# RANDOM MATRIX THEORY AND ZETA FUNCTIONS

By

Nina Claire Snaith

School of Mathematics

June 2000

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science

# Abstract

Functions with zeros displaying the statistics of the eigenvalues of random matrices occur both in number theory, for example the Riemann zeta function and $L$-functions, and in the semiclassical study of systems which have chaotic classical counterparts; the spectral determinant of such a system is an example of this type of function.

In this thesis we study the characteristic polynomial $Z(U,\theta)$ of an $N \times N$ matrix $U$ belonging to an ensemble of random matrices. We determine the mean values of $|Z|^s$ and $(Z/Z^*)^s$ averaged over such an ensemble. From these we show that the asymptotic distributions, as $N \to \infty$, of the real and imaginary parts of the logarithm of $Z$ are independent and Gaussian.

Our random matrix calculations are compared with analytical results and numerical computations for the number theoretical and spectral functions mentioned above. It is found that in the appropriate limit the logarithm of such functions have a distribution which, when normalized to have mean zero and unit variance, matches that of the logarithm of the random matrix function $Z$. This implies that in this limit the value distribution of the logarithm of these functions depends only on the random matrix distribution of their zeros. We also formulate a conjecture for the manner in which the mean values of the Riemann zeta function, $L$-functions and spectral determinants divide, again in the limit equivalent to large matrix size $N$, into a product of a factor which derives purely from random matrix theory, and a component which is specific to the particular function under scrutiny.

# Acknowledgements

If these acknowledgements are to appear in chronological order, then this first thank you is due to the patient and enthusiastic teaching of David Goodings and Julie Lefebvre who inspired my initial interest in the field of quantum chaos.

The term *super*viser was never applied more aptly to anyone than to Jon Keating, who is nothing if not super. Thanks, Jon, for your perpetual encouragement and for putting up with me for four years - green ink and all. I couldn't have asked for a better adviser.

Many thanks also to all members of the Bristol quantum chaos group, past and present, who have been so helpful and created such a pleasant, comfortable atmosphere to work in. More than a word is owed to my good friends and fellow graduate students in the department who kept me company during my time here.

Perhaps my first acknowledgements should have been to Mum and Dad for passing on the mathematical genes in the first place, and also to Anna, Dan and Francesco for lots of love and laughs.

The numerical data for the Riemann zeta function was kindly given to me by Andrew Odlyzko, to whom I am very grateful, as I am to Joachim Weber, creator of the program which produced data for the kicked tops.

Finally I'd like to thank the Natural Sciences and Engineering Research Council of Canada and the Canadian Federation of University Women for generously supporting my studies.

# Author's Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and in no way represent those of the University of Bristol. The thesis has not been presented to any other University for examination either in the United Kingdom or overseas.

Nina Claire Snaith

Date: 7th June 2000

# Contents

# List of Figures

# Chapter 1

# Setting the Scene

Here is a question: what can be said about the value distribution of a function if all that is known is that its zeros show a certain behaviour, and in particular display the statistics characteristic of the eigenvalues of an ensemble of random matrices? Further, if we have a particular function in mind, where do the features specific to that given function enter the value distribution and its moments and how are they reconciled with the behaviour common to all functions sharing the same zero statistics?

These questions, posed to serve the interests of quantum chaos, but eventually crossing over into the realms of number theory, will be the subject of the ensuing chapters. This chapter will lay the groundwork for the pursuit of the answers, while in the following one we will launch into the calculation in earnest, studying the value distribution of a general function with zero statistics belonging to a random matrix ensemble. In Chapter 3 we compare the results of Chapter 2 with the mean values of the Riemann zeta function, a much-studied example of a function with the appropriate zero statistics, extending this to other L-functions in Chapter 4. In the final chapter we retire again to the domain of quantum chaos to investigate the spectral determinant of classically chaotic systems.

## 1.1 Quantum chaos and periodic orbit theory

Quantum chaos is the name by which the investigation of the quantum analogues of classically chaotic systems has come to be known. In a quantum system it is not possible to define chaotic versus regular behaviour as one can within the realms of classical mechanics; for a start phase space trajectories do not exist, so the definition that nearby paths diverge exponentially quickly has no meaning. However, in the semiclassical limit, as a quantum system approaches its classical counterpart, one might expect the chaoticity (or integrability) of the classical version to be somehow reflected in the eigenenergies and wavefunctions of the semiclassical system. Later in this chapter we will see that there does indeed seem to be such a signature of chaos in that the eigenvalues for classically integrable systems are uncorrelated while those corresponding to chaotic classical behaviour show the distinctive statistics of the eigenvalues of large random matrices. For the moment, however, we will explore the tools of the semiclassical trade.

The backbone of such semiclassical studies are formulae, asymptotic in the limit as the quantum parameter $h$ (Planck's constant) tends to zero, which relate the quantum eigenvalues to strictly classical properties of the analogous classical system. This limit $h = 2\pi\hbar \to 0$ is called the semiclassical limit.

Quantum eigenvalues, $E_n$, and wavefunctions, $\psi_n(\mathbf{r})$, arise as the solutions to the time-independent Schrödinger equation

$$H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}). \tag{1.1.1}$$

$H$ is the Hamiltonian operator, and for example for motion in a scalar potential $V(\mathbf{r},t)$ with no other influences $H = \frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)$. It should be noted that the $E_n$ are real as they are the eigenvalues of the Hermitian operator, $H$. The eigenvalue equation (1.1.1) derives from the time-dependent Schödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = H\Psi(\mathbf{r},t) \tag{1.1.2}$$

by separation of the space and time variables if the potential $V(\mathbf{r},t)$ does not depend on time.

For bound systems, the above eigenvalues are discrete, and a great deal of study has been made of the statistics of their values. A useful quantity in this work is the density of states,

$$d(E) = \sum_{n=1}^{\infty}\delta(E - E_n), \tag{1.1.3}$$

where $\delta$ is the Dirac delta distribution. As the density of states picks out the positions of the eigenvalues in this way, it is the starting point for the study of eigenvalue statistics.

A key step in the progress of semiclassical asymptotics was Gutzwiller's development in the 1960's and 70's of the trace formula [Gut67, Gut69, Gut70, Gut71], soon after explored by Balian and Bloch [BB70, BB71, BB72, BB74]. This formula provides an asymptotic (for small $\hbar$) expression for the density of states and it involves *only* information about the periodic orbits of the classical system.

The route to the trace formula involves the Green function $G(\mathbf{q}'',\mathbf{q}',E)$ (here written in three dimensions), which is the solution to the equation

$$(H - E)G(\mathbf{q}'',\mathbf{q}',E) = \delta(\mathbf{q}'' - \mathbf{q}'). \tag{1.1.4}$$

This function can be written as

$$G(\mathbf{q}'',\mathbf{q}',E) = \sum_{n=1}^{\infty}\frac{\psi_n(\mathbf{q}'')\psi_n^*(\mathbf{q}')}{E_n - E}, \tag{1.1.5}$$

where the $\psi_n$ and $E_n$ are the same eigenfunctions and eigenvalues belonging to $H$ as in (1.1.1). Thanks to the orthonormality of the wavefunctions $\psi_n$, we see that setting $\mathbf{q}'' = \mathbf{q}'$ and integrating over all space, we have

$$\int d^3\mathbf{q}G(\mathbf{q}, \mathbf{q}, E) = \sum_{n=1}^{\infty} \frac{1}{E_n - E}. \tag{1.1.6}$$

Thus we see that we can create the delta functions at the energy levels in the density of states by allowing for imaginary energy:

$$d(E) = \frac{1}{\pi} \lim_{\epsilon \to 0} \operatorname{Im} \int d^3\mathbf{q}G(\mathbf{q}, \mathbf{q}, E + i\epsilon). \tag{1.1.7}$$

Now, the Green function has the physical interpretation that it represents the amplitude at $\mathbf{q}''$ of waves with energy $E$ which are continually being emitted at $\mathbf{q}'$. Since quantum mechanics is a wave theory, at the short-wave limit of which lies classical mechanics, it is not surprising that in the semiclassical limit (that is, in the approach to the classical regime) the Green function can be written just as a sum over all the classical routes which could be taken by a particle starting at $\mathbf{q}'$ with energy $E$ in order to end up at $\mathbf{q}''$. The evolution of the waves emanating from $\mathbf{q}'$ are related more and more closely to the classical trajectories as the semiclassical limit is approached.

In order to obtain a semiclassical approximation to the density of states, this semiclassical Green function is inserted into (1.1.7). Since it is the trace of the Green function which appears in that equation, the sum is now over *closed* classical trajectories from $\mathbf{q}$ back to $\mathbf{q}$. It turns out that the trajectories which shrink to zero length as $\mathbf{q}''$ and $\mathbf{q}'$ approach $\mathbf{q}$ (as we take the trace) will contribute to a smooth mean density of states, $\overline{d}(E)$. In addition to this, performing the integral in (1.1.7) by the method of stationary phase picks out just the periodic orbits (that is, those which leave and return to $\mathbf{q}$ with the same momentum), and each of these contribute an oscillatory term to the density of states. Thus we have a sum over the primitive periodic orbits $p$ and their repetitions (represented by the sum over $k$):

$$d(E) \sim \overline{d}(E) + \frac{2}{\hbar^{1+\eta}} \sum_p \sum_{k=1}^{\infty} A_{p,k} \cos\left[\frac{k}{\hbar}(S_p + \mu_p)\right], \tag{1.1.8}$$

where $\eta = 0$ if the system being studied is classically chaotic, and $\eta = (n-1)/2$ for a system with $n$ degrees of freedom which has integrable classical dynamics.

All the quantities on the right hand side of this equation are classical - not dependent on $\hbar$. $S_p$ is the action of the periodic orbit $p$, defined by

$$S_p(E) = \oint \mathbf{p} \cdot d\mathbf{q}; \tag{1.1.9}$$

that is, the integral of the momentum along the orbit. $A_{p,k}$ is an amplitude related to the stability of the $k$th repetition of the orbit $p$ and is given by

$$A_{p,k} = \frac{T_p}{2\pi[\det(M_p^k - I)]^{1/2}}, \tag{1.1.10}$$

$T_p(E) = dS_p/dE$ being the period of the orbit and $M_p$ being the monodromy matrix which describes flow linearized around the orbit. The Maslov index $\mu_p$ counts points along the orbit where focusing of neighbouring trajectories occurs, and from here on we will just incorporate it into the action variable, $S_p$.

For the mean density of states, we have

$$\overline{d}(E) \sim \frac{1}{h^n} \frac{d\Omega(E)}{dE}, \tag{1.1.11}$$

where $\Omega(E)$ is the volume in phase space with energy less than $E$. This follows from the general rule that for a system with $n$ degrees of freedom there exists about one quantum state for each volume $h^n$ in the energy shell $\Omega(E)$, and can be seen more easily if we consider the spectral staircase

$$\mathcal{N}(E) = \int_0^E d(x)dx, \tag{1.1.12}$$

which consists of a unit step at the position of each eigenvalue. $\mathcal{N}(E)$ therefore represents the number of $E_n$ such that $E_n \leq E$. Semiclassically (once more this means looking at just the leading-order term as $\hbar \to 0$) we can write

$$\mathcal{N}(E) \sim \overline{\mathcal{N}}(E) + \frac{2}{\hbar^\eta} \sum_p \sum_{k=1}^{\infty} \frac{A_{p,k}}{kT_p} \sin(kS_p/\hbar), \tag{1.1.13}$$

where $\overline{\mathcal{N}}(E) = \Omega(E)/h^n$. We see that the derivative of $\overline{\mathcal{N}}(E)$ gives the expression for $\overline{d}(E)$, as expected.

As we are particularly interested for the purposes of this thesis in functions which have zeros at the positions of the eigenvalues, we will also consider the spectral determinant, which is constructed to be just such a function:

$$\Delta(E) \equiv \det[A(E,H)(E-H)] = \prod_j [A(E,E_j)(E-E_j)], \tag{1.1.14}$$

where $A$ has no real zeros and is included so as to make the product converge. For classically chaotic systems we have semiclassically [BK90]

$$\Delta(E) \sim B(E) \exp(-i\pi\overline{\mathcal{N}}(E)) \prod_p \exp\left( - \sum_{k=1}^{\infty} \frac{\exp(ikS_p/\hbar)}{k\sqrt{|\det(M_p^k - 1)|}} \right), \tag{1.1.15}$$

where $B(E)$ is a real function which will not worry us as it has no zeros.

As we mentioned at the beginning of this section, one of the goals of quantum chaology was to attempt to detect in the statistical distribution of a set of energy eigenvalues, the semiclassical positions of which are described by the above formulae, an indication of whether the corresponding classical system behaves chaoticly or integrably. The answer, suggested by Berry and Tabor [BT77] and later examined in depth by Bohigas, Giannoni and Schmit [BGS84, BG84], involves the theory of random matrices.

## 1.2 Random Matrix Theory

Random matrix theory had its beginnings in the 1950's and 60's in the field of nuclear physics. Physicists were attempting to create a model which would predict the excitation states of atomic nuclei. To do this, it would be necessary to work out the quantum Hamiltonians, $H$, of the particular nuclear systems under consideration and then solve Schrödinger's equation (1.1.1) to find its energy eigenvalues. This proved to be very difficult due to the complexity of the interactions in a nuclear system, and only the lowest few energy levels could be modelled accurately. In fact, the Hamiltonians are so complicated that, when written as matrices with respect to an appropriate basis, the matrix elements show essentially no correlations; they look like random numbers.

When the problem of predicting specific energy levels proved to be intractable, attention turned to considering the statistics of the distribution of the energy levels. It was Wigner who had the idea that since the matrix Hamiltonian of each individual nucleus is so complicated and there are no obvious correlations between the various matrix elements, perhaps the best way to approximate the energy level statistics of a given nucleus is by the average statistics of a collection of matrices all sharing some kind of symmetry, but with otherwise uncorrelated, random entries. This is the basis of random matrix theory.

However, such an idea is not limited to the study of nuclear physics. If a quantum system and the related classical counterpart are considered, and the classical version displays chaotic behaviour, then Wigner's conjecture is likely to be useful here as well. In this case, the effective randomness of the Hamiltonian matrix elements arises from the complexity of the wavefunctions of the quantum system, since in the semiclassical limit they inherit their behaviour from the chaotic trajectories of the classical system. The lack of correlations in the matrix elements implies that, as in the nuclear case, it might be expected that the energy level statistics of the quantum system would be predicted by the average statistics of a number of random matrices, just as an average over all possible states provides general information about a system in a particular state in ordinary statistical mechanics.

It should be noted that whenever we discuss the statistics of energy levels being predicted by random matrix theory, we will always be implying that the system under consideration has been desymmetrized. If a system has any exact integrals of the motion or discrete symmetries then the corresponding quantum numbers retain their meaning and so allow the Hamiltonian matrix to be written as a series of blocks along the diagonal with zeros elsewhere. As the eigenvalues of the different blocks are uncorrelated (possessing different conserved quantum numbers, for instance those corresponding to total spin or parity) their *combined* spectrum cannot be expected to show the distinctive statistics which are found in the ensembles of random matrices discussed below. Thus when we speak of a desymmetrized system, we mean that we are considering just one of the blocks on the diagonal of the full Hamiltonian.

### 1.2.1 The Gaussian ensembles

The problem next is to determine how to group the matrices according to their particular symmetries. These groupings are prescribed by the symmetries occurring in physical systems and all that follows is explained in much greater detail in Mehta's book on random matrix theory [Meh91]. A very readable review of random matrices can be found in [Boh91].

As a first example, the Hamiltonian matrix of any system with integral total angular momentum and with time-reversal symmetry can be written as a symmetric matrix. Since all Hamiltonians are Hermitian, these symmetric matrices must also be real. Systems with integral *or* half-integral total angular momentum, but which have time-reversal invariance and rotational symmetry can also be represented by real, symmetric matrices. Thus it was proposed that an average over the eigenvalue statistics of an ensemble of real, symmetric matrices might be equivalent to the distribution of the eigenvalues of any one particular real, symmetric, random matrix, and thus predict the distribution of the energy levels of a physical system with the symmetries mentioned in this paragraph: time-reversal symmetry, an integer total angular momentum or rotational symmetry, and a chaotic classical counterpart.

Thus an ensemble $E_{1G}$ is defined in the space of all $N \times N$ real symmetric matrices, $T_{1G}$. The term ensemble implies that a probability density $P$ is applied to the space $T_{1G}$ which specifies how each matrix will be weighted in an ensemble average. The two conditions which determine the form of this probability density are [Meh91]:

1. The ensemble is invariant under every transformation

   $$H \to O^T HO$$

   of $T_{1G}$ into itself, where $O$ is any orthogonal matrix.

2. The various elements $H_{kj}$, $k \leq j$, are statistically independent .

The first condition is necessary because the basis chosen in order to represent the quantum Hamiltonian operator as a matrix was arbitrary. However, the final result should be the same no matter what basis was chosen - it should not make a difference what representation was used for the Hamiltonian. In general, if the basis chosen is changed by a unitary transformation $\psi \to U\psi$ then the Hamiltonian changes as

$$H \to UHU^{-1}. \tag{1.2.1}$$

However, since we are dealing only with matrix representations which are real and symmetric, the unitary transformations are restricted to orthogonal ones. Thus the first condition above says that if

$$H' = O^{-1}HO = O^T HO \tag{1.2.2}$$

then the probability $P(H')dH'$ equals $P(H)dH$, ie. the probability that a matrix in $E_{1G}$ belongs to the volume element $dH$ is invariant under real orthogonal transformations. This makes sense as $H$ and $H'$ can be thought of as representing the same physical system in two different representations.

The two conditions above determine the probability density function for the matrices in the ensemble. The statistical invariance of the matrix elements, condition (2), implies that the probability density is a product of functions, each depending on one matrix element only, and the form turns out to be [Meh91]

$$P(H) \propto \prod_j \exp[bH_{jj}] \prod_{k\leq j} \exp[-a(H_{kj})^2], \tag{1.2.3}$$

where $a$ is real and positive and $b$ is real. Because of the Gaussian factors above and the invariance of the ensemble under orthogonal transformations, $E_{1G}$ has been named the Gaussian Orthogonal Ensemble or GOE.

While the GOE dealt with systems which had time-reversal symmetry and integer total angular momentum or time-reversal symmetry and some rotational symmetry, a different collection of random matrices is needed to represent systems with time-reversal (but no rotational) symmetry and a half-integral total angular momentum. This ensemble is, in many ways, similar to the previous one, but this time it is defined in the space, $T_{4G}$, of all self-dual Hermitian matrices. A self-dual matrix is merely one which is equal to its time-reverse, so $T_{4G}$ contains $T_{1G}$, but the latter is only a very small fraction of the former. The Gaussian Symplectic Ensemble is defined in $T_{4G}$ by the two conditions [Meh91]:

1. The ensemble is invariant under every transformation

$$H \to S^{-1}HS$$

of $T_{4G}$ into itself, where $S$ is any unitary symplectic matrix.

2. The various linearly independent components of $H$ are statistically independent.

We note the definition that a unitary symplectic matrix $U$ satisfies $UU^{\dagger} = 1$ and $U^t JU = J$, where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ and $I$ is the identity matrix. Again, the probability density is the product of factors like those in (1.2.3), each one depending on one of the independent components of $H$.

The third, and final, ensemble contains matrices which could represent the quantum Hamiltonians of systems with no time-reversal symmetry. Such Hamiltonians have no restrictions other than their hermiticity. Thus, in order that the choice of basis with which to represent a Hamiltonian operator as a matrix does not affect the outcome, the probabilities $P(H')dH'$ and $P(H)dH$ for two matrices related by any unitary transformation $H' = U^{-1}HU$ must be equal. So, for the Gaussian Unitary Ensemble (GUE), if $T_{2G}$ is the space of Hermitian matrices:

1. The ensemble is invariant under every transformation

$$H \to U^{-1}HU$$

of $T_{2G}$ into itself, where $U$ is any unitary matrix.

2. The various linearly independent components of $H$ are statistically independent .

Since the GUE matrices can be complex, the linearly independent components in condition (2) are the real and imaginary parts of the off-diagonal elements $H_{kj}$, $k < j$, and the real parts of the diagonal elements, so the probability density is the product of factors containing each of these components.

### 1.2.2 Circular ensembles

The Gaussian ensembles of matrices have the drawback that because they are defined on a space of matrices which is not compact, there is no way to assign the same weight to every matrix; matrices representing different quantum systems are not treated equally. So, Dyson came up with three similar ensembles to those described above, but which are mathematically simpler than the Gaussian case. The difference is that whereas the matrices in the Gaussian ensembles were Hermitian, the matrices in Dyson's circular ensembles are unitary. The ensembles are called circular orthogonal (COE), circular symplectic (CSE) and circular unitary (CUE) and they correspond to systems sharing time-reversal symmetry, etc, in exactly the same way as the GOE, GSE and GUE respectively. Also, in a similar manner to the Gaussian ensembles, the COE is defined in the space of symmetric unitary matrices, the CSE contains unitary, self-dual matrices, and all arbitrary unitary matrices are in the CUE.

In the Gaussian ensembles, the eigenvalues of the Hermitian matrices could easily be connected with the energy levels of real systems because at least some of the matrices represented physical quantum Hamiltonians. The argument for using the circular ensembles is that we can imagine that a unitary matrix, $S$, is a function of a Hermitian matrix, $H$, which could be a quantum Hamiltonian. The exact form of the function does not matter, but the function could be imagined to be something like:

$$S = \exp(-iH\tau) \quad \text{or} \quad S = \frac{1 - i\tau H}{1 + i\tau H}. \tag{1.2.4}$$

The eigenvalues of $S$ have the form $e^{i\theta_j}$, $j = 1, \ldots N$, since $S$ is unitary.

The reason for imagining a relation like (1.2.4) is that if $S$ is a function of $H$, then the angles $\theta_1, \dots, \theta_N$ will be functions of the eigenvalues, $E_1, \dots, E_N$ of $H$. For a restricted energy range, this function should be approximately linear, and so it is reasonable to make the conjecture [Dys62] that the behaviour of a small number of the $N$ eigenvalues of $S$ will be statistically the same as a small number of the $N$ eigenvalues of the Hermitian matrix, $H$.

We will consider first the CUE, as this is the ensemble which will be dealt with most extensively in this thesis. The CUE is defined in the space of all unitary $N \times N$ matrices by the statement that the probability that a matrix is in the neighbourhood $dS$ of $S$ is given by

$$P(S)dS = \frac{\mu(dS)}{V}, \tag{1.2.5}$$

where $V$ is the total volume of the space of unitary matrices and $\mu(dS)$ is the volume of $dS$.

For the definition of $\mu(dS)$ the fact is used that any unitary matrix $S$ can be written as $S = UW$ for some pair of unitary matrices $U$ and $W$. A neighbourhood of $S$ is defined as

$$S + dS = U(1 + idM)W, \tag{1.2.6}$$

where $dM$ is a Hermitian matrix with infinitesimal elements $dM_{ij} = dM_{ij}^{(0)} + idM_{ij}^{(1)}$. Here we follow Mehta's notation for the real and imaginary components of $dM_{ij}$. $dM_{ij}^{(0)}$ and $dM_{ij}^{(1)}$ are thus real and vary independently over small intervals of length $d\mu_{ij}^{(0)}$ or $d\mu_{ij}^{(1)}$. $\mu(dS)$ is then defined as

$$\mu(dS) = \prod_{i \le j} d\mu_{ij}^{(0)} \prod_{i<j} d\mu_{ij}^{(1)}. \tag{1.2.7}$$

(1.2.5) then gives the statistical weight of each neighbourhood $dS$, which would be needed for example in an average over the CUE.

This definition for the CUE leads to the theorem, which is straightforward to prove [Meh91],

**Theorem 1.** *The CUE is uniquely defined in the space* $T_{2c}$ *of all* $N \times N$ *unitary matrices by the property of being invariant under every automorphism*

$$S \to USW \tag{1.2.8}$$

*of* $T_{2c}$ *into itself, where* $U$ *and* $W$ *are any two* $N \times N$ *unitary matrices.*

The probability $P(S)dS$ is therefore invariant over all matrices $S$ in $T_{2c}$, and so $\mu(dS)$ is just Haar measure on the group $U(N)$ of all $N \times N$ unitary matrices.

For the other two circular ensembles similar discussions lead to the following theorems.

**Theorem 2.** *The orthogonal ensemble (COE) is uniquely defined in the space* $T_{1c}$ *of unitary symmetric matrices of order* $N \times N$ *by the property of being invariant under every automorphism*

$$S \to W^T SW \tag{1.2.9}$$

*of* $T_{1c}$ *into itself where* $W$ *is any* $N \times N$ *unitary matrix.*

**Theorem 3.** *The symplectic ensemble (CSE) is uniquely defined in the space* $T_{4c}$ *of self-dual unitary quaternion matrices of order* $N \times N$ *by the property of being invariant under every automorphism*

$$S \to W^R SW \tag{1.2.10}$$

*of* $T_{4c}$ *into itself, where* $W$ *is any* $N \times N$ *unitary quaternion matrix.*

Here we recall that a self-dual matrix is one which is equal to its time-reverse.

### 1.2.3 Eigenvalue joint probability density function

In the preceding sections, the probability density function was discussed for the matrices in the various Gaussian and circular ensembles. As the aim of random matrix theory is to predict eigenvalue statistics for physical systems, it is useful to know the joint probability density function for the eigenvalues of the matrices in a given ensemble. The joint probability density function (JPDF) can be calculated from the density function of the matrices (eg. (1.2.3, 1.2.5)) by a change of variables. $P(x_1, \ldots, x_N)dx_1 \cdots dx_N$, where $P$ is the JPDF for a given ensemble, is the probability of choosing from that ensemble a matrix with eigenvalues lying between $x_1$ and $x_1 + dx_1$, $x_2$ and $x_2 + dx_2$, etc. Thus, to find the average over a particular ensemble of a function $f(x_1, x_2, \ldots, x_N)$, we multiply the function by the JPDF and integrate over the full range of the eigenvalues:

$$< f(x_1, \ldots, x_N) >= \int \cdots \int dx_1 \cdots dx_N P(x_1, \ldots, x_N) f(x_1, \ldots x_N). \tag{1.2.11}$$

If the orthogonal, symplectic and unitary ensembles are labelled by $\beta =$ 1,4 and 2 respectively (these numbers are related to the number of linearly independent components in the matrices) then the form of the JPDF for the Gaussian ensembles is

$$P_\beta(x_1, \ldots, x_N) = C_\beta \exp\left(-\frac{1}{2}\beta \sum_{j=1}^{N} x_j^2\right) \prod_{1 \leq j < k \leq N} |x_j - x_k|^\beta. \tag{1.2.12}$$

In the equation above, and the one following, $C_\beta$ is a normalization constant.

For the circular ensembles, the JPDF is

$$P_\beta(\theta_1, \ldots, \theta_N) = C'_\beta \prod_{1 \leq j < k \leq N} |\exp(i\theta_j) - \exp(i\theta_k)|^\beta, \tag{1.2.13}$$

where again $\beta$ distinguishes between the orthogonal, symplectic and unitary ensembles and the eigenvalues are $e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_N}$.

The statistics of the distribution of the matrix eigenvalues for a given ensemble can be defined in terms of the JPDF. For instance, the $n$-point correlation functions are given by

$$R_n(x_1, \dots, x_n) = \frac{N!}{(N-n)!} \int \cdots \int P(x_1, \dots, x_N) dx_{n+1} \cdots dx_N, \qquad (1.2.14)$$

and therefore are the probability densities of finding eigenvalues near the points $x_1, \dots, x_n$, while the remaining $N - n$ eigenvalues can be anywhere. Other statistics can also be defined, and the calculations pertaining to many of those can be found in the book on random matrix theory by Mehta [Meh91].

One common statistic is that of the nearest neighbour spacing distribution. This is the probability density $\mathcal{P}(s)$ of the spacings between consecutive eigenvalues and it illustrates the distinctive 'level repulsion' which exists between the eigenvalues of random matrices; that is, the probability density drops to zero as the spacing distance tends to zero. This statistic is shown for the three Gaussian ensembles in Figure 1.1. The repulsion is strongest for the GSE where $\mathcal{P}(s)$ is a quartic in $s$ near the origin, whereas the repulsion is quadratic for the GUE and linear for the GOE.

As a contrast to the eigenvalues of random matrices, a set of completely uncorrelated points displays Poisson statistics. In this case the nearest neighbour spacing distribution, also shown in Figure 1.1, displays no level repulsion.

The random matrix conjecture claims that for an $N \times N$ matrix chosen from one of the ensembles mentioned in the previous sections, the statistical properties of any finite number of its eigenvalues will tend with probability one, as $N \to \infty$, to the average ensemble statistics.

In the limit as $N$, the matrix size, tends to infinity the limit of any of these statistics involving only a finite number of eigenvalues is identical for the two ensembles CUE and GUE. The same is true between the COE and the GOE,

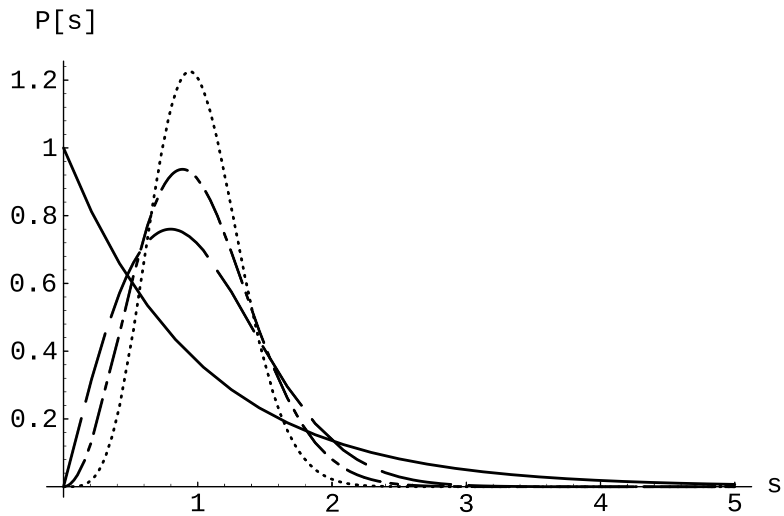


Figure 1.1: The nearest neighbour spacing distribution for the GOE (long dashed), GUE (dot-dashed), GSE (dotted) and Poisson (solid) statistics. The eigenvalues have been normalized so as to have unit mean consecutive spacing.

and also CSE and GSE statistics are the same in this limit. We will make use of this fact and perform calculations with the circular ensembles, which are mathematically simpler to work with.

## 1.3 Universality

In the previous sections we have discussed some random matrix ensembles and the random matrix conjecture (see, for example, [Meh91]) which suggests that the average eigenvalue statistics of such an ensemble might in general be asymptotically the same as the statistics of any particular matrix drawn from the ensemble. Thus these statistics are universal in that they are observed in almost all of the matrices from any given ensemble.

In the same way, in the context of physical systems, universality refers to the extent to which most systems having a Hamiltonian matrix which can be considered random, follow the eigenvalue statistics of the appropriate random matrix ensemble. It is clear that random matrix theory (RMT) can only predict features which are common to the majority of systems because the ensemble average is effectively over all possible systems. Thus nothing which is particular to a specific system can be predicted by RMT - it is beyond the range of universality. Not all eigenenergy statistics are fully universal, however; they tend to be universal - that is, for chaotic systems they match the average statistics of one of the matrix ensembles mentioned earlier - over short energy scales. We will see later in this section that this is related to the universality of the long periodic orbits of the systems. The natural scale with which to compare long and short correlations is the mean level spacing $\overline{d}(E)^{-1}$ (see 1.1.8). Over ranges much longer than this mean spacing, eigenenergy correlations cease to be universal, as we will see in the following example.

The form factor, $K(\tau)$, is the Fourier transform of the two point correlation function of a spectrum of eigenvalues. We write a density of states, with a delta function at each eigenvalue like (1.1.3), for a set of eigenvalues $e_n$, scaled so as to have an average distance of unity between consecutive levels (ie. $\overline{d}(e) = 1$). Thus

$$\begin{aligned} K(\tau) &= \int_{-\infty}^{\infty} dx \exp(2\pi i x\tau) \langle [d(e-x/2)-1][d(e+x/2)-1] \rangle \\ &\sim \frac{1}{h\overline{d}} \Big\langle \sum_i \sum_j A_i A_j \exp\left[\frac{i}{\hbar}(S_i - S_j)\right] \delta[T - 1/2(T_i+T_j)] \Big\rangle \end{aligned} \tag{1.3.1}$$

$$\approx \begin{cases} \frac{1}{h\overline{d}} \sum_j A_j^2 \delta(T-T_j) & \text{if} \quad \tau < \tau^* \\ \tau & \text{if} \quad \tau^* < \tau < 1 \\ 1 & \text{if} \quad \tau > 1 \end{cases} . \tag{1.3.2}$$

In the first line the average denoted by the angled brackets is an average in $e$ over a range which, while still being small compared to $e$, covers many energy levels. This average picks up a contribution whenever it encounters a pair of eigenvalues a distance $x$ apart; these do not have to be consecutive levels. The second line is the semiclassical asymptotic result; that is, the leading-order term in $h$. This arises directly from (1.1.8). In the final line, various approximations are made in different regions ($T = \tau h \bar{d}$). This last line (1.3.2) is an ansatz due to Berry [Ber85, Ber88, Ber91] for the form factor of a system fitting into the GUE category of no time-reversal invariance. A sketch of the form factor is shown below.

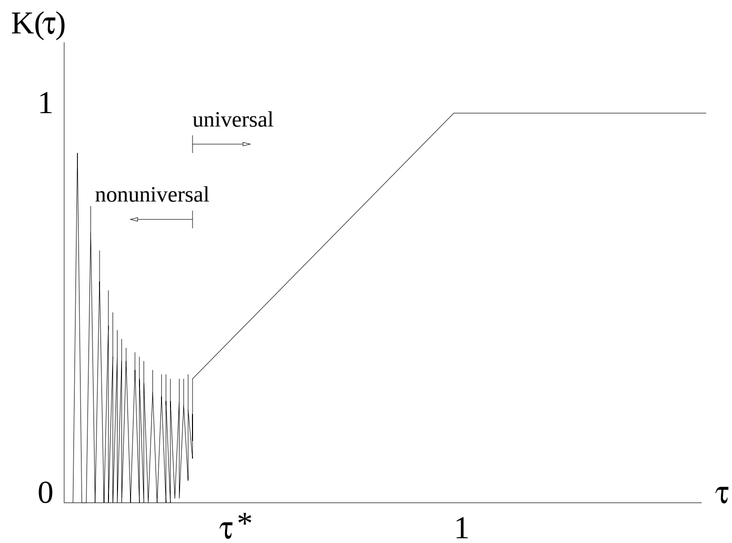


Figure 1.2: A sketch of the form factor described in Equation 1.3.2.

It is clear that at very small $\tau$ there is a delta spike at the period of each periodic orbit. At some unknown time $\tau^*$ between the period of the shortest orbit and $\tau = 1$, these delta functions average to a universal curve: in this case that of the GUE random matrix ensemble. For $\tau$ less than one, the GUE curve is just the straight line, $K(\tau) = \tau$, whereas $K(\tau) = 1$ when $\tau > 1$.

Thus we can see that while there is universality on some scales, at very short times (which correspond to long correlation scales on the energy axis due to the fact that $\tau$ and $x$ are Fourier conjugate variables) there are features which are specific to the individual system under study. Another way of viewing this is that in all the semiclassical formulae in Section 1.1 we see that the oscillatory term belonging to a given periodic orbit contains the factor $\exp(ikS_p/\hbar)$. Since statistical measures of correlation in the eigenvalue positions involve averaging (see, for example (1.3.1)) on a scale which is small compared to the energy $E$, we can make the approximation $S_p(E+\epsilon) \approx S_p(E) + \epsilon dS_p/dE = S_p(E) + \epsilon T_p(E)$. Thus we see that the wavelength of the oscillatory term connected with the periodic orbit $p$ is $h/T_p$, that is it is inversely proportional to the period of the orbit. The wavelength of the oscillation defines the scale on which the term belonging to that periodic orbit can affect correlations between eigenvalues in (1.1.8). For example, when $\tau = 1$ then $T = h\overline{d}$ and we see that it is those orbits with this period that just resolve eigenvalues separated by the mean level spacing. Thus Berry's ansatz implies that the periodic orbits with periods much shorter than $h\overline{d}$, which can vary greatly from one system to the next, lead to non-universal long-range correlations between the eigenvalues, while the longer orbits together contribute an effect on the short-range statistics which is universal. This universality of the long orbits is expressed explicitly in the Hannay-Ozorio de Almeida sum rule [HA84], which says that in general for any chaotic system, as $T \to \infty$

$$\sum_p \sum_{k=1}^{\infty} A_{p,k}^2 \delta(T - kT_p) \approx \frac{T}{4\pi^2}. \tag{1.3.3}$$

## 1.4 The Riemann zeta function

The Riemann zeta function is of interest to pure mathematicians because of its connection with prime numbers, but it is also a hugely important tool in quantum chaos because many calculations involving the Riemann zeta function mirror the most fundamental manipulations in semiclassical work, those concerning the energy eigenvalues of semiclassical systems and the action of the periodic orbits of those systems. Whereas the semiclassical calculations involve sums over periodic orbits of the system in question, as illustrated in

Section 1.1, the Riemann zeta function version contains sums over prime numbers. As much knowledge has built up about prime numbers over the years, the Riemann zeta calculations are often more tractable than the periodic orbit ones, and so can provide insight as to how the semiclassical calculations ought to proceed.

The aspect of the Riemann zeta function which is important to this work is the distribution of its non-trivial zeros. The Riemann zeta function is defined by the Euler product over prime numbers

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \tag{1.4.1}$$

or, equivalently, by a sum over integers

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{1.4.2}$$

The above definitions are only valid for $\mathrm{Re}\, s > 1$, but an analytic continuation allows the function to be defined on the whole complex plane. There then exists a functional equation (see [Tit86])

$$\zeta(s) = 2^s \pi^{s-1} \sin(s\pi/2)\Gamma(1-s)\zeta(1-s), \tag{1.4.3}$$

which relates the zeta function for $\mathrm{Re}\, s > 1/2$ to its values in the other half-plane.

The Riemann zeta function has an infinite number of 'trivial' zeros, which lie at the negative even integers, a pole at $s = 1$, and more importantly, an infinite number of complex zeros in the critical strip where $\mathrm{Re}\, s$ lies between zero and one. The Riemann Hypothesis (RH), as yet unproven but generally believed to be true, is that these complex zeros all lie on the line $\mathrm{Re}\, s = 1/2$. This is called the critical line. As there is much interest in the Riemann zeta function evaluated on this critical line, we sometimes write $\zeta(1/2+it)$ so that $t$ measures the distance up the line $\mathrm{Re}\, s = 1/2$.

The connection between the Riemann zeta function and random matrix theory is that in the limit as $t$ approaches infinity (that is, very high up the critical line) the statistics of the distribution of the complex Riemann zeros appear to tend exactly to those of the eigenvalues of the GUE random matrix ensemble. In 1973, Montgomery [Mon73] proved that in the limit as $t \to \infty$ up the critical line, the form factor statistic for the Riemann zeros was identical to that of the GUE up to the point $\tau = 1$, see Section 1.3. He further conjectured that the similarity held for all $\tau$. Many more recent results strongly support Montgomery's conjecture, for instance Rudnick and Sarnak have shown that all $n$-point correlations between the Riemann zeros high on the critical line agree with GUE statistics [RS96], though still just in the region equivalent to $\tau < 1$. The same result on the $n$-point correlations has been shown by Bogomolny and Keating [BK96] using a different method which impinges no restrictions on the range of validity, but which relies on a conjecture by Hardy and Littlewood on the distribution of prime numbers. Added to this is a vast quantity of numerical evidence generated by Odlyzko [Odl97].

Number theorists are very interested in mean values of the zeta function on the critical line and this is where the motivation for Chapter 3 lies. We will find that the universal aspect in the value distribution of a function with GUE zeros (or CUE, for the statistics are the same between these two in the limit of large matrix size that we are dealing with), will allow us propose a conjecture for these Riemann mean values which have proven so elusive when approached from the number theoretical perspective.

While this is our specific interest in the Riemann zeta function, it is the subject of much wider study in the field of quantum chaos because the positions of the prime numbers determine the Riemann zero statistics in a completely analogous manner to the connection between periodic orbits and eigenvalues discussed in Section 1.1. We will here present a taste of the argument, but for a very thorough discussion see [Kea93].

Since we are interested in the positions of zeros, we will start with the staircase function $N_R(T)$ (analogous to the spectral staircase of Section 1.1) which is a function of $T$, the distance up the critical line, and which increases by one at the position of each zero. Its value at $T$ is therefore the number of zeros on the critical line between $s = 1/2$ and $s = 1/2 + iT$. (Since the complex zeros occur in complex conjugate pairs, we will consider just $T > 0$ when we speak of counting zeros.) The change in the argument of the zeta function round a closed curve counts the zeros contained within that curve, so choosing a rectangle symmetric across the critical line in order that the functional equation (1.4.3) can be used to our advantage, we find

$$\begin{aligned} N_R(T) &= \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + O(T^{-1}) + \frac{1}{\pi}\operatorname{Im}\log\zeta(1/2+iT) \\ &\equiv \overline{N}_R(T) + \frac{1}{\pi}\operatorname{Im}\log\zeta(1/2+iT), \end{aligned} \tag{1.4.4}$$

where the branch of the log is defined continuously from $+\infty + iT$ to $1/2 + iT$ starting with the value 0.

For the argument of $\zeta(1/2+iT)$ we now use the Euler product. It should be noted that this is not entirely legitimate as the product does not converge on the critical line where we are evaluating the zeta function. However, it is through this substitution that we see the remarkable similarity with the formulae of Section 1.1 so we will continue and find

$$N_R(T) = \overline{N}_R(T) - \frac{1}{\pi}\sum_p \sum_{k=1}^{\infty} \frac{1}{k}\exp\left(-\frac{1}{2}k\log p\right)\sin(Tk\log p), \tag{1.4.5}$$

where we have expanded the logarithm of each factor in the Euler product.

From this we can determine, by differentiating, the density of states [Ber85]. Assuming that the complex zeros of the Riemann zeta function lie at $1/2+i\gamma_n$, $0 < \gamma_1 \le \gamma_2 \le \cdots$, we have

$$\begin{aligned} d_R(T) &= \sum_n \delta(T - \gamma_n) \qquad (1.4.6) \\ &= \overline{d}_R(T) - 2\sum_p \sum_{k=1}^{\infty} \frac{\log p}{2\pi} \exp\left(-\frac{1}{2}k\log p\right)\cos(Tk\log p), \end{aligned}$$

where $\overline{d}_R(T) = d\overline{N}_R/dT \sim \frac{1}{2\pi}\log\left(\frac{T}{2\pi}\right)$ is the mean density of states.

It is this formula which we need to compare with the density of states for physical systems, expressed in terms of periodic orbits in (1.1.8). In that equation we let $\eta = 0$ because the distribution of the zeros of the Riemann zeta function follow GUE statistics, and it was discussed in Section 1.2 why it is the classically chaotic systems rather than the classically integrable ones which are predicted to display such eigenvalue statistics. It has often been suggested that there may be an underlying physical system, the eigenvalues of which are the Riemann zeros (for example [BK99]), and if such a system exists it is believed that it would have a chaotic classical limit with no time-reversal invariance, as suggested by the GUE zero distribution.

We therefore set about comparing (1.1.8) and (1.4.6) and see that if we set $\hbar = 1$ then the sum over primes is just like a sum over orbits with Maslov indices $\mu_p = 0$ and actions, periods and amplitudes identified as

$$S_p(T) = T\log p \qquad (1.4.7a)$$

$$T_p = \log p \qquad (1.4.7b)$$

$$A_{p,k} = \frac{T_p}{2\pi}\exp\left(-\frac{1}{2}kT_p\right). \qquad (1.4.7c)$$

This similarity in the roles of prime numbers and periodic orbits is not only interesting to those in search of the elusive physical system lurking behind the Riemann zeta function, but it has also led to the development of techniques for dealing with periodic orbit sums through analogous methods already used on the Riemann zeta function [BK92, Kea92].

We can see from (1.4.6) that the term associated with the $k$th repetition of the primitive periodic orbit $p$ has wavelength $2\pi/\log p^k$. Thus, for example, since we know the mean spacing of the Riemann zeros at a large height $T$ up the critical line is asymptotically $2\pi/\log\left(\frac{T}{2\pi}\right)$, prime powers such that $\log p^k > \log\left(\frac{T}{2\pi}\right)$ determine the correlations in the positions of the zeros on scales shorter than the mean spacing. Also, since there is a smallest prime number (just as there is a shortest periodic orbit for a given physical system) there is an upper limit to the distance over which the positions of the Riemann zeros can be correlated, as illustrated by [Ber88].

We have now seen why the Riemann zeta function is of great interest to quantum chaologists, but it has long been studied from a purely number theoretical point of view. Of the many aspects which have been explored, it is the mean values, also called moments, which are relevant to this thesis. These moments, defined as

$$\frac{1}{T}I_\lambda(T) = \frac{1}{T}\int_0^T |\zeta(1/2+it)|^{2\lambda}dt, \tag{1.4.8}$$

are discussed in detail in Section 3.3 and it is these mean values which motivated this thesis project from the beginning. Here is $\zeta(1/2+it)$, for large $t$ a prime example of a function with zeros displaying random matrix statistics: how are these mean values (1.4.8) affected by the distribution of zeros?

## 1.5 Overview

In the attempt to answer and extend the above question, we commence in the following chapter to study the characteristic polynomial of a random $N \times N$ matrix, since this function clearly has zeros with random matrix statistics. We calculate the moments of this characteristic polynomial, where the average corresponding to the integral in (1.4.8) is performed over an ensemble of random matrices. Thus these mean values really are the moments of an *average* matrix from the ensemble, in the same way as an eigenvalue statistic such as (1.2.14) is an average statistic of the ensemble. In the same manner we study the value distribution of the logarithm of the characteristic polynomial, finding that the real and imaginary parts tend to independent Gaussian distributions as $N \to \infty$. This Gaussian result for just the imaginary part of the logarithm has already been determined, by a different method, by Costin and Lebowitz [CL95]. There is also overlap with the work of Haake et al. [HKS$^+$96] who, in the course of studying the statistics of the coefficients of the random matrix characteristic polynomials, derived expressions for the second moment of the modulus of this function for circular ensembles and for matrices showing Poisson eigenvalue statistics. These second moments also figure in the work of Kettemann, Klakow and Smilansky [KKS97], who studied autocorrelation functions of characteristic polynomials for Gaussian and circular ensembles of random matrices and compared them with the corresponding semiclassical expressions for spectral determinants. They note the connection between the quantum autocorrelation function and the classical Ruelle zeta function (related to the Fredholm determinant of the Perron-Frobenius operator - the classical evolution operator) and in [Smi99] Smilansky illustrates how the semiclassical expressions for the autocorrelation functions coincide with the equivalent random matrix results when the eigenvalues $w_2, w_3, \ldots$ of the Perron-Frobenius operator are distant from the main eigenvalue $w_1 = 1$.

There is also a field-theoretic technique for calculating autocorrelations functions of spectral determinants using Grassman variables [AS95, KKS97], but these expressions are normalized with respect to the correlators with coinciding correlation points and these latter are precisely the moments we consider in this thesis. Thus this technique does not yield an alternate method for the calculation of these moments.

Armed with the random matrix results developed in Chapter 2, a comparison is made with the Riemann zeta function in Chapter 3. We use the known values of the Riemann moments (1.4.8) and a theorem of Selberg's which states that the real and imaginary parts of the logarithm of $\zeta(1/2+it)$ are independently Gaussian in the limit as $t \to \infty$ to seek out the similarities with the random matrix calculations.

$L$-functions figure in Chapter 4 because when grouped together in families their zeros follow the statistics of a further pair of matrix ensembles [KS99b]. There is number theoretical interest in mean values of these families of $L$-functions [CF99], and again we compare the known results with random matrix theory.

In the final chapter we consider the value distribution of the spectral determinant (1.1.14) and its logarithm for a periodically kicked top, a chaotic system with eigenvalues displaying random matrix statistics. The value distribution of the imaginary part of the log of the spectral determinant was investigated for various chaotic billiard systems by Aurich, Bolte and Steiner [ABS94] (and later in [ABS97]) and found to be Gaussian in the semiclassical limit, as predicted by random matrix theory. For the kicked tops we found that even the approach of the distributions of the real and imaginary parts of the logarithm of the spectral determinant to this Gaussian limit is predicted by random matrix theory, the contributions specific to each kicked top being pushed to the extremes of the distribution in the semiclassical limit. We also conjecture that in this limit the moments of the spectral determinant factor into a random matrix component, dependent only on the zero statistics, and a part specific to the function under consideration, which takes the form of a product over periodic orbits.

# Chapter 2

# Random Matrix Theory takes Centre Stage

We now turn our attention to calculating mean values and the value distribution of the characteristic polynomial,

$$Z(U,\theta) = \prod_{n=1}^{N} \left(1 - e^{i(\theta_n - \theta)}\right), \tag{2.0.1}$$

where the $\theta_n$ are the eigenphases of a random $N \times N$ unitary matrix $U$ belonging to one of the circular ensembles mentioned in the introduction. This function then has zeros at points on the unit circle distributed with random matrix statistics, so it is just what we need to begin our calculations. In particular, mean values of this function where the average is performed over the ensemble of matrices will provide us with a random matrix moment with which to compare the moments of other functions, such as the Riemann zeta function, which have zeros with random matrix statistics. These random matrix mean values are expected to display the properties which surface in the moments of any function with the correct zero statistics, but not to contain any information specific to one particular function, since they are by construction an average over many such functions. Thus only characteristics which are common to all are expected to survive the average.

In the course of the chapter we will examine closely the distribution of values of the real and imaginary parts of the logarithm of $Z(U,\theta)$ and discover them to be Gaussian in the limit as $N\to\infty$, as well as considering the ensemble average values of powers of $|Z(U,\theta)|$ in the same limit.

In the final section we will briefly consider, by way of comparison, the value distribution of a function with uncorrelated zeros.

## 2.1 The CUE generating functions

The goal is to calculate average values of the function $Z(U,\theta)$ where the average is performed over an ensemble of random matrices. We start with the CUE ensemble described in the previous chapter.

The average over the ensemble is performed as follows. For any such average it is necessary to know the weighting assigned to each matrix in the average. This is well known as the ensemble has been extensively studied (for example [Meh91] [Dys62]). It is also identical to the group of unitary matrices $U(N)$ endowed with Haar measure, which implies that in terms of the eigenvalues of the matrix, this weighting is [Wey46] $\frac{1}{(2\pi)^N N!}\prod_{j<m}\left|e^{i\theta_j}-e^{i\theta_m}\right|^2$, so the $s^{th}$ moment of $|Z(U,\theta)|$ (where $s$ may be considered to be a complex variable) is

$$\begin{aligned}\langle|Z(U,\theta)|^s\rangle_{CUE} &= \frac{1}{(2\pi)^N N!}\times\int_0^{2\pi}\cdots\int_0^{2\pi}d\theta_1\cdots d\theta_N \\ &\quad\times\prod_{1\le j<m\le N}\left|e^{i\theta_j}-e^{i\theta_m}\right|^2\left|\prod_{n=1}^{N}(1-e^{i(\theta_n-\theta)})\right|^s .\end{aligned} \tag{2.1.1}$$

We will subdue all the integrals of this form which we encounter by the use of variations of Selberg's integral (described at length in Chapter 17 of [Meh91]). The particular integral which we need at the moment is

$$\begin{aligned} &J(a,b,\alpha,\beta,\gamma,N) \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left|\prod_{1\le j<\ell\le N}(x_j-x_\ell)\right|^{2\gamma}\times\prod_{j=1}^{N}(a+ix_j)^{-\alpha}(b-ix_j)^{-\beta}dx_j \qquad (2.1.2)\\ &= \frac{(2\pi)^N}{(a+b)^{(\alpha+\beta)N-\gamma N(N-1)-N}}\cdot\prod_{j=0}^{N-1}\frac{\Gamma(1+\gamma+j\gamma)\Gamma(\alpha+\beta-(N+j-1)\gamma-1)}{\Gamma(1+\gamma)\Gamma(\alpha-j\gamma)\Gamma(\beta-j\gamma)}. \end{aligned}$$

In the above formula, $a$, $b$, $\alpha$, $\beta$ and $\gamma$ are complex numbers, Re $a$, Re $b$, Re $\alpha$ and Re $\beta$ are all greater than zero, Re $(\alpha+\beta)>1$ and

$$-\frac{1}{N} < \text{Re }\gamma < \min\left(\frac{\text{Re }\alpha}{N-1},\frac{\text{Re }\beta}{N-1},\frac{\text{Re }(\alpha+\beta-1)}{2(N-1)}\right). \qquad (2.1.3)$$

In attempting to coerce (2.1.1) into the form of Selberg's integral, we note that

$$\left|e^{i\theta_j}-e^{i\theta_m}\right| = 2\left|\sin(\theta_j/2-\theta_m/2)\right|, \qquad (2.1.4a)$$

and similarly

$$\left|1-e^{i(\theta_p-\theta)}\right| = 2\left|\sin(\theta_p/2-\theta/2)\right|. \qquad (2.1.4b)$$

Therefore we can write (2.1.1) as

$$\begin{aligned} \langle|Z(U,\theta)|^s\rangle_{CUE} &= \frac{2^{N(N-1)}2^{sN}}{N!(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi}d\theta_1\cdots d\theta_N \qquad (2.1.5)\\ &\quad\times\prod_{1\le j<m\le N}\left|\sin(\theta_j/2-\theta_m/2)\right|^2\prod_{n=1}^{N}\left|\sin(\theta_n/2-\theta/2)\right|^s, \end{aligned}$$

and we note that this integral is in fact independent of $\theta$, which we eliminate to obtain

$$\begin{aligned}\langle |Z|^s \rangle_{CUE} &= \frac{2^{N(N-1)}2^{sN}2^N}{N!(2\pi)^N} \int_0^\pi \cdots \int_0^\pi d\theta_1 \cdots d\theta_N \prod_{1\le j<m\le N} |\sin(\theta_j-\theta_m)|^2 \\ &\quad \times \prod_{n=1}^N |\sin\theta_n|^s . \end{aligned} \tag{2.1.6}$$

If we write

$$\begin{aligned}\sin(\theta_j-\theta_m) &= \sin\theta_j\cos\theta_m - \cos\theta_j\sin\theta_m \\ \frac{\sin(\theta_j-\theta_m)}{\sin\theta_j\sin\theta_m} &= \cot\theta_m - \cot\theta_j,\end{aligned} \tag{2.1.7}$$

then the moment can be written as

$$\begin{aligned}\langle |Z|^s \rangle_{CUE} &= \frac{2^{N^2+sN}}{N!(2\pi)^N} \int_0^\pi \cdots \int_0^\pi d\theta_1 \cdots d\theta_N \prod_{1\le j<m\le N} |\cot\theta_m - \cot\theta_j|^2 \\ &\quad \times \prod_{n=1}^N \left(\sin^2\theta_n\right)^{N-1} \prod_{n=1}^N |\sin\theta_n|^s . \end{aligned} \tag{2.1.8}$$

A change of variables, $x_n = \cot\theta_n$, gives us

$$\begin{aligned}\langle |Z|^s \rangle_{CUE} &= \frac{2^{N^2+sN}}{N!(2\pi)^N} \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty dx_1 \cdots dx_N \prod_{1\le j<m\le N} |x_m - x_j|^2 \\ &\quad \times \prod_{n=1}^N ((1+ix_n)(1-ix_n))^{-N-s/2} \\ &= \frac{2^{N^2+sN}}{N!(2\pi)^N} J(1,1,N+s/2,N+s/2,1,N) \\ &= \prod_{j=1}^N \frac{\Gamma(j)\Gamma(s+j)}{(\Gamma(j+s/2))^2} \equiv M_N(s), \end{aligned} \tag{2.1.9}$$

where $J$ is the form of Selberg's integral defined in (2.1.2). Considering the conditions, listed in (2.1.3), on the various parameters, (2.1.9) holds if $-1 < \mathrm{Re}\, s$.

This result is of interest not only because it yields the form of the moments of $|Z|$, but also because $M_N(s)$, in (2.1.9), is the generating function for the moments of the real part of the logarithm of $Z$, which we will study in the next section.

We will return shortly to the implications of the generating function (2.1.9) but for the moment we will ask the natural question, which is what is the moment generating function for the imaginary part of the logarithm of $Z$. To answer this, we will consider $\langle (Z/Z^*)^{s/2} \rangle$. Here $\arg Z(U,\theta)$ is defined by continuous variation along $\theta - i\epsilon$, starting at $-i\epsilon$ and taking the limit $\epsilon \to 0$ (assuming that $\theta$ is not equal to any of the eigenphases $\theta_n$), with the further specification that $\log Z(U, \theta - i\epsilon) \to 0$ as $\epsilon \to \infty$. Thus $\mathrm{Im} \log Z(U,\theta)$ has a jump discontinuity of size $\pi$ when $\theta = \theta_n$. Now we have

$$\begin{aligned} \left(\frac{Z}{Z^*}\right)^{\frac{1}{2}} &= \exp\left(i \,\mathrm{Im} \log Z(U,\theta)\right) \\ &= \exp\left(-i \sum_{n=1}^{N} \sum_{m=1}^{\infty} \frac{\sin[(\theta_n - \theta)m]}{m}\right), \end{aligned} \tag{2.1.10}$$

where for each value of $n$, the sum of sine functions lies in $(-\pi, \pi]$.

We want to evaluate the $s^{th}$ moment of this quantity, so we use again the definition of the joint probability density function of CUE eigenvalues to write

$$\begin{aligned} \left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} &= \frac{1}{N!(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \prod_{1 \le j < m \le N} \left| e^{i\theta_j} - e^{i\theta_m} \right|^2 \\ &\quad \times \prod_{n=1}^{N} \exp\left(-is \sum_{m=1}^{\infty} \frac{\sin[(\theta_n - \theta)m]}{m}\right). \end{aligned} \tag{2.1.11}$$

As (2.1.11) is independent of $\theta$, we can set $\theta$ to zero, and as all the $\theta_n$'s lie in the interval $(0, 2\pi)$ we can replace the sum of sine functions in the exponent with a saw-tooth wave:

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k} = \frac{\pi - x}{2}, \quad \text{for} \quad 0 < x < 2\pi. \tag{2.1.12}$$

We note that this relation keeps the sine sum within the range $(-\pi, \pi]$ prescribed by the definition of the logarithm.

Substituting (2.1.12) into (2.1.11) yields

$$\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} = \frac{1}{N!(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \prod_{1\le j<m\le N} \left|e^{i\theta_j} - e^{i\theta_m}\right|^2 \times \prod_{n=1}^{N} \exp\left(-is\left(\frac{\pi-\theta_n}{2}\right)\right). \quad (2.1.13)$$

However, $|e^{i\theta_j} - e^{i\theta_m}| = 2|\sin(\theta_j/2 - \theta_m/2)|$, so

$$\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} = \frac{2^{N(N-1)}}{N!(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \times \prod_{1\le j<m\le N} \left|\sin(\theta_j/2 - \theta_m/2)\right|^2 \prod_{n=1}^{N} \exp\left(-\frac{is}{2}(\pi-\theta_n)\right). \quad (2.1.14)$$

Now let $\phi_j = \theta_j/2 - \pi/2$, so that we obtain

$$\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} = \frac{2^{N(N-1)}}{N!(2\pi)^N} 2^N \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} d\phi_1 \cdots d\phi_N \times \prod_{1\le j<m\le N} \left|\sin(\phi_j - \phi_m)\right|^2 \prod_{n=1}^{N} \exp(is\phi_n). \quad (2.1.15)$$

The trick now is to use the relation $\sin(\phi_j - \phi_m) = (\tan\phi_j - \tan\phi_m) \times \cos\phi_j \cos\phi_m$ to obtain

$$\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} = \frac{2^{N^2}}{N!(2\pi)^N} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} d\phi_1 \cdots d\phi_N \times \prod_{1\le j<m\le N} \left|\tan\phi_j - \tan\phi_m\right|^2 \times \prod_{n=1}^{N} (\cos^2\phi_n)^{N-1} \prod_{n=1}^{N} (\cos\phi_n + i\sin\phi_n)^s, \quad (2.1.16)$$

into which we can substitute $x_j = \tan\phi_j$:

$$\begin{aligned}\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} &= \frac{2^{N^2}}{N!(2\pi)^N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_N \prod_{1\le j<m\le N} \left|x_j - x_m\right|^2 \\ &\times \prod_{n=1}^{N} \left(\frac{1}{1+x_n^2}\right)^N \times \prod_{n=1}^{N} \left(\frac{1}{\sqrt{1+x_n^2}} + i\frac{x_n}{\sqrt{1+x_n^2}}\right)^s \\ &= \frac{2^{N^2}}{N!(2\pi)^N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_N \prod_{1\le j<m\le N} \left|x_j - x_m\right|^2 \\ &\times \prod_{n=1}^{N} \left(\frac{1}{1+x_n^2}\right)^N \times \prod_{n=1}^{N} \left(\frac{\sqrt{1+ix_n}}{\sqrt{1-ix_n}}\right)^s . \end{aligned} \tag{2.1.17}$$

This is now in the perfect form to apply the variation of Selberg's integral already introduced in (2.1.2), but this time we let $a = b = 1$, $\alpha = N - s/2$, $\beta = N + s/2$ and $\gamma = 1$. The condition (2.1.3) is only satisfied when $|\operatorname{Re} s| < 2$.

We now have

$$\left\langle \left(\frac{Z}{Z^*}\right)^{\frac{s}{2}} \right\rangle_{CUE} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j)}{\Gamma(j+s/2)\Gamma(j-s/2)} \equiv L_N(s). \tag{2.1.18}$$

As we intimated earlier, the interest in these functions $L_N(s)$ and $M_N(s)$ is partly due to the fact that from them we obtain the generating functions for the moments of the imaginary and real parts, respectively, of the log of $Z$. The real log moments are generated by $M_N(s)$, so we have

$$M_N(s) = \sum_{j=0}^{\infty} \frac{\langle(\log|Z|)^j\rangle_{CUE}}{j!} s^j, \tag{2.1.19}$$

and, defining $Q_j$ as the cumulants of the real part of the log,

$$\log M_N(s) = \sum_{j=1}^{\infty} \frac{Q_j}{j!} s^j. \tag{2.1.20}$$

For the imaginary part of the log if we define $\tilde{L}_N(s) \equiv L_N(-is)$, then $\tilde{L}_N(s)$ is our generating function. In this case we let $R_j$ denote the cumulants, and we obtain

$$\tilde{L}_N(s) = \sum_{j=0}^{\infty} \frac{\langle (\mathrm{Im}\log Z)^j \rangle_{CUE}}{j!} s^j, \tag{2.1.21}$$

and

$$\log \tilde{L}_N(s) = \sum_{j=1}^{\infty} \frac{R_j}{j!} s^j. \tag{2.1.22}$$

We can see now that by taking derivatives of $M_N(s)$ or $\tilde{L}_N(s)$ at $s = 0$ we recover the moments of the real and imaginary parts of the log of $Z$. The cumulants are obtainable by performing the same procedure on the log of $M$ or $\tilde{L}$.

## 2.2 Moments of $\mathrm{Re}\log Z$

In the previous section we applied Selberg's integral to tease out the result

$$M_N(s) \equiv \langle |Z|^s \rangle_{CUE} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+s)}{(\Gamma(j+s/2))^2}. \tag{2.2.1}$$

We turn our attention now to its role as a generating function for the moments of the real part of the logarithm of $Z$. The principle is straightforward; firstly we note that

$$\langle |Z|^s \rangle_{CUE} = \langle e^{s\log|Z|} \rangle_{CUE}, \tag{2.2.2}$$

so we can write

$$\begin{aligned} \langle (\log|Z|)^n \rangle_{CUE} &= \left\langle \frac{d^n}{ds^n} e^{s\log|Z|} \Big|_{s=0} \right\rangle_{CUE} \\ &= \frac{d^n}{ds^n} \langle |Z|^s \rangle_{CUE} \Big|_{s=0} \\ &= \frac{d^n}{ds^n} M_N(s) \Big|_{s=0}. \end{aligned} \tag{2.2.3}$$

Thus the moments of $\log|Z|$ can be calculated merely by taking derivatives at $s=0$ of the moments of $|Z|$.

Trivially, for $n=0$ we have just

$$\langle(\log|Z|)^0\rangle_{CUE} = M_N(s)\big|_{s=0} = 1. \tag{2.2.4}$$

For the first moment of $\log|Z|$ we need the first derivative of $M_N(s)$:

$$\begin{aligned}\frac{dM_N(s)}{ds} &= M_N(s)\left(\sum_{j=1}^{N}\frac{\frac{d}{ds}\Gamma(j+s)}{\Gamma(j+s)} - \frac{\frac{d}{d(s/2)}\Gamma(j+s/2)}{\Gamma(j+s/2)}\right) \\ &= M_N(s)\left(\sum_{j=1}^{N}\psi(j+s)-\psi(j+s/2)\right),\end{aligned} \tag{2.2.5}$$

where we have introduced the polygamma functions $\psi^{(n)}(z) = \frac{d^{n+1}}{dz^{n+1}}\log\Gamma(z)$. The digamma function is $\psi(z)\equiv\psi^{(0)}(z)$.

Evaluating (2.2.5) at $s=0$ gives

$$< (\log|Z|)^1 >_{CUE} = \frac{dM_N(s)}{ds}\bigg|_{s=0} = 0. \tag{2.2.6}$$

Further derivatives produce the pattern

$$\begin{aligned}\frac{d^nM_N(s)}{ds^n} &= \frac{d^{n-1}M_N(s)}{ds^{n-1}}Q_1(N,s) + \binom{n-1}{1}\frac{d^{n-2}M_N(s)}{ds^{n-2}}Q_2(N,s) \\ &\quad + \cdots + \binom{n-1}{n-2}\frac{dM_N(s)}{ds}Q_{n-1}(N,s) + \binom{n-1}{n-1}M_N(s)Q_n(N,s),\end{aligned} \tag{2.2.7}$$

where

$$\begin{aligned}Q_n(N,s) &= \frac{d^n}{ds^n}\log M_N(s) \\ &= \frac{d^{n-1}}{ds^{n-1}}\sum_{j=1}^{N}\left(\psi(j+s)-\psi(j+s/2)\right) \\ &= \sum_{j=1}^{N}\left(\psi^{(n-1)}(j+s) - \frac{1}{2^{n-1}}\psi^{(n-1)}(j+s/2)\right)\end{aligned} \tag{2.2.8}$$

for $n = 1, 2, 3, \ldots$. The cumulants, $Q_n(N,0) \equiv Q_n$, are exactly the derivatives of the logarithm of $M_N(s)$ defined in (2.1.20), and prior to moving on to the second moment, it is useful to consider these quantities in more depth.

The relation between the $Q_n(N,s)$ for consecutive values of $n$ is

$$\frac{d}{ds} Q_n(N,s) = Q_{n+1}(N,s). \tag{2.2.9}$$

To calculate the moments, we need the values of $Q_n(N,0)$. These are

$$Q_n(N,0) = \frac{2^{n-1}-1}{2^{n-1}} \sum_{j=1}^{N} \psi^{(n-1)}(j). \tag{2.2.10}$$

This is a beautifully concise formula for the cumulants for finite $N$, but we would like to know what the behaviour of these quantities is to leading order as $N$ becomes large.

It is clear from (2.2.10) that $Q_1(N,0) = 0$ but a little more work is needed to prise the asymptotic behaviour out of the higher cumulants.

We start with the second cumulant. We use the asymptotic expansion for the polygamma functions [AS65]

$$\psi(z) \sim \log z - \frac{1}{2z} - \sum_{n=1}^{\infty} \frac{B_{2n}}{2nz^{2n}} \tag{2.2.11}$$

and

$$\psi^{(n)}(z) \sim (-1)^{n-1} \left[ \frac{(n-1)!}{z^n} + \frac{n!}{2z^{n+1}} + \sum_{k=1}^{\infty} B_{2k} \frac{(2k+n-1)!}{(2k)!\,z^{2k+n}} \right] \tag{2.2.12}$$

for $|z| \to \infty$ with $|\arg z| < \pi$. Here the $B_{2k}$ are Bernoulli numbers. Also useful are the integral forms of the digamma function

$$\psi(z) + \gamma = \int_0^{\infty} \frac{e^{-t} - e^{-zt}}{1 - e^{-t}} dt \tag{2.2.13}$$

and the polygamma functions

$$\psi^{(n)}(z) = (-1)^{n-1} \int_0^\infty \frac{t^n e^{-zt}}{1-e^{-t}} dt. \tag{2.2.14}$$

From (2.2.12) we see that the second cumulant is of order $\log N$ as $N$ becomes large:

$$\begin{aligned} Q_2(N,0) &= \frac{1}{2} \sum_{j=1}^{N} \psi^{(1)}(j) \\ &= \frac{1}{2} \sum_{j=1}^{N} \left( \frac{1}{j} + \frac{1}{2j^2} + O\left(\frac{1}{j^3}\right)\right) \\ &\sim \frac{1}{2} \log N. \end{aligned} \tag{2.2.15}$$

The lower order terms can be found with the help of (2.2.14):

$$Q_2 = \frac{1}{2} \sum_{j=1}^{N} \int_0^\infty \frac{te^{-jt}}{1-e^{-t}} dt, \tag{2.2.16}$$

so interchanging the order of integration and summation yields

$$\begin{aligned} Q_2 &= \frac{1}{2} \int_0^\infty \frac{te^{-t}}{1-e^{-t}} \frac{1-e^{-Nt}}{1-e^{-t}} dt \\ &= \frac{1}{2} \int_0^\infty \frac{te^{-t}}{(1-e^{-t})^2} - \frac{te^{-(N+1)t}}{(1-e^{-t})^2} dt \end{aligned} \tag{2.2.17}$$

We now integrate by parts, using $\frac{d}{dt}\frac{e^{-t}}{1-e^{-t}} = \frac{-e^{-t}}{(1-e^{-t})^2}$. The result is three integrals which can again be written in terms of polygamma functions using (2.2.13) and (2.2.14).

$$\begin{aligned} Q_2 &= \frac{1}{2} \left[ \int_0^\infty \frac{e^{-t}}{1-e^{-t}} dt - \int_0^\infty \frac{e^{-(N+1)t}}{1-e^{-t}} dt + N \int_0^\infty \frac{te^{-(N+1)t}}{1-e^{-t}} dt \right] \\ &= \frac{1}{2} [-\psi(1) + \psi(N+1) + N\psi^{(1)}(N+1)] \\ &= \frac{1}{2} \log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} - \frac{1}{80N^4} + O\left(\frac{1}{N^6}\right), \end{aligned} \tag{2.2.18}$$

where the final line utilizes (2.2.11) and (2.2.12). This same result can be achieved using another method often found to be very useful for random matrix calculations, shown in Appendix A.

While $Q_2(N,0)$ diverges like $\log N$, the higher cumulants tend to a constant in the limit as $N \to \infty$. This can be seen by proceeding as in (2.2.15) but with the asymptotic expansions for the higher polygamma functions. This method, however, would not give the value of the constant leading order term of each cumulant, so instead we again use (2.2.14) and write, as for $Q_2$,

$$\begin{aligned} Q_n &= \frac{2^{n-1}-1}{2^{n-1}} \sum_{j=1}^{N} (-1)^n \int_0^\infty \frac{t^{n-1}e^{-jt}}{1-e^{-t}} dt \qquad (2.2.19) \\ &= \frac{2^{n-1}-1}{2^{n-1}} (-1)^n \left( \int_0^\infty \frac{t^{n-1}e^{-t}}{(1-e^{-t})^2} dt - \int_0^\infty \frac{t^{n-1}e^{-(N+1)t}}{(1-e^{-t})^2} dt \right). \end{aligned}$$

The first integral integrates by parts to yield the integral representation of the Riemann zeta function,

$$\Gamma(m)\zeta(m) = \int_0^\infty \frac{t^{m-1}}{e^t - 1} dt, \qquad (2.2.20)$$

valid for $m > 1$, and the second is expressible in terms of polygamma functions:

$$\begin{aligned} Q_n = \frac{2^{n-1}-1}{2^{n-1}} (-1)^n \big[ &(n-1)!\zeta(n-1) - (n-1)(-1)^{n-1}\psi^{(n-2)}(N+1) \\ &+ N(-1)^n \psi^{(n-1)}(N+1) \big]. \qquad (2.2.21) \end{aligned}$$

Applying (2.2.12) once more we obtain

$$Q_n = \frac{2^{n-1}-1}{2^{n-1}} (-1)^n \left( \Gamma(n)\zeta(n-1) - \frac{(n-3)!}{N^{n-2}} \right) + O(N^{1-n}), \qquad (2.2.22)$$

for $n \geq 3$.

Turning now to the moments, via (2.2.7), and remembering that $dM_N(s)/ds|_{s=0} = Q_1(N,0) = 0$, we find

$$\langle(\log|Z|)^2\rangle_{CUE} = Q_2(N,0) = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} - \frac{1}{80N^4} + O\left(\frac{1}{N^6}\right). \tag{2.2.23}$$

For the third moment,

$$\langle(\log|Z|)^3\rangle_{CUE} = \left.\frac{d^3M_N(s)}{ds^3}\right|_{s=0} = Q_3(N,0) \tag{2.2.24}$$

and so, asymptotically, we have (using (2.2.22))

$$\begin{aligned}\langle\log|Z|^3\rangle_{CUE} &= (-1)^3\frac{2^2-1}{2^2}\zeta(2)\Gamma(3) + (-1)^4\frac{2^2-1}{2^2}\frac{1}{N} + O\left(\frac{1}{N^2}\right)\\ &= \frac{-\pi^2}{4} + \frac{3}{4N} + O\left(\frac{1}{N^2}\right).\end{aligned} \tag{2.2.25}$$

Similarly,

$$\begin{aligned}\frac{d^4M_N(s)}{ds^4} &= \frac{d^3M_N(s)}{ds^3}Q_1(N,s) + 3\frac{d^2M_N(s)}{ds^2}Q_2(N,s)\\ &\quad +3\frac{dM_N(s)}{ds}Q_3(N,s) + M_N(s)Q_4(N,s).\end{aligned} \tag{2.2.26}$$

Evaluating this at $s=0$, using all the previous results for the lower moments and cumulants, yields the fourth moment asymptotically as

$$\begin{aligned}\langle(\log|Z|)^4\rangle_{CUE} &= 3\left(\frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} + \cdots\right)^2\\ &\quad +(-1)^4\frac{7}{8}\zeta(3)\Gamma(4) + (-1)^5\frac{7}{8}\frac{1}{N^2} + O(N^{-3}),\end{aligned} \tag{2.2.27}$$

because $Q_2(N,0) = d^2M_N(s)/ds^2\big|_{s=0}$ and $Q_1(N,0) = dM_N(s)/ds\big|_{s=0} = 0$.

The knowledge of the leading-order behaviour of each cumulant, summarized here,

$$\begin{aligned} Q_1 &= 0 \\ Q_2 &= \langle(\log|Z|)^2\rangle_{CUE} = \frac{1}{2}\ln N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} + O(N^{-4}) \\ Q_n &= (-1)^n \frac{2^{n-1}-1}{2^{n-1}}\zeta(n-1)\Gamma(n) + O(N^{2-n}), \quad n \geq 3, \end{aligned} \tag{2.2.28}$$

is all that is necessary to deduce the asymptotic behaviour of the moments themselves:

$$\begin{aligned} \langle(\log|Z|)^n\rangle_{CUE} &= \frac{d^n}{ds^n} M_N(s) \mid_{s=0} \\ &= \begin{cases} (2k-1)!!\langle(\log|Z|)^2\rangle^k + O((\ln N)^{k-2}) & \text{if } n = 2k \\ O((\ln N)^{k-1}) & \text{if } n = 2k+1 \end{cases} \end{aligned} \tag{2.2.29}$$

When the moments are written in this way it can be seen that the distribution of $\log|Z|/\sqrt{(1/2)\log N}$ is asymptotic to a Gaussian, as $N$ becomes large, because the $2k^{th}$ moment of this latter distribution is equal to $(2k-1)!!$ times the $k^{th}$ power of the second moment, while the odd moments are zero. The same result is evident from the cumulants (again normalized by $\sqrt{(1/2)\log N}$) as the cumulants of a Gaussian distribution are all zero except for the second.

One last comment concerning computations of the above moments and cumulants: when $N$ becomes very large, the time necessary to evaluate $Q_j$ using the sum over polygamma functions (2.2.10) becomes infeasible. As we know the leading-order behaviour of these parameters, see (2.2.22), we can determine the order in $N$ of the error involved in using an asymptotic form for the $Q$'s in the calculation of the moments.

If we use just the first two leading-order terms from (2.2.22) for the third cumulant and higher, and the terms down to and including $N^{-4}$ for the second cumulant, ( 2.2.18), then the error due to leaving out the lower order terms is $O((\log N)^{k-3})/N^2)$ for the $2k^{th}$ moment and $O((\log N)^{k-1}/N^2)$ for the $2k+1^{th}$ moment. These errors become negligible around $N = 100$ and immense savings are made in computation time.

## 2.3 Value distribution of $\mathrm{Re}\log Z$

It was seen in the preceding section that the function $M_N(s)$, the $s^{th}$ moment of $|Z|$, acted as a generating function for the moments of the real part of the log of $Z$, but its usefulness is not yet exhausted. The information about all possible moments is concealed within the probability density function for the value of a given function, and it is this which we will next determine for $\log|Z|$.

The probability density function for the values taken on by $\log|Z|$ is given by

$$\rho_N(x) = \langle \delta(\log|Z| - x)\rangle_{CUE}. \tag{2.3.1}$$

Taking the Fourier transform of this,

$$\begin{aligned} \widehat{\rho_N}(s) &= \left\langle \int_{-\infty}^{\infty} e^{isx}\delta(\log|Z| - x)dx \right\rangle_{CUE} \\ &= \left\langle e^{is\log|Z|}\right\rangle_{CUE}, \end{aligned} \tag{2.3.2}$$

we obtain the moments of $|Z|$ itself, $M_N(is)$. Therefore, the distribution of $\log|Z|$ is

$$\begin{aligned} \rho_N(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx} M_N(is)ds \qquad (2.3.3) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+is)}{(\Gamma(j+is/2))^2} e^{-isx} ds. \end{aligned}$$

This is the exact distribution function for $\log|Z|$ - no asymptotic approximations have been made. Unfortunately, it does not appear to be a simple matter to perform the Fourier transform of $M_N(is)$ analytically; so far the transform has only been computed numerically. This is not entirely satisfactory, as the oscillatory nature of the integrand means that it is only feasible to integrate over a finite range of $s$. For instance we integrated over $s$ in (-10,10) for the value of $N$ and the range of $x$ featured in Figure 2.1. The integrand is very small, though, when $|s|$ is large, so the error is not visible when the distribution is plotted.

On the other hand, more can be done analytically if we consider asymptotic approximations to the distribution. We now write $\rho_N(x)$ as

$$\begin{aligned}\rho_N(x) &= \frac{\prod_{j=1}^{N}\Gamma(j)}{2\pi}\int_{-\infty}^{\infty}\exp\left(-isx+\sum_{j=1}^{N}(\log\Gamma(j+is)\right.\\ &\qquad \left. -2\log\Gamma(j+is/2))\right)ds, \qquad (2.3.4)\end{aligned}$$

and then expand the exponent as a power series.

The coefficients in this expansion are just the cumulants studied in the previous section, multiplied by some extra factors of $i$. So, letting $Q_j \equiv Q_j(N,0)$,

$$\begin{aligned}\rho_N(x) &= \frac{\prod_{j=1}^{N}\Gamma(j)}{2\pi}\int_{-\infty}^{\infty}\exp\left(-isx-\log\prod_{j=1}^{N}\Gamma(j)+0\cdot s+Q_2\frac{(is)^2}{2}\right.\\ &\qquad \left. +Q_3\frac{(is)^3}{3!}+Q_4\frac{(is)^4}{4!}+\cdots\right)ds \qquad (2.3.5)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-isx-Q_2s^2/2-iQ_3s^3/3!+Q_4s^4/4!+\cdots\right)ds.\end{aligned}$$

Now let $y = s\sqrt{Q_2}$,

$$\begin{aligned}\rho_N(x) &= \frac{1}{2\pi\sqrt{Q_2}}\int_{-\infty}^{\infty}\exp\left(\frac{-iyx}{\sqrt{Q_2}}-\frac{y^2}{2}-\frac{iQ_3y^3}{3!Q_2^{3/2}}+\frac{Q_4y^4}{4!Q_2^2}+\cdots\right)dy\\ &= \frac{1}{2\pi\sqrt{Q_2}}\int_{-\infty}^{\infty}e^{-iyx/\sqrt{Q_2}}e^{-y^2/2}\times\left(1-\frac{iQ_3y^3}{3!Q_2^{3/2}}+\frac{Q_4y^4}{4!Q_2^2}+\cdots\right.\\ &\qquad \left. +\left(\frac{-iQ_3y^3}{3!Q_2^{3/2}}+\cdots\right)^2\Big/2+\cdots\right)dy. \qquad (2.3.6)\end{aligned}$$

For large $N$, $Q_2$ tends to $(1/2)\log N$ and $Q_j$ tends to a constant for $j \geq 3$. Therefore,

$$\frac{Q_j}{Q_2^{j/2}} \to 0 \quad \text{as} \quad N\to\infty, \quad \text{for} \quad j \geq 3. \qquad (2.3.7)$$

Thus as $N \to \infty$,

$$\begin{aligned} \rho_N(x) &\sim \frac{1}{2\pi\sqrt{Q_2}} \int_{-\infty}^{\infty} e^{-iyx/\sqrt{Q_2}} e^{-y^2/2} dy \\ &\sim \frac{1}{\sqrt{2\pi Q_2}} \exp\left(\frac{-x^2}{2Q_2}\right) \equiv \rho_N^G(x). \end{aligned} \tag{2.3.8}$$

Thus, to leading order, the distribution is a Gaussian. Terms of lower order in $N$ can be obtained by returning to (2.3.6), separating off the Gaussian term, then grouping the remaining terms according to their powers in $y$:

$$\begin{aligned} \rho_N(x) &= \rho_N^G(x) + \frac{1}{2\pi\sqrt{Q_2}} \int_{-\infty}^{\infty} e^{-iyx/\sqrt{Q_2}} e^{-y^2/2} \left( \frac{Q_3(iy)^3}{3!Q_2^{3/2}} + \frac{Q_4(iy)^4}{4!Q_2^2} + \cdots \right. \\ &\quad + \left( \frac{Q_3(iy)^3}{3!Q_2^{3/2}} + \frac{Q_4(iy)^4}{4!Q_2^2} + \cdots \right)^2 \Big/ 2! \\ &\quad \left. + \left( \frac{Q_3(iy)^3}{3!Q_2^{3/2}} + \frac{Q_4(iy)^4}{4!Q_2^2} + \cdots \right)^3 \Big/ 3! + \cdots \right) dy \\ &= \rho_N^G(x) + \frac{1}{2\pi\sqrt{Q_2}} \int_{-\infty}^{\infty} e^{-iyx/\sqrt{Q_2}} e^{-y^2/2} \left( \frac{A_3(iy)^3}{Q_2^{3/2}} + \frac{A_4(iy)^4}{Q_2^2} \right. \\ &\quad \left. + \frac{A_5(iy)^5}{Q_2^{5/2}} + \cdots \right) dy, \end{aligned} \tag{2.3.9}$$

where $A_3 = Q_3/3!$, $A_4 = Q_4/4!$, $A_5 = Q_5/5!$, but $A_6$ and higher are more complicated combinations of the cumulants.

The standard Gaussian distribution (mean 0, variance 1) is

$$\tilde{\rho}_N^G(x) = \sqrt{Q_2}\rho_N^G(\sqrt{Q_2}x), \tag{2.3.10}$$

so we will standardize the distribution $\rho_N(x)$ in the same way.

$$\begin{aligned}\tilde{\rho}_N(x) &= \tilde{\rho}_N^G(x) + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iyx}e^{-y^2/2} \\ &\quad \times \left(\frac{A_3(iy)^3}{Q_2^{3/2}} + \frac{A_4(iy)^4}{Q_2^2} + \frac{A_5(iy)^5}{Q_2^{5/2}} + \cdots\right) dy \\ &= \tilde{\rho}_N^G(x) + \frac{1}{2\pi}\sum_{m=3}^{\infty}\frac{A_m i^m}{Q_2^{m/2}}\int_{-\infty}^{\infty} e^{-iyx}e^{-y^2/2}y^m dy. \qquad (2.3.11)\end{aligned}$$

The remaining integrals can themselves be expanded as a sum, and the final result for the standardized distribution is therefore

$$\begin{aligned}\tilde{\rho}_N(x) &= \tilde{\rho}_N^G(x) + \frac{1}{\sqrt{2\pi}}\sum_{m=3}^{\infty}\frac{A_m}{Q_2^{m/2}}e^{-x^2/2} \\ &\quad \times \sum_{p=0}^{m}\binom{m}{p}x^p \begin{cases} i^{m-p}(m-p-1)!!, & m-p \text{ even} \\ 0, & m-p \text{ odd}\end{cases} \qquad (2.3.12)\end{aligned}$$

From this expression it can be seen that after the Gaussian term, the next correction is of order $(\log N)^{-3/2}$. The size of the correction, however, is dependent on $x$.

An alternate method for retaining lower order terms than the Gaussian, and one which displays more clearly than ( 2.3.12) the role of these terms, requires that we go back to (2.3.6) and this time approximate $\rho_N(x)$ with

$$\rho_N(x) \simeq \rho_N^A(x) \equiv \frac{1}{2\pi\sqrt{Q_2}}\int_{-\infty}^{\infty}\exp\left(\frac{-ixy}{\sqrt{Q_2}} - \frac{y^2}{2} - \frac{iQ_3y^3}{3!Q_2^{3/2}}\right) dy. \qquad (2.3.13)$$

It is possible to write $a_0z + a_1z^2 + a_2z^3 = a_2\left(z + \frac{a_1}{3a_2}\right)^3 - \frac{a_1^2}{3a_2}z + a_0z - \frac{a_1^3}{27a_2^2}$, so we let

$$a_0 = \frac{-ix}{\sqrt{Q_2}} \quad a_1 = \frac{-1}{2} \quad a_2 = \frac{-iQ_3}{3!Q_2^{3/2}}, \qquad (2.3.14)$$

and write

$$\rho_N^A(x) = \frac{1}{2\pi\sqrt{Q_2}} e^{-Q_2^3/(6Q_3^2)} \int_{-\infty}^{\infty} \exp\left( \frac{-iQ_3}{3!Q_2^{3/2}} \left( y + \frac{Q_2^{3/2}}{iQ_3} \right)^3 - \left( \frac{ix}{\sqrt{Q_2}} + \frac{iQ_2^{3/2}}{2Q_3} \right) y \right) dy. \quad (2.3.15)$$

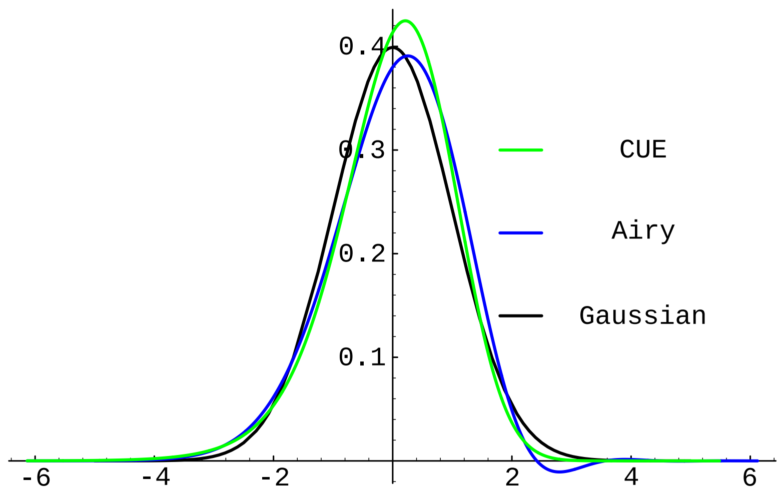


Figure 2.1: A comparison of the standardized exact distribution of $\log|Z|$ (green), the Airy approximation (blue) and the Gaussian (black) for $N = 42$.

Let $z = y - iQ_2^{3/2}/Q_3$:

$$\rho_N^A(x) = \frac{1}{2\pi\sqrt{Q_2}} e^{-Q_2^3/(6Q_3^2)} e^{-\left(\frac{ix}{\sqrt{Q_2}} + \frac{iQ_2^{3/2}}{2Q_3}\right)\left(\frac{iQ_2^{3/2}}{Q_3}\right)} \times \int_{-\infty - iQ_2^{3/2}/Q_3}^{\infty - iQ_2^{3/2}/Q_3} \exp\left( \frac{-iQ_3}{3!Q_2^{3/2}} z^3 - \left( \frac{ix}{\sqrt{Q_2}} + \frac{iQ_2^{3/2}}{2Q_3} \right) z \right) dz.$$

As the integrand has no poles and tends exponentially to zero as the real part of $z$ approaches $\pm\infty$ in the upper half-plane (and $Q_3 < 0$) the line of integration can be shifted back to the real axis. Then, as the exponent in the integrand is an odd function, we can discard the sine term and are left with

$$\begin{aligned}
\rho_N^A(x) &= \frac{1}{2\pi\sqrt{Q_2}} e^{-\frac{Q_2^3}{6Q_3^2}+\frac{xQ_2}{Q_3}+\frac{Q_2^3}{2Q_3^2}} \int_{-\infty}^{\infty} \cos\left(\frac{-Q_3}{3!Q_2^{3/2}}z^3 - \left(\frac{x}{\sqrt{Q_2}} + \frac{Q_2^{3/2}}{2Q_3}\right) z\right) dz \\
&= \left(\frac{-2}{Q_3}\right)^{1/3} e^{\frac{Q_2^3}{3Q_3^2}+\frac{xQ_2}{Q_3}} \operatorname{Ai}\left(\frac{2^{1/3}x}{Q_3^{1/3}} + \frac{Q_2^2}{2^{2/3}Q_3^{4/3}}\right). \qquad (2.3.16)
\end{aligned}$$

This is what we will call the Airy function approximation to the distribution of $\log|Z|$. The order of the error in this approximation can be determined in an identical manner to the Gaussian case and the result is $O((\log N)^{-2})$.

## 2.4 Moments of $\operatorname{Im}\log Z$

From Section 2.1 we remember that

$$\tilde{L}_N(s) \equiv \left\langle \left(\frac{Z}{Z^*}\right)^{\frac{-is}{2}} \right\rangle_{CUE} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j)}{\Gamma(j+is/2)\Gamma(j-is/2)}. \qquad (2.4.1)$$

In this section we intend to use the function $\tilde{L}_N(s)$ to carry out a study of the moments of the imaginary part of the log of $Z$. We will show that the distribution of the imaginary part of the log is Gaussian, just as that of the real part was. This result is of wider interest because $\frac{1}{\pi}\operatorname{Im}\log Z(U,\theta) = \tilde{N}(\theta)$, the fluctuating part of the spectral staircase function. The full staircase is created by summing a series of unit step functions, one at each CUE eigenphase, $\theta_j$, $1 \le j \le N$, and when the mean is subtracted from this the fluctuating function results. This relation between the imaginary part of the log of $Z$ and the staircase function is detailed in Appendix B. Our results show that the distribution of values of the fluctuating staircase function is Gaussian. This has already been determined for the Gaussian matrix ensembles, GUE, GOE and GSE, in terms of the random variable which gives the number of eigenvalues in an interval of length $L$ of a matrix from one of the above ensembles [CL95].

This statistic is very closely related to the staircase function and was found by Costin and Lebowitz to have Gaussian distribution.

Returning to our generating function (2.4.1),

$$\langle(\operatorname{Im}\log Z)^k\rangle_{CUE} = \frac{d^k}{ds^k}\left\langle e^{(\operatorname{Im}\log Z)s}\right\rangle_{CUE}\bigg|_{s=0} = \frac{d^k}{ds^k}\tilde{L}_N(s)\bigg|_{s=0}. \tag{2.4.2}$$

When writing the imaginary part of the logarithm of $Z$ it is to be remembered that we are using the definition given in Section 2.1 .

The zeroth moment is one, as expected, and the first moment of $\operatorname{Im}\log Z$ turns out to be zero in a similar manner as for the real case, see Section 2.2. The $n^{th}$ derivative of $\tilde{L}_N(s)$ is

$$\begin{aligned}\frac{d^n\tilde{L}_N(s)}{ds^n} &= \frac{d^{n-1}\tilde{L}_N(s)}{ds^{n-1}}R_1(N,s) + \binom{n-1}{1}\frac{d^{n-2}\tilde{L}_N(s)}{ds^{n-2}}R_2(N,s)+\cdots \\ &\quad + \binom{n-1}{n-2}\frac{d\tilde{L}_N(s)}{ds}R_{n-1}(N,s) + \tilde{L}_N(s)R_n(N,s),\end{aligned} \tag{2.4.3}$$

where

$$\begin{aligned} R_n(N,s) &= \frac{d^n}{ds^n}\log\tilde{L}_N(s) \\ &= \frac{d^{n-1}}{ds^{n-1}}\left[\frac{1}{2}\sum_{j=1}^{N}\left(-i\psi(j+is/2)+i\psi(j-is/2)\right)\right], \\ R_n(N,0) &= \frac{i^n}{2^n}\sum_{j=1}^{N}\left[-\psi^{n-1}(j)+(-1)^{n-1}\psi^{n-1}(j)\right] \\ &= \begin{cases} 0 & \text{if n odd} \\ \frac{(-1)^{n/2+1}}{2^{n-1}}\sum_{j=1}^{N}\psi^{n-1}(j) & \text{if n even}\end{cases}. \end{aligned} \tag{2.4.4}$$

These parameters, $R_n(N,0)\equiv R_n$, are the cumulants defined in (2.1.22). Note that for the even cumulants

$$R_{2m}(N,0) = \frac{(-1)^{m+1}}{2^{2m-1}-1}Q_{2m}(N,0). \tag{2.4.5}$$

As all the odd $R(N,0)$'s are zero and the first moment is also zero, it is clear by induction that at $s=0$ all odd derivatives of $\tilde{L}_N(s)$, and therefore all odd moments, are zero. The form of the even moments reduces to

$$\langle(\operatorname{Im}\log Z)^{2k}\rangle_{CUE} = \sum_{j=1}^{k}\binom{2k-1}{2j-1}\frac{d^{2k-2j}\tilde{L}_N(s)}{ds^{2k-2j}}\bigg|_{s=0} R_{2j}(N,0). \quad (2.4.6)$$

From (2.4.5) and (2.2.22) we see that the asymptotic form of the even $R$'s must be

$$R_{2j}(N,0) = \frac{(-1)^{j+1}}{2^{2j-1}}(2j-1)!\zeta(2j-1) + O(N^{2-2j}), \quad (2.4.7)$$

for $j \geq 2$, and $R_2(N,0) = Q_2(N,0)$.

The asymptotic form of the moments of Im log $Z$ therefore follows as

$$\langle(\operatorname{Im}\log Z)^{2k}\rangle_{CUE} = (2k-1)!!\langle(\operatorname{Im}\log Z)^2\rangle^k_{CUE} + O((\log N)^{k-2}), \quad (2.4.8)$$

where the second moment is just the same as the second moment of the real part of the log, given in (2.2.18),

$$\langle(\operatorname{Im}\log Z)^2\rangle_{CUE} = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} - \frac{1}{80N^4} + O\left(\frac{1}{N^6}\right). \quad (2.4.9)$$

Thus the even moments are Gaussian to leading order, and the error for the $2k^{th}$ moment is of order $(\log N)^{k-2}$. Being zero, the odd moments are exactly Gaussian.

As mentioned in the discussion on the real log moments, much time can be saved in the computation of the moments if an asymptotic form is used for the cumulants. This time, due to the lack of odd $R$'s, it is sufficient to keep just the leading-order (constant) term in the expansion of $R_j$, $j \geq 4$, (and terms up to $O(N^{-4})$ for $R_2(N,0)$, as in (2.4.9)) and still the error on the $2k^{th}$ moment is $O\left(\frac{(\log N)^{k-2}}{N^2}\right)$.

## 2.5 Value distribution of $\mathrm{Im}\log Z$

In the manner of Section 2.3, we now consider the distribution of the values taken by $\mathrm{Im}\log Z$. The distribution is again Gaussian in the limit as $N\to\infty$. We could see this from the moments and cumulants calculated in the previous section, but it is informative to study the distribution itself and the corrections to the Gaussian limit.

The distribution of $\mathrm{Im}\log Z$ is

$$\sigma(x) = \langle \delta(\mathrm{Im}\log Z - x)\rangle_{CUE}. \tag{2.5.1}$$

Taking the Fourier transform of this,

$$\begin{aligned} \widehat{\sigma}(k) &= \left\langle \int_{-\infty}^{\infty} e^{ikx}\delta(\mathrm{Im}\log Z - x)\right\rangle_{CUE} \\ &= \left\langle e^{ik\,\mathrm{Im}\log Z}\right\rangle_{CUE}, \end{aligned} \tag{2.5.2}$$

we obtain the moments of $\exp(i\,\mathrm{Im}\log Z)$ which we have already calculated in (2.1.18). Therefore, the distribution $\sigma(x)$ is

$$\begin{aligned} \sigma(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx}\left\langle e^{is\,\mathrm{Im}\log Z}\right\rangle_{CUE} ds \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx}\prod_{j=1}^{N}\frac{(\Gamma(j))^2}{\Gamma(j+s/2)\Gamma(j-s/2)}ds. \end{aligned} \tag{2.5.3}$$

As in the case of the distribution of $\log|Z|$, this transform has not been calculated analytically, so instead a numerical approximation has been used to plot $\sigma(x)$, in which only integration from -8 to +8 was possible. The error due to this truncation of the range of integration is bounded for all values of x by $(1/\pi)\int_8^\infty e^{-(R_2 s^2/2)}ds$. This bound evaluates to $3.19\times 10^{-5}$ when $N=42$, and becomes even smaller as $N$ increases.

Looking at the distribution asymptotically, however, approximations can be made to a certain extent analytically, as for $\log|Z|$. We begin with

$$\sigma(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\Bigg( - isx + \sum_{j=1}^{N} (2\log\Gamma(j) - \log\Gamma(j+s/2) - \log\Gamma(j-s/2))\Bigg) ds. \tag{2.5.4}$$

We want the coefficients of the power series,

$$\sum_{j=1}^{N} (2\log\Gamma(j) - \log\Gamma(j+s/2) - \log\Gamma(j-s/2)) = R_0 + R_1 is + R_2(is)^2/2! + R_3(is)^3/3! + \cdots, \tag{2.5.5}$$

but these are just the cumulants studied in Section 2.4. Thus $R_0 = 0$ and

$$R_n = \begin{cases} 0 & \text{n odd} \\ \frac{(-1)^{n/2+1}}{2^{n-1}} \sum_{j=1}^{N} \psi^{n-1}(j) & \text{n even} \end{cases}. \tag{2.5.6}$$

Therefore,

$$\sigma(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx - R_2 s^2/2! + R_4 s^4/4! + \cdots} ds. \tag{2.5.7}$$

If we let $y = s\sqrt{R_2}$, then

$$\sigma(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(\frac{-iyx}{\sqrt{R_2}} - \frac{y^2}{2} + \frac{R_4 y^4}{R_2^2 4!} + \cdots\right) \frac{dy}{\sqrt{R_2}}. \tag{2.5.8}$$

Since $R_2 = O(\log N)$ and $R_{2k} = O(1)$, $k = 2, 3, \ldots$, $R_{2k}/R_2^k$ will be small as $N$ grows. We therefore approximate $\exp(R_4 y^4/(4! R_2^2) + \cdots)$ by 1. Thus for large $N$ we have

$$\begin{aligned} \sigma(x) &\sim \frac{1}{2\pi\sqrt{R_2}}\int_{-\infty}^{\infty} \exp\left(\frac{-iyx}{\sqrt{R_2}} - \frac{y^2}{2}\right) dy \\ &\sim \frac{1}{\sqrt{2\pi R_2}} \exp\left(\frac{-x^2}{2R_2}\right) \equiv \sigma^G(x). \end{aligned} \tag{2.5.9}$$

This is exactly the asymptotic result we expected, a Gaussian with variance $R_2 = \langle(\mathrm{Im}\log Z)^2\rangle_{CUE}$.

As in the $\log|Z|$ case, we would like to know what error is incurred by making the Gaussian approximation. Expanding the second exponential in the following integral,

$$\begin{aligned}\sigma(x) &= \frac{1}{2\pi\sqrt{R_2}}\int_{-\infty}^{\infty}\exp\left(\frac{-iyx}{\sqrt{R_2}}-\frac{y^2}{2}\right)\\ &\qquad\times\exp\left(\frac{R_4y^4}{R_2^2 4!}-\frac{R_6y^6}{R_2^3 6!}+\cdots\right)dy \\ &= \sigma^G(x)+\frac{1}{2\pi\sqrt{R_2}}\int_{-\infty}^{\infty}\exp\left(\frac{-iyx}{\sqrt{R_2}}-\frac{y^2}{2}\right)\\ &\qquad\times\left[\frac{C_4y^4}{R_2^2}+\frac{C_6y^6}{R_2^3}+\frac{C_8y^8}{R_2^4}+\cdots\right]dy,\end{aligned} \tag{2.5.10}$$

where $C_4 = R_4/4!$, $C_6 = -R_6/6!$ and the higher coefficients are more complicated combinations of the $R$'s.

The standardized Gaussian distribution (mean zero, unit variance) is

$$\tilde{\sigma}^G(x) = \sqrt{R_2}\sigma^G(\sqrt{R_2}x), \tag{2.5.11}$$

so changing variables in $\sigma(x)$ in the same manner yields

$$\tilde{\sigma}(x) = \tilde{\sigma}^G(x) + \frac{1}{2\pi}\sum_{m=2}^{\infty}\frac{C_{2m}}{R_2^m}\int_{-\infty}^{\infty}e^{-iyx}e^{-y^2/2}y^{2m}dy. \tag{2.5.12}$$

This expression is similar in form to the distribution of $\log|Z|$, so

$$\begin{aligned}\tilde{\sigma}(x) &= \tilde{\sigma}^G(x)+\frac{1}{\sqrt{2\pi}}\sum_{m=2}^{\infty}\frac{C_{2m}}{R_2^m}e^{-x^2/2} \\ &\qquad\times\sum_{p=0}^{2m}\binom{2m}{p}(-ix)^p\begin{cases}(2m-p-1)!! & \text{if } 2\mathrm{m}-\mathrm{p} \text{ is even}\\ 0 & \text{if } 2\mathrm{m}-\mathrm{p} \text{ is odd}\end{cases}.\end{aligned} \tag{2.5.13}$$

If we keep one more term in the exponent in (2.5.8), then we end up with a distribution in the form of a Pearcey integral. This will be called the Pearcey approximation:

$$\sigma^P(x) = \frac{1}{2\pi\sqrt{R_2}} \int_{-\infty}^{\infty} \exp\left( \frac{-iyx}{\sqrt{R_2}} - \frac{y^2}{2} + \frac{R_4 y^4}{R_2^2 4!} \right) dy. \tag{2.5.14}$$

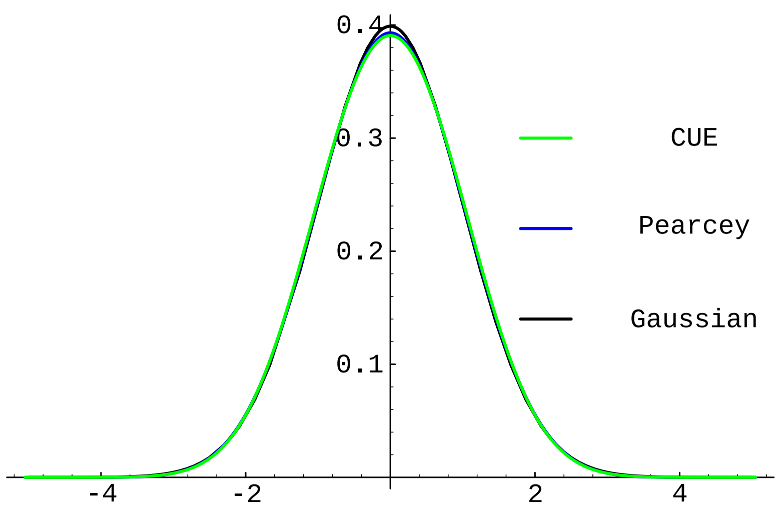


Figure 2.2: A comparison of the standardized exact distribution of $\operatorname{Im}\log Z$ (green), the Pearcey approximation (blue) and the Gaussian (black) for $N = 42$.

The correction to this approximation can be found in the usual way.

$$\begin{aligned}\tilde{\sigma}(x) &= \tilde{\sigma}^P(x) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-iyx - \frac{y^2}{2} + \frac{R^4 y^4}{R_2^2 4!}\right)\left(-\frac{R_6 y^6}{R_2^3 6!} + \cdots\right. \\ &\quad + \left(-\frac{R_6 y^6}{R_2^3 6!} + \cdots\right)^2 /2 + \left(-\frac{R_6 y^6}{R_2^3 6!} + \cdots\right)^3 /3! + \cdots \left.\right) dy \\ &= \tilde{\sigma}^P(x) + \frac{1}{2\pi}\sum_{m=3}^{\infty}\frac{D_{2m}}{R_2^m}\int_{-\infty}^{\infty} y^{2m}\exp\left(-iyx - \frac{y^2}{2} + \frac{R_4 y^4}{R_2^2 4!}\right) dy,\end{aligned} \tag{2.5.15}$$

where the $D$'s follow the usual pattern,

$$D_6 = -R_6/6!,\ \ D_8 = R_8/8! \text{ and } D_{10} = -R_{10}/10!, \tag{2.5.16}$$

while the higher $D$'s are more complicated.

## 2.6 Joint distribution

We are looking here at the joint distribution of the real and imaginary part of the log of $Z$. The aim is to show that the distributions are independent in the limit $N \to \infty$ where they each tend individually to a Gaussian.

Firstly we devise a generating function for these joint moments. The method here is identical to that used in the earlier sections in this chapter.

$$\begin{aligned}&\langle |Z|^t e^{is(\mathrm{Im}\log Z)}\rangle_{CUE} \\ &= \frac{1}{N!(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \prod_{1\le j<k\le N}\left|e^{i\theta_j} - e^{i\theta_k}\right|^2 \\ &\quad \times\prod_{n=1}^{N}\left|1 - e^{i(\theta_n-\theta)}\right|^t \prod_{n=1}^{N}\exp\left(-is\sum_{m=1}^{\infty}\frac{\sin[(\theta_n-\theta)m]}{m}\right) \\ &= \frac{1}{N!(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \prod_{1\le j<k\le N}\left|e^{i\theta_j} - e^{i\theta_k}\right|^2 \\ &\quad \times\prod_{n=1}^{N}\left|1 - e^{i\theta_n}\right|^t \prod_{n=1}^{N}\exp\left(-is\sum_{m=1}^{\infty}\frac{\sin(\theta_n m)}{m}\right).\end{aligned} \tag{2.6.1}$$

We know, as in Section 2.1, that

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{\pi - x}{2} \quad 0 < x < 2\pi, \tag{2.6.2}$$

so

$$\begin{aligned}
&\langle |Z|^t e^{is(\operatorname{Im}\log Z)} \rangle_{CUE} \\
&= \frac{1}{N!(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \prod_{1 \le j < k \le N} \left| e^{i\theta_j} - e^{i\theta_k} \right|^2 \\
&\quad \prod_{n=1}^{N} \left| 1 - e^{i\theta_n} \right|^t \prod_{n=1}^{N} \exp\left( -is \frac{\pi - \theta_n}{2} \right).
\end{aligned} \tag{2.6.3}$$

Now, using the relations in (2.1.4), we have

$$\begin{aligned}
\langle |Z|^t e^{is(\operatorname{Im}\log Z)} \rangle_{CUE} &= \frac{2^{N(N-1)} 2^{tN}}{N!(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \\
&\quad \times \prod_{1 \le j < k \le N} \left| \sin(\theta_j/2 - \theta_k/2) \right|^2 \prod_{n=1}^{N} \left| \sin(\theta_n/2) \right|^t \\
&\quad \times \prod_{n=1}^{N} \exp\left( -is \left( \frac{\pi}{2} - \frac{\theta_n}{2} \right) \right).
\end{aligned} \tag{2.6.4}$$

Then, letting $\phi_j = \theta_j/2 - \pi/2$,

$$\begin{aligned}
\langle |Z|^t e^{is(\operatorname{Im}\log Z)} \rangle_{CUE} &= \frac{2^{N^2} 2^{tN}}{N!(2\pi)^N} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} d\phi_1 \cdots d\phi_N \\
&\quad \times \prod_{1 \le j < k \le N} \left| \sin(\phi_j - \phi_k) \right|^2 \prod_{n=1}^{N} \left| \cos\phi_n \right|^t \\
&\quad \times \prod_{n=1}^{N} (\cos\phi_n + i \sin\phi_n)^s.
\end{aligned} \tag{2.6.5}$$

Noting that $\sin(\phi_j - \phi_k) = (\tan\phi_j - \tan\phi_k)\cos\phi_j \cos\phi_k$, we obtain

$$\begin{aligned}\langle |Z|^t e^{is(\mathrm{Im}\log Z)}\rangle_{CUE} &= \frac{2^{N^2}2^{tN}}{N!(2\pi)^N}\int_{-\pi/2}^{\pi/2}\cdots\int_{-\pi/2}^{\pi/2} d\phi_1\cdots d\phi_N \\ &\times \prod_{1\le j<k\le N}\left|\tan\phi_j-\tan\phi_k\right|^2\prod_{n=1}^{N}(\cos\phi_n)^{2(N-1)} \\ &\times \prod_{n=1}^{N}(\cos\phi_n)^t\prod_{n=1}^{N}(\cos\phi_n+i\sin\phi_n)^s. \end{aligned} \tag{2.6.6}$$

Now we change variables and let $x_j=\tan\phi_j$:

$$\begin{aligned}&\langle |Z|^t e^{is(\mathrm{Im}\log Z)}\rangle_{CUE} \\ &= \frac{2^{N^2}2^{tN}}{N!(2\pi)^N}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} dx_1\cdots dx_N\prod_{1\le j<k\le N}|x_j-x_k|^2 \\ &\times\prod_{n=1}^{N}\left(\frac{1}{1+x_n^2}\right)^{N+t/2}\prod_{n=1}^{N}\left(\frac{1}{\sqrt{1+x_n^2}}+i\frac{x_n}{\sqrt{1+x_n^2}}\right)^s \\ &= \frac{2^{N^2}2^{tN}}{N!(2\pi)^N}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} dx_1\cdots dx_N\prod_{1\le j<k\le N}|x_j-x_k|^2 \\ &\times\prod_{n=1}^{N}(1+ix_n)^{-N-t/2+s/2}(1-ix_n)^{-N-t/2-s/2} \\ &= \frac{2^{N^2}2^{tN}}{N!(2\pi)^N}J(1,1,N+t/2-s/2,N+t/2+s/2,1,N) \\ &= \prod_{j=1}^{N}\frac{\Gamma(j)\Gamma(t+j)}{\Gamma(j+t/2+s/2)\Gamma(j+t/2-s/2)}. \end{aligned} \tag{2.6.7}$$

The conditions on the validity of Selberg's integral translate into the restrictions $\mathrm{Re}\,t/2+\mathrm{Re}\,s/2>-1$, $\mathrm{Re}\,t/2-\mathrm{Re}\,s/2>-1$ and $\mathrm{Re}\,t>-1$.

In our quest for the joint distribution of the real and imaginary parts of the log, we need to determine the joint cumulants of the distribution. This necessitates expanding the log of our generating function around $t=0$ and $s=0$.

Firstly, we write

$$\begin{aligned}
&\langle |Z|^t e^{is(\mathrm{Im}\log Z)}\rangle_{CUE} \\
&= \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(t+j)}{\Gamma(j+t/2+s/2)\Gamma(j+t/2-s/2)} \\
&= \exp\left(\sum_{j=1}^{N} \log\Gamma(j) + \log\Gamma(t+j) - \log\Gamma(j+t/2+s/2)\right. \\
&\qquad \left. - \log\Gamma(j+t/2-s/2)\right),
\end{aligned} \tag{2.6.8}$$

and then we want to express the exponent as a power series in $s$ and $t$. Let

$$\begin{aligned}
f(t,s) &= \sum_{j=1}^{N} \log\Gamma(j) + \log\Gamma(t+j) - \log\Gamma(j+t/2+s/2) \\
&\qquad - \log\Gamma(j+t/2-s/2) \\
&= \alpha_{00} + \alpha_{10}t + \alpha_{01}s + \frac{\alpha_{20}}{2}t^2 + \alpha_{11}ts + \frac{\alpha_{02}}{2}s^2 + \frac{\alpha_{30}}{3!}t^3 + \frac{\alpha_{21}}{2!1!}t^2 s \\
&\qquad + \frac{\alpha_{12}}{2!1!}ts^2 + \frac{\alpha_{03}}{3!}s^3 + \cdots \\
&= f(0,0) + \left.\frac{\partial f}{\partial t}\right|_{(0,0)} t + \left.\frac{\partial f}{\partial s}\right|_{(0,0)} s + \frac{1}{2!}\left[\left.\frac{\partial^2 f}{\partial t^2}\right|_{(0,0)} t^2 + 2\left.\frac{\partial^2 f}{\partial t\partial s}\right|_{(0,0)} ts\right. \\
&\qquad \left. + \left.\frac{\partial^2 f}{\partial s^2}\right|_{(0,0)} s^2\right] + \frac{1}{3!}\left[\left.\frac{\partial^3 f}{\partial t^3}\right|_{(0,0)} t^3 + 3\left.\frac{\partial^3 f}{\partial t^2\partial s}\right|_{(0,0)} t^2 s\right. \\
&\qquad \left. + 3\left.\frac{\partial^3 f}{\partial t\partial s^2}\right|_{(0,0)} ts^2 + \left.\frac{\partial^3 f}{\partial s^3}\right|_{(0,0)} s^3\right] + \cdots,
\end{aligned} \tag{2.6.9}$$

so

$$\alpha_{ij} = \left.\frac{\partial^{i+j} f(t,s)}{\partial t^i \partial s^j}\right|_{(0,0)}. \tag{2.6.10}$$

The result is that

$$\alpha_{n0} = Q_n, \tag{2.6.11a}$$

$$\alpha_{0n} = i^n R_n, \tag{2.6.11b}$$

and if $n \neq 0$ and $m \neq 0$,

$$\begin{aligned}\alpha_{mn} &= \frac{\partial^m}{\partial t^m}\left[\sum_{j=1}^{N}\frac{1}{2^n}\left(-\psi^{(n-1)}(j+t/2+s/2)\right.\right.\\&\qquad \left.\left.+(-1)^{n-1}\psi^{(n-1)}(j+t/2-s/2)\right)\right]_{(0,0)}\\&= \sum_{j=1}^{N}\frac{1}{2^n}\frac{1}{2^m}\left[-\psi^{(n+m-1)}(j+t/2+s/2)\right.\\&\qquad \left.+(-1)^{n-1}\psi^{(n+m-1)}(j+t/2-s/2)\right]_{(0,0)}\\&= \begin{cases} 0 & \text{if n odd} \\ \frac{-1}{2^{n+m-1}}\sum_{j=1}^{N}\psi^{(n+m-1)}(j) & \text{if n even}\end{cases}. \qquad (2.6.12)\end{aligned}$$

We now turn our attention directly to the joint distribution of the real and imaginary parts of the log of $Z$. We carry out the following manipulations:

$$\begin{aligned}\langle |Z|^{it}e^{is(\operatorname{Im}\log Z)}\rangle_{CUE} &= \left\langle\int_{-\infty}^{\infty} e^{itx}\delta(\log|Z|-x)dx\right.\\&\qquad \left.\times\int_{-\infty}^{\infty} e^{isy}\delta(\operatorname{Im}\log Z-y)dy\right\rangle_{CUE}\\&= \int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\; e^{itx+isy}\langle\delta(\log|Z|-x)\\&\qquad \times\delta(\operatorname{Im}\log Z-y)\rangle_{CUE}. \qquad (2.6.13)\end{aligned}$$

This is the Fourier transform of the joint probability density function $\langle\delta(\log|Z|-x)\delta(\operatorname{Im}\log Z-y)\rangle_{CUE}$. So,

$$\begin{aligned}&\tau_N(x,y) \equiv \langle\delta(\log|Z|-x)\delta(\operatorname{Im}\log Z-y)\rangle_{CUE} \qquad (2.6.14)\\&= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt\, ds\, e^{-(itx+isy)}\left\langle e^{it\log|Z|}e^{is\operatorname{Im}\log Z}\right\rangle_{CUE}\\&= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt\, ds\, e^{-(itx+isy)}\prod_{j=1}^{N}\frac{\Gamma(j)\Gamma(it+j)}{\Gamma(j+it/2+s/2)\Gamma(j+it/2-s/2)}\\&= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt\, ds\, e^{-(itx+isy)}\exp\left(\alpha_{10}it+\alpha_{01}s+\frac{\alpha_{20}}{2}(it)^2\right.\\&\qquad \left.+\alpha_{11}its+\frac{\alpha_{02}}{2}s^2+\frac{\alpha_{30}}{3!}(it)^3+\frac{\alpha_{21}}{2!1!}(it)^2s+\frac{\alpha_{12}}{2!1!}its^2+\frac{\alpha_{03}}{3!}s^3+\cdots\right).\end{aligned}$$

As a start, we know that $\alpha_{10} = \alpha_{01} = \alpha_{11} = 0$. Next, the second cumulants are related to those for the individual real or imaginary distributions, so

$$\alpha_{20} = -\alpha_{02} = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} - \frac{1}{80N^4} + O\left(\frac{1}{N^6}\right). \qquad (2.6.15)$$

Lastly, from comparison with the cumulants of the real distribution alone, $\alpha_{mn} = O(1)$, for $m+n \geq 3$, so we will make the change of variables $v = t\sqrt{|\alpha_{20}|}$ and $w = s\sqrt{|\alpha_{20}|}$ and then keep just the leading-order contributions to the integral:

$$\begin{aligned}\tau(x,y) &= \frac{1}{4\pi^2\alpha_{20}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dv\, dw \exp\left(-\frac{ivx}{\sqrt{\alpha_{20}}} - \frac{iwy}{\sqrt{\alpha_{20}}} - \frac{v^2}{2} - \frac{w^2}{2}\right.\\ &\quad + \frac{\alpha_{30}}{3!\alpha_{20}^{3/2}}(iv)^3 + \frac{\alpha_{21}}{2!\alpha_{20}^{3/2}}(iv)^2 w + \frac{\alpha_{12}}{2!\alpha_{20}^{3/2}} ivw^2\\ &\quad \left. + \frac{\alpha_{03}}{3!\alpha_{20}^{3/2}} w^3 + \cdots\right)\\ &= \frac{1}{4\pi^2\alpha_{20}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dv\, dw \exp\left(-\frac{ivx}{\sqrt{\alpha_{20}}} - \frac{iwy}{\sqrt{\alpha_{20}}} - \frac{v^2}{2} - \frac{w^2}{2}\right)\\ &\quad \times\left(1 + O\left(\frac{1}{(\log N)^{3/2}}\right)\right). \end{aligned} \qquad (2.6.16)$$

To leading order we ignore the $O\left(\frac{1}{(\log N)^{3/2}}\right)$ term. Thus

$$\begin{aligned}\tau(x,y) &\sim \frac{1}{4\pi^2\alpha_{20}}\int_{-\infty}^{\infty} dv\, e^{-\frac{ivx}{\sqrt{\alpha_{20}}} - \frac{v^2}{2}} \int_{-\infty}^{\infty} dw\, e^{-\frac{iwy}{\sqrt{\alpha_{20}}} - \frac{w^2}{2}} \\ &\sim \left(\frac{1}{\sqrt{2\pi\alpha_{20}}}\exp\left(\frac{-x^2}{2\alpha_{20}}\right)\right)\left(\frac{1}{\sqrt{2\pi\alpha_{20}}}\exp\left(\frac{-y^2}{2\alpha_{20}}\right)\right).\end{aligned} \qquad (2.6.17)$$

We see, therefore, that to leading order $\tau(x,y)$ is the product of two Gaussian distributions. Thus the real and imaginary distributions are independent in this limit, but at the next order there will be cross terms.

## 2.7 Asymptotics of the generating functions

In preparation for our eventual comparison of these random matrix calculations with the moments of other relevant functions we now consider the leading-order behaviour of $M_N(s)$ and $L_N(s)$ as $N$ becomes large.

Starting with $M_N(s)$, when $s$ is an even integer we can work out the leading order behaviour very simply. (We note that the second moment was calculated prior to this work by [HKS$^+$96] using a method similar to that in Appendix A.) We see that if $k \leq N-1$

$$\begin{aligned}
\langle |Z|^{2k} \rangle_{CUE} &= \prod_{j=1}^{N} \frac{(j-1)!(2k+j-1)!}{(j+k-1)!^2} \\
&= \frac{0!1!\cdots(N-1)!(2k)!(2k+1)!\cdots(2k+N-1)!}{k!(k+1)!\cdots(k+N-1)!k!(k+1)!\cdots(k+N-1)!} \\
&= \frac{0!1!\cdots(k-1)!(k+N)!(k+N+1)!\cdots(2k+N-1)!}{N!(N+1)!\cdots(k+N-1)!k!(k+1)!\cdots(2k-1)!} \\
&= \frac{0!1!\cdots(k-1)!}{k!(k+1)!\cdots(2k-1)!}[(N+1)\cdots(k+N)]\,[(N+2)\cdots \\
&\qquad (k+N+1)]\cdots\cdots[(k+N)\cdots(2k+N-1)] \\
&= \prod_{j=0}^{k-1} \frac{j!}{(k+j)!} N^{k^2} + O(N^{k^2-1}). \qquad (2.7.1)
\end{aligned}$$

Following this work, Brézin and Hikami have determined that the above formula holds for the integer moments of characteristic polynomials of matrices belonging to other random matrix ensembles, including the GUE [BH99].

For general $s$, using (2.1.20) with the asymptotic expressions for the cumulants, (2.2.18) and (2.2.22), we define the leading-order coefficient, for large $N$, as

$$\begin{aligned}
f_{CUE}(s/2) &\equiv \lim_{N\to\infty} \frac{M_N(s)}{N^{(s/2)^2}} \qquad (2.7.2) \\
&= \exp\left( (s/2)^2(\gamma+1) + \sum_{j=3}^{\infty} (-s)^j \left(\frac{2^{j-1}-1}{2^{j-1}}\right) \frac{\zeta(j-1)}{j} \right).
\end{aligned}$$

We can see here that the $N$ dependence of the leading-order term comes from the order $\log N$ term in the second cumulant, $Q_2(N,0)$, and it is the constant term of each of the cumulants which features in $f_{CUE}(s/2)$.

It turns out that $f_{CUE}(s/2)$ can be expressed in terms of the Barnes G-function [Bar00, Vor87, Vig79],

$$G(1+z) = (2\pi)^{z/2} e^{-[(1+\gamma)z^2+z]/2} \prod_{n=1}^{\infty} \left[ (1+z/n)^n e^{-z+z^2/(2n)} \right], \tag{2.7.3}$$

which has zeros at the negative integers, $-n$, with multiplicity $n$ $(n = 1, 2, 3 \ldots)$. Other properties useful to us are that

$$\begin{aligned} G(1) &= 1, \\ G(z+1) &= \Gamma(z)\, G(z), \end{aligned} \tag{2.7.4}$$

and furthermore, for $|z| < 1$,

$$\log G(1+z) = (\log(2\pi) - 1)\frac{z}{2} - (1+\gamma)\frac{z^2}{2} + \sum_{n=3}^{\infty} (-1)^{n-1} \zeta(n-1) \frac{z^n}{n}. \tag{2.7.5}$$

Comparing the final line of (2.7.2) with (2.7.5) we see that for $|s| < 1$

$$f_{CUE}(s/2) = \frac{(G(1+s/2))^2}{G(1+s)}, \tag{2.7.6}$$

and so this equality holds by analytic continuation for all $s$.

Through the use of (2.7.4) we see that

$$G(n) = \prod_{j=1}^{n-1} \Gamma(j), \quad n = 2, 3, 4 \ldots, \tag{2.7.7}$$

and so we can check that for an even integer $s = 2n$:

$$\begin{aligned}\frac{(G(1+n))^2}{G(1+2n)} &= \frac{\prod_{j=1}^{n}\Gamma(j)^2}{\prod_{m=1}^{2n}\Gamma(m)}\\ &= \frac{\prod_{j=1}^{n}\Gamma(j)}{\prod_{m=n+1}^{2n}\Gamma(m)}\\ &= \frac{\prod_{j=0}^{n-1}j!}{\prod_{m=n}^{2n-1}m!}\\ &= \prod_{j=0}^{n-1}\frac{j!}{(j+n)!} = f_{CUE}(n),\end{aligned} \tag{2.7.8}$$

which is exactly the same as (2.7.1).

We note here a very important point on the deviation of individual matrices from the average ensemble behaviour. If we consider a function $f(U)$ dependent an a matrix $U$, and examine the variance of $f^2(U)$ from the ensemble average value $\overline{f}_2 = \langle f^2(U)\rangle_{CUE}$, we see that

$$\langle (f^2(U) - \overline{f}_2)^2\rangle_{CUE} = \langle f^4(U)\rangle_{CUE} - \overline{f}_2^2. \tag{2.7.9}$$

Normalizing with respect to $\overline{f}_2$,

$$\left\langle \left(\frac{f^2(U)}{\overline{f}_2} - 1\right)^2\right\rangle = \frac{\langle f^4(U)\rangle_{CUE}}{\overline{f}_2^2} - 1. \tag{2.7.10}$$

If $f(U) = \mathrm{Re}\log Z(U,\theta)$ for some fixed $\theta$, or if $f(U) = \mathrm{Im}\log Z(U,\theta)$, we see from (2.2.29) and (2.4.8) that (2.7.10) tends to a constant as $N\to\infty$. However, if $f(U) = |Z(U,\theta)|$, since $\langle |Z|^s\rangle_{CUE} \propto N^{(s/2)^2}$, (2.7.10) diverges as $N\to\infty$ and the deviation of the behaviour of $|Z|$ for one particular matrix from the ensemble mean could be huge. This is a warning to be heeded in later attempts to compare average random matrix results to individual functions with zeros displaying random matrix statistics.

Returning now to $L_N(s)$, and following the same procedure as was applied earlier in this section to $M_N(s)$,

$$\lim_{N\to\infty} L_N(s)N^{s^2/4} = \lim_{N\to\infty}\prod_{j=1}^{N}\frac{(\Gamma(j))^2}{\Gamma(j+s/2)\Gamma(j-s/2)}N^{s^2/4} \tag{2.7.11}$$

$$= \exp\left(-(\gamma+1)\left(\frac{s}{2}\right)^2 - \sum_{j=2}^{\infty}\frac{\zeta(2j-1)s^{2j}}{2^{2j}j}\right).$$

This has a similar form to (2.7.2). This time we guess that the leading-order coefficient of $L_N(s)$ is $G(1-s/2)G(1+s/2)$, which has zeros of order $n$ at the points $\pm 2n$ for $n = 1, 2, \ldots$.

$$\begin{aligned}
&\log(G(1-s/2)G(1+s/2))\\
&= \log G(1-s/2) + \log G(1+s/2)\\
&= -(\log(2\pi)-1)\frac{s}{4} - (1+\gamma)\frac{s^2}{8} + \sum_{n=3}^{\infty}(-1)^{n-1}\zeta(n-1)\frac{(-s)^n}{2^n n}\\
&\quad +(\log(2\pi)-1)\frac{s}{4} - (1+\gamma)\frac{s^2}{8} + \sum_{n=3}^{\infty}(-1)^{n-1}\zeta(n-1)\frac{s^n}{2^n n}\\
&= -(1+\gamma)\frac{s^2}{4} + 2\sum_{n=2}^{\infty}(-1)^{2n-1}\zeta(2n-1)\frac{(-s)^{2n}}{2^{2n}(2n)}\\
&= -(1+\gamma)\frac{s^2}{4} - \sum_{n=2}^{\infty}\frac{\zeta(2n-1)s^{2n}}{2^{2n}n}.
\end{aligned} \tag{2.7.12}$$

As the above holds for $|s/2| < 1$, $\lim_{N\to\infty} L_N(s)\times N^{s^2/4} = G(1-s/2)G(1+s/2)$ for $|s| < 2$ and so, by analytic continuation, for all s.

## 2.8 The COE and CSE ensembles

Although this work was motivated by connections between random matrix theory and the Riemann zeta function and so the initial work was done on the CUE ensemble, it can be extended to include two other random matrix ensembles: the Circular Orthogonal Ensemble and the Circular Symplectic Ensemble. In fact it is fairly straightforward to repeat the calculations for these two cases as the method involving Selberg's integral is equally applicable here as for the CUE.

For a general circular ensemble, the integral corresponding to (2.1.1) is

$$\begin{aligned}\langle|Z(U,\theta)|^s\rangle_{RMT} &\\ = \frac{(\beta/2)!^N}{(N\beta/2)!(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \prod_{1\le j<m\le N}\left|e^{i\theta_j}-e^{i\theta_m}\right|^\beta &\\ \times\left|\prod_{n=1}^N\left(1-e^{i(\theta_n-\theta)}\right)\right|^s, &\end{aligned} \tag{2.8.1}$$

where $\beta=1$ if $Z(U,\theta)=\prod_{n=1}^N(1-e^{i(\theta_n-\theta)})$ is to have zeros distributed like the eigenphases of a COE matrix, $\beta=4$ for CSE statistics and $\beta=2$ is the CUE result already considered. Exactly the same method as was applied in Section 2.1 is used here to obtain

$$\begin{aligned}\langle|Z|^s\rangle_{RMT} &= \prod_{j=0}^{N-1}\frac{\Gamma(1+j\beta/2)\Gamma(1+s+j\beta/2)}{(\Gamma(1+s/2+j\beta/2))^2}\\ &\equiv M_N(\beta,s).\end{aligned} \tag{2.8.2}$$

This leads to an equivalent version of (2.2.7),

$$\langle(\log|Z|)^n\rangle_{RMT}=\sum_{j=1}^n\binom{n-1}{j-1}\left.\frac{d^{n-j}M_N(\beta,s)}{ds^{n-j}}\right|_{s=0}Q_j^\beta(N,0), \tag{2.8.3}$$

where

$$Q_n^\beta(N,0)=\frac{2^{n-1}-1}{2^{n-1}}\sum_{j=0}^{N-1}\psi^{(n-1)}(1+j\beta/2). \tag{2.8.4}$$

As in the CUE case, $Q_1^\beta(N,0)=0$. Applying the recurrence formula for the polygamma function,

$$\psi^{(1)}(z+1)=\psi^{(1)}(z)-\frac{1}{z^2}, \tag{2.8.5}$$

we have in the CSE case that

$$\begin{aligned}
Q_2^4(N,0) &= \frac{1}{2}\sum_{j=0}^{N-1}\psi^{(1)}(1+2j) \\
&= \frac{1}{2}\left(\psi^{(1)}(1)+\sum_{j=1}^{N-1}\left(\psi^{(1)}(1)-\sum_{m=1}^{2j}\frac{1}{m^2}\right)\right) \\
&= \frac{N}{2}\psi^{(1)}(1)-\frac{1}{2}\sum_{k=1}^{N-1}\frac{N-k}{(2k-1)^2}-\frac{1}{2}\sum_{k=1}^{N-1}\frac{N-k}{(2k)^2} \\
&= \frac{1}{4}\log N+\frac{1}{4}(1+\gamma)+\frac{1}{4}\log 2+\frac{3}{16}\zeta(2)+O(N^{-1}). \quad (2.8.6)
\end{aligned}$$

In the COE case we follow a very similar procedure, except that as we now have polygamma functions of half-integers, we need to consider the case of even and odd $N$ separately. We start with $N$ even, relating the polygamma functions of integers back to $\psi^{(1)}(1)$ and those with half-integer argument to $\psi^{(1)}(1/2)$, and find that

$$\begin{aligned}
Q_2^1(N,0) &= \frac{1}{2}\sum_{j=0}^{N-1}\psi^{(1)}(1+j/2) \\
&= \frac{1}{2}\left(\frac{N}{2}\psi^{(1)}(1)+\frac{N}{2}\psi^{(1)}(1/2)-\sum_{k=1}^{N/2}\frac{4(N/2-k+1)}{(2k-1)^2}\right. \\
&\qquad \left.-\sum_{k=1}^{(N-2)/2}\frac{N/2-k}{k^2}\right) \\
&= \log N+1+\gamma-\frac{3}{4}\zeta(2)+O(N^{-1}). \quad (2.8.7)
\end{aligned}$$

The calculation for odd $N$ is very similar and the result is the same.

For $Q_n^\beta(N,0)$, $n \geq 3$, the sum converges as $N \to \infty$. To find this limiting value $Q_n^{\beta\,\infty}$, we write the polygamma function in its integral form (2.2.14) and interchange the order of summation and integration, as we did for the cumulants in Section 2.2, to obtain

$$
\begin{aligned}
Q_n^{\beta\,\infty} &\equiv \lim_{N\to\infty} Q_n^{\beta}(N,0) = \frac{2^{n-1}-1}{2^{n-1}}(-1)^n \int_0^\infty \frac{e^{-t}t^{n-1}}{(1-e^{-t})(1-e^{-\beta t/2})}dt \\
&= \frac{2^{n-1}-1}{2^{n-1}}(-1)^n \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\int_0^\infty e^{-(1+s+\beta r/2)t}t^{n-1}dt \\
&= \frac{2^{n-1}-1}{2^{n-1}}(-1)^n \sum_{r=0}^{\infty}\sum_{s=1}^{\infty}\Gamma(n)(s+\beta r/2)^{-n}. \qquad (2.8.8)
\end{aligned}
$$

Now if $\beta=4$, then the number of ways in which $s+2r=k$ is $k/2$ if $k$ is even and $(k+1)/2$ if $k$ is odd. This leads us to

$$
\begin{aligned}
Q_n^{4\,\infty} &= \frac{2^{n-1}-1}{2^{n-1}}(-1)^n\Gamma(n)\left(\sum_{k=1}^{\infty}\frac{k}{(2k-1)^n}+\sum_{k=1}^{\infty}\frac{k}{(2k)^n}\right) \\
&= \frac{2^{n-1}-1}{2^{n-1}}(-1)^n\Gamma(n)\frac{1}{2}\left(\sum_{k=1}^{\infty}\frac{2k-1}{(2k-1)^n}+\sum_{k=1}^{\infty}\frac{1}{(2k-1)^n}+\sum_{k=1}^{\infty}\frac{2k}{(2k)^n}\right) \\
&= \frac{2^{n-1}-1}{2^{n}}(-1)^n\Gamma(n)\left(\zeta(n-1)+\left(1-\frac{1}{2^n}\right)\zeta(n)\right). \qquad (2.8.9)
\end{aligned}
$$

Similarly,

$$
Q_n^{1\,\infty} = (2^{n-1}-1)(-1)^n\Gamma(n)\left(\zeta(n-1)-\left(1-\frac{1}{2^n}\right)\zeta(n)\right). \qquad (2.8.10)
$$

The asymptotic convergence of these $Q$ parameters ensures that the distribution of the real part of the log of $Z(U,\theta)$ is Gaussian for large $N$ when the zeros are distributed with COE or CSE statistics, just as it was for the CUE. All the calculations carried out for the CUE transfer immediately to the other two ensembles by replacing $Q_n$ with $Q_n^{\beta}$.

Among other things, we can once more write the leading-order (for large $N$) coefficient of $M_N(\beta,s)$ in terms of the Barnes G-function (2.7.3) and the gamma function. Using the asymptotics of the cumulants $Q_n^4$ just derived, we have for the CSE

$$
\begin{aligned}
M_N(4,s) &= \exp\Bigg(\left(\frac{1}{4}\log N + \frac{1}{4}(1+\gamma) + \frac{1}{4}\log 2 + \frac{3}{16}\zeta(2) + O\left(\frac{1}{N}\right)\right)\frac{s^2}{2} \\
&\quad + \sum_{n=3}^{\infty} \frac{2^{n-1}-1}{2^n}(-1)^n(n-1)!\,(\zeta(n-1) \\
&\quad + \left(1-\frac{1}{2^n}\right)\zeta(n) + o(1)\bigg)\frac{s^n}{n!}\Bigg), \qquad (2.8.11)
\end{aligned}
$$

which gives us

$$
\begin{aligned}
\lim_{N\to\infty}\frac{M_N(4,s)}{N^{s^2/8}} &= \exp\Bigg(\left(\frac{1}{4}(1+\gamma) + \frac{1}{4}\log 2 + \frac{3}{16}\zeta(2)\right)\frac{s^2}{2} \qquad (2.8.12) \\
&\quad + \sum_{n=3}^{\infty}(-1)^n\bigg(\frac{1}{2}\zeta(n-1) - \frac{1}{2^n}\zeta(n-1) + \frac{1}{2}\zeta(n) \\
&\quad - \frac{1}{2^n}\zeta(n) - \frac{1}{2^{n+1}}\zeta(n) + \frac{1}{2^{2n}}\zeta(n)\bigg)\frac{s^n}{n}\Bigg).
\end{aligned}
$$

The result is therefore, using (2.7.5) and

$$
\log\Gamma(1+z) = -\gamma z + \sum_{n=2}^{\infty}\zeta(n)\frac{(-z)^n}{n}, \qquad (2.8.13)
$$

valid for $|z| < 1$, that

$$
\lim_{N\to\infty}\frac{M_N(4,s)}{N^{s^2/8}} = 2^{s^2/8}\frac{G(1+s/2)\sqrt{\Gamma(1+s)}\,\Gamma(1+s/4)}{\sqrt{G(1+s)\Gamma(1+s/2)}\,\Gamma(1+s/2)}, \qquad (2.8.14)
$$

which has poles of order $k$ at $s = -(2k-1)$ and simple zeros at $s = -(4k-2)$ for $k = 1, 2, 3, \ldots$.

This reduces to a simpler expression when $s = 4k$, for integer $k$:

$$
\begin{aligned}
&2^{2k^2}\frac{G(1+2k)\sqrt{\Gamma(1+4k)}\Gamma(1+k)}{\sqrt{G(1+4k)\Gamma(1+2k)}\Gamma(1+2k)}\\
&=2^{2k^2}\frac{\left(\prod_{j=1}^{2k}\Gamma(j)\right)\Gamma(1+k)\sqrt{\Gamma(1+4k)}}{\sqrt{\left(\prod_{j=1}^{4k}\Gamma(j)\right)\Gamma(1+2k)}\Gamma(1+2k)}\\
&=2^{2k^2}\frac{1^{2k-1}2^{2k-2}\cdots(2k-1)k!\sqrt{(4k)!}}{\sqrt{1^{4k-1}2^{4k-2}\cdots(4k-1)(2k)!}(2k)!}\\
&=2^{2k^2}\frac{1^{2k-1}2^{2k-2}\cdots(2k-1)k!\sqrt{2^{2k}(2k)!}\sqrt{(4k-1)!!}}{\sqrt{1^{4k-1}3^{4k-3}5^{4k-5}\cdots(4k-1)(2k)!2^{2k-1}4^{2k-2}\cdots(4k-2)(2k)!}}\\
&=2^{2k^2}\frac{1^{2k-1}2^{2k-2}\cdots(2k-1)k!2^k}{\sqrt{1^{4k-2}3^{4k-4}5^{4k-6}\cdots(4k-3)^2 2^{\sum_{j=1}^{2k-1}j}1^{2k-1}2^{2k-2}\cdots(2k-1)(2k)!}}\\
&=\frac{2^{2k^2}k!2^k}{1^{2k-1}3^{2k-2}5^{2k-3}\cdots(4k-3)(2k)!2^{(2k-1)(2k)/2}}\\
&=\frac{2^k}{\left[\prod_{j=1}^{2k-1}(2k-1)!!\right](2k-1)!!}.
\end{aligned}
\tag{2.8.15}
$$

We can check this by examining $M_N(4,4k)$ directly:

$$
\begin{aligned}
&\prod_{j=0}^{N-1}\frac{\Gamma(1+2j)\Gamma(1+4k+2j)}{(\Gamma(1+2k+2j))^2}\\
&=\frac{\Gamma(1)\Gamma(3)\Gamma(5)\cdots\Gamma(2N-1)\Gamma(1+4k)\Gamma(3+4k)\cdots\Gamma(2N-1+4k)}{\left(\Gamma(1+2k)\Gamma(3+2k)\cdots\Gamma(2N-1+2k)\right)^2}\\
&=\frac{0!2!4!\cdots(2N-2)!(4k)!(4k+2)!\cdots(2N-2+4k)!}{\left((2k)!(2k+2)!\cdots(2N-2+2k)!\right)^2}\\
&=\frac{0!2!4!\cdots(2k-2)!(2N+2k)!(2N+2+2k)!\cdots(2N-2+4k)!}{(2N)!(2N+2)!\cdots(2N-2+2k)!(2k)!(2k+2)!\cdots(4k-2)!}
\end{aligned}
$$

$$= \frac{0!2!4!\cdots(2k-2)!}{(2k)!(2k+2)!\cdots(4k-2)!}[(2N+1)(2N+2)\cdots(2N+2k)]$$
$$[(2N+3)(2N+4)\cdots(2N+2+2k)]\cdots$$
$$[(2N-1+2k)(2N+2k)\cdots(2N-2+4k)]$$

$$\sim \frac{((2N)^{2k})^k 1^{2k-2}2^{2k-2}3^{2k-4}4^{2k-4}\cdots(2k-3)^2(2k-2)^2}{1^{2k-1}2^{2k-1}3^{2k-2}4^{2k-2}\cdots(2k-1)^k(2k)^k\cdots(4k-3)(4k-2)}$$

$$= \frac{(2N)^{2k^2}1^{2k-1}2^{2k-2}3^{2k-3}\cdots(2k-2)^2(2k-1)/(2k-1)!!}{2^{(2k)(2k-1)/2}1^{2k-1}2^{2k-2}\cdots(2k-1)1^{2k-1}3^{2k-2}5^{2k-3}\cdots(4k-3)}$$

$$= N^{2k^2}\frac{2^k}{\left[\prod_{j=1}^{2k-1}(2j-1)!!\right](2k-1)!!}. \tag{2.8.16}$$

We see that this agrees perfectly with (2.8.15), as expected.

In a similar manner, we can consider the leading-order behaviour of $M_N(1,s)$.

$$M_N(1,s) = \exp\Bigg(\left(\log N + 1 + \gamma - \frac{3}{4}\zeta(2) + O(N^{-1})\right)\frac{s^2}{2} \tag{2.8.17}$$
$$+\sum_{n=3}^{\infty}(2^{n-1}-1)(-1)^n(n-1)!\,(\zeta(n-1)$$
$$-\left(1-\frac{1}{2^n}\right)\zeta(n)+o(1)\Big)\frac{s^n}{n!}\Bigg),$$

giving

$$\lim_{N\to\infty}\frac{M_N(1,s)}{N^{s^2/2}} = \exp\Bigg(\left(1+\gamma-\frac{3}{4}\zeta(2)\right)\frac{s^2}{2} \tag{2.8.18}$$
$$+\sum_{n=3}^{\infty}(-1)^n\Big(2^{n-1}\zeta(n-1)-\zeta(n-1)$$
$$-2^{n-1}\zeta(n)+\frac{3}{2}\zeta(n)-\frac{1}{2^n}\zeta(n)\Big)\frac{s^n}{n}\Bigg).$$

In this case,

$$\lim_{N\to\infty}\frac{M_N(1,s)}{N^{s^2/2}} = \frac{G(1+s)\Gamma(1+s)\sqrt{\Gamma(1+s)}}{\Gamma(1+s/2)\sqrt{G(1+2s)\Gamma(1+2s)}}. \tag{2.8.19}$$

This combination has a $k$th order pole at $s=-(2k+1)/2$ and simple poles at $s=-(2k-1)$ for $k=1,2,3\dots$.

In our usual manner, we will now consider moments where $s=2k$.

$$\begin{aligned}
&\frac{G(1+2k)\Gamma(1+2k)\sqrt{\Gamma(1+2k)}}{\Gamma(1+k)\sqrt{G(1+4k)\Gamma(1+4k)}}\\
&\quad=\frac{\prod_{j=1}^{2k}\Gamma(j)\,\Gamma(1+2k)\sqrt{\Gamma(1+2k)}}{\Gamma(1+k)\sqrt{\prod_{j=1}^{4k}\Gamma(j)\,\Gamma(1+4k)}}\\
&\quad=\frac{\prod_{j=1}^{2k}(j-1)!(2k)!\sqrt{(2k)!}}{k!\sqrt{\prod_{j=1}^{4k+1}(j-1)!}}\\
&\quad=\frac{1^{2k-1}2^{2k-2}\cdots(2k-1)(2k)!\sqrt{(2k)!}}{k!\sqrt{1^{4k}2^{4k-1}3^{4k-2}\cdots(4k-2)^3(4k-1)^2(4k)}}\\
&\quad=\frac{1^{2k-1}2^{2k-2}\cdots(2k-1)(2k)!\sqrt{(2k)!}}{k!1^{2k}2^{2k-1}3^{2k-1}4^{2k-2}\cdots(4k-2)(4k-1)\sqrt{2\cdot4\cdots(4k-2)4k}}\\
&\quad=\frac{1^{2k-1}2^{2k-2}3^{2k-3}\cdots(2k-1)(2k)!}{k!1^{2k}2^{2k-1}3^{2k-1}4^{2k-2}\cdots(4k-2)(4k-1)2^k}\\
&\quad=\frac{1^{2k}2^{2k-2}3^{2k-2}4^{2k-4}5^{2k-4}\cdots(2k-3)^4(2k-2)^2(2k-1)^2}{1^{2k}2^{2k-1}3^{2k-1}4^{2k-2}\cdots(4k-3)^2(4k-2)(4k-1)}\\
&\quad=\frac{(2k-1)!^2(2k-3)!^2(2k-5)!^2\cdots3!3!1!1!}{(4k-1)!(4k-3)!(4k-5)!\cdots5!3!1!}\\
&\quad=\frac{(2k-1)!(2k-3)!(2k-5)!\cdots3!1!}{(4k-1)!(4k-3)!(4k-5)!\cdots(2k+1)!}\\
&\quad=\prod_{j=1}^{k}\frac{(2j-1)!}{(2k+2j-1)!}.
\end{aligned}\tag{2.8.20}$$

This agrees with a determination of the same leading-order coefficient directly from the formula (2.8.2) for $M_N(1,2k)$.

If we now consider the generating function for the imaginary part of the log of $Z(U,\theta)$, we have for the three circular ensembles that

$$\langle e^{is(\operatorname{Im}\log Z)}\rangle_{RMT}=\prod_{j=0}^{N-1}\frac{(\Gamma(1+j\beta/2))^2}{\Gamma(j\beta/2+1+s/2)\Gamma(j\beta/2+1-s/2)}\equiv L_N(\beta,s),\tag{2.8.21}$$

and again the odd log moments are zero and the even ones are

$$\langle (\mathrm{Im} \log Z)^{2k} \rangle_{RMT} = \sum_{j=1}^{k} \binom{2k-1}{2j-1} \frac{d^{2k-2j} L_N(\beta, -is)}{ds^{2k-2j}} \bigg|_{s=0} R_{2j}^{\beta}(N, 0), \quad (2.8.22)$$

where

$$R_n^{\beta}(N,0) = \begin{cases} 0 & \text{if n odd} \\ \frac{(-1)^{n/2+1}}{2^{n-1}} \sum_{j=0}^{N-1} \psi^{n-1}(1 + j\beta/2) & \text{if n even} \end{cases} . \quad (2.8.23)$$

We see from (2.8.4) that $R_2^{\beta}(N,0) = Q_2^{\beta}(N,0)$ and that $R_{2n}^{\beta} = \frac{(-1)^{n+1}}{2^{2n-1}-1} Q_{2n}^{\beta}$, so the value distribution of $\mathrm{Im} \log Z$ is Gaussian for the COE and CSE as well as for the CUE.

The leading-order behaviour of $L_N(4, s)$ is

$$\begin{aligned} L_N(4,s) &= \exp\Bigg( \left( \frac{1}{4} \log N + \frac{1}{4}(1+\gamma) + \frac{1}{4} \log 2 + \frac{3}{16}\zeta(2) + O\left(\frac{1}{N}\right) \right) \frac{(is)^2}{2} \\ &\quad + \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{2^{2n}} (2n-1)! \, (\zeta(2n-1) \\ &\quad + \left(1 - \frac{1}{2^{2n}}\right) \zeta(2n) + o(1) \Big) \frac{(is)^{2n}}{(2n)!} \Bigg), \end{aligned} \quad (2.8.24)$$

so

$$\begin{aligned} \lim_{N\to\infty} L_N(4,s) \times N^{s^2/8} &= \exp\Bigg( - \left( \frac{1}{4}(1+\gamma) + \frac{1}{4} \log 2 + \frac{3}{16} \zeta(2) \right) \frac{s^2}{2} \\ &\quad - \sum_{n=2}^{\infty} \Bigg( \frac{1}{2^{2n}} \zeta(2n-1) + \frac{1}{2^{2n}} \zeta(2n) \\ &\quad - \frac{1}{2^{4n}} \zeta(2n) \Bigg) \frac{s^{2n}}{2n} \Bigg). \end{aligned} \quad (2.8.25)$$

From our usual expansions of the G-function (2.7.5) and the gamma functions (2.8.13), we find that

$$\lim_{N\to\infty} L_N(4,s) \times N^{s^2/8}$$
$$= 2^{-s^2/8}\sqrt{\frac{G(1+s/2)G(1-s/2)\Gamma(1+s/4)\Gamma(1-s/4)}{\Gamma(1+s/2)\Gamma(1-s/2)}}, \quad (2.8.26)$$

which has zeros of order $k$ at $s = \pm(4k-2)$ and also at $s = \pm 4k$ for $k = 1, 2, 3, \ldots$.

For the COE, we have

$$\begin{aligned} L_N(1,s) &= \exp\Bigg(\left(\log N + 1 + \gamma - \frac{3}{4}\zeta(2) + O\left(\frac{1}{N}\right)\right)\frac{(is)^2}{2} \\ &\quad + \sum_{n=2}^{\infty}(-1)^{n+1}(2n-1)!\,\Big(\zeta(2n-1) \\ &\quad - \left(1 - \frac{1}{2^{2n}}\right)\zeta(2n) + o(1)\Big)\frac{(is)^{2n}}{(2n)!}\Bigg), \end{aligned} \quad (2.8.27)$$

so

$$\begin{aligned} \lim_{N\to\infty} N^{s^2/2} L_N(1,s) &= \exp\Bigg(-\left(1+\gamma - \frac{3}{4}\zeta(2)\right)\frac{s^2}{2} \quad (2.8.28) \\ &\quad - \sum_{n=2}^{\infty}\left(\zeta(2n-1) - \zeta(2n) + \frac{1}{2^{2n}}\zeta(2n)\right)\frac{s^{2n}}{2n}\Bigg). \end{aligned}$$

Thus we see that

$$\lim_{N\to\infty} N^{s^2/2} L_N(1,s) = \sqrt{\frac{G(1+s)G(1-s)\Gamma(1+s)\Gamma(1-s)}{\Gamma(1+s/2)\Gamma(1-s/2)}}, \quad (2.8.29)$$

a combination which has zeros of order $k$ at $s = \pm 2k$ as well as zeros also of order $k$ at $s = \pm(2k+1)$.

## 2.9 Uncorrelated eigenvalues

In this chapter so far we have considered functions with zeros distributed like the eigenvalues of matrices from the CUE, COE and CSE ensembles described in Chapter 1. The positions of these zeros were not mutually independent; they displayed features such as level repulsion (see Figure 1.1) characteristic of the random matrix ensembles. However, in this section we explore mean values and the value distribution of a function with completely uncorrelated zeros. We are still considering a function with $N$ zeros on the unit circle, but each phase is independent of the positions of the others, and has a uniform probability of lying anywhere in $[0, 2\pi)$. Such zeros would be expected to display Poisson statistics, for example with a nearest neighbour spacing as in Figure 1.1.

We will once more call the function $Z(\theta) = \prod_{n=1}^{N}(1 - e^{i(\theta_n - \theta)})$, but here the set of phases $\theta_n$, $n = 1, 2, \ldots, N$ are uncorrelated. We examine the same mean values as in the circular ensemble cases, so we begin with

$$\langle |Z(\theta)|^s \rangle = \frac{1}{(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{n=1}^{N} |1 - e^{i(\theta_n - \theta)}|^s d\theta_1 \cdots d\theta_N. \quad (2.9.1)$$

The second moment, $\langle |Z(\theta)|^2 \rangle$ has previously been calculated by [HKS+96].

Due to the lack of correlation between the phases, this integral separates into a product of $N$ identical integrals. Each integral is independent of $\theta$, which can be removed by a change of variables.

$$
\begin{aligned}
\langle |Z(\theta)|^s \rangle &= \left( \frac{1}{2\pi} \int_0^{2\pi} |1 - e^{i(\theta_n - \theta)}|^s d\theta_n \right)^N \\
&= \left( \frac{1}{2\pi} \int_0^{2\pi} |1 - e^{i\theta_n}|^s d\theta_n \right)^N \\
&= \left( \frac{1}{2\pi} \int_0^{2\pi} (2\sin(\theta_n/2))^s d\theta_n \right)^N \\
&= \left( \frac{2^{s+2}}{2\pi} \int_0^{\pi/2} (\sin\theta_n)^s d\theta_n \right)^N \\
&= \left( \frac{2^{s+2}}{2\pi} 2^{s-1} \frac{\Gamma\left(\frac{s+1}{2}\right)\Gamma\left(\frac{s+1}{2}\right)}{\Gamma(s+1)} \right)^N \\
&= \left( \frac{\Gamma(1+s)}{\Gamma(1+s/2)\Gamma(1+s/2)} \right)^N \\
&\equiv M_P(N,s).
\end{aligned} \tag{2.9.2}
$$

We notice that $M_P(N,s)$ has an $N$th-order pole at $s = -1, -3, -5, \ldots$ and an $N$th-order zero at $s = -2, -4, -6, \ldots$.

Once again, we are interested in the cumulants $b_i$ of the distribution of $\log|Z|$ and so we write

$$
M_P(N,s) = \exp(b_1 s + b_2 s^2/2 + b_3 s^3/3! + \cdots), \tag{2.9.3}
$$

and we calculate $b_i$ by taking the $i$th derivative at $s = 0$ of

$$
\log M_P(N,s) = N(\log\Gamma(1+s) - 2\log\Gamma(1+s/2)). \tag{2.9.4}
$$

Thus

$$
\begin{aligned}
b_1 &= N(\psi(1+s) - \psi(1+s/2))\big|_{s=0} = 0, \\
b_2 &= N\big(\psi^{(1)}(1) - \frac{1}{2}\psi^{(1)}(1)\big), \\
b_3 &= N\big(\psi^{(2)}(1) - \frac{1}{2^2}\psi^{(2)}(1)\big),
\end{aligned} \tag{2.9.5}
$$

and

$$\begin{aligned} b_n &= N\big(1-\frac{1}{2^{n-1}}\big)\psi^{(n-1)}(1) \\ &= N(-1)^n(n-1)!\big(1-\frac{1}{2^{n-1}}\big)\zeta(n). \end{aligned} \quad (2.9.6)$$

We see that these cumulants differ from those related to functions with zeros displaying random matrix statistics in that all the cumulants are of exactly the same order in $N$. We also note that as $N$ grows $M_P(N,s)$ is not of the order of any power of $N$, but rather varies exponentially with $N$.

However, if we normalize by the second moment and consider $\log|Z|/\sqrt{\frac{1}{2}\zeta(2)N}$, we find that the distribution as $N\to\infty$ is once again Gaussian.

$$\begin{aligned} &\langle\delta(x-\log|Z|/\sqrt{\zeta(2)N/2})\rangle \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iyx} M_P(N, iy/\sqrt{\zeta(2)N/2})dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iyx}\exp\left(\frac{(iy)^2}{2} - N\frac{\zeta(3)}{4}\frac{(iy)^3}{(\frac{1}{2}\zeta(2)N)^{3/2}}+\cdots\right)dy \\ &\sim \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right). \end{aligned} \quad (2.9.7)$$

Now we consider the mean values of $Z(\theta)/Z^*(\theta)$, as these will lead to the distribution of values of the imaginary part of the log of $Z$,

$$\operatorname{Im}\log Z(\theta) = -\sum_{n=1}^{N}\sum_{m=1}^{\infty}\frac{\sin[(\theta_n-\theta)m]}{m}. \quad (2.9.8)$$

The mean values are therefore

$$\begin{aligned}
&\left\langle \left( \frac{Z(\theta)}{Z^*(\theta)} \right)^{s/2} \right\rangle \\
&= \langle e^{is\mathrm{Im}\log Z(\theta)} \rangle \\
&= \frac{1}{(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{n=1}^{N} \exp\left( -is \sum_{m=1}^{\infty} \frac{\sin[(\theta_n - \theta)m]}{m} \right) d\theta_1 \cdots d\theta_N \\
&= \left( \frac{1}{2\pi} \int_0^{2\pi} \exp\left( -is \sum_{m=1}^{\infty} \frac{\sin[(\theta_n - \theta)m]}{m} \right) d\theta_n \right)^N \\
&= \left( \frac{1}{2\pi} \int_0^{2\pi} \exp\left( -is \sum_{m=1}^{\infty} \frac{\sin[\theta_n m]}{m} \right) d\theta_n \right)^N .
\end{aligned} \tag{2.9.9}$$

With the help of

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{\pi - x}{2}, \tag{2.9.10}$$

which holds for $0 < x < 2\pi$, we obtain

$$\left\langle \left( \frac{Z}{Z^*} \right)^{s/2} \right\rangle = \left( \frac{1}{2\pi} \int_0^{2\pi} \exp\left( -is \frac{\pi - \theta_n}{2} \right) d\theta_n \right)^N . \tag{2.9.11}$$

Next, with a change of variables $\phi_n = \theta_n/2 - \pi/2$, we find

$$\begin{aligned}
\left\langle \left( \frac{Z}{Z^*} \right)^{s/2} \right\rangle &= \left( \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \exp(is\phi_n) 2 d\phi_n \right)^N \\
&= \left( \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \exp(is\phi_n) d\phi_n \right)^N \\
&= \left( \frac{2}{\pi} \frac{\sin(s\pi/2)}{s} \right)^N .
\end{aligned} \tag{2.9.12}$$

We can write this answer in terms of gamma functions using

$$\Gamma(1+z)\Gamma(1-z) = \frac{\pi z}{\sin(\pi z)}, \tag{2.9.13}$$

so that finally we have

$$\left\langle \left( \frac{Z}{Z^*} \right)^{s/2} \right\rangle = \left( \frac{1}{\Gamma(1+s/2)\Gamma(1-s/2)} \right)^N \equiv L_P(N,s). \tag{2.9.14}$$

To determine the cumulants $B_j$ of the distribution of the imaginary part of the logarithm of $Z$, a function with uncorrelated zeros, we take successive derivatives at $s=0$ of

$$\log L_P(N,-is) = N(-\log\Gamma(1-is/2) - \log\Gamma(1+is/2)). \tag{2.9.15}$$

This results in

$$\begin{aligned}
B_1 &= N\big( -\frac{-i}{2}\psi(1-is/2) - \frac{i}{2}\psi(1+is/2)\big)\big|_{s=0} = 0, \qquad (2.9.16)\\
B_2 &= N\big( -\frac{(-i)^2}{2^2}\psi^{(1)}(1-is/2) - \frac{i^2}{2^2}\psi^{(1)}(1+is/2)\big)\big|_{s=0} = \frac{N}{2}\psi^{(1)}(1),\\
B_3 &= N\big( -\frac{(-i)^3}{2^3}\psi^{(2)}(1-is/2) - \frac{i^3}{2^3}\psi^{(2)}(1+is/2)\big)\big|_{s=0} = 0,\\
B_4 &= N\big( -\frac{(-i)^4}{2^4}\psi^{(3)}(1-is/2) - \frac{i^4}{2^4}\psi^{(3)}(1+is/2)\big)\big|_{s=0} = -\frac{N}{2^3}\psi^{(3)}(1).
\end{aligned}$$

Thus the general form is

$$B_n = \begin{cases} 0 & \text{if } n \text{ odd} \\ \frac{(-1)^{n/2+1}N}{2^{n-1}}\psi^{(n-1)}(1) = \frac{N}{2^{n-1}}(-1)^{n/2+1}(n-1)!\zeta(n) & \text{if } n \text{ even} \end{cases} . \tag{2.9.17}$$

Once more we notice that each of the non-zero cumulants is of order $N$. In this case of the imaginary part of the logarithm, as opposed to the real part, the odd cumulants are zero, implying that the distribution of values of $\operatorname{Im}\log Z$ is symmetric. If we normalize by the second log moment, we discover our usual Gaussian distribution in the limit as $N \to \infty$.

$$\begin{aligned}
&\langle\delta(x - \operatorname{Im}\log Z(\theta)/\sqrt{\zeta(2)N/2})\rangle \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iyx} L_P(N, y/\sqrt{\zeta(2)N/2})dy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iyx}\exp\left(-\frac{y^2}{2} - \frac{N}{32}\zeta(4)\frac{y^4}{(\frac{1}{2}\zeta(2)N)^2} - \cdots\right) dy \\
&\sim \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right). \qquad (2.9.18)
\end{aligned}$$

# Chapter 3

# Enter the Riemann Zeta Function

We now want to put to use the random matrix calculations developed in Chapter 2. As was mentioned in Chapter 1, the Riemann zeta function, $\zeta(s)$, has zeros which, high on the critical line, have statistics tending to those of the eigenvalues of the CUE ensemble. Thus the zeta function is comparable to the characteristic polynomial, $Z(U, \theta)$, introduced in the previous chapter, in the sense that they have the same zero statistics. Much study has been made of mean values of the zeta function and its logarithm along the critical line and it is our purpose to show that we can obtain a great deal of insight into these mean values by comparisons with the random matrix calculations. Mean values of the logarithm of $\zeta(s)$ are predicted asymptotically by random matrix theory (RMT) and moments of zeta itself show a distinct split into a contribution specific to the Riemann zeta function and one which is purely derived from random matrix calculations and depends only on the zero distribution of the zeta function. Again this result is in the asymptotic limit as we move very high up on the critical line.

## 3.1 The value distribution of $\log\zeta(1/2+it)$

We will examine first what is known about the distribution of $\log\zeta(1/2+it)$. Of most interest to us at this point is an unpublished theorem by Selberg. It states that for a rectangle $E$ in $\mathbb{R}^2$,

$$\lim_{T\to\infty}\frac{1}{T}\left|\left\{t : T\leq t\leq 2T, \frac{\log\zeta(1/2+it)}{\sqrt{1/2\log\log T}}\in E\right\}\right| = \frac{1}{2\pi}\int\int_E e^{-(x^2+y^2)/2}dx\,dy, \tag{3.1.1}$$

that is, in the limit as $T$, the height up the critical line, tends to infinity, the value distribution of the real and imaginary part of $\log\zeta(1/2+iT)/\sqrt{(1/2)\log\log T}$ each tend independently to a Gaussian.

As we pass to lower than leading-order results, however, the amount known analytically for the Riemann zeta function decreases. We turn, therefore, to the numerical work of Odlyzko [Odl97]. Inspired by Selberg's theorem, Odlyzko has carried out extensive numerical computations on the value distributions of the real and imaginary part of the log of the Riemann zeta function, expecting to see agreement with the Gaussian curve. Although this work is performed extraordinarily high up the critical line (where $T\sim 10^{19}$), still the distributions have not converged to the Gaussian, as can be seen in Figure 3.1.

Odlyzko also studied the moments of the distribution. If $P(x)$ is one of the curves in Figure 3.1, then the $k^{th}$ moment is $\int_{-\infty}^{\infty}x^kP(x)dx$. These are shown below in Table 3.1 as a comparison between the moments of the Gaussian distribution (with mean zero and unit standard deviation) and the numerical Riemann zeta moments normalized in the same way. Each of the two sets of Odlyzko's Riemann zeta function moments is calculated using a slightly different interval along the critical line, so the extent to which they disagree gives an estimate of the error due to the choice of the range over which the zeta function is averaged.

It can be seen that the odd moments for the Riemann zeta function are not

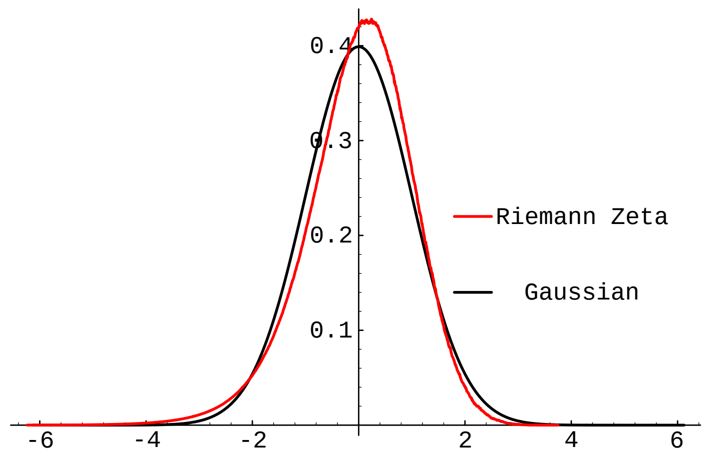


Figure 3.1: Numerical data (due to Odlyzko) for the probability density function for the values of $\log|\zeta(1/2+it)|$ near the $10^{20th}$ Riemann zero compared to the standard Gaussian

| Moment | Odlyzko a) | Odlyzko b) | Normal |
|---|---|---|---|
| 1 | 0.0 | 0.0 | 0 |
| 2 | 1.0 | 1.0 | 1 |
| 3 | -0.53625 | -0.55069 | 0 |
| 4 | 3.9233 | 3.9647 | 3 |
| 5 | -7.6238 | -7.8839 | 0 |
| 6 | 38.434 | 39.393 | 15 |
| 7 | -144.78 | -148.77 | 0 |
| 8 | 758.57 | 765.54 | 105 |
| 9 | -4002.5 | -3934.7 | 0 |
| 10 | 24060.5 | 22722.9 | 945 |

Table 3.1: Moments of the distribution of $\log|\zeta(1/2+it)|$ near the $10^{20}$th zero ($T \approx 1.520 \times 10^{19}$)

zero, as they are for the Gaussian. In fact the $9^{th}$ moment is hugely different, indicating again that convergence to the Gaussian distribution is very slow.

For the imaginary part of the logarithm of zeta Odlyzko has made similar computations. The distribution is shown in Figure 3.2.

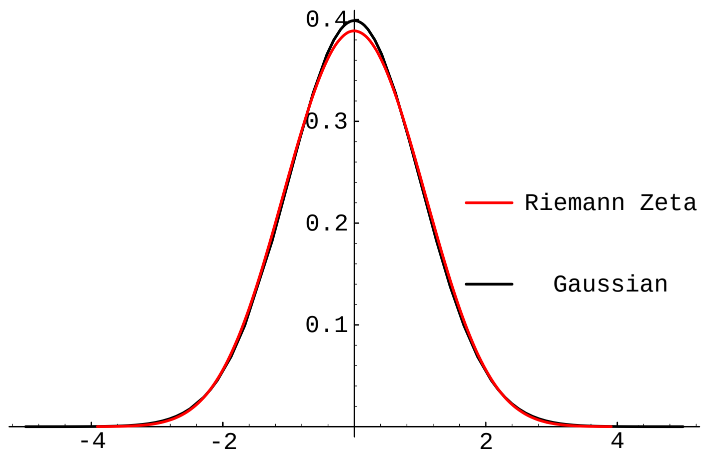


Figure 3.2: Numerical data (due to Odlyzko) for the probability density function for the values of $\operatorname{Im}\log\zeta(1/2+it)$ near the $10^{20th}$ Riemann zero compared to the standard Gaussian

In this case the distribution appears symmetric, and so looks more similar to the Gaussian, but it can be seen that here also the curve has not converged to its limiting form, despite the height up the critical line at which the computation was made. The table of moments in Table 3.2 also indicates how far the moments are from being Gaussian.

| Moment | $\zeta$ | Normal |
|---|---|---|
| 1 | $-6.3 \times 10^{-6}$ | 0 |
| 2 | 1.0 | 1 |
| 3 | $-4.7 \times 10^{-4}$ | 0 |
| 4 | 2.831 | 3 |
| 5 | $-9.1 \times 10^{-3}$ | 0 |
| 6 | 12.71 | 15 |
| 7 | -0.140 | 0 |
| 8 | 76.57 | 105 |

Table 3.2: Moments of $\operatorname{Im}\log\zeta(1/2+it)$ near the $10^{20}$th zero ($T \approx 1.520\times10^{19}$).

## 3.2 Comparison with the random matrix calculations

We now wish to test the predictions of random matrix theory against the Riemann zeta function results of the previous section. To do this we need to know how to compare the asymptotic variables $T$, for the Riemann zeta function, and $N$, for random matrices. The limits as each of these parameters tends to infinity are equivalent, as in each case we are considering a growing density of zeros, but the relative rate of approach to the limit remains to be determined. We accomplish this by equating the density of zeros in the two cases. For the Riemann zeta function, this density is asymptotic to $(1/2\pi)\log(T/2\pi)$, while there are $N$ CUE zeros on the unit circle, yielding a density of $N/2\pi$. Thus the equivalence is

$$N \sim \log(T/2\pi). \tag{3.2.1}$$

The results of the previous section already look promising when compared with the random matrix results of the Chapter 2. Selberg's result in (3.1.1) agrees with the fact that in the CUE case the distribution of the real and imaginary part of the logarithm of $Z$ are independently Gaussian in the large $N$ limit. It seems, therefore, that the log distribution is universal in the asymptotic limit; the random matrix calculation predicts exactly the known value distribution for the log of the zeta function.

To consider lower order, non-Gaussian, contributions we plot the exact distribution of $\log|Z|$, as shown in (2.3.3), against Odlyzko's data. As Odlyzko's calculations are performed around the $10^{20}$th zero (T=$1.520\times 10^{19}$), the proper value of $N$ is 42 (although the distribution does not change visibly if $N$ is varied by $\pm 1$). The remarkable result is shown in Figure 3.3. We can also produce a table for the moments of the real part of the log, Table 3.3. This is the same as Table 3.1 except that the random matrix moments are included.

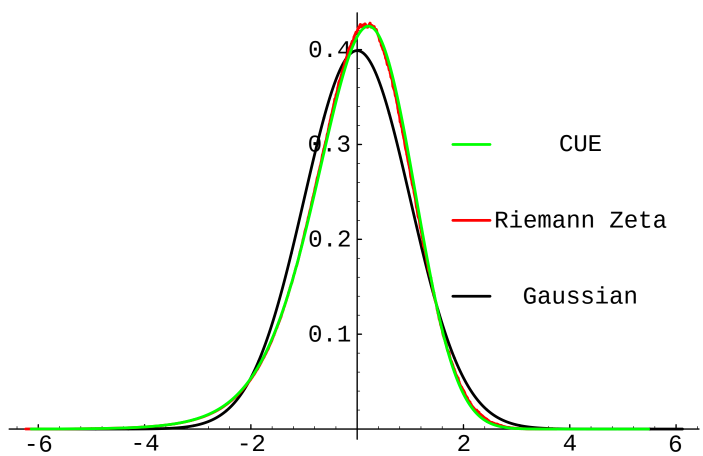


Figure 3.3: A comparison of the CUE value distribution for $\log|Z|$ with N=42, Odlyzko's data for the distribution of $\log|\zeta|$ near the $10^{20th}$ zero and the standard Gaussian

Both the value distribution and its moments show that although the common asymptotic Gaussian limit has not been reached, still the value distribution of the log of the real part of the zeta function and that of the CUE characteristic polynomial agree convincingly well.

| Moment | $\zeta$ a) | $\zeta$ b) | CUE | Normal |
|---|---|---|---|---|
| 1 | 0.0 | 0.0 | 0.0 | 0 |
| 2 | 1.0 | 1.0 | 1.0 | 1 |
| 3 | -0.53625 | -0.55069 | -0.56544 | 0 |
| 4 | 3.9233 | 3.9647 | 3.89354 | 3 |
| 5 | -7.6238 | -7.8839 | -7.76965 | 0 |
| 6 | 38.434 | 39.393 | 38.0233 | 15 |
| 7 | -144.78 | -148.77 | -145.043 | 0 |
| 8 | 758.57 | 765.54 | 758.036 | 105 |
| 9 | -4002.5 | -3934.7 | -4086.92 | 0 |
| 10 | 24060.5 | 22722.9 | 25347.77 | 945 |

Table 3.3: Moments of $\log|\zeta|$ near the $10^{20}$th zero ($T \approx 1.520 \times 10^{19}$) compared with the real part of the log for the CUE characteristic polynomial with $N = 42$.

If we now move on to the imaginary part of the logarithm, the results are not as impressive, as the distributions are already very close to Gaussian, but none the less they suggest that the Riemann zeta distribution agrees with the ensemble average at lower than just the Gaussian leading order. The probability density functions for the values achieved by the imaginary part of the logarithm are shown in Figure 3.4. The moments of this distribution are tabulated in Table 3.4. Once again, these are all standardized so as to be compared with the Gaussian of mean zero and variance one.

| Moment | $\zeta$ | CUE | Normal |
|---|---|---|---|
| 1 | $-6.3 \times 10^{-6}$ | 0.0 | 0 |
| 2 | 1.0 | 1.0 | 1 |
| 3 | $-4.7 \times 10^{-4}$ | 0.0 | 0 |
| 4 | 2.831 | 2.87235 | 3 |
| 5 | $-9.1 \times 10^{-3}$ | 0.0 | 0 |
| 6 | 12.71 | 13.29246 | 15 |
| 7 | -0.140 | 0.0 | 0 |
| 8 | 76.57 | 83.76939 | 105 |

Table 3.4: Moments of $\operatorname{Im}\log\zeta$ near the $10^{20}$th zero ($T \approx 1.520 \times 10^{19}$) compared with $\operatorname{Im}\log Z$ for $N = 42$.

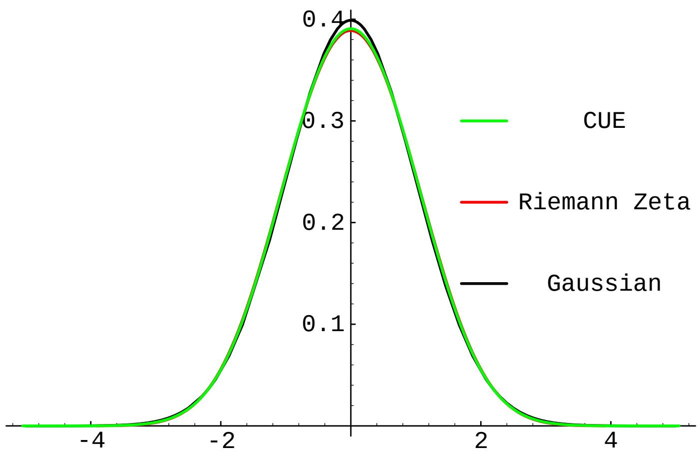


Figure 3.4: A comparison of the CUE distribution for $\operatorname{Im}\log Z$ with N=42, Odlyzko's data for the Riemann Zeta function, $\operatorname{Im}\log\zeta$, near the $10^{20th}$ zero and the standard Gaussian

The excellent agreement found between the Riemann zeta function and random matrix theory is surprising not only because it implies that merely the manner of distribution of zeros determines the statistics of the values attained by a function's logarithm, but also because it implies that not only the Gaussian leading-order term in the distribution is universal, but something of the lower orders too.

We know, however, that the moments of the log of the Riemann zeta function are not universal at all orders because the next to leading order term of the second moment of the imaginary part of the log has been calculated by Goldston [Gol87] assuming Montgomery's pair correlation conjecture for the Riemann zeros (the sign on the $1/m$ term is incorrect in the above reference):

$$
\begin{aligned}
&\frac{1}{T}\int_0^T (\operatorname{Im}\log\zeta(1/2+it))^2 dt \\
&\sim \frac{1}{2}\log\log\frac{T}{2\pi} + \frac{1}{2}(\gamma+1) + \sum_{m=2}^{\infty}\sum_p \left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}.
\end{aligned} \tag{3.2.2}
$$

Here we see immediately that there are non-universal contributions in the constant term. Prime numbers are a feature specific to the Riemann zeta function and we cannot expect that our random matrix calculations will predict these prime sums; they are beyond the universal range of random matrix theory.

If we compare, however, (3.2.2) with the random matrix second moment,

$$
\langle(\operatorname{Im}\log Z)^2\rangle_{CUE} = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + O(N^{-2}), \tag{3.2.3}
$$

remembering the equivalence $N \sim \log(T/2\pi)$, we see that RMT predicts everything *but* the prime sums. There appears to be a neat division between the universal and the non-universal components of these moments.

In order to study these non-universal terms a little better, we will turn to mean values of combinations of the zeta function itself, rather than its logarithm.

## 3.3 Mean values of the Riemann zeta function

In this section we will review the current knowledge in the number theoretical community on the moments of $|\zeta(1/2+it)|$. It is conjectured that the following limit exists:

$$
\begin{aligned}
&\lim_{T\to\infty}\frac{1}{(\log T)^{\lambda^2}}\frac{1}{T}\int_0^T |\zeta(1/2+it)|^{2\lambda}dt \\
&\equiv \lim_{T\to\infty}\frac{1}{(\log T)^{\lambda^2}}\frac{1}{T}I_\lambda(T) = c(\lambda).
\end{aligned} \tag{3.3.1}
$$

The evidence for this follows. If we start our history with the integer moments, the second moment, evaluated to first order by Hardy and Littlewood [HL18], has been calculated in the form

$$\int_0^T |\zeta(1/2+it)|^2 dt = T\log(T/2\pi) + (2\gamma - 1)T + O(T^{1/2+\epsilon}) \qquad (3.3.2)$$

by Atkinson [Atk39], and with an error term of $O(T^{1/2}\log T)$ by Ingham [Ing26].

Headway has also been made with the fourth moment. While Ingham [Ing26] first came up with the leading-order term, Heath-Brown [HB79] gives us

$$\int_0^T |\zeta(1/2+it)|^4 dt = \sum_{n=0}^{4} c_n T(\log T)^n + O(T^{7/8+\epsilon}), \qquad (3.3.3)$$
$$\text{where } c_4 = \frac{1}{2\pi^2}, \qquad c_3 = \frac{2}{\pi^2}\left(4\gamma - 1 - \log(2\pi) - 12\frac{\zeta'(2)}{\pi^2}\right).$$

It is interesting to note that if the moments are defined slightly differently, with a smooth cut-off instead of finite limits to the integral, then the fourth moments agree at leading order, but not at next-to-leading order [Atk41].

$$\int_0^\infty |\zeta(1/2+it)|^4 e^{-\delta t} dt = \frac{1}{\delta}\left(A\log^4\frac{1}{\delta} + B\log^3\frac{1}{\delta} + C\log^2\frac{1}{\delta} + D\log\frac{1}{\delta} + E + O\left(\left(\frac{1}{\delta}\right)^{13/14+\epsilon}\right)\right),$$
$$A = \frac{1}{2\pi^2}, \qquad B = -\frac{2}{\pi^2}\left(\log 2\pi - 3\gamma + 12\frac{\zeta'(2)}{\pi^2}\right). \qquad (3.3.4)$$

Although nothing has been proven about the higher integer moments, there is a conjecture by Conrey and Ghosh [CG92] that

$$\int_0^T |\zeta(1/2+it)|^6 dt \sim \frac{42}{9!} a_3 T \log^9 T, \qquad (3.3.5)$$

where

$$a_\lambda = \prod_p (1-1/p)^{\lambda^2} \left( \sum_{m=0}^{\infty} \left( \frac{\Gamma(\lambda+m)}{m!\Gamma(\lambda)} \right)^2 p^{-m} \right), \tag{3.3.6}$$

and one for the 8th moment by Conrey and Gonek [CG98] that

$$\int_0^T |\zeta(1/2+it)|^8 dt \sim \frac{24024}{16!} a_4 T \log^{16} T. \tag{3.3.7}$$

In addition, there is a whole range of results supporting the conjecture (3.3.1). For a start, Conrey and Ghosh [CG84], assuming the Riemann hypothesis (RH), deduced a lower bound for moments with $k \geq 0$, which was extended to $k > -1/2$ by Gonek (with the extra assumption that the zeros are all simple):

$$\begin{aligned} I_k(T) &\geq [C_k + o(1)] T (\log T)^{k^2} \quad (3.3.8) \\ \text{where} \quad C_k &= (\Gamma(k^2+1))^{-1} \prod_p \left[ (1-1/p)^{k^2} \sum_{m=0}^{\infty} \left( \frac{\Gamma(k+m)}{m!\Gamma(k)} \right)^2 p^{-m} \right]. \end{aligned}$$

If we write (noting that if conjecture (3.3.1) holds then $\lim_{T\to\infty} g_k(T)$ will converge to a constant)

$$g_k(T) \equiv \left( \frac{a_k}{\Gamma(1+k^2)} T \log^{k^2} T \right)^{-1} I_k(T), \tag{3.3.9}$$

then Heath-Brown [HB93] proved an upper bound, again with RH, for $0 \leq k < 2$: $g_k(T) \leq \frac{2}{(k^2+1)(2-k)}$.

Heath-Brown also proved that [HB81]

$$T(\log T)^{k^2} << I_k(T) << T(\log T)^{k^2} \tag{3.3.10}$$

for $k = 1/n$ with $n$ an integer. Here we define $A << B$ to maen that there exists a positive constant $c$ such that $|A| \leq c|B|$ for all $T$. Jutila [Jut83] noted that the constants implied by the inequalities in (3.3.10) are independent of k. Also, $I_k(T) >> T(\log T)^{k^2}$ for half-integers [Ram80] and for all positive rationals [HB81].

The results in the previous paragraph are regardless of the truth of RH, but if it is assumed to be correct then Ramachandra showed ([Ram80] and [Ram78] respectively)

$$\begin{aligned} I_k &<< T(\log T)^{k^2} \quad 0 < k < 2 \\ I_k &>> T(\log T)^{k^2} \quad k \geq 0. \end{aligned} \tag{3.3.11}$$

There is also a result for the moment with the smooth cut-off [Tit86] p174, for integer $k \geq 0$

$$\int_0^\infty |\zeta(1/2+it)|^{2k} e^{-\delta t} dt >> \frac{1}{\delta} \log^{k^2} \frac{1}{\delta}. \tag{3.3.12}$$

Two more results of Conrey and Ghosh [CG92] will be very useful to us as we propose a conjecture for $g_\lambda$ (see (3.3.9)). They are

**Theorem 4.** *If $\kappa = 0$ or if the Riemann Hypothesis holds and $\kappa = 1$, then*

$$\begin{aligned} &\lim_{T\to\infty} \frac{d}{dk} \left( \frac{1}{T(\log T)^{k^2}} \int_1^T |\zeta(1/2+it)|^{2k} dt \right) \Big|_{k=\kappa} \qquad (3.3.13) \\ &= \frac{d}{dk} \left( \frac{1}{\Gamma(1+k^2)} \prod_p (1-1/p)^{k^2} \left( \sum_{m=0}^\infty \left( \frac{\Gamma(k+m)}{m!\Gamma(k)} \right)^2 p^{-m} \right) \right) \Big|_{k=\kappa} \\ &= \begin{cases} o(1) & \text{if } \kappa = 0 \\ 2\gamma - 2 & \text{if } \kappa = 1 \end{cases} \end{aligned}$$

and

**Theorem 5.** *Assuming RH and the pair correlation conjecture*

$$\begin{aligned} &\lim_{T\to\infty} \frac{d^2}{dk^2} \left( \frac{1}{T(\log T)^{k^2}} \int_1^T |\zeta(1/2+it)|^{2k} dt \right) \Big|_{k=0} \qquad (3.3.14) \\ &= \frac{d^2}{dk^2} \left( \frac{1}{\Gamma(1+k^2)} \prod_p (1-1/p)^{k^2} \left( \sum_{m=0}^\infty \left( \frac{\Gamma(k+m)}{m!\Gamma(k)} \right)^2 p^{-m} \right) \right) \Big|_{k=0} + 2 \\ &= 2 \sum_p \sum_{m=1}^\infty \left( \frac{1}{m^2} - \frac{1}{m} \right) p^{-m} + 2\gamma + 2 \end{aligned}$$

These theorems are of interest for Conrey and Ghosh in the study of their lower bound for the moments $I_k$, but they help us test our conjecture on the mean values of the Riemann zeta function, described in the following section, because the derivatives of our conjectured moments agree with those of the Riemann zeta function at the points mentioned above.

## 3.4 A conjecture on the mean values of the Riemann zeta function

We have already calculated the $s^{th}$ moment of $|Z|$; it is embodied in the function $M_N(s)$ featuring in (2.1.9). This is very useful as $M_N(2\lambda)$ is exactly what we need to compare to the Riemann zeta moment,

$$\frac{1}{T} I_\lambda(T) = \frac{1}{T} \int_0^T |\zeta(1/2 + it)|^{2\lambda} dt, \tag{3.4.1}$$

discussed in the preceding section.

Conrey and Ghosh have expressed the conjecture (3.3.1) in the form

$$\frac{1}{T} I_\lambda(T) \sim g_\lambda \frac{a_\lambda}{\Gamma(1+\lambda^2)} \log^{\lambda^2} T, \tag{3.4.2}$$

where

$$a_\lambda = \left\{ \prod_p (1 - 1/p)^{\lambda^2} \left( \sum_{m=0}^{\infty} \left( \frac{\Gamma(\lambda + m)}{m! \Gamma(\lambda)} \right)^2 p^{-m} \right) \right\}. \tag{3.4.3}$$

To simplify our notation, we will define

$$g_\lambda / \Gamma(1 + \lambda^2) = f_\lambda. \tag{3.4.4}$$

We now consider what random matrix theory might be expected to contribute to the pursuit of these Riemann moments.

Before we begin, it must be stated that, unlike the case of the log moments, it is not to be expected that random matrix theory (RMT) will predict even the leading-order term of these moments of $|\zeta|$. For a start, we know that there are non-universal contributions, ie. the prime sums, in the next-to-leading-order terms of the log moments, and upon exponentiation, additive contributions become multiplicative. Even more important, however, RMT deals with statistics on *short* energy scales; we saw in Section 1.3 that eigenvalue correlations over energy scales long compared to the mean level spacing depend on the distinctive short orbits of a system, or in the case of the Riemann zeta function, long range correlations depend on the low prime numbers in a non-universal way. Because of this, if the value of a function is determined, at least asymptotically, by the positions of just the nearest zeros to the point under consideration and the zeros have the statistics of the eigenvalues of a RMT ensemble, then perhaps RMT might predict the value distribution of that function. However, if the function depends on much more distant zeros, then the long range correlations of those zeros are likely to induce non-universal contributions in the value distribution of the function. For the *log* of the zeta function there is an explicit asymptotic formula involving a sum over just the zeros close to the point at which the function is being evaluated, but there is no such formula for the zeta function itself; its value depends on all the zeros.

Despite this warning remark, we will forge ahead to examine the moments of $|Z|$ for any relation to the moments of $|\zeta|$. The first thing to notice is that the asymptotic result, studied at length in Section 2.7, is

$$M_N(2\lambda) = \langle |Z|^{2\lambda} \rangle_{CUE} \sim f_\lambda^{CUE} N^{\lambda^2}, \tag{3.4.5}$$

where $f^{CUE}$ doesn't depend on $N$. Remembering the equivalence between $N$ and $\log T$, (3.2.1), we see that this implies asymptotic behaviour of the same order as (3.4.2). We recall that for integer moments, this coefficient, $f^{CUE}$ has the simple form (2.7.1):

$$f_k^{CUE} = \prod_{n=0}^{k-1} \frac{n!}{(k+n)!}. \tag{3.4.6}$$

The conjecture which we propose concerning this coefficient (not only for integer $\lambda$) is that

$$f_\lambda = f_\lambda^{CUE}. \tag{3.4.7}$$

That is, we conjecture that the coefficient of the Riemann zeta function moments breaks neatly into two factors: one containing the product over primes, $a_\lambda$, certainly non-universal information, and one derived directly from random matrix theory, $f_\lambda$. We can also write this conjecture in another manner

$$\begin{aligned} &\lim_{T\to\infty} (\log T)^{-\lambda^2} \frac{1}{T} \int_0^T |\zeta(1/2+it)|^{2\lambda} dt \\ &= \lim_{N\to\infty} N^{-\lambda^2} M_N(2\lambda) \times a_\lambda, \end{aligned} \tag{3.4.8}$$

where $a_\lambda$ is defined by (3.4.3).

Before we examine the evidence, both numerical and analytic, for this conjecture, a brief justification adds to its credibility. Remembering that it is low primes, in analogy with the short periodic orbits, that are responsible for non-universal contributions, we note that if we replaced each prime $p$ with $p^t$ and increased $t$, the prime product $a_\lambda$ would tend to one. That is, if we somehow imagine all the primes growing in size, and thus leaving less and less non-universal information, then the limit $\lim_{T\to\infty}(1/T)(\log T)^{-\lambda^2} I_\lambda(T)$ tends to the random matrix coefficient, $f_\lambda^{CUE}$, just as it should do if the conjecture is true.

More substantial evidence, however, is provided by comparing the random matrix results with the few known values of $f_\lambda$. The result of such an exercise is

$$\begin{aligned} f_0^{CUE} &= 1 = f_0 \\ f_1^{CUE} &= 1 = f_1 \\ f_2^{CUE} &= \frac{1}{12} = f_2 \\ f_3^{CUE} &= \frac{42}{9!} = f_3 \ ? \\ f_4^{CUE} &= \frac{24024}{16!} = f_4 \ ? \end{aligned} \tag{3.4.9}$$

where the question mark indicates those values of $f_\lambda$ which are conjectures. Another encouraging observation is that at integer values of $k$, $f_k^{CUE} \geq 1/\Gamma(1+k^2)$, where the right hand side is Conrey and Ghosh's lower bound.

This is as far as we can go in checking the conjecture against analytical results for the mean values of the Riemann zeta function. However, Odlyzko [Odl97] has computed mean values for non-integral $\lambda$. These are shown in Table 3.5. The CUE result, without the prime product $a_\lambda$ is included simply to illustrate how clear it is that that alone, without the non-universal factor, does not even come close to predicting the Riemann zeta value. It should be noted that the limits $T \to \infty$ and $N \to \infty$ in (3.4.8) are far from being realized in these numerics; $N$ is only 42. This is why it cannot be expected that the column headed 'CUE with prime product' exactly match the zeta function column; even apart from any purely numerical errors. However, despite this the agreement is quite convincing.

We can also gain support for the conjecture by comparing the distribution of values of $|Z|$ and $|\zeta(1/2+it)|$.

Starting with $|Z|$, let us call the distribution of values $P_N(x)$. So

$$M_N(s) = \int_0^\infty x^s P_N(x) dx, \tag{3.4.10}$$

which gives us

| $\lambda$ | CUE with prime product | $r(\lambda, H)$ | $C_\lambda$ (lower bound) | CUE | % error CUE with primes | % error CUE |
|---|---|---|---|---|---|---|
| 0.1 | 1.011 | 1.004 | 1.0042 | 1.0129 | 0.741 | 0.886 |
| 0.2 | 1.038 | 1.034 | 1.0172 | 1.0430 | 0.395 | 0.870 |
| 0.3 | 1.071 | 1.067 | 1.0381 | 1.0803 | 0.423 | 1.25 |
| 0.4 | 1.105 | 1.098 | 1.064 | 1.1171 | 0.649 | 1.74 |
| 0.5 | 1.133 | 1.123 | 1.0904 | 1.1466 | 0.914 | 2.10 |
| 0.6 | 1.151 | 1.135 | 1.1113 | 1.1631 | 1.37 | 2.25 |
| 0.7 | 1.152 | 1.132 | 1.1195 | 1.1616 | 1.77 | 2.26 |
| 0.8 | 1.133 | 1.107 | 1.1076 | 1.1386 | 2.38 | 2.85 |
| 0.9 | 1.091 | 1.06 | 1.069 | 1.0925 | 2.92 | 3.07 |
| 1 . | 1.024 | 0.989 | 1. | 1.0238 | 3.52 | 3.52 |
| 1.1 | 0.933 | 0.896 | 0.901 | 0.9350 | 4.16 | 4.35 |
| 1.2 | 0.822 | 0.787 | 0.776 | 0.8307 | 4.48 | 5.55 |
| 1.3 | 0.699 | 0.667 | 0.637 | 0.7167 | 4.89 | 7.45 |
| 1.4 | 0.571 | 0.544 | 0.494 | 0.5996 | 4.99 | 10.2 |
| 1.5 | 0.446 | 0.426 | 0.36 | 0.4858 | 4.65 | 14.0 |
| 1.6 | 0.333 | 0.319 | 0.246 | 0.3806 | 4.27 | 19.3 |
| 1.7 | 0.237 | 0.229 | 0.157 | 0.2880 | 3.37 | 25.8 |
| 1.8 | 0.158 | 0.156 | 0.092 | 0.2103 | 1.41 | 34.8 |
| 1.9 | 0.100 | 0.101 | 0.05 | 0.1480 | 0.542 | 46.5 |
| 2. | 0.0602 | 0.0624 | 0.025 | 0.1003 | 3.53 | 60.7 |

Table 3.5: Comparison of the mean value of $r(\lambda, H) = H^{-1}(\log T)^{-\lambda^2}\int_T^{T+H}|\zeta(1/2+it)|^{2\lambda}dt$ calculated numerically for the Riemann zeta function near the $10^{20th}$ zero, the equivalent quantity using random matrix theory, with and without the prime product, ($N=42$) and the lower bound on the leading-order coefficient (Conrey and Ghosh), $C_\lambda$.

$$M_N(is) = \int_0^\infty e^{is\log x} P_N(x)dx. \tag{3.4.11}$$

Taking a Fourier transform of each side we have

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-isy} M_N(is)ds &= \int_{-\infty}^{\infty}\int_0^\infty e^{-isy} e^{is\log x} P_N(x)dxds \\
&= 2\pi \int_0^\infty \delta(\log x - y)P_N(x)dx \\
&= 2\pi \int_{-\infty}^{\infty} \delta(z - y)P_N(e^z)e^z dz \\
&= 2\pi P_N(e^y)e^y. \qquad (3.4.12)
\end{aligned}$$

Thus

$$P_N(x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-is} M_N(is)ds. \tag{3.4.13}$$

For any finite $N$ we can plot $P_N(x)$ by numerical evaluation of (3.4.13). This is done in Figure 3.5 together with data for the value distribution of $|\zeta(1/2+it)|$ when $t \approx 10^6$, which corresponds via (3.2.1) to $N = 12$.

If we consider large $N$, we can make the approximation

$$\begin{aligned}
P_N(x) &= \frac{1}{2\pi x}\int_{-\infty}^{\infty} \exp(-is\log x - Q_2 s^2/2! - iQ_3 s^3/3! + Q_4 s^4/4! + \cdots)ds \\
&= \frac{1}{2\pi x\sqrt{Q_2}}\int_{-\infty}^{\infty} \exp\left(\frac{-is\log x}{\sqrt{Q_2}} - \frac{s^2}{2} - \frac{iQ_3 s^3}{Q_2^{3/2}3!} + \cdots\right) ds \\
&\sim \frac{1}{2\pi x\sqrt{Q_2}}\int_{-\infty}^{\infty} \exp\left(\frac{-is\log x}{\sqrt{Q_2}} - \frac{s^2}{2}\right) ds \qquad (3.4.14) \\
&= \frac{1}{x\sqrt{2\pi Q_2}} \exp\left(\frac{-\log^2 x}{2Q_2}\right). \qquad (3.4.15)
\end{aligned}$$

Here the parameters $Q_j$ are the cumulants studied in Section 2.2.

This approximation is valid if $\log x >> -\frac{1}{2}\log N$, this region being that in which the stationary point of (3.4.14), at $s^* = -i\log x/\sqrt{Q_2}$, stays well away from the first pole of $M_N(is/\sqrt{Q_2})$ at $s = i\sqrt{Q_2}$.

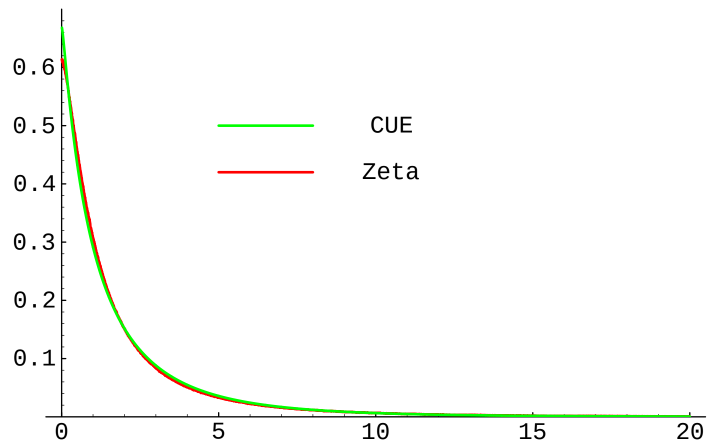


Figure 3.5: The CUE value distribution of $|Z|$ (green), corresponding to $N = 12$, with numerical data for the value distribution of $|\zeta(1/2+it)|$ (red) near $t = 10^6$.

As $x \to 0$, on the other hand, we note that the integral (3.4.13) is dominated by the pole of $M_N(is)$ at $s = i$. The poles further from the real axis result in higher powers of $x$, due to the factor $\exp(-is \log x)$, but the single power of $x$ resulting from the lowest pole cancels the $1/x$ outside the integral, yielding

$$\lim_{x \to 0} P_N(x) = \frac{1}{\Gamma(N)} \prod_{j=1}^{N} \left( \frac{\Gamma(j)}{\Gamma(j-1/2)} \right)^2 . \tag{3.4.16}$$

If $N$ is large, this is asymptotic to [HKO]

$$N^{1/4} (G(1/2))^2 = \exp\left( \frac{1}{12} \log 2 + 3\zeta'(-1) - \frac{1}{2} \log \pi \right) N^{1/4}. \tag{3.4.17}$$

All of this is relevant because when we write the conjecture as in (3.4.8), it suggests that as $t \to \infty$, the value distribution of $|\zeta(1/2+it)|$ might tend to

$$\tilde{P}_N(x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} e^{-is\log x} a(is/2) M_N(is) ds. \tag{3.4.18}$$

This would imply

$$\tilde{P}_N(0) = a(-1/2) P_N(0), \tag{3.4.19}$$

which agrees with numerical computations which show that $a(-1/2) \approx 0.919$ and $P_{12}(0) \approx 0.671$, yielding $a(-1/2)P_{12}(0) \approx 0.617$. This is indeed close to 0.613, which is the numerically computed value at zero of the probability density function for values of $|\zeta(1/2+it)|$.

Away from $x = 0$, in the region where (3.4.15) is valid, the stationary point is at $s^* = -i\log x/\sqrt{Q_2}$ so $a(is^*/2\sqrt{Q_2}) = a(\log x/(2Q_2))$. Since $a(0) = 1$, as long as $|\log x| << Q_2$ then $a$ is close to 1 and so the contribution from the prime product recedes to the tail of the distribution when $N$ is large. This is born out by the good agreement of the two curves in Figure 3.5.

Further evidence for the conjecture (3.4.7) is found by looking at the derivatives of the leading-order coefficient of the moments of $|\zeta(1/2+it)|$. The values which are known are stated in (3.3.13) and (3.3.14). Since we believe that

$$\lim_{N\to\infty} N^{-(s/2)^2} a_{s/2} M_N(s) = \lim_{T\to\infty} \frac{1}{(\log T)^{(s/2)^2}} \frac{1}{T}\int_0^T |\zeta(1/2+it)|^s dt, \tag{3.4.20}$$

we will take derivatives of the left hand side and see if they tally with the Riemann zeta function results. We will use the notation that

$$\begin{aligned}\langle |Z|^s\rangle_{CUE}^P &= M_N(s) a_{s/2} \qquad (3.4.21)\\ &= \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+s)}{(\Gamma(j+s/2))^2} \cdot \prod_p \left\{ (1-1/p)^{(s/2)^2} \sum_{m=0}^{\infty} \left(\frac{\Gamma(m+s/2)}{m!\Gamma(s/2)}\right)^2 p^{-m} \right\},\end{aligned}$$

where the notation $P$ has nothing to do with the average, but just reminds us that we are artificially introducing prime contributions into the CUE average. Conrey and Gonek [CG98] have written $a_k$ ($k$ an integer) as

$$M_N(2k)a_k = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2k)}{(\Gamma(j+k))^2} \cdot \prod_p \left\{ (1-1/p)^{(k-1)^2} \sum_{r=0}^{k-1} \frac{\binom{k-1}{r}^2}{p^r} \right\}, \quad (3.4.22)$$

which makes calculating $a_k$ easier for integer values in the following work.

The quantities we want are

$$\begin{aligned}
&\lim_{N\to\infty} \frac{d}{dk} \frac{1}{N^{k^2}} \langle |Z|^{2k} \rangle_{CUE}^{P} \\
&= \lim_{N\to\infty} \left( \frac{1}{N^{k^2}} \langle 2\log|Z||Z|^{2k} \rangle_{CUE}^{P} - \frac{2k\log N}{N^{k^2}} \langle |Z|^{2k} \rangle_{CUE}^{P} \right) \\
&= \lim_{N\to\infty} \left( 2D_{CUE}^{P}(k,1) - 2k\log N D_{CUE}^{P}(k,0) \right), \qquad (3.4.23)
\end{aligned}$$

evaluated at $k=0$ and $k=1$, and

$$\begin{aligned}
&\lim_{N\to\infty} \frac{d^2}{dk^2} \frac{1}{N^{k^2}} \langle |Z|^{2k} \rangle_{CUE}^{P} \\
&= \lim_{N\to\infty} \left( \frac{1}{N^{k^2}} \langle 4(\log|Z|)^2 |Z|^{2k} \rangle_{CUE}^{P} - \frac{4k\log N}{N^{k^2}} \langle 2\log|Z||Z|^{2k} \rangle_{CUE}^{P} \right. \\
&\quad \left. -\frac{2\log N}{N^{k^2}} \langle |Z|^{2k} \rangle_{CUE}^{P} + \frac{4k^2(\log N)^2}{N^{k^2}} \langle |Z|^{2k} \rangle_{CUE}^{P} \right) \\
&= \lim_{N\to\infty} \left( 4D_{CUE}^{P}(k,2) - 8k\log N D_{CUE}^{P}(k,1) \right. \\
&\quad \left. +(4k^2(\log N)^2 - 2\log N) D_{CUE}^{P}(k,0) \right) \qquad (3.4.24)
\end{aligned}$$

at $k=0$, where

$$D_{CUE}^{P}(k,n) = \frac{1}{N^{k^2}} \langle |Z|^{2k} (\log|Z|)^n \rangle_{CUE}^{P}. \qquad (3.4.25)$$

To calculate these $D$ functions, we notice that

$$\begin{aligned}
\frac{d^n}{ds^n} a_{s/2} M_N(s) &= \frac{d^n}{ds^n} \langle |Z|^s \rangle_{CUE}^{P} \qquad (3.4.26) \\
&= \left\langle \frac{d^n}{ds^n} |Z|^s \right\rangle_{CUE}^{P} \\
&= \langle |Z|^s (\log|Z|)^n \rangle_{CUE}^{P}.
\end{aligned}$$

In the same vein as the calculations in Section 2.2, we have that

$$\frac{d^n}{ds^n} a_{s/2} M_N(s) = \sum_{m=1}^{n} \binom{n-1}{m-1} \frac{d^{n-m}(a_{s/2} M_N(s))}{ds^{n-m}} K_m(N,s), \tag{3.4.27}$$

where the cumulants are now

$$\begin{aligned} K_m(N,s) &= Q_m(N,s) + \frac{d^{m-1}}{ds^{m-1}} \left( \frac{d}{ds} \log a_{s/2} \right) \\ &= \sum_{j=1}^{N} \left( \psi^{(n-1)}(j+s) - \frac{1}{2^{n-1}} \psi^{(n-1)}(j+s/2) \right) \\ &+ \frac{d^{m-1}}{ds^{m-1}} \left( \frac{d}{ds} \log a_{s/2} \right), \end{aligned} \tag{3.4.28}$$

and

$$\frac{d}{ds} \log a_{s/2} = \tag{3.4.29}$$
$$\sum_p \left( \frac{\sum_{m=0}^{\infty} \frac{1}{p^m} \left( \frac{\Gamma(s/2+m)}{\Gamma(s/2)m!} \right)^2 \left( \psi(s/2+m) - \psi(s/2) + \frac{s}{2} \log(1-1/p) \right)}{\sum_{m=0}^{\infty} \left( \frac{\Gamma(s/2+m)}{\Gamma(s/2)m!} \right)^2 p^{-m}} \right).$$

In the above expression, $Q_m(N,s)$ is identical to the cumulant (2.2.8) derived for the real part of the log of $Z$ in Chapter 2.

We now want to calculate $D^P_{CUE}(0,0)$, $D^P_{CUE}(0,1)$, $D^P_{CUE}(1,0)$, $D^P_{CUE}(1,1)$ and $D^P_{CUE}(0,2)$. Again making use of what we know already,

$$D^P_{CUE}(0,0) = 1, \tag{3.4.30a}$$

$$\begin{aligned} D^P_{CUE}(1,0) &= \frac{1}{N} a_1 M_N(2) \\ &= \frac{1}{N}(1 \cdot (N+1)) = 1 + \frac{1}{N}, \end{aligned} \tag{3.4.30b}$$

$$D^P_{CUE}(0,1) = 0, \tag{3.4.30c}$$

$$
\begin{aligned}
&D^P_{CUE}(1,1) = \frac{1}{N}(a_1 M_N(2) K_1(N,2)) \\
&= \frac{1}{N}(1\cdot(N+1))\left(\log N + \gamma - 1 + \frac{3}{2N} + O\left(\frac{1}{N^2}\right)\right. \\
&\qquad \left.+\sum_p\left(\frac{\sum_{m=0}^{\infty}\frac{1}{p^m}(\psi(m+1)-\psi(1)+\log(1-1/p))}{\sum_{m=0}^{\infty}\frac{1}{p^m}}\right)\right) \\
&= (1+1/N)\left(\log N + \gamma - 1 + O\left(\frac{1}{N}\right)\right. \\
&\qquad \left.+\sum_p\left(\frac{\sum_{m=1}^{\infty}\frac{1}{p^m}\sum_{k=1}^{m}\frac{1}{k}}{\sum_{m=0}^{\infty}\frac{1}{p^m}}+\log(1-1/p)\right)\right) \\
&= (1+1/N)\Bigg(\log N + \gamma - 1 + O\left(\frac{1}{N}\right). \\
&\qquad +\sum_p\Bigg(\frac{\frac{1}{p}+\frac{1}{p^2}\left(1+\frac{1}{2}\right)+\frac{1}{p^3}\left(1+\frac{1}{2}+\frac{1}{3}\right)+\frac{1}{p^4}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right)+\cdots}{\sum_{m=0}^{\infty}\frac{1}{p^m}} \\
&\qquad +\log(1-1/p)\Bigg)\Bigg) \\
&= (1+1/N)\Bigg(\log N + \gamma - 1 + O\left(\frac{1}{N}\right) \\
&\qquad +\sum_p\Bigg(\frac{-\log(1-1/p)-\frac{1}{p}\log(1-1/p)-\frac{1}{p^2}\log(1-1/p)-\cdots}{\sum_{m=0}^{\infty}\frac{1}{p^m}} \\
&\qquad +\log(1-1/p)\Bigg)\Bigg) \\
&= (1+1/N)(\log N + \gamma - 1 + O\left(\frac{1}{N}\right) + 0) \\
&= \log N + \gamma - 1 + O\left(\frac{\log N}{N}\right),
\end{aligned}
\tag{3.4.30d}
$$

$$
\begin{aligned}
D^P_{CUE}(0,2) &= \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) \\
&+\frac{1}{2}\sum_p\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}+o(1).
\end{aligned}
\tag{3.4.30e}
$$

Therefore we have, from (3.4.23) and (3.4.24),

$$\lim_{N\to\infty} \frac{d}{dk}\frac{1}{N^{k^2}}\langle|Z|^{2k}\rangle^P_{CUE}\Big|_{k=0} = 0 \tag{3.4.31a}$$

$$\lim_{N\to\infty} \frac{d}{dk}\frac{1}{N^{k^2}}\langle|Z|^{2k}\rangle^P_{CUE}\Big|_{k=1} = 2\gamma - 2 \tag{3.4.31b}$$

$$\lim_{N\to\infty} \frac{d^2}{dk^2}\frac{1}{N^{k^2}}\langle|Z|^{2k}\rangle^P_{CUE}\Big|_{k=0} = 2\gamma + 2 + 2\sum_p\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}, \tag{3.4.31c}$$

which are precisely the Riemann zeta function results from (3.3.13) and (3.3.14). So, the derivative of our conjectured leading-order coefficient for the Riemann moment at $\lambda = 0$ and $\lambda = 1$, and the second derivative at $\lambda = 0$ agree down to the constant term with the derivatives of the moments of the Riemann zeta function itself, which is what we would expect if the conjecture in (3.4.20) is correct.

Before we leave this topic, it is instructive to see how the form of the prime product proposed in (3.4.3) can be derived heuristicly. It appears straightforwardly if we use the Euler product for the Riemann zeta function and assume that the primes are uncorrelated. We should note that the prime product does not converge on the critical line and so we must be very careful over our truncation of the product.

If we warm up by tackling the second moment, then

$$\begin{aligned}
&\langle|\zeta(1/2+it)|^2\rangle_t \\
&= \left\langle \prod_p \left(1-\frac{1}{p^{1/2+it}}\right)^{-1}\left(1-\frac{1}{p^{1/2-it}}\right)^{-1}\right\rangle_t \\
&\approx \prod_p \left\langle \frac{1}{\left(1-\frac{1}{p^{1/2}e^{it\log p}}\right)\left(1-\frac{1}{p^{1/2}e^{-it\log p}}\right)}\right\rangle_t .
\end{aligned} \tag{3.4.32}$$

The extent of the average over $t$ can be defined in various ways, but to average over one cycle of the oscillatory component, $\exp(it\log p)$, would imply integration from $t = 0$ to $t = 2\pi/\log p$.

$$
\begin{aligned}
\langle|\zeta(1/2+it)|^2\rangle_t &\approx \prod_p \frac{\log p}{2\pi}\int_0^{2\pi/\log p} \frac{d\theta}{\left(1-\frac{1}{p^{1/2}e^{i\theta\log p}}\right)\left(1-\frac{1}{p^{1/2}e^{-i\theta\log p}}\right)} \\
&= \prod_p \frac{\log p}{2\pi}\oint_{|z|=1} \frac{dz}{zi\log p\left(1-\frac{1}{p^{1/2}z}\right)\left(1-\frac{z}{p^{1/2}}\right)} \\
&= \prod_p \frac{1}{1-1/p}. \qquad (3.4.33)
\end{aligned}
$$

As this product does not converge, we look for a sensible place to truncate it. Since the mean spacing of the zeros of the Riemann zeta function is $2\pi/\log t$, and the terms in (3.4.32) oscillate like $\exp(it\log p)$, we truncate the above product at $p = T$, where $T$ is the height up the critical line at which we are calculating the moments. This truncation effects a loss of information about the correlations between the Riemann zeros on a scale smaller than the mean level spacing. Information about long-range correlations is retained, and as this encodes material specific to the Riemann zeta function, this is exactly what we want. However, given that the high primes, those related to the short-range, universal distribution statistics of the zeros, have been removed, it is not surprising that the method being followed here needs some help from random matrix theory in order to predict correct-to-leading-order moments of $|\zeta(1/2+it)|$.

If we continue, and try to express the result to leading order in $T$, we obtain

$$
\begin{aligned}
\langle|\zeta(1/2+it)|^2\rangle_t &\approx \prod_{p<T} \frac{1}{1-1/p} \\
&= \exp\left(-\sum_{p<T}\log(1-1/p)\right) \\
&\approx \exp\left(\sum_{p<T}\frac{1}{p}\right) \\
&\approx \exp\left(\int_{T_o}^{T}\frac{dx}{x\log x}\right) \\
&= \exp(\log\log T - \log\log T_o), \qquad (3.4.34)
\end{aligned}
$$

where the transition from a sum over primes to an integral makes use of the knowledge that the density of prime numbers tends to $1/\log p$ for large primes $p$. This is a consequence of the prime number theorem (see [Tit86]).

There is some ambiguity in what the lower limit $T_o$ should be. 1 or 2 might be reasonable choices, but instead we will *define* the lower limit so that

$$\prod_{p<T} \frac{1}{1-1/p} = \log T \tag{3.4.35}$$

because we know from (3.3.2) that the leading-order term of the second moment is precisely $\log T$. Having made this choice, we will continue to use it consistently and treat (3.4.35) as a definition.

We now apply the same reasoning to the case of the general $2k^{th}$ moment. Below we will use the expansion $(1-x)^{-k} = \sum_{n=0}^{\infty} \binom{-k}{n} (-1)^n x^n$.

$$\begin{aligned}
\langle |\zeta(1/2+it)|^{2k} \rangle_t &= \prod_p \left\langle \frac{1}{\left(1-\frac{1}{p^{1/2+it}}\right)^k \left(1-\frac{1}{p^{1/2-it}}\right)^k} \right\rangle_t \\
&\approx \prod_p \left\langle \sum_{n=0}^{\infty} \frac{(-1)^n \binom{-k}{n}}{p^{n/2+int}} \sum_{m=0}^{\infty} \frac{(-1)^m \binom{-k}{m}}{p^{m/2-imt}} \right\rangle_t \\
&= \prod_p \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{n+m} \binom{-k}{n} \binom{-k}{m}}{p^{(n+m)/2}} \left\langle \frac{1}{p^{it(n-m)}} \right\rangle_t \\
&\approx \prod_p \sum_{n=0}^{\infty} \frac{\binom{-k}{n}^2}{p^n} \\
&= \prod_p \left(1 + \frac{k^2}{p} + \frac{k^2(k+1)^2}{2p^2} + \cdots \right).
\end{aligned} \tag{3.4.36}$$

The sum of $1/p$ terms diverges, so to cancel this we wish to multiply by

$$\prod_p (1-1/p)^{k^2} = \prod_p \left(1 - \frac{k^2}{p} + \frac{k^2(k^2-1)}{2p^2} + \cdots \right). \tag{3.4.37}$$

Now we have

$$\langle|\zeta(1/2+it)|^{2k}\rangle_t = \frac{\prod_p (1-1/p)^{k^2} \sum_{n=0}^{\infty} \frac{\binom{-k}{n}^2}{p^n}}{\prod_p (1-1/p)^{k^2}}, \tag{3.4.38}$$

where the product in the numerator converges, and so can be considered as running up to infinity.

Using (3.4.35) to cope with the non-convergent $\prod_p (1-1/p)^{-1}$, we proceed to

$$\begin{aligned} &\langle|\zeta(1/2+it)|^{2k}\rangle \\ &= \left(\prod_p (1-1/p)^{k^2} \sum_{n=0}^{\infty} \left(\frac{\Gamma(k+n)}{\Gamma(k)n!}\right)^2 \frac{1}{p^n}\right) (\log T)^{k^2}, \end{aligned} \tag{3.4.39}$$

because

$$\begin{aligned} \binom{-k}{n}^2 &= \left(\frac{(-k)(-k-1)(-k-2)\cdots(-k-n+1)}{n!}\right)^2 \\ &= \left(\frac{(k+n-1)\cdots(k+2)(k+1)k}{n!}\right)^2 \\ &= \left(\frac{\Gamma(k+n)}{\Gamma(k)n!}\right)^2. \end{aligned} \tag{3.4.40}$$

Thus we can gain, even from this heuristic derivation, a little insight into the origin of the the product over primes $a_k$ figuring in the conjecture on the moments of $|\zeta(1/2+it)|$, as it is precisely the product in (3.4.39).

## 3.5 A second mean value conjecture

To maintain the balance between mean values of products and quotients of the zeta function adhered to in the previous chapter, let us consider the moments of $(\zeta(1/2+it)/\zeta(1/2-it))^{1/2}$.

To start with the second moment,

$$\begin{aligned}
\left\langle \left( \frac{\zeta(1/2+it)}{\zeta(1/2-it)} \right) \right\rangle_t &= \left\langle \prod_p \frac{\left(1-\frac{1}{p^{1/2-it}}\right)}{\left(1-\frac{1}{p^{1/2+it}}\right)} \right\rangle_t \\
&\approx \prod_p \left\langle \left(1-\frac{1}{p^{1/2-it}}\right) \sum_{n=0}^{\infty} \left(\frac{1}{p^{1/2+it}}\right)^n \right\rangle_t \\
&= \prod_p \left\langle \sum_{n=0}^{\infty} \frac{e^{-itn\log p}}{p^{n/2}} - \sum_{n=0}^{\infty} \frac{e^{it\log p} e^{-itn\log p}}{p^{(n+1)/2}} \right\rangle_t \\
&\approx \prod_p (1-1/p) \;\sim\; \frac{1}{\log T}. \qquad (3.5.1)
\end{aligned}$$

This last step follows from the definition (3.4.35) we made in the previous section.

Now we continue on to the $2k^{th}$ moment ($k$ need not be an integer):

$$\begin{aligned}
&\left\langle \left( \frac{\zeta(1/2+it)}{\zeta(1/2-it)} \right)^k \right\rangle_t \\
&\approx \prod_p \left\langle \frac{\left(1-\frac{1}{p^{1/2-it}}\right)^k}{\left(1-\frac{1}{p^{1/2+it}}\right)^k} \right\rangle_t \\
&= \prod_p \left\langle \left( \sum_{n=0}^{\infty} \binom{k}{n} \frac{(-1)^n}{p^{n/2}} e^{itn\log p} \right) \left( \sum_{m=0}^{\infty} \binom{-k}{m} \frac{(-1)^m}{p^{m/2}} e^{-itm\log p} \right) \right\rangle_t \\
&\approx \prod_p \sum_{n=0}^{\infty} \binom{k}{n} \binom{-k}{n} \frac{1}{p^n} \\
&= \prod_p \left( 1 - \frac{k^2}{p} + \frac{(k+1)k^2(k-1)}{2!2!p^2} + \cdots \right). \qquad (3.5.2)
\end{aligned}$$

As in the previous section, we cancel out the $1/p$ term, but this time by multiplying by $(1-1/p)^{-k^2}$:

$$\begin{aligned}
\left\langle \left( \frac{\zeta(1/2+it)}{\zeta(1/2-it)} \right)^k \right\rangle_t &\approx \prod_p \left( \frac{\sum_{n=0}^{\infty} \binom{k}{n}\binom{-k}{n}\frac{1}{p^n}}{(1-1/p)^{k^2}} \right) \cdot \prod_p (1-1/p)^{k^2} \\
&\sim \prod_p \left( \frac{\sum_{n=0}^{\infty} \binom{k}{n}\binom{-k}{n}\frac{1}{p^n}}{(1-1/p)^{k^2}} \right) \frac{1}{(\log T)^{k^2}}. \qquad (3.5.3)
\end{aligned}$$

This result suggests that the $2k^{th}$ moment of $(\zeta(1/2+it)/\zeta(1/2-it))^{1/2}$ behaves asymptotically like $O((\log T)^{-k^2})$. We expect from this, using (3.2.1), that the leading-order term of the moment $L_N(2k) = \langle (Z(U,\theta)/Z^*(U,\theta))^k\rangle_{CUE}$ should be $O(N^{-k^2})$.

This is exactly right, as we recall from Chapter 2 that as $N$ becomes large

$$\lim_{N\to\infty} L_N(s) \times N^{s^2/4} = G(1-s/2)G(1+s/2). \tag{3.5.4}$$

This agreement in the asymptotic behaviour of the random matrix case versus the Riemann zeta mean values (noting that the $k$ exponent used above when working with the Riemann zeta function is equal to $s/2$) prompts us to make the following conjecture:

$$\lim_{T\to\infty} (\log T)^{s^2/4} \times \left\langle \left(\frac{\zeta(1/2+it)}{\zeta(1/2-it)}\right)^{s/2} \right\rangle_t = \lim_{N\to\infty} N^{s^2/4} b_{s/2} L_N(s), \tag{3.5.5}$$

where the $t$ average is around the value $t \approx T$ and the notation is

$$\begin{aligned} b_{s/2} &= \prod_p \Bigg( (1-1/p)^{-(s/2)^2} \\ &\qquad \sum_{n=0}^{\infty} \frac{\Gamma(1+s/2)\Gamma(1-s/2)}{\Gamma(1+s/2-n)\Gamma(1-s/2-n)n!n!} \frac{1}{p^n} \Bigg). \end{aligned} \tag{3.5.6}$$

As there are no analytical or numerical results for the Riemann zeta function with which to test this, the only evidence at the moment for this conjecture is the success of the similar proposal for $|\zeta|$ and the work on the log moments in the following section.

## 3.6 Modified ζ-specific generating functions

We have one last method up our sleeve for checking the validity of our conjectures. In (3.4.8) we see that the function $M_N(2\lambda)$ is modified by multiplication by the prime product $a_\lambda$ in order to predict the asymptotics of the mean values of the zeta function. $M_N(s)$ is also, from Section 2.1, the generating function of the moments of the real part of the log of $Z$. We now propose the use of $M_N(s)a_{s/2}$ as a generating function, to discover if this will produce the correct non-universal prime sums in the log moments; for example, that in (3.2.2). As we did when just $M_N(s)$ was the generating function, we can take successive derivatives of $M_N(s)a_{s/2}$ evaluated at $s=0$ and so obtain the moments of the real part of the log of $Z$, plus extra non-random matrix contributions which we hope will predict the non-universal terms of the Riemann log moments. We will call these moments with primes included $\langle(\mathrm{Re}\log Z)^k\rangle^P_{CUE}$ to distinguish them from the purely RMT ones. As the calculations are messy and it is truly just a matter of differentiating, we will just set out the results here. They are all asymptotic for large $N$.

$$\begin{aligned}
\langle\log|Z|\rangle^P_{CUE} &\sim 0 \qquad (3.6.1)\\
\langle(\log|Z|)^2\rangle^P_{CUE} &\sim \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \left\{\frac{1}{2}\sum_{m=2}^{\infty}\sum_{p}\left(-\frac{1}{m}+\frac{1}{m^2}\right)\frac{1}{p^m}\right\}\\
\langle(\log|Z|)^3\rangle^P_{CUE} &\sim -\frac{3}{2}\zeta(2) + \left\{\frac{3}{2}\sum_{m=2}^{\infty}\sum_{p}(\psi(m)+\gamma)\frac{1}{p^m m^2}\right\}\\
\langle(\log|Z|)^4\rangle^P_{CUE} &\sim 3\Bigg(\frac{1}{2}\log N + \frac{1}{2}(\gamma+1) +\\
&\qquad \left\{\frac{1}{2}\sum_{m=2}^{\infty}\sum_{p}\left(-\frac{1}{m}+\frac{1}{m^2}\right)\frac{1}{p^m}\right\}\Bigg)^2 + \frac{21}{4}\zeta(3)\\
&\quad + \Bigg\{3\sum_{m=1}^{\infty}\sum_{p}\Bigg(-\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{p^n n^2} + \psi(m)^2 + 2\gamma\psi(m)\\
&\quad +\gamma^2 + \frac{1}{2}\psi^{(1)}(m) - \frac{1}{2}\psi^{(1)}(1)\Bigg)\frac{1}{p^m m^2}\Bigg\}
\end{aligned}$$

In the above moments, the quantities in curly brackets are the extra contributions obtained because we are multiplying $M_N(s)$ by the product over primes. The purely RMT generating function yields precisely everything but these bracketed terms.

Once again, there is little analytical work from the number theorists with which to compare these, but in non-rigorous calculations of our own the moments $\langle(\log|\zeta(1/2+it)|)^k\rangle_t$ contain precisely the prime sums listed above. These Riemann zeta function calculations use techniques which have proved reliable in other important work ([BK95], [BK96]) but again involve the use of the Euler product on the critical line where it doesn't converge, so care must be taken to provide an appropriate truncation of the product. The prime sums in the moments above come directly from the most significant terms in sums resulting from taking a logarithm of the Euler product, while the random matrix terms (those outside the curly brackets) appear after applying to multiple sums over primes the Hardy-Littlewood conjecture on the correlations between primes, in the manner of the two references by Bogomolny and Keating cited above.

These techniques have been used to check the prime sums in all four moments in (3.6.1), while the Hardy-Littlewood method vindicated our confidence in the random matrix (unbracketed) portion of the first, second and third moment. The fourth moment would proceed in the same manner, but the technique is more time-consuming than the result merits, as at this stage it is the prime contributions which we wish to check. The details of these calculations appear in Appendix C.

Table 3.6 displays numerically how much the log moments are improved when the prime product is included in the generating function. Since prime sum contributions have not been calculated for the fifth cumulant and higher, the moments higher than the fourth do not contain the full allowance of prime contributions granted by the conjecture in (3.4.8). (This conjecture would imply that the prime-modified cumulants, $K$, agree exactly with the cumulants of $\log|\zeta(1/2+it)|$ down to and including constant terms). However, the most significant non-universal contributions result from the lowest cumulants, as can be seen if we write the moments in terms of the cumulants alone instead of relating them to lower moments in the manner of (3.4.27). Borrowing notation from (3.4.28),

$$\begin{aligned}\langle(\log|Z|)^k\rangle^P_{CUE} &= \sum_{\sum_{j=1}^k j\,n_j=k}\left(\frac{K_1}{1!}\right)^{n_1}\left(\frac{K_2}{2!}\right)^{n_2}\cdots\left(\frac{K_k}{k!}\right)^{n_k}\frac{k!}{n_1!n_2!\cdots n_k!}\\ &= \sum_{\sum_{j=2}^k j\,n_j=k}\left(\frac{K_2}{2!}\right)^{n_2}\cdots\left(\frac{K_k}{k!}\right)^{n_k}\frac{k!}{n_2!\cdots n_k!}.\end{aligned} \tag{3.6.2}$$

The first line is a general relation between moments and cumulants, and the second holds because $K_1 = 0$ in the case of the real log moments. Remembering that the second cumulant behaves asymptotically as $(1/2)\log N$ and the higher cumulants are $O(1)$, we see that the prime contributions to low cumulants are more significant asymptotically than those connected with the higher cumulants. Even so, the lack of the high cumulant prime sums is the reason that the upper moments in Table 3.6 are not as close to the Riemann moments as the lower ones. Still, the inclusion of the primes certainly effects a definite improvement over the purely CUE moments. This table, apart from column 4, is the same as Table 3.3, only the moments haven't been scaled to be comparable to a Gaussian with mean zero and unit variance. This is merely in order that we have as many low moments as possible for comparison with the Riemann zeta function, as moments one to four contain the complete prime sum contribution due to the generating function $M_N(s)a_{s/2}$.

Since the modification of the generating function proved such a success for the real part of the log, it is natural to attempt to reproduce it for the imaginary part as well.

We have not discovered any number theoretical work on moments of this type, so our best test of the conjecture is to calculate the moments of the imaginary part of the log using the product over primes multiplied by the product of gamma functions $L_N(s)$, as in (3.5.5), as the generating function. These moments will be denoted $\langle(\operatorname{Im}\log Z)^k\rangle^P_{CUE}$ with the $P$ to remind us that they include the prime contributions.

The resulting moments are

| Moment | ζ a) | ζ b) | CUE + primes | CUE |
|---|---|---|---|---|
| 1 | -0.001595 | 0.000549 | 0.0 | 0.0 |
| 2 | 2.5736 | 2.51778 | 2.56939 | 2.65747 |
| 3 | -2.2263 | -2.19591 | -2.21609 | -2.44955 |
| 4 | 25.998 | 25.1283 | 26.017 | 27.4967 |
| 5 | -81.2144 | -79.2332 | -81.2922 | -89.4481 |
| 6 | 655.921 | 628.48 | 663.493 | 713.597 |
| 7 | -3966.46 | -3765.29 | -4052.98 | -4437.47 |
| 8 | 33328.6 | 30385.5 | 34808.2 | 37806 |
| 9 | -282163 | -250744 | -304267 | -332278 |
| 10 | $2.271\times10^6$ | $2.298\times10^6$ | $3.082\times10^6$ | $3.359\times10^6$ |

Table 3.6: Moments of $\mathrm{Re}\log\zeta$ near the $10^{20}$th zero ($T \approx 1.520\times10^{19}$) (averages in a) and b) taken over different intervals) compared with $\mathrm{Re}\log Z$ with and without the prime contributions, with $N = 42$.

$$\begin{aligned}
\langle \mathrm{Im}\log Z\rangle_{CUE}^{P} &\sim 0 \qquad (3.6.3)\\
\langle (\mathrm{Im}\log Z)^2\rangle_{CUE}^{P} &\sim \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \left\{\frac{1}{2}\sum_{m=2}^{\infty}\sum_{p}\left(-\frac{1}{m}+\frac{1}{m^2}\right)\frac{1}{p^m}\right\}\\
\langle (\mathrm{Im}\log Z)^3\rangle_{CUE}^{P} &\sim 0\\
\langle (\mathrm{Im}\log Z)^4\rangle_{CUE}^{P} &\sim 3\left(\frac{1}{2}\log N + \frac{1}{2}(\gamma+1)\right.\\
&\quad \left.+\left\{\frac{1}{2}\sum_{m=2}^{\infty}\sum_{p}\left(-\frac{1}{m}+\frac{1}{m^2}\right)\frac{1}{p^m}\right\}\right)^2 - \frac{3}{4}\zeta(3)\\
&\quad +\left\{3\sum_{n=1}^{\infty}\sum_{p}\frac{1}{p^n n^2}\left(-\frac{1}{4}\sum_{m=1}^{\infty}\frac{1}{p^m m^2}\right.\right.\\
&\quad \left.\left.-\frac{1}{2}(\psi^{(1)}(n)-\psi^{(1)}(1))\right)\right\}.
\end{aligned}$$

The second moment is exactly the rigorously calculated second moment of the imaginary part of the log of zeta due to Goldston (3.2.2) and the prime sums in the other moments again agree exactly with our own heuristic developments of the moments for the Riemann zeta function itself (see Appendix C).

There is again numerical evidence, in the form of comparison with Odlyzko's computations with the zeta function. These are laid out in Table 3.7, where again the moments are unnormalized so that we have the maximum number available for comparison.

| Moment | $\zeta$ | CUE + primes | CUE |
|---|---|---|---|
| 1 | $-1.0 \times 10^{-5}$ | 0.0 | 0.0 |
| 2 | 2.573 | 2.569 | 2.657 |
| 3 | $-1.9 \times 10^{-3}$ | 0.0 | 0.0 |
| 4 | 18.74 | 18.69 | 20.28 |
| 5 | -0.097 | 0.0 | 0.0 |
| 6 | 216.5 | 215.6 | 249.5 |
| 7 | -3.8 | 0.0 | 0.0 |
| 8 | 3355 | 3321 | 4178 |

Table 3.7: Moments of $\operatorname{Im}\log\zeta$ near the $10^{20}$th zero ($T \approx 1.520 \times 10^{19}$) compared with $\operatorname{Im}\log Z$ for $N = 42$ with and without the conjectured prime contributions.

We see that for the imaginary log moments, as for the real ones, there is a marked improvement if prime contributions are included, which both supports conjecture (3.5.5) and adds evidence that when dealing with mean values of $\zeta$, as opposed to its logarithm, the universal and non-universal components appear to split neatly into separate factors, at least in the asymptotic limit.

# Chapter 4

# A chorus of L-functions

Whereas in the previous chapter we concentrated exclusively on the Riemann zeta function, here we will consider a wider range of functions: the L-functions. These are closely related to the zeta function already discussed. In particular, it is conjectured that their zeros lie on a line and that asymptotically high on this line the zeros display CUE statistics in the same manner as the zeta function [RS96, Rum93, Rub98]. We could, therefore, choose an individual L-function and carry out the identical comparisons with random matrix theory as were undergone in Chapter 3. However, much more interesting than this is the fact that L-functions come in families, averages over which reveal [KS99a, KS99b] that the low-lying zeros on the critical line appear to obey the statistics of one of the random matrix ensembles defined by Haar measure on the compact groups $U(N)$, $O(N)$ and $USp(2N)$. Haar measure is the natural choice as it provides a uniform weighting of the matrices in each group. Thus in this chapter we perform averages of the characteristic polynomials of matrices from the above ensembles and compare them with averages over families of $L$-functions. This work can also be found published in [KS00a].

## 4.1 Zero statistics of L-functions

$L$-functions occur in various forms, but they share many of the properties of the Riemann zeta function; they can be written as a Dirichlet series over the positive integers or as an Euler product over the primes, plus they have a functional equation. Examples are the Dirichlet $L$-functions, which take the form of a Dirichlet sum but with the replacement of the 1 in the numerator of (1.4.2) by a Dirichlet character (see [Dav80]), as well as more general functions (for example see [RS96]). Each obeys a functional equation similar to that of the Riemann zeta function, which relates the function at points in one half-plane to points in the remainder of the complex plane. The $L$-functions can be normalized so that the critical line dividing the plane in this way lies at $\mathrm{Re}\, s = 1/2$. Then for $L$-functions, as for the Riemann zeta function, it is conjectured, and widely believed, that the non-trivial zeros lie on the line $\mathrm{Re}\, s = 1/2$; this is the generalized Riemann hypothesis (GRH). Whereas high on this critical line the zeros of any given $L$-function appear to have the same statistical distribution as the eigenvalues of the CUE ensemble, or equivalently the group $U(N)$ of (large) $N \times N$ unitary matrices endowed with Haar measure [RS96, Rum93, Rub98], Katz and Sarnak [KS99a, KS99b] have proposed that the positions of the lowest zeros - those nearest to the point $s = 1/2$ - averaged over families of $L$-functions follow the statistical distribution of the eigenvalues not always of $U(N)$, but in some cases of the compact groups $O(N)$ or $USp(2N)$. This is supported by the fact that for the finite function field analogue the equivalent of the Riemann Hypothesis is known to be true, and Katz and Sarnak have discovered that zeta functions over function fields have zero statistics which show exactly the behaviour just described.

Conrey and Farmer [CF99] have extended this idea that the low-lying zeros of families of $L$-functions show particular statistics to the study of the mean values of the $L$-functions $L_f(s)$ within families at the central point $s = 1/2$. They have found evidence that the symmetry type to which the low-lying zeros subscribe also determines the behaviour of these mean values. In particular, they conjecture that in general, as $\mathcal{Q} \to \infty$,

$$\frac{1}{\mathcal{Q}^*} \sum_{\substack{f \in \mathcal{F} \\ c(f) \leq \mathcal{Q}}} V(L_f(\tfrac{1}{2}))^k \sim g_k \frac{a(k)}{\Gamma(1+B(k))} (\log \mathcal{Q}^A)^{B(k)}, \tag{4.1.1}$$

where they choose $V(z)$ depending on the symmetry type ($V(z) = |z|^2$ for unitary symmetry and $V(z) = z$ for the orthogonal or symplectic case); $A$ is a symmetry-dependent constant; the family, $\mathcal{F}$, over which the average is performed is considered to be partially ordered by the conductor, $c(f)$, of each $L$-function; and the sum is over the $\mathcal{Q}^*$ elements with $c(f) \leq \mathcal{Q}$. The symmetry type of the family manifests itself in the expectation that it alone determines the values of $g_k$ and $B(k)$. These functions are thus universal, being independent of the details of the particular family in question. $a(k)$, on the other hand, is expected to depend on the specific family involved.

As an example we can consider one of the simplest families of $L$-functions, those with real quadratic Dirichlet characters $\chi_d$ (defined below). Thus we have $L$-functions of the form

$$L(s, \chi_d) = \sum_{n=1}^{\infty} \frac{\chi_d(n)}{n^s} = \prod_p \left(1 - \frac{\chi_d(p)}{p^s}\right)^{-1}. \tag{4.1.2}$$

These $L$-functions cannot be constructed for all conductors $|d|$, because the Dirichlet characters are not real for every choice of $d$, but where real characters do exist, they can be written as (for $n > 0$)

$$\chi_d(n) = \left(\frac{d}{n}\right) \tag{4.1.3}$$

where the right hand side is Kronecker's extension of Legendre's symbol. This latter symbol is defined for prime $p$ as

$$\left(\frac{d}{p}\right) = \begin{cases} +1 & \text{if } p \nmid d \text{ and } x^2 \equiv d(\text{mod} p) \text{ is soluble} \\ 0 & \text{if } p|d \\ -1 & \text{if } p \nmid d \text{ and } x^2 \equiv d(\text{mod} p) \text{ is not soluble} \end{cases}, \tag{4.1.4}$$

and the Kronecker symbol extends this to non-prime $p$. For some conductors $|d|$ there exist two real characters, one corresponding to $|d|$ and one to $-|d|$, and for some conductors there is zero or one appropriate character. Explicitly, a real Dirichlet character exists where $d$ is a product of relatively prime factors chosen from -4, 8, -8, $(-1)^{\frac{1}{2}(p-1)}p$, with $p > 2$.

So, the mean value of Conrey and Farmer [CF99] sums over all such real characters with conductor $|d| \leq D$ and as $D \to \infty$ they find that

$$\frac{1}{D^*} \sum_{|d| \leq D} L\left(\frac{1}{2}, \chi_d\right)^k \sim g_k \frac{a(k)}{\Gamma(1 + \frac{1}{2}k(k+1))} (\log D^{\frac{1}{2}})^{\frac{1}{2}k(k+1)}, \tag{4.1.5}$$

where $D^*$ is the number of quadratic characters included in the sum.

The family of $L$-functions with real quadratic Dirichlet characters are believed to belong to the symplectic variety of $L$-functions in that the statistical behaviour of the zeros near to $s = 1/2$ is that of the eigenvalues of the unitary symplectic matrices of $USp(2N)$ which lie on the unit circle close to one. The mean value (4.1.5) has the characteristic form for symplectic families of $L$-functions as (using the notation of (4.1.1)) for this symmetry type $B(k) = \frac{1}{2}k(k+1)$ and $A = 1/2$.

For orthogonal symmetry, on the other hand, $B(k) = \frac{1}{2}k(k-1)$ although $A$ still equals $1/2$ [CF99]. The low zeros of these families show the statistics of the eigenvalues of $O(N)$, with Haar measure once again. In the random matrix calculations to follow we will consider only $SO(2N)$, orthogonal matrices with even dimension and determinant $+1$, because a family of $L$-functions governed by $O(N)$ falls approximately into two halves; one displaying even symmetry about $s = 1/2$, and the other odd symmetry. This latter class contributes zero to averages at the central value, while the zero statistics of the former are expected to follow those of $SO(2N)$.

## 4.2 Symplectic symmetry

### 4.2.1 Random matrices in $USp(2N)$

We are interested here in the group of symplectic unitary matrices, $USp(2N)$. These are $2N \times 2N$ matrices, $U$, with $UU^\dagger = 1$ and $U^tJU = J$, where $J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix}$ and $I_N$ is the $N \times N$ identity matrix. For these matrices, the eigenvalues lie on the unit circle and come in complex conjugate pairs. Thus the characteristic polynomial of such a matrix with eigenvalues $e^{i\theta_1}, e^{-i\theta_1}, e^{i\theta_2}, e^{-i\theta_2}, \ldots e^{i\theta_N}, e^{-i\theta_N}$ is

$$Z(U,\theta) = \prod_{n=1}^{N} \left(1 - e^{i(\theta_n - \theta)}\right)\left(1 - e^{i(-\theta_n-\theta)}\right). \tag{4.2.1}$$

As we are interested in the mean values of L-functions at the point $s = 1/2$ (the symmetry point for the zeros on the critical line), we look at $Z$ at the corresponding point, $\theta = 0$, this being the point around which the eigenvalues are symmetric. The first task is to calculate the moments of $Z(U,0)$ over $USp(2N)$. Taking Haar measure on the group, the joint probability density function of the eigenvalues is [Wey46]

$$N_{Sp} \prod_{1\le i<j\le N} \left(\sin\left(\frac{\theta_i - \theta_j}{2}\right)\sin\left(\frac{\theta_i+\theta_j}{2}\right)\right)^2 \prod_{k=1}^{N} \sin^2\theta_k, \tag{4.2.2}$$

so noting that

$$\cos(A-B) - \cos(A+B) \;=\; 2\sin A \sin B, \tag{4.2.3}$$

the joint probability density function of the $N$ independent eigenvalues is

$$N_{Sp} \prod_{1\le i<j\le N} \left(\frac{1}{2}(\cos\theta_j - \cos\theta_i)\right)^2 \prod_{k=1}^{N} \sin^2\theta_k. \tag{4.2.4}$$

$N_{Sp}$ is the normalization constant and can be determined by averaging (4.2.4) over the full range, 0 to $2\pi$, of each $\theta_j$ and setting the result equal to unity. Thus

$$N_{Sp} = 2^{-3N} \prod_{j=1}^{N} \frac{\Gamma(1+N+j)}{\Gamma(1+j)(\Gamma(1/2+j))^2} = \frac{2^{2N^2-2N}}{\pi^N N!}. \tag{4.2.5}$$

The product of gamma functions is the most convenient form for us, but after a little manipulation, it can be seen to be equal to the right side, which is the more common form of this constant.

Armed with this normalization constant, we tackle the moments of $Z$ itself. Firstly,

$$\begin{aligned} Z(U,0) &= \Big|\prod_{n=1}^{N}(1-e^{i\theta_n})\Big|^2 \\ &= 2^{2N}\prod_{n=1}^{N}\sin^2(\theta_n/2) \\ &= 2^N\prod_{n=1}^{N}(1-\cos\theta_n). \end{aligned} \tag{4.2.6}$$

Now we have that

$$\begin{aligned} &\langle Z(U,0)^s\rangle_{USp(2N)} \\ &= N_{Sp}2^{Ns}\int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \prod_{1\le i<j\le N}\left(\frac{1}{2}(\cos\theta_j-\cos\theta_i)\right)^2 \\ &\quad\times\prod_{k=1}^{N}\sin^2\theta_k\prod_{n=1}^{N}(1-\cos\theta_n)^s \\ &= N_{Sp}2^{Ns+2N-N^2}\int_0^{\pi}\cdots\int_0^{\pi} d\theta_1\cdots d\theta_N \prod_{1\le i<j\le N}(\cos\theta_j-\cos\theta_i)^2 \\ &\quad\times\prod_{k=1}^{N}\sin^2\theta_k\prod_{n=1}^{N}(1-\cos\theta_n)^s, \end{aligned}$$

which, after the transformation $x_j = \cos\theta_j$, becomes

$$\begin{aligned}
&\langle Z(U,0)^s\rangle_{USp(2N)} \\
&= N_{Sp} 2^{Ns+2N-N^2} \int_{-1}^{1}\cdots\int_{-1}^{1} dx_1\cdots dx_N \prod_{1\le i<j\le N}(x_j-x_i)^2 \\
&\quad\times \prod_{k=1}^{N}(1-x_k^2)^{1/2}\prod_{n=1}^{N}(1-x_n)^s \\
&= N_{Sp} 2^{Ns+2N-N^2} \int_{-1}^{1}\cdots\int_{-1}^{1} dx_1\cdots dx_N \prod_{1\le i<j\le N}(x_j-x_i)^2 \\
&\quad\times \prod_{k=1}^{N}(1-x_k)^{1/2+s}(1+x_k)^{1/2}.
\end{aligned} \tag{4.2.7}$$

There is a form of Selberg's integral (detailed in [Meh91]) which states

$$\begin{aligned}
&\int_{-1}^{1}\cdots\int_{-1}^{1}\prod_{1\le j<l\le n}|(x_j-x_l)|^{2\gamma}\prod_{j=1}^{n}(1-x_j)^{\alpha-1}(1+x_j)^{\beta-1}dx_j \\
&= 2^{\gamma n(n-1)+n(\alpha+\beta-1)}\prod_{j=0}^{n-1}\frac{\Gamma(1+\gamma+j\gamma)\Gamma(\alpha+j\gamma)\Gamma(\beta+j\gamma)}{\Gamma(1+\gamma)\Gamma(\alpha+\beta+\gamma(n+j-1))},
\end{aligned} \tag{4.2.8}$$

if $\mathrm{Re}\alpha > 0$, $\mathrm{Re}\beta > 0$ and $\mathrm{Re}\gamma > -\min\left(\frac{1}{n}, \frac{\mathrm{Re}\alpha}{n-1}, \frac{\mathrm{Re}\beta}{n-1}\right)$.

In our case $\gamma = 1$, $\alpha = 3/2 + s$ and $\beta = 3/2$, so

$$\begin{aligned}
\langle Z(U,0)^s\rangle_{USp(2N)} &= N_{Sp}2^{Ns+2N-N^2}2^{N^2-N+3N/2+Ns+3N/2-N} \\
&\quad\times\prod_{j=0}^{N-1}\frac{\Gamma(2+j)\Gamma(3/2+s+j)\Gamma(3/2+j)}{\Gamma(2)\Gamma(3+s+N+j-1)} \\
&= N_{Sp}2^{2Ns+3N}\prod_{j=1}^{N}\frac{\Gamma(1+j)\Gamma(1/2+s+j)\Gamma(1/2+j)}{\Gamma(1+s+N+j)} \\
&= 2^{2Ns}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+s+j)}{\Gamma(1/2+j)\Gamma(1+s+N+j)} \\
&\equiv M_{Sp}(N,s).
\end{aligned} \tag{4.2.9}$$

We now consider the cumulants $c_j$ in the expansion

$$M_{Sp}(N,s) = e^{c_1 s + c_2 s^2/2 + c_3 s^3/3! + c_4 s^4/4! + \cdots}. \tag{4.2.10}$$

As $\log M_{Sp}(N,s)$ is

$$\begin{aligned}\log M_{Sp}(N,s) &= 2sN\log 2 + \sum_{j=1}^{N}(\log\Gamma(1/2+s+j) + \log\Gamma(1+N+j) \\ &\quad - \log\Gamma(1/2+j) - \log\Gamma(1+s+N+j)),\end{aligned} \tag{4.2.11}$$

we find that

$$\begin{aligned}c_1 &= \frac{d}{ds}\log M_{Sp}(N,s)\Big|_{s=0} \\ &= 2N\log 2 + \sum_{j=1}^{N}\left(\frac{d}{ds}\log\Gamma(1/2+s+j)\right. \\ &\qquad \left. -\frac{d}{ds}\log\Gamma(1+s+N+j)\right)\Big|_{s=0} \\ &= 2N\log 2 + \sum_{j=1}^{N}(\psi(1/2+j) - \psi(1+N+j))\end{aligned} \tag{4.2.12}$$

is the first cumulant, while the higher ones are given by

$$\begin{aligned}c_n &= \frac{d^n}{ds^n}\log M_{Sp}(N,s)\Big|_{s=0} \\ &= \sum_{j=1}^{N}\left(\psi^{(n-1)}(1/2+j) - \psi^{(n-1)}(1+N+j)\right),\end{aligned} \tag{4.2.13}$$

where $\psi^{(j)}(z) = \frac{d^{j+1}}{dz^{j+1}}\log\Gamma(z)$ is a polygamma function.

We seek the behaviour of these cumulants for large $N$. For the first cumulant we use the asymptotic formula

$$\psi(z) \sim \log z - \frac{1}{2z} - \sum_{n=1}^{\infty}\frac{B_{2n}}{2nz^{2n}}, \tag{4.2.14}$$

which holds when $z \to \infty$ with $|\arg z| < \pi$, and the $B_{2n}$ are the Bernoulli numbers. Also, we need the integral form of the digamma function, previously given in (2.2.13),

$$\psi(z) + \gamma = \int_0^\infty \frac{e^{-t} - e^{-zt}}{1 - e^{-t}} dt. \tag{4.2.15}$$

Applying (4.2.15), we obtain

$$\begin{aligned} c_1 &= \frac{d}{ds} \log M_{Sp}(N,s)\Big|_{s=0} \\ &= 2N \log 2 + \sum_{j=1}^{N} \left(\psi(j+1/2) - \psi(j+N+1)\right) \qquad (4.2.16) \\ &= 2N \log 2 + \sum_{j=1}^{N} \left( \int_0^\infty \frac{e^{-t} - e^{-(j+1/2)t}}{1-e^{-t}} dt - \gamma \right. \\ &\qquad \left. - \int_0^\infty \frac{e^{-t} - e^{-(j+N+1)t}}{1-e^{-t}} dt + \gamma \right) \\ &= 2N \log 2 + \sum_{j=1}^{N} \int_0^\infty \frac{e^{-(j+N+1)t} - e^{-(j+1/2)t}}{1-e^{-t}} dt. \end{aligned}$$

Now we interchange the order of the summation and integration so that we can perform the sum explicitly and then integrate by parts.

$$\begin{aligned} c_1 &= 2N \log 2 + \int_0^\infty \frac{e^{-(N+2)t}(1-e^{-Nt}) - e^{-3t/2}(1-e^{-Nt})}{(1-e^{-t})^2} dt \\ &= 2N \log 2 + \left[ \frac{\left(-e^{-(N+1)t} + e^{-(2N+1)t} + e^{-t/2} - e^{-(N+1/2)t}\right) e^{-t}}{1-e^{-t}} \Bigg|_0^\infty \right. \\ &\qquad - \int_0^\infty \frac{\left((N+1)e^{-(N+1)t} - (2N+1)e^{-(2N+1)t}\right) e^{-t}}{1-e^{-t}} dt \\ &\qquad \left. - \int_0^\infty \frac{\left(-\frac{1}{2}e^{-t/2} + (N+1/2)e^{-(N+1/2)t}\right) e^{-t}}{1-e^{-t}} dt \right] \\ &= 2N \log 2 - (N+1) \int_0^\infty \frac{e^{-(N+2)t}}{1-e^{-t}} dt + (2N+1) \int_0^\infty \frac{e^{-(2N+2)t}}{1-e^{-t}} dt \\ &\qquad + \frac{1}{2} \int_0^\infty \frac{e^{-3t/2}}{1-e^{-t}} dt - (N+1/2) \int_0^\infty \frac{e^{-(N+3/2)t}}{1-e^{-t}} dt. \end{aligned} \tag{4.2.17}$$

Using (4.2.15) in reverse,

$$\begin{aligned}
c_1 &= 2N\log 2 + (N+1)\psi(N+2) - (2N+1)\psi(2N+2) - \frac{1}{2}\psi(3/2) \\
&\quad +(N+1/2)\psi(N+3/2) \\
&= 2N\log 2 + (N+1)\left(\log(N+2) - \frac{1}{2N} + O(N^{-2})\right) \\
&\quad -(2N+1)\left(\log(2N+2) + \frac{1}{4N} + O(N^{-2})\right) \\
&\quad -\frac{1}{2}(-\gamma - 2\log 2 + 2) + (N+1/2)\left(\log N + \log\left(1+\frac{3}{2N}\right)\right. \\
&\quad \left. + \frac{1}{2N+3} + O(N^{-2})\right) \\
&= \frac{1}{2}\log N + \frac{\gamma}{2} + O(N^{-1}). \qquad (4.2.18)
\end{aligned}$$

For the second cumulant we need to use the asymptotic formula for higher polygamma functions, valid as $z \to \infty$ with $|\arg z| < \pi$,

$$\psi^{(n)}(z) \sim (-1)^{n-1}\left[\frac{(n-1)!}{z^n} + \frac{n!}{2z^{n+1}} + \sum_{k=1}^{\infty} B_{2k}\frac{(2k+n-1)!}{(2k)!z^{2k+n}}\right]. \qquad (4.2.19)$$

There is also an integral formula for the higher polygamma functions which we met in (2.2.14) and which will prove useful:

$$\psi^{(n)}(z) = (-1)^{n-1}\int_0^\infty \frac{t^n e^{-zt}}{1-e^{-t}}dt. \qquad (4.2.20)$$

This leads us to

$$\begin{aligned}
c_2 &= \sum_{j=1}^{N}\left(\psi^{(1)}(j+1/2) - \psi^{(1)}(1+N+j)\right) \\
&= \sum_{j=1}^{N}\left(\int_0^\infty \frac{te^{-(j+1/2)t}}{1-e^{-t}}dt - \int_0^\infty \frac{te^{-(1+N+j)t}}{1-e^{-t}}dt\right). \qquad (4.2.21)
\end{aligned}$$

Again we interchange the order of the summation and integration, perform the sum and integrate by parts. The result, expressed in terms of polygamma functions, is

$$\begin{aligned} c_2 &= -\psi(3/2) - \frac{1}{2}\psi^{(1)}(3/2) + \psi(N+3/2) + (N+1/2)\psi^{(1)}(N+3/2) \\ &\quad +\psi(N+2) + (N+1)\psi^{(1)}(N+2) - \psi(2N+2) \\ &\quad -(2N+1)\psi^{(1)}(2N+2) \\ &= \log N + 1 + \gamma + \log 2 - \frac{3}{2}\zeta(2) + O(N^{-1}). \end{aligned} \tag{4.2.22}$$

The higher cumulants follow in a similar manner

$$\begin{aligned} c_n &= \sum_{j=1}^{N}\left((-1)^n \int_0^\infty \frac{t^{n-1}e^{-(1/2+j)t}}{1-e^{-t}}dt - (-1)^n \int_0^\infty \frac{t^{n-1}e^{-(1+N+j)t}}{1-e^{-t}}dt\right) \\ &= (-1)^n \int_0^\infty \frac{t^{n-1}e^{-(3/2)t}}{1-e^{-t}}\,\frac{1-e^{-Nt}}{1-e^{-t}}dt \\ &\quad -(-1)^n \int_0^\infty \frac{t^{n-1}e^{-2-N}}{1-e^{-t}}\,\frac{1-e^{-Nt}}{1-e^{-t}}dt, \end{aligned} \tag{4.2.23}$$

and so

$$\begin{aligned} \lim_{N\to\infty} c_n &= (-1)^n \int_0^\infty \frac{t^{n-1}e^{-(3/2)t}}{(1-e^{-t})(1-e^{-t})}dt \\ &= (-1)^n \left[\frac{-t^{n-1}e^{-(3/2)t}}{1-e^{-t}}\Big|_0^\infty + (n-1)\int_0^\infty \frac{t^{n-2}e^{-(1/2)t}}{1-e^{-t}}dt \right. \\ &\quad \left. -\frac{1}{2}\int_0^\infty \frac{t^{n-1}e^{-(1/2)t}}{1-e^{-t}}dt\right] \\ &= -(n-1)\psi^{(n-2)}(1/2) - \frac{1}{2}\psi^{(n-1)}(1/2) \qquad (4.2.24) \\ &= (-1)^n(n-1)!\left[(2^{n-1}-1)\zeta(n-1) - \frac{1}{2}(2^n-1)\zeta(n)\right]. \end{aligned}$$

We now turn our attention back to the moments of $Z$ itself. We would like to know the coefficient of the leading-order term for integer moments, as it is the corresponding quantity for the $L$-functions that Conrey and Farmer have conjectured the form of in (4.1.1).

Starting with the first moment,

$$
\begin{aligned}
\langle Z(U,0)\rangle_{USp(2N)} &= 2^{2N}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+1+j)}{\Gamma(1/2+j)\Gamma(2+N+j)} \\
&= 2^{2N}\frac{\Gamma(N+2)\Gamma(N+3)\cdots\Gamma(2N+1)}{\Gamma(1/2+1)\Gamma(1/2+2)\cdots\Gamma(1/2+N)} \\
&\quad\times\frac{\Gamma(1/2+2)\Gamma(1/2+3)\cdots\Gamma(1/2+1+N)}{\Gamma(3+N)\Gamma(4+N)\cdots\Gamma(2+2N)} \\
&= 2^{2N}\frac{\Gamma(N+2)\Gamma(1/2+1+N)}{\Gamma(2+2N)\Gamma(1/2+1)} \\
&= 2^{2N}\frac{(N+1)!(N+1/2)!}{(1/2)!(2N+1)!} \\
&= 2^{2N}\frac{(N+1/2)(N-1/2)(N-3/2)\cdots\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}}{(2N+1)(2N)(2N-1)\cdots(N+3)(N+2)} \\
&= 2^{N}\frac{(2N+1)(2N-1)(2N-3)\cdots7\cdot5\cdot3}{(2N+1)(2N)(2N-1)\cdots(N+3)(N+2)} \\
&= \begin{cases} 2^{N}\frac{(N+1)(N-1)\cdots7\cdot5\cdot3}{2N(2N-2)\cdots(N+4)(N+2)} & N \text{ even} \\ 2^{N}\frac{N(N-2)(N-4)\cdots7\cdot5\cdot3}{2N(2N-2)\cdots(N+3)} & N \text{ odd}\end{cases} \\
&= \begin{cases} 2^{N/2}\frac{(N+1)(N-1)\cdots7\cdot5\cdot3}{N(N-1)\cdots(N/2+2)(N/2+1)} & N \text{ even} \\ 2^{(N+1)/2}\frac{N(N-2)(N-4)\cdots7\cdot5\cdot3}{N(N-1)\cdots(N/2+5/2)(N/2+3/2)} & N \text{ odd}\end{cases} \\
&= \begin{cases} 2^{N/2}\frac{(N+1)!(N/2)!}{(N/2)!2^{N/2}N!} & N \text{ even} \\ 2^{(N+1)/2}\frac{N!(N/2+1/2)!}{(N/2-1/2)!2^{N/2-1/2}N!} & N \text{ odd}\end{cases} \\
&= N+1. \qquad (4.2.25)
\end{aligned}
$$

The second moment can be related back to the first in the following way:

$$
\begin{aligned}
\langle Z(U,0)^2\rangle_{USp(2N)} &= 2^{4N}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+2+j)}{\Gamma(1/2+j)\Gamma(3+N+j)} \\
&= 2^{4N}\frac{\Gamma(N+2)\Gamma(N+3)\Gamma(1/2+N+1)\Gamma(1/2+N+2)}{\Gamma(2N+2)\Gamma(2N+3)\Gamma(1/2+1)\Gamma(1/2+2)} \\
&= 2^{4N}\frac{(N+1)!(N+2)!(N+1/2)!(N+3/2)!}{(2N+1)!(2N+2)!(1/2)!(3/2)!} \\
&= 2^{4N}\frac{(N+2)(N+3/2)}{(3/2)(2N+2)}\left(\frac{(N+1)!(N+1/2)!}{(1/2)!(2N+1)!}\right)^2 \\
&= (\langle Z(U,0)\rangle_{USp(2N)})^2\ \frac{N^2+\frac{7}{2}N+3}{3N+3}
\end{aligned}
$$

$$
\begin{aligned}
\langle Z(U,0)^2\rangle_{USp(2N)} &= (N+1)^2\frac{N^2+\frac{7}{2}N+3}{3N+3} \\
&= \frac{1}{3}(N+1)(N^2+\frac{7}{2}N+3) \\
&\sim \frac{1}{3}N^3. \qquad (4.2.26)
\end{aligned}
$$

The third moment follows in a similar manner:

$$
\begin{aligned}
\langle Z(U,0)^3\rangle_{USp(2N)} &= 2^{6N}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+3+j)}{\Gamma(1/2+j)\Gamma(4+N+j)} \\
&= 2^{6N}\frac{\Gamma(2+N)\Gamma(3+N)\Gamma(4+N)}{\Gamma(3/2)\Gamma(5/2)\Gamma(7/2)} \\
&\quad\times\frac{\Gamma(1/2+N+1)\Gamma(1/2+N+2)\Gamma(1/2+N+3)}{\Gamma(2N+2)\Gamma(2N+3)\Gamma(2N+4)} \\
&= 2^{6N}\left(\frac{(N+1)!(N+1/2)!}{(1/2)!(2N+1)!}\right)^3\frac{(N+2)(N+2)(N+3)}{(3/2)(3/2)(5/2)} \\
&\quad\times\frac{(N+3/2)(N+3/2)(N+5/2)}{(2N+2)(2N+2)(2N+3)} \\
&= (N+1)^3\frac{(N+2)(N+2)(N+3)(N+3/2)(N+5/2)}{3\cdot 3\cdot 5(N+1)(N+1)} \\
&\sim \frac{1}{45}N^6. \qquad (4.2.27)
\end{aligned}
$$

The higher moments follow in exactly the same way, and the general leading-order term appears to be

$$
\langle Z(U,0)^n\rangle_{USp(2N)} \sim \left(\prod_{j=1}^{n}(2j-1)!!\right)^{-1} N^{\frac{1}{2}n(n+1)}, \qquad (4.2.28)
$$

and this can be proven by induction.

This is a useful result for the integer moments, but the asymptotic expressions for the cumulants, (4.2.18), (4.2.22) and (4.2.24), can be inserted into (4.2.10) to allow us to write the leading-order coefficient of the moment $M_{Sp}(N,s)$ as

$$
\begin{aligned}
f_{Sp}(s) &\equiv \lim_{N\to\infty} \frac{M_{Sp}(N,s)}{N^{s/2+s^2/2}} \\
&= \lim_{N\to\infty} \exp\left(\left(\frac{1}{2}\log N + \frac{1}{2}\gamma + O(N^{-1})\right) s + (\log N + 1 + \gamma\right. \\
&\quad + \log 2 - \frac{3}{2}\zeta(2) + O(N^{-1})\Big) \frac{s^2}{2} \\
&\quad + \sum_{n=3}^{\infty} \Big((-1)^n (n-1)!(2^{n-1}-1)\zeta(n-1) \\
&\quad \left. - \frac{1}{2}(-1)^n (n-1)!(2^n-1)\zeta(n) + o(1)\Big) \frac{s^n}{n!}\right) / N^{s/2+s^2/2} \\
&= \exp\left(\frac{\gamma}{2}s + \left(1+\gamma+\log 2 - \frac{3}{2}\zeta(2)\right)\frac{s^2}{2}\right. \\
&\quad + \sum_{n=3}^{\infty}\Big((-1)^n(2^{n-1}-1)\zeta(n-1) \\
&\quad \left. - \frac{1}{2}(-1)^n(2^n-1)\zeta(n)\Big)\frac{s^n}{n}\right). \qquad (4.2.29)
\end{aligned}
$$

This coefficient can be expressed as a combination of gamma functions and the Barnes G-function [Bar00, Vig79] in much the same way as we dealt with the circular ensembles in Chapter 2. We remember from that chapter the recurrence relation

$$
\begin{aligned}
G(1) &= 1, \qquad (4.2.30) \\
G(z+1) &= \Gamma(z)\, G(z),
\end{aligned}
$$

and furthermore, for $|z| < 1$ we have the expansion

$$
\log G(1+z) = (\log(2\pi) - 1)\frac{z}{2} - (1+\gamma)\frac{z^2}{2} + \sum_{n=3}^{\infty}(-1)^{n-1}\zeta(n-1)\frac{z^n}{n}. \qquad (4.2.31)
$$

Combining this with

$$
\log\Gamma(1+z) = -\gamma z + \sum_{n=2}^{\infty}\zeta(n)\frac{(-z)^n}{n}, \qquad (4.2.32)
$$

which holds for $|z|<1$, we see that, for $|s|<1/2$,

$$\begin{aligned}&\log G(1+s)-\frac{1}{2}\log G(1+2s)-\frac{1}{2}\log\Gamma(1+2s)+\frac{1}{2}\log\Gamma(1+s)\\&=\frac{\gamma}{2}s+(1+\gamma)\frac{s^2}{2}-\frac{3}{4}\zeta(2)s^2+\sum_{n=3}^{\infty}\Big((-1)^n(2^{n-1}-1)\zeta(n-1)\\&\quad-\frac{1}{2}(-1)^n(2^n-1)\zeta(n)\Big)\frac{s^n}{n}.\end{aligned}\tag{4.2.33}$$

A comparison with (4.2.29) shows that

$$f_{Sp}(s) \;=\; 2^{s^2/2}\times\frac{G(1+s)\sqrt{\Gamma(1+s)}}{\sqrt{G(1+2s)\Gamma(1+2s)}},\tag{4.2.34}$$

for $|s|<1/2$, and hence by analytic continuation for all $s$.

As a check we compare the result (4.2.34), when $s$ is an integer, with (4.2.28). Using (4.2.30) we see that

$$G(n)=\prod_{j=1}^{n-1}\Gamma(j),\tag{4.2.35}$$

and so for integer $n$,

$$\begin{aligned}f_{Sp}(n) &= 2^{n^2/2}\frac{G(1+n)\sqrt{\Gamma(1+n)}}{\sqrt{G(1+2n)\Gamma(1+2n)}}\\&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}\Gamma(j)\right)\sqrt{\Gamma(1+n)}}{\sqrt{\prod_{j=1}^{2n}\Gamma(j)\,\Gamma(1+2n)}}\\&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}\Gamma(j)\right)\sqrt{n!}}{\sqrt{\prod_{j=1}^{2n+1}(j-1)!}}\\&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}(j-1)!\right)\sqrt{n!}}{\sqrt{2^{2n-1}\,3^{2n-2}\,4^{2n-3}\cdots(2n-1)^2\,2n}}\end{aligned}$$

$$\begin{aligned}
f_{Sp}(n) &= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}(j-1)!\right)\sqrt{1}\sqrt{2}\cdots\sqrt{n-1}\sqrt{n}}{\sqrt{2}\sqrt{4}\cdots\sqrt{2n}\sqrt{2^{2n-2}3^{2n-2}4^{2n-4}\cdots(2n-2)^2(2n-1)^2}}\\
&= 2^{n^2/2}\frac{1^{n-1}2^{n-2}\cdots(n-2)^2(n-1)}{2^{n/2}2^{n-1}3^{n-1}4^{n-2}5^{n-2}\cdots(2n-2)(2n-1)}\\
&= 2^{n^2/2}\frac{1}{2^{n/2}2^{\sum_{j=1}^{n-1}j}\prod_{j=1}^{n}(2j-1)!!}\\
&= \left(\prod_{j=1}^{n}(2j-1)!!\right)^{-1}. \qquad (4.2.36)
\end{aligned}$$

We see that this agrees precisely with (4.2.28).

Following the ideas developed in [KS00b], these integer coefficients have also been calculated independently by Brézin and Hikami [BH99].

Having the generating function, $M_{Sp}(N,s)$, it is a short step to find the value distributions of both $\log Z(U,0)$ and $Z(U,0)$ itself. The distribution of $\log Z(U,0)/\log N$ is

$$\begin{aligned}
&\langle\delta(x-\log Z(U,0)/\log N)\rangle_{USp(2N)}\\
&\quad=\left\langle\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iy(x-\log Z(U,0)/\log N)}dy\right\rangle_{USp(2N)}\\
&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx}M_{Sp}(N,iy/\log N)dy \qquad (4.2.37)\\
&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx}e^{c_1 iy/\log N+c_2(iy/\log N)^2/2+c_3(iy/\log N)^3/3!+\cdots}dy \qquad (4.2.38)\\
&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx}\exp\left[\left(\frac{1}{2}\log N+O(1)\right)iy/\log N\right.\\
&\qquad\qquad\left.-(\log N+O(1))\frac{y^2}{2(\log N)^2}-i(O(1))\frac{y^3}{3!(\log N)^3}+\cdots\right]dy.
\end{aligned}$$

Thus

$$\begin{aligned}
&\lim_{N\to\infty}\langle\delta(x-\log Z(U,0)/\log N)\rangle_{USp(2N)}\\
&=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx+iy/2}dy \qquad (4.2.39)\\
&=\delta(x-1/2),
\end{aligned}$$

and so the distribution of values of $\log Z(U,0)/\log N$ tends to a delta function centred at $x = 1/2$.

If we instead retain the $y^2$ term in the exponent in (4.2.38), we have the central limit theorem

$$\lim_{N\to\infty}\left\langle \delta\left(x + \frac{c_1}{\sqrt{c_2}} - \frac{\log Z(U,0)}{\sqrt{c_2}}\right)\right\rangle_{USp(2N)} = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{x^2}{2}\right), \quad (4.2.40)$$

where $c_1$ and $c_2$ are related to $N$ by (4.2.18) and (4.2.22), respectively. For finite $N$, the exact distribution is of course given by (4.2.37), where $M_{Sp}(N,s)$ is defined by (4.2.9).

It is not difficult to determine as well the distribution of the values of $Z(U,0)$ itself; we will call this $P_{Sp}(N,x)$. As

$$M_{Sp}(N,s) = \int_0^\infty x^s P_{Sp}(N,x)dx, \quad (4.2.41)$$

we have, in exact analogy with (3.4.13),

$$\begin{aligned} P_{Sp}(N,x) &= \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-is} M_{Sp}(N,is)ds \quad (4.2.42)\\ &= \frac{1}{2\pi x}\int_{-\infty}^{\infty} e^{-is\log x + c_1 is - c_2 s^2/2 - ic_3 s^3/3! + \cdots} ds. \end{aligned}$$

Although $P_{Sp}(N,x)$ does not have a limiting distribution as $N \to \infty$, we suggest the approximation

$$\begin{aligned} P_{Sp}(N,x) &\approx \frac{1}{2\pi x}\int_{-\infty}^{\infty}\exp\left[-is\log x + ic_1 s - \frac{c_2 s^2}{2}\right]ds \quad (4.2.43)\\ &= \frac{1}{x\sqrt{2\pi c_2}}\exp\left(-\frac{(\log x - c_1)^2}{2c_2}\right), \quad (4.2.44) \end{aligned}$$

and plot it, for two values of $N$, in Figure 4.1 along with the exact distribution (4.2.42). It should be noted that the approximation (4.2.44) is valid when $x$ is fixed and $N \to \infty$. In this limit the terms in the exponent containing cumulants higher than the second become negligible, and also the stationary point of (4.2.43), $s = (ic_1 - i\log x)/c_2 \sim i/2 - i\log x/\log N$, tends to $i/2$ and so stays well away from the nearest pole of the integrand of (4.2.42) at $i3/2$. We can in fact be more explicit and say that (4.2.44) is expected to be a good approximation when $\log x >> -\log N$ and $N$ is large, which follows from an identical argument to that pertaining to (3.4.15).

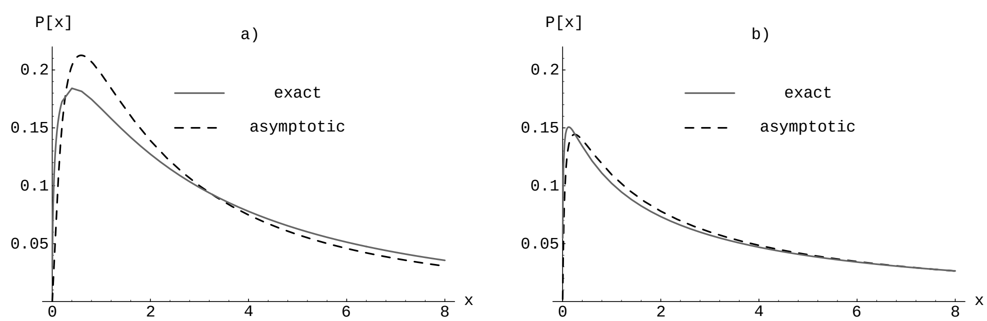


Figure 4.1: Distribution of the values of $Z(U,0)$ for matrices in $USp(2N)$, a)$N = 6$, b) $N = 42$. The solid curve is the exact distribution (4.2.42) and the dashed curve is the large $N$ approximation (4.2.44).

It may be seen from Figure 4.1 that $P_{Sp}(N,0) = 0$. Although the approximation (4.2.44) also tends to zero as $x \to 0$, it does not predict the correct rate of approach. This may instead be obtained by examining the poles of the integrand of

$$P_{Sp}(N,x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-is}2^{2iNs}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+is+j)}{\Gamma(1/2+j)\Gamma(1+is+N+j)}ds. \quad (4.2.45)$$

These poles are those of the factors $\Gamma(1/2+is+j)$ so they occur at the points $s = i(2k+1)/2$ and are of order $k$, for $k = 1, 2, \ldots, N$, then of order $N$ for all higher $k$.

First we consider the residue of the simple pole at $s = \frac{3}{2}i$. We can write

$$\Gamma(1/2+is+1) = \frac{\Gamma(1/2+is+2)}{1/2+is+1} = -i\frac{\Gamma(is+5/2)}{s-3i/2}, \tag{4.2.46}$$

so the residue of $\Gamma(1/2+is+1)$ at $s=\frac{3}{2}i$ is $-i$. As a result, and because none of the other factors have a singularity at $s=\frac{3}{2}i$, the residue of the integrand at $s=\frac{3}{2}i$ is

$$x^{3/2}2^{-3N}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)}{\Gamma(1/2+j)\Gamma(N+j-1/2)}\cdot\prod_{j=2}^{N}\Gamma(j-1)\cdot -i. \tag{4.2.47}$$

Due to the factor $x^{-is}$, in the limit $x\to 0$ the lowest pole (that at $s=\frac{3}{2}i$) gives the dominant contribution. This is because at $s=i(2k+1)/2$ there is a $k$th order pole of $\prod_{j=1}^{N}\Gamma(1/2+is+j)$ and the other factors are analytic. However, the expansion of $x^{-is}$ around $s=i(2k+1)/2$ produces a factor of $x^{(2k+1)/2}$. Thus the residue from the pole at $s=i(2k+1)/2$ will be of order $x^{(2k+1)/2}$ in $x$ and as $x\to 0$, the significant contribution will be from $k=1$.

From the residue at this lowest pole we thus find that as $x\to 0$

$$P_{Sp}(N,x)\sim x^{1/2}2^{-3N}\frac{1}{\Gamma(N)}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(j)}{\Gamma(1/2+j)\Gamma(N+j-1/2)}. \tag{4.2.48}$$

It is demonstrated in Appendix D that the contour of integration of (4.2.45) can be closed around the poles with negligible effect.

Alternatively, we can examine the value distribution of $Z(U,0)^{\frac{1}{\log N}}$. Changing variables in (4.2.42), results in

$$\langle\delta(x-Z(U,0)^{\frac{1}{\log N}})\rangle_{USp(2N)} = \frac{1}{2\pi x}\int_{-\infty}^{\infty}x^{-iy}M_{Sp}(N,iy/\log N)dy, \tag{4.2.49}$$

and so

$$\begin{aligned}\lim_{N\to\infty}\langle\delta(x-Z(U,0)^{\frac{1}{\log N}})\rangle_{USp(2N)} &= \frac{1}{2\pi x}\int_{-\infty}^{\infty}e^{-iy\log x}e^{iy/2}dy\\ &= \frac{1}{x}\delta(\log x-1/2).\end{aligned} \tag{4.2.50}$$

### 4.2.2 $L$-functions with symplectic symmetry

In Section 4.1 we gave a brief description of the mean values at $s = 1/2$ for families of $L$-functions and the relation of these to the symmetry type displayed by the low-lying zeros. Here we consider the case of symplectic symmetry in more depth.

If we again use Conrey and Farmer's notation, as in (4.1.1), then in the symplectic case they have $V(z) = z$ and find that $B(k) = \frac{1}{2}k(k+1)$ [CF99]. They also list several families which are conjectured to have low-lying zeros with symplectic symmetry, the simplest of which consists of the Dirichlet $L$-functions, $L(s, \chi_d)$, where $\chi_d$ is a quadratic Dirichlet character. These were discussed in Section 4.1. The conjectured form for the moments was given in (4.1.5), and for this case the first few values of $g_k$ for integer $k$ have been found using number-theoretic techniques to be [Jut81, Sou99, CF99] $g_1 = 1$, $g_2 = 2$, $g_3 = 2^4$ and, by conjecture, $g_4 = 3 \cdot 2^8$. It seems very likely, as was the case for the Riemann zeta function in Chapter 3, that $g_k$ should be related to the random matrix moment values calculated in Section 4.2.1 because $g_k$ is believed to be purely symmetry-determined. Thus we make a conjecture similar to (3.4.7) in the chapter on the Riemann zeta function.

Making the identification

$$N = \log(\mathcal{Q}^A), \tag{4.2.51}$$

and recalling that as $N \to \infty$ $M_{Sp}(N,k) \sim f_{Sp}(k) N^{\frac{1}{2}k(k+1)}$, we conjecture that for symplectic families of $L$-functions

$$\frac{g_k}{\Gamma(1 + \frac{1}{2}k(k+1))} = f_{Sp}(k). \tag{4.2.52}$$

Following the arguments of Chapter 3 for the case of the Riemann zeta function, the relation between $N$ and $\mathcal{Q}$ should arise from equating the mean densities of zeros. For the $L$-functions we need the density near $s = 1/2$

because we are dealing with the $L$-functions just at this point. The counting function for the zeros, that is the number of zeros in the critical strip up to certain height $T$, can be calculated by integrating the argument of a function $\xi(s)$ which has the same non-trivial zeros as the related $L$-function and also obeys a simple functional equation, around a rectangle containing the zeros which are being counted. The density of $L$-function zeros near the critical point can then be determined by differentiating the staircase function at some finite height $T$. In the case of $L$-functions with quadratic Dirichlet characters $\chi_d$, (4.1.5), the mean density at a fixed height up the critical line increases like $\frac{1}{2\pi}\log|d|$ as $|d| \to \infty$. Since the mean density of eigenvalues of a matrix in $USp(2N)$ is $N/\pi$, we equate $N = (1/2)\log D$, and obtain exactly the proposed relation (4.2.51), since $A = 1/2$ in this case.

It is then striking that the first few values of $f_{Sp}$ at the integers, $f_{Sp}(1) = 1$, $f_{Sp}(2) = \frac{1}{3}$, $f_{Sp}(3) = \frac{1}{45}$ and $f_{Sp}(4) = \frac{1}{4725}$, agree precisely, via (4.2.52), with the values that Conrey and Farmer report for the symplectic $L$-functions.

Thus the mean values of $L$-functions appear to be partially predicted by RMT in much the same way as for the Riemann zeta function. The only difference is that in the case of $\zeta(s)$ the average was along the critical line rather than over a family of functions. This is not a significant difference, however, and Conrey and Farmer in fact suggest that we think of the Riemann zeta function as a unitary family (with zeros showing the statistics of the eigenvalues of matrices from $U(N)$) in its own right, where we are averaging over special values of the family $\{\zeta(1/2+it)\}$ as $t$ ranges over the real numbers.

The validity of the conjecture (4.2.52) would imply many results on the value distribution of the central values of symplectic $L$-functions. The distribution for the logarithm of symplectic families of $L$-functions, for example, is expected to behave for asymptotically large $\mathcal{Q}$ in the same way as that of the characteristic polynomial $Z$, always remembering that $N$ must be related to the $L$-function parameter via the density of zeros. This is because the conjecture (4.2.52) can also be written as

$$\lim_{\mathcal{Q}\to\infty} \frac{1}{(\log \mathcal{Q}^A)^{\frac{1}{2}k(k+1)}\mathcal{Q}^*} \sum_{\substack{f\in\mathcal{F} \\ c(f)\le\mathcal{Q}}} L_f(\frac{1}{2})^k = a(k)\times\left(\lim_{N\to\infty}\frac{M_{Sp}(N,k)}{N^{\frac{1}{2}k(k+1)}}\right), \tag{4.2.53}$$

so the value distribution of $\log L_f(\frac{1}{2})/\log\log\mathcal{Q}^A$ defined by averages with $c(f)\le\mathcal{Q}$, would be, for large $\mathcal{Q}$ and making the identification (4.2.51),

$$V_{Sp}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixy} a(iy/\log N) M_{Sp}(N, iy/\log N)dy, \tag{4.2.54}$$

leading to

$$\lim_{N\to\infty} V_{Sp}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} a(0)e^{-ixy+iy/2}dy. \tag{4.2.55}$$

Since $a(0)=1$, we see that this would imply that the distribution of $\log L_f(\frac{1}{2})/\log\log\mathcal{Q}^A$ is asymptotic to $\delta(x-1/2)$, in just the same way as for $\log Z(U,0)/\log N$.

Following the same line of argument, we suggest that

$$\begin{aligned}
&\lim_{\mathcal{Q}\to\infty}\left\langle \delta\left(x+\frac{\tilde{c_1}}{\sqrt{\tilde{c_2}}}-\frac{\log L_f\left(\frac{1}{2}\right)}{\sqrt{\tilde{c_2}}}\right)\right\rangle_{\mathcal{F}} \\
&= \lim_{N\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty} a\left(\frac{iy}{\sqrt{c_2}}\right) e^{-iyx-iyc_1/\sqrt{c_2}} e^{iyc_1/\sqrt{c_2}-y^2/2+c_3(iy)^3/(c_2^{3/2}3!)+\cdots}dy \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right),
\end{aligned} \tag{4.2.56}$$

where $\langle\cdot\rangle_{\mathcal{F}}$ denotes an average over a family $\mathcal{F}$ of $L$-functions, as in (4.1.1), and $\tilde{c_1}$ and $\tilde{c_2}$ are given by (4.2.18) and (4.2.22), respectively, again with the identification (4.2.51).

If we now turn to the distribution of values of $L_f(\frac{1}{2})$ itself, $W_{Sp}(x)$, we can close the contour of

$$W_{Sp}(x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-is} a(is) M_{Sp}(N,is)ds \tag{4.2.57}$$

around the poles and obtain, as $x \to 0$, the dominant contribution from the pole at $s = (3i)/2$:

$$W_{Sp}(x) \sim x^{1/2} a(-3/2) 2^{-3N} \frac{1}{\Gamma(N)} \prod_{j=1}^{N} \frac{\Gamma(1+N+j)\Gamma(j)}{\Gamma(1/2+j)\Gamma(N+j-1/2)}. \tag{4.2.58}$$

This is of particular note in the light of recent interest in the non-vanishing of the central values of $L$-functions, see for example [Sou99, IS97, ILS99] and references therein. Clearly (4.2.58) implies that as long as $a(-3/2)$ is finite for a particular family of symplectic $L$-functions, the probability that the central value of those $L$-functions lies in the range $(0,x)$ decreases like $x^{3/2}$ as $x \to 0$.

## 4.3 Orthogonal symmetry

### 4.3.1 Random matrices in $SO(2N)$

To perform the same calculations for the orthogonal matrices, we must choose the group of matrices carefully. The behaviour of the eigenvalues of the orthogonal matrices depend on which of four categories we are dealing with. While the symplectic matrices always had determinant 1, an orthogonal matrix can have 1 or -1 as a determinant. If the dimension of the matrix is even and the determinant is 1, then the eigenvalues come in complex conjugate pairs excluding 1 and -1. If, however, the determinant is -1 when the dimension is still even, -1 must be an eigenvalue. As all the others must come in conjugate pairs, there must be an eigenvalue at 1 to pair with -1. If the dimension of the matrix is odd, though, 1 is the extra (unpaired) eigenvalue if the determinant is positive, and it is -1 for matrices with negative determinant.

The calculations below pertain to $2N \times 2N$ orthogonal matrices with positive determinant, that is, to the group $SO(2N)$. The joint probability density function for the eigenvalues of this group is very similar to that of the group of symplectic matrices, so the calculations below follow much the same lines as in Section 4.2.1.

The joint probability density function for the eigenvalues dictated by Haar measure on $SO(2N)$ is [Wey46],

$$N_O \prod_{1\leq i<j\leq N} \left(\frac{1}{2}(\cos\theta_j - \cos\theta_i)\right)^2, \tag{4.3.1}$$

where

$$N_O = 2^{-N} \prod_{j=1}^{N} \frac{\Gamma(N+j-1)}{\Gamma(1+j)(\Gamma(j-1/2))^2} = \frac{2^{2N^2-4N+1}}{\pi^N N!}. \tag{4.3.2}$$

For matrices in $SO(2N)$ the characteristic polynomial takes the same form as for $USp(2N)$,

$$Z(U,\theta) = \prod_{n=1}^{N} \left(1 - e^{i(\theta_n-\theta)}\right)\left(1 - e^{i(-\theta_n-\theta)}\right), \tag{4.3.3}$$

so the $s^{th}$ moment is

$$\begin{aligned}
&\langle Z(U,0)^s\rangle_{SO(2N)} \\
&= N_O \int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \prod_{1\leq i<j\leq N}\left(\frac{1}{2}(\cos\theta_j-\cos\theta_i)\right)^2 \\
&\quad \times 2^{Ns}\prod_{n=1}^{N}(1-\cos\theta_n)^s \\
&= N_O\, 2^{2N-N^2+Ns}\int_0^{\pi}\cdots\int_0^{\pi} d\theta_1\cdots d\theta_N \prod_{1\leq i<j\leq N}(\cos\theta_j-\cos\theta_i)^2 \\
&\quad \times \prod_{n=1}^{N}(1-\cos\theta_n)^s \\
&= N_O\, 2^{2N-N^2+Ns}\int_{-1}^{1}\cdots\int_{-1}^{1} dx_1\cdots dx_N \prod_{1\leq i<j\leq N}(x_j-x_i)^2 \\
&\quad \times \prod_{n=1}^{N}(1-x_n^2)^{-1/2}(1-x_p)^s.
\end{aligned} \tag{4.3.4}$$

We use the Selberg integral (4.2.8) again, this time with $\gamma = 1$, $\alpha = s + 1/2$ and $\beta = 1/2$.

$$\begin{aligned}
\langle Z(U,0)^s \rangle_{SO(2N)} &= N_O\, 2^{2N-N^2+Ns} \cdot 2^{N^2-N+sN+N/2+N/2-N} \\
&\quad \times \prod_{j=0}^{N-1} \frac{\Gamma(2+j)\Gamma(s+1/2+j)\Gamma(1/2+j)}{\Gamma(2)\Gamma(s+1+N+j-1)} \\
&= N_O\, 2^{N+2Ns} \prod_{j=1}^{N} \frac{\Gamma(1+j)\Gamma(s+j-1/2)\Gamma(j-1/2)}{\Gamma(s+N+j-1)} \\
&= 2^{2Ns} \prod_{j=1}^{N} \frac{\Gamma(N+j-1)\Gamma(s+j-1/2)}{\Gamma(j-1/2)\Gamma(s+j+N-1)} \\
&\equiv M_O(N,s). \qquad (4.3.5)
\end{aligned}$$

We now start looking at the cumulants of $M_O$. As for the symplectic case, we write

$$M_O(N,s) = e^{q_1 s + q_2 s^2/2 + q_3 s^3/3! + q_4 s^4/4! + \cdots} \qquad (4.3.6)$$

so the cumulants are derivatives of $\log M_O(N,s)$ evaluated at $s = 0$. The logarithm of $M_O$ is

$$\begin{aligned}
\log M_O(N,s) &= 2Ns\log 2 + \sum_{j=1}^{N} (\log\Gamma(N+j-1) + \log\Gamma(s+j-1/2) \\
&\quad - \log\Gamma(j-1/2) - \log\Gamma(s+j+N-1)). \qquad (4.3.7)
\end{aligned}$$

The first derivative is

$$\begin{aligned}
&\frac{d}{ds}\log M_O(N,s) \\
&= 2N\log 2 + \sum_{j=1}^{N} \frac{d}{ds}(\log\Gamma(s+j-1/2) - \log\Gamma(s+j+N-1)) \\
&= 2N\log 2 + \sum_{j=1}^{N} (\psi(s+j-1/2) - \psi(s+j+N-1)), \qquad (4.3.8)
\end{aligned}$$

and so the cumulants are

$$\begin{aligned} q_1 &= \frac{d}{ds} \log M_O(N,s)\Big|_{s=0} \\ &= 2N \log 2 + \sum_{j=1}^{N} (\psi(j-1/2) - \psi(j+N-1)) \end{aligned} \tag{4.3.9}$$

for the first, and for higher cumulants

$$\begin{aligned} q_n &= \frac{d^n}{ds^n} \log M_O(N,s)\Big|_{s=0} \\ &= \sum_{j=1}^{N} \left(\psi^{(n-1)}(j-1/2) - \psi^{(n-1)}(j+N-1)\right). \end{aligned} \tag{4.3.10}$$

To discover the leading-order behaviour of these cumulants, we repeat very nearly what we did in the Section 4.2.1. Starting with $q_1$ and making use again of (4.2.15) and (4.2.14),

$$\begin{aligned} q_1 &= 2N \log 2 + \sum_{j=1}^{N} \left( \int_0^\infty \frac{e^{-t} - e^{-(j-1/2)t}}{1-e^{-t}} dt - \gamma \right. \\ &\qquad \left. - \int_0^\infty \frac{e^{-t} - e^{-(j+N-1)t}}{1-e^{-t}} dt + \gamma \right) \\ &= 2N \log 2 + \sum_{j=1}^{N} \int_0^\infty \frac{e^{-(j+N-1)t} - e^{-(j-1/2)t}}{1-e^{-t}} dt \\ &= 2N \log 2 + \int_0^\infty \frac{e^{-Nt}(1-e^{-Nt}) - e^{-t/2}(1-e^{-Nt})}{(1-e^{-t})^2} dt \\ &= 2N \log 2 + \int_0^\infty \frac{e^{-Nt} - e^{-2Nt} - e^{-t/2} + e^{-(N+1/2)t}}{(1-e^{-t})^2} dt. \end{aligned}$$

Now we integrate by parts,

$$\begin{aligned}
q_1 &= 2N\log 2 - (N-1)\int_0^\infty \frac{e^{-Nt}}{1-e^{-t}}dt + (2N-1)\int_0^\infty \frac{e^{-2Nt}}{1-e^{-t}}dt \\
&\quad -\frac{1}{2}\int_0^\infty \frac{e^{-t/2}}{1-e^{-t}}dt - (N-1/2)\int_0^\infty \frac{e^{-(N+1/2)t}}{1-e^{-t}}dt \\
&= 2N\log 2 + (N-1)\psi(N) - (2N-1)\psi(2N) + \frac{1}{2}\psi(1/2) \\
&\quad +(N-1/2)\psi(1/2+N) \\
&= 2N\log 2 + (N-1)\left(\log N - \frac{1}{2N} + O(N^{-2})\right) \\
&\quad -(2N-1)\left(\log 2 + \log N - \frac{1}{4N} + O(N^{-2})\right) - \frac{\gamma}{2} \\
&\quad -\log 2 + (N-1/2)\left(\log(N+1/2) - \frac{1}{1+2N} + O(N^{-2})\right) \\
&= -\frac{1}{2}\log N - \frac{\gamma}{2} + O\left(\frac{1}{N}\right). \qquad (4.3.11)
\end{aligned}$$

Now moving on to the second derivative, this time calling upon (4.2.20) and (4.2.19),

$$\begin{aligned}
q_2 &= \sum_{j=1}^{N}\left(\psi^{(1)}(j-1/2) - \psi^{(1)}(j+N-1)\right) \\
&= \sum_{j=1}^{N}\left(\int_0^\infty \frac{te^{-(j-1/2)t}}{1-e^{-t}}dt - \int_0^\infty \frac{te^{-(j+N-1)t}}{1-e^{-t}}dt\right) \\
&= \int_0^\infty \frac{te^{-t/2}(1-e^{-Nt})}{(1-e^{-t})^2}dt - \int_0^\infty \frac{te^{-Nt}(1-e^{-Nt})}{(1-e^{-t})^2}dt.
\end{aligned}$$

Again integration by parts helps us to

$$\begin{aligned}
q_2 &= \int_0^\infty \frac{e^{-t/2}}{1-e^{-t}}dt + \frac{1}{2}\int_0^\infty \frac{te^{-t/2}}{1-e^{-t}}dt - \int_0^\infty \frac{e^{-(N+1/2)t}}{1-e^{-t}}dt \\
&\quad +(N-1/2)\int_0^\infty \frac{te^{-(N+1/2)t}}{1-e^{-t}}dt - \int_0^\infty \frac{e^{-Nt}}{1-e^{-t}}dt \\
&\quad +(N-1)\int_0^\infty \frac{te^{-Nt}}{1-e^{-t}}dt + \int_0^\infty \frac{e^{-2Nt}}{1-e^{-t}}dt \\
&\quad -(2N-1)\int_0^\infty \frac{te^{-2Nt}}{1-e^{-t}}dt \\
&= -\psi(1/2) + \frac{1}{2}\psi^{(1)}(1/2) + \psi(N+1/2) + (N-1/2)\psi^{(1)}(N+1/2) \\
&\quad +\psi(N) + (N-1)\psi^{(1)}(N) - \psi(2N) - (2N-1)\psi^{(1)}(2N) \\
&= \log N + 1 + \gamma + \log 2 + \frac{3}{2}\zeta(2) + O\left(\frac{1}{N}\right). \qquad (4.3.12)
\end{aligned}$$

The higher cumulants converge to a constant as $N$ increases, and following our usual routine we obtain,

$$\begin{aligned}
q_n &= \sum_{j=1}^N \left(\psi^{(n-1)}(j-1/2) - \psi^{(n-1)}(j+N-1)\right) \\
&= \sum_{j=1}^N \left((-1)^n \int_0^\infty \frac{t^{n-1}e^{-(j-1/2)t}}{1-e^{-t}}dt - (-1)^n \int_0^\infty \frac{t^{n-1}e^{-(j+N-1)t}}{1-e^{-t}}dt\right) \\
&= (-1)^n \int_0^\infty \frac{t^{n-1}e^{-t/2}}{1-e^{-t}}\frac{1-e^{-Nt}}{1-e^{-t}}dt - (-1)^n \int_0^\infty \frac{t^{n-1}e^{-Nt}}{1-e^{-t}}\frac{1-e^{-Nt}}{1-e^{-t}}dt
\end{aligned}$$

so,

$$\begin{aligned}
\lim_{N\to\infty} q_n &= (-1)^n \int_0^\infty \frac{t^{n-1}e^{-t/2}}{(1-e^{-t})^2}dt \\
&= (-1)^n \int_0^\infty \frac{-t^{n-1}e^{t/2}(-e^{-t})}{(1-e^{-t})^2}dt \\
&= (-1)^n \left[(n-1)\int_0^\infty t^{n-2}\frac{e^{t/2}}{1-e^{-t}}dt + \frac{1}{2}\int_0^\infty t^{n-1}\frac{e^{t/2}}{1-e^{-t}}dt\right] \\
&= -(n-1)\psi^{(n-2)}(-1/2) + \frac{1}{2}\psi^{(n-1)}(-1/2). \qquad (4.3.13)
\end{aligned}$$

We move now from the cumulants back to the moments of $Z(U,0)$. As in the symplectic case, we can examine the first few integer moments individually.

$$\begin{aligned}
&\langle Z(U,0)\rangle_{SO(2N)} \\
&= 2^{2N}\prod_{j=1}^{N}\frac{\Gamma(N+j-1)\Gamma(j+1/2)}{\Gamma(j-1/2)\Gamma(j+N)} \\
&= 2^{2N}\frac{\Gamma(N)\Gamma(N+1)\cdots\Gamma(2N-1)\Gamma(3/2)\cdots\Gamma(N-1/2)\Gamma(N+1/2)}{\Gamma(1/2)\Gamma(3/2)\cdots\Gamma(N-1/2)\Gamma(N+1)\cdots\Gamma(2N-1)\Gamma(2N)} \\
&= 2^{2N}\frac{\Gamma(N)\Gamma(N+1/2)}{\Gamma(1/2)\Gamma(2N)} \\
&= 2^{2N}\frac{(N-1)!(N-1/2)!}{(-1/2)!(2N-1)!} \\
&= 2^{2N}\frac{(N-1/2)(N-3/2)\cdots(5/2)(3/2)(1/2)}{(2N-1)(2N-2)\cdots(N+2)(N+1)N} \\
&= 2^{N}\frac{(2N-1)(2N-3)\cdots 5\cdot 3\cdot 1}{(2N-1)(2N-2)\cdots(N+2)(N+1)N} \\
&= \begin{cases} 2^{N}\frac{(N-1)(N-3)\cdots 5\cdot 3\cdot 1}{(2N-2)(2N-4)\cdots(N+2)N} & N \text{ even} \\ 2^{N}\frac{(N-2)(N-4)\cdots 5\cdot 3\cdot 1}{(2N-2)(2N-4)\cdots(N+3)(N+1)} & N \text{ odd}\end{cases} \\
&= \begin{cases} 2^{N/2}\frac{(N-1)!}{(N/2-1)!2^{N/2-1}}\frac{1}{(N-1)(N-2)\cdots(N/2+1)(N/2)} & N \text{ even} \\ 2^{N/2+1/2}\frac{(N-2)!}{(N/2-3/2)!2^{N/2-3/2}}\frac{1}{(N-1)(N-2)\cdots(N/2+3/2)(N/2+1/2)} & N \text{ odd}\end{cases} \\
&= 2. \qquad (4.3.14)
\end{aligned}$$

Therefore the second moment becomes

$$\begin{aligned}
\langle Z(U,0)^2\rangle_{SO(2N)} &= 2^{4N}\prod_{j=1}^{N}\frac{\Gamma(N+j-1)\Gamma(3/2+j)}{\Gamma(j-1/2)\Gamma(1+j+N)} \\
&= 2^{4N}\frac{\Gamma(N)\Gamma(N+1)\Gamma(N+1/2)\Gamma(N+3/2)}{\Gamma(1/2)\Gamma(3/2)\Gamma(2N)\Gamma(2N+1)} \\
&= 2^{4N}\frac{(N-1)!N!(N-1/2)!(N+1/2)!}{(-1/2)!(1/2)!(2N-1)!(2N)!} \\
&= \left(2^{2N}\frac{(N-1)!(N-1/2)!}{(-1/2)!(2N-1)!}\right)^2\frac{N(N+1/2)}{(1/2)(2N)} \\
&\sim 4N, \qquad (4.3.15)
\end{aligned}$$

and the third comes out as

$$
\begin{aligned}
&\langle Z(U,0)^3\rangle_{SO(2N)} \\
&= 2^{6N}\prod_{j=1}^{N}\frac{\Gamma(N+j-1)\Gamma(5/2+j)}{\Gamma(j-1/2)\Gamma(2+j+N)} \\
&= 2^{6N}\frac{\Gamma(N)\Gamma(N+1)\Gamma(N+2)\Gamma(N+1/2)\Gamma(N+3/2)\Gamma(N+5/2)}{\Gamma(1/2)\Gamma(3/2)\Gamma(5/2)\Gamma(2N)\Gamma(2N+1)\Gamma(2N+2)} \\
&= 2^{6N}\frac{(N-1)!N!(N+1)!(N-1/2)!(N+1/2)!(N+3/2)!}{(-1/2)!(1/2)!(3/2)!(2N-1)!(2N)!(2N+1)!} \\
&= \left(2^{2N}\frac{(N-1)!(N-1/2)!}{(-1/2)!(2N-1)!}\right)^3\frac{N^2(N+1)(N+1/2)(N+1/2)(N+3/2)}{(1/2)(1/2)(3/2)(2N)(2N)(2N+1)} \\
&\sim 8\frac{N^3}{3}. \qquad (4.3.16)
\end{aligned}
$$

The general leading-order behaviour seems to be (where the product in the brackets is just 1 for $n=1$)

$$
\langle Z(U,0)^n\rangle_{SO(2N)} \sim 2^n\left(\prod_{j=1}^{n-1}(2j-1)!!\right)^{-1} N^{\frac{1}{2}n(n-1)}, \qquad (4.3.17)
$$

and this is easily proved by induction. This result was also obtained independently in [BH99].

This yields the coefficient of the leading-order term for integer moments, but we can do as in the symplectic case and express this coefficient in terms of the Barnes G-function and gamma functions.

Using our knowledge of the cumulants, this leading-order coefficient is

$$\begin{aligned}
f_O(s) &\equiv \lim_{N\to\infty} \frac{M_O(N,s)}{N^{s^2/2-s/2}} \\
&= \lim_{N\to\infty} \exp\left[\left(-\frac{1}{2}\log N - \frac{1}{2}\gamma + O(N^{-1})\right) s \right. \\
&\qquad + \left(\log N + 1 + \gamma + \log 2 + \frac{3}{2}\zeta(2) + O(N^{-1})\right)\frac{s^2}{2} \\
&\qquad + \sum_{n=3}^{\infty}\left(-(n-1)\psi^{(n-2)}(-1/2)\right. \\
&\qquad \left.\left. + \frac{1}{2}\psi^{(n-1)}(-1/2) + o(1)\right)\frac{s^n}{n}\right] / N^{s^2/2-s/2} \\
&= \exp\left[-\frac{\gamma}{2}s + \left(1 + \gamma + \log 2 + \frac{3}{2}\zeta(2)\right)\frac{s^2}{2}\right. \\
&\qquad + \sum_{n=3}^{\infty}\left((-1)^n(2^{n-1}-1)\zeta(n-1)\right. \\
&\qquad \left.\left. + \frac{1}{2}(-1)^n(2^n-1)\zeta(n)\right)\frac{s^n}{n}\right]. \qquad (4.3.18)
\end{aligned}$$

Examining the product form of $M_O(N,s)$ we see that the coefficient is expected to have poles of order $k$ at $s = -(2k-1)/2$, for $k = 1, 2, 3 \ldots$. Using (4.2.31) and (4.2.32), we see that a combination with the correct poles is (for $|s| < 1/2$)

$$\begin{aligned}
&\log G(1+s) - \frac{1}{2}\log G(1+2s) + \frac{1}{2}\log\Gamma(1+2s) - \frac{1}{2}\log\Gamma(1+s) \\
&= -\frac{\gamma}{2} + \left(1 + \gamma + \frac{3}{2}\zeta(2)\right)\frac{s^2}{2} + \sum_{n=3}^{\infty}\left((-1)^n(2^{n-1}-1)\zeta(n-1)\right. \\
&\qquad \left. + \frac{1}{2}(-1)^n(2^n-1)\zeta(n)\right)\frac{s^n}{n}, \qquad (4.3.19)
\end{aligned}$$

and comparing with (4.3.18) we thus find that

$$f_O(s) = 2^{s^2/2} \times \frac{G(1+s)\sqrt{\Gamma(1+2s)}}{\sqrt{G(1+2s)\Gamma(1+s)}}, \qquad (4.3.20)$$

for $|s| < 1/2$, and hence by analytic continuation in the rest of the complex plane.

This leading-order coefficient reduces for integer moments, again using (4.2.30), to

$$
\begin{aligned}
f_O(n) &= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}\Gamma(j)\right)\sqrt{\Gamma(1+2n)}}{\sqrt{\left(\prod_{j=1}^{2n}\Gamma(j)\right)\Gamma(1+n)}}\\
&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}(j-1)!\right)\sqrt{(2n)!}}{\sqrt{2^{2n-2}3^{2n-3}\cdots(2n-2)^2(2n-1)n!}}\\
&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}(j-1)!\right)\sqrt{2^n n!}\sqrt{(2n-1)!!}}{2^{n-1}4^{n-2}\cdots(2n-2)\sqrt{3^{2n-3}5^{2n-5}\cdots(2n-1)n!}}\\
&= 2^{n^2/2}\frac{\left(\prod_{j=1}^{n}(j-1)!\right)\sqrt{2^n}}{2^{\sum_{j=1}^{n-1}j}1^{n-1}2^{n-2}\cdots(n-1)\sqrt{3^{2n-4}5^{2n-6}\cdots(2n-3)^2}}\\
&= 2^{n^2/2}\frac{1^{n-1}2^{n-2}\cdots(n-2)^2(n-1)2^{n/2}}{2^{n(n-1)/2}3^{n-2}5^{n-3}\cdots(2n-3)1^{n-1}2^{n-2}\cdots(n-1)}\\
&= 2^n\left(\prod_{j=1}^{n-1}(2j-1)!!\right)^{-1}, \qquad (4.3.21)
\end{aligned}
$$

which agrees with (4.3.17), as it should.

Once more, we can examine the value distribution of $Z(U,0)$ and its logarithm. The value distribution of $\log Z(U,0)/\log N$ is

$$
\begin{aligned}
&\langle\delta(x-\log Z(U,0)/\log N)\rangle_{SO(2N)}\\
&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx}M_O(N,iy/\log N)dy \qquad (4.3.22)\\
&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[-iyx+\left(-\frac{1}{2}\log N+O(1)\right)iy/\log N\right.\\
&\qquad\left.-(\log N+O(1))\frac{y^2}{2(\log N)^2}-i(O(1))\frac{y^3}{3!(\log N)^3}+\cdots\right]dy,
\end{aligned}
$$

yielding the limiting distribution

$$
\begin{aligned}
\lim_{N\to\infty}\langle\delta(x-\log Z(U,0)/\log N)\rangle_{SO(2N)} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyx-iy/2}dy\\
&= \delta(x+1/2). \qquad (4.3.23)
\end{aligned}
$$

This is a delta distribution as in the symplectic case, but this time centred at $x = -1/2$.

Keeping the $y^2$ term in the exponent in the integral above leads to the central limit theorem:

$$\lim_{N\to\infty}\left\langle \delta\left(x+\frac{q_1}{\sqrt{q_2}}-\frac{\log Z(U,0)}{\sqrt{q_2}}\right)\right\rangle_{SO(2N)} = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{x^2}{2}\right). \quad (4.3.24)$$

The value distribution of $Z(U,0)^{\frac{1}{\log N}}$ is similarly straightforward to compute. We see that, just as in (4.2.49),

$$\langle\delta(x-Z(U,0)^{\frac{1}{\log N}})\rangle_{SO(2N)} = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-iy}M_O(N,iy/\log N)dy, \quad (4.3.25)$$

and so

$$\begin{aligned}\lim_{N\to\infty}\langle\delta(x-Z(U,0)^{\frac{1}{\log N}})\rangle_{SO(2N)} &= \frac{1}{2\pi x}\int_{-\infty}^{\infty} e^{-iy\log x}e^{-iy/2}dy \\ &= \frac{1}{x}\delta(\log x+1/2). \quad (4.3.26)\end{aligned}$$

We also examine the distribution simply of $Z(U,0)$, $P_O(N,x)$. As

$$P_O(N,x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-iy}M_O(N,iy)dy, \quad (4.3.27)$$

we can make the approximation

$$\begin{aligned}P_O(N,x) &\approx \frac{1}{2\pi x}\int_{-\infty}^{\infty}\exp\left[-iy\log x+iq_1y-\frac{q_2y^2}{2}\right]dy \\ &= \frac{1}{x\sqrt{2\pi q_2}}\exp\left(\frac{-(\log x-q_1)^2}{2q_2}\right), \quad (4.3.28)\end{aligned}$$

valid as $N\to\infty$ when $x$ is fixed (and, like (4.2.44), expected to be a good approximation when $\log x >> -\log N$ and $N$ is large). The result (4.3.28)

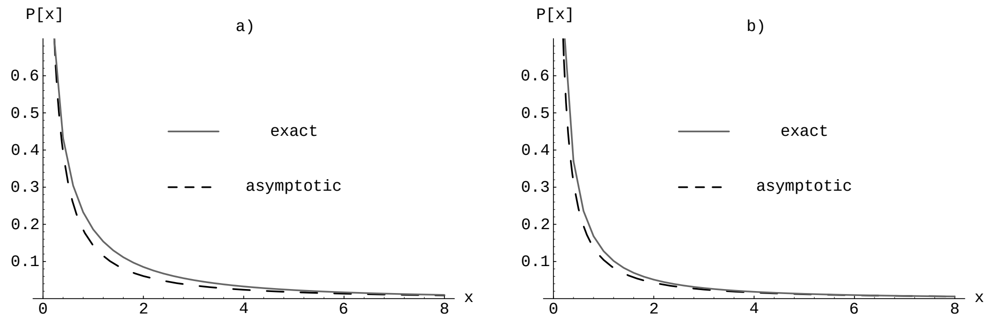


Figure 4.2: Distribution of the values of $Z(U,0)$ for matrices in the group $SO(2N)$, with a) $N=6$ and b) $N=42$. The solid curve is the exact distribution (4.3.27) and the dashed curve is the large $N$ approximation in (4.3.28).

is plotted in Figure 4.2 for $N=6$ and $N=42$ along with the numerically calculated exact distribution, from (4.3.27).

Unlike the symplectic case (and unlike the approximation (4.3.28)), $P_O(N,x)$ diverges as $x\to 0$. This can be seen by considering the poles of the integrand, which occur at $i/2, 3i/2, 5i/2, \ldots$. Once again it is the lowest pole, the simple one at $i/2$, that dominates the integral as $x\to 0$. In this case we find that

$$P_O(N,x)\sim x^{-1/2}2^{-N}\frac{1}{\Gamma(N)}\prod_{j=1}^{N}\frac{\Gamma(N+j-1)\Gamma(j)}{\Gamma(j-1/2)\Gamma(j+N-3/2)}, \qquad (4.3.29)$$

in that limit.

### 4.3.2 $L$-functions with orthogonal symmetry

We now turn our attention to families of $L$-functions with a symmetry governed by an ensemble of orthogonal matrices. $L$-functions of this type fall into two categories, even and odd, which are related to the ensembles $SO(2N)$ and $SO(2N+1)$ respectively. Of the $L$-functions comprising an orthogonal family, approximately one half will have even symmetry, and the other half odd symmetry, these latter vanishing at $s = 1/2$.

Examples of such families are given in [CF99]. Referring to (4.1.1), in the orthogonal case $V(z) = z$ and $B(k) = \frac{1}{2}k(k-1)$. As in the symplectic case, the first few of the coefficients $g_k$ with integer coefficients have been calculated. The known values are $g_1 = 1$, $g_2 = 2$, $g_3 = 2^3$ and it is conjectured that $g_4 = 2^7$ [CF99].

With $N$ taking the place of $\log(Q^A)$, we conjecture this time that

$$\lim_{\mathcal{Q}\to\infty} \frac{1}{(\log \mathcal{Q}^A)^{\frac{1}{2}k(k-1)}\mathcal{Q}^*} \sum_{\substack{f \in \mathcal{F} \\ c(f) \leq \mathcal{Q}}} L_f(\frac{1}{2})^k = a(k) \times f_O(k)/2. \tag{4.3.30}$$

The right hand side is divided by two because the random matrix average was just over $SO(2N)$, whereas the sum over central values of the $L$-functions contains an equal number of functions contributing zero to the average; namely the $L$-functions with odd symmetry about $s = 1/2$. Once again, we expect the relation (4.2.51) to follow from equating the density of zeros of the $L$-functions and the density of eigenphases of the matrices.

Having posed the conjecture (4.3.30), we check it against the known values of $g_k$. It is clear that the first four coefficients $f_O(1) = 2$, $f_O(2) = 4$, $f_O(3) = \frac{8}{3}$ and $f_O(4) = \frac{16}{45}$ satisfy conjecture (4.3.30); that is $g_k/\Gamma(1 + \frac{1}{2}k(k-1)) = f_O(k)/2$, for $k = 1, 2, 3, 4$.

As for the symplectic case, we can examine what (4.3.30) implies about the value distributions of $L$-functions and their logarithms. Since the $L$-functions with odd symmetry are zero at $s = 1/2$, we now restrict ourselves to averages over the orthogonal $L$-functions with even symmetry. These are expected to satisfy (4.3.30) without the factor 1/2 on the right hand side.

The value distribution of $\log L_f\left(\frac{1}{2}\right)/\log\log \mathcal{Q}^A$ for $L$-functions with even symmetry (defined by averaging as in (4.3.30)) is expected to be given, for large $\mathcal{Q}$, and with the identification (4.2.51), by

$$V_O(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixs} a(is/\log N) M_O(N, is/\log N) ds, \tag{4.3.31}$$

and following the argument laid out for the symplectic case, this converges to $\delta(x+1/2)$ as $N \to \infty$.

We can once again state a conjectural central limit theorem, this time for averages over a family $\mathcal{F}$ of $L$-functions with $c(f) \leq \mathcal{Q}$ governed by the symmetry $SO(2N)$:

$$\begin{aligned}
&\lim_{\mathcal{Q}\to\infty} \left\langle \delta\left( x + \frac{\tilde{q}_1}{\sqrt{\tilde{q}_2}} - \frac{\log L_f\left(\frac{1}{2}\right)}{\sqrt{\tilde{q}_2}} \right) \right\rangle_{\mathcal{F}} \\
&\quad = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} a\left( \frac{iy}{\sqrt{q_2}} \right) e^{-iyx - iyq_1/\sqrt{q_2}} e^{iyq_1/\sqrt{q_2} - y^2/2 + q_3(iy)^3/(q_2^{3/2}3!) + \cdots} dy \\
&\quad = \frac{1}{\sqrt{2\pi}} \exp\left( -\frac{x^2}{2} \right),
\end{aligned} \tag{4.3.32}$$

where $\tilde{q}_1$ and $\tilde{q}_2$ are related to (4.3.11) and (4.3.12), respectively, via (4.2.51).

For the value distribution of $L_f(\frac{1}{2})$ itself, which the conjecture suggests for large $\mathcal{Q}$ is

$$W_O(x) = \frac{1}{2\pi x}\int_{-\infty}^{\infty} x^{-is} a(is) M_O(N, is) ds, \tag{4.3.33}$$

we expect that near $x = 0$,

$$W_O(x) \sim x^{-1/2} a(-1/2) 2^{-N} \frac{1}{\Gamma(N)} \prod_{j=1}^{N} \frac{\Gamma(N+j-1)\Gamma(j)}{\Gamma(j-1/2)\Gamma(j+N-3/2)}; \tag{4.3.34}$$

the contribution to the integral (4.3.33) from the simple pole at $s = i/2$. For $L$-functions with even symmetry from an orthogonal family for which $a(-1/2) \neq 0$, this analysis therefore suggests that the likelihood that the central value vanishes is integrably singular, and that the probability of a value in the range $(0, x)$ vanishes as $x^{1/2}$ when $x \to 0$.

# Chapter 5

# Encore: Dynamical Zeta Functions

Until this point we have concentrated on number theoretical functions, the Riemann zeta function and $L$-functions, as examples of functions which have zeros displaying random matrix statistics. As was described in Chapter 1, however, such statistics also occur in physics; the eigenvalues of quantum systems with a chaotic classical counterpart are expected to show either COE, CSE or CUE statistics in the semiclassical limit. An example of a function with appropriate zeros is therefore the spectral determinant (1.1.14) of such a system, as this function has a zero at each eigenvalue.

Through the analogy (1.4.7) between prime numbers and the role of classical periodic orbits in semiclassical expressions, a glance at the form of the moments conjectured for the Riemann zeta function (3.4.8) suggests that in the semiclassical limit similar mean values of the spectral determinant might take the form of a product over periodic orbits multiplied by the appropriate random matrix moment from Chapter 2.

## 5.1 Value distributions of $\mathcal{Z}(E)$

The spectral determinant defined in the Introduction (1.1.14) can be written using the notation

$$\Delta(E) \sim B(E)\exp(-i\pi\overline{\mathcal{N}}(E))\mathcal{Z}(E), \tag{5.1.1}$$

where

$$\mathcal{Z}(E) = \prod_p \exp\left(-\sum_{m=1}^{\infty} \frac{\exp(imS_p/\hbar)}{m\sqrt{|\det(M_p^m - 1)|}}\right) \tag{5.1.2}$$

is called a dynamical zeta function.

It is this function $\mathcal{Z}(E)$ in which we will be interested. There is always ambiguity as to how comparison should be made with the random matrix function $Z(U,\theta)$. Clearly multiplication by a function without zeros would not change the zero distribution of $\mathcal{Z}$, but it would change the moments and value distribution. To avert this problem, we note that the real and imaginary part of the logarithm of $Z(U,\theta)$ both have zero mean; that is, the first moment of $\mathrm{Re}\log Z$ and $\mathrm{Im}\log Z$ are both zero. We see that $\mathcal{Z}$ shares just this property as $\log\mathcal{Z}(E)$ has the form

$$\log\mathcal{Z}(E) = -\sum_p \sum_{m=1}^{\infty} \frac{\exp(imS_p/\hbar)}{m\sqrt{|\det(M_p^m - 1)|}}. \tag{5.1.3}$$

Since $M_p^m$ is slowly varying compared to the exponential in the numerator (semiclassically $S_p(E) >> \hbar$), upon an average over $E$ the real and imaginary parts of $\log\mathcal{Z}$ will oscillate to zero. Thus $\mathcal{Z}$ is the function upon which we will concentrate, rather than $\Delta(E)$.

First of all our aim is to compare the value distribution of the real and imaginary parts of $\log\mathcal{Z}$ with the random matrix distributions. The system which we choose for this task is the kicked top (see for example [HKS87]). This is a three-dimensional angular momentum, the square of which is conserved, $\mathbf{J}^2 = j(j+1)$. This fixes the dimension of the Hilbert space as $N = 2j+1$, where $j$ is a positive integer or half-integer. A simple kicked top might be a spin precessing around the $y$ axis while being kicked periodically. We consider the $N \times N$ unitary time evolution operator

$$U = e^{-i\frac{\tau_z}{2j+1}J_z^2}e^{-i\alpha_y J_y}, \tag{5.1.4}$$

which allows the spin to rotate by an angle $\alpha_y$ around the $y$ axis before being kicked round the $z$ axis through an angle proportional to $J_z$ with a kicking strength $\tau_z$. Subsequent applications of $U$, then, give us a stroboscopic view of the evolution of a quantum state, at points in time separated by the kick period. The behaviour of this system is chaotic in the classical limit ($j \to \infty$) when $\tau_z$ is not too small (see [HKS87] for details).

The evolution operator which we will actually perform calculations with is

$$U = e^{-i\frac{\tau_z}{2j+1}J_z^2 - i\alpha_z J_z}e^{-i\alpha_y J_y}e^{-i\frac{\tau_x}{2j+1}J_x^2 - i\alpha_x J_x}, \tag{5.1.5}$$

but it follows the same principle as (5.1.4). This top has no time-reversal symmetry if the torsion parameters $\tau$ and the rotation angles $\alpha$ are of order one and non-zero, but if we were to set $\tau_x = \alpha_x = 0$, we would obtain a time-reversal symmetric top; the eigenvalues of $U$ would then be expected to obey COE statistics. As it is, we will use the eigenvalues of (5.1.5) to construct a spectral determinant, the value distribution of which we will compare with the CUE results of Chapter 2.

The reason that the kicked top is so appropriate for our task is that the $\tau$ and $\alpha$ parameters can be varied to obtain a whole family of tops with CUE eigenvalue statistics. In order to create smooth value distributions for matrices of dimension $N \times N$ where $N$ is of a reasonable size for computation, we take the average of the value distributions of the spectral determinants of many such tops. This is in analogy with Odlyzko's numerical computations which lead to Figures 3.1 and 3.2. He averaged the Riemann zeta function over a range on the critical line which covered at least $10^6$ zeros. This can be considered as equivalent to averaging over thousands of sets of $N = \log T$ zeros to obtain a smooth value distribution for the zeta function at height $T$ up the critical line. Thus we set $\alpha_z = 1$, $\alpha_y = 1$, and let $10 \leq \tau_z \leq 15$, $4 \leq \tau_x \leq 7$ and $1.1 \leq \alpha_x \leq 1.3$. These parameters were varied enough between the various kicked tops being averaged over that the sets of eigenvalues produced appeared to be independent from each other.

We note here that with reference to (2.7.10) it is possible that with large enough matrix size the value distribution of the logarithm of the spectral determinant of one single kicked top might show good agreement with random matrix theory. However, it is probable in the case of the value distribution of the modulus of the spectral determinant that some averaging over more than one system would be necessary even for large matrix size, as (2.7.10) indicates a strong likelihood that the value distribution for one kicked top, for example, will not lie close to the ensemble average.

Computing the distribution of values of

$$\mathrm{Re}\log\prod_{n=1}^{N}(1-e^{i(\theta_n-\theta)}) \tag{5.1.6}$$

as $\theta$ varies from 0 to $2\pi$, where the $e^{i\theta_n}$ are the eigenvalues of the matrix (5.1.5), for 800 sets of parameter values results in an average distribution which is shown in Figure 5.1. It is compared with the value distribution of $\mathrm{Re}\log Z$, (2.3.3). Both curves have been normalized to have mean zero and unit variance.

If the same process is performed for the imaginary part of the spectral determinant, we obtain Figure 5.2.

Both Figure 5.1 and 5.2 display the same excellent agreement with the random matrix distribution of the real and imaginary parts of the logarithm

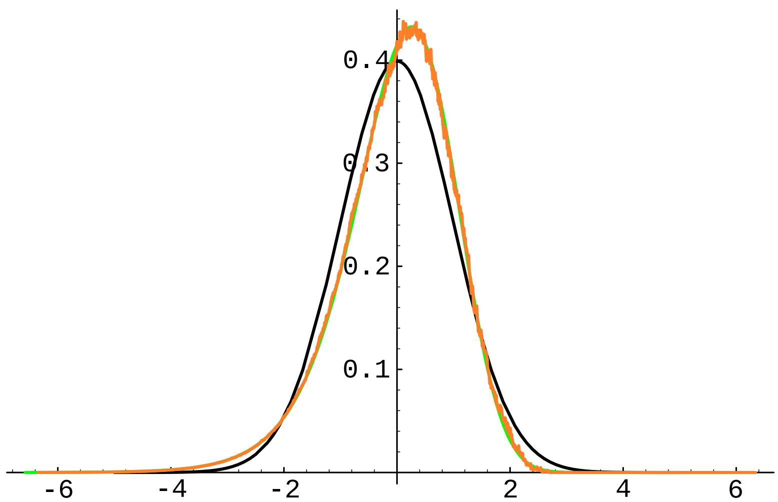


Figure 5.1: The distribution (orange) of the real part of the log of the spectral determinant for 800 values of the torsion parameters and rotation angles for kicked tops of dimension $N = 21$ ($j = 10$) compared with the equivalent CUE distribution also with $N = 21$ (green) and a Gaussian (black).

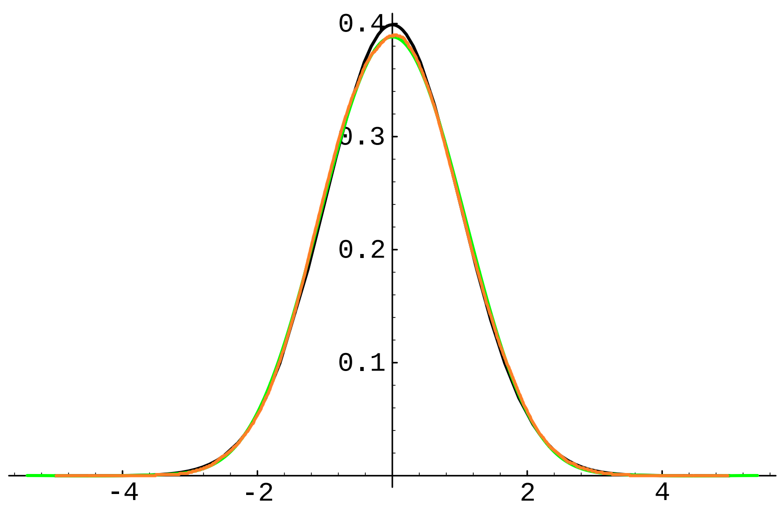


Figure 5.2: The distribution (orange) of the imaginary part of the log of the spectral determinant for 800 values of the torsion parameters and rotation angles for kicked tops of dimension $N = 21$ ($j = 10$) compared with the equivalent CUE distribution also with $N = 21$ (green) and a Gaussian (black).

of $Z(U,\theta)$.

We can also consider the average value distribution of

$$\prod_{n=1}^{N} |1 - e^{i(\theta_n - \theta)}|, \tag{5.1.7}$$

where again the $\theta_n$ are the eigenphases of the time evolution matrix of the kicked top (5.1.5) for a given set of parameters. The result is compared with the value distribution of $|Z|$ averaged over the CUE ensemble (3.4.13) in Figure 5.3. The first part of the figure shows the good general agreement between the kicked top data and the CUE calculation over a large range of values, but if we compare the distributions near zero, as in the second part of the figure, we see a clear disagreement which is reminiscent of Figure 3.5 where the value distribution of $|\zeta(1/2+it)|$ near zero contained a non-universal contribution from the product over primes (3.4.3) evaluated at the pole of $M_N(is)$. This discrepancy of the value distributions at zero for the Riemann zeta function is predicted in (3.4.19).

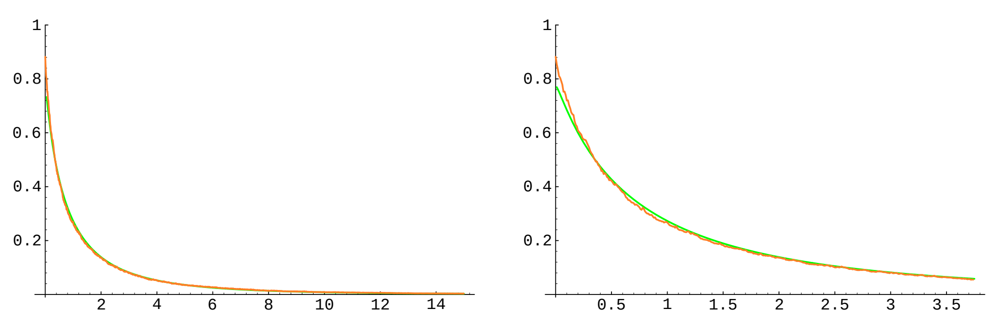


Figure 5.3: The distribution of the modulus of the spectral determinant for 800 values of the torsion parameters and rotation angles for kicked tops (orange) of dimension $N = 21$ ($j = 10$), compared to the CUE distribution, also for $N = 21$ (green).

This suggests that there is a non-universal contribution to the value distribution near zero of the modulus of the spectral determinant for the kicked tops. It should be noted that in an average over the value distributions of many diverse maps it is likely that a deviation from random matrix theory such as that seen in Figure 5.3 would disappear as all the non-universal contributions would average to zero, but the tops used here are not representative of all possible maps; their classical behaviour is too similar.

## 5.2 Moments of $|\mathcal{Z}(E)|$

To investigate this non-universal contribution encountered in the previous section we will return to the dynamical zeta function (5.1.2) and restrict ourselves to flows with two degrees of freedom and two-dimensional maps. Thus the monodromy matrix $M_p$ has dimension two and for hyperbolic systems the two eigenvalues are $\exp(\pm\lambda_p T_p)$, where $\lambda_p$ is the instability exponent of the primitive orbit $p$. Following [BK90] we see that

$$\begin{aligned}\frac{1}{\sqrt{|\det(M_p^m - 1)|}} &= \frac{1}{\exp(\frac{1}{2}m\lambda_p T_p)[1-\exp(-m\lambda_p T_p)]} \\ &= \exp(-\frac{1}{2}m\lambda_p T_p)\sum_{k=0}^{\infty}\exp(-mk\lambda_p T_p), \qquad (5.2.1)\end{aligned}$$

allowing us to write the dynamical zeta function as

$$\begin{aligned}\mathcal{Z}(E) &= \prod_p \exp\left(-\sum_{k=0}^{\infty}\sum_{m=1}^{\infty}\frac{\exp(-m(k+\frac{1}{2})\lambda_p T_p)\exp(imS_p/\hbar)}{m}\right) \\ &= \prod_p\prod_{k=0}^{\infty}\left(1-\exp(-(k+\frac{1}{2})\lambda_p T_p)\exp(iS_p/\hbar)\right). \qquad (5.2.2)\end{aligned}$$

Turning now to the $2\lambda$th moment of $|\mathcal{Z}|$, we examine

$$\begin{aligned}
&\langle \mathcal{Z}(E)^\lambda \mathcal{Z}^*(E)^\lambda \rangle \\
&= \langle \prod_p \prod_{k=0}^{\infty} [1 - \exp(-(k+\frac{1}{2})\lambda_p T_p) \exp(iS_p/\hbar)]^\lambda \\
&\quad \times [1 - \exp(-(k+\frac{1}{2})\lambda_p T_p) \exp(-iS_p/\hbar)]^\lambda \rangle \\
&= \langle \prod_p \prod_{k=0}^{\infty} \left( \sum_{m=0}^{\infty} \frac{\Gamma(\lambda+1)(-1)^m}{\Gamma(\lambda-m+1)m!} \exp(-m(k+\frac{1}{2})\lambda_p T_p + imS_p/\hbar) \right) \\
&\quad \times \left( \sum_{n=0}^{\infty} \frac{\Gamma(\lambda+1)(-1)^n}{\Gamma(\lambda-n+1)n!} \exp(-n(k+\frac{1}{2})\lambda_p T_p - inS_p/\hbar) \right) \rangle. \qquad (5.2.3)
\end{aligned}$$

We now assume the periodic orbit actions are uncorrelated, and thus the system has no symmetries, so that we can bring the average inside the orbit product. Expanding the product over $k$ and the sum over $m$ we obtain

$$\begin{aligned}
&\langle \mathcal{Z}(E)^\lambda \mathcal{Z}^*(E)^\lambda \rangle \qquad (5.2.4) \\
&= \prod_p \langle \left( 1 - \lambda \exp(-\frac{1}{2}\lambda_p T_p + iS_p/\hbar) + \cdots \right) \\
&\quad \times \left( 1 - \lambda \exp(-\frac{1}{2}\lambda_p T_p - iS_p/\hbar) + \cdots \right) \\
&\times \left( 1 - \lambda \exp(-\frac{3}{2}\lambda_p T_p + iS_p/\hbar) + \cdots \right) \\
&\quad \times \left( 1 - \lambda \exp(-\frac{3}{2}\lambda_p T_p - iS_p/\hbar) + \cdots \right) \\
&\times \cdots \rangle.
\end{aligned}$$

From this expression we see that the divergent terms (those containing a factor $\exp(-\beta\lambda_p T_p)$ with $\beta \leq 1$, as we shall verify below) arise from the $k=0$ factor in the $k$ product. These terms simplify because any exponential with a non-zero complex argument will disappear upon performing the average. The convergent terms are more complicated, but as far as the divergent terms are concerned, upon applying the average over $E$ these reduce to

$$\langle \mathcal{Z}(E)^{\lambda} \mathcal{Z}^*(E)^{\lambda} \rangle \approx \prod_p \left(1 + \lambda^2 \exp(-\lambda_p T_p) + \cdots \right), \tag{5.2.5}$$

where all the later terms will converge under the product over periodic orbits $p$.

The divergence of the expression (5.2.3) means it is not very useful for calculating moments of the spectral determinant. So, to tame the divergence, we will make a comparison with the second moment, given by

$$\begin{aligned} &\langle \mathcal{Z}(E) \mathcal{Z}^*(E) \rangle \\ &= \langle \prod_p \prod_{k=0}^{\infty} \left(1 - \exp(-(k+\frac{1}{2})\lambda_p T_p) \exp(iS_p/\hbar)\right) \\ &\quad \times \left(1 - \exp(-(k+\frac{1}{2})\lambda_p T_p) \exp(-iS_p/\hbar)\right) \rangle \\ &\approx \prod_p \left(1 + \exp(-\lambda_p T_p) + \cdots \right). \end{aligned} \tag{5.2.6}$$

Once again the divergent term arises from the $k = 0$ factor. The more complicated higher terms represented by the $\cdots$ above are all convergent.

A comparison of (5.2.5) and (5.2.6) indicates that

$$\mu_\lambda = \frac{\langle \prod_p \prod_{k=0}^{\infty} \left| \sum_{m=0}^{\infty} \frac{\Gamma(\lambda+1)(-1)^m}{\Gamma(\lambda-m+1)m!} \exp(-m(k+\frac{1}{2})\lambda_p T_p + imS_p/\hbar) \right|^2 \rangle}{\langle \prod_p \prod_{k=0}^{\infty} \left|1 - \exp(-(k+\frac{1}{2})\lambda_p T_p) \exp(iS_p/\hbar)\right|^2 \rangle^{\lambda^2}} \tag{5.2.7}$$

is a convergent quantity. This leads us to conjecture that the moments of the spectral determinant have the form

$$\frac{\langle \mathcal{Z}(E)^{\lambda} \mathcal{Z}^*(E)^{\lambda} \rangle}{\langle \mathcal{Z}(E) \mathcal{Z}^*(E) \rangle^{\lambda^2}} = \mu_\lambda \frac{M_N(\beta, 2\lambda)}{(M_N(\beta, 2))^{\lambda^2}}, \tag{5.2.8}$$

where the product over the periodic orbits $\mu_\lambda$ supplies the system-specific information, whereas the universal component comes directly from the random matrix moments $M_N(\beta, 2\lambda)$, (2.8.2), with the value of $\beta$ chosen for the matrix ensemble having the same symmetries as the system under consideration.

For the purposes of computation, if the instability exponents $\lambda_p$ are large, the product over $k$ in (5.2.7) will be dominated by the $k = 0$ factor. In such a case a diagonal approximation can be made, simplifying things considerably, and $\mu_\lambda$ is expected to be well approximated by

$$\mu_\lambda \approx \prod_p \frac{\left(\sum_{m=0}^{\infty}\left(\frac{\Gamma(\lambda+1)}{\Gamma(\lambda-m+1)m!}\right)^2 \exp(-m\lambda_p T_p)\right)^{\overline{g}}}{(1+\exp(-\lambda_p T_p))^{\overline{g}\lambda^2}}, \tag{5.2.9}$$

where we generalize to systems with symmetries, so $\overline{g}$ is the average multiplicity of the periodic orbit actions. For example, a system without time-reversal symmetry is likely to have very few orbits with the same action, making $\overline{g} = 1$. On the other hand, for a system with time-reversal symmetry most orbits, the exceptions being those which are self-retracing, come paired with a time-reverse orbit which has the same period and action, and so $\overline{g} = 2$.

Now we will consider just the second moment of $|\mathcal{Z}|$ and check whether its rate of divergence is in keeping with the result expected from the random matrix results from Section 2.8

$$M_N(\beta, 2\lambda) \propto N^{\frac{2}{\beta}\lambda^2} \tag{5.2.10}$$

for large $N$, remembering that $\beta = 1, 2$ and 4 for the COE, CUE and CSE respectively.

Making the approximation, valid for large $\lambda_p$, that the $k = 0$ factor in (5.2.6) dominates, we have

$$\langle \mathcal{Z}(E)\mathcal{Z}^*(E)\rangle \approx \prod_p (1+\exp(-\lambda_p T_p))^{\overline{g}}. \tag{5.2.11}$$

As we have said, this product diverges, so we must truncate it at a suitable point. We remember from Section 1.3 that an orbit of length $T_p$ contributes to correlations in the eigenvalue positions on a scale of about $h/T_p$. If we want to limit our product over periodic orbits to those which affect eigenvalue correlations on a scale longer than the mean spacing, $1/\overline{d}$, then we need to truncate the product at $T_p \leq h\overline{d}$. Thus, in analogy with (3.4.34), we write

$$\begin{aligned}\langle \mathcal{Z}(E)\mathcal{Z}^*(E)\rangle &\approx \prod_{T_p \leq h\overline{d}} (1+\exp(-\lambda_p T_p))^{\overline{g}} \\ &= \exp\left(\overline{g} \sum_{T_p \leq h\overline{d}} \log(1+\exp(-\lambda_p T_p))\right). \qquad (5.2.12)\end{aligned}$$

The sum rule of Hannay and Ozorio de Almeida (1.3.3) implies that the number of periodic orbits in a chaotic system with period less than or equal to $T$ is of order $\exp(h_t T)/(h_t T)$ for large $T$. Here $h_t$ is the topological entropy, which is equal to the average $\lambda_p$ for strongly chaotic systems. For large $\lambda_p$ we continue from (5.2.12) by approximating $\log(1+\exp(-\lambda_p T_p))$ with $\exp(-\lambda_p T_p)$, then we turn the sum into an integral using the density of orbits of length about $T$, $\exp(h_t T)/T$:

$$\begin{aligned}\langle \mathcal{Z}(E)\mathcal{Z}^*(E)\rangle &\approx \exp\left(\overline{g}\int_{T \leq h\overline{d}} e^{-\lambda T}\frac{e^{h_t T}}{T}dT\right) \\ &\approx \exp\left(\overline{g}\int_{T \leq h\overline{d}} \frac{dT}{T}\right) \\ &\propto (h\overline{d})^{\overline{g}}. \qquad (5.2.13)\end{aligned}$$

As in the Riemann zeta function case, there is ambiguity about the lower limit of the integral, but equating $\overline{d}$ with the density of random matrix eigenvalues $N/(2\pi)$, gives us a second moment of order $N^2$ for a system with time-reversal symmetry - agreeing perfectly with the COE result - and a second moment of order $N$ for the CUE, again exactly what we expect. The same calculation can be made for the higher moments, yielding a result equivalent to (5.2.10).

# Chapter 6

# Finale

In answer to the question posed at the beginning of Chapter 1, the results of the intervening pages suggest that in an appropriately chosen limit, which corresponds to large matrix size in the random matrix case, the correctly normalized value distribution of the logarithm of a function with random matrix distributed zeros depends only on those zero statistics; the features specific to the individual function are pushed to the extremes of the value distribution in this limit. On the other hand, the moments $|Z|^s$ and $(Z/Z^*)^s$ of the characteristic polynomial $Z(U, \theta)$ of an $N \times N$ matrix $U$ belonging to an ensemble of random matrices do not on their own predict the equivalent mean values of a particular function displaying the zero statistics of that ensemble. Instead the mean values of $Z$, once again in the limit of large $N$, must be multiplied by a contribution specific to the function being studied. Thus the moments of a function with random matrix zeros divide, in this limit, into two separate factors, one universal and the other specific to the given function. The conclusions of each of the cases studied are detailed below.

For the Riemann zeta function, comparison with random matrix calculations led to a conjecture that the Riemann moments $\frac{1}{T}\int_0^T |\zeta(1/2+it)|^{2\lambda} dt$ take the form, in the limit of large $T$, of a product over the prime numbers (a contribution specific to the Riemann zeta function) multiplied by a factor which is equivalent to the ensemble mean value of $|Z|^{2\lambda}$ as $N$ becomes large.

This conjecture is supported by the known Riemann moments. It also implies that, for large $T$, the distribution of values of $\zeta(1/2 + iT)$ and its logarithm will agree over the major part of the range of values with the corresponding value distributions of $Z$, as the contributions from the prime numbers will be pushed to the extremes of the distribution as $T \to \infty$.

Recent results of Katz and Sarnak [KS99b, KS99a] suggest that the low-lying zeros of families of $L$-functions display the statistics of various ensembles of random matrices, so we compared mean values over these families of $L$-functions, evaluated at the central point $s = 1/2$, with those of $Z(U, 0)$ averaged over the appropriate ensembles. Conrey and Farmer [CF99] have conjectured that $L$-function mean values factor into a product over primes which is specific to the family of $L$-functions multiplied by a component which they expected to be determined by the random matrix distribution of the zeros. We find that the asymptotic results for the integer moments of $Z(U, 0)$ agree precisely with the few known values of this factor expected by Conrey and Farmer to be determined by random matrix theory.

Finally, in the case of spectral determinants of classically chaotic systems we find, as expected, that the value distribution of the spectral determinant and its logarithm show very close agreement with random matrix theory as the system-specific characteristics are pushed to the extremes of the distribution in the semiclassical limit. We also conjecture, in a similar manner to the case of the Riemann zeta function, that in the semiclassical limit mean values of the modulus of the spectral determinant divide into a component containing specific information about the periodic orbits of the system, and a contribution which amounts to the leading order (in $N$) random matrix moments of $|Z|$.

# Appendix A

# The second log moment

The method embarked upon here becomes unfeasibly cumbersome for even the third moment, but it involves a technique which is often very useful for these types of calculations. Therefore it is worth stating here, even though a more tidy method was subsequently developed to deal with this particular problem, see Section 2.2. We start with the expansion of the logarithm of $Z(U,\theta)$,

$$\log Z(U,\theta) = -\sum_{p=1}^{N}\sum_{m=1}^{\infty}\frac{1}{m}e^{i(\theta_p-\theta)m}. \tag{A.1}$$

If we consider the real part of this expression, and take the second moment of that, then

$$\begin{aligned}\langle(\log|Z(U,\theta)|)^2\rangle_{CUE} &= \int_0^{2\pi}\cdots\int_0^{2\pi} d\theta_1\cdots d\theta_N \frac{1}{(2\pi)^N}\det\left(e^{i(j-k)\theta_j}\right)\\ &\times\left(\sum_{p,q=1}^{N}\sum_{m_1,m_2=1}^{\infty}\frac{1}{m_1m_2}\cos[(\theta_p-\theta)m_1]\cos[(\theta_q-\theta)m_2]\right).\end{aligned} \tag{A.2}$$

Here the joint probability density function for the CUE eigenphases is expressed in a different form from (1.2.13). This determinantal form is allowed because the function being averaged over the ensemble is symmetric amongst all the eigenphases (as elucidated very clearly in [Haa90]). Interchanging the order of the integration and the summation over the $m$'s,

$$\langle(\log|Z(U,\theta)|)^2\rangle_{CUE} = \sum_{m_1,m_2=1}^{\infty} \frac{1}{m_1 m_2} \frac{1}{(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} d\theta_1 \cdots d\theta_N \times \quad \text{(A.3)}$$

$$\left[ \sum_{p=q=1}^{N} \det \begin{pmatrix} 1 & e^{-i\theta_1} & \cdots & e^{-i(N-1)\theta_1} \\ e^{i\theta_2} & 1 & \cdots & e^{-i(N-2)\theta_2} \\ \vdots & \vdots & \cdots & \vdots \\ C_p^{(1)}\,C_p^{(2)} e^{i(p-1)\theta_p} & C_p^{(1)}\,C_p^{(2)} e^{i(p-2)\theta_p} & \cdots & C_p^{(1)}\,C_p^{(2)} e^{i(p-N)\theta_p} \\ \vdots & \vdots & \cdots & \vdots \\ e^{i(N-1)\theta_N} & e^{i(N-2)\theta_N} & \cdots & 1 \end{pmatrix} \right.$$

$$\left. + \sum_{p\neq q} \det \begin{pmatrix} 1 & e^{-i\theta_1} & \cdots & e^{-i(N-1)\theta_1} \\ e^{i\theta_2} & 1 & \cdots & e^{-i(N-2)\theta_2} \\ \vdots & \vdots & \cdots & \vdots \\ C_p^{(1)} e^{i(p-1)\theta_p} & C_p^{(1)} e^{i(p-2)\theta_p} & \cdots & C_p^{(1)} e^{i(p-N)\theta_p} \\ \vdots & \vdots & \cdots & \vdots \\ C_q^{(2)} e^{i(q-1)\theta_q} & C_q^{(2)} e^{i(q-2)\theta_q} & \cdots & C_q^{(2)} e^{i(q-N)\theta_q} \\ \vdots & \vdots & \cdots & \vdots \\ e^{i(N-1)\theta_N} & e^{i(N-2)\theta_N} & \cdots & 1 \end{pmatrix} \right].$$

In the above formula, we are using the notation $C_x^{(i)} = \cos[(\theta_x - \theta)m_i]$, where $x = p, q$ and $i = 1, 2$. The elements containing the cosine functions are in the $p^{th}$ row in the first matrix, and in the $p^{th}$ and the $q^{th}$ rows in the second matrix.

The next step is to multiply the $j^{th}$ column of each matrix by $e^{ij\theta}$ and the $k^{th}$ row of each matrix by $e^{-ik\theta}$. As all of these factors multiply to one, this does not affect the determinant. As the variable $\theta_k$ appears only in the $k^{th}$ row of the matrix, the integration can be brought inside the determinant, and

$$\langle(\log|Z|)^2\rangle_{CUE} = \sum_{m_1,m_2=1}^{\infty} \frac{1}{m_1 m_2} \frac{1}{(2\pi)^N} \times \quad \text{(A.4)}$$

$$\left[\sum_{p=1}^{N}\det\begin{pmatrix} 2\pi & 0 & \cdots \\ 0 & 2\pi & \cdots \\ \vdots & \vdots & \cdots \\ \int_0^{2\pi} C_p^{(1)}\, C_p^{(2)} e^{i(p-1)(\theta_p-\theta)}d\theta_p & \int_0^{2\pi} C_p^{(1)}\, C_p^{(2)} e^{i(p-2)(\theta_p-\theta)}d\theta_p & \cdots \\ \vdots & \vdots & \cdots \\ 0 & 0 & \cdots \end{pmatrix}\right.$$

$$\left.+\sum_{p\neq q}\det\begin{pmatrix} 2\pi & 0 & \cdots \\ 0 & 2\pi & \cdots \\ \vdots & \vdots & \cdots \\ \int_0^{2\pi} C_p^{(1)} e^{i(p-1)(\theta_p-\theta)}d\theta_p & \int_0^{2\pi} C_p^{(1)} e^{i(p-2)(\theta_p-\theta)}d\theta_p & \cdots \\ \vdots & \vdots & \cdots \\ \int_0^{2\pi} C_q^{(2)} e^{i(q-1)(\theta_q-\theta)}d\theta_q & \int_0^{2\pi} C_q^{(2)} e^{i(q-2)(\theta_q-\theta)}d\theta_q & \cdots \\ \vdots & \vdots & \cdots \\ 0 & 0 & \cdots \end{pmatrix}\right].$$

After performing the integrals over the cosine functions (which imposes constraints on the indices being summed over), the first matrix can be diagonalized. We are left with

$$\langle(\log|Z|)^2\rangle_{CUE} = \sum_{m=1}^{\infty}\sum_{p=1}^{N}\frac{1}{2m^2} + \sum_{m_1,m_2=1}^{\infty}\sum_{p\neq q}\frac{1}{m_1 m_2} \tag{A.5}$$

$$\times\det\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots \\ \frac{\delta(m_1-p+1)}{2} & \cdots & \frac{\delta(m_1-1)}{2} & 0 & \frac{\delta(-m_1+1)}{2} & \frac{\delta(-m_1+2)}{2} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \cdots & \cdots \\ \frac{\delta(m_2-q+1)}{2} & \cdots & \cdots & \cdots & \frac{\delta(m_2-1)}{2} & 0 & \frac{\delta(-m_2+1)}{2} & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

In the above matrix, at most two of the elements in the $p^{th}$ row and two in the $q^{th}$ row will be non-zero in any one term of the multiple sum over the $m$'s and $p$ and $q$. If such a non-zero element does not occur in both the $(p,q)^{th}$ *and* in the $(q,p)^{th}$ elements, then one or both of the $p^{th}$ and $q^{th}$ rows can be turned into a row of zeros by adding appropriate multiples of the identity matrix rows to it. This imposes the condition that $m_1 = m_2 = |p-q|$. Therefore, (A.5) reduces to

$$\langle (\log|Z|)^2 \rangle_{CUE} = \frac{N}{2}\sum_{m=1}^{\infty}\frac{1}{m^2} + \sum_{p\neq q}\frac{1}{(p-q)^2}\det\begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & \frac{1}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\ 0 & 0 & \cdots & \frac{1}{2} & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{A.6}$$

If we interchange the $p^{th}$ and $q^{th}$ rows, then we acquire a negative sign, and the matrix becomes diagonal.

$$\begin{aligned}\langle (\log|Z|)^2 \rangle_{CUE} &= \frac{N\pi^2}{12} - \sum_{n=1}^{N-1}\frac{2(N-n)}{n^2}\left(\frac{1}{2}\cdot\frac{1}{2}\right) \\ &= \frac{N\pi^2}{12} - \frac{1}{2}\sum_{n=1}^{N-1}\frac{N-n}{n^2} \\ &= \frac{1}{2}\sum_{n=1}^{N-1}\frac{1}{n} + \frac{N}{2}\sum_{n=N}^{\infty}\frac{1}{n^2}. \end{aligned} \tag{A.7}$$

We are interested in the large $N$ asymptotics of this expression, so to deal with the first sum we substitute the first few terms of

$$\sum_{k=1}^{n}\frac{1}{k} = \gamma + \log n + \frac{1}{2n} - \sum_{k=2}^{\infty}\frac{A_k}{n(n+1)\cdots(n+k-1)}, \tag{A.8}$$

where $A_k = \frac{1}{k}\int_0^1 x(1-x)(2-x)(3-x)\cdots(k-1-x)dx$, and apply the Euler-Maclaurin formula to the second sum. The result is

$$\langle(\log|Z|)^2\rangle_{CUE} = \frac{1}{2}\log N + \frac{1}{2}(\gamma+1) + \frac{1}{24N^2} - \frac{1}{80N^4} + O\left(\frac{1}{N^6}\right), \qquad \text{(A.9)}$$

although more terms could be retained if necessary.

# Appendix B

# Determining the relation between the fluctuating part of the staircase function and the spectral determinant function

We are going to make use of the following theorem [Mar65]:

$$\frac{1}{2\pi i}\int_L \frac{d}{dz}\mathrm{Ln}f(z)dz = \frac{1}{2\pi i}\int_L \frac{f'(z)}{f(z)}dz = N - P, \tag{B.1}$$

where $L$ is a closed, rectifiable Jordan curve, $N$ is the number of zeros of $f(z)$ (counted a number of times equal to the order) enclosed in $L$ and $P$ is the number of enclosed poles counted in the same manner.

The function to be considered will be

$$\mathcal{Z}(s) = \prod_{p=1}^{N}(1 - \frac{e^{i\theta_p}}{s}), \tag{B.2}$$

where the product is over $N$ CUE eigenphases, (this is just the spectral determinant, $Z(U,\theta) = \prod_{p=1}^{N}(1-e^{i(\theta_p-\theta)})$ extended to the whole complex plane) and we note that it has the property

$$\begin{aligned}
\mathcal{Z}^*(s) &= \prod_{p=1}^{N}(1-\frac{e^{-i\theta_p}}{s^*}) \\
&= \left(\frac{-1}{s^*}\right)^N e^{-i\sum_{p=1}^{N}\theta_p}\prod_{p=1}^{N}(1-e^{i\theta_p}s^*) \\
&= \left(\frac{-1}{s^*}\right)^N e^{-i\sum_{p=1}^{N}\theta_p}\mathcal{Z}(1/s^*). \qquad \text{(B.3)}
\end{aligned}$$

This means that

$$\mathcal{Z}(s) = (-1/s)^N e^{i\sum_p \theta_p}\mathcal{Z}^*(1/s^*). \qquad \text{(B.4)}$$

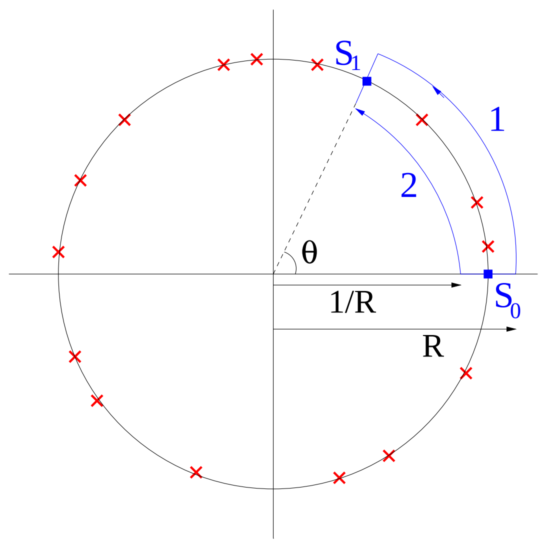


Figure B.1: A sketch illustrating the eigenvalues (red crosses) of a CUE matrix, along with the path of integration (in blue).

In the statement (B.1) of the theorem above, let $f(z)$ be replaced by $\mathcal{Z}(s)$. $L$ is the blue curve in Figure B.1; it runs counter-clockwise from $s_o$ all the way round to $s_o$ again (ie. curve 1 - curve 2). $\mathcal{Z}(s)$ has no poles, and its zeros are on the unit circle at $\theta_1$, $\theta_2$, ..., $\theta_N$. Therefore, the right-hand side of (B.1) is just $N(\theta)$, the number of zeros less than $\theta$. This is just the staircase function for the $\theta_p$'s. We will assume that $\theta_j \neq 0$ and $\theta_j \neq \theta$ for any $j$.

So, applying the theorem, we have

$$\begin{aligned} N(\theta) &= \frac{1}{2\pi i}\int_L \frac{d}{ds}\mathrm{Ln}(\mathcal{Z}(s))ds \\ &= \frac{1}{2\pi i}\left(\int_1 \frac{d}{ds}\mathrm{Ln}(\mathcal{Z}(s))ds - \int_2 \frac{d}{ds}\mathrm{Ln}(\mathcal{Z}(s))ds\right) \\ &= \frac{1}{2\pi i}\left(\int_1 \frac{d}{ds}\mathrm{Ln}(\mathcal{Z}(s))ds - \int_2 \frac{d}{ds}\mathrm{Ln}\left((-1/s)^N e^{i\sum_p \theta_p}\mathcal{Z}^*(1/s^*)\right)ds\right). \end{aligned} \tag{B.5}$$

We now change the variables in the second integral so that $t = 1/s^*$. If we write $t = Re^{i\phi}$, then $(1/t^*) = (1/R)e^{i\phi}$, so curve 2 is transformed into curve 1.

$$\begin{aligned} N(\theta) &= \frac{1}{2\pi i}\left(\int_1 \frac{d}{ds}\mathrm{Ln}(\mathcal{Z}(s))ds - \int_1 \frac{d}{dt}\mathrm{Ln}\left((-t^*)^N e^{i\sum_p \theta_p}\mathcal{Z}^*(t)\right)dt\right) \\ &= \frac{1}{2\pi i}\left([\mathrm{Ln}(\mathcal{Z}(s))]_{s_o}^{s_1} - \left[\mathrm{Ln}\left((-t^*)^N e^{i\sum_p \theta_p}\mathcal{Z}^*(t)\right)\right]_{s_o}^{s_1}\right) \\ &= \frac{1}{2\pi i}\Bigg([\mathrm{Ln}(\mathcal{Z}(s))]_{s_o}^{s_1} - N\,[\mathrm{Ln}(-1)]_{s_o}^{s_1} - N\,[\mathrm{Ln}(t^*)]_{s_o}^{s_1} \\ &\qquad -i\left[\sum_p \theta_p\right]_{s_o}^{s_1} - [\mathrm{Ln}\,(\mathcal{Z}^*(t))]_{s_o}^{s_1}\Bigg) \\ &= \frac{1}{2\pi i}\left(2i\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(s_1))] - 2i\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(s_0))] - N[\mathrm{Ln}(t^*)]_{s_o}^{s_1}\right). \end{aligned} \tag{B.6}$$

Since $s_o = 1$ and $s_1 = e^{i\theta}$,

$$\begin{aligned} N(\theta) &= \frac{1}{2\pi i}\left(2i\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(e^{i\theta}))] + N\mathrm{Ln}(e^{i\theta}) - 2i\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(1))]\right) \\ &= \frac{1}{\pi}\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(e^{i\theta}))] + \frac{N\theta}{2\pi} - \frac{1}{\pi}\mathrm{Im}[\mathrm{Ln}(\mathcal{Z}(1))]. \end{aligned} \tag{B.7}$$

As stated in [BKP98], and easily confirmed by integrating $N(\theta) - N\theta/(2\pi)$ over the integral $0 < \theta < 2\pi$, the fluctuating part (ie. with mean zero) of the staircase function is given by

$$\tilde{N}(\theta) \equiv N(\theta) - \frac{N\theta}{2\pi} + \frac{1}{\pi}\mathrm{Im}[\mathrm{Ln}(\prod_p (1 - e^{i\theta_p}))]. \tag{B.8}$$

Therefore, from (B.7) and (B.8), the fluctuating part of the staircase is given by

$$\tilde{N}(\theta) = \frac{1}{\pi}\mathrm{Im}[\mathrm{Ln}(Z(U,\theta))], \tag{B.9}$$

where $Z(U,\theta)$ is given by

$$Z(U,\theta) = \prod_{p=1}^{N} \left(1 - e^{(i\theta_p - \theta)}\right). \tag{B.10}$$

Thus in studying the imaginary part of the logarithm of $Z(U,\theta)$, we are also investigating the much studied spectral staircase function.

# Appendix C

# A heuristic study of Riemann log moments

## C.1 The second moment of $\operatorname{Re}\log\zeta$

Before starting we note that in the following we will be making use of the Euler product formula for the Riemann zeta function on the critical line where it fails to converge. The results of this misdemeanour, however, are eye-opening, and we will therefore proceed.

To begin, then,

$$\begin{aligned}
\operatorname{Re}\log\zeta(1/2+it) &= \operatorname{Re}\log\prod_p\left(1-\frac{1}{p^{1/2+it}}\right)^{-1} \\
&= -\operatorname{Re}\sum_p\log\left(1-\frac{1}{p^{1/2+it}}\right) \\
&= \operatorname{Re}\sum_p\sum_{m=1}^{\infty}\frac{1}{m\sqrt{p^m}}e^{itm\log p} \\
&= \sum_p\sum_{m=1}^{\infty}\frac{1}{m\sqrt{p^m}}\cos(tm\log p).
\end{aligned} \tag{C.1}$$

So, the second moment looks like

$$
\begin{aligned}
&\langle (\operatorname{Re}\log\zeta(1/2+it))^2\rangle \\
&\quad = \sum_p \sum_q \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{mn\sqrt{p^m q^n}} \langle \cos(tm\log p)\cos(tn\log q)\rangle \\
&\quad = \frac{1}{2}\sum_p \sum_q \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{mn\sqrt{p^m q^n}} \langle \cos(t[m\log p + n\log q]) \\
&\qquad\qquad + \cos(t[m\log p - n\log q])\rangle. \qquad \text{(C.2)}
\end{aligned}
$$

The first of the cosine terms can be neglected because $p$ and $q$ cannot be chosen so as to make $m\log p + n\log q$ vanish, implying that this term will always be oscillatory and will therefore average to zero.

There are now several ways to proceed. One of these is to utilize the similarity between the above expression and the form factor for the Riemann zeros, scaled so that there is a mean spacing of unity between the zeros. We saw in Chapter 1 that this is the Fourier transform of the two-point correlation function $\langle (d(t+\frac{y}{2\overline{d}})-1)(d(t-\frac{y}{2\overline{d}})-1)/\overline{d}^2\rangle_t$, where $\overline{d}\sim\frac{1}{2\pi}\log\frac{t}{2\pi}$. So we have [Kea93], as $t\to\infty$,

$$
\begin{aligned}
\tilde{K}_R(T) &\equiv K_R(\tau) \qquad \text{(C.3)} \\
&\sim \frac{1}{\log(t/2\pi)} \sum_p \sum_q \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{\log p\log q}{\sqrt{p^m q^n}} \langle \cos\left(t\log\left(\frac{p^m}{q^n}\right)\right)\rangle \\
&\qquad \times\delta(T - \frac{1}{2}\log(p^m q^n)),
\end{aligned}
$$

where $p$ and $q$ are summing over prime numbers and $T=\tau\log(t/(2\pi))$.

From this we see that

$$
\begin{aligned}
&\log(t/2\pi)\int_0^{\infty} \frac{\tilde{K}_R(T)}{T^2}dT \\
&\quad = \sum_p \sum_q \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{\log p\log q}{m^2(\log p)^2\sqrt{p^m q^n}} \left\langle \cos(t\log\left(\frac{p^m}{q^n}\right)\right\rangle \\
&\quad = 2\langle (\operatorname{Re}\log\zeta(1/2+it))^2\rangle. \qquad \text{(C.4)}
\end{aligned}
$$

This last line is a good approximation because the average in the angled brackets will only be non-zero when $p^m \approx q^n$.

Now we apply the Berry Ansatz for the form factor, equivalent to (1.3.2), which is:

$$\tilde{K}_R(T) = \begin{cases} \begin{cases} \frac{T}{\log(t/2\pi)} & T \le \log(t/2\pi) \\ 1 & T > \log(t/2\pi) \end{cases} & \text{if} \quad T > T^* \\ \frac{1}{\log(t/2\pi)} \sum_p \sum_{m=1}^{\infty} \frac{(\log p)^2}{p^m} \delta(T - m \log p) & \text{if} \quad T \le T^*. \end{cases} \tag{C.5}$$

That is, the form factor of the Riemann zeros follows that of GUE distributed zeros for $T$ greater than some $T^* < \log(t/2\pi)$, whereas below that value the individual delta spikes are significant. Thus we have

$$\begin{aligned} &\log(t/2\pi) \int_0^\infty \frac{\tilde{K}_R(T)}{T^2} dT \\ &= \sum_p \sum_{m=1}^{\infty} \frac{(\log p)^2}{p^m} \int_0^{T^*} \frac{\delta(T - m \log p)}{T^2} dT \\ &\quad + \int_{T^*}^{\log(t/2\pi)} \frac{dT}{T} + \log(t/2\pi) \int_{\log(t/2\pi)}^{\infty} \frac{dT}{T^2} \\ &= \sum_{\substack{p \\ \log p^m < T^*}} \sum_{m=1}^{\infty} \frac{1}{m^2 p^m} + \log\log(t/2\pi) - \log T^* + 1. \end{aligned} \tag{C.6}$$

A trick we will use again later on to deal with the sum above is to write

$$\begin{aligned} \sum_{p<P} \sum_{m=1}^{\infty} \frac{1}{m^2 p^m} &= \sum_{p<P} \sum_{m=1}^{\infty} \frac{1}{m p^m} - \sum_{p<P} \sum_{m=1}^{\infty} \frac{m-1}{m^2} \frac{1}{p^m} \\ &= -\log \prod_p (1 - 1/p) - \sum_{p<P} \sum_{m=1}^{\infty} \frac{m-1}{m^2} \frac{1}{p^m}. \end{aligned} \tag{C.7}$$

In Titchmarsh [Tit86] we find that for large $P$,

$$\log \prod_{p<P} (1 - 1/p)^{-1} \sim \log(e^\gamma \log P) = \gamma + \log\log P, \tag{C.8}$$

where $\gamma$ is Euler's constant, so

$$\sum_{p<P}\sum_{m=1}^{\infty}\frac{1}{m^2p^m}=\gamma+\log\log P-\sum_{p<P}\sum_{m=1}^{\infty}\frac{m-1}{m^2}\frac{1}{p^m}. \tag{C.9}$$

Thus, in (C.6) (converting the condition $\log p^m < T^*$ into $p < e^{T^*}$ because the $m = 1$ terms contribute most to the sum)

$$\begin{aligned}
&\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle\\
&=\frac{1}{2}\log(t/2\pi)\int_0^{\infty}\frac{\tilde{K}_R(T)}{T^2}dT \qquad (C.10)\\
&=\frac{1}{2}\log\log(t/2\pi)+\frac{1}{2}(\gamma+1)+\frac{1}{2}\sum_{p}\sum_{m=1}^{\infty}\left(\frac{-1}{m}+\frac{1}{m^2}\right)\frac{1}{p^m}.
\end{aligned}$$

With the identification $N = \log T$, this supports the result of calculating the second moment of $\log|Z|$ (3.6.1) via the prime-modified generating function, see Section 3.6.

Method number two for calculating the second moment is the method which we will continue to use for the higher moments. We begin again with

$$\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle=\frac{1}{2}\sum_{p}\sum_{q}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{e^{-\epsilon\log(p^mq^n)}}{mn\sqrt{p^mq^n}}\left\langle\cos\left(t\log\left(\frac{p^m}{q^n}\right)\right)\right\rangle, \tag{C.11}$$

where we have replaced $t$ with $t+i\epsilon$ and so have shifted ourselves just slightly off the critical line, thereby missing the zeros as we perform the average. In the end we will take the limit $\epsilon \to 0$.

We expect that there will be a contribution to the log moments which is specific to the Riemann zeta function (containing sums over primes) as well as the universal, random matrix contribution. As was mentioned in Chapter 1, it is expected that the low primes will contribute to the non-universal part, so we split the sum at primes around $t/2\pi$ so that the primes below the cut-off will contribute to correlations between zeros on a scale larger than the mean spacing of zeros. Therefore we write

$$\begin{aligned}
&\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle \\
&\quad=\frac{1}{2}\sum_{p<t/2\pi}\sum_{q<t/2\pi}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{mn\sqrt{p^m q^n}}\langle\cos(t\log(p^m/q^n))\rangle \qquad \text{(C.12)}\\
&\qquad+\frac{1}{2}\sum_{\text{other } p,q}\sum\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{mn\sqrt{p^m q^n}}e^{-\epsilon\log(p^m q^n)}\langle\cos(t\log(p^m/q^n))\rangle \\
&\quad=\frac{1}{2}\sum_{p<t/2\pi}\sum_{m=1}^{\infty}\frac{1}{m^2p^m}+\frac{1}{2}\sum_{\text{other } p,q}\sum\frac{1}{\sqrt{pq}}e^{-\epsilon\log(pq)}\langle\cos(t\log(p/q))\rangle.
\end{aligned}$$

Above we have made the assumption that the contribution from the low primes is mostly diagonal and the $n$ and $m$ sum can be dropped from the high prime sums because squared primes are less dense than primes, so the likelihood of finding a pair of squared primes close together in a certain range is less likely than finding a close pair of primes. The relevance of this is that unless $p^m \approx q^n$, $\cos(t\log(p^m/q^n))$ will average to zero.

Continuing from (C.12), the method we apply to the second sum is that described in [BK95] and we will see that it introduces a step function which effectively selects only large primes. This is why it is being applied to the sum containing high primes. To the sum over lower primes, we apply (C.9).

$$\begin{aligned}
&\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle \\
&\quad=\frac{1}{2}\log\log(t/(2\pi))+\frac{1}{2}\gamma-\frac{1}{2}\sum_{p}\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \qquad \text{(C.13)}\\
&\qquad+\frac{1}{2}\sum_{p\geq t/(2\pi)}\sum_{h\neq 0}e^{-2\epsilon\log p}\frac{1}{p}\cos(t\log((p+h)/p))P(1,1,h),
\end{aligned}$$

where we are using the notation of [BK95] in that $P(m,n,k)$ is the probability that $q=(mp-k)/n$ is a prime if $p$ is one, and $k$ has no common factors with $m$, $p$ or $n$: $(k,m)=(k,n)=(k,p)=1$. We expect contributions to the average in (C.12) only from $h << p$ because larger $h$ terms will be washed out in the average. Therefore, we can write $\log((p+h)/p)\sim h/p$. This leads us, following Bogomolny and Keating [BK95], to

$$\begin{aligned}
&\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \\
&\qquad+\frac{1}{2}\mathrm{Re}\sum_{p\geq t/(2\pi)}\frac{1}{p}\sum_{h\neq 0}e^{-2\epsilon\log p}\exp(ith/p)P(1,1,h) \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \\
&\qquad+\frac{1}{2}\mathrm{Re}\sum_{p\geq t/(2\pi)}\frac{1}{p\log p}\log(t/p)\Theta(\log(p/t))e^{-2\epsilon\log p} \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \\
&\qquad+\frac{1}{2}\mathrm{Re}\int_{t/(2\pi)}^{\infty}e^{-2\epsilon\log p}\frac{\log t-\log p}{p\log^2 p}\Theta(\log(p/t))dp \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \\
&\qquad+\frac{1}{2}\mathrm{Re}\int_{\log t}^{\infty}e^{-2\epsilon\alpha}\frac{\log t}{\alpha^2}d\alpha - \frac{1}{2}\mathrm{Re}\int_{\log t}^{\infty}\frac{d\alpha}{\alpha},
\end{aligned}\tag{C.14}$$

where $\Theta$ is the unit step function.

Noting that the integral of $1/\alpha$ disappears in the limit as $t\to\infty$,

$$\begin{aligned}
&\langle(\mathrm{Re}\log\zeta(1/2+it))^2\rangle \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} \\
&\qquad+\frac{1}{2}\left[-e^{-2\epsilon\alpha}\frac{\log t}{\alpha}\Big|_{\log(t/2\pi)}^{\infty} - O(\epsilon)\right] \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}\gamma - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m} + \frac{1}{2} \\
&\quad\sim \frac{1}{2}\log\log(t/(2\pi)) + \frac{1}{2}(\gamma+1) - \frac{1}{2}\sum_p\sum_{m=1}^{\infty}\frac{m-1}{m^2p^m}.
\end{aligned}\tag{C.15}$$

Again, this is the desired answer.

## C.2 The third moment of $\operatorname{Re}\log\zeta$

We now study the third moment of the real part of the log of the Riemann zeta function.

$$\begin{aligned}
&\langle(\operatorname{Re}\log\zeta(1/2+it))^3\rangle \\
&= \sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\frac{1}{m_1m_2m_3\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}}} \\
&\times\langle\operatorname{Re}(e^{itm_1\log p_1})\operatorname{Re}(e^{itm_2\log p_1})\operatorname{Re}(e^{itm_3\log p_3})\rangle. \qquad \text{(C.16)}
\end{aligned}$$

However, instead of making $t$ real, let $t = t + i\epsilon$, where $\epsilon$ is a small positive quantity, so that we are running just beside the zeros along the critical line instead of through them. At the end we will take the limit $\epsilon \to 0$.

$$\begin{aligned}
&\langle(\operatorname{Re}\log\zeta(1/2+it))^3\rangle \\
&= \sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\frac{e^{-\epsilon\log(p_1^{m_1}p_2^{m_2}p_3^{m_3})}}{m_1m_2m_3\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}}} \\
&\times\langle\cos(tm_1\log p_1)\cos(tm_2\log p_2)\cos(tm_3\log p_3)\rangle \\
&= \frac{1}{2}\sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\frac{e^{-\epsilon\log(p_1^{m_1}p_2^{m_2}p_3^{m_3})}}{m_1m_2m_3\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}}} \\
&\langle(\cos[t\log(p_1^{m_1}p_2^{m_2})] + \cos[t\log(p_1^{m_1}/p_2^{m_2})])\cos(tm_3\log p_3)\rangle \\
&= \frac{1}{4}\sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\frac{e^{-\epsilon\log(p_1^{m_1}p_2^{m_2}p_3^{m_3})}}{m_1m_2m_3\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}}} \\
&\times\langle\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3})] + \cos[t\log(p_1^{m_1}p_2^{m_2}/p_3^{m_3})] \\
&+\cos[t\log(p_1^{m_1}p_3^{m_3}/p_2^{m_2})] + \cos[t\log(p_1^{m_1}/(p_2^{m_2}p_3^{m_3}))]\rangle \\
&= \frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\frac{e^{-\epsilon\log(p_1^{m_1}p_2^{m_2}p_3^{m_3})}}{m_1m_2m_3\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}}} \\
&\times\langle\cos[t\log(p_1^{m_1}/(p_2^{m_2}p_3^{m_3}))]\rangle. \qquad \text{(C.17)}
\end{aligned}$$

The last line of (C.17) is due to the symmetry amongst the prime sums, and to the fact that the first of the four cosines will average to zero as its argument will never be close to zero.

Following Bogomolny and Keating [BK95], we retain only the $m_1 = m_2 = m_3 = 1$ terms as those corresponding to higher $m$'s are generally smaller due to the factors $1/\sqrt{p^m}$.

Thus

$$\langle(\mathrm{Re}\log\zeta(1/2+it))^3\rangle \approx \frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{p_3}\frac{e^{-\epsilon\log(p_1p_2p_3)}}{\sqrt{p_1p_2p_3}}\langle\cos[t\log p_1/(p_2p_3)]\rangle. \tag{C.18}$$

The biggest contribution from the average will occur when $p_1 \approx p_2p_3$. Let $p_1 = p_2p_3+k$, where $k << p_2p_3$. In the notation of [BK95], $(k,p_2) = (k,p_3) = 1$ and we want the probability as $k$ varies that $p_1$ is a prime. Again from [BK95], this is $P(p_2,1,k) = P(p_3,1,k)$.

So, now we have (making the approximation $p_1 \approx p_2p_3$ in the slowly varying factors)

$$\begin{aligned}&\langle(\mathrm{Re}\log\zeta(1/2+it))^3\rangle\\ &\quad\approx \frac{3}{4}\sum_{p_2}\sum_{p_3}\frac{e^{-\epsilon\log(p_1p_2p_3)}}{\sqrt{p_2p_3p_2p_3}}\mathrm{Re}\sum_{\substack{k\neq 0\\(k,p_2)=(k,p_3)=1}}\exp\left(\frac{itk}{p_2p_3}\right)P(p_2,1,k).\end{aligned}$$

We perform this sum, in the manner of Bogomolny and Keating, by first evaluating the sum over all non-zero $k$, then subtracting off those terms for which $k$ is a multiple of $p_2$ or $p_3$. Using Bogomolny and Keating's approximations, this results in,

$$\begin{aligned}
&\langle (\mathrm{Re}\log\zeta(1/2+it))^3\rangle \\
&\approx \frac{3}{4}\mathrm{Re}\sum_{p_2}\sum_{p_3}\frac{e^{-2\epsilon\log(p_2p_3)}}{p_2p_3\log(p_2p_3)}\left(\log\left(\frac{t}{p_2p_3}\right)\Theta\left(\log\left(\frac{p_2p_3}{t}\right)\right)\right. \\
&\quad \left.-\log\left(\frac{t}{p_3}\right)\Theta\left(\log\left(\frac{p_3}{t}\right)\right)-\log\left(\frac{t}{p_2}\right)\Theta\left(\log\left(\frac{p_2}{t}\right)\right)\right) \qquad \text{(C.19)} \\
&\approx \frac{3}{4}\mathrm{Re}\int_1^\infty\int_1^\infty\frac{e^{-2\epsilon\log(p_2p_3)}}{p_2p_3\log(p_2p_3)\log p_2\log p_3}((\log t-\log p_2-\log p_3) \\
&\quad \times\Theta(\log p_2+\log p_3-\log t)-(\log t-\log p_3)\Theta(\log p_3-\log t) \\
&\quad -(\log t-\log p_2)\Theta(\log p_2-\log t))dp_2dp_3 \\
&\approx \frac{3}{4}\mathrm{Re}\int_0^\infty\int_0^\infty\frac{e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}((\log t-\alpha_2-\alpha_3)\Theta(\alpha_2+\alpha_3-\log t) \\
&\quad -(\log t-\alpha_3)\Theta(\alpha_3-\log t)-(\log t-\alpha_2)\Theta(\alpha_2-\log t))d\alpha_2d\alpha_3.
\end{aligned}$$

Now, because this is symmetric between the two $\alpha$'s, we integrate over just half of the positive quadrant, $\alpha_3<\alpha_2$, and then divide the integral up to look at each of the step functions separately.

$$\begin{aligned}
&\langle (\mathrm{Re}\log\zeta(1/2+it))^3\rangle \\
&= \frac{3}{2}\int_0^\infty\int_0^{\alpha_2}\frac{e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}((\log t-\alpha_2-\alpha_3)\Theta(\alpha_2+\alpha_3-\log t) \\
&\quad -(\log t-\alpha_3)\Theta(\alpha_3-\log t)-(\log t-\alpha_2)\Theta(\alpha_2-\log t))d\alpha_3d\alpha_2 \\
&= \frac{3}{2}\int_{(\log t)/2}^{\log t}\int_{\log t-\alpha_2}^{\alpha_2}\frac{(\log t-\alpha_2-\alpha_3)e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}d\alpha_3d\alpha_2 \\
&\quad +\frac{3}{2}\int_{\log t}^\infty\int_0^{\alpha_2}\frac{(\log t-\alpha_2-\alpha_3)e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}d\alpha_3d\alpha_2 \\
&\quad -\frac{3}{2}\int_{\log t}^\infty\int_0^{\alpha_2}\frac{(\log t-\alpha_2)e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}d\alpha_3d\alpha_2 \\
&\quad -\frac{3}{2}\int_{\log t}^\infty\int_{\log t}^{\alpha_2}\frac{(\log t-\alpha_3)e^{-2\epsilon(\alpha_2+\alpha_3)}}{(\alpha_2+\alpha_3)\alpha_2\alpha_3}d\alpha_3d\alpha_2 \\
&= -\frac{3}{2}\frac{\pi^2}{6}. \qquad \text{(C.20)}
\end{aligned}$$

In the above, the singularities in the various integrals cancel each other out, and we have neglected integrals of order $\epsilon$. (C.20) is now exactly the constant, leading-order term $-\frac{3}{4}\zeta(2)\Gamma(3)$ in the moment $\langle(\log|Z|)^3\rangle_{CUE}$.

In the above calculation, however, we have lost the contributions from the small primes. These lower primes, in analogy with the short periodic orbits of a classically chaotic system, constitute the basis for non-universal terms, and for the small primes it is not so clear that only the $m_1 = m_2 = m_3 = 1$ terms of (C.17) are significant. However, we still want terms where $p_1^{m_1} p_2^{m_2}/p_3^{m_3} \approx 1$, so we'll consider the case when $p_1 = p_2 = p_3$ and $m_1 + m_2 = m_3$. Then we have

$$\begin{aligned}
&\frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_3 m_2 (m_3 - m_2)\sqrt{p^{2m_3}}} \\
&= \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1 - m_2/m_3}{m_2 m_3^2 (1 - m_2/m_3) p^{m_3}} \\
&\quad + \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{m_2/m_3}{m_2 m_3 (m_3 - m_2) p^{m_3}} \\
&= \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_2 m_3^2 p^{m_3}} + \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_3^2 (m_3 - m_2) p^{m_3}} \\
&= \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_2 m_3^2 p^{m_3}} + \frac{3}{4}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_3^2 m_2 p^{m_3}} \\
&= \frac{3}{2}\sum_p \sum_{m_3=2}^{\infty}\sum_{m_2=1}^{m_3-1} \frac{1}{m_2 m_3^2 p^{m_3}} = \frac{3}{2}\sum_p \sum_{m_3=2}^{\infty} (\psi(m) + \gamma)\frac{1}{m_3^2 p_3^m}. \qquad \text{(C.21)}
\end{aligned}$$

This is exactly the non-universal contribution seen in (3.6.1) which resulted from applying the conjecture (3.4.8) on the moments of $|\zeta(1/2 + it)|$ to the question of calculating the log moments in Section 3.6.

## C.3 The fourth moment of $\mathrm{Re}\log\zeta$

We treat the fourth moment in much the same way, although we will restrict ourselves to hunting out the non-universal sums over primes. It is absolutely expected that the Bogomolny-Keating technique would reproduce the random matrix result ((3.6.1) without the quantities in curly brackets), but the effort necessary to do so would be immense.

We begin, as usual, by expanding the moment to a sum of cosine functions:

$$\begin{aligned}
&\langle(\operatorname{Re}\log\zeta(1/2+it))^4 \\
&\quad=\sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{p_4}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\sum_{m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle\operatorname{Re}\exp(itm_1\log p_1)\operatorname{Re}\exp(itm_2\log p_2) \\
&\qquad\times\operatorname{Re}\exp(itm_3\log p_3)\operatorname{Re}\exp(itm_4\log p_4)\rangle \\
&=\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle\cos(tm_1\log p_1)\cos(tm_2\log p_2)\cos(tm_3\log p_3)\cos(tm_4\log p_4)\rangle \\
&=\frac{1}{4}\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle(\cos[t\log(p_1^{m_1}p_2^{m_2})]+\cos[t\log(p_1^{m_1}/p_2^{m_2})]) \\
&\qquad\times(\cos[t\log(p_3^{m_3}p_4^{m_4})]+\cos[t\log(p_3^{m_3}/p_4^{m_4})])\rangle \\
&=\frac{1}{8}\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4})]+\cos[t\log(p_1^{m_1}p_2^{m_2}/(p_3^{m_3}p_4^{m_4}))] \\
&\qquad+\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3}/p_4^{m_4})]+\cos[t\log(p_1^{m_1}p_2^{m_2}p_4^{m_4}/p_3^{m_3})] \\
&\qquad+\cos[t\log(p_1^{m_1}p_3^{m_3}p_4^{m_4}/p_2^{m_2})]+\cos[t\log(p_1^{m_1}/(p_2^{m_2}p_3^{m_3}p_4^{m_4}))] \\
&\qquad+\cos[t\log(p_1^{m_1}p_3^{m_3}/(p_2^{m_2}p_4^{m_4}))]+\cos[t\log(p_1^{m_1}p_4^{m_4}/(p_2^{m_2}p_3^{m_3}))]\rangle \\
&=\frac{1}{8}\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \qquad \text{(C.22)} \\
&\qquad\times\langle 3\cos[t\log(p_1^{m_1}p_2^{m_2}/(p_3^{m_3}p_4^{m_4}))]+4\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3}/p_4^{m_4})]\rangle
\end{aligned}$$

where the last line holds due to the symmetry of the sums and the fact that the first cosine in the second to last line averages to zero.

We can see immediately that there will be a significant contribution to the above average when either $p_1^{m_1}p_2^{m_3}=p_3^{m_3}p_4^{m_4}$, in the case of the first cosine, or $p_1^{m_1}p_2^{m_2}p_3^{m_3}=p_4^{m_4}$ from the second cosine.

As the $p$'s are primes, the only way to obtain the second of these situations is to have $p_1 = p_2 = p_3 = p_4$ and $m_1 + m_2 + m_3 = m_4$. So we look at

$$\begin{aligned} T_1 &= \frac{1}{2} \sum_{\substack{p_1,p_2,p_3,p_4 \\ p_1=p_2=p_3=p_4}} \sum_{\substack{m_1,m_2,m_3,m_4=1 \\ m_1+m_2+m_3=m_4}}^{\infty} \frac{1}{m_1 m_2 m_3 m_4 \sqrt{p_1^{m_1} p_2^{m_2} p_3^{m_3} p_4^{m_4}}} \\ &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{1}{m_1(n-m_1)(m_4-n)m_4} \right) \frac{1}{p^{m_4}}. \end{aligned} \tag{C.23}$$

In the above rearrangement, $n = m_1 + m_2$. We now perform partial fraction decomposition twice in a row, then apply the fact that once $m_4$ and $n$ are fixed, a sum over $m_1$ is equivalent to a sum over $n - m_1$ if $m_1$ is summed from 1 to $n-1$.

$$\begin{aligned} T_1 &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{1}{m_1 n(m_4-n)m_4} + \frac{1}{n(n-m_1)(m_4-n)m_4} \right) \frac{1}{p^{m_4}} \\ &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{1}{m_1 n m_4^2} + \frac{1}{m_1(m_4-n)m_4^2} + \frac{1}{n(n-m_1)m_4^2} \right. \\ &\qquad \left. + \frac{1}{(n-m_1)(m_4-n)m_4^2} \right) \frac{1}{p^{m_4}} \\ &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{2}{m_1 n} + \frac{2}{(n-m_1)(m_4-n)} \right) \frac{1}{m_4^2 p^{m_4}} \\ &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{2}{m_1 n} + \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{2}{(n-m_1)(m_4-m_1)} \right. \\ &\qquad \left. + \frac{2}{(m_4-n)(m_4-m_1)} \right) \frac{1}{m_4^2 p^{m_4}} \\ &= \frac{1}{2} \sum_p \sum_{m_4=3}^{\infty} \left( \left( \sum_{n=2}^{m_4-1} \sum_{m_1=1}^{n-1} \frac{1}{m_1 n} + \sum_{m_1=1}^{m_4-2} \sum_{n=m_1+1}^{m_4-1} \frac{1}{m_1 n} \right) \right. \\ &\qquad \left. + \left( \sum_{n=2}^{m_4-1} \sum_{k=1}^{n-1} \frac{2}{k(m_4-n+k)} + \sum_{j=1}^{m_4-2} \sum_{m_1=1}^{m_4-j-1} \frac{2}{j(m_4-m_1)} \right) \right) \frac{1}{m_4^2 p^{m_4}} \end{aligned}$$

$$\begin{aligned} T_1 &= \frac{1}{2}\sum_p \sum_{m_4=3}^{\infty}\left(\left(\sum_{n=2}^{m_4-1}\sum_{m_1=1}^{n-1}\frac{1}{m_1 n}+\sum_{n=1}^{m_4-2}\sum_{m_1=n+1}^{m_4-1}\frac{1}{m_1 n}\right)\right. \\ &\qquad \left.+\left(\sum_{k=1}^{m_4-2}\sum_{n=k+1}^{m_4-1}\frac{2}{k(m_4-n+k)}+\sum_{j=1}^{m_4-2}\sum_{k=j+1}^{m_4-1}\frac{2}{jk}\right)\right)\frac{1}{m_4^2 p^{m_4}} \\ &= \frac{1}{2}\sum_p \sum_{m_4=3}^{\infty}\left(\left(\sum_{n=1}^{m_4-1}\sum_{m_1=1}^{m_4-1}\frac{1}{m_1 n}-\sum_{n=1}^{m_4-1}\frac{1}{n^2}\right)\right. \\ &\qquad \left.+\left(\sum_{k=1}^{m_4-2}\sum_{j=k+1}^{m_4-1}\frac{2}{jk}+\sum_{k=2}^{m_4-1}\sum_{j=1}^{k-1}\frac{2}{jk}\right)\right)\frac{1}{m_4^2 p^{m_4}} \\ &= \frac{3}{2}\sum_p \sum_{m_4=3}^{\infty}\left(\sum_{n=1}^{m_4-1}\sum_{m_1=1}^{m_4-1}\frac{1}{m_1 n}-\sum_{n=1}^{m_4-1}\frac{1}{n^2}\right)\frac{1}{m_4^2 p^{m_4}} \\ &= \frac{3}{2}\left(\sum_p \sum_{m=2}^{\infty}\left((\psi(m)-\psi(1))^2+\psi^{(1)}(m)-\psi^{(1)}(1)\right)\frac{1}{m^2 p^m}\right). \end{aligned} \tag{C.24}$$

The second contribution from (C.22) is a little more complicated because, as for the second moment, the sums do not converge. We are considering the sum

$$T_2 = \frac{3}{8}\sum_{p_1,p_2,p_3,p_4}\ \sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1 m_2 m_3 m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}}, \tag{C.25}$$

where the restrictions are that

$$\begin{array}{llllll} \text{either} & p_1=p_3 & p_2=p_4 & m_1=m_3 & m_2=m_4 & (a) \\ \text{or} & p_1=p_4 & p_2=p_3 & m_1=m_4 & m_2=m_3 & (b) \\ \text{or} & p_1=p_2 = p_3=p_4 & & m_1+m_2 = m_3+m_4 & & (c) \end{array}$$

Cases (a) and (b) are equivalent due to the symmetry of the sums, so we will include the contribution of (a) twice, but we must be careful not to overcount the $p_1 = p_2 = p_3 = p_4$ contributions. This caution also applies to case (c), which replicates the $p_1 = p_2 = p_3 = p_4$ contributions yet again. We will get around this problem by including *none* of these contributions in the sums arising from cases (a) and (b), as they are all covered by (c). Therefore we have the following, where it can be seen that, as in the case of the second moment, the sums are sometimes divergent.

$$\begin{aligned}
T_2 &= \frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\frac{1}{m_1^2 m_2^2 p_1^{m_1} p_2^{m_2}} - \frac{3}{4}\sum_{p}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\frac{1}{m_1^2 m_2^2 p^{m_1} p^{m_2}} \\
&\quad +\frac{3}{8}\sum_{p}\sum_{n=2}^{\infty}\sum_{m_1=1}^{n-1}\sum_{m_3=1}^{n-1}\frac{1}{m_1(n-m_1)(n-m_3)m_3 p^n} \\
&= \frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{m_1=2}^{\infty}\sum_{m_2=1}^{\infty}\frac{(1-m_1)(1-m_2)}{m_1^2 m_2^2 p_1^{m_1} p_2^{m_2}} \\
&\quad +\frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\frac{m_1(1-m_2)}{m_1^2 m_2^2 p_1^{m_1} p_2^{m_2}} \\
&\quad +\frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\frac{m_2(1-m_1)}{m_1^2 m_2^2 p_1^{m_1} p_2^{m_2}} \\
&\quad +\frac{3}{4}\sum_{p_1}\sum_{p_2}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\frac{1}{m_1 m_2 p_1^{m_1} p_2^{m_2}} \\
&\quad -\frac{3}{4}\sum_{p}\left(\sum_{m_1=1}^{\infty}\frac{1}{m_1^2 p^{m_1}}\right)\left(\sum_{m_2=1}^{\infty}\frac{1}{m_2^2 p^{m_2}}\right) \\
&\quad +\frac{3}{8}\sum_{p}\sum_{n=2}^{\infty}\sum_{m_1=1}^{n-1}\sum_{m_3=1}^{n-1}\left(\frac{1}{m_1 m_3}+\frac{1}{n(n-m_1)}\right) \\
&\qquad \times\left(\frac{1}{m_3 n}+\frac{1}{n(n-m_3)}\right)\frac{1}{p^n} \\
&= \frac{3}{4}\left(\sum_{p}\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right)^2 \\
&\quad +\frac{3}{2}\left(\sum_{p_1}\sum_{m_1=1}^{\infty}\frac{1}{m_1 p_1^{m_1}}\right)\left(\sum_{p_2}\sum_{m_2=1}^{\infty}\frac{1-m_2}{m_2^2 p_2^{m_2}}\right) \\
&\quad +\frac{3}{4}\left(\sum_{p}\sum_{m=1}^{\infty}\frac{1}{mp^m}\right)^2 - \frac{3}{4}\sum_{p}\left(\sum_{m_1=1}^{\infty}\frac{1}{m_1^2 p^{m_1}}\right)\left(\sum_{m_2=1}^{\infty}\frac{1}{m_2^2 p^{m_2}}\right) \\
&\quad +\frac{3}{8}\sum_{p}\sum_{n=2}^{\infty}\sum_{m_1=1}^{n-1}\sum_{m_3=1}^{n-1}\left(\frac{1}{m_1 m_3}+\frac{1}{m_1(n-m_3)}+\frac{1}{m_3(n-m_1)}\right. \\
&\quad \left.+\frac{1}{(n-m_1)(n-m_3)}\right)\frac{1}{n^2 p^n}. \qquad \text{(C.26)}
\end{aligned}$$

We have seen sums like those over $1/(mp^m)$ before. If we imagine truncating them at $p < t/(2\pi)$, then we have seen in Section C.1, on the second moment, that $\sum_p \sum_{m=1}^{\infty} 1/(mp^m) \simeq \log\log(t/(2\pi)) + \gamma$. Thus

$$\begin{aligned}
T_2 &= \frac{3}{4}\left(\sum_p \sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right)^2 + 3\left(\frac{1}{2}\log\log\left(\frac{t}{2\pi}\right)+\frac{1}{2}\gamma\right) \\
&\quad \left(\sum_p \sum_{m=1}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right) + 3\left(\frac{1}{2}\log\log\left(\frac{t}{2\pi}\right)+\frac{1}{2}\gamma\right)^2 \\
&\quad -\frac{3}{4}\sum_p\left(\sum_{m_1=1}^{\infty}\frac{1}{m_1^2 p^{m_1}}\right)\left(\sum_{m_2=1}^{\infty}\frac{1}{m_2^2 p^{m_2}}\right) \\
&\quad +\frac{3}{2}\sum_p\sum_{n=2}^{\infty}\left(\sum_{m_1=1}^{n-1}\sum_{m_3=1}^{n-1}\frac{1}{m_1 m_3}\right)\frac{1}{n^2 p^n} \\
&= 3\left(\frac{1}{2}\log\log\left(\frac{t}{2\pi}\right)+\frac{1}{2}\gamma+\frac{1}{2}\sum_p\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right)^2 \\
&\quad -\frac{3}{4}\sum_p\left(\sum_{m=1}^{\infty}\frac{1}{m^2 p^m}\right)^2 + \frac{3}{2}\sum_p\sum_{n=2}^{\infty}(\psi(n)-\psi(1))^2\frac{1}{n^2 p^n}. \qquad \text{(C.27)}
\end{aligned}$$

The total contribution to the fourth moment from the terms $T_1$ and $T_2$ is therefore

$$\begin{aligned}
T_1 + T_2 &= 3\left(\frac{1}{2}\log\log\left(\frac{t}{2\pi}\right)+\frac{1}{2}\gamma+\frac{1}{2}\sum_p\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right)^2 \\
&\quad -\frac{3}{4}\sum_p\left(\sum_{m=1}^{\infty}\frac{1}{m^2 p^m}\right)^2 \qquad \text{(C.28)} \\
&\quad +3\sum_p\sum_{n=2}^{\infty}\left((\psi(n)-\psi(1))^2+\frac{1}{2}\left(\psi^{(1)}(m)-\psi^{(1)}(1)\right)\right)\frac{1}{n^2 p^n}.
\end{aligned}$$

## C.4 The second moment of $\operatorname{Im}\log\zeta$

In an identical manner to the section on the second moment of the real part of the log, we deal with the imaginary second log moment.

$$
\begin{aligned}
&\langle (\operatorname{Im} \log \zeta(1/2 + it))^2 \rangle \\
&= \sum_p \sum_q \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{mn\sqrt{p^m q^n}} \langle \sin[tm \log p] \sin[tn \log q] \rangle \\
&= \frac{1}{2} \sum_p \sum_q \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{mn\sqrt{p^m q^n}} \langle \cos[t \log(p^m/p^n)] - \cos[t \log(p^m q^n)] \rangle \\
&= \frac{1}{2} \sum_p \sum_q \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{mn\sqrt{p^m q^n}} \langle \cos[t \log(p^m/q^n)] \rangle \\
&= \langle (\operatorname{Re} \log \zeta(1/2 + it))^2 \rangle. && \text{(C.29)}
\end{aligned}
$$

Therefore all the discussion for the second moment of the real part of the log holds equally well here.

## C.5 The third moment of $\operatorname{Im} \log \zeta$

We treat the third moment similarly:

$$
\begin{aligned}
&\langle (\operatorname{Im} \log \zeta(1/2 + it))^3 \rangle \\
&= \sum_{p_1} \sum_{p_2} \sum_{p_3} \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \sum_{m_3=1}^{\infty} \frac{1}{m_1 m_2 m_3 \sqrt{p_1^{m_1} p_2^{m_2} p_3^{m_3}}} \\
&\quad \times \langle \sin[tm_1 \log p_1] \sin[tm_2 \log p_2] \sin[tm_3 \log p_3] \rangle \\
&= \frac{1}{2} \sum_{p_1,p_2,p_3} \sum_{m_1,m_2,m_3=1}^{\infty} \frac{1}{m_1 m_2 m_3 \sqrt{p_1^{m_1} p_2^{m_2} p_3^{m_3}}} \\
&\quad \times \langle (\cos[t \log(p_1^{m_1}/p_2^{m_2})] - \cos[t \log(p_1^{m_1} p_2^{m_2})]) \sin[tm_3 \log p_3] \rangle \\
&= \frac{1}{4} \sum_{p_1,p_2,p_3} \sum_{m_1,m_2,m_3=1}^{\infty} \frac{1}{m_1 m_2 m_3 \sqrt{p_1^{m_1} p_2^{m_2} p_3^{m_3}}} \\
&\quad \times \langle \sin[t \log(p_2^{m_2} p_3^{m_3}/p_1^{m_1})] + \sin[t \log(p_1^{m_1} p_3^{m_3}/p_2^{m_2})] \\
&\quad - \sin[t \log(p_3^{m_3}/(p_1^{m_1} p_2^{m_2}))] - \sin[t \log(p_1^{m_1} p_2^{m_2} p_3^{m_3})] \rangle \\
&= \frac{3}{4} \sum_{p_1,p_2,p_3} \sum_{m_1,m_2,m_3=1}^{\infty} \frac{1}{m_1 m_2 m_3 \sqrt{p_1^{m_1} p_2^{m_2} p_3^{m_3}}} \langle \sin[t \log(p_1^{m_1} p_2^{m_2}/p_3^{m_3})] \rangle \\
&\sim 0. && \text{(C.30)}
\end{aligned}
$$

This is exactly the result supported by numerical evidence and the random matrix calculations.

## C.6 The fourth moment of $\operatorname{Im}\log\zeta$

Now the fourth moment:

$$\begin{aligned}
&\langle(\operatorname{Im}\log\zeta(1/2+it))^4\rangle \\
&\quad=\sum_{p_1}\sum_{p_2}\sum_{p_3}\sum_{p_4}\sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty}\sum_{m_3=1}^{\infty}\sum_{m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle\sin[t\log p_1^{m_1}]\sin[t\log p_2^{m_2}]\sin[t\log p_3^{m_3}]\sin[t\log p_4^{m_4}]\rangle \\
&\quad=\frac{1}{4}\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle(\cos[t\log(p_1^{m_1}/p_2^{m_2})]-\cos[t\log(p_1^{m_1}p_2^{m_2})]) \\
&\qquad(\cos[t\log(p_3^{m_3}/p_4^{m_4})]-\cos[t\log(p_3^{m_3}p_4^{m_4})])\rangle \\
&\quad=\frac{1}{8}\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \\
&\qquad\times\langle\cos[t\log(p_1^{m_1}p_4^{m_4}/(p_2^{m_2}p_3^{m_3}))]+\cos[t\log(p_1^{m_1}p_3^{m_3}/(p_2^{m_2}p_4^{m_4}))] \\
&\qquad-\cos[t\log(p_1^{m_1}/(p_2^{m_2}p_3^{m_3}p_4^{m_4}))]-\cos[t\log(p_1^{m_1}p_3^{m_3}p_4^{m_4}/p_2^{m_2})] \\
&\qquad-\cos[t\log(p_1^{m_1}p_2^{m_2}p_4^{m_4}/p_3^{m_3})]-\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3}/p_4^{m_4})] \\
&\qquad+\cos[t\log(p_1^{m_1}p_2^{m_2}/(p_3^{m_3}p_4^{m_4}))]+\cos[t\log(p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4})]\rangle \\
&\quad=\sum_{p_1,p_2,p_3,p_4}\sum_{m_1,m_2,m_3,m_4=1}^{\infty}\frac{1}{m_1m_2m_3m_4\sqrt{p_1^{m_1}p_2^{m_2}p_3^{m_3}p_4^{m_4}}} \qquad\qquad \text{(C.31)}\\
&\qquad\times\left\langle\frac{3}{8}\cos[t\log(p_1^{m_1}p_2^{m_2}/(p_3^{m_3}p_4^{m_4}))]-\frac{1}{2}\cos[t\log(p_1^{m_1}/(p_2^{m_2}p_3^{m_3}p_4^{m_4}))]\right\rangle.
\end{aligned}$$

We know all about the evaluation of the sums over these cosine averages from the fourth moment of the real part of the log, so we know that the contribution from the terms where the log factor in the argument of the cosines is zero is

$$\begin{aligned}
T_2-T_1 &= 3\left(\frac{1}{2}\log\log\left(\frac{t}{2\pi}\right)+\frac{1}{2}\gamma+\frac{1}{2}\sum_p\sum_{m=2}^{\infty}\left(\frac{1}{m^2}-\frac{1}{m}\right)\frac{1}{p^m}\right)^2 \qquad \text{(C.32)}\\
&\quad-\frac{3}{4}\sum_p\left(\sum_{m=1}^{\infty}\frac{1}{m^2p^m}\right)^2-\frac{3}{2}\sum_p\sum_{n=2}^{\infty}\left(\psi^{(1)}(m)-\psi^{(1)}(1)\right)\frac{1}{n^2p^n}.
\end{aligned}$$

This is expected to give the prime sum contributions to $\langle(\operatorname{Im}\log\zeta(1/2+it))^4\rangle$, and indeed it agrees precisely with the prime sums in (3.6.3).

# Appendix D

# Closing the contour of integration around the poles of $M_{Sp}(N,s)$ and $M_O(N,s)$

## D.1 $USp(2N)$

If we define a contour of integration as composed of three lines in the complex plane, $S_1$, $S_2$ and $S_3$, where $S_1$ runs from $(-n,0)$ to $(-n,in)$, $S_2$ from $(-n,in)$ to $(n,in)$ and $S_3$ from $(n,in)$ to $(n,0)$, then we want to show that

$$\lim_{n\to\infty}\int_{s_m(n)} e^{-iy\log x} M_{Sp}(N,iy)dy = 0 \tag{D.1}$$

for $m = 1,2,3$ and $x < 1$. Here $M_{Sp}(N,s)$ is the moment of the characteristic polynomial of matrices in $USp(2N)$,

$$M_{Sp}(N,s) = 2^{2Ns}\prod_{j=1}^{N}\frac{\Gamma(1+N+j)\Gamma(1/2+s+j)}{\Gamma(1/2+j)\Gamma(1+s+N+j)}. \tag{D.2}$$

Let us define, for positive integers $j$,

$$f(j,y)=\frac{\Gamma(1/2+iy+j)}{\Gamma(1+iy+N+j)}, \tag{D.3}$$

which has poles at half integers greater than $j$ on the positive imaginary axis, and zeros at integers on the imaginary axis greater than or equal to $1+N+j$.

We now have

$$M_{Sp}(N,iy)=\text{const}2^{2Niy}\prod_{j=1}^{N}f(j,y). \tag{D.4}$$

If we start with $S_3$ then we can make use of Stirling's formula,

$$\Gamma(z+j)\sim\sqrt{2\pi}e^{-z}z^{z+j-1/2}, \tag{D.5}$$

which holds as $|z|\to\infty$ and $|\arg z|<\pi$. We consider $y=n+it$ on $S_3$, so

$$\begin{aligned}|f(j,n+it)| &= \left|\frac{\Gamma(1/2+in-t+j)}{\Gamma(1+in-t+N+j)}\right| \\ &= \left|\frac{\Gamma(1/2+z+j)}{\Gamma(1+z+N+j)}\right|,\end{aligned} \tag{D.6}$$

where we have substituted $z=in-t$. As $t$ runs from $n$ to 0, $\arg z\in[\pi/2,3\pi/4]$ and so we can apply (D.5). For large $n$ the result is

$$\begin{aligned}|f(j,n+it)| &\sim \left|\frac{e^{-z}z^{z+j}}{e^{-z}z^{z+N+j+1/2}}\right| \\ &= \left|e^{-(N+1/2)\log z}\right| \\ &= \left|e^{-(N+1/2)\log(in-t)}\right| \\ &= \left|e^{-1/2(N+1/2)\log(n^2+t^2)}\right|.\end{aligned} \tag{D.7}$$

Since $N>0$, as $n\to\infty$, $|f(j,n+it)|\to 0$ uniformly for $n+it$ on $S_3$.

As

$$\begin{aligned} |2^{2Niy}| &= |2^{2Ni(n+it)}| \\ &= |2^{-2Nt}| \\ &\leq 1, \end{aligned} \tag{D.8}$$

$|M_{Sp}(N,iy)| \to 0$ uniformly as $n \to \infty$ on $S_3$. Thus we can choose any $\epsilon > 0$ and find an $n^*$ such that for every $n > n^*$ $|M_{Sp}(N,iy)| < \epsilon$ for all $y \in S_3(n)$. Therefore we have, for $\log x < 0$,

$$\begin{aligned} \Big|\int_{S_3(n)} e^{-iy\log x} M_{Sp}(N,iy)dy\Big| &< \epsilon\Big|\int_n^0 e^{-in\log x} e^{\log xt} i\ dt\Big| \\ &< \epsilon\int_0^n e^{\log xt} dt \\ &= \frac{\epsilon}{\log x} e^{t\log x}\Big|_0^n \\ &= \frac{\epsilon}{\log x}(e^{n\log x} - 1). \end{aligned} \tag{D.9}$$

Thus, since $\epsilon$ is arbitrarily small, the integral tends to zero on $S_3$ as $n \to \infty$.

Next we work on $S_1$. Here we want to examine

$$\begin{aligned} |f(j,-n+it)| &= \Big|\frac{\Gamma(1/2 - in - t + j)}{\Gamma(1 - in - t + N + j)}\Big| \\ &= \Big|\frac{\Gamma(1/2 + z + j)}{\Gamma(1 + z + N + j)}\Big|. \end{aligned} \tag{D.10}$$

This time $z = -in - t$, so as $t$ runs from 0 to $n$, $\arg z \in [-\pi/2, -3\pi/4]$ and we can use (D.5) to obtain

$$\begin{aligned} |f(j,-n+it)| &\sim \big|e^{-(N+1/2)\log z}\big| \\ &= \big|e^{-1/2(N+1/2)\log(n^2+t^2)}\big|. \end{aligned} \tag{D.11}$$

So $|f(j,-n+it)| \to 0$ uniformly for $-n+it$ on $S_1$ as $n \to \infty$. Also $|2^{2Niy}| \leq 1$ on $S_1$, so $|M_{Sp}(N,iy)| \to 0$ uniformly as $n \to \infty$ on $S_1$ and the argument follows just as for $S_3$ that (D.1) holds for $m = 1$.

The above was based on a similar calculation performed by Hughes (PhD thesis, Bristol University) for the function $M_N(s)$.

For $S_2$ the method is just slightly different since we cannot use Stirling's approximation along this line.

Instead we use [GR65]

$$\lim_{|z|\to\infty} \frac{\Gamma(z+a)}{\Gamma(z)} e^{-a\log z} = 1. \tag{D.12}$$

On $S_2$ we have $y = t + in$, $t \in [-n, n]$, so

$$\begin{aligned} |f(j,y)| &= \left|\frac{\Gamma(1/2+iy+j)}{\Gamma(1+iy+N+j)}\right| \\ &= \left|\frac{\Gamma(1/2+it-n+j)}{\Gamma(1+it-n+N+j)}\right| \\ &= \left|\frac{\Gamma(z)}{\Gamma(z+1/2+N)}\right|, \end{aligned} \tag{D.13}$$

where $z = 1/2 + it - n + j$. Therefore we have

$$\begin{aligned} \lim_{n\to\infty} \left|\frac{\Gamma(z)}{\Gamma(z+1/2+N)}\right| &= \lim_{n\to\infty} \left|\frac{1}{e^{(1/2+N)\log z}}\right| \\ &= \lim_{n\to\infty} \left|e^{-(1/2+N)\log(1/2+it-n+j)}\right| \\ &= \lim_{n\to\infty} \left|e^{-(1/2+N)\log[(1/2-n+j)^2+t^2]/2}\right| \\ &= 0. \end{aligned} \tag{D.14}$$

Thus $|M_{Sp}(N,iy)| \to 0$ uniformly as $n \to \infty$ on $S_2$ as here we again have $|2^{2Niy}| \le 1$. We can therefore find for each $\epsilon > 0$ an $n^*$ such that for every $n > n^*$, $|M_{Sp}(N,iy)| < \epsilon$ for all $y$ on $S_2$. So,

$$\begin{aligned} \left|\int_{S_2} e^{-iy\log x} M_{Sp}(N,iy)dy\right| &< \epsilon \int_{-n}^{n} \left|e^{-i(t+in)\log x}\right| dt \\ &\le \epsilon \int_{-n}^{n} e^{n\log x} dt \\ &= \epsilon(2n)e^{n\log x}. \end{aligned} \tag{D.15}$$

Since $\log x < 0$, (D.1) is true for $m = 2$.

## D.2 $SO(2N)$

The method here is identical to that in the previous section. First we remember that

$$M_O(N,s) = 2^{2Ns} \prod_{j=1}^{N} \frac{\Gamma(N+j-1)\Gamma(s+j-1/2)}{\Gamma(j-1/2)\Gamma(s+j+N-1)}, \qquad (D.16)$$

and so we define

$$f(j,y) = \frac{\Gamma(iy+j-1/2)}{\Gamma(iy+j+N-1)}, \qquad (D.17)$$

which has poles at the half integers on the imaginary axis at values greater than or equal to $i(j-1/2)$ and zeros at all integers on the positive imaginary axis above, and including, $i(j+N-1)$.

We have, therefore, that

$$M_O(N,iy) = \text{const} 2^{2Niy} \prod_{j=1}^{N} f(j,y), \qquad (D.18)$$

and we want to show that

$$\lim_{n\to\infty} \int_{S_m(n)} e^{-iy\log x} M_O(N,iy)dy = 0 \qquad (D.19)$$

for $m = 1,2,3$, where once again $S_1$ runs from $(-n,0)$ to $(-n,in)$, $S_2$ from $(-n,in)$ to $(n,in)$ and $S_3$ from $(n,in)$ to $(n,0)$ and $\log x < 0$.

Starting with $S_3$, and using (D.5), we let $y = n + it$ so that $t \in [0,n]$, and

$$|f(j,y)| = |\frac{\Gamma(in-t+j-1/2)}{\Gamma(in-t+j+N-1)}|. \qquad (D.20)$$

We now set $z = in - t$ and notice that $\arg z \in [\pi/2, 3\pi/4]$ so that

$$\begin{aligned}
|f(j,y)| &= \left|\frac{\Gamma(z+j-1/2)}{\Gamma(z+j+N-1)}\right| \\
&\sim \left|\frac{e^{-z}z^{z+j-1}}{e^{-z}z^{z+j+N-3/2}}\right| \\
&= \left|e^{\log z(-N+1/2)}\right| \\
&= \left|e^{(-N+1/2)\log(n^2+t^2)/2}\right|. \qquad \text{(D.21)}
\end{aligned}$$

Thus, since $|2^{2Niy}| \leq 1$, $|M_O(N,iy)| \to 0$ uniformly on $S_3$ as $n \to \infty$, and so we can write for some arbitrarily small $\epsilon$, as long as $n$ is sufficiently large,

$$\begin{aligned}
\left|\int_{S_3(n)} e^{-iyx} M_O(N,iy)dy\right| &< \epsilon\left|\int_n^0 e^{-in\log x} e^{\log xt} i\, dt\right| \\
&\leq \epsilon \int_0^n e^{\log xt} dt \\
&= \frac{\epsilon}{\log x}(e^{n\log x} - 1). \qquad \text{(D.22)}
\end{aligned}$$

$\epsilon$ is arbitrarily small and $\log x < 0$, so (D.19) holds for $S_3$ as $n \to \infty$.

Precisely the same technique works to prove (D.19) for $S_1$, and $S_2$ is tackled in the same manner as for the symplectic case in the previous section. Using (D.12) and $z = it - n + j - 1/2$, where $t \in [-n, n]$,

$$\begin{aligned}
|f(j,t+in)| &= \left|\frac{\Gamma(z)}{\Gamma(z+N-1/2)}\right| \\
\lim_{n\to\infty}\left|\frac{\Gamma(z)}{\Gamma(z+N-1/2)}\right| &= \lim_{n\to\infty}\left|\frac{1}{e^{(N-1/2)\log z}}\right| \\
&= \lim_{n\to\infty}\left|e^{-(N-1/2)\log(it-n+j-1/2)}\right| \\
&= \lim_{n\to\infty}\left|e^{-(N-1/2)\log(t^2+(n-j+1/2)^2)/2}\right| \\
&= 0. \qquad \text{(D.23)}
\end{aligned}$$

As usual $|M_O(N,iy)| \to 0$ uniformly as $n \to \infty$ on $S_2$, so for any $\epsilon > 0$ $|M_O(N,iy)| < \epsilon$ on $S_2(n)$ for large enough $n$. Thus we write

$$\begin{aligned}\left|\int_{S_2} e^{-iy\log x} M_O(N,iy)dy\right| &< \epsilon\int_{-n}^{n}\left|e^{-i(t+in)\log x}\right|dt \\ &\leq \epsilon\int_{-n}^{n} e^{n\log x}dt \\ &= \epsilon(2n)e^{n\log x}, \end{aligned} \tag{D.24}$$

which tends to zero as $n \to \infty$ since $\log x < 0$. Thus (D.19) is proven.

# Bibliography


[ABS94] R. Aurich, J. Bolte, and F. Steiner. Universal signatures of quantum chaos. *Phys. Rev. Lett.*, **73**:1356–1359, 1994.

[ABS97] R. Aurich, A. Bäcker, and F. Steiner. Mode fluctuations as fingerprints of chaotic and non-chaotic systems. *International Journal of Modern Physics B*, **11**(7):805–849, 1997.

[AS65] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, Inc., New York, 1965.

[AS95] A.V. Andreev and B.D. Simons. Correlators of spectral determinants in quantum chaos. *Phys. Rev. Lett.*, **75**(12):2304–7, 1995.

[Atk39] F.V. Atkinson. The mean value of the zeta-function on the critical line. *Quarterly Journal of Mathematics (Oxford Series)*, **10**:122–128, 1939.

[Atk41] F.V. Atkinson. The mean value of the zeta-function on the critical line. *Proceedings of the London Mathematical Society (2)*, **47**:174–200, 1941.

[Bar00] E.W. Barnes. The theory of the $G$-function. *Q. J. Math.*, **31**:264–314, 1900.

[BB70] R. Balian and C. Bloch. Distribution of eigenfrequencies for the wave equation in a finite domain. I. three-dimensional problem with smooth boundary surface. *Ann. Phys.*, **60**:401–47, 1970.

[BB71] R. Balian and C. Bloch. Distribution of eigenfrequencies for the wave equation in a finite domain. II. electromagnetic field. riemannian spaces. *Ann. Phys.*, **64**:271–307, 1971.

[BB72] R. Balian and C. Bloch. Distribution of eigenfrequencies of the wave equation in a finite domain: III. eigenfrequency density oscillations. *Ann. Phys.*, **69**:76–160, 1972.

[BB74] R. Balian and C. Bloch. Solution of the Schrödinger equation in terms of classical paths. *Ann. Phys.*, **85**:514–45, 1974.

[Ber85] M.V. Berry. Semiclassical theory of spectral rigidity. *Proc. R. Soc. Lond. A*, **400**:229–251, 1985.

[Ber88] M.V. Berry. Semiclassical formula for the number variance of the Riemann zeros. *Nonlinearity*, **1**:399–407, 1988.

[Ber91] M.V. Berry. Some quantum-to-classical asymptotics. In M.-J. Giannoni, A. Voros, and J. Zinn-Justin, editors, *Les Houches, Session LII, 1989*, pages 87–199. North Holland: Elsevier Science Publishers, Amsterdam, 1991.

[BG84] O. Bohigas and M.J. Giannoni. In J.M.G.Gomez J.S. Dehesa and A. Polls, editors, *Lecture Notes in Physics 209: Mathematical and computational methods in nuclear physics*, pages 1–99. Springer Verlag, New York, 1984.

[BGS84] O. Bohigas, M.J. Giannoni, and C. Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation. *Phys. Rev. Lett.*, **52**:1–4, 1984.

[BH99] E. Brézin and S. Hikami. Characteristic polynomials of random matrices. *Preprint*, 1999.

[BK90] M.V. Berry and J.P. Keating. A rule for quantizing chaos? *J.Phys. A: Math. Gen.*, **23**:4839–49, 1990.

[BK92] M.V. Berry and J.P. Keating. A new asymptotic representation for $\zeta(1/2+it)$ and quantum spectral determinants. *Proc. R. Soc. Lond. A*, **437**:151–73, 1992.

[BK95] E.B. Bogomolny and J.P. Keating. Random matrix theory and the Riemann zeros I: three- and four-point correlations. *Nonlinearity*, **8**:1115–1131, 1995.

[BK96] E.B. Bogomolny and J.P. Keating. Random matrix theory and the Riemann zeros II:n-point correlations. *Nonlinearity*, **9**:911–935, 1996.

[BK99] M.V. Berry and J.P. Keating. The Riemann zeros and eigenvalue asymptotics. *SIAM Rev.*, **41**(2):236—266, 1999.

[BKP98] M.V. Berry, J.P. Keating, and S.D. Prado. Orbit bifurcations and spectral statistics. *J.Phys. A: Math. Gen.*, **31**:L245–L254, 1998.

[Boh91] O. Bohigas. Random matrix theories and chaotic dynamics. In A. Voros M.-J. Giannoni and and J. Zinn-Justin, editors, *Les Houches, Session LII, 1989*, pages 252–303. North Holland: Elsevier Science Publishers, Amsterdam, 1991.

[BT77] M.V. Berry and M. Tabor. Level clustering in the regular spectrum. *Proc. R. Soc. Lond. A.*, **356**:375–394, 1977.

[CF99] J.B. Conrey and D.W. Farmer. Mean values of $L$-functions and symmetry. *Preprint*, 1999.

[CG84] J.B. Conrey and A. Ghosh. On mean values of the zeta-function. *Mathematika*, **31**:159–161, 1984.

[CG92] J.B. Conrey and A. Ghosh. On mean values of the zeta-function, iii. *Proceedings of the Amalfi Conference on Analytic Number Theory, Università di Salerno*, 1992.

[CG98] J.B. Conrey and S.M. Gonek. High moments of the Riemann zeta-function. *Preprint*, 1998.

[CL95] O. Costin and J.L. Lebowitz. Gaussian fluctuation in random matrices. *Physical Review Letters*, **75**(1):69–72, 1995.

[Dav80] H. Davenport. *Multiplicative Number Theory*. Springer-Verlag, New York, second edition, 1980.

[Dys62] F.J. Dyson. Statistical theory of the energy levels of complex systems, i, ii and iii. *J. Math. Phys.*, **3**:140–175, 1962.

[Gol87] D.A. Goldston. On the function $S(T)$ in the theory of the Riemann zeta-function. *Journal of Number Theory*, **27**:149–177, 1987.

[GR65] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 1965.

[Gut67] M.C. Gutzwiller. Phase-integral approximation in momentum space and the bound states of an atom. *J. Math. Phys.*, **8**(10):1979–2000, 1967.

[Gut69] M.C. Gutzwiller. Phase-integral approximations in momentum space and the bound states of an atom, II. *J. Math. Phys.*, **10**:1004–20, 1969.

[Gut70] M.C. Gutzwiller. Energy spectrum according to classical mechanics. *J. Math. Phys.*, **11**(6):1791–1806, 1970.

[Gut71] M.C. Gutzwiller. Periodic orbits and classical quantization conditions. *J. Math. Phys.*, **12**(3):343–58, 1971.

[HA84] J.H. Hannay and A. M. Ozorio De Almeida. Periodic orbits and a correlation function for the semiclassical density of states. *J. Phys. A: Math. Gen.*, **17**:3429–40, 1984.

[Haa90] F. Haake. *Quantum Signatures of Chaos*. Springer, Berlin, 1990.

[HB79] D.R. Heath-Brown. The fourth power moment of the Riemann zeta function. *Proc. London Math. Soc. (3)*, **38**:385–422, 1979.

[HB81] D.R. Heath-Brown. Fractional moments of the Riemann zeta-function. *J.London Math. Soc. (2)*, **24**:65–78, 1981.

[HB93] D.R. Heath-Brown. Fractional moments of the Riemann zeta-function, ii. *Quart. J. Math. Oxford (2)*, **44**:185–197, 1993.

[HKO] C.P. Hughes, J.P. Keating, and N. O'Connell. On the characteristic polynomial of a random unitary matrix, in preparation.

[HKS87] F. Haake, M. Kuś, and R. Scharf. Classical and quantum chaos for a kicked top. *Z. Phys. B - Condensed Matter*, **65**:381–395, 1987.

[HKS+96] F. Haake, M. Kuś, H.-J. Sommers, H. Schomerus, and K. Życzkowski. Secular determinants of random unitary matrices. *J. Phys. A: Math. Gen.*, **29**:3641–58, 1996.

[HL18] 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.

[ILS99] H. Iwaniec, W. Luo, and P. Sarnak. Low lying zeros of families of $L$-functions. *Preprint*, 1999.

[Ing26] A.E. Ingham. Mean-value theorems in the theory of the Riemann zeta-function. *Proceedings of the London Mathematical Society (2)*, **27**:273–300, 1926.

[IS97] H. Iwaniec and P. Sarnak. The non-vanishing of central values of automorphic $L$-functions and Siegel's zeros. *Preprint*, 1997.

[Jut81] M. Jutila. On the mean value of $L(1/2,\chi)$ for real characters. *Analysis*, **1**:149–161, 1981.

[Jut83] M. Jutila. On the value distribution of the zeta-function on the critical line. *Bull. London Math. Soc.*, **15**:513–518, 1983.

[Kea92] J.P. Keating. Periodic orbit resummation and the quantization of chaos. *Proc. R. Soc. Lond. A*, **436**:99–108, 1992.

[Kea93] J.P. Keating. The Riemann zeta function and quantum chaology. In G. Casati, I Guarneri, and U. Smilansky, editors, *Quantum Chaos*, pages 145–85. North-Holland, Amsterdam, 1993.

[KKS97] S. Kettemann, D. Klakow, and U. Smilansky. Characterization of quantum chaos by the autocorrelation function of spectral determinants. *J. Phys. A: Math. Gen.*, **30**(10):3643–62, 1997.

[KS99a] N.M. Katz and P. Sarnak. *Random Matrices, Frobenius Eigenvalues and Monodromy.* AMS, Providence, Rhode Island, 1999.

[KS99b] N.M. Katz and P. Sarnak. Zeros of zeta functions and symmetry. *Bull. Amer. Math. Soc.*, **36**:1–26, 1999.

[KS00a] J.P. Keating and N.C. Snaith. Random matrix theory and $L$-functions at $s = 1/2$. *Commun. Math. Phys*, **214**:91–110, 2000.

[KS00b] J.P. Keating and N.C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Commun. Math. Phys.*, **214**:57–89, 2000.

[Mar65] A.I. Markushevich. *Theory of Functions of a Complex Variable, Vol. II.* Prentice-Hall Inc., Englewood Cliffs, NJ, 1965.

[Meh91] M.L. Mehta. *Random Matrices.* Academic Press, London, second edition, 1991.

[Mon73] H.L. Montgomery. The pair correlation of the zeta function. *Proc. Symp. Pure Math*, **24**:181–93, 1973.

[Odl97] A.M. Odlyzko. The $10^{20}$th zero of the Riemann zeta function and 70 million of its neighbors. *Preprint*, 1997.

[Ram78] K. Ramachandra. Some remarks on the mean value of the Riemann zeta-function and other Dirichlet series - i. *Hardy-Ramanujan J.*, **1**:1–15, 1978.

[Ram80] K. Ramachandra. Some remarks on the mean value of the Riemann zeta-function and other Dirichlet series - iii. *Ann. Acad. Sci. Fenn. Ser. AI Math.*, **5**:145–158, 1980.

[RS96] Z. Rudnick and P. Sarnak. Principal $L$-functions and random matrix theory. *Duke Mathematical Journal*, **81**(2):269–322, 1996.

[Rub98] M. Rubinstein. *Evidence for a Spectral Interpretation of Zeros of L-functions.* PhD thesis, Princeton University, 1998.

[Rum93] R. Rumely. Numerical computations concerning ERH. *Math. Comp.*, **61**:415–440, 1993.

[Smi99] U. Smilansky. Semiclassical quantization of maps and spectral correlations. In I.V. Lerner, J.P. Keating, and D.E. Khmelnitskii, editors, *Supersymmetry and trace formulae: chaos and disorder*, pages 173–191. Plenum, New York, 1999.

[Sou99] K. Soundararajan. Non-vanishing of quadratic Dirichlet $L$-functions at $s = \frac{1}{2}$. *Preprint*, 1999.

[Tit86] E.C. Titchmarsh. *The Theory of the Riemann Zeta Function.* Clarendon Press, Oxford, second edition, 1986.

[Vig79] M.-F. Vignéras. L'equation fonctionnelle de la fonction zeta de Selberg du groupe modulaire $PSL(2, Z)$. *Asterisque*, **61**:235–249, 1979.

[Vor87] A. Voros. Spectral functions, special functions and the selberg zeta function. *Commun. Math. Phys.*, **110**:439–465, 1987.

[Wey46] H. Weyl. *Classical Groups.* Princeton University Press, 1946.