.2.27). Using (4.15) of [54], namely

$$
f_t^{(j)}(L) = \left(\frac{L}{t}\right)^{\frac{2k-1-j}{2}} I_{2k-1-j}(2\sqrt{Lt}), \tag{7.2.29}
$$

we have

$$
\begin{aligned}
&\int_{U(N)} |Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h} dA \\
&\quad = (-1)^{\frac{k(k-1)}{2}+k-h} N^{k^2+2k-2h} \\
&\quad\quad \times \left(\frac{d}{dx}\right)^{2k-2h} \left[e^{-\frac{x}{2}} \det_{k\times k}\left(\frac{I_{2k+1-(i+j)}(2\sqrt{x})}{\sqrt{x^{2k+1-(i+j)}}}\right) \left(1+O\left(\tfrac{1}{N}\right)\right)\right]\Bigg|_{x=0}.
\end{aligned} \tag{7.2.30}
$$

For example, the determinant for $k=3$ is

$$\det\begin{pmatrix} \frac{I_5(2\sqrt{x})}{\sqrt{x^5}} & \frac{I_4(2\sqrt{x})}{\sqrt{x^4}} & \frac{I_3(2\sqrt{x})}{\sqrt{x^3}} \\ \frac{I_4(2\sqrt{x})}{\sqrt{x^4}} & \frac{I_3(2\sqrt{x})}{\sqrt{x^3}} & \frac{I_2(2\sqrt{x})}{\sqrt{x^2}} \\ \frac{I_3(2\sqrt{x})}{\sqrt{x^3}} & \frac{I_2(2\sqrt{x})}{\sqrt{x^2}} & \frac{I_1(2\sqrt{x})}{\sqrt{x}} \end{pmatrix} = x^{\frac{3}{2}}\cdot x^{\frac{2}{2}}\cdot x^{\frac{1}{2}}\cdot\det\begin{pmatrix} \frac{I_5(2\sqrt{x})}{\sqrt{x^2}} & \frac{I_4(2\sqrt{x})}{\sqrt{x}} & I_3(2\sqrt{x}) \\ \frac{I_4(2\sqrt{x})}{\sqrt{x^2}} & \frac{I_3(2\sqrt{x})}{\sqrt{x}} & I_2(2\sqrt{x}) \\ \frac{I_3(2\sqrt{x})}{\sqrt{x^2}} & \frac{I_2(2\sqrt{x})}{\sqrt{x}} & I_1(2\sqrt{x}) \end{pmatrix} \tag{7.2.31}$$

$$= x^{\frac{6}{2}}\cdot x^{\frac{1}{2}}\cdot x^{\frac{2}{2}}\cdot\det\begin{pmatrix} I_5(2\sqrt{x}) & I_4(2\sqrt{x}) & I_3(2\sqrt{x}) \\ I_4(2\sqrt{x}) & I_3(2\sqrt{x}) & I_2(2\sqrt{x}) \\ I_3(2\sqrt{x}) & I_2(2\sqrt{x}) & I_1(2\sqrt{x}) \end{pmatrix} \tag{7.2.32}$$

$$= x^{\frac{9}{2}}\det\begin{pmatrix} I_1(2\sqrt{x}) & I_2(2\sqrt{x}) & I_3(2\sqrt{x}) \\ I_2(2\sqrt{x}) & I_3(2\sqrt{x}) & I_4(2\sqrt{x}) \\ I_3(2\sqrt{x}) & I_4(2\sqrt{x}) & I_5(2\sqrt{x}) \end{pmatrix}. \tag{7.2.33}$$

Hence, for $k=3$,

$$\det_{3\times 3}\left(f_x^{(i+j-2)}(1)\right) = \det_{3\times 3}\left(\frac{I_{2k+1-(i+j)}(2\sqrt{x})}{\sqrt{x^{2k+1-(i+j)}}}\right) \tag{7.2.34}$$

$$= x^{\frac{9}{2}}\det_{3\times 3}\left(I_{i+j-1}(2\sqrt{x})\right). \tag{7.2.35}$$

It is simple to generalize the argument above for all $k\in\mathbb{N}$, hence

$$\int_{U(N)}|Z_A'(0)|^{2k-2h}|Z_A(0)|^{2h}dA$$
$$= (-1)^{\frac{k(k-1)}{2}+k-h}N^{k^2+2k-2h}$$
$$\times\left(\frac{d}{dx}\right)^{2k-2h}\left(e^{-\frac{x}{2}}x^{\frac{k^2}{2}}\det_{k\times k}\left(I_{i+j-1}(2\sqrt{x})\right)\right)\left(1+O\left(\tfrac{1}{N}\right)\right)\Bigg|_{x=0}, \tag{7.2.36}$$

which is equation (7.1.10) of theorem 7.1.1.

Finally, we notice that the determinant in (7.2.28) is reminiscent of the determinant in (7.1.37). Thus, from [77] we have that

$$\exp\left(-\int_0^{4x}(\sigma_{\mathrm{III}'}(s)+k^2)\frac{ds}{s}\right) = (-1)^{\frac{k(k-1)}{2}}$$
$$\times\prod_{j=0}^{k-1}\frac{(j+k)!}{j!}x^{-\frac{k^2}{2}}e^{-x}\det_{k\times k}\left(I_{i+j-1}(2\sqrt{x})\right), \tag{7.2.37}$$

where $\sigma_{\mathrm{III}'}(s)$ is the solution of the Painlevé equation

$$(s\ \sigma_{\mathrm{III}'}'')^2+\sigma_{\mathrm{III}'}'(4\,\sigma_{\mathrm{III}'}'-1)(\sigma_{\mathrm{III}'}-s\ \sigma_{\mathrm{III}'}')-\frac{k^2}{16}=0, \tag{7.2.38}$$

satisfying the boundary condition

$$\sigma_{\mathrm{III}'}(s)\overset{s\to 0}{\sim}-k^2+\frac{s}{8}+O(s^2), \tag{7.2.39}$$

for $k\in\mathbb{N}$. This means that the average we are looking at is related to the solution of the Painlevé equation in the following manner

$$\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA = (-1)^{k-h} N^{k^2+2k-2h} \prod_{j=0}^{k-1} \frac{j!}{(j+k)!} \\ \times \left(\frac{d}{dx}\right)^{2k-2h} e^{\frac{x}{2}} \exp\left(-\int_0^{4x} (\sigma_{\mathrm{III}'}(s) + k^2) \frac{ds}{s}\right) \left(1 + O\left(\tfrac{1}{N}\right)\right)\Bigg|_{x=0}. \tag{7.2.40}$$

Finally, since

$$\int_{U(N)} |Z_A'(0)|^{2k-2h} |Z_A(0)|^{2h} dA = \int_{U(N)} |P_N'(A,0)|^{2k-2h} |P_N(A,0)|^{2h} dA \tag{7.2.41}$$

it is good to confirm that the result above reduces to the known case proved by Keating and Snaith, see theorem 2.1.1. This is true since for $k = h$, theorem 7.1.1 reads

$$\int_{U(N)} |P_N(A,0)|^{2k} dA = \int_{U(N)} |Z_A(0)|^{2k} dA \sim c_k N^{k^2}, \tag{7.2.42}$$

where

$$c_k := \prod_{j=0}^{k-1} \frac{j!}{(j+k)!}. \tag{7.2.43}$$

This exactly matches the asymptotic form of theorem 2.1.1.

7.3 Proof of theorem 7.1.2

We now turn to proving our result concerning logarithmic derivatives. Recall that theorem 7.1.2 gives the leading order asymptotic behaviour of the moments of the logarithmic derivative of P_N, evaluated near the unit circle. In particular, we will need to take derivatives of $\det(I - Ae^{\alpha}) = P_N(A, i\alpha)$ and $\det(I - A^* e^{-\alpha}) = P_N(A^*, -i\alpha)$ for some $\alpha \in \mathbb{C}$ with $\mathrm{Re}(\alpha) > 0$.

Hence we record some useful equalities. These are similar in flavour to (7.2.8) and (7.2.13) used in the proof of theorem 7.1.2.

$$\frac{d}{d\alpha} P_N(A^*, -i\alpha) = \frac{d}{d\alpha} \det(I - A^* e^{-\alpha}) = -e^{-\alpha} P_N'(A^*, -i\alpha), \tag{7.3.1}$$

and

$$\frac{d}{d\alpha} P_N(A, i\alpha) = \frac{d}{d\alpha} \det(I - Ae^{\alpha}) = e^{\alpha} P_N'(A, i\alpha). \tag{7.3.2}$$

We will use theorem 7.1.6 to evaluate the moments of the logarithmic derivative. Recall the statement of the theorem. Take $K, L, N, Q, R \in \mathbb{N}$ and provided that $N \geq \max\{Q - K, R - L\}$, and

$\mathrm{Re}(\gamma_q), \mathrm{Re}(\delta_r) > 0$ we have

$$\int_{U(N)} \frac{\prod_{j=1}^{K} P_N(A^*, -i\alpha_j) \prod_{l=K+1}^{K+L} P_N(A, i\alpha_l)}{\prod_{q=1}^{Q} P_N(A^*, -i\gamma_q) \prod_{r=1}^{R} P_N(A, -i\delta_r)} dA$$
$$= e^{\frac{N}{2}(\sum_{l=1}^{L} \alpha_{K+l} - \sum_{k=1}^{K} \alpha_k)} \frac{(-1)^{(K+L)(K+L-1)/2}}{K!L!(2\pi i)^{K+L}}$$
$$\times \oint \cdots \oint e^{\frac{N}{2}(\sum_{k=1}^{K} w_k - \sum_{l=1}^{L} w_{K+l})} \frac{\prod_{j=1}^{K} \prod_{l=1}^{L} z(w_j - w_{K+l}) \prod_{q=1}^{Q} \prod_{r=1}^{R} z(\gamma_q + \delta_r)}{\prod_{j=1}^{K} \prod_{r=1}^{R} z(w_j + \delta_r) \prod_{l=1}^{L} \prod_{q=1}^{Q} z(-w_{K+l} + \delta_q)}$$
$$\times \frac{\Delta(w_1, \ldots, w_{K+L})^2 \prod_{j=1}^{K+L} dw_j}{\prod_{j=1}^{K+L} \prod_{k=1}^{K+L} (w_k - \alpha_j)}, \tag{7.3.3}$$

where the w contours enclose the poles at $\alpha_1, \ldots, \alpha_{K+L}$. The function $z(x)$ is as defined in (7.1.47).

We set $L = K = Q = R = k$ in (7.3.3) and γ_j, δ_j satisfying the requirements of the theorem. By differentiating the left hand side with respect to all the $\alpha_1, \ldots, \alpha_{2k}$, and using (7.3.1) and (7.3.2) we find that

$$\int_{U(N)} \prod_{m=1}^{2k} \frac{d}{d\alpha_m} \frac{\prod_{j=1}^{k} P_N(A^*, -i\alpha_j) \prod_{l=k+1}^{2k} P_N(A, i\alpha_l)}{\prod_{q=1}^{k} P_N(A^*, -i\gamma_q) \prod_{r=1}^{k} P_N(A, -i\delta_r)} dA$$
$$= \int_{U(N)} \frac{\left(\prod_{j=1}^{k} \frac{d}{d\alpha_j} P_N(A^*, -i\alpha_j)\right) \left(\prod_{l=k+1}^{2k} \frac{d}{d\alpha_l} P_N(A, i\alpha_l)\right)}{\prod_{q=1}^{k} P_N(A^*, -i\gamma_q) \prod_{r=1}^{k} P_N(A, i\delta_r)} dA \tag{7.3.4}$$
$$= (-1)^k \int_{U(N)} \frac{\prod_{j=1}^{k} e^{-\alpha_j} P_N'(A^*, -i\alpha_j) \prod_{l=k+1}^{2k} e^{\alpha_l} P_N'(A, i\alpha_l)}{\prod_{q=1}^{k} P_N(A^*, -i\gamma_q) \prod_{r=1}^{k} P_N(A, i\delta_r)} dA. \tag{7.3.5}$$

Similarly (trivially) the right hand side becomes

$$\prod_{j=1}^{2k} \frac{d}{d\alpha_j} \left[e^{\frac{N}{2}(\sum_{l=1}^{k} \alpha_{k+l} - \sum_{j=1}^{k} \alpha_j)} \frac{(-1)^{k(2k-1)}}{k!^2 (2\pi i)^{2k}} \right.$$
$$\times \oint \cdots \oint e^{\frac{N}{2}(\sum_{j=1}^{k} w_j - \sum_{l=1}^{k} w_{k+l})} \frac{\prod_{j=1}^{k} \prod_{l=1}^{k} z(w_j - w_{k+l}) \prod_{q=1}^{k} \prod_{r=1}^{k} z(\gamma_q + \delta_r)}{\prod_{j=1}^{k} \prod_{r=1}^{k} z(w_j + \delta_r) \prod_{l=1}^{k} \prod_{q=1}^{k} z(-w_{k+l} + \delta_q)}$$
$$\left. \times \frac{\Delta(w_1, \ldots, w_{2k})^2 \prod_{j=1}^{2k} dw_j}{\prod_{j=1}^{2k} \prod_{l=1}^{2k} (w_l - \alpha_j)} \right]. \tag{7.3.6}$$

Subsequently, equating (7.3.5) with (7.3.6) and setting $\alpha_1 = \cdots = \alpha_k = \alpha$, $\alpha_{k+1} = \cdots = \alpha_{2k} = -\alpha$, and all $\gamma, \delta = \alpha$, with $\mathrm{Re}(\alpha) > 0$, we find that

$$(-1)^k \int_{U(N)} \left(\frac{P_N'(A^*, -i\alpha) P_N'(A, -i\alpha)}{P_N(A^*, -i\alpha) P_N(A, -i\alpha)} \right)^k e^{-2\alpha k} dA$$
$$= \prod_{j=1}^{2k} \frac{d}{d\alpha_j} \left(e^{\frac{N}{2}(\sum_{l=k+1}^{2k} \alpha_l - \sum_{j=1}^{k} \alpha_j)} \frac{(-1)^{k(2k-1)}}{k!^2 (2\pi i)^{2k}} \right.$$
$$\times \oint \cdots \oint e^{\frac{N}{2}(\sum_{j=1}^{k} w_j - \sum_{l=1}^{k} w_{k+l})} \frac{\prod_{j=1}^{k} \prod_{l=1}^{k} z(w_j - w_{k+l}) \prod_{q=1}^{k} \prod_{r=1}^{k} z(\gamma_q + \delta_r)}{\prod_{j=1}^{k} \prod_{r=1}^{k} z(w_j + \delta_r) \prod_{l=1}^{k} \prod_{q=1}^{k} z(-w_{k+l} + \delta_q)}$$
$$\left. \times \frac{\Delta(w_1, \ldots, w_{2k})^2 \prod_{j=1}^{2k} dw_j}{\prod_{j=1}^{2k} \prod_{l=1}^{2k} (w_l - \alpha_j)} \right)_{\substack{\alpha_1 = \cdots = \alpha_k \\ \alpha_{k+1} = \cdots = \alpha_{2k} = -\alpha \\ \gamma_1 = \cdots = \gamma_k = \alpha \\ \delta_1 = \cdots = \delta_k = \alpha}}. \tag{7.3.7}$$

Notice that the left hand side of (7.3.7) is essentially the logarithmic derivative we wish to evaluate.

Now we compute the derivative with respect to the α_j in the right hand side of (7.3.7) using (see also the procedure (7.2.21) used to prove theorem 7.1.1)

$$\frac{d}{d\alpha}\frac{e^{\pm N\alpha/2}}{\prod_j (w_j-\alpha)} = \frac{e^{\pm N\alpha/2}}{\prod_j (w_j-\alpha)}\left(\sum_j \frac{1}{(w_j-\alpha)} \pm \frac{N}{2}\right). \tag{7.3.8}$$

Employing (7.3.8) to the right hand side of (7.3.7) and applying the substitutions of the α_j, γ_j, δ_j, we obtain

$$\begin{aligned}(-1)^k &\int_{U(N)} \left(\frac{P_N'(A^*,-i\alpha)}{P_N(A^*,-i\alpha)}\frac{P_N'(A,-i\alpha)}{P_N(A,-i\alpha)}\right)^k e^{-2\alpha k} dA \\ &= \frac{(-1)^k}{k!^2(2\pi i)^{2k}} e^{-Nk\alpha} z(2\alpha)^{k^2} \\ &\quad\times \oint\cdots\oint e^{\frac{N}{2}\left(\sum_{j=1}^k w_j - \sum_{l=1}^k w_{k+l}\right)} \frac{\prod_{j=1}^k \prod_{l=1}^k z(w_j - w_{k+l})}{\prod_{j=1}^k z(w_j+\alpha)^k \prod_{l=1}^k z(-w_{k+l}+\alpha)^k} \\ &\quad\times \frac{\Delta(w_1,\ldots,w_{2k})^2 \left(\sum_{j=1}^{2k}\frac{1}{w_j-\alpha} - \frac{N}{2}\right)^k \left(\sum_{j=1}^{2k}\frac{1}{w_j+\alpha} + \frac{N}{2}\right)^k \prod_{j=1}^{2k} dw_j}{\prod_{j=1}^{2k}(w_j-\alpha)^k (w_j+\alpha)^k}. \end{aligned} \tag{7.3.9}$$

To determine the leading order asymptotics in N, let $\alpha = a/N$, with $a = o(1)$ as $N \to \infty$, and substitute $w_j = au_j/N$. For large N, we can now simplify the integrand by replacing each occurrence of the function $z(x)$ by $1/x$, see (7.2.2). The double product involving $z(w_j - w_{k+l})$ thus cancels a portion of the $\Delta(w_1,\ldots,w_{2k})^2$, up to a factor of $(-1)^{k^2}$, and so we let

$$q(w_1,\ldots,w_{2k}) \coloneqq \Delta(w_1,\ldots,w_{2k})\Delta(w_1,\ldots,w_k)\Delta(w_{k+1},\ldots,w_{2k}), \tag{7.3.10}$$

so q represents the surviving parts of the Vandermonde squared term in the integrand and $q(\frac{d}{dL})$ has the equivalent meaning to (7.2.23). As in the previous section, we can introduce extra variables L_j and pull out the polynomial q from the integrand as a differential operator, see the calculations after (7.2.23). Similarly, the factors containing $z(x)$ in the denominator cancel some of the $(w_j-\alpha)(w_j+\alpha)$ factors in the denominator, again up to a $(-1)^{k^2}$. Carrying out these steps, and cancelling out the powers of a that can be pulled outside the integral, (7.3.9) becomes

$$\begin{aligned}(-1)^k &\int_{U(N)} \left(\frac{P_N'(A^*,-i\frac{a}{N})}{P_N(A^*,-i\frac{a}{N})}\frac{P_N'(A,-i\frac{a}{N})}{P_N(A,-i\frac{a}{N})}\right)^k e^{-2k\frac{a}{N}} dA \\ &= \frac{(-1)^k}{k!^2(2\pi i)^{2k}} e^{-ak}\frac{N^{2k}}{2^{k^2}}\left(1+O\left(\tfrac{a}{N}\right)\right) \\ &\quad\times q\left(\frac{d}{dL}\right)\oint\cdots\oint\Bigg(e^{\sum_{j=1}^{2k}u_jL_j}\Big(\sum_{j=1}^{2k}\frac{1}{au_j-a}-\frac{1}{2}\Big)^k\Big(\sum_{j=1}^{2k}\frac{1}{au_j+a}+\frac{1}{2}\Big)^k \\ &\qquad\times \frac{du_1\cdots du_{2k}}{\prod_{j=1}^k(u_j-1)^k\prod_{j=k+1}^{2k}(u_j+1)^k}\Bigg)\Bigg|_{\substack{L_1,\ldots,L_k=\frac{a}{2}\\ L_{k+1},\ldots,L_{2k}=-\frac{a}{2}}} \end{aligned} \tag{7.3.11}$$

where the contours of integration enclose ± 1.

Introducing more variables t_1, t_2, similarly again to the method used in proving theorem 7.1.1, see

(7.2.26), the right hand side of (7.3.11) can be written as

$$\frac{(-1)^k e^{-ak}}{k!^2(2\pi i)^{2k}2^{k^2}}\left(\frac{N}{a}\right)^{2k}\left(\frac{d}{dt_1}\right)^k\left(\frac{d}{dt_2}\right)^k\Bigg[e^{\frac{a(t_2-t_1)}{2}}\left(1+O\left(\tfrac{a}{N}\right)\right)$$
$$\times\, q\left(\frac{d}{dL}\right)\oint\cdots\oint\frac{\exp\left(\sum_{j=1}^{2k}u_jL_j+\frac{t_1}{u_j-1}+\frac{t_2}{u_j+1}\right)}{\prod_{j=1}^{k}(u_j-1)^k\prod_{j=k+1}^{2k}(u_j+1)^k}du_1\cdots du_{2k}\Bigg]\Bigg|_{\substack{L_1,\dots,L_k=\frac{a}{2}\\ L_{k+1},\dots,L_{2k}=-\frac{a}{2}\\ t_1,t_2=0}} \tag{7.3.12}$$

where the contours still encircle ±1. Now, the $2k$ dimensional residue above can be separated into a product:

$$\frac{1}{(2\pi i)^{2k}}\oint\cdots\oint\frac{\exp\left(\sum_{j=1}^{2k}u_jL_j+\frac{t_1}{u_j-1}+\frac{t_2}{u_j+1}\right)}{\prod_{j=1}^{k}(u_j-1)^k\prod_{j=k+1}^{2k}(u_j+1)^k}du_1\cdots du_{2k}=:\prod_{j=1}^{k}f(L_j)\prod_{j=k+1}^{2k}g(L_j), \tag{7.3.13}$$

where

$$f(L)=\frac{1}{2\pi i}\oint\frac{\exp\left(uL+\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du, \tag{7.3.14}$$

and

$$g(L)=\frac{1}{2\pi i}\oint\frac{\exp\left(uL+\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du. \tag{7.3.15}$$

We now use the technique of lemma 2.2 from [47]. That lemma reads as follows (using consistent notation)

$$q\left(\frac{d}{dL}\right)\prod_{j=1}^{2k}f(z_j)\Bigg|_{\substack{z_1=\cdots=z_k=1\\ z_{k+1}=\cdots=z_{2k}=-1}}=k!^2\det_{2k\times2k}\begin{pmatrix} f(1) & f^{(1)}(1) & \cdots & f^{(2k-1)}(1)\\ f^{(1)}(1) & f^{(2)}(1) & \cdots & f^{(2k)}(1)\\ \vdots & \vdots & \ddots & \vdots\\ f^{(k-1)}(1) & f^{(k)}(1) & \cdots & f^{(3k-2)}(1)\\ f(-1) & f^{(1)}(-1) & \cdots & f^{(2k-1)}(-1)\\ f^{(1)}(-1) & f^{(2)}(-1) & \cdots & f^{(2k)}(-1)\\ \vdots & \vdots & \ddots & \vdots\\ f^{(k-1)}(-1) & f^{(k)}(-1) & \cdots & f^{(3k-2)}(-1)\end{pmatrix}. \tag{7.3.16}$$

Hence, adapting (7.3.16) we find that (7.3.11) becomes

$$(-1)^k\int_{U(N)}\left(\frac{P_N'(A^*,-i\frac{a}{N})}{P_N(A^*,-i\frac{a}{N})}\frac{P_N'(A,-i\frac{a}{N})}{P_N(A,-i\frac{a}{N})}\right)^k e^{-2k\frac{a}{N}}dA$$
$$=\frac{(-1)^k e^{-ak}}{2^{k^2}}\left(\frac{N}{a}\right)^{2k}$$
$$\times\left(\frac{d}{dt_1}\right)^k\left(\frac{d}{dt_2}\right)^k\left[e^{a\frac{(t_2-t_1)}{2}}\left(1+O\left(\tfrac{a}{N}\right)\right)\det_{2k\times2k}\begin{pmatrix}f^{(i+j-2)}(\frac{a}{2})\\ g^{(i+j-2)}(-\frac{a}{2})\end{pmatrix}\right]\Bigg|_{t_1,t_2=0} \tag{7.3.17}$$

where the first k rows of the above matrix ($1\le i\le k$) have entries $f^{(i+j-2)}(\frac{a}{2})$, in column $1\le j\le 2k$, and the last k rows (rows $i+k$, with $1\le i\le k$) have entries $g^{(i+j-2)}(-\frac{a}{2})$, in column $1\le j\le 2k$ (see the pattern evident in the determinant in (7.3.16)).

In [14], we also study the determinant obtained here from the point of view of Riemann–Hilbert problems, see section 7 in that paper.

Next, we drop the factors $e^{a(t_2-t_1)/2}$ and e^{-ak} in the right hand side of (7.3.17) as they do not affect the asymptotic since a becomes small as $N \to \infty$. We also approximate, in the integrands, $e^{\pm au/2}$ by $1 \pm au/2$. One might think that to obtain just the leading order asymptotic for small a, the $\pm au/2$ would not be needed. However, this turns out to be incorrect, since if we just approximate by 1, i.e. without the term $\pm au/2$, the resulting determinant is magically independent of t_1 and t_2, and does not survive the differentiation (lemma 7.3.1) below.

Thus, the leading order term can be simplified to

$$\frac{(-1)^k}{2^{k^2}} \left(\frac{N}{a}\right)^{2k} \left(\frac{d}{dt_1}\right)^k \left(\frac{d}{dt_2}\right)^k \det_{2k\times 2k} \mathrm{M}\Big|_{t_1,t_2=0}, \tag{7.3.18}$$

where the matrix M has entries in the top k rows of

$$\frac{1}{2\pi i} \oint \frac{(1+\frac{au}{2})u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u-1)^k} du, \tag{7.3.19}$$

and entries in the final k rows of

$$\frac{1}{2\pi i} \oint \frac{(1-\frac{au}{2})u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u+1)^k} du. \tag{7.3.20}$$

Returning to (7.3.17), dropping the $\exp(-2ak/N)$ in the left hand side as it does not impact the leading asymptotic when N is large, we have determined,

$$\int_{U(N)} \left|\frac{P_N'}{P_N}(A, \tfrac{ia}{N})\right|^{2k} dA = \frac{1}{2^{k^2}} \left(\frac{N}{a}\right)^{2k} \left(\frac{d}{dt_1}\right)^k \left(\frac{d}{dt_2}\right)^k \left[\det_{2k\times 2k} \mathrm{M}\right]\Bigg|_{t_1,t_2=0} \times (1+O(a)) \tag{7.3.21}$$

as $N \to \infty$ with $a = \alpha N \to 0$.

Consider now the factor $1 \pm au/2$ that appears in the entries (7.3.19) and (7.3.20). The role of this factor can be analyzed using the following multi-linearity property of determinants. Let A be an $n \times n$ matrix, let $a_1, \ldots, a_n$ denote the rows (or columns) of A, and let v be an n-dimensional vector. Then for any scalar x,

$$\det(a_1, \ldots, a_j + xv, \ldots, a_n) = \det(A) + x\det(a_1, \ldots, v, \ldots, a_n). \tag{7.3.22}$$

Expanding in this fashion, the $1 \pm au/2$ results in two determinants for each row, so 2^{2k} determinants altogether. We can encode the choice of term for the row in the vector $\underline{r} = (r_1, \ldots, r_{2k})$, where r_j is the choice of either 1 or $au/2$ for $1 \le j \le k$, and r_j similarly is 1 or $-au/2$ for $k+1 \le j \le 2k$. The lemma below describes what happens in the simplest of cases, where we select, for each row, just the 1 from $1 \pm au/2$, i.e.

$$\underline{r} = (1, 1, \ldots, 1). \tag{7.3.23}$$

The proof, along with that of lemma 7.3.3 which deals with the more complicated question of mixing choices of 1 and $\pm au/2$, will be given in the next section.

Lemma 7.3.1. *We have*

$$\det_{2k\times 2k} \begin{pmatrix} \frac{1}{2\pi i} \oint \frac{u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u-1)^k} du \\ \frac{1}{2\pi i} \oint \frac{u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u+1)^k} du \end{pmatrix} = (-2)^{k^2}. \tag{7.3.24}$$

Thus the determinant in the above lemma, does not depend on t_1 or t_2, and, on applying $\left(\frac{d}{dt_1}\right)^k \left(\frac{d}{dt_2}\right)^k$, does not contribute to (7.3.21).

Now that we understand what happens when just the 1 is selected from $1 \pm au/2$, we next examine the contribution from the $\pm au/2$ terms. We will only consider those determinants that are obtained by a single selection of these terms along exactly one of the rows (as in expansion (7.3.22)), as these are the determinants that will result in the main asymptotics of size N^{2k}/a^{2k-1} (each row for which we select $\pm au/2$ increases the power of a by 1, by (7.3.22), which will contribute to the asymptotic described in (7.3.21)). Hence, we are only considering

$$\underline{r} = (\overbrace{1, \ldots, 1, \tfrac{au}{2}, 1, \ldots, 1}^{k}, \overbrace{1, \ldots, 1}^{k}), \text{ or,} \tag{7.3.25}$$

$$\underline{r} = (\underbrace{1, \ldots, 1}_{k}, \underbrace{1, \ldots, 1, -\tfrac{au}{2}, 1, \ldots, 1}_{k}) \tag{7.3.26}$$

Selecting this term for each entry in a specific row has the effect of incrementing the power of u in the numerator of the corresponding integrands from $i+j-2$ to $i+j-1$. This then matches with the entries in the row below, giving a zero value for the determinant, unless the selected row is row k or $2k$. The next lemma summarizes what happens in either of these two cases. First, however, we give an example for small k.

Example 7.3.2. *Let $k = 3$. Then select $au/2$ for the second row, i.e choose $\underline{r} = (1, au/2, 1, 1, 1, 1)$. The yields the following determinant (multiplied by $a/2$)*

$$\det \begin{pmatrix}
\frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^4\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^4\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^4\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du
\end{pmatrix} = 0, \tag{7.3.27}$$

since the second and third rows are identical. However, if we chose $-au/2$ in the last row (or equally, $au/2$ in the third row) then the determinant is no longer zero. In terms of $\underline{r}$, this choice is encoded as $\underline{r} = (1, 1, 1, 1, 1, -au/2)$ and the determinant now equals (multiplied by $-a/2$)

$$\det \begin{pmatrix}
\frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du & \frac{1}{2\pi i}\oint \frac{u^4\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^3}du \\
\frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^2\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du \\
\frac{1}{2\pi i}\oint \frac{u^3\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^4\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du & \frac{1}{2\pi i}\oint \frac{u^5\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^3}du
\end{pmatrix}. \tag{7.3.28}$$

The next lemma gives the value of the determinant in both the cases that

$$\underline{r} = (\overbrace{1, \ldots, 1, \tfrac{au}{2}}^{k}, \overbrace{1, \ldots, 1}^{k}), \text{ or,} \tag{7.3.29}$$

$$\underline{r} = (\underbrace{1, \ldots, 1}_{k}, \underbrace{1, \ldots, 1, -\tfrac{au}{2}}_{k}) \tag{7.3.30}$$

Lemma 7.3.3. *Let* $\mathrm{M}^{(\mathrm{k})}$ *be the matrix identical to the one displayed in* (7.3.24)*, except that the power of* u *in the numerator of each integrand is* $i+j-1$ *along its* k*th row, rather than* $i+j-2$ *(i.e.* $\underline{r} = (1, \ldots, 1, au/2, 1, \ldots, 1)$*). Then,* $\det \mathrm{M}^{(\mathrm{k})}$ *is a polynomial in* t_1 *and* t_2 *of degree* $2k$*, and satisfies:*

$$\det \mathrm{M}^{(\mathrm{k})} = \frac{k\, 2^{k^2-2k}}{k!^2(2k-1)}(t_1+t_2)^{2k} + O((|t_1|+|t_2|)^{2k-1}). \tag{7.3.31}$$

Furthermore, let $\mathrm{M}^{(2\mathrm{k})}$ *be the matrix identical to* (7.3.24)*, except for its* $2k$*th row features* u *instead of the power* $i+j-1$ *(i.e.* $\underline{r} = (1, \ldots, 1, 1, \ldots, 1, -au/2)$*), then*

$$\det \mathrm{M}^{(2\mathrm{k})} = -\frac{k\, 2^{k^2-2k}}{k!^2(2k-1)}(t_1+t_2)^{2k} + O((|t_1|+|t_2|)^{2k-1}). \tag{7.3.32}$$

We have therefore proved that, to leading order, the only contribution to the main term comes from selecting either the kth row or the $2k$th row from the multi-linearity decomposition of the determinant of M. Hence, as $N \to \infty$ with $a = \alpha N \to 0$

$$\begin{aligned}
\int_{U(N)} \left| \frac{P_N'}{P_N}(A, i\alpha) \right|^{2k} dA &= \frac{1}{2^{k^2}} \left(\frac{N}{a}\right)^{2k} \\
&\times \left(\frac{d}{dt_1}\right)^k \left(\frac{d}{dt_2}\right)^k \left[\frac{a}{2} \det_{2k\times 2k} \mathrm{M}^{(\mathrm{k})} - \frac{a}{2} \det_{2k \times 2k} \mathrm{M}^{(2\mathrm{k})} \right]\Bigg|_{t_1,t_2=0} \times (1+O(a)) \quad (7.3.33)\\
&= \frac{1}{2^{k^2}} \left(\frac{N}{a}\right)^{2k} \frac{a}{2} \frac{k\, 2^{k^2-2k}}{(k!)^2(2k-1)} \\
&\times \left(\frac{d}{dt_1}\right)^k \left(\frac{d}{dt_2}\right)^k \left[2(t_1+t_2)^{2k} + O((|t_1|+|t_2|)^{2k-1}) \right]\Bigg|_{t_1,t_2=0} \\
&\times (1+O(a)) \quad (7.3.34)\\
&= \left(\frac{N}{a}\right)^{2k} \frac{a}{2} \frac{k\, 2^{-2k}}{(k!)^2(2k-1)}\, 2(2k)! \times (1+O(a)) \quad (7.3.35)\\
&= \frac{2k(2k-2)!}{(k-1)!k!} \frac{N^{2k}}{2^{2k}a^{2k-1}} \times (1+O(a)) \quad (7.3.36)\\
&= \binom{2k-2}{k-1} \frac{N^{2k}}{(2a)^{2k-1}} \times (1+O(a)). \quad (7.3.37)
\end{aligned}$$

So we have (7.1.13).

7.3.1 Proof of lemma 7.3.1

Recall the statement of the lemma,

$$\det_{2k\times 2k} \begin{pmatrix} \frac{1}{2\pi i} \oint \frac{u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u-1)^k} du \\ \frac{1}{2\pi i} \oint \frac{u^{i+j-2} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u+1)^k} du \end{pmatrix} = (-2)^{k^2}. \tag{7.3.38}$$

We first establish that the determinant in (7.3.38) is independent of t_1 and t_2 by showing that its derivative with respect to either variable is 0.

When we differentiate with respect, say, to t_1 we get a sum of $2k$ determinants of the $2k$ matrices formed by differentiating the entries of a specific column of the original matrix. We will show that each of these $2k$ determinants is 0. The jth of these determinants has the entries of its jth column differentiated with respect to t_1, and they are equal, in the top half of the matrix (in the ith row, with $1 \le i \le k$), to

$$\frac{1}{2\pi i}\int_{|u|=2} \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k+1}}du \tag{7.3.39}$$

and, in the bottom half (in the $(k+i)$th row, with $1 \le i \le k$),

$$\frac{1}{2\pi i}\int_{|u|=2} \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{k}(u-1)}du. \tag{7.3.40}$$

If we select the first column, i.e. $j = 1$, the integrand of each entry in this column is of size $O(|u|^{-2})$, as $|u| \to \infty$. This is because the numerator in (7.3.39) or (7.3.40) is at most $O(|u|^{k-1})$, and the denominator is $O(|u|^{k+1})$. As $|u| \to \infty$, the length of the contour grows proportionally to $|u|$, hence taking a large contour shows that each entry in this column is 0, and hence the determinant is 0.

Otherwise, if the column being differentiated has index $j > 1$, we can show that the resulting column is a linear combination of columns $1, \dots, j-1$ (and therefore again, the determinant is 0). For, if we add the first $j-1$ entries in the ith row of the top half of the matrix, we get

$$\frac{1}{2\pi i}\oint \sum_{l=1}^{j-1} \frac{u^{i+l-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k}}du = \frac{1}{2\pi i}\oint \frac{u^{i-1}(u^{j-1}-1)\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k+1}}du. \tag{7.3.41}$$

This nearly matches (7.3.39), the difference being

$$\frac{1}{2\pi i}\oint \frac{u^{i-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k+1}}du, \tag{7.3.42}$$

but by argument above, the integrand is $O(|u|^{-2})$, hence (7.3.42) equals 0.

Similarly, the sum of the first $j-1$ entries in row i in the bottom half equals

$$\frac{1}{2\pi i}\oint \sum_{l=1}^{j-1} \frac{u^{i+l-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{k}}du = \frac{1}{2\pi i}\oint \frac{u^{i-1}(u^{j-1}-1)\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{k}(u-1)}du. \tag{7.3.43}$$

Again, the difference between the right hand side above and (7.3.40),

$$\frac{1}{2\pi i}\oint \frac{u^{i-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{k}(u-1)}du, \tag{7.3.44}$$

equals 0, because the integrand is $O(|u|^{-2})$.

We have thus shown that the jth differentiated column is equal to the sum of the first $j-1$ non-differentiated columns, and hence by multi-linearity, the corresponding determinant is 0, as claimed.

A similar computation shows the derivative with respect to t_2 of the determinant in the lemma equals 0. This also shows that it is not enough to approximate the $\exp(\pm au/2)$ term in (7.3.17) by just 1.

Having established that the left hand side of (7.3.24) is independent of t_1 and t_2, we can determine its value by specializing $t_1 = t_2 = 0$, in which case we can evaluate the residue at $u = 1$ and the top k

rows have entries

$$\frac{1}{2\pi i}\oint \frac{u^{i+j-2}}{(u-1)^k}du = \binom{i+j-2}{k-1}, \quad 1 \le i \le k,\ 1 \le j \le 2k, \tag{7.3.45}$$

and the bottom k rows have entries

$$\frac{1}{2\pi i}\oint \frac{u^{i+j-2}}{(u+1)^k}du = (-1)^{i+j-k-1}\binom{i+j-2}{k-1}, \qquad 1 \le i \le k,\ 1 \le j \le 2k. \tag{7.3.46}$$

The first identity is easily obtained by writing $u^{i+j-2} = ((u-1)+1)^{i+j-2}$ and extracting the coefficient of $(u-1)^{k-1}$. The second identity follows similarly by writing $u^{i+j-2} = ((u+1)-1)^{i+j-2}$.

Next, we can pull out $(-1)^{-k-1}$ from each of the bottom k rows of the determinant, and as $k(k+1)$ is even, these powers of -1 altogether give 1. Hence we have shown that

$$\det_{2k\times 2k}\begin{pmatrix} \frac{1}{2\pi i}\oint \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du \\ \frac{1}{2\pi i}\oint \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du \end{pmatrix} = \det\begin{pmatrix} \binom{i+j-2}{k-1} \\ {\scriptstyle 1\le i\le k, 1\le j\le 2k} \\ \hline \\ (-1)^{i+j}\binom{i+j-2}{k-1} \\ {\scriptstyle 1\le i\le k, 1\le j\le 2k} \end{pmatrix}. \tag{7.3.47}$$

For example, the matrix (7.3.47) for $k=3$ is the following.

$$\begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} \\ \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} & \binom{7}{2} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} \\ \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} & -\binom{7}{2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 3 & 6 & 10 \\ 0 & 1 & 3 & 6 & 10 & 15 \\ 1 & 3 & 6 & 10 & 15 & 21 \\ 0 & 0 & 1 & -3 & 6 & -10 \\ 0 & 1 & -3 & 6 & -10 & 15 \\ 1 & -3 & 6 & -10 & 15 & -21 \end{pmatrix} \tag{7.3.48}$$

Notice the symmetry between the top and bottom half of the matrix, as well as the 'chequerboard' pattern of -1s in the bottom half.

The aim is now to manipulate the determinants of matrices of the type shown in (7.3.47) in to a determinant that has already been computed. Namely, the result will follow from the following calculation of Conrey et al. [46] (c.f. (2.7.14) in that paper),

$$\det\begin{pmatrix} \binom{0}{0} & \binom{0}{1} & \cdots & \binom{0}{k-1} & \binom{0}{0} & -\binom{0}{1} & \cdots & (-1)^{k-1}\binom{0}{k-1} \\ \binom{1}{0} & \binom{1}{1} & \cdots & \binom{1}{k-1} & -\binom{1}{0} & \binom{1}{1} & \cdots & (-1)^{k}\binom{1}{k-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{2k-1}{0} & \binom{2k-1}{1} & \cdots & \binom{2k-1}{k-1} & -\binom{2k-1}{0} & \binom{2k-1}{1} & \cdots & (-1)^{k}\binom{2k-1}{k-1} \end{pmatrix} = (-2)^{k^2}. \tag{7.3.49}$$

In the top half of the matrix from (7.3.47), starting from row k and working up, we subtract row $i-1$ from row i, $i = k, \ldots, 2$ and use Pascal's identity:

$$\binom{n}{r} - \binom{n-1}{r} = \binom{n-1}{r-1}. \tag{7.3.50}$$

This decreases by 1 both indices of the binomial coefficients in all elements of rows 2 to k but does not change the determinant. The first row remains unchanged. An example of this method can be found below, see example 7.3.4. In the bottom half, instead of subtracting, we add row $i-1$ to row i for $i = k, k-1, \ldots, 2$.

We then repeat the procedure, but this time on rows $i = k, k-1, \ldots, 3$, (this time reducing both indices of the binomial coefficients in all except the first *two* rows) and so on, until we have row reduced the matrix to the following form:

$$\det \begin{pmatrix} \binom{0}{k-1} & \cdots & \binom{2k-1}{k-1} \\ \vdots & \ddots & \vdots \\ \binom{0}{0} & \cdots & \binom{2k-1}{0} \\ \binom{0}{k-1} & \cdots & -\binom{2k-1}{k-1} \\ \vdots & \ddots & \vdots \\ (-1)^{k+1}\binom{0}{0} & \cdots & (-1)^{3k}\binom{2k-1}{0} \end{pmatrix}. \tag{7.3.51}$$

Example 7.3.4. *Let $k = 3$. Then by (7.3.48) the matrix (7.3.47) has the form*

$$\begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} \\ \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} & \binom{7}{2} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} \\ \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} & -\binom{7}{2} \end{pmatrix}. \tag{7.3.52}$$

We subtract row 2 from row 3, add row 5 and row 6, and apply Pascal's identity,

$$\begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} \\ \binom{2}{2}-\binom{1}{2} & \binom{3}{2}-\binom{2}{2} & \binom{4}{2}-\binom{3}{2} & \binom{5}{2}-\binom{4}{2} & \binom{6}{2}-\binom{5}{2} & \binom{7}{2}-\binom{6}{2} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} \\ \binom{2}{2}-\binom{1}{2} & -\binom{3}{2}+\binom{2}{2} & \binom{4}{2}-\binom{3}{2} & -\binom{5}{2}+\binom{4}{2} & \binom{6}{2}-\binom{5}{2} & -\binom{7}{2}+\binom{6}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} & \binom{6}{2} \\ \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} & \binom{5}{1} & \binom{6}{1} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} & \binom{6}{2} \\ \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1} & \binom{5}{1} & -\binom{6}{1} \end{pmatrix}, \tag{7.3.53}$$

repeating this procedure, we can subtract row 1 from row 2, add row 4 and row 5 and apply Pascal's identity again, giving

$$= \begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{0}{1} & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} & \binom{6}{2} \\ \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} & \binom{5}{1} & \binom{6}{1} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{0}{1} & \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1} & \binom{5}{1} \\ \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1} & \binom{5}{1} & -\binom{6}{1} \end{pmatrix}, \tag{7.3.54}$$

and finally, we can repeat the whole procedure once to reduce rows 3 *and* 6 *(i.e. take the new row* 2 *from row* 3 *etc.) to find*

$$= \begin{pmatrix} \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \binom{5}{2} \\ \binom{0}{1} & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} & \binom{6}{2} \\ \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \binom{4}{0} & \binom{6}{1} \\ \binom{0}{2} & -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} & -\binom{5}{2} \\ -\binom{0}{1} & \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1} & \binom{5}{1} \\ \binom{0}{0} & -\binom{1}{0} & \binom{2}{0} & -\binom{3}{0} & \binom{4}{0} & -\binom{5}{0} \end{pmatrix}. \tag{7.3.55}$$

This is in the form of (7.3.51)*.*

We now rearrange the rows so to match (7.3.49). An interchange of any two rows changes the determinant by a factor of -1. An even number of row swaps (the same for the top and bottom halves), and pulling out $(-1)^{k-1}$ from each of the k bottom rows therefore does not change the determinant, but transforms it in to the form of (7.3.49). Hence

$$\det_{2k\times 2k}\begin{pmatrix} \frac{1}{2\pi i}\oint \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du \\ \frac{1}{2\pi i}\oint \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du \end{pmatrix} = (-2)^{k^2}, \tag{7.3.56}$$

as required.

7.3.2 Proof of lemma 7.3.3

The method for proving lemma 7.3.3 begins similarly to that of lemma 7.3.1.We first recall the statement of the lemma. Let $\mathrm{M^{(k)}}$ be the following $2k \times 2k$ matrix

$$\mathrm{M^{(k)}} := \begin{pmatrix} \frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du & \cdots & \frac{1}{2\pi i}\oint \frac{u^{k-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du \\ \vdots & \ddots & \vdots \\ \frac{1}{2\pi i}\oint \frac{u^{k-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du & \cdots & \frac{1}{2\pi i}\oint \frac{u^{2k-3}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du \\ \frac{1}{2\pi i}\oint \frac{u^{k}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du & \cdots & \frac{1}{2\pi i}\oint \frac{u^{2k-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^k}du \\ \frac{1}{2\pi i}\oint \frac{\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du & \cdots & \frac{1}{2\pi i}\oint \frac{u^{k-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du \\ \vdots & \ddots & \vdots \\ \frac{1}{2\pi i}\oint \frac{u^{k-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du & \cdots & \frac{1}{2\pi i}\oint \frac{u^{2k-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^k}du \end{pmatrix}, \tag{7.3.57}$$

(i.e. the matrix (7.3.24) but with the power of u in the integrand of the kth row replaced by $i+j-1$). Then we have to prove that

$$\det \mathrm{M^{(k)}} = \frac{k2^{k^2-2k}}{(k!)^2(2k-1)}(t_1+t_2)^{2k} + O((|t_1|+|t_2|)^{2k-1}). \tag{7.3.58}$$

Similarly, for the equivalent statement and definition of $\mathrm{M^{(2k)}}$.

We will first prove that $\det \mathrm{M^{(k)}}$ is a polynomial of degree $2k$ in t_1 and t_2. Our strategy is to show that the $(2k+1)$-st and higher partial derivatives are all 0. This is achieved below with the help of lemma 7.3.6, proposition 7.3.7 and proposition 7.3.8. Then in proposition 7.3.9 we determine the value of the coefficients of the terms of order $2k$.

Differentiating our $2k \times 2k$ determinant with respect to either variable produces, as in the proof of the previous lemma, a sum of $2k$ determinants where the entries of the resulting matrix are identical to the original, except that the jth determinant has the entries of its jth column differentiated. If we repeatedly differentiate at least $2k+1$ times in total with respect to the two t variables, we get a sum of determinants, each one specified by two lists of non-negative integers

$$\{m_1, \dots m_{2k}\} \text{ and } \{n_1, \dots n_{2k}\}, \tag{7.3.59}$$

such that

$$m_1 + \dots + m_{2k} + n_1 + \dots + n_{2k} > 2k. \tag{7.3.60}$$

Here m_j is the number of times that column j has been differentiated with respect to t_1 and n_j is the number of times column j has been differentiated with respect to t_2.

Thus we are looking at the determinant of the matrix with upper entries

$$\frac{1}{2\pi i}\int_{|u|=2} \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k+m_j}(u+1)^{n_j}}\,du, \qquad \text{for } 1 \le i < k, 1 \le j \le 2k; \tag{7.3.61}$$

$$\frac{1}{2\pi i}\int_{|u|=2} \frac{u^{k-1+j}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{k+m_j}(u+1)^{n_j}}\,du, \qquad \text{for } i = k, 1 \le j \le 2k; \tag{7.3.62}$$

and lower entries

$$\frac{1}{2\pi i}\int_{|u|=2} \frac{u^{i+j-2}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{m_j}(u+1)^{k+n_j}}\,du, \qquad \text{for } 1 \le i \le k, 1 \le j \le 2k. \tag{7.3.63}$$

To facilitate this discussion it is helpful to let

$$I(r,E,G) := \frac{1}{2\pi i}\int_{|u|=2} \frac{u^{r}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u-1)^{E}(u+1)^{G}}\,du. \tag{7.3.64}$$

Note that if $E+G \ge r+2$ then, as in the proof of the previous lemma, $I(r,E,G) = 0$ – see the argument after (7.3.39). Also, we have two easily proved recursion formulas:

$$I(r,E,G) = I(r-1,E-1,G) + I(r-1,E,G) \tag{7.3.65}$$

and

$$I(r,E,G) = I(r-1,E,G-1) - I(r-1,E,G). \tag{7.3.66}$$

For example, one shows (7.3.65) by observing

$$\begin{aligned} &I(r-1,E-1,G) + I(r-1,E,G) \\ &\qquad = \frac{1}{2\pi i}\int_{|u|=2} \frac{u^{r-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{G}}\left(\frac{1}{(u-1)^{E-1}}+\frac{1}{(u-1)^{E}}\right)du && (7.3.67) \\ &\qquad = \frac{1}{2\pi i}\int_{|u|=2} \frac{u^{r-1}\exp\left(\frac{t_1}{u-1}+\frac{t_2}{u+1}\right)}{(u+1)^{G}}\frac{u}{(u-1)^{E}}\,du && (7.3.68) \\ &\qquad = I(r,E,G). && (7.3.69) \end{aligned}$$

In general we are interested in the following collection of matrices

Definition 7.3.5. *Let $\mathcal{M} := \mathcal{M}_{2k}$ be the set of $2k \times 2k$ matrices $M = (M_{i,j})$ where each entry $M_{i,j}$ is* *given by the integrals*

$$M_{i,j} := I(r_{i,j}, E_j, G_j) \tag{7.3.70}$$

$$= \frac{1}{2\pi i} \int_{|u|=2} \frac{u^{r_{i,j}} \exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u-1)^{E_j}(u+1)^{G_j}}\, du. \tag{7.3.71}$$

Note that the exponents in the denominator of the integrand, E_j and G_j, depend only on the column index, j. Moreover, for the definition of $\mathcal{M}$ we require that each column have a similar structure regarding the exponents $r_{i,j}$, namely that

$$r_{i,j} := c_j + r_i \tag{7.3.72}$$

where $c_j, r_i \in \mathbb{Z}$.

For the particular form of matrix we are interested in, the derivatives of $\mathrm{M}^{(\mathrm{k})}$ given by (7.3.61) to (7.3.63), we could define, for example,

$$r_i = i - 2 \qquad \text{for } 1 \le i < k, \tag{7.3.73}$$

$$r_k = k - 1 \tag{7.3.74}$$

$$r_i = i - k - 2 \quad \text{for } k + 1 \le i \le 2k, \tag{7.3.75}$$

$$c_j = j \qquad \text{for } 1 \le j \le 2k \tag{7.3.76}$$

in the definition of $M_{i,j}$.

Let us define the *degree* of $M_{i,j}$ as

$$d_{i,j} := r_{i,j} - E_j - G_j. \tag{7.3.77}$$

We will also sometimes refer equivalently to the "degree" of $I(r_{i,j}, E_j, G_j)$.

We call the degree of the Jth column

$$D_J \equiv D_J(M) := \max_i d_{i,J}, \tag{7.3.78}$$

i.e. the maximal degree of any entry in the column. We define the total degree of M to be

$$D \equiv D(M) := \sum_{J=1}^{2k} D_J. \tag{7.3.79}$$

Note that any column with $D_J \le -2$ is a column of zeros by the usual argument of considering the size of the integrand.

If we apply one of our recursion formulae, (7.3.65) or (7.3.66), to each entry in a particular column then each entry in that column is a sum and we can split our determinant into a sum of two determinants using multi-linearity (see (7.3.22)) along that column. One determinant will be of a matrix with the same degree as the original matrix and one will have a degree that is less by 1. This idea is utilised in the following lemma.

Lemma 7.3.6. *If the matrix $M \in \mathcal{M}$ (see definition 7.3.5) has two equal column degrees, say $D_J(M) = D_{J'}(M)$ for some $J \ne J'$, then there exists a matrix $M' \in \mathcal{M}$ such that $\det M = \det M'$ and $D(M') < D(M)$. This means that we can replace the determinant in question by a determinant of a matrix of lower degree.*

Proof. Assume that we have $D_J(M) = D_{J'}(M)$ for some columns J and J'. Due to the structure of the exponent $r_{i,j} = r_i + c_j$ in (7.3.71), this means that $d_{i,J} = d_{i,J'}$ for all $1 \leq i \leq 2k$. Using the definition of the degree $d_{i,j}$, see (7.3.77), this means that $c_J + r_i - E_J - G_J = c_{J'} + r_i - E_{J'} - G_{J'}$ or equivalently

$$c_{J'} = c_J - (E_J - E_{J'}) - (G_J - G_{J'}). \tag{7.3.80}$$

Assume for convenience that $E_J > E_{J'}$ and $G_J > G_{J'}$, but all other orderings follow in exactly the same way. Then using (7.3.65) and (7.3.66) we act on each element, indexed by $1 \leq i \leq 2k$, in column J in the following way

$$\begin{aligned}
I(r_{i,J}, E_J, G_J) &= I(c_J + r_i, E_J, G_J) && (7.3.81)\\
&= I(c_J + r_i - 1, E_J - 1, G_J) + \mathit{lower} && (7.3.82)\\
&= I(c_J + r_i - 2, E_J - 2, G_J) + \mathit{lower} && (7.3.83)\\
&= \cdots \\
&= I(c_J + r_i - (E_J - E_{J'}), E_{J'}, G_J) + \mathit{lower} && (7.3.84)\\
&= I(c_J + r_i - (E_J - E_{J'}) - 1, E_{J'}, G_J - 1) + \mathit{lower} && (7.3.85)\\
&= I(c_J + r_i - (E_J - E_{J'}) - 2, E_{J'}, G_J - 2) + \mathit{lower} && (7.3.86)\\
&= \cdots \\
&= I(c_J + r_i - (E_J - E_{J'}) - (G_J - G_{J'}), E_{J'}, G_{J'}) + \mathit{lower} && (7.3.87)\\
&= I(c_{J'} + r_i, E_{J'}, G_{J'}) + \mathit{lower}, && (7.3.88)
\end{aligned}$$

where *lower* denotes a matrix element of lower degree. This is true for any row i, so we separate the determinant using multi-linearity so that we have the sum of two determinants, one with $I(c_J + r_i, E_J, G_J)$ replaced with $I(c_{J'} + r_i, E_{J'}, G_{J'})$ in each element (i, J) and the other with the $(i, J)^{th}$ element replaced by something of lower degree. The former determinant is zero because it has two equal columns (namely J and J') and the latter is a determinant of a matrix of lower degree than M. □

We continue, in the following two propositions, to eliminate cases where the determinant is zero.

Proposition 7.3.7. *For $M \in \mathcal{M}$, suppose that the total degree of M, $D(M) < 2k^2 - 3k$. Then $\det(M) = 0$. Furthermore, if $D(M) = 2k^2 - 3k$ and $\det(M) \neq 0$ then it follows that the column degrees, in some order, take distinct values from $-1, 0, 1, 2, \ldots, 2k - 2$.*

Proof. We may assume that no two columns have equal degrees or else we apply lemma 7.3.6 and reduce out of that situation. Next, if $D_J \leq -2$ for any J then we have a column of zeros and the determinant is zero by the usual argument. Then the minimal total degree for a matrix with non-zero determinant will occur when the column degrees are (in some order) $-1, 0, 1, 2, \ldots, 2k - 2$. However,

$$-1 + 0 + 1 + \cdots + (2k - 2) = 2k^2 - 3k. \tag{7.3.89}$$

□

Now we specialise to the case described by (7.3.61)–(7.3.63) with the following proposition.

Proposition 7.3.8. *Suppose that m_j and n_j are non-negative integers for $j = 1, \ldots, 2k$ such that*

$$m_1 + \cdots + m_{2k} + n_1 + \cdots + n_{2k} > 2k \tag{7.3.90}$$

and let $M = (M_{i,j})_{1\le i,j\le 2k}$ *with*

$$M_{i,j} = \begin{cases} I(i+j-2, k+m_j, n_j) & \text{if } 1 \le i \le k-1, 1 \le j \le 2k \\ I(i+j-1, k+m_j, n_j) & \text{if } i = k, 1 \le j \le 2k \\ I(i-k+j-2, m_j, k+n_j) & \text{if } k+1 \le i \le 2k, 1 \le j \le 2k. \end{cases} \tag{7.3.91}$$

Then $\det M = 0$.

The same is true if the matrix in question is

$$M_{i,j} = \begin{cases} I(i+j-2, k+m_j, n_j) & \text{if } 1 \le i \le k, 1 \le j \le 2k \\ I(i-k+j-2, m_j, k+n_j) & \text{if } k+1 \le i \le 2k-1, 1 \le j \le 2k \\ I(i-k+j-1, m_j, k+n_j) & \text{if } i = 2k, 1 \le j \le 2k. \end{cases} \tag{7.3.92}$$

Proof. Notice that the matrix $M = (M_{i,j})$ defined by (7.3.91) is exactly $\mathrm{M}^{(\mathrm{k})}$, and the matrix given in (7.3.92) is $\mathrm{M}^{(2\mathrm{k})}$ (see for example (7.3.57)). Thus, the statement of the lemma concerns differentiating $\mathrm{M}^{(\mathrm{k})}$ (resp. $\mathrm{M}^{(2\mathrm{k})}$) more than k times with respect to either t_1 or t_2.

As the degree of $I(r_{i,j}, E_j, G_j)$ is $r_{i,j} - E_j - G_j$ by (7.3.77), it is easy to check in (7.3.91) that the maximal degree for each column comes from the entries in the kth row, and in (7.3.92) the maximal degree comes from entries in the $2k$th row. In either case, for the Jth column, the largest entry is in either the kth or the $2k$th position and we have

$$D_J(M) = J - 1 - m_J - n_J \tag{7.3.93}$$

and

$$D(M) = \sum_{J=1}^{2k} D_J(M) = 2k^2 - k - \sum_{J=1}^{2k} (m_J + n_J) < 2k^2 - 3k. \tag{7.3.94}$$

By proposition 7.3.7 we have $\det(M) = 0$. □

Remembering that m_j is the number of times that column j has been differentiated with respect to t_1 and n_j is the number of times column j has been differentiated with respect to t_2, we have thus shown that all $(2k+1)$st and higher partial derivatives of $\det \mathrm{M}^{(\mathrm{k})}$ in lemma 7.3.3 are 0. Therefore $\det \mathrm{M}^{(\mathrm{k})}$ is a polynomial of degree at most $2k$ in t_1 and t_2.

Next we determine that $\det \mathrm{M}^{(\mathrm{k})}$ is a polynomial of degree $2k$ in t_1 and t_2 by identifying the coefficients of the terms $t_1^a t_2^b$ of degree $a+b = 2k$. Consider a mixed derivative $\frac{d^a}{dt_1^a} \frac{d^b}{dt_2^b}$ of $\det \mathrm{M}^{(\mathrm{k})}$ and set $t_1 = t_2 = 0$. As before, we get a sum of determinants, where each determinant is associated to one of the ways in which we can differentiate the columns of $\det \mathrm{M}^{(\mathrm{k})}$ with respect to t_1 (a times) and with respect to t_2 (b times). The following proposition describes what happens to a single one of these determinants.

Proposition 7.3.9. *Now suppose that we have the same matrix M defined at* (7.3.91) *except with*

$$m_1 + \cdots + m_{2k} + n_1 + \cdots + n_{2k} = 2k, \tag{7.3.95}$$

i.e. the total degree is $2k^2 - 3k$. *If the determinant is not zero, then*

$$\det M = \pm \binom{2k-2}{k-1} 2^{(k-1)^2}. \tag{7.3.96}$$

The same is true for a matrix of form (7.3.92).

Proof. We begin with M defined by (7.3.91) (so $M = \mathrm{M}^{(\mathrm{k})}$). Let $p_J := m_J + n_J$ for each $1 \le J \le 2k$. Consider the top half of the matrix, $1 \le i \le k$ and apply the usual decomposition of the Jth column,

$$\begin{aligned}
M_{i,J} &= I(r_{i,J}, k + m_J, n_J) && (7.3.97)\\
&= I(r_{i,J} - 1, k + m_J, n_J - 1) + \mathit{lower} && (7.3.98)\\
&= \cdots\\
&= I(r_{i,J} - n_J, k + m_J, 0) + \mathit{lower} && (7.3.99)\\
&= I(r_{i,J} - n_J - 1, k + m_J - 1, 0) + \mathit{lower} && (7.3.100)\\
&= \cdots\\
&= I(r_{i,J} - n_J - m_J, k, 0) + \mathit{lower} && (7.3.101)\\
&= I(r_{i,J} - p_J, k, 0) + \mathit{lower} && (7.3.102)\\
&= \binom{r_{i,J} - p_J}{k-1} + \mathit{lower}, && (7.3.103)
\end{aligned}$$

where $r_{i,J}$ is as defined by (7.3.91) and the final evaluation is done by a simple residue calculation of the integral I,

$$\begin{aligned}
I(r_{i,J} - p_J, k, 0) &= \frac{1}{2\pi i}\int_{|u|=2} \frac{u^{r_{i,J}-p_J}\exp\left(\frac{t_1}{u-1} + \frac{t_2}{u+1}\right)}{(u-1)^k} du && (7.3.104)\\
&= \frac{1}{2\pi i}\int_{|u|=2} \frac{u^{r_{i,J}-p_J}}{(u-1)^k} du && (7.3.105)\\
&= \binom{r_{i,J} - p_J}{k-1} && (7.3.106)
\end{aligned}$$

where we set $t_1, t_2 = 0$ to determine the coefficient, and use (7.3.45) to evaluate the integral.

In the lower half of the matrix, for $k+1 \le i \le 2k$, we have similarly

$$\begin{aligned}
M_{i,J} &= I(r_{i,J}, m_J, k + n_J) && (7.3.107)\\
&= I(r_{i,J} - 1, m_J, k + n_J - 1) + \mathit{lower} && (7.3.108)\\
&= \cdots\\
&= I(r_{i,J} - n_J, m_J, k) + \mathit{lower} && (7.3.109)\\
&= I(r_{i,J} - n_J - 1, m_J - 1, k) + \mathit{lower} && (7.3.110)\\
&= \cdots\\
&= I(r_{i,J} - n_J - m_J, 0, k) + \mathit{lower} && (7.3.111)\\
&= I(r_{i,J} - p_J, 0, k) + \mathit{lower} && (7.3.112)\\
&= (-1)^{r_{i,J}-p_J-k-1}\binom{r_{i,J} - p_J}{k-1} + \mathit{lower}. && (7.3.113)
\end{aligned}$$

Now we separate the determinant, as described at (7.3.22), so that we have the sum of two determinants, one with the binomial coefficients down column J and the other with a lower degree integral. However, in the latter matrix, the degree of column J will be lower than the degree of the original matrix M. Since the degree of M is $2k^2 - 3k$ and we ascertained in proposition 7.3.7 that any matrix with lower degree has zero determinant, we are simply left with the determinant of the matrix with column J replaced with the binomial coefficients given in (7.3.97) and (7.3.107). We repeat this process for each of the columns of M to end up with a matrix of binomial coefficients. Thus, we have shown that the determinant of M given by (7.3.91) is equal to the determinant of

$$\begin{pmatrix} \binom{-m_1-n_1}{k-1} & \binom{1-m_2-n_2}{k-1} & \cdots & \binom{k-1-m_k-n_k}{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{k-2-m_1-n_1}{k-1} & \binom{k-1-m_2-n_2}{k-1} & \cdots & \binom{2k-3-m_k-n_k}{k-1} \\ \binom{k-m_1-n_1}{k-1} & \binom{k+1-m_2-n_2}{k-1} & \cdots & \binom{2k-1-m_k-n_k}{k-1} \\ (-1)^{m_1+n_1+k+1}\binom{-m_1-n_1}{k-1} & (-1)^{m_2+n_2+k}\binom{1-m_2-n_2}{k-1} & \cdots & (-1)^{m_k+n_k+k}\binom{k-1-m_k-n_k}{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{m_1+n_1}\binom{k-1-m_1-n_1}{k-1} & (-1)^{m_2+n_2+1}\binom{k-m_2-n_2}{k-1} & \cdots & (-1)^{m_k+n_k+k-1}\binom{2k-2-m_k-n_k}{k-1} \end{pmatrix}. \tag{7.3.114}$$

Similarly, the matrix M given by (7.3.92) will be as above, except that the exceptional row will occur in the $2k$th row rather than the kth.

We will now refer to the "degree" of a binomial coefficient as being the degree of the integral I that it came from. So, the degree of $\binom{r_{i,J}-p_J}{k-1}$ is $r_{i,J} - p_J - k$, and recall $p_J = m_J + n_J$. Our matrix M has the structure (7.3.91) (or (7.3.92)) and no degrees have been changed by the processes of turning it into a matrix of binomial coefficients using (7.3.97) and (7.3.107). Therefore the degree of a column is determined by the degree of the element in the kth row (respectively $2k$th). In the kth (respectively $2k$th) row - which, recall, always has the largest degree - $r_{k,J} = k + J - 1$ (resp. $r_{2k,J} = k + J - 1$) so in either (7.3.91) or (7.3.92) the degree of the Jth column is $D_J = J - 1 - p_J$.

Since the degree of the matrix is still $2k^2 - 3k$, and the determinant is not zero, by proposition 7.3.7 we have that the column degrees for $J = 1, 2, \ldots, 2k$ (which are governed by the value of the kth, resp $2k$th row) must take the distinct values in

$$\{-1, 0, 1, 2, \ldots, 2k-2\}. \tag{7.3.115}$$

For example, one way to arrange this would be to have $m_J + n_j = p_J = 1$ for all J, but this is not the only solution. Thus the elements of the kth (see (7.3.114) for reference) (resp. $2k$th) row *must*, in some order, take values $\binom{k-1}{k-1}$, $\binom{k}{k-1}, \ldots, \binom{3k-2}{k-1}$ (resp. $\binom{k-1}{k-1}$, $-\binom{k}{k-1}$, $\binom{k+1}{k-1}, \ldots, -\binom{3k-2}{k-1}$ for matrix (7.3.92)), so as to achieve the required set of column degrees. If all the $m_J + n_J = p_J = 1$ then the elements occur in this order across row k ($2k$, respectively), but for other combinations of the p_Js they will occur in a different order. Once the set of p_J's are fixed, then all the matrix entries are determined and we end up with a column-wise permutation (implying an over all factor of ± 1 that we haven't determined) of the matrix with entries that for an initial matrix (7.3.91) look like

$$m_{i,j} = \begin{cases} \binom{i+j-3}{k-1} & \text{if } 1 \le i \le k-1, 1 \le j \le 2k \\ \binom{i+j-2}{k-1} & \text{if } i = k, 1 \le j \le 2k \\ (-1)^{i+j}\binom{i+j-3-k}{k-1} & \text{if } k+1 \le i \le 2k, 1 \le j \le 2k. \end{cases} \tag{7.3.116}$$

or for an initial matrix (7.3.92) look like

$$m_{i,j} = \begin{cases} \binom{i+j-3}{k-1} & \text{if } 1 \le i \le k, 1 \le j \le 2k \\ (-1)^{i+j}\binom{i+j-3-k}{k-1} & \text{if } k+1 \le i \le 2k-1, 1 \le j \le 2k \\ (-1)^{j-1}\binom{i+j-2-k}{k-1} & \text{if } i = 2k, 1 \le j \le 2k. \end{cases} \tag{7.3.117}$$

Note that we take $\binom{-1}{k-1} = 0$, so that the $(1,1)$ and $(k+1,1)$ entries of the matrix are 0. For example, the matrix defined by (7.3.116) is

$$\begin{pmatrix} \binom{-1}{k-1} & \binom{0}{k-1} & \cdots & \binom{k-2}{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{k-3}{k-1} & \binom{k-2}{k-1} & \cdots & \binom{2k-4}{k-1} \\ \binom{k-1}{k-1} & \binom{k}{k-1} & \cdots & \binom{2k-2}{k-1} \\ (-1)^k\binom{-1}{k-1} & (-1)^{k+1}\binom{0}{k-1} & \cdots & (-1)^{k+1}\binom{k-2}{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\binom{k-2}{k-1} & \binom{k-1}{k-1} & \cdots & (-1)^k\binom{2k-3}{k-1} \end{pmatrix}. \tag{7.3.118}$$

From (7.3.118), it is clear that the $(k,1)$st entry of the matrix (7.3.116) equals 1 while all the other entries in the first column are zero. Expanding the determinant of the above matrix along the first column thus gives:

$$(-1)^{k+1}\det\begin{pmatrix} \binom{0}{k-1} & \cdots & \binom{2k-2}{k-1} \\ \vdots & \ddots & \vdots \\ \binom{k-2}{k-1} & \cdots & \binom{3k-4}{k-1} \\ (-1)^{k+1}\binom{0}{k-1} & \cdots & (-1)^{k+1}\binom{2k-2}{k-1} \\ \vdots & \ddots & \vdots \\ -\binom{k-2}{k-1} & \cdots & -\binom{3k-4}{k-1} \\ \binom{k-1}{k-1} & \cdots & \binom{3k-3}{k-1} \end{pmatrix}. \tag{7.3.119}$$

Next, we notice that the new first column is zero except the last entry. Expanding along that column we get the following $(2k-2)\times(2k-2)$ determinant:

$$(-1)^{k+1}\det\begin{pmatrix} \binom{1}{k-1} & \cdots & \binom{2k-2}{k-1} \\ \vdots & \ddots & \vdots \\ \binom{k-1}{k-1} & \cdots & \binom{3k-4}{k-1} \\ (-1)^k\binom{1}{k-1} & \cdots & (-1)^{k+1}\binom{2k-2}{k-1} \\ \vdots & \ddots & \vdots \\ \binom{k-1}{k-1} & \cdots & -\binom{3k-4}{k-1} \end{pmatrix}. \tag{7.3.120}$$

We arrive at the above matrix also for an initial matrix of form (7.3.92), but in that case the only non-zero element of column 1 of (7.3.117) is the $+1$ in the $2k$th row, so expanding around that gives -1 times the resulting $(2k-1)\times(2k-1)$ minor. In the subsequent minor from this first expansion, the non-zero element of the new first column is a $+1$ in the kth row, so expanding round this element give a sign of $(-1)^{k+1}$. Thus in the (7.3.92) case we end up with the above determinant, but with the overall factor of $(-1)^{k+1}$ replaced by $(-1)^k$.

Working now from (7.3.120) we apply row reductions using the identity (7.3.50) and exactly the same procedure as in equations (7.3.47) to (7.3.56) so that we arrive at

$$(-1)^{k+1}\det\begin{pmatrix} \binom{1}{k-1} & \cdots & \binom{2k-2}{k-1} \\ \binom{1}{k-2} & \cdots & \binom{2k-2}{k-2} \\ \vdots & \ddots & \vdots \\ \binom{1}{1} & \cdots & \binom{2k-2}{1} \\ (-1)^k\binom{1}{k-1} & \cdots & (-1)^{k+1}\binom{2k-2}{k-1} \\ (-1)^{k+1}\binom{1}{k-2} & \cdots & (-1)^k\binom{2k-2}{k-2} \\ \vdots & \ddots & \vdots \\ \binom{1}{1} & \cdots & -\binom{2k-2}{1} \end{pmatrix}. \tag{7.3.121}$$

Example 7.3.10. *For example, for $k = 3$, the $(2k-2)\times(2k-2)$ determinant is*

$$\det\begin{pmatrix} \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} \\ \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} \\ -\binom{1}{2} & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2} \\ \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1} \end{pmatrix} = \det\begin{pmatrix} 0 & \frac{2}{2} & \frac{3\times 2}{2} & \frac{4\times3\times2}{2\times2} \\ \frac{1}{1} & \frac{2}{1} & \frac{3}{1} & \frac{4}{1} \\ 0 & \frac{2}{2} & -\frac{3\times2}{2} & \frac{4\times3\times2}{2\times2} \\ \frac{1}{1} & -\frac{2}{1} & \frac{3}{1} & -\frac{4}{1} \end{pmatrix}. \tag{7.3.122}$$

It is clear from both (7.3.121) and example 7.3.10 that we can factor j out of the jth column, and $\frac{1}{(k-i)}$ out of both the ith row and the $(i+k-1)$th row for $i = 1,\dots k-1$. This gives the following

$$(-1)^{k+1}\binom{2k-2}{k-1}\det\begin{pmatrix} \binom{0}{k-2} & \cdots & \binom{2k-3}{k-2} \\ \binom{0}{k-3} & \cdots & \binom{2k-3}{k-3} \\ \vdots & \ddots & \vdots \\ \binom{0}{0} & \cdots & \binom{2k-3}{0} \\ (-1)^k\binom{0}{k-2} & \cdots & (-1)^{k+1}\binom{2k-3}{k-2} \\ (-1)^{k+1}\binom{0}{k-3} & \cdots & (-1)^k\binom{2k-3}{k-3} \\ \vdots & \ddots & \vdots \\ \binom{0}{0} & \cdots & -\binom{2k-3}{0} \end{pmatrix} = \binom{2k-2}{k-1}2^{(k-1)^2}, \tag{7.3.123}$$

where the last step follows by pulling out $(-1)^k$ from each of the bottom $k-1$ rows (hence an even power of -1), and then applying (7.3.49) with $k-1$ rather than k.

Hence the determinant of the matrix with entries given in (7.3.116) is equal to

$$\binom{2k-2}{k-1}2^{(k-1)^2}. \tag{7.3.124}$$

The matrix with entries given in (7.3.117) is equal to

$$-\binom{2k-2}{k-1}2^{(k-1)^2}. \tag{7.3.125}$$

□

For a given $\frac{d^a}{dt_1^a}\frac{d^b}{dt_2^b}$ (so $\sum_{j=1}^{2k} m_j = a$ and $\sum_{j=1}^{2k} n_j = b$, with $a+b = 2k$), we now wish to determine the multiplicity of a given $(p_1,\dots,p_{2k})$, where as usual $p_j = m_j + n_j$.

For example, if $k = 2$, and $a = b = 2$, the vector $(p_1, p_2, p_3, p_4) = (1, 1, 1, 1)$ can arise in 24 ways.

We have these patterns for (m_1, m_2, m_3, m_4) (n_1, n_2, n_3, n_4):

(m_1, m_2, m_3, m_4)	(n_1, n_2, n_3, n_4)
$(1, 1, 0, 0)$	$(0, 0, 1, 1)$
$(1, 0, 1, 0)$	$(0, 1, 0, 1)$
$(1, 0, 0, 1)$	$(0, 1, 1, 0)$
$(0, 1, 1, 0)$	$(1, 0, 0, 1)$
$(0, 1, 0, 1)$	$(1, 0, 1, 0)$
$(0, 0, 1, 1)$	$(1, 1, 0, 0)$

However, each vector appearing here occurs twice when we carry out the partial derivative $\frac{d^a}{dt_1^a}\frac{d^b}{dt_2^b}$ on the matrix $\mathrm{M}^{(\mathrm{k})}$. For example, $(1, 1, 0, 0)$ gets counted twice, as we can differentiate the first column and then the second, or else the second column and then the first. All the following arguments hold equally well if instead of matrix $\mathrm{M}^{(\mathrm{k})}$ we use the matrix with the modified $2k$th row mentioned in lemma 7.3.3.

Generally, the number of occurrences of $(p_1, \ldots, p_{2k})$ obtained by applying $\frac{d^a}{dt_1^a}\frac{d^b}{dt_2^b}$ to $\det \mathrm{M}^{(\mathrm{k})}$, is equal to the coefficient of $c_1^{p_1} \cdots c_{2k}^{p_{2k}}$ in

$$(c_1 + \cdots + c_{2k})^a (c_1 + \cdots + c_{2k})^b = (c_1 + \cdots + c_{2k})^{2k}. \tag{7.3.126}$$

The resulting coefficient therefore equals the multinomial coefficient

$$\frac{(2k)!}{\prod_{j=1}^{2k} p_j!}. \tag{7.3.127}$$

Next we show that all of the ± 1 add up to 1. Recall that once the choice of $(p_1, \ldots, p_{2k})$, i.e. the number of times we differentiate with respect to t_1, t_2 in each column, are fixed, then the determinant of the corresponding matrix is equal up to ± 1, the determinant for $p_J = 1$ for all J. We showed that, if all the $m_J + n_J = p_J = 1$, then column degrees were, in order, $\{-1, 0, 1, \ldots, 2k-2\}$. However, for other values of $(p_1, \ldots, p_{2k})$ they will be some permutation of this ordered set. If it is an *even* permutation then the overall sign will be plus; if it is an *odd* permutation the sign will be minus.

Example 7.3.11. *As mentioned in the proof of proposition 7.3.9, any given permutation σ of the sequence $-1, \ldots, 2k-2$ completely determines the sequence of p_j. For example, when $k = 2$ in total there are 24 permutations of $\{-1, 0, 1, 2\}$. Of these there are 8 permutations each of which give a determinant value of ± 4. For these permutations, the lists of column degrees, D_j, the corresponding sequence of p_j that produce that permutation, along with their signs and multiplicities (from (7.3.127)) are listed below. The sum of the multiplicity times the sign gives $+1$ as desired.*

D_j				σ				p_j				*sign*	*mult*
-1	0	1	2	1	2	3	4	1	1	1	1	$+$	24
-1	0	2	1	1	2	4	3	1	1	0	2	$-$	12
-1	1	0	2	1	3	2	4	1	0	2	1	$-$	12
-1	1	2	0	1	3	4	2	1	0	0	3	$+$	4
0	-1	1	2	2	1	3	4	0	2	1	1	$-$	12
0	-1	2	1	2	1	4	3	0	2	0	2	$+$	6
0	1	-1	2	2	3	1	4	0	0	3	1	$+$	4
0	1	2	-1	2	3	4	1	0	0	0	4	-1	1

From the column of permutations denoted by σ, we see that the legal permutations of $\{1,2,3,4\}$ are $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ such that $\sigma_1 \leq 2; \sigma_2 \leq 3; \sigma_3 \leq 4;$. The reason for this is that $p_j \geq 0$ and

$$p_j = j - 1 - D_j = j + 1 - \sigma_j. \tag{7.3.128}$$

The remaining 16 permutation give a determinant of 0. One can see this by the following argument. Take for example, the permutation $\sigma = (4,3,2,1)$. This is not a valid choice (i.e. it yields a zero determinant) because, using (7.3.128)

$$(p_1, p_2, p_3, p_4) = (-2, 0, 2, 4) \tag{7.3.129}$$

contains a negative *value of $p_j := m_j + n_j$ (recall that m_j, n_j count the number of times a derivative is performed, c.f. (7.3.59)).*

Using example 7.3.11, we can see more generally that the legal permutations of $\{1, 2, \ldots, 2k\}$ are $\sigma_1, \sigma_2, \ldots, \sigma_{2k}$ with $\sigma_1 \leq 2; \sigma_2 \leq 3; \sigma_3 \leq 4; \ldots$. The reason is as described within example 7.3.11: the other permutations would force at least one $m_j + n_j = p_j < 0$.

So what we still have to prove is that

$$(2k)! \sum_{\substack{\sigma \in S_{2k} \\ \sigma_j \leq j+1}} \frac{\operatorname{sgn}(\sigma)}{\prod_{j=1}^{2k} (j + 1 - \sigma_j)!} = 1. \tag{7.3.130}$$

However, the sum in the above equation is the determinant of the $2k \times 2k$ matrix, denoted by C_{2k}, whose (i,j)th entry is $1/(j+1-i)!$ if $i \leq j+1$, and 0 otherwise (because of the restriction $\sigma_j \leq j+1$). For example, the matrix C_6 equals:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{24} & \frac{1}{120} & \frac{1}{720} \\ 1 & 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{24} & \frac{1}{120} \\ 0 & 1 & 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{24} \\ 0 & 0 & 1 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \tag{7.3.131}$$

We can put C_{2k} into triangular form by subtracting i times row i from row $i+1$, for $i = 1, \ldots, 2k-1$. Letting $l = j - i$, one can prove inductively, one row at a time, that the resulting entries are equal to $1/(l!(l+i))$ if $j \geq i$ and 0 otherwise. In particular the (i,i) diagonal entry equals $1/i$, and hence

$$\det C_{2k} = \frac{1}{(2k)!}, \tag{7.3.132}$$

thus establishing the identity (7.3.130).

We have thus proven that $\frac{d^a}{dt_1^a} \frac{d^b}{dt_2^b}$ (where $\sum_{j=1}^{2k} m_j = a$ and $\sum_{j=1}^{2k} n_j = b$) applied to the matrix $\mathrm{M}^{(\mathrm{k})}$, and setting $t_1 = t_2 = 0$, is equal to (7.3.124). Hence, the coefficient of $t_1^a t_2^b$ in $\det \mathrm{M}^{(\mathrm{k})}$ is equal to

$$\frac{1}{a!b!} \binom{2k-2}{k-1} 2^{(k-1)^2}. \tag{7.3.133}$$

If instead we consider the matrix $\mathrm{M}^{(2\mathrm{k})}$ (i.e. the matrix with the modified $2k$th row), then the coefficient of $t_1^a t_2^b$ in the determinant of that matrix is equal to, using (7.3.125),

$$-\frac{1}{a!b!}\binom{2k-2}{k-1}2^{(k-1)^2}. \tag{7.3.134}$$

Comparing coefficients, lemma 7.3.3 follows.

7.3.3 Comparison with exact formula for $k = 1$ and $k = 2$

Recall that Conrey and Snaith determined an exact formula for the logarithmic-derivative moments, see theorem 7.1.8. In this section, we use their result to determine the leading order of the first two moments (i.e. $k = 1$ and $k = 2$). We will demonstrate firstly that to show that their result agrees with the (7.1.13) in the appropriate limiting regime. Additionally, we emphasise that their alternative method does not provide a clear path to generally determining the leading order, hence showing the utility of our theorem.

In order to calculate the first log-derivative moment, we let $A = \{a\}$ and $B = \{\beta\}$ with $\mathrm{Re}(\alpha) > 0$ and $\mathrm{Re}(\beta) > 0$. Then the result of Conrey and Snaith gives that

$$J(A;B) = J^*(A;B) \tag{7.3.135}$$

where

$$J(\{\alpha\};\{\beta\}) := \int_{U(N)} (-e^{-\alpha})\frac{P_N'}{P_N}(A^*,-i\alpha)\,(-e^{-\beta})\frac{P_N'}{P_N}(A,-i\beta)dA, \tag{7.3.136}$$

and

$$\begin{aligned} J^*(\{\alpha\};\{\beta\}) &:= H_{\{\varnothing\},\{\varnothing\}}(\{\alpha\})H_{\{\varnothing\},\{\varnothing\}}(\{\beta\}) + H_{\{\varnothing\},\{\varnothing\}}(\{\alpha,\beta\}) \\ &\quad + e^{-N(\alpha+\beta)}z(\alpha+\beta)z(-\alpha-\beta) &(7.3.137)\\ &= 0 + \left(\frac{z'}{z}\right)'(\alpha+\beta) + e^{-N(\alpha+\beta)}z(\alpha+\beta)z(-\alpha-\beta), &(7.3.138)\end{aligned}$$

where $H_{S,T}$ is defined by (7.1.55).

Now let $\alpha = a/N$ and $\beta = b/N$ where $a, b \to 0$ as $N \to \infty$. It is useful for this calculation, and for the subsequent calculation for the second moment, to write down the behaviour of $z(x)$ and its derivatives for small x:

$$z(x) = \frac{1}{1-e^{-x}} = \frac{1}{x} + \frac{1}{2} + \frac{x}{12} - \frac{x^3}{720} + O(x^4), \tag{7.3.139}$$

$$\frac{z'(x)}{z(x)} = \frac{1}{1-e^{x}} = -\frac{1}{x} + \frac{1}{2} - \frac{x}{12} + \frac{x^3}{720} + O(x^4), \tag{7.3.140}$$

$$\left(\frac{z'(x)}{z(x)}\right)' = \frac{e^x}{(1-e^x)^2} = \frac{1}{x^2} - \frac{1}{12} + \frac{x^2}{240} + O(x^4). \tag{7.3.141}$$

Thus we have

$$\begin{aligned} J^*(\{\tfrac{a}{N}\};\{\tfrac{b}{N}\}) &= \left(\frac{1}{(\frac{a}{N}+\frac{b}{N})^2} + \frac{e^{-a-b}}{(\frac{a}{N}+\frac{b}{N})^2}\right)\left(1+O\left(\frac{a+b}{N}\right)\right) &(7.3.142)\\ &= \left(\frac{N^2}{(a+b)^2} - (1-a-b)\frac{N^2}{(a+b)^2}\right)(1+O(a+b)) &(7.3.143)\\ &= \left(\frac{N^2}{a+b}\right)(1+O(a+b)). &(7.3.144)\end{aligned}$$

So when $a = b$

$$J^*(\{\tfrac{a}{N}\};\{\tfrac{a}{N}\}) = \frac{N^2}{2a}(1 + O(a)). \tag{7.3.145}$$

Thus, by (7.3.135),

$$\int_{U(N)} \left|\frac{P_N'}{P_N}(A, i\alpha)\right|^2 dA = \frac{N^2}{2a}(1 + O(a)), \tag{7.3.146}$$

matching (7.1.13) at $k = 1$.

Now we consider the second moment, i.e. $k = 2$. Let $A = \{\alpha_1, \alpha_2\}$ and $B = \{\beta_1, \beta_2\}$. Then the theorem of Conrey and Snaith gives that $J(A; B) = J^*(A; B)$ where

$$\begin{aligned} &J(\{\alpha_1, \alpha_2\}; \{\beta_1, \beta_2\}) \\ &\quad= \int_{U(N)} e^{-\alpha_1-\alpha_2-\beta_1-\beta_2} \frac{P_N'}{P_N}(A^*, -i\alpha_1)\frac{P_N'}{P_N}(A, -i\alpha_2)\frac{P_N'}{P_N}(A, -i\beta_1)\frac{P_N'}{P_N}(A, -i\beta_2)dA \end{aligned} \tag{7.3.147}$$

$$\begin{aligned} &J^*(\{\alpha_1, \alpha_2\}; \{\beta_1, \beta_2\}) \\ &\quad= \sum_{\substack{S\subset A, T\subset B\\ |S|=|T|}} \left(e^{-N(\sum_{\hat\alpha\in S}\hat\alpha+\sum_{\hat\beta\in T}\hat\beta)} \frac{Z(S,T)Z(S^-,T^-)}{Z^\dagger(S,S^-)Z^\dagger(T,T^-)} \sum_{\substack{(A-S)+(B-T)\\ =U_1+\cdots+U_R\\ |U_r|\le 2}} \prod_{r=1}^{R} H_{S,T}(U_r)\right), \end{aligned} \tag{7.3.148}$$

where Z, Z^*, and $H_{S,T}$ are all defined in the statement of theorem 7.1.8.

We have to take a little care in setting all the αs and βs equal here because we will encounter factors of $\frac{z'}{z}(\alpha_2 - \alpha_1)$ and $\frac{z'}{z}(\beta_2 - \beta_1)$. These divergent terms will cancel as $\alpha_2 \to \alpha_1$ and $\beta_2 \to \beta_1$, but in order to control this we will set $\alpha_1 = \beta_1 = \alpha$ and $\alpha_2 = \beta_2 = \alpha + h$, with a view to letting $h \to 0$ later. This gives

$$\begin{aligned} J(\{\alpha,\alpha+h\};&\{\alpha,\alpha+h\}) \\ &= \left(\frac{z'}{z}\right)'(2\alpha)\left(\frac{z'}{z}\right)'(2\alpha+2h) + \left(\frac{z'}{z}\right)'(2\alpha+h)\left(\frac{z'}{z}\right)'(2\alpha+h) \\ &+ e^{-N(2\alpha)}z(2\alpha)z(-2\alpha) \\ &\qquad\times\left(\left(\frac{z'}{z}\right)'(2\alpha+2h) + \left(\frac{z'}{z}(h) - \frac{z'}{z}(2\alpha+h)\right)\left(\frac{z'}{z}(h) - \frac{z'}{z}(2\alpha+h)\right)\right) \\ &+ e^{-N(2\alpha+h)}z(2\alpha+h)z(-2\alpha-h) \\ &\qquad\times\left(\left(\frac{z'}{z}\right)'(2\alpha+h) + \left(\frac{z'}{z}(h) - \frac{z'}{z}(2\alpha+2h)\right)\left(\frac{z'}{z}(-h) - \frac{z'}{z}(2\alpha)\right)\right) \\ &+ e^{-N(2\alpha+h)}z(2\alpha+h)z(-2\alpha-h) \\ &\qquad\times\left(\left(\frac{z'}{z}\right)'(2\alpha+h) + \left(\frac{z'}{z}(-h) - \frac{z'}{z}(2\alpha)\right)\left(\frac{z'}{z}(h) - \frac{z'}{z}(2\alpha+2h)\right)\right) \\ &+ e^{-N(2\alpha+2h)}z(2\alpha+2h)z(-2\alpha-2h) \\ &\qquad\times\left(\left(\frac{z'}{z}\right)'(2\alpha) + \left(\frac{z'}{z}(-h) - \frac{z'}{z}(2\alpha+h)\right)\left(\frac{z'}{z}(-h) - \frac{z'}{z}(2\alpha+h)\right)\right) \\ &+ e^{-N(4\alpha+2h)}\frac{z(2\alpha)z^2(2\alpha+h)z(2\alpha+2h)z(-2\alpha)z^2(-2\alpha-h)z(-2\alpha-2h)}{(z(-h)z(h))^2}. \end{aligned} \tag{7.3.149}$$

The final term above is zero in the $h \to 0$ limit, but there are also terms of order h^{-2} and order h^{-1}. Using Mathematica to expand to order h^2 anything multiplying the divergent terms, we can confirm that all divergent terms cancel. In the $h \to 0$ limit we are left with

$$J(\{\alpha,\alpha\};\{\alpha,\alpha\}) = \lim_{h\to 0} J(\{\alpha,\alpha+h\};\{\alpha,\alpha+h\}) \tag{7.3.150}$$

$$= \frac{2e^{4\alpha} + e^{-2\alpha N}(-e^{2\alpha}N^2 + 2e^{4\alpha}N^2 - e^{6\alpha}N^2 - 2e^{4\alpha})}{(1-e^{2\alpha})^4}. \tag{7.3.151}$$

Now we scale $\alpha = a/N$ where $a \to 0$ as $N \to \infty$. Expanding the exponentials of the form $e^{ka/N}$, $k = 2, 4, 6$, in powers of a/N, we find that terms in the numerator of order N^2 and N cancel and we are left with:

$$J(\{\tfrac{a}{N},\tfrac{a}{N}\};\{\tfrac{a}{N},\tfrac{a}{N}\}) = \frac{\left(2 + e^{-2a}(-4a^2 + 32a^2 - 36a^2 - 2)\right)N^4}{16a^4}\left(1 + O\left(\frac{a}{N}\right)\right)$$

$$= \frac{(2 + (1-2a)(-8a^2-2))N^4}{16a^4}\,(1+O(a)) \tag{7.3.152}$$

$$= \frac{N^4}{4a^3}(1+O(a)). \tag{7.3.153}$$

Again, this is identical to (7.1.13) when $k = 2$.

We see that, with the help of Mathematica, the leading order term of theorem 7.1.8 can be extracted for a specific k. However, obtaining a formula for a general k seems very tricky from the complicated theorem 7.1.8, illustrating the value of the alternate method detailed in section 7.3 of this paper.

Bibliography

[1] G. Akemann, J. Baik, and P. Di Francesco. *The Oxford Handbook of Random Matrix Theory.* Oxford University Press, 2011.

[2] S. A. Altuğ, S. Bettin, I. Petrow, R. Rishikesh, and I. Whitehead. A recursion formula for moments of derivatives of random matrix polynomials. *Quarterly Journal of Mathematics*, 65(4):1111–1125, 2014.

[3] E. Alvarez and N. C. Snaith. Moments of the logarithmic derivative of characteristic polynomials from $SO(N)$ and $USp(2N)$. *arXiv preprint arXiv:2003.05906*, 2020.

[4] L.-P. Arguin. Extrema of Log-correlated Random Variables: Principles and Examples. In *Advances in Disordered Systems, Random Processes and Some Applications*, pages 166–204. Cambridge University Press, 2017.

[5] L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Communications in Mathematical Physics*, 349(2):703–751, 2017.

[6] L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Communications on Pure and Applied Mathematics*, 72(3):500–535, 2019.

[7] L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *The Annals of Applied Probability*, 27(1):178–215, 2017.

[8] J. Aru, E. Powell, and A. Sepúlveda. Critical Liouville measure as a limit of subcritical measures. *Electronic Communications in Probability*, 24, 2019.

[9] T. Assiotis, E. C. Bailey, and J. P. Keating. On the moments of the moments of the characteristic polynomials of Haar distributed symplectic and orthogonal matrices. *Annales de l'Institut Henri Poincaré D*, page To appear, 2019. arXiv preprint arXiv:1910.12576.

[10] T. Assiotis and J. P. Keating. Moments of moments of characteristic polynomials of random unitary matrices and lattice point counts. *Random Matrices: Theory and Applications*, page 2150019, 2020.

[11] A. Ayyer and I. Fischer. Bijective proofs of skew Schur polynomial factorizations. *Journal of Combinatorial Theory, Series A*, 174:105241, 2020.

[12] J. Baik, A. Borodin, P. Deift, and T. Suidan. A model for the bus system in Cuernavaca (Mexico). *Journal of Physics A: Mathematical and General*, 39(28):8965, 2006.

[13] J. Baik and E. M. Rains. Algebraic aspects of increasing subsequences. *Duke Mathematical Journal*, 109(1):1–65, 2001.

[14] E. C. Bailey, S. Bettin, G. Blower, J. B. Conrey, A. Prokhorov, M. O. Rubinstein, and N. C. Snaith. Mixed moments of characteristic polynomials of random unitary matrices. *Journal of Mathematical Physics*, 60(8):083509, 2019.

[15] E. C. Bailey and J. P. Keating. On the moments of the moments of the characteristic polynomials of random unitary matrices. *Communications in Mathematical Physics*, 371(2):689–726, 2019.

[16] E. C. Bailey and J. P. Keating. On the moments of the moments of $\zeta(1/2 + it)$. *arXiv preprint arXiv:2006.04503*, 2020.

[17] T. H. Baker and P. J. Forrester. Finite-N fluctuation formulas for random matrices. *Journal of Statistical Physics*, 88(5-6):1371–1386, 1997.

[18] E. Basor, P. Bleher, R. Buckingham, T. Grava, A. Its, E. Its, and J. P. Keating. A representation of joint moments of CUE characteristic polynomials in terms of painlevé functions. *Nonlinearity*, 32(10):4033, 2019.

[19] E. Basor, F. Ge, and M. O. Rubinstein. Some multidimensional integrals in number theory and connections with the Painlevé V equation. *Journal of Mathematical Physics*, 59(9):091404, 2018.

[20] N. Berestycki. An elementary approach to Gaussian multiplicative chaos. *Electronic communications in Probability*, 22, 2017.

[21] M. Biskup and O. Louidor. Extreme local extrema of two-dimensional discrete Gaussian free field. *Communications in Mathematical Physics*, 345(1):271–304, 2016.

[22] P. M. Bleher, P. Bleher, and A. Its. *Random matrix models and their applications*, volume 40. Cambridge university press, 2001.

[23] E. B. Bogomolny and J. P. Keating. Random matrix theory and the Riemann zeros. I. Three-and four-point correlations. *Nonlinearity*, 8(6):1115, 1995.

[24] E. B. Bogomolny and J. P. Keating. Random matrix theory and the Riemann zeros II: n-point correlations. *Nonlinearity*, 9(4):911, 1996.

[25] A. Bondarenko and K. Seip. Large greatest common divisor sums and extreme values of the Riemann zeta function. *Duke Mathematical Journal*, 166(9):1685–1701, 2017.

[26] A. Bondarenko and K. Seip. Extreme values of the Riemann zeta function and its argument. *Mathematische Annalen*, 372(3-4):999–1015, 2018.

[27] T. Bothner. On the origins of Riemann–Hilbert problems in mathematics. *arXiv preprint arXiv:2003.14374*, 2020.

[28] P. Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probability Theory and Related Fields*, 148(3-4):479–500, 2010.

[29] M. D. Bramson. Maximal displacement of branching Brownian motion. *Communications on Pure and Applied Mathematics*, 31(5):531–581, 1978.

[30] R. d. l. Bretèche and G. Tenenbaum. Sommes de Gál et applications. *Proceedings of the London Mathematical Society*, 119(1):104–134, 2019.

[31] H. Bui, B. Conrey, and M. Young. More than 41% of the zeros of the zeta function are on the critical line. *Acta Arithmetica*, 150(1):35–64, 2011.

[32] D. Bump. *Lie groups*. Springer, 2004.

[33] D. Bump and A. Gamburd. On the averages of characteristic polynomials from classical groups. *Communications in Mathematical Physics*, 265(1):227–274, 2006.

[34] D. Carpentier and P. Le Doussal. Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models. *Physical Review E*, 63(2):026110, 2001.

[35] R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Mathematical Journal*, 167(12):2243–2345, 2018.

[36] T. Claeys and I. Krasovsky. Toeplitz determinants with merging singularities. *Duke Mathematical Journal*, 164(15):2897–2987, 2015.

[37] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: I. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2040):20140313, 2015.

[38] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: II. In *Advances in the Theory of Numbers*, pages 75–85. Springer, 2015.

[39] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: III. *Indagationes Mathematicae*, 26(5):736–747, 2015.

[40] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: IV. *Research in Number Theory*, 2(1):24, 2016.

[41] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: V. *Proceedings of the London Mathematical Society*, 118(4):729–752, 2019.

[42] J. B. Conrey. The fourth moment of derivatives of the Riemann zeta-function. *The Quarterly Journal of Mathematics*, 39(1):21–36, 1988.

[43] J. B. Conrey. More than two fifths of the zeros of the Riemann zeta function are on the critical line. *J. Reine Angew. Math*, 399(1):1–26, 1989.

[44] J. B. Conrey and D. W. Farmer. Mean values of L-functions and symmetry. *International Mathematics Research Notices*, 2000(17):883–908, 2000.

[45] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. Autocorrelation of random matrix polynomials. *Communications in Mathematical Physics*, 237(3):365–395, 2003.

[46] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. Integral moments of l-functions. *Proceedings of the London Mathematical Society*, 91(1):33–104, 2005.

[47] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. Lower order terms in the full moment conjecture for the Riemann zeta function. *Journal of Number Theory*, 128(6):1516–1554, 2008.

[48] J. B. Conrey, D. W. Farmer, and M. R. Zirnbauer. Howe pairs, supersymmetry, and ratios of random characteristic polynomials for the unitary groups $u(n)$. *arXiv preprint math-ph/0511024*, 2005.

[49] J. B. Conrey, D. W. Farmer, and M. R. Zirnbauer. Autocorrelation of ratios of L-functions. *Communications in Number Theory and Physics*, 2(3):593–636, 2008.

[50] J. B. Conrey, P. J. Forrester, and N. C. Snaith. Averages of ratios of characteristic polynomials for the compact classical groups. *International Mathematics Research Notices*, 2005(7):397–431, 2005.

[51] J. B. Conrey and A. Ghosh. On mean values of the zeta-function, II. *Acta Arithmetica*, 52(4):367–371, 1989.

[52] J. B. Conrey and A. Ghosh. Mean values of the zeta-function, III. In *Proceedings of the Amalfi Conference on Analytic Number Theory, Università di Salerno*, 1992.

[53] J. B. Conrey and S. M. Gonek. High moments of the Riemann zeta-function. *Duke Mathematical Journal*, 107(3):577–604, 2001.

[54] J. B. Conrey, M. O. Rubinstein, and N. C. Snaith. Moments of the derivative of characteristic polynomials with an application to the Riemann zeta function. *Communications in Mathematical Physics*, 267(3):611–629, 2006.

[55] J. B. Conrey and N. C. Snaith. Correlations of eigenvalues and Riemann zeros. *Communications in Number Theory and Physics*, 2(3):477–536, 2008.

[56] I. Cooper. *Random matrix thoery and families of elliptic curve L-functions*. PhD thesis, University of Bristol, 2018.

[57] L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007.

[58] P.-O. Dehaye. Joint moments of derivatives of characteristic polynomials. *Algebra & Number Theory*, 2(1):31–68, 2008.

[59] P.-O. Dehaye. A note on moments of derivatives of characteristic polynomials. *arXiv preprint arXiv:1001.3820*, 2010.

[60] P. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, volume 3. American Mathematical Soc., 1999.

[61] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation. *Annals of Mathematics*, 137(2):295–368, 1993.

[62] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Physical Review B*, 24(5):2613, 1981.

[63] B. Derrida. A generalization of the random energy model which includes correlations between energies. *Journal de Physique Lettres*, 46(9):401–407, 1985.

[64] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *Journal of Applied Probability*, 31(A):49–62, 1994.

[65] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated Gaussian fields. *The Annals of Probability*, 45(6A):3886–3928, 2017.

[66] E. Dueñez, D. W. Farmer, S. Froehlich, C. P. Hughes, F. Mezzadri, and T. Phan. Roots of the derivative of the Riemann-zeta function and of characteristic polynomials. *Nonlinearity*, 23(10):2599, 2010.

[67] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Critical Gaussian multiplicative chaos: convergence of the derivative martingale. *The Annals of Probability*, 42(5):1769–1808, 2014.

[68] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Renormalization of critical Gaussian multiplicative chaos and KPZ relation. *Communications in Mathematical Physics*, 330(1):283–330, 2014.

[69] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Log-correlated Gaussian fields: an overview. In *Geometry, analysis and probability*, pages 191–216. Springer, 2017.

[70] F. J. Dyson. Statistical theory of the energy levels of complex systems. I. *Journal of Mathematical Physics*, 3(1):140–156, 157–165, 166–175, 1962.

[71] H. M. Edwards. *Riemann's Zeta Function.* Academic press, 1974.

[72] B. Fahs. Uniform asymptotics of Toeplitz determinants with Fisher–Hartwig singularities. *arXiv preprint arXiv:1909.07362*, 2019.

[73] D. W. Farmer, S. M. Gonek, and C. P. Hughes. The maximum size of l-functions. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2007(609):215–236, 2007.

[74] P. J. Forrester and N. E. Frankel. Applications and generalizations of Fisher–Hartwig asymptotics. *Journal of Mathematical Physics*, 45(5):2003–2028, 2004.

[75] P. J. Forrester and J. P. Keating. Singularity dominated strong fluctuations for some random matrix averages. *Communications in Mathematical Physics*, 250(1):119–131, 2004.

[76] P. J. Forrester and N. S. Witte. Application of the τ-function theory of Painlevé equations to random matrices: $\mathrm{P_V}$, $\mathrm{P_{III}}$, the LUE, JUE, and CUE. *Communications on Pure and Applied Mathematics*, 55(6):679–727, 2002.

[77] P. J. Forrester and N. S. Witte. Boundary conditions associated with the painlevé III$'$ and V evaluations of some random matrix averages. *Journal of Physics A: Mathematical and General*, 39(28):8983, 2006.

[78] Y. V. Fyodorov and J.-P. Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *Journal of Physics A: Mathematical and Theoretical*, 41(37):372001, 2008.

[79] Y. V. Fyodorov and P. L. Doussal. Statistics of extremes in eigenvalue-counting staircases. *arXiv preprint arXiv:2001.04135*, 2020.

[80] Y. V. Fyodorov, S. Gnutzmann, and J. P. Keating. Extreme values of CUE characteristic polynomials: a numerical study. *Journal of Physics A: Mathematical and Theoretical*, 51(46):464001, 2018.

[81] Y. V. Fyodorov, G. A. Hiary, and J. P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. *Physical Review Letters*, 108(17):170601, 2012.

[82] Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2007):20120503, 2014.

[83] Y. V. Fyodorov, B. A. Khoruzhenko, and N. J. Simm. Fractional Brownian motion with Hurst index $h = 0$ and the Gaussian unitary ensemble. *The Annals of Probability*, 44(4):2980–3031, 2016.

[84] Y. V. Fyodorov and P. Le Doussal. Moments of the position of the maximum for GUE characteristic polynomials and for log-correlated Gaussian processes. *Journal of Statistical Physics*, 164(1):190–240, 2016.

[85] Y. V. Fyodorov and N. J. Simm. On the distribution of the maximum value of the characteristic polynomial of GUE random matrices. *Nonlinearity*, 29(9):2837, 2016.

[86] B. Gambier. *Sur les équations différentielles du second ordre et du premier degré dont l'intégrale générale est à points critiques fixes.* PhD thesis, Paris, 1909.

[87] B. Gambier. Sur les équations différentielles du second ordre et du premier degré dont l'intégrale générale est à points critiques fixes. *Acta Mathematica*, 33:1–55, 1910.

[88] B. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, pages 423–453, 1943.

[89] V. Gorin and M. Rahman. Random sorting networks: local statistics via random matrix laws. *Probability Theory and Related Fields*, 175(1-2):45–96, 2019.

[90] V. I. Gromak, I. Laine, and S. Shimomura. *Painlevé differential equations in the complex plane*, volume 28. Walter de Gruyter, 2008.

[91] G. H. Hardy and J. E. Littlewood. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. *Acta Mathematica*, 41:119–196, 1918.

[92] A. J. Harper. A note on the maximum of the Riemann zeta function, and log–correlated random variables. *arXiv preprint arXiv:1304.0677*, 2013.

[93] A. J. Harper. Sharp conditional bounds for moments of the Riemann zeta function. *arXiv preprint arXiv:1305.4618*, 2013.

[94] A. J. Harper. On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture. *arXiv preprint arXiv:1906.05783*, 2019.

[95] A. J. Harper. The Riemann zeta function in short intervals [after Najnudel, and Arguin, Belius, Bourgade, Radziwiłł, and Soundararajan]. *arXiv preprint arXiv:1904.08204*, 2019.

[96] D. R. Heath-Brown. Fractional Moments of the Riemann Zeta-Function. *Journal of the London Mathematical Society*, 2(1):65–78, 1981.

[97] D. A. Hejhal. On the triple correlation of zeros of the zeta function. *International Mathematics Research Notices*, 1994(7):293–302, 1994.

[98] C. P. Hughes. *On the characteristic polynomial of a random unitary matrix and the Riemann zeta function.* PhD thesis, University of Bristol, 2001.

[99] C. P. Hughes, J. P. Keating, and N. O'Connell. Random matrix theory and the derivative of the Riemann zeta function. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 456(2003):2611–2627, 2000.

[100] C. P. Hughes, J. P. Keating, and N. O'Connell. On the Characteristic Polynomial of a Random Unitary matrix. *Communications in Mathematical Physics*, 220(2):429–451, 2001.

[101] E. L. Ince. *Ordinary Differential Equations.* Dover Publications, 1956.

[102] A. E. Ingham. Mean-value theorems in the theory of the Riemann zeta-function. *Proceedings of the London Mathematical Society*, 2(1):273–300, 1926.

[103] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53. American Mathematical Soc., 2004.

[104] H. Iwaniec and P. Sarnak. Perspectives on the analytic theory of L-functions. In *Visions in Mathematics*, pages 705–741. Springer, 2000.

[105] K. Iwasaki, H. Kimura, S. Shimemura, and M. Yoshida. *From Gauss to Painlevé: a modern theory of special functions*, volume 16. Springer Science & Business Media, 2013.

[106] K. Johansson. On random matrices from the compact classical groups. *Annals of Mathematics*, pages 519–545, 1997.

[107] M. Jutila. On the Mean Value of $L(1/2, \chi)$ for Real Characters. *Analysis*, 1(2):149–161, 1981.

[108] J.-P. Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985.

[109] N. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bulletin of the American Mathematical Society*, 36(1):1–26, 1999.

[110] N. M. Katz and P. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45. American Mathematical Soc., 1999.

[111] J. P. Keating. Random Matrices and Number Theory: Some Recent Themes. In *Stochastic Processes and Random Matrices: Lecture Notes of the Les Houches Summer School*, volume 104, pages 348–381. Oxford University Press, 2017.

[112] J. P. Keating and B. E. Odgers. Symmetry transitions in random matrix theory & L-functions. *Communications in Mathematical Physics*, 281(2):499–528, 2008.

[113] J. P. Keating, B. Rodgers, E. Roditty-Gershon, and Z. Rudnick. Sums of divisor functions in $\mathbb{F}_q[t]$ and matrix integrals. *Mathematische Zeitschrift*, 288(1-2):167–198, 2018.

[114] J. P. Keating and E. Scott. Undergraduate summer research project. Unpublished, 2015.

[115] J. P. Keating and N. C. Snaith. Random matrix theory and L-functions at $s = 1/2$. *Communications in Mathematical Physics*, 214(1):91–100, 2000.

[116] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Communications in Mathematical Physics*, 214(1):57–89, 2000.

[117] N. Kistler. Derrida's random energy models. From spin glasses to the extremes of correlated random fields. *Springer Lecture Notes in Mathematics*, 2143, 2015.

[118] M. Krbálek and P. Šeba. The statistical properties of the city transport in Cuernavaca (Mexico) and random matrix ensembles. *Journal of Physics A: Mathematical and General*, 33(26):L229, 2000.

[119] A. Kundu, S. N. Majumdar, and G. Schehr. Exact distributions of the number of distinct and common sites visited by N independent random walkers. *Physical Review Letters*, 110(22):220602, 2013.

[120] A. Kupiainen, R. Rhodes, and V. Vargas. Integrability of Liouville theory: proof of the DOZZ Formula. *Annals of Mathematics*, 191(1):81–166, 2020.

[121] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer Science & Business Media, 2012.

[122] I. G. Macdonald. *Symmetric functions and Hall polynomials.* Oxford university press, 1998.

[123] A. Mason and N. Snaith. *Orthogonal and symplectic n-level densities*, volume 251. American Mathematical Society, 2018.

[124] M. L. Mehta. *Random matrices.* Elsevier, 2004.

[125] F. Mezzadri. Random matrix theory and the zeros of $\zeta'(s)$. *Journal of Physics A: Mathematical and General*, 36(12):2945, 2003.

[126] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Proc. Symp. Pure Math*, volume 24, pages 181–193, 1973.

[127] J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probability Theory and Related Fields*, 172(1-2):387–452, 2018.

[128] M. Nikula, E. Saksman, and C. Webb. Multiplicative chaos and the characteristic polynomial of the CUE: the L^1–phase. *Transactions of the American Mathematical Society*, 2020.

[129] A. M. Odlyzko. The 10^{20}th zero of the Riemann zeta function and 70 million of its neighbors. Preprint, 1989.

[130] P. Painlevé. Mémoire sur les équations différentielles dont l'intégrale générale est uniforme. *Bulletin de la Société Mathématique de France*, 28:201–261, 1900.

[131] P. Painlevé. Sur les équations différentielles du second ordre et d'ordre supérieur dont l'intégrale générale est uniforme. *Acta Mathematica*, 25:1–85, 1902.

[132] E. Paquette and O. Zeitouni. The maximum of the CUE field. *International Mathematics Research Notices*, 2018(16):5028–5119, 2017.

[133] E. Powell. Critical Gaussian chaos: convergence and uniqueness in the derivative normalisation. *Electronic Journal of Probability*, 23, 2018.

[134] K. Pratt, N. Robles, A. Zaharescu, and D. Zeindler. More than five-twelfths of the zeros of ζ are on the critical line. *Research in the Mathematical Sciences*, 7(2):1–74, 2020.

[135] R. A. Proctor. Young tableaux, Gelfand patterns, and branching rules for classical groups. *Journal of Algebra*, 164(2):299–360, 1994.

[136] K. Ramachandra. Some remarks on the mean value of the Riemann zeta-function and other Dirichlet series-II. *Hardy-Ramanujan Journal*, 3, 1980.

[137] G. Remy. The Fyodorov–Bouchaud formula and Liouville conformal field theory. *Duke Mathematical Journal*, 2019.

[138] R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probability Surveys*, 11, 2014.

[139] R. Rhodes and V. Vargas. Gaussian Multiplicative Chaos and Liouville Quantum Gravity. In *Stochastic Processes and Random Matrices: Lecture Notes of the Les Houches Summer School*, volume 104, pages 548–577. Oxford University Press, 2017.

[140] B. Riemann. Über die Anzahl der Primzahlen unter einer gegebenen Grösse. *Ges. Math. Werke und Wissenschaftlicher Nachlaß*, 2:145–155, 1859.

[141] Z. Rudnick and P. Sarnak. Zeros of principal L-functions and random matrix theory. *Duke Mathematical Journal*, 81(2):269–322, 1996.

[142] E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *arXiv preprint arXiv:1609.00027*, 2016.

[143] W. M. Schmidt. Northcott's theorem on heights II. the quadratic case. *Acta Arithmetica*, 70(4):343–375, 1995.

[144] A. Selberg. Contributions to the theory of the Riemann zeta-function. *Arch. for Math. og Naturv. B*, 48(5), 1946.

[145] K. Soundararajan. Nonvanishing of quadratic Dirichlet L-functions at $s = 1/2$. *Annals of Mathematics*, 152:447–488, 2000.

[146] K. Soundararajan. Moments of the Riemann zeta function. *Annals of Mathematics*, 170:981–993, 2009.

[147] R. P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, 1999.

[148] R. P. Stanley. *Enumerative Combinatorics*, volume 1. Cambridge University Press, 2 edition, 2011.

[149] E. Subag and O. Zeitouni. Freezing and decorated Poisson point processes. *Communications in Mathematical Physics*, 337(1):55–92, 2015.

[150] G. Szegö. Ein Grenzwertsatz über die Toeplitzschen Determinanten einer reellen positiven Funktion. *Mathematische Annalen*, 76(4):490–503, 1915.

[151] G. Szegö. *Orthogonal polynomials*, volume 23. American Mathematical Soc., 1939.

[152] G. Szegö. On certain Hermitian forms associated with the Fourier series of a positive function. *Comm. Sém. Math. Univ. Lund [Medd. Lunds Univ. Mat. Sem.]*, 1952(Tome Supplémentaire):228–238, 1952.

[153] The LMFDB Collaboration. The L-functions and modular forms database. `http://www.lmfdb.org`, 2019. [Online; accessed 24 February 2020].

[154] E. C. Titchmarsh. *The theory of the Riemann zeta-function.* Oxford University Press, 2 edition, 1986.

[155] C. Webb. The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos – the L^2–phase. *Electronic Journal of Probability*, 20, 2015.

[156] H. Weyl. *The classical groups: their invariants and representations*, volume 45. Princeton University Press, 1946.

[157] H. Widom. Toeplitz determinants with singular generating functions. *American Journal of Mathematics*, 95(2):333–383, 1973.

[158] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions I. In *The Collected Works of Eugene Paul Wigner*, pages 524–540. Springer, 1993.

[159] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions II. In *The Collected Works of Eugene Paul Wigner*, pages 541–545. Springer, 1993.

[160] B. Winn. Derivative moments for characteristic polynomials from the CUE. *Communications in Mathematical Physics*, 315(2):531–562, 2012.