Title
The statistics of the zeros of the Riemann zeta-function and related topics

Author
Rodgers, Brad

Publication Date
2013

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

The statistics of the zeros of the Riemann zeta-function and related topics

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Bradley William Rodgers

2013

Abstract of the Dissertation

The statistics of the zeros of the Riemann zeta-function and related topics

by

Bradley William Rodgers

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2013

Professor Terence Tao, Chair

This thesis concerns statistical patterns among the zeros of the Riemann zeta function, and conditioned on the Riemann hypothesis proves several related original results. Among these:

By extending a well known result of H. Montgomery, we show, at an only microscopically blurred resolution, that the distance between two randomly selected zeros of the zeta function tends to weakly repel away from the location of low-lying zeros of the zeta function.

For random collections of consecutive zeros that are not so large as to see this resurgence effect, we support the view that they resemble the bulk eigenvalues of a random matrix by in particular proving an analogue of the strong Szegő theorem.

Concerning even smaller collections of zeros, we show that a statement that the zeros of the Riemann zeta function locally resemble the eigenvalues of a random matrix (the GUE Conjecture) is logically equivalent to a statement about the distribution of primes. On this basis, we make a conjecture for the covariance in short intervals of integers with fixed numbers of prime factors, weighted by the higher order von Mangoldt function. This is related to the so-called ratio conjecture. The covariance pattern is surprisingly simple to write down.

We finally include a rigorous derivation that uniform variants of the Hardy-Littlewood conjectures agree with the GUE Conjecture. Even thus conditioned, the range of correlation test functions against which we may confirm the GUE pattern for zeta zeros remains limited. We consider in detail the case of two, three, and four point correlations, the two point case being due to Mongtomery.

The dissertation of Bradley William Rodgers is approved.

Sudip Chakravarty

William Duke

Rowan Killip

Terence Tao, Committee Chair

University of California, Los Angeles

2013

To Dwain Wall, from whom I got my first glimpse of mathematics.

Table of Contents

List of Figures

List of Tables

Acknowledgments

This thesis would have been impossible without a large number of people, and no doubt I will only be able to thank a small portion of them here:

I want to thank first my advisor, Terry Tao, who has taught me a great deal of mathematics during weekly meetings and through blog posts and emails. I owe him a debt of gratitude for looking through the material that follows, though of course the responsibility for any errors that remain is my own. I am grateful for his feedback, creativity, and his optimistic and humanistic outlook.

I thank Rowan Killip as well, from whom I have also learned a great deal of mathematics and who has contributed to my aesthetic outlook on the subject. My interest in the strong Szegő theorem is due to him, of which chapter 3 of this thesis is an outcome.

I also thank John Garnett, who was responsible for convincing me to attend UCLA for graduate school.

Finally among the faculty at UCLA I want to thank Tamara Kucherenko, who went well beyond her responsibilities to run a reading course with me in Banach space and convex geometry during my first year and more generally offer me advice during a time I had difficulty adapting to a new environment. Some discussions with Leonard Lipshitz were also a great help during this time.

I want also to thank my parents Bill and Lynne and my sister Lindsay for their unwavering love and support, and also my friends who have made my time the past few years more enjoyable and the work that went into this thesis more possible.

I have also especially benefitted from discussions with a number of mathematicians outside UCLA, and for their role in the document that follows I want to thank in particular: Zeev Rudnick for reading an early draft of what was to become chapter 3, pointing out to me the work of M. Berry along with his own work with D. Faifman, and for expressing interest and feedback in my work on the covariance of almost primes; Jeff Stopple for reading an early version of what was to become chapter 2; Steve Gonek for supplying me with a copy of the preprint [25] before it had appeared in print; and Jonathan Bober for a very helpful comment in a discussion hosted on the website MathOverflow, available at `http://mathoverflow.net/questions/83027/`.

Chapter 2 is drawn from an article of the author's to appear in *Quart. J. of Math.* I thank the anonymous referee of this article, and an anonymous referee who has looked over the material of Chapter 3 and 4 as well.

I am grateful to have been supported financially at various times during my PhD by a Radcliffe Dee Fellowship, an NSF RTG Grant, and a UCLA Dissertation Year Fellowship.

Los Angeles, May 2013

Vita

2007	Phi Beta Kappa
2008	B.S. (Mathematics), Purdue University
2008-2009	Radcliffe Dee Fellowship, UCLA
2012-2013	Dissertation Year Fellowship, UCLA

Publications

"Macroscopic Pair Correlation of the Riemann Zeroes for Smooth Test Functions," to appear in *Quart. J. of Math.*

CHAPTER 1

Introduction

1.1 Background material

This thesis is a compilation and extension of three papers that the author has written in the past two years regarding the well-known but still conjectural connection between the spacing of the zeros of the Riemann zeta-function and the bulk spacing of eigenvalues of a wide variety of random matrix ensembles – the Gaussian Unitary Ensemble being the best known. These papers are

1. "Macroscopic Pair Correlation of the Riemann Zeroes for Smooth Test Functions"

2. "A central limit theorem for the zeroes of the Riemann zeta function"

3. "Arithmetic consequences of the GUE Conjecture for zeta zeros."

The first of these has been accepted for publication in *Quart. J. of Math.* They have been lightly edited and rearranged to fit more cohesively together, though I hope they still may be read independently. I sometimes restate in full (though never reprove) results from earlier chapters for this reason. The content of each chapter is summarized below.

For the sake of completeness, we now briefly recall the very basic properties of the Riemann zeta-function. A number of texts such as [79], [56] serve as a more complete introduction. $\zeta(s)$ is defined for $\Re s > 1$ by

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{\substack{p \\ \text{prime}}} \frac{1}{1-p^{-s}}, \tag{1.1}$$

and by analytic continuation elsewhere. If we define $\xi(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s)$, we have the functional equation

$$\xi(s) = \xi(1-s),$$

and from this, the relation (1.1), and basic analytic properties of the gamma factor $\pi^{-s/2}\Gamma(s/2)$, one sees that $\zeta(s)$ has a simple pole at $s = 1$, zeros at the negative even integers $s = -2, -4, ...$, and all other zeros confined to the critical strip $\Re s \in [0, 1]$. These latter zeros are the so-called nontrivial zeros, and we label them with multiplicity by $\rho = 1/2 + i\gamma$.

To first order the distribution of zeros is described by a formula first stated by Riemann and proved rigorously by von Mangoldt.

Theorem 1.1.1 (Riemann - von Mangoldt)**.** *Let*

$$N(T) := \#\{\rho : \zeta(\rho) = 0, \Re\rho \in (0, 1), \Im\rho \in (0, T]\}$$

be the number of zeros of the zeta function in the critical strip of imaginary height no more than T. Then

$$N(T) = \frac{T}{2\pi}\log T - \frac{T}{2\pi} + O(\log T).$$

That is to say, the density of zeros at an imaginary height T is roughly $\log T/2\pi$, or said another way the average spacing between two zeros at a height T is $2\pi/\log T$.

A more detailed investigation of the spacing of zeros of the zeta function depends typically on the assumption of the Riemann hypothesis (RH), that all non-trivial zeros lie on the line $\Re s = 1/2$ – that is, that for all zeros $\gamma \in \mathbb{R}$. *We will assume the Riemann hypothesis in what follows.*

It was on the basis of RH that Hugh Mongtomery first investigated the effect the placement of one zero has on the likelihood other zeros are nearby, at the scale of mean spacing. He was led to the following:

Conjecture 1.1.2 (Pair correlation)**.** *For any fixed Schwartz test function η,*

$$\frac{2\pi}{T\log T}\sum_{\substack{T\le\gamma,\gamma'\le 2T\\ \text{distinct}}}\eta\big(\tfrac{\log T}{2\pi}(\gamma-\gamma')\big)\sim\int_{\mathbb{R}}\eta(x)\Big(1-\big(\tfrac{\sin\pi x}{\pi x}\big)^2\Big)\,dx.$$

In a now famous conversation over tea at the Institute for Advanced Study with Freeman Dyson, Montgomery learned the pair correlation function $1-\left(\frac{\sin\pi x}{\pi x}\right)^2$ he had conjectured for zeta zeros is known to be the pair correlation function for eigenvalues of certain classes of random matrices. On this basis, for reasons explained shortly one may also conjecture

Conjecture 1.1.3 (GUE)**.** *For any fixed n and any fixed $\eta \in C_c(\mathbb{R}^n)$,*

$$\lim_{T\to\infty}\frac{1}{T}\int_T^{2T}\sum_{\substack{\gamma_1,\ldots,\gamma_n\\ \text{distinct}}}\eta\Big(\tfrac{\log T}{2\pi}(\gamma_1-t),\ldots,\tfrac{\log T}{2\pi}(\gamma_n-t)\Big)\,dt=\int_{\mathbb{R}^n}\eta(x)\det_{n\times n}\Big(K(x_i-x_j)\Big)\,d^nx \tag{1.2}$$

where the entries of the $n\times n$ determinant are formed from the function $K(x) = \frac{\sin\pi x}{\pi x}$.

The left hand side of (1.2) is referred to as an n-level joint intensity. The reader familiar with probability may note that the GUE Conjecture is the statement that the zeros of the zeta function near a random translate t, stretched out to have mean unit density, tend in distribution to the determinantal point process with sine-kernel.

A wide variety of random matrix ensembles display this same pattern, perhaps the best known of which is the *Gaussian Unitary Ensemble* (GUE). This is an ensemble of random $n \times n$ Hermitian matrices $M_{n\times n}$ with upper triangular entries composed of independent and identically distributed complex normal variables of mean 0 and complex variance 1, lower triangular entries defined by Hermitian symmetry, and diagonal entries composed of independent and identically distributed real normal variables of mean 0 and variance 1.

Theorem 1.1.4 (Wigner)**.** *For $n \times n$ GUE-distributed random matrices $M_{n\times n}$, for any test function η the eigenvalues $\lambda_1, ..., \lambda_n$ of $M_{n\times n}$ satisfy*

$$\mathbf{E}\frac{1}{n}\sum_{i=1}^{n}\eta\Big(\frac{\lambda_i}{\sqrt{n}}\Big) \sim \int_{\mathbb{R}} \eta(x)\rho_{sc}(x)\,dx \tag{1.3}$$

where $\rho_{sc}(x) := \frac{1}{2\pi}(4-x^2)_+^{1/2}$.

Theorem 1.1.5 (Gaudin-Mehta)**.** *For any $E \in (-2,2)$, any $k \geq 1$, and any $\eta \in C_c(\mathbb{R}^k)$,*

$$\mathbf{E}\sum_{\substack{i_1,...,i_k\\ \text{distinct}}} \eta\Big(n\rho_{sc}(E)\big(\frac{\lambda_{i_1}}{\sqrt{n}}-E\big), ..., n\rho_{sc}\big(\frac{\lambda_{i_k}}{\sqrt{n}}-E\big)\Big) \sim \int_{\mathbb{R}^k} \eta(x) \det_{k\times k}\Big(K(x_i-x_j)\Big) \tag{1.4}$$

where K is defined in Theorem 1.1.3.

Theorem 1.1.4 has long been known to hold for more general classes of random Hermitian matrices than GUE; recently Theorem 1.1.5 as well was shown to hold in the case that gaussian entries are replaced by more general i.i.d variables with bounded m^{th} moment, for m sufficiently large [78], [22].

Here the semi-circular law, Theorem 1.1.4 is a (rather weaker) analogue of the Riemann-von Mangoldt Theorem 1.1.1, while Theorem 1.1.5 is exactly the analogue of the GUE Conjecture. Note that in our counts we have stretched out the points $\lambda_i/\sqrt{n}$ near E by a factor of $n\rho_{sc}(E)$ so that thus rescaled they locally have a unit mean density, exactly as with zeta zeros.

Although the relationship between the GUE Conjecture and Theorem 1.1.5 is especially striking in light of a conjecture of Hilbert and Polya that the reason γ are always real is that they have an interpretation as eigenvalues of some still-to-be-described self-adjoint operator, for us it will be more convenient computationally to work with the eigenvalues of random $n \times n$ unitary matrices $U(N)$ with Haar probability-measure. In this case the eigenvalues follow a uniform distribution around the unit circle, rather than the somewhat more complicated semicircular distribution of Hermitian eigenvalues. Locally they too tend to a determinantal point process with sine-kernel; the analogue of Theorem 1.1.5 for $U(N)$ is given by Proposition 4.0.7.

Quite independent of speculations like Hilbert and Polya's, there is by now a great deal of concrete evidence in favor of the GUE Conjecture, both numerical and theoretical. Numerically, computations begun by Odlyzko [57] concerning the spacing between consecutive entries in large tables of zeros leave little doubt in the truth of the GUE Conjecture. Theoretically, conditioned on the Riemann hypothesis, results of Montgomery [53], Hejhal [38], and Rudnick & Sarnak [63] successively established (in the cases $k = 2$, $k = 3$, and $k \geq 3$ respectively),

Theorem 1.1.6 (Montgomery, Hejhal, Rudnick & Sarnak)**.** *For Schwartz* $\eta : \mathbb{R}^n \to \mathbb{R}$ *so that*

$$\text{supp } \hat{\eta} \subset \{x \in \mathbb{R}^n \ : |x_1| + \cdots + |x_n| \leq 2\},$$

equation (1.2) *is true.*

That is, provided we restrict ourselves to counting with stringently smooth test functions, the GUE Conjecture can be verified. We have adopted a convention here of labeling zeros with multiplicity in Conjecture 1.1.3, in the (likely vacuous) case that some zeros of the zeta function are not simple.

In a slightly different direction, beginning with the work of Katz and Sarnak [45], [46], a large number of analogous results have been proven for the statistics of zeros of *families* of L-functions over function fields. Here of course the zeros of L-functions do have a spectral interpretation, as the inverses of eigenvalues of the Frobenius operator acting on a suitable cohomology group. This perspective of families has been profitably applied in the number field case as well to show that the low-lying zeros of certain families of L-functions fall into random matrix patterns (again, in the number field case, always counted by stringently band-limited test functions as in Theorem 1.1.6). See for instance [62].

1.2 A summary of chapters

The conjectured convergence of zeros to a determinantal point process with sine-kernel described by the GUE Conjecture is really only a statement about the effect that zeros at a height T have on other zeros within a distance of $O(1/\log T)$. In chapters 2 through 4 we will discuss the extent to which the zeros at a larger scale resemble or cease to resemble random matrices in this way.

It will be convenient to make use of the following terminology, meaningful as $T \to \infty$: the interaction of zeros at a height T seperated by a distance of $O(1/\log T)$ we refer to as *microscopic*. The GUE Conjecture therefore describes the statistical interaction of zeros at a microscopic level. The statistical interaction between zeros at a height T separated by a distance larger in order than $1/\log T$ but less in order than $o(1)$ as $T \to \infty$ we refer to as *mesoscopic*. Finally the interaction between zeros at a scale larger than $\lesssim 1$, we refer to as *macroscopic*. The mesoscopic and macroscopic regime of interaction will be illustrated by examples presently.

In chapter 2, we consider the pair correlation function of the zeta zeros in the macroscopic regime: the sums we consider are of the sort

$$\sum_{\substack{0\leq\gamma,\gamma'\leq T\\ \text{distinct}}} f(\gamma-\gamma'). \tag{1.5}$$

Note that we have not dilated the difference $\gamma-\gamma'$ by a factor of $\log T/2\pi$ here, as in Conjecture 1.1.2. We evaluate these sums uniformly over a range of functions f, however, that may oscillate on a microscopic scale of $1/\log T$. In chapter 2 we explain how it is that Montgomery's work on the pair correlation function can be recovered from this information, and that we are thus able to see the statistical behavior of $\gamma-\gamma'$ near any point s up to an only microscopically blurred resolution. This is enough to rigorously see, for instance, a resurgence phenomenon in which $\gamma-\gamma'$ tends to weakly repel away from low-lying γ themselves (see figure 2.1), first noticed heuristically by Bogomolny and Keating [4], and recently rediscovered by Pérez-Marco [58].

Around a given $\gamma\in(0,T]$, the sums 1.5 will tend to collect on the order of $\log T$ γ' near γ. In chapter 3, we consider the linear statistics of slightly smaller mesoscopic collections of consecutive zeros. That is, for random t uniformly distributed in the interval $[T,2T]$, we consider the statistics

$$\Delta_\eta := \sum_\gamma \eta\Big(\tfrac{\log T}{2\pi n(T)}(\gamma-t)\Big),$$

where η is some test function of compact support and bounded variation, and $n(T)$ is some function that grows with T such that $n(T)\to\infty$ but $n(T)=o(\log T)$. The linear statistics take the expected value

$$\mathop{\mathbf{E}}_{t\in[T,2T]} \Delta_\eta = n(T)\int_{\mathbb{R}} \eta(x)\,dx + o(1)$$

and we show that for all but certain pathological η, the random variable

$$\frac{\Delta_\eta - \mathbf{E}\,\Delta_\eta}{\sqrt{\mathrm{Var}\ \Delta_\eta}}$$

tends in distribution to a normal variable of mean 0 and variance 1, where

$$\mathrm{Var}\ \Delta_\eta \sim \int_{-n(T)}^{n(T)} |x| \cdot |\hat{\eta}(x)|^2\, dx.$$

This is an analogue of the strong Szegő theorem, a matter of classical interest in random matrix theory. The reason this result so closely mirrors a result of random matrix theory is that in the mesoscopic regime, we do not see the macroscopic resurgence phenomenon of chapter 2, which acts as the only substantial obstruction differentiating the statistics of zeros from random matrix statistics.

In chapter 4, we provide a proof of Theorem 1.1.6, along the lines of [42], and develop a more precise account of the resemblance in the mesoscopic regime of the statistics of the zeta function to the statistics of random matrices.

In Chapter 5, we show that the GUE Conjecture as stated above is logically equivalent to a statement about the distribution of the primes. This overlaps with the recent work [25], who consider similar questions conditioned on hypotheses we do not require here. This statement about the primes, though complicated, has at least one elegant consequence, that as weighted by the higher-order von Mangoldt functions, the counts almost primes of various orders in short intervals have a curiously simple covariance pattern. To arrive at this statement we will pass through some random matrix statistics of independent interest. The methods used in this chapter also yield a simple proof, conditioned on RH, of new bounds for the moments of the logarithmic derivative of the zeta function.

Finally, in Chapter 6, we show that conditioned on a uniform variant of conjectures made by Hardy & Littlewood regarding the spacing of prime pairs, one can enlarge slightly the class of test functions against which the correlation functions of zeta zeros agrees with the GUE pattern in Theorem 1.1.6. This extends well-known work of Montgomery in the pair correlation case (an account of which can be found in [54]).

1.3 Notation

$f(x) \lesssim g(x)$	There is a constant C such that $\lvert f(x)\rvert \leq Cg(x)$. Used interchangeably with $f(x) = O(g(x))$.
$f_k(x) \lesssim_k g_k(x)$	There is a constant C_k depending on k so that $\lvert f_k(x)\rvert \leq C_k g_k(x)$
$e(x)$	$e(x) = e^{i2\pi x\xi}$
$\hat{f}(\xi)$	$\hat{f}(\xi) := \int_{-\infty}^{\infty} f(x)e(-x\xi)\,dx$
$\check{f}(x)$	$\check{f}(x) := \int_{-\infty}^{\infty} f(\xi)e(x\xi)\,d\xi$
$\mathbb{N}_+$	$\mathbb{N}_+ := \{1, 2, 3, ...\}$
γ	imaginary ordinate of a nontrivial zeta zero, $\zeta(1/2+i\gamma) = 0$
$C_c(\mathbb{R}^k)$	the set of continuous and compactly supported functions on $\mathbb{R}^n$
$\det_{n\times n}$	an $n \times n$ determinant
$K(x)$	$K(x) := \frac{\sin \pi x}{\pi x}$
$\Lambda(n)$	von Mangoldt function, $\log p$ if n is p^k the power of a prime, 0 otherwise
ψ	$\psi(x) := \sum_{n\leq x} \Lambda(n)$
$U(N)$	the group of $N \times N$ unitary matrices u, with Haar probability measure du
$\tilde{d}(\xi)$	$\tilde{d}(\xi) := \sum \delta_\gamma(\xi)$
$S(t)$	$S(t) := \frac{1}{\pi} \arg \zeta(1/2 + it)$
$\Omega(t)$	$\Omega(t) := \frac{1}{2}\frac{\Gamma'}{\Gamma}\big(\frac{1}{4} + i\frac{t}{2}\big) + \frac{1}{2}\frac{\Gamma'}{\Gamma}\big(\frac{1}{4} - i\frac{t}{2}\big) - \log \pi$
$Z(\beta)$	$Z(\beta) := \det(1 - e^{-\beta}u)$
$dz(x)$	$dz(x) := e^{-x/2}d\big(\psi(e^x) - e^x\big)$
admissible	See definition 5.2.3
$\upsilon_T(x, y)$	$\upsilon_T(x, y) := \big(1 - T\lvert x - y\rvert\big)_+$

Ψ_T	See equation (5.10)
Θ_T	See equation (5.11)
$\Lambda_j(n)$	$\Lambda_j(n) := \sum_{d\mid n} \mu(d) \log^k(n/d)$
ψ_j	$\psi_j(x) := \sum_{n \le x} \Lambda_j(n)$
$\tilde{\psi}_j$	See equation (5.22)
$\tilde{\psi}_j(x; H)$	$\tilde{\psi}_j(x; H) = \tilde{\psi}_j(x + H) - \tilde{\psi}_j(x)$
$H_j(r)$	See equations (5.26) and (5.27)
$\int_{\to -\infty}^{\to \infty}$	An improper integral
α, α_R	Bump functions centered at 0 of width 2 and $2R$; see equations (5.29), (5.30)
$Z_T, Z_T(\sigma)$	Point processes induced by zeta zeros, see definitions 5.4.1 and 5.4.3
GUE(σ)	The GUE Conjecture with averaging σ, see definition 5.4.4
$\mathcal{S}$	The sine-kernel determinantal point process (See Appendix B)
$\mathcal{S}_N$	See definition 4.0.7
$\mathcal{S}'_N$	See definition 5.4.6
G_T	$G_T(\eta, t) := \sum_\gamma \eta\Big(\frac{\log T}{2\pi}(\gamma - t)\Big)$
L_T	$L_T(\eta, t) := \int_{-\infty}^{\infty} \eta\Big(\frac{\log T}{2\pi}(\xi - t)\Big) \frac{\log(\lvert\xi\rvert+2)}{2\pi}\, d\xi$
$\widetilde{G}_T$	$\widetilde{G}_T(\eta, t) := \int_{-\infty}^{\infty} \eta\Big(\frac{\log T}{2\pi}(\xi - t)\Big)\, dS(\xi)$
M_k	See equation (5.39)
$d\lambda_k(t)$	$d\lambda_k(t) := \log^k(\lvert t\rvert + 2)\, dt$
ω_ϵ	$\omega_\epsilon(x) := 1 - \alpha_\epsilon(x)$
Ω_ϵ	$\Omega_\epsilon(x) := \omega_\epsilon(x) \mathbf{1}_{\mathbb{R}_+}(x)$
$f\rvert_\epsilon$	$f\rvert_\epsilon(x) := f(x)\Omega_\epsilon(x)$
$f\rvert_a^b$	$f\rvert_a^b(x) := f\rvert_a(x) - f\rvert_b(x)$

CHAPTER 2

Macroscopic statistics: Macroscopic pair correlation of the Riemann zeros for smooth test functions

2.1 Introduction

This chapter is an account of the pair correlation function of the zeros of the Riemann zeta function. In particular we will consider the pair correlation function at a macroscopic scale. In this way, we shall see how far the macroscopic statistics of the Riemann zeta function can be understood in a rigorous fashion. By this we mean especially those numerical statistics that seem to indicate a statistical repulsion of the differences of zeros of Riemann's zeta function away from low lying zeros. (See Figure 2.1.) Such statistics were first noticed by Bogomolny and Keating [4], who predicted them heuristically, and recently rediscovered by Pérez-Marco in [58]. [69] contains a relatively recent survey.

We will make one concession of rigor, which is to assume as throughout the thesis that the Riemann Hypothesis is true. This is not entirely necessary, but without it the results (and conjectures) that follow would not be particularly meaningful.[1] We will show that a formula first conjectured by Bogomolny and Keating indicating the observed repulsion is true for sufficiently smooth test func-

[1]If for a sufficiently nice function f one understands $f(x+iy)$ in the harmonic sense to be $\int \hat{f}(\xi)e((x+iy)\xi)d\xi$, then what follows can made unconditional. This observation, which has been made before for similar problems, would seem to be of rather secondary interest.

tions. Our approach in some ways consists in nothing more than carefully coloring in their heuristic computations.

If the non-trivial zeros of $\zeta(s)$ are written in the form $1/2 + i\gamma$, then the Riemann Hypothesis is the statement that γ is always real. We will slightly abuse terminology by sometimes referring to the γ's as 'zeros'; this should cause no confusion. The number of such γ in an interval $[T\,,\,T+1]$ is known by the Riemann-von Mangoldt formula, Theorem 1.1.1 to be roughly $\frac{\log T}{2\pi}$, so that the spacing between consecutive zeros in this interval is on average $\frac{2\pi}{\log T}$. (A slightly better approximation to density near T is $\frac{\log(T/2\pi)}{2\pi}$, which will make an appearance later.)

We recall Montgomery's [53] more precise conjecture concerning these spacings

Conjecture 2.1.1 (Pair Correlation Conjecture)**.** *For a fixed Schwartz class test function f,*

$$\sum_{\substack{\gamma,\gamma'\in[0,T]\\ \text{distinct}}} f\big(\tfrac{\log T}{2\pi}(\gamma-\gamma')\big) = T\frac{\log T}{2\pi}\bigg(\int_{-\infty}^{\infty} f(u)\Big[1-\Big(\frac{\sin \pi u}{\pi u}\Big)^2\Big]du + o(1)\bigg).$$

More informally, for a random γ' of height roughly T, the expected number of distinct γ to lie in an interval $[\gamma' + \frac{2\pi\alpha}{\log T}\,,\,\gamma' + \frac{2\pi\beta}{\log T}]$ is $\int_\alpha^\beta 1 - (\frac{\sin \pi u}{\pi u})^2 du$. Since the integrand is small when u is small, there is very little chance that the distance between zeros will be orders less than $1/\log T$, and in this sense zeros repel one another.

This conjecture Montgomery derived from a slightly stronger conjecture:

Conjecture 2.1.2 (Strong Pair Correlation Conjecture)**.** *For fixed M and $w(u) = \frac{4}{4+u^2}$,*

$$\frac{2\pi}{T\log T}\sum_{0\le\gamma,\gamma'\le T} T^{i\alpha(\gamma-\gamma')}w(\gamma-\gamma') = 1-(1-|\alpha|)_+ + o(1) + (1+o(1))T^{-2\alpha}\log T, \tag{2.1}$$

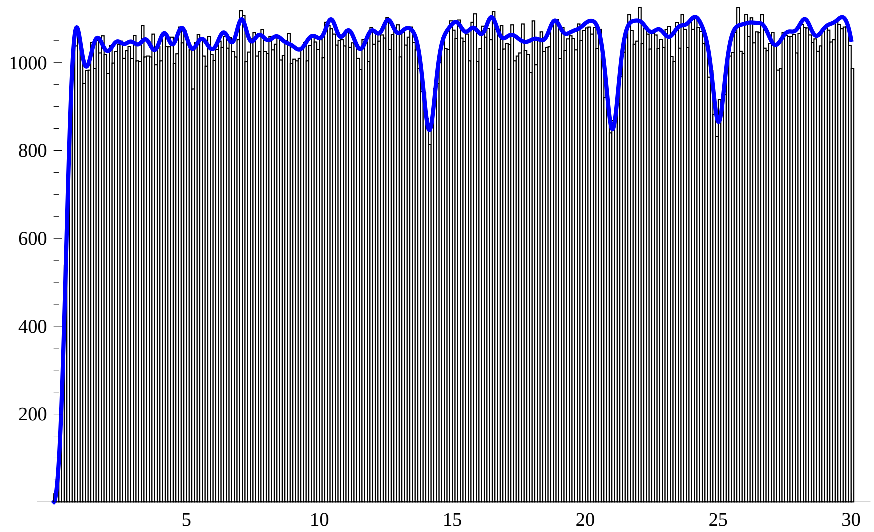

Figure 2.1: A histogram of $\gamma - \gamma'$ for the first 10,000 zeros, counting the number of such differences to lie in intervals of size 0.1, compared to the appropriately scaled Bogomolny-Keating prediction in Theorem 2.1.4. Note that the smaller troughs around 14.13, 21.02, and 25.01 occur approximately at locations of zeros themselves (although this pattern becomes less prominent around higher-lying zeros; see the discussion at the end of section 4), and that the larger trough at the origin would, if stretched out by a factor of $\log T/2\pi$, resemble the GUE measure in Montgomery's conjecture, where T is the number of zeros included in the histogram. Generated with Mathematica.

uniformly for $\alpha \in [-M\,,\,M]$.

Note that here γ, γ' need not be distinct, and for α away from 0, the term $T^{-2\alpha}\log T$ is unimportant.

For any $\epsilon > 0$, he was able to prove this for $M = 1-\epsilon$, and by integration in α he could prove Conjecture 2.1.1 for f with $\text{supp}\hat{f} \subset [-M\,,\,M]$. That this holds uniformly across α and that the weight function $w(\gamma-\gamma')$ collects γ in the vicinity of γ' numbering not just $O(1)$ but $O(\log T)$ has a certain statistical significance which is not typically remarked upon but which extends beyond Conjecture 2.1.1.

We may draw out the point – and motivate our computations that follow – by putting the Strong Pair Correlation Conjecture in a slightly different form. Montgomery made use of the fact that the left hand side of (2.1) is equal to

$$\begin{aligned}
&\frac{4}{T\log T}\int_0^T \Big|\sum_{\gamma} \frac{T^{i\alpha\gamma}}{1+(\gamma-t)^2}\Big|^2\,dt\, + O\Big(\frac{\log^2 T}{T}\Big)\\
&= \frac{4}{T\log T}\int_0^T \sum_{\gamma,\gamma'} \rho(\gamma-t)\rho(\gamma'-t)e\big(\alpha\tfrac{\log T}{2\pi}(\gamma-t) - \alpha\tfrac{\log T}{2\pi}(\gamma'-t)\big)\,dt\\
&\quad + O\Big(\frac{\log^2 T}{T}\Big),
\end{aligned}$$

where $\rho(u) = 1/(1+u^2)$. The passage to the first line is made possible by knowing what is in effect the one-level density of Zeta zeros, that

$$N(T) := \#\{\gamma \in (0,T)\} = \tfrac{T}{2\pi}\log\big(\tfrac{T}{2\pi}\big) - \tfrac{T}{2\pi} + O(\log T).$$

The reader is referred to Montgomery's paper for details. On the other hand, recalling the Fourier pair,

$$g(\nu) := \Big(\frac{\sin \pi\nu}{\pi\nu}\Big)^2$$

$$\hat{g}(x) = (1-|x|)_+,$$

we have by Parseval,

$$\begin{aligned}
&\int_{\mathbb{R}^2} \rho\big(\tfrac{2\pi\nu_1}{\log T}\big)\rho\big(\tfrac{2\pi\nu_1}{\log T}\big)e\big(\alpha\nu_1-\alpha\nu_2\big)\Big[1+\delta(\nu_1-\nu_2)-\Big(\frac{\sin\pi(\nu_1-\nu_2)}{\pi(\nu_1-\nu_2)}\Big)^2\Big]d\nu_1 d\nu_2\\
&=\frac{\log^2 T}{4\pi^2}\int_{\mathbb{R}^2}\hat{\rho}\big(-\tfrac{\log T}{2\pi}\xi_1\big)\hat{\rho}\big(-\tfrac{\log T}{2\pi}\xi_2\big)\delta(\xi_1+\xi_2)\\
&\qquad\qquad\times\Big[1+\delta\big(\tfrac{\xi_1-\xi_2}{2}+\alpha\big)-\big(1-\big|\tfrac{\xi_1-\xi_2}{2}+\alpha\big|\big)_+\Big]\,d\xi_1 d\xi_2\\
&=\frac{\log^2 T}{4}\int_{\mathbb{R}} e^{-2\log T|\xi|}\big[1+\delta(\xi+\alpha)-(1-|\xi+\alpha|)_+\big]\,d\xi\\
&=\frac{\log^2 T}{4}\Big(T^{-2\alpha}+\tfrac{1}{\log T}[1-(1-|\alpha|)_+ + o(1)]\Big)
\end{aligned}$$

We may conclude that Montgomery's Strong Pair Correlation Conjecture is equivalent to the conjecture that

$$\begin{aligned}
&\frac{1}{T}\int_0^T\sum_{\gamma,\gamma'}\rho(\gamma-t)\rho(\gamma'-t)e\big(\alpha\tfrac{\log T}{2\pi}(\gamma-t)-\alpha\tfrac{\log T}{2\pi}(\gamma'-t)\big)\,dt\\
&\quad=(1+o(1))\int_{\mathbb{R}^2}\rho\big(\tfrac{2\pi\nu_1}{\log T}\big)\rho\big(\tfrac{2\pi\nu_2}{\log T}\big)e\big(\alpha(\nu_1-\nu_2)\big)\\
&\qquad\qquad\times\Big[1+\delta(\nu_1-\nu_2)-\Big(\frac{\sin\pi(\nu_1-\nu_2)}{\pi(\nu_1-\nu_2)}\Big)^2\Big]\,d\nu_1 d\nu_2.
\end{aligned}$$

Here the δ function corresponds with those terms on the left in which $\gamma=\gamma'$ and both can be removed accordingly. In fact, this is equivalent to the claim that

$$\begin{aligned}
&\frac{1}{T}\int_0^T\sum_{\gamma\neq\gamma'}\rho(\gamma-t)\rho(\gamma'-t)e\big(\alpha_1\tfrac{\log T}{2\pi}(\gamma-t)+\alpha_2\tfrac{\log T}{2\pi}(\gamma'-t)\big)\,dt\\
&=o(1)+(1+o(1))\int_{\mathbb{R}^2}\rho\big(\tfrac{2\pi\nu_1}{\log T}\big)\rho\big(\tfrac{2\pi\nu_2}{\log T}\big)e(\alpha_1\nu_1+\alpha_2\nu_2)\Big[1-\Big(\frac{\sin\pi(\nu_1-\nu_2)}{\pi(\nu_1-\nu_2)}\Big)^2\Big]\,d\nu_1 d\nu_2.
\end{aligned}\tag{2.2}$$

uniformly for α_1 and α_2 in a fixed bounded region. This can be shown by modifying the previous argument, although we leave the (somewhat tedious and secondary) details to the reader.

Said somewhat more informally: because we may integrate in α_1 and α_2, the measures

$$\left[\frac{1}{T}\int_0^T \sum_{\gamma\neq\gamma'} \delta\big(\nu_1 - \tfrac{\log T}{2\pi}(\gamma - t)\big)\delta\big(\nu_2 - \tfrac{\log T}{2\pi}(\gamma' - t)\big)\right]\rho\big(\tfrac{2\pi\nu_1}{\log T}\big)\rho\big(\tfrac{2\pi\nu_2}{\log T}\big)\, d\nu_1 d\nu_2, \tag{2.3}$$

and

$$\left[1 - \Big(\frac{\sin\pi(\nu_1-\nu_2)}{\pi(\nu_1-\nu_2)}\Big)^2\right]\rho\big(\tfrac{2\pi\nu_1}{\log T}\big)\rho\big(\tfrac{2\pi\nu_2}{\log T}\big)\, d\nu_1 d\nu_2, \tag{2.4}$$

are asymptotically indistinguishable with respect to test functions that have a fixed compact Fourier support and therefore do not concentrate themselves too narrowly. This latter measure, of course, is the limiting pair correlation measure associated to the Gaussian Unitary Ensemble (GUE).

There is nothing special about our use of the function ρ here. Its placement serves only to cutoff measures (2.3) and (2.4) so they are (close to being) supported in a square region with dimensions of order $\log T$. This is an important feature of Montgomery's conjecture – outside of this region the random-matrix-theory measure given by (2.4) ceases to be as good an approximation to (2.3).

Even inside this region – to which we will restrict our attention in this paper – the measure (2.4) misses important phenomena that will be important if we are to achieve a stronger estimate than (2.2). These are the troughs near low-lying zeros in histograms of $\gamma - \gamma'$, seen in Figure 2.1.

Because the 1-level density of zeros near T is not stationary but grows like $\frac{\log(T/2\pi)}{2\pi}$, measure (2.3) will even more closely resemble measure (2.4) if the average from 0 to T is replaced by an average from T to $T+H$ for $H = o(T)$. We will do so in the sequel, and in addition, for technical reasons, we will use smoothed averages; we replace

$$\frac{1}{H}\int_T^{T+H} \cdots dt = \frac{1}{H}\int_{\mathbb{R}} \mathbf{1}_{[0,1]}\Big(\tfrac{t-T}{H}\Big)\cdots\, dt$$

by

$$\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \cdots \, dt,$$

where σ is some smooth bump function with mass 1. Making use of such smoothed averages makes the computations that follows easier than they would otherwise be.

We prove

Theorem 2.1.3. *For fixed σ and h with $\hat{\sigma}, \hat{h}$ smooth and compactly supported, and σ of mass 1; fixed $\epsilon, \kappa > 0$; L within a bounded distance κ of $\log T$; and $H \leq T$,*

$$\begin{aligned}
&\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \sum_{\gamma \neq \gamma'} h(\gamma - t, \gamma' - t) e\big(\alpha_1 \tfrac{L}{2\pi}(\gamma - t) + \alpha_2 \tfrac{L}{2\pi}(\gamma' - t)\big)\, dt \\
&= O\Big(\frac{T^{|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|}}{H}\Big) + O\Big(\log T\Big(\frac{H}{T} + \frac{1}{H}\Big)\Big) \qquad (2.5)\\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big[\Big(\frac{\log(T/2\pi)}{2\pi}\Big)^2 + Q_T(\nu_1 - \nu_2)\Big] h(\nu_1, \nu_2) e\big(\alpha_1 \tfrac{L}{2\pi}\nu_1 + \alpha_2 \tfrac{L}{2\pi}\nu_2\big)\, d\nu_1 d\nu_2
\end{aligned}$$

uniformly for $(|\alpha_1| + |\alpha_2|)/2 \leq 1 - \epsilon$.

Here

$$\begin{aligned}
Q_t(u) := &\frac{1}{4\pi^2}\bigg(\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) - B(iu) + \Big(\frac{\zeta'}{\zeta}\Big)'(1-iu) - B(-iu) \\
&+ \Big(\frac{t}{2\pi}\Big)^{-iu}\zeta(1-iu)\zeta(1+iu)A(iu) + \Big(\frac{t}{2\pi}\Big)^{iu}\zeta(1+iu)\zeta(1-iu)A(-iu)\bigg),
\end{aligned}$$

defined by continuity at $u = 0$, and

$$A(s) := \prod_p \frac{(1 - \frac{1}{p^{1+s}})(1 - \frac{2}{p} + \frac{1}{p^{1+s}})}{\big(1 - \frac{1}{p}\big)^2} = \prod_p \bigg(1 - \frac{(1-p^{-s})^2}{(p-1)^2}\bigg) = 1 + O(s^2),$$

and

$$B(s) := \sum_p \frac{\log^2 p}{(p^{1+s} - 1)^2}.$$

Remark: One may, of course, optimize by setting $H = T^{1-\epsilon/2}/\sqrt{\log T}$. The reason we have written our error terms in this manner is that the second error term $O(\log T(H/T + 1/H))$ is somewhat artificial. This will become clear in the proof; at the cost of a somewhat more baroque result it could be eliminated. The remaining error term is somewhat more fundamental – and at any rate both effectively amount to an inverse power-of-T error term.

Remark: Here and in what follows we adopt the convention of counting zeros with multiplicity, so that in particular for a function f, a sum

$$\sum_{\gamma \neq \gamma'} f(\gamma, \gamma')$$

is really

$$\sum_{\gamma, \gamma'} f(\gamma, \gamma') - \sum_{\gamma} f(\gamma, \gamma)$$

In all likelihood every zero of the zeta function occurs with multiplicity 1, but provided we adopt this notational convention we do not need to assume so.

Remark: This theorem is of interest mainly in the case that $\alpha_1 = -\alpha_2$, as in Conjecture 2.1.2. We do not specialize to this case only because in the computation of higher order correlation functions such a specialization ceases to be as natural.

A pair correlation function of this form was first conjectured by Bogomolny and Keating in [4], on part on an analogy from the field of quantum chaos. Further support for this form was offered by Conrey and Snaith [17], who showed that it can be derived from the ratio conjecture of Conrey, Farmer, and Zirnbauer [13]. It bears remarking that, in the semiclassical language, we shall only really see the diagonal terms $(\zeta'/\zeta)'(1 + iu) - B(iu)$ and conjugate because of the restrictions we place on α_1 and α_2. Indeed, apart from some analytic devices which mimic the large sieve, we recover these terms in much the same fashion as Bogomolny and

Keating. There are no new arithmetical ideas required; to rigorously extend the range in which α_1, α_2 lie and effectively see the off-diagonal terms that remain is more difficult and will almost certainly require a breakthrough.

By re-averaging Theorem 2.1.3 in T, and using a 1-level density estimate as before, we show that

Theorem 2.1.4. *For fixed $\epsilon > 0$ and fixed ω with a smooth and compactly supported Fourier transform,*

$$\begin{aligned}\frac{1}{T}\sum_{0<\gamma\neq\gamma'\leq T}&\omega(\gamma-\gamma')e\big(\alpha\tfrac{\log T}{2\pi}(\gamma-\gamma')\big) \\ &= O_\delta\Big(\frac{1}{T^\delta}\Big) + \int_{\mathbb{R}}\omega(u)e\big(\alpha\tfrac{\log T}{2\pi}u\big)\left[\frac{1}{T}\int_0^T \Big(\frac{\log(t/2\pi)}{2\pi}\Big)^2 + Q_t(u)\,dt\right]du\end{aligned} \tag{2.6}$$

for any $\delta < \epsilon/2$, uniformly for $|\alpha| < 1-\epsilon$.

In fact, by proceeding carefully in the analysis that follows, one can show this even for Montgomery's choice of test function $\omega(u) = w(u)$, owing to the rapid decrease of this function's Fourier transform. We leave this refined computation to the reader. That the integral on the right extends over all of $\mathbb{R}$, rather than simply $[-T, T]$ may seem surprising since from the left hand side we inherit only information about ω in the latter interval, but the decay of ω is in all cases sufficient that the difference between these two regions of integration is absorbed into the error term.

The sum over zeros in (2.6) is of the same form as that in the Strong Pair Correlation Conjecture's (2.1). (Should we push our analysis to include $\omega = w$, they would be the same sum.) Our result therefore refines Montgomery's original work. Indeed, one can show that the difference between this prediction and a somewhat naive extension of the GUE prediction in which

$$Q_t(u)$$

has been replaced by

$$K_t(u) = -\Big(\frac{\log(t/2\pi)}{2\pi}\Big)^2 K\Big(\frac{\log(t/2\pi)}{2\pi}u\Big),$$

where $K(u) = \big(\frac{\sin \pi u}{\pi u}\big)^2$, is larger than the error term. We conclude our paper with a demonstration of the difference between these two predictions.

For α not restricted so stringently: there is no reason not to believe that for any fixed compact region M, Theorem 2.1.4 is true for all $\alpha \in M$ for any $\delta < 1/2$.

Remark: By integrating in α in Theorem 2.1.4, one can see the statistics

$$\sum_{0<\gamma\neq\gamma'\leq T} f\big(\tfrac{\log T}{2\pi}(\gamma-\gamma'-s)\big)\omega(\gamma-\gamma'),$$

to a uniform error of $O(\|\hat{f}\|_{L^1}/T^\delta)$ for any s, where f is supported in $[-1+\epsilon, 1-\epsilon]$ and $\delta < \epsilon/2$. This is the precise sense in which one can see the troughs in the pair correlation measure. If one wants these statistics only for fixed s, without uniformity, the computations that follow can be simplified slightly.

Remark: The above statistics say nothing about the case that our macroscopic interval grows. For instance, they say nothing about the sums

$$\frac{1}{T}\sum_{0<\gamma\neq\gamma'\leq T} \omega\Big(\frac{\gamma-\gamma'}{T^\nu}\Big)e\big(\alpha\tfrac{\log T}{2\pi}(\gamma-\gamma')\big) \tag{2.7}$$

when $0 < \nu < 1$. The statistics over such intervals would be expected to deviate from GUE statistics in the same way as above, and bear relevance to the variance statistics first obtained by Fujii [26] for the number of zeros lying in an interval $[t, t+h]$, where $h \to \infty$ as $t \to \infty$.

In fact, using the method below, one can evaluate the sums (2.7) to a high degree of accuracy, at the cost of imposing additional band-limits on the Fourier variable α, beyond those of Theorems 2.1.3 and 2.1.4. However, we do not pursue the matter here.

2.2 The explicit formula and an outline of a proof

Our main tool in establishing this will be the well known explicit formula relating the zeros of the zeta-function to the primes, which we will make use of several times in this thesis.

We define

$$\Omega(\xi) := \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\big(\tfrac{1}{4} + i\tfrac{\xi}{2}\big) + \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\big(\tfrac{1}{4} - i\tfrac{\xi}{2}\big) - \log \pi.$$

Theorem 2.2.1 (The explicit formula)**.** *Let g a measurable function such that $g(x) = \frac{g(x+)+g(x-)}{2}$, and for some $\delta > 0$,*

$$(a) \qquad \int_{-\infty}^{\infty} e^{(\frac{1}{2}+\delta)|x|}|g(x)|dx < +\infty,$$

$$(b) \qquad \int_{-\infty}^{\infty} e^{(\frac{1}{2}+\delta)|x|}|dg(x)| < +\infty.$$

Then we have

$$\lim_{L\to\infty} \sum_{|\gamma|<L} \hat{g}\big(\tfrac{\gamma}{2\pi}\big) - \int_{-L}^{L} \frac{\Omega(\xi)}{2\pi}\hat{g}\big(\tfrac{\xi}{2\pi}\big)\, d\xi = \int_{-\infty}^{\infty} [g(x) + g(-x)]e^{-x/2}\, d\big(e^x - \psi(e^x)\big), \tag{2.8}$$

where here $\psi(x) = \sum_{n\le x} \Lambda(n)$, for the von Mangoldt function Λ.

The explicit formula in this generality is due to Weil, [81] but before him something very much like it was written down in varying degrees by Riemann [59], and Guinand [31]. Note that if

$$\tilde{d}(\xi) = \sum_{\gamma} \delta_\gamma(\xi),$$

the left hand side of the explicit formula becomes the principal value integral

$$\int_{-\infty}^{\infty} \hat{g}\left(\tfrac{\xi}{2\pi}\right)\left[\tilde{d}(\xi) - \tfrac{\Omega(\xi)}{2\pi}\right] d\xi.$$

(In what follows we will be working with nice enough functions that it will not matter that this is a principal value integral.) If we define $S(T)$ in the standard way (see [56] pp. 452) so that

$$N(T) = \frac{1}{\pi} \arg \Gamma\left(\tfrac{1}{4} + i\tfrac{T}{2}\right) - \tfrac{T}{2\pi} \log \pi + 1 + S(T),$$

then (only on the Riemann Hypothesis)

$$\left[\tilde{d}(\xi) - \tfrac{\Omega(\xi)}{2\pi}\right] d\xi = dS(\xi)$$

This is, if nothing else, a notational convenience, and we will make use of it for that reason in what follows.

$S(T)$ is relatively small and oscillatory, so that $\frac{\Omega(\xi)}{2\pi}$ is an expression for the mean density of zeros around ξ. By Stirling's formula,

$$\frac{\Omega(\xi)}{2\pi} = \frac{\log\left((|\xi| + 2)/2\pi\right)}{2\pi} + O\Big(\frac{1}{|\xi| + 2}\Big).$$

It is therefore clear in this formulation that the explicit formula expresses a Fourier duality between the error term in the prime number theorem and the error term of the zero-counting function.

The explicit formula is proven by a simple contour integration argument, making use of the the reflection formula to evaluate one-half of the contour. (For the outline of a proof, with the final result stated slightly differently, see [43] pp. 108 or [56] pp. 410.) Note that here we have assumed the Riemann hypothesis, so that γ is always real, but even if the Riemann hypothesis is false, the identity

(2.8) remains true, for $1/2 + i\gamma$ labeling the nontrivial zeros of the zeta function, and $\hat{g}(\gamma/2\pi)$ interpreted harmonically in the case that γ has imaginary part.

We will also need another result which makes an appearance in [56] – in fact it is what is used by the authors to state the explicit formula differently. This is

Lemma 2.2.2. *Let $a > 0$ and $b > 0$ be fixed. If $J \in L^1(\mathbb{R})$, J is of bounded variation on $\mathbb{R}$, and if $J(x) = J(0) + O(|x|)$, then*

$$\lim_{T\to\infty} \int_{-T}^{T} \tfrac{\Gamma'}{\Gamma}(a \pm ibt)\hat{J}(t)dt = \tfrac{\Gamma'}{\Gamma}(a)J(0) + \int_0^\infty \frac{e^{-ay}}{1-e^{-y}}\Big[J(0) - J\Big(\mp \tfrac{by}{2\pi}\Big)\Big]dx.$$

Proof. See [56] pp. 414. □

An outline of a proof of Theorem 2.1.3: This outline will be expanded upon in section 4. We prove our result with a series of computational Lemmas. To simplify exposition, we will deal only with the case that $h(\nu_1, \nu_2) = r(\nu_1/2\pi)r(\nu_2/2\pi)$ for $\hat{r}$ smooth and compactly supported. The proof in general is identical, but writing h as a product of two functions draws some parts of the proof into clearer light. We dilate the functions r for the sake of tidying up some of the formulas that follow.

Because $\big[\tilde{d}(\xi) - \frac{\Omega(\xi)}{2\pi}\big]d\xi = dS(\xi)$, and because we do not expect that $\Omega(\xi)/2\pi$ to correlate at a local scale with $\tilde{d}(\xi)$, to understand the pair correlation statistics $\tilde{d}(\xi_1+t)\tilde{d}(\xi_2+t)$ it is enough to understand those of $dS(\xi_1+t)dS(\xi_2+t)$. In fact, in many ways, the statistics of the latter are a more natural quantity to consider. The second error term in Theorem 2.1.3 alluded to earlier is due to the passage back to the former.

It will follow from computation with the explicit formula that the smoothed average

$$\Big\langle \int\int_{\mathbb{R}^2} r\Big(\frac{\xi_1}{2\pi}\Big)r\Big(\frac{\xi_2}{2\pi}\Big)e\big(\alpha_1\tfrac{L}{2\pi}\xi_1 + \alpha_2\tfrac{L}{2\pi}\xi_2\big)dS(\xi_1+t)dS(\xi_2+t)\Big\rangle_{t\in[T,T+H]}$$

is very close to

$$\sum_{\log n - \log m = O(1/T)} \big(\hat{r}(-\log n - \alpha_1 L)\hat{r}(\log m - \alpha_2 L) \\ + \hat{r}(\log n - \alpha_1 L)\hat{r}(-\log m - \alpha_2 L)\big)\frac{\Lambda^2(n)}{n}$$

$$= \sum_{n=1}^{\infty} \big(\hat{r}(-\log n - \alpha_1 L)\hat{r}(\log n - \alpha_2 L) \\ + \hat{r}(\log n - \alpha_1 L)\hat{r}(-\log n - \alpha_2 L)\big)\frac{\Lambda^2(n)}{n}$$

as in passing from the first line both n and m must be less than T, for $\alpha_1, \alpha_2 \leq 1$ by the compact support of $\hat{r}$. (Recall L is near $\log T$.) These are the terms inherited from the measure $d\psi(e^x)$ in the explicit formula; those terms which are inherited from $d(e^x)$ drop out, roughly speaking, because they have no discrete part. These ideas date back to Montgomery's original proof in [53]. We are able to obtain slightly better error terms only by using smoothed averages. This suppresses the appearance of any large-sieve-type inequalities. This is the content of Lemma 2.3.1.

One can already see the diagonal terms in this, but care is needed to do so rigorously. This is done in Lemma 2.3.2.

We do not recover the off-diagonal terms in this way, only a proxy for them, but we show in Lemma 2.3.3 that for the restricted class of test functions we consider there is no difference between the two.

Finally, in Lemmas 2.3.5 and 2.3.6 we show that our intuition about being able to recover $\tilde{d}(\xi_1 + t)\tilde{d}(\xi_2 + t)$ from $dS(\xi_1 + t)dS(\xi_2 + t)$ is correct.

2.3 Lemmata

Lemma 2.3.1. *For fixed r and σ, with $\hat{r}$ and $\hat{\sigma}$ smooth and compactly supported and σ of mass 1, we have*

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r\Big(\tfrac{\xi_1-t}{2\pi}\Big)r\Big(\tfrac{\xi_2-t}{2\pi}\Big)e\Big(\alpha_1\tfrac{L}{2\pi}(\xi_1-t)+\alpha_2\tfrac{L}{2\pi}(\xi_2-t)\Big)\,dS(\xi_1)dS(\xi_2)\,dt$$
$$=\sum_{n=1}^{\infty}\big[\hat{r}(-\log n-\alpha_1 L)\hat{r}(\log n-\alpha_2 L)+\hat{r}(\log n-\alpha_1 L)\hat{r}(-\log n-\alpha_2 L)\big]\frac{\Lambda^2(n)}{n} \tag{2.9}$$
$$+\,O\bigg(\frac{e^{(|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|)L}}{H}\bigg).$$

Here the implied constant depends upon r and σ.

Proof. Note that $\hat{g}(\frac{\xi}{2\pi}) = r\big(\frac{\xi-t}{2\pi}\big)e\big(\alpha\frac{L}{2\pi}(\xi-t)\big)$ when $g(x) = e\big(\frac{xt}{2\pi}\big)\hat{r}(-(x+\alpha L))$. As this function has compact support, we are justified in using the explicit formula to evaluate the left hand side of (2.9). It is equal to

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\sum_{\varepsilon\in\{-1,1\}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e(-\tfrac{t}{2\pi}(\varepsilon_1x_1+\varepsilon_2x_2))\hat{r}(\varepsilon_1x_1-\alpha_1L)$$
$$\times\,\hat{r}(\varepsilon_2x_2-\alpha_2L)e^{-(x_1+x_2)/2}d\big(e^{x_1}-\psi(e^{x_1})\big)d\big(e^{x_2}-\psi(e^{x_2})\big)$$
$$=\sum_{\varepsilon\in\{-1,1\}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e(-\tfrac{T}{2\pi}(\varepsilon_1x_1+\varepsilon_2x_2))\hat{\sigma}\big(\tfrac{H}{2\pi}(\varepsilon_1x_1+\varepsilon_2x_2)\big)\hat{r}(\varepsilon_1x_1-\alpha_1L) \tag{2.10}$$
$$\times\,\hat{r}(\varepsilon_2x_2-\alpha_2L)e^{-(x_1+x_2)/2}d\big(e^{x_1}-\psi(e^{x_1})\big)d\big(e^{x_2}-\psi(e^{x_2})\big),$$

where the interchange of integrals is justified by Fubini's theorem.

Note that, for instance,

$$\int_{-\infty}^{\infty} \hat{\sigma}\big(\tfrac{H}{2\pi}(\varepsilon_1 x_1 + \varepsilon_2 x_2)\big)\hat{r}(\varepsilon_1 x_1 - \alpha_1 L)\hat{r}(\varepsilon_2 x_2 - \alpha_2 L)e^{-(x_1+x_2)/2} d(e^{x_1})$$
$$\lesssim_{r,\sigma} \hat{r}(\varepsilon_2 x_2 - \alpha_2 L)e^{-x_2/2}\frac{e^{\min(|\alpha_1|L, x_2)/2}}{H},$$

and if $M(x)$ is either of the functions x or $\psi(x)$,

$$\int_{-\infty}^{\infty} \hat{r}\big((\varepsilon_2 x_2 - \alpha_2 L)\big)e^{-x_2/2}e^{|\alpha_1|L/2} dM(e^{x_2}) \lesssim_r e^{(\frac{|\alpha_1|+|\alpha_2|}{2})L}.$$

In this way, expanding the measure

$$d\big(e^{x_1} - \psi(e^{x_1})\big)d\big(e^{x_2} - \psi(e^{x_2})\big)$$
$$= d(e^{x_1})d(e^{x_2}) - d(x^{x_2})d\psi(e^{x_1}) - d(e^{x_1})d\psi(e^{x_2}) + d\psi(x^{x_1})d\psi(e^{x_2}),$$

the integrand in (2.9) integrated with respect to all terms but $d\psi(e^{x_1})d\psi(e^{x_2})$ is bound by

$$O\bigg(\frac{e^{(\frac{\alpha_1}{2}+\frac{\alpha_2}{2})L}}{H}\bigg),$$

for any $\varepsilon \in \{-1, 1\}^2$.

On the other hand,

$$\sum_{\varepsilon\in\{-1,1\}^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e(-\tfrac{T}{2\pi}(\varepsilon_1 x_1 + \varepsilon_2 x_2))\hat{\sigma}\big(\tfrac{H}{2\pi}(\varepsilon_1 x_1 + \varepsilon_2 x_2)\big)\hat{r}(\varepsilon_1 x_1 - \alpha_1 L)$$
$$\times \hat{r}(\varepsilon_2 x_2 - \alpha_2 L)e^{-(x_1+x_2)/2} d\psi(e^{x_1})d\psi(e^{x_2})$$
$$= \sum_{\varepsilon\in\{-1,1\}^2}\sum_{n_1,n_2=1}^{\infty} e(-\tfrac{T}{2\pi}(\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2))\hat{\sigma}\big(\tfrac{H}{2\pi}(\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2)\big)$$
$$\times \hat{r}(\varepsilon_1 \log n_1 - \alpha_1 L)\hat{r}(\varepsilon_2 \log n_2 - \alpha_2 L)\frac{\Lambda(n_1)\Lambda(n_2)}{\sqrt{n_1 n_2}}$$

We will consider each $\varepsilon \in \{-1, 1\}^2$ in turn. That $\hat{\sigma}$ is compactly supported implies that we may restrict our sum to those (n_1, n_2) with $\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2 = O\big(\frac{1}{H}\big)$. But that $\hat{r}$ is compactly supported implies that we may restrict our sum to those $n_1 = \exp[\varepsilon_1\alpha_1 L + O(1)], n_2 = \exp[\varepsilon_2\alpha_2 L + O(1)]$. Since n_1, n_2 are positive integers, for sufficiently large H the terms with $\varepsilon_1 = \varepsilon_2 = 1$ or $\varepsilon_1 = \varepsilon_2 = -1$ will be null. On the other hand, for $\varepsilon_1 = 1, \varepsilon_2 = -1$ or $\varepsilon_1 = -1, \varepsilon_2 = 1$, if $n_1 \neq n_2$, we have $|n_1 - n_2| \geq 1$, so

$$|\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2| \gtrsim \frac{1}{\sqrt{n_1 n_2}},$$

and so if $H/e^{(|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|)L}$ is sufficiently large, the only n_1, n_2 for which $\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2 = O\big(\frac{1}{H}\big)$ are those for which $n_1 = n_2$ so that $\varepsilon_1 \log n_1 + \varepsilon_2 \log n_2 = 0$. This gives (2.9), as the left hand side is always $O(1)$, to cover the case that $e^{(|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|)L}/H$ is large. □

It is easy to formally discern the term

$$\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) - B(iu) = \sum \frac{\log n\, \Lambda(n)}{n^{1+iu}} - \sum_p \frac{\log^2 p}{(p^{1+iu}-1)^2} = \sum \frac{\Lambda^2(n)}{n^{1+iu}}$$

here, but some care is needed to deal with issues of convergence, near $u = 0$ especially. One can either make use of the explicit formula, or in some manner reprove it, and we follow the first path.

Lemma 2.3.2. *For $\min(|\alpha_1|, |\alpha_2|) \leq 1 - \epsilon$ and r with $\hat{r}$ smooth and compactly supported, and for $\lambda - (1-\epsilon)L$ sufficiently large (where 'sufficiently large' depends only on ϵ and the region in which $\hat{r}$ can be supported),*

$$\sum_{n=1}^{\infty} \big[\hat{r}(-\log n - \alpha_1 L)\hat{r}(\log n - \alpha_2 L) + \hat{r}(\log n - \alpha_1 L)\hat{r}(-\log n - \alpha_2 L)\big] \frac{\log(n)\Lambda(n)}{n}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2)\Big[\lambda\delta(\nu_1-\nu_2) + \Big(\frac{\zeta'}{\zeta}\Big)'\big(1+i2\pi(\nu_1-\nu_2)\big)$$
$$+ \Big(\frac{\zeta'}{\zeta}\Big)'\big(1-i2\pi(\nu_1-\nu_2)\big) + \frac{e((\nu_1-\nu_2)\lambda) + e(-(\nu_1-\nu_2)\lambda)}{(2\pi(\nu_1-\nu_2))^2}\Big] d\nu_1 d\nu_2 \qquad (2.11)$$

Remark: It may seem odd that we have such freedom in choosing λ in this Lemma. We are able to do so because our restrictions on α_1 and α_2 mean that we will not see a large part of the measure against which we integrate $r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2)$. We will eventually want $\lambda = \log(T/2\pi)$ and $L = \log(T) + O(1)$, and it is worthwhile to keep this in mind.

Proof. The left hand side of (2.11) is

$$\int_{-\infty}^{\infty} \left[\hat{r}(-x-\alpha_1 L)\hat{r}(x-\alpha_2 L)\tfrac{|x|}{e^{|x|/2}} + \hat{r}(x-\alpha_1 L)\hat{r}(-x-\alpha_2 L)\tfrac{|x|}{e^{|x|/2}}\right] e^{-x/2} d\psi(e^x).$$

For $q(x) = \hat{r}(-x-\alpha_1 L)\hat{r}(x-\alpha_2 L)\frac{|x|}{e^{|x|/2}}$,

$$\begin{aligned}\hat{q}(\xi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} & r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2) \\ & \cdot \left(\frac{1}{\left(\frac{1}{2}+i2\pi(\nu_1-\nu_2-\xi)\right)^2} + \frac{1}{\left(\frac{1}{2}-i2\pi(\nu_1-\nu_2-\xi)\right)^2}\right) d\nu_1 d\nu_2\end{aligned}$$

so that by the explict formula the left hand side of (2.11) is

$$\begin{aligned}&\int_{-\infty}^{\infty} \left[\hat{r}(-x-\alpha_1 L)\hat{r}(x-\alpha_2 L) + \hat{r}(x-\alpha_1 L)\hat{r}(-x-\alpha_2 L)\right]|x|e^{-(\frac{x}{2}+\frac{|x|}{2})} d(e^x) \\ &\quad - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2) \\ &\quad\quad \times \left(\frac{1}{\left(\frac{1}{2}+i2\pi(\nu_1-\nu_2-\xi)\right)^2} + \frac{1}{\left(\frac{1}{2}-i2\pi(\nu_1-\nu_2-\xi)\right)^2}\right) d\nu_1 d\nu_2\, dS(\xi)\end{aligned}$$

The first of these terms is

$$\int_{-\infty}^{\infty} |x|\hat{r}(-x-\alpha_1 L)\hat{r}(x-\alpha_2 L)\, dx + \int_{-\infty}^{\infty} |x|e^{-|x|}\hat{r}(-x-\alpha_1 L)\hat{r}(x-\alpha_2 L)\, dx,$$

and because $\hat{r}$ is compactly supported and one of $|\alpha_1|$, $|\alpha_2|$ is no greater than $1-\epsilon$, for sufficiently large $(\lambda - (1-\epsilon)L)$ the first integral is equal to

$$\int_{-\infty}^{\infty} \lambda \cdot \Big[1 - \big(1 - |\tfrac{x}{\lambda}|\big)_+\Big] \hat{r}(-x - \alpha_1 L)\hat{r}(x - \alpha_2 L)\, dx$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2)\big[\lambda\delta(\nu_1-\nu_2) - \lambda^2\Big(\frac{\sin \pi\lambda(\nu_1-\nu_2)}{\pi\lambda(\nu_1-\nu_2)}\Big)^2\big]\, d\nu_1 d\nu_2.$$

It is worth remarking that we can make this transition only because of the restricted range of α_1, α_2.

Evaluation of the second integral is routine, and we have that the left hand side of (2.11) is

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2) \\
&\quad\times \Bigg[\lambda\delta(\nu_1-\nu_2) - \lambda^2\Big(\frac{\sin\pi(\nu_1-\nu_2)}{\pi(\nu_1-\nu_2)}\Big)^2 \\
&\quad + \Bigg(\frac{1}{\big(1+i2\pi(\nu_1-\nu_2)\big)^2} + \frac{1}{\big(1-i2\pi(\nu_1-\nu_2)\big)^2}\Bigg) \\
&\quad - \int_{-\infty}^{\infty}\Bigg(\frac{1}{\big(\frac12 + i2\pi(\nu_1-\nu_2-\xi)\big)^2} + \frac{1}{\big(\frac12 - i2\pi(\nu_1-\nu_2-\xi)\big)^2}\Bigg)dS(\xi)\Bigg]\, d\nu_1 d\nu_2.
\end{aligned} \tag{2.12}$$

Note that for s real, by Lemma 2.2.2 (or contour integration),

$$\int_{-\infty}^{\infty}\Bigg(\frac{1}{\big(\frac12 + i(s-\xi)\big)^2} + \frac{1}{\big(\frac12 - i(s-\xi)\big)^2}\Bigg)\frac{\Omega(\xi)}{2\pi}\, d\xi$$
$$= -\frac{1}{4}\Big((\tfrac{\Gamma'}{\Gamma})'\big(\tfrac12(1+is)\big) + (\tfrac{\Gamma'}{\Gamma})'\big(\tfrac12(1-is)\big)\Big)$$

Recall that for any s, (differentiating the classical representation 2.12.7 in [79]),

$$\Big(\frac{\zeta'}{\zeta}\Big)'(s) = \frac{1}{(s-1)^2} + \frac{1}{s^2} - \frac{1}{4}(\tfrac{\Gamma'}{\Gamma})'\big(\tfrac12 s\big) - \sum_{\gamma}\frac{1}{\big(s - (\frac12 + i\gamma)\big)^2}, \tag{2.13}$$

On expanding $\left(\frac{\sin u}{u}\right)^2 = \frac{2}{(2u)^2} - \frac{e^{i2u}+e^{-i2u}}{(2u)^2}$, we see the expression (2.12) is just

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\nu_1)r(\nu_2)e(\alpha_1 L\nu_1 + \alpha_2 L\nu_2)\Big[\lambda\delta(\nu_1-\nu_2) + \Big(\frac{\zeta'}{\zeta}\Big)'\big(1+i2\pi(\nu_1-\nu_2)\big) + \Big(\frac{\zeta'}{\zeta}\Big)'\big(1-i2\pi(\nu_1-\nu_2)\big) + \frac{e((\nu_1-\nu_2)\lambda)+e(-(\nu_1-\nu_2)\lambda)}{(2\pi(\nu_1-\nu_2))^2}\Big]d\nu_1 d\nu_2$$

□

We need that the prediction in (2.11) does not differ so much from that in Theorem 2.1.3:

Lemma 2.3.3. *For h with $\hat{h}$ smooth and compactly supported, so long as $\lambda \geq (\frac{|\alpha_1|}{2}+\frac{|\alpha_2|}{2})L$,*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\nu_1,\nu_2)e(\alpha_1 L\nu_1+\alpha_2 L\nu_2)\Big[-\frac{e(-(\nu_1-\nu_2)\lambda)}{(2\pi(\nu_1-\nu_2))^2} + e(-(\nu_1-\nu_2)\lambda)\zeta(1-i2\pi(\nu_1-\nu_2))\zeta(1+i2\pi(\nu_1-\nu_2))A(i2\pi(\nu_1-\nu_2))\Big]d\nu_1 d\nu_2$$
$$= O_h\Big(\frac{1}{e^{\lambda-(\frac{|\alpha_1|}{2}+\frac{|\alpha_2|}{2})L}}\Big).$$

We note that by symmetry, the integral of $h(\nu_1,\nu_2)e(\alpha_1 L\nu_1+\alpha_2 L\nu_2)$ with respect to the conjugate measure will be similarly bounded. And of course, for $\lambda < (\frac{|\alpha_1|}{2}+\frac{|\alpha_2|}{2})L$ this estimate is true, but trivially.

Plainly, to prove this we need only show that for f Schwartz on $\mathbb{R}$ with compact Fourier support, and $P \geq 0$,

$$\int_{-\infty}^{\infty} f(u)e(-Pu)\Big(-\frac{1}{4\pi^2u^2}+\zeta(1-i2\pi u)\zeta(1+i2\pi u)A(i2\pi u)\Big)du = O_f\Big(\frac{1}{e^P}\Big). \tag{2.14}$$

We will use the formula

Proposition 2.3.4. *For f Schwartz on $\mathbb{R}$,*

$$\int_{-\infty}^{\infty} f(u)\Big(-\frac{1}{4\pi^2 u^2} + \zeta(1-i2\pi u)\zeta(1+i2\pi u)\Big)\, du$$
$$= \hat{f}(0) + \int_1^\infty \int_0^\infty (\hat{f}+\hat{f}'')\Big(\log\big(\tfrac{y}{x}\big)\Big)\cdot \big[\mathbf{1}_{[0,1]}(x)\cdot xy - \lfloor y \rfloor (x - \lfloor x \rfloor)\big]\,\frac{dx}{x^2}\frac{dy}{y^2}$$

Proof of Proposition 2.3.4. Note that

$$\frac{\zeta(s)}{s} = \int_1^\infty \frac{\lfloor x \rfloor}{x^{s+1}}dx \quad \text{for } \Re s > 1,$$

$$\frac{\zeta(s)}{s} = -\int_0^\infty \frac{x - \lfloor x \rfloor}{x^{s+1}}dx \quad \text{for } 0 < \Re s < 1,$$
$$\frac{1}{s-1} = \int_1^\infty \frac{x}{x^{s+1}}dx \quad \text{for } \Re s > 1,$$
$$\frac{1}{s-1} = -\int_0^1 \frac{x}{x^{s+1}}dx \quad \text{for } 0 < \Re s < 1.$$

Therefore for $\epsilon > 0$ and u real,

$$\frac{\zeta(1+\epsilon+i2\pi u)\zeta(1-\epsilon-i2\pi u)}{1-(\epsilon+i2\pi u)^2} - \frac{1}{-\epsilon-i2\pi u}\frac{1}{\epsilon+i2\pi u}$$
$$= \int_1^\infty \int_0^\infty \big[\mathbf{1}_{[0,1]}(x)\cdot xy - \lfloor y \rfloor (x - \lfloor x \rfloor)\big]\Big(\frac{x}{y}\Big)^{\epsilon+i2\pi u}\frac{dx}{x^2}\frac{dy}{y^2}$$

This expression will remain bounded for u near 0 as $\epsilon \to 0^+$. If we fix ϵ and integrate in u with respect to $f(u)du$, the the bottom line becomes

$$\int_1^\infty \int_0^\infty \hat{f}\Big(\log\big(\tfrac{y}{x}\big)\Big)\cdot\Big(\frac{x}{y}\Big)^{\epsilon}\big[\mathbf{1}_{[0,1]}(x)\cdot xy - \lfloor y \rfloor (x - \lfloor x \rfloor)\big]\,\frac{dx}{x^2}\frac{dy}{y^2},$$

where we are justified in changing the order of integration by a trivial application of Fubini's theorem. Now letting $\epsilon \to 0^+$, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} f(u)\Big(- \frac{1}{4\pi^2 u^2} + \frac{\zeta(1-i2\pi u)\zeta(1+i2\pi u)}{1+4\pi^2u^2}\Big)\, du \\
&= \int_1^{\infty}\int_0^{\infty} \hat{f}\Big(\log\big(\tfrac{y}{x}\big)\Big) \cdot \big[\mathbf{1}_{[0,1]}(x)\cdot xy - \lfloor y \rfloor (x - \lfloor x \rfloor)\big]\, \frac{dx}{x^2}\frac{dy}{y^2}.
\end{aligned}$$

Replacing $f(u)$ with $(1+4\pi^2u^2)f(u)$ completes the proof. □

Proof of Lemma 2.3.3. Returning to Lemma 2.3.3, we note that

$$A(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{\phi(n)^2} \sum_{d,\delta | n} \mu(d)\mu(\delta)(d\delta)^{-s}.$$

For f of compact Fourier support, and P a sufficiently large positive number

$$\begin{aligned}
&\int_{-\infty}^{\infty} f(u)e(-Pu)\Big(- \frac{1}{4\pi^2 u^2} + \zeta(1-i2\pi u)\zeta(1+i2\pi u)\Big)\, du \\
&= \hat{f}(P) + \int_1^{\infty}\int_0^{\infty} [\hat{f} + \hat{f}''] \big(\log\big(\tfrac{y}{x}\big) + P\big) \cdot \big[\mathbf{1}_{[0,1]}(x)\cdot xy - \lfloor y \rfloor (x - \lfloor x \rfloor)\big]\, \frac{dx}{x^2}\frac{dy}{y^2} \\
&= O_f\Big(\int_1^{\infty} \frac{1}{e^p y}\frac{\lfloor y \rfloor}{y^2} dy \Big) \\
&= O_f\Big(\frac{1}{e^P}\Big),
\end{aligned}$$

since, going from the second line to the third, $\log(x) = \log(y) + P + O(1)$, otherwise $[\hat{f} + \hat{f}''](\log(y/x) + P)$ is null. Hence,

$$\begin{aligned}
&\int_{-\infty}^{\infty} f(u)A(i2\pi u)e(-Pu)\Big(- \frac{1}{4\pi^2 u^2} + \zeta(1-i2\pi u)\zeta(1+i2\pi u)\Big)\, du \\
&\lesssim_f \frac{1}{e^P} \sum_{n=1}^{\infty} \frac{\mu(n)}{\phi(n)^2} \sum_{d,\delta|n} \frac{\mu(d)\mu(\delta)}{d\delta} \qquad (2.15)\\
&\lesssim_f \frac{1}{e^P}
\end{aligned}$$

Finally,

$$\int_{-\infty}^{\infty} f(u)(A(i2\pi u)-1)e(-xu)du = \sum_{n>1}\frac{\mu(n)}{\phi(n)^2}\sum_{d,\delta|n}\mu(d)\mu(\delta)\hat{f}(x+\log(d\delta))$$
$$= 0$$

for sufficiently large positive x. Hence

$$\begin{aligned}
&\int_{-\infty}^{\infty}\frac{f(u)(A(i2\pi u)-1)}{(2\pi u)^2}e(-Pu)\,du \\
&= \int_0^{\infty}\int_x^{\infty}\int_{-\infty}^{\infty} f(u)(A(i2\pi u)-1)e(-Pu-yu)\,du\,dy\,dx \\
&= 0,
\end{aligned}$$

for sufficiently large P. Combining this with (2.15) proves the lemma. □

These lemmas give us very accurate information about the statistics of $dS(\xi_1)dS(\xi_2)$. Our last two lemmas will allow us to unravel this product into the pair correlation measure we are after.

Lemma 2.3.5. *For r and σ with $\hat{r}, \hat{\sigma}$ smooth and compactly supported,*

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}h\Big(\tfrac{\xi-t}{2\pi}\Big)e\big(\alpha\tfrac{L}{2\pi}(\xi-t)\big)\,dS(\xi)\,dt = O_{h,\sigma}\Big(\frac{1}{H}\Big),$$

uniformly in L.

Remark: This is in essence a 1-level density estimate.

Proof. We have by the explicit formula,

$$\begin{aligned}
&\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \int_{-\infty}^{\infty} h\Big(\tfrac{\xi-t}{2\pi}\Big) e\big(\alpha \tfrac{L}{2\pi}(\xi - t)\big)\, dS(\xi)\, dt \\
&= \frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \int_{-\infty}^{\infty} \big[e\big(-\tfrac{xt}{2\pi}\big)\hat{h}(x-\alpha L) + e\big(\tfrac{xt}{2\pi}\big)\hat{h}(-(x+\alpha L))\big] e^{-x/2}\, d\big(e^x - \psi(e^x)\big) \\
&\lesssim_h \int_{-\infty}^{\infty} \hat{\sigma}\big(\tfrac{Hx}{2\pi}\big) e^{-x/2}\, d\big(e^x - \psi(e^x)\big) \\
&\lesssim_{h,\sigma} \frac{1}{H}.
\end{aligned}$$

□

Lemma 2.3.6. *For r and σ with $\hat{r}, \hat{\sigma}$ smooth and compactly supported,*

$$\begin{aligned}
&\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(\xi_1 - t) r(\xi_2 - t) e\big(\alpha_1 \tfrac{L}{2\pi}(\xi_1 - t) + \alpha_2 \tfrac{L}{2\pi}(\xi_2 - t)\big) dS(\xi_1)\, \frac{\Omega(\xi_2)}{2\pi}\, d\xi_2\, dt \\
&\qquad = O_{r,\sigma}\Big(\frac{e^{|\frac{\alpha_1}{2}|L}}{H}\Big)
\end{aligned}$$

Proof. By Lemma 2.2.2,

$$\begin{aligned}
&\int_{-\infty}^{\infty} r\Big(\tfrac{\xi_2 - t}{2\pi}\Big) e\big(\alpha_2 \tfrac{L}{2\pi}(\xi_2 - t)\big) \frac{\Omega(\xi_2)}{2\pi}\, d\xi_2 \\
&= \Big(\tfrac{\Gamma'}{\Gamma}\Big(\tfrac{1}{4}\Big) - \log \pi\Big) \hat{r}(-\alpha_2 L) \\
&\quad + \int_0^{\infty} \frac{e^{-y/4}}{1 - e^{-y}} \big(\hat{r}(-\alpha_2 L) - \tfrac{1}{2}\big[e\big(\tfrac{yt}{2\pi}\big)\hat{r}(-(y + \alpha_2 L)) + e\big(-\tfrac{yt}{2\pi}\big)\hat{r}(y - \alpha_2 L)\big]\big)\, dy.
\end{aligned}$$

We may use Lemma 2.3.5 to deal with the terms attached to $\hat{r}(-\alpha_2 L)$, and employing the explicit formula to deal with the terms that remain,

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}r(\xi_1-t)r(\xi_2-t)e\Big(\alpha_1\tfrac{L}{2\pi}(\xi_1-t)+\alpha_2\tfrac{L}{2\pi}(\xi_2-t)\Big)dS(\xi_1)\,\frac{\Omega(\xi_2)}{2\pi}\,d\xi_2\,dt$$
$$=O_{r,\sigma}\Big(\frac{1}{H}\Big)-\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\cdot\tfrac{1}{2}\sum_{\varepsilon}\int_0^{\infty}\frac{e^{-y/4}}{1-e^{-y}}\int_{-\infty}^{\infty}e\Big(-\tfrac{t}{2\pi}(\varepsilon_1x+\varepsilon_2y)\Big)$$
$$\times\,\hat{r}(\varepsilon_1x-\alpha_1L)\hat{r}(\varepsilon_2y-\alpha_2L)e^{-x/2}\,d\Big(e^x-\psi(e^x)\Big)\,dy\,dt$$
$$=O\bigg(\bigg|\sum_{\varepsilon}\int_0^{\infty}\frac{e^{-y/4}}{1-e^{-y}}\int_{-\infty}^{\infty}\hat{\sigma}\Big(\tfrac{H}{2\pi}(\varepsilon_1x+\varepsilon_2y)\Big)\hat{r}(\varepsilon_1x-\alpha_1L)\hat{r}(\varepsilon_2y-\alpha_2L)$$
$$\times\,e^{-x/2}\,d\Big(e^x-\psi(e^x)\Big)dy\bigg|\bigg)$$
$$=O_{r,\sigma}\Big(\frac{e^{|\frac{\alpha_1}{2}|L}}{H}\Big).$$

□

2.4 Proof of Theorem 2.1.3 and Theorem 2.1.4

Proof of Theorem 2.1.3. We set $\lambda=\log(T/2\pi)/2\pi$. Using Lemmas 2.3.1, 2.3.2, and 2.3.3 we have

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}r\Big(\tfrac{\xi_1-t}{2\pi}\Big)r\Big(\tfrac{\xi_2-t}{2\pi}\Big)e\Big(\alpha_1\tfrac{L}{2\pi}(\xi_1-t)+\alpha_2\tfrac{L}{2\pi}(\xi_2-t)\Big)\,dS(\xi_1)dS(\xi_2)\,dt$$
$$=O\Big(\frac{T^{|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|}}{H}\Big)+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[\frac{\log(T/2\pi)}{2\pi}\delta(\nu_1-\nu_2)+(2\pi)^2Q_T(2\pi(\nu_1-\nu_2))\Big]\tag{2.16}$$
$$\times\,r(\nu_1)r(\nu_2)e(\alpha_1L\nu_1+\alpha_2L\nu_2)\,d\nu_1d\nu_2.$$

On the other hand,

$$\Big(\tilde{d}(\xi_1)-\frac{\Omega(\xi_1)}{2\pi}\Big)\Big(\tilde{d}(\xi_2)-\frac{\Omega(\xi_2)}{2\pi}\Big)$$
$$=\tilde{d}(\xi_1)\tilde{d}(\xi_2)-\Big(\tilde{d}(\xi_1)-\frac{\Omega(\xi_1)}{2\pi}\Big)\cdot\frac{\Omega(\xi_2)}{2\pi}-\Big(\tilde{d}(\xi_2)-\frac{\Omega(\xi_2)}{2\pi}\Big)\cdot\frac{\Omega(\xi_1)}{2\pi}-\frac{\Omega(\xi_1)}{2\pi}\frac{\Omega(\xi_2)}{2\pi},$$

so by Lemma 11, the left hand side of (2.16) is

$$O\Big(\frac{T^{|\frac{\alpha_1}{2}|}+T^{|\frac{\alpha_2}{2}|}}{H}\Big)+\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}r\Big(\tfrac{\xi_1-t}{2\pi}\Big)r\Big(\tfrac{\xi_2-t}{2\pi}\Big)$$
$$\times\, e\big(\alpha_1\tfrac{L}{2\pi}(\xi_1-t)+\alpha_2\tfrac{L}{2\pi}(\xi_2-t)\big)\Big(\tilde{d}(\xi_1)\tilde{d}(\xi_2)-\frac{\Omega(\xi_1)}{2\pi}\frac{\Omega(\xi_2)}{2\pi}\Big)\,d\xi_1 d\xi_2\,dt$$

However,

$$\begin{aligned}
&\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}r\Big(\tfrac{\xi_1-t}{2\pi}\Big)r\Big(\tfrac{\xi_2-t}{2\pi}\Big)e\big(\alpha_1\tfrac{L}{2\pi}(\xi_1-t)+\alpha_2\tfrac{L}{2\pi}(\xi_2-t)\big)\tilde{d}(\xi_1)\tilde{d}(\xi_2)\,d\xi_1 d\xi_2\,dt\\
&=\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\sum_{\gamma\neq\gamma'}r\Big(\tfrac{\gamma-t}{2\pi}\Big)r\Big(\tfrac{\gamma'-t}{2\pi}\Big)e\big(\alpha_1\tfrac{L}{2\pi}(\gamma-t)+\alpha_2\tfrac{L}{2\pi}(\gamma'-t)\big)\,dt\\
&\quad+\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\sum_{\gamma}r^2\Big(\tfrac{\gamma-t}{2\pi}\Big)e\big((\alpha_1+\alpha_2)\tfrac{L}{2\pi}(\gamma-t)\big)\,dt\\
&=\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\sum_{\gamma\neq\gamma'}r\Big(\tfrac{\gamma-t}{2\pi}\Big)r\Big(\tfrac{\gamma'-t}{2\pi}\Big)e\big(\alpha_1\tfrac{L}{2\pi}(\gamma-t)+\alpha_2\tfrac{L}{2\pi}(\gamma'-t)\big)\,dt\\
&\quad+O_{r,\sigma}\Big(\frac{1}{H}\Big)+\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\int_{-\infty}^{\infty}r^2\Big(\tfrac{\xi-t}{2\pi}\Big)e\big((\alpha_1+\alpha_2)\tfrac{L}{2\pi}(\xi-t)\big)\frac{\Omega(\xi)}{2\pi}\,d\xi\,dt,
\end{aligned}$$

using Lemma 2.3.5 in the last line.

By Stirling's formula, we have that

$$\Omega(\xi+t+T)=\log(T/2\pi)+O\Big(\frac{1}{|\xi+t+T|+2}\Big)+O\big(\log(1+\tfrac{|\xi+t|}{T})\big)$$

so that

$$\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\frac{\Omega(\xi+t)}{2\pi}dt=\frac{\log(T/2\pi)}{2\pi}+O\Big(\frac{1}{H}\Big)+O\Big(\frac{|\xi|+H}{T}\Big)$$

and

$$\begin{aligned}
&\frac{1}{H}\int_{\mathbb{R}}\sigma\Big(\tfrac{t-T}{H}\Big)\frac{\Omega(\xi_1+t)}{2\pi}\frac{\Omega(\xi_1+t)}{2\pi}dt\\
&=\Big(\frac{\log(T/2\pi)}{2\pi}\Big)^2+O\Big(\frac{\log T}{H}\Big)+O\Big(\frac{|\xi_1|+|\xi_2|+H}{T}\Big).
\end{aligned}$$

By removing the δ function, (2.16) therefore implies that

$$\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big) \sum_{\gamma \neq \gamma'} r\big(\tfrac{\gamma-t}{2\pi}\big) r\big(\tfrac{\gamma'-t}{2\pi}\big) e\big(\alpha_1 \tfrac{L}{2\pi}(\gamma-t) + \alpha_2 \tfrac{L}{2\pi}(\gamma'-t)\big)\, dt + O\Big(\log T\Big(\frac{H}{T} + \frac{1}{H}\Big)\Big)$$

$$= O\Big(\frac{T^{|\frac{\alpha_1}{2}|+|\frac{\alpha_2}{2}|}}{H}\Big)$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big[\Big(\frac{\log(T/2\pi)}{2\pi}\Big)^2 + Q_T(\nu_1-\nu_2)\Big] r\big(\tfrac{\nu_1}{2\pi}\big) r\big(\tfrac{\nu_2}{2\pi}\big) e\big(\alpha_1 \tfrac{L}{2\pi}\nu_1 + \alpha_2 \tfrac{L}{2\pi}\nu_2\big)\, d\nu_1 d\nu_2,$$

as claimed. □

Proof of Theorem 4. In Theorem 2.1.3, we let $h(\nu_1,\nu_2) = \omega(\nu_1-\nu_2)\eta\big(\frac{\nu_1+\nu_2}{2}\big)$ where η is any function with a smooth and compactly supported Fourier transform and $\hat{\eta}(0) = 1$ say, and we set $\alpha_1 = -\alpha_2 = \alpha$. The left hand side of (2.5) is

$$\sum_{\gamma\neq\gamma'} \omega(\gamma-\gamma') e\big(\alpha \tfrac{L}{2\pi}(\gamma-\gamma')\big) \frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big)\eta\big(\tfrac{\gamma+\gamma'}{2} - t\big)\, dt.$$

On the other hand, averaging T from 0 to R:

$$\frac{1}{R}\int_0^R \frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big)\eta(\nu-t)\, dt\, dT = \frac{1}{R}\int_{-\infty}^{\infty} \eta(\tau)\int_{-\frac{R}{H}+\frac{\nu-\tau}{H}}^{\frac{\nu-\tau}{H}} \sigma(Q)\, dQ\, dt.$$

If $H = R^{1-\epsilon/2}$ and we fix δ' less than $\epsilon/2$, we have that for $\nu \in [R^{1-\delta'}, R - R^{1-\delta'}]$,

$$\int_{-\infty}^{\infty} \eta(\tau)\int_{-\frac{R}{H}+\frac{\nu-\tau}{H}}^{\frac{\nu-\tau}{H}} \sigma(Q)\, dQ\, dt = \int_{-\infty}^{\infty} \eta(\tau)\, d\tau + O\Big(\frac{1}{R^j}\Big),$$

for any $j > 0$ as both g and σ are Schwartz. (The implied constant will vary with j.) On the other hand, for $\nu \geq R + R^{1-\delta'}$ or $\nu \leq -R^{1-\delta'}$ this quantity is

$$O\Big(\frac{1}{R^j}\Big)$$

for the same reason. For typographical reasons we use the notation

$$E_1 = \{(\gamma,\gamma') \ : \ \gamma \neq \gamma',\, R^{1-\delta'} \leq \tfrac{\gamma+\gamma'}{2} \leq R - R^{1-\delta'}\},$$

and

$$\begin{aligned} E_2 = \{(\gamma,\gamma') \ :\gamma \neq \gamma', \text{ and either} & \\ & R - R^{1-\delta'} \leq \tfrac{\gamma+\gamma'}{2} \leq R + R^{1-\delta'} \text{ or } -R^{1-\delta'} \leq \tfrac{\gamma+\gamma'}{2} \leq R^{1-\delta'}\}, \end{aligned}$$

and we therefore have,

$$\begin{aligned} &\frac{1}{R}\int_0^R \sum_{\gamma\neq\gamma'} \omega(\gamma-\gamma')e\big(\alpha\tfrac{L}{2\pi}(\gamma-\gamma')\big)\frac{1}{H}\int_{\mathbb{R}} \sigma\Big(\tfrac{t-T}{H}\Big)\eta\big(\tfrac{\gamma+\gamma'}{2}-t\big)\,dt \\ &= \frac{1}{R}\sum_{(\gamma,\gamma')\in E_1} \omega(\gamma-\gamma')e\big(\alpha\tfrac{L}{2\pi}(\gamma-\gamma')\big)\int_{-\infty}^{\infty}\eta(\tau)\,d\tau \\ &\quad + O\Bigg(\frac{1}{R}\sum_{(\gamma,\gamma')\in E_2}\omega(\gamma-\gamma')\Bigg) + O\Big(\frac{1}{R^j}\Big) \\ &= \frac{1}{R}\sum_{0<\gamma\neq\gamma'\leq R} \omega(\gamma-\gamma')e\big(\alpha\tfrac{L}{2\pi}(\gamma-\gamma')\big)\int_{-\infty}^{\infty}\eta(\tau)\,d\tau + O\Big(\frac{\log R}{R^{\delta'}}\Big), \end{aligned}$$

as ω is Schwartz and our sums are therefore effectively limited to γ,γ' with $\gamma-\gamma' = O(1)$. On the other hand, Theorem 2.1.3 implies this is

$$O\Big(\frac{\log R}{R^{\epsilon/2}}\Big)+\int_{-\infty}^{\infty}\eta(\tau)\,d\tau\int_{-\infty}^{\infty}\omega(u)e\big(\alpha\tfrac{\log T}{2\pi}u\big)\left[\frac{1}{R}\int_0^R \Big(\frac{\log(T/2\pi)}{2\pi}\Big)^2 + Q_T(u)\,dT\right]du.$$

Selecting $\delta' > \delta$ gives us the theorem. □

It is worthwhile finally to discuss the difference between this prediction and one in which $Q_t(u)$ has been replaced by

$$K_t(u) = -\frac{\sin(\pi\frac{\log(T/2\pi)}{2\pi}u)}{(\pi u)^2}$$

This replacement would amount to a naïve extension of the GUE conjecture. We want at least to check that

$$\Delta_1 := \int_{\mathbb{R}} \omega(u) e\big(\alpha \tfrac{\log T}{2\pi} u\big) \left[\frac{1}{T}\int_0^T Q_t(u) - K_t(u)\, dt\right] du$$

is not negligible, since otherwise the rather more recondite expression involved in defining Q_t would be unnecessary.

In the first place, we showed above, through Lemma 2.3.3, that

$$\Delta_2 := \int_{\mathbb{R}} \omega(u) e\big(\alpha \tfrac{\log T}{2\pi} u\big) \left[\frac{1}{T}\int_0^T Q_t(u) - \tilde{Q}_t(u)\, dt\right] du = O_\delta\big(\frac{1}{T^\delta}\big),$$

for δ and α as in Theorem 2.1.4, where

$$\tilde{Q}_t(u) = \frac{1}{4\pi^2}\bigg(\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) - B(iu) + \Big(\frac{\zeta'}{\zeta}\Big)'(1-iu) - B(-iu) \\ + e(-\tfrac{\log(t/2\pi)}{2\pi} u) + e(\tfrac{\log(t/2\pi)}{2\pi} u)\bigg).$$

Indeed, if we content ourselves with Δ_2 decaying like $O_k(1/\log^k T)$ for any k, instead of $O_\delta(1/T^\delta)$, this is obviously true as long as α is bound away from -1 and 1 – including α larger than 1. (And quite false for α either -1 or 1.)

On the other hand,

$$\tilde{Q}_t(u) - K_t(u) = \frac{1}{4\pi^2}\bigg(2\Re\Big[\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) - B(iu)\Big] + \frac{2}{u^2}\bigg),$$

so that for α bound away from 0

$$\Delta_3 := \int_{\mathbb{R}} \omega(u) e\big(\alpha \tfrac{\log T}{2\pi} u\big) \left[\frac{1}{T}\int_0^T \tilde{Q}_t(u) - K_t(u)\, dt\right] du = O_k\Big(\frac{1}{\log^k T}\Big),$$

but for $\alpha = 0$, Δ_3 is not even $o(1)$ as $\tilde{Q}_t(u) - K_t(u)$ has no dependence on t and is not identically 0 (this would falsely imply that $(\zeta'/\zeta)'(1+iu)$ is bounded, for instance). This is enough to see that K_t cannot replace Q_t in Theorem 2.1.3 or 2.1.4.

We can examine the difference between the two predictions in greater detail. Note that for $\Re s > 1$,

$$\sum_{n=1}^{\infty} \frac{\Lambda^2(n)}{n^s} = \sum_{k=1}^{\infty} c_k \Big(\frac{\zeta'}{\zeta}\Big)'(ks),$$

where

$$c_k = \sum_{d|k} \mu(d) d,$$

so that

$$\tilde{Q}_t(u) - K_t(u) = \frac{1}{2\pi^2}\Bigg[\Re\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) + \frac{1}{u^2} + \sum_{k=2}^{\infty} c_k \Re\Big(\frac{\zeta'}{\zeta}\Big)'(k+iku)\Bigg].$$

By (2.13) the function $\Re\Big(\frac{\zeta'}{\zeta}\Big)'(1+iu) + \frac{1}{u^2}$ has troughs corresponding to each γ, however these troughs are all of the same width and height and so their appearance is only really striking for low-lying zeros which are spread far apart. For higher zeros with less space between them, the troughs interfere with one another. In addition, the functions $\Re\Big(\frac{\zeta'}{\zeta}\Big)'(k+iku)$ have troughs, smaller in depth and width, around γ/k. Since, for instance, $c_2 = -1$, this corresponds to a small bump in the pair correlation function around the values $\gamma/2$ – for instance around $14.132/2 = 7.066$ – at least while tested against sufficiently smooth test functions. One can discern this bump in Figure 2.1.

We note in passing that for $\alpha > 1 + \epsilon$, the above discussion shows that the

Bogomolny-Keating heuristics are consistent with a conjecture that

$$\begin{aligned}\frac{1}{T}\sum_{0<\gamma\neq\gamma'<T} w(\gamma-\gamma')e\big(\alpha\tfrac{\log T}{2\pi}(\gamma-\gamma')\big) &= o(1) + \int_{\mathbb{R}} w(u)e\big(\alpha\tfrac{\log T}{2\pi}u\big) \\ &\qquad \times \left[\frac{1}{T}\int_0^T \Big(\frac{\log(t/2\pi)}{2\pi}\Big)^2 + K_t(u)\,dt\right] du \\ &= o(1),\end{aligned}$$

shown by Chan [11] to be essentially equivalent to a conjecture of Montgomery and Soundararajan [54] for the second moment of primes in short intervals. Their conjecture was based upon the Hardy-Littlewood conjectures, but even on the assumption of these conjectures, to rigorously extend the domain of α in Theorem 2.1.4 while maintaining an inverse power-of-T error term will require much more work.

CHAPTER 3

A central limit theorem for the zeros of the zeta function

3.1 Introduction

In chapter 2 we saw that macroscopically the pair correlation function of zeta zeros deviates in an interesting fashion from pair correlation function of a determinantal point process with sine kernel. In particular, the distance separating two random zeros tends repel weakly but noticeably away from the value of low-lying zeros of the zeta function.

This chapter is an account of a mesoscopic central limit theorem for the number of zeros of the Riemann zeta function as counted by a (possibly) smoothed counting function. We remind the reader that we assume RH. Our concern is the statistical distribution of γ near some large (random) height T. We will give evidence through this central limit theorem for an observation first put forward by Berry [8] in heuristic language: that while the zeros at a macroscopic scale deviate away from the universal sine-kernel determinantal pattern, at a mesoscopic scale collections of zeros still resemble the points of a sine-kernel determinantal point process in almost all meaningful ways. The limit theorem we prove has the advantage of a simple and satisfying statement, mirroring results for the unitary group that go back to Szegő ([68], ch. 6). Later in chapter 4 of this thesis we will provide a more comprehensive account of this resemblance, insofar as our rigorous knowledge extends, at the price of requiring a more ornate statement.

If $N(T)$ is the number of nontrivial zeros in the upper half plane with height no more than T, then the number of zeros $N(t+h) - N(t)$ to occur in an interval $[t, t+h]$ is expected to be roughly $h\frac{\log t}{2\pi}$ [79]. It was first shown by Fujii [26], [27] that the oscillation of this quantity is Gaussian, with a variance depending upon the number of zeros expected to lie in the interval.

Theorem 3.1.1 (Fujii's mesoscopic central limit theorem)**.** *Let $n(T)$ be a fixed function tending to infinity as $T \to \infty$ in such a way that $n(T) = o(\log T)$, and let t be a random variable uniformly distributed on the interval $[T, 2T]$. For notational reasons we label by X_T the probability space from which the random variable t is drawn. Then, letting $\Delta = \Delta(t,T) := N(t + \frac{2\pi n(T)}{\log T}) - N(t)$,*

$$\mathop{\mathbf{E}}_{X_T} \Delta = n(T) + o(1),$$

$$\mathrm{Var}_{X_T}(\Delta) \sim \frac{1}{\pi^2} \log n(T),$$

and in distribution

$$\frac{\Delta - \mathbf{E}\,\Delta}{\sqrt{\mathrm{Var}\Delta}} \Rightarrow N(0,1)$$

as $T \to \infty$.

The main purpose of this note is to generalize Fujii's theorem in the following way:

Theorem 3.1.2 (A general mesoscopic central limit theorem)**.** *Let $n(T)$ and X_T be as in Theorem 3.1.1. For a fixed real valued function η with compact support and bounded variation, define*

$$\Delta_\eta = \Delta_\eta(t,T) = \sum_\gamma \eta\big(\tfrac{\log T}{2\pi n(T)}(\gamma - t)\big),$$

where the sum is over all zeros γ, counted with multiplicity. In the case that

$\int |x||\hat{\eta}(x)|^2\,dx$ diverges, we have

$$\mathop{\mathbf{E}}_{X_T} \Delta_\eta = n(T) \int_{\mathbb{R}} \eta(\xi) d\xi + o(1),$$

$$\mathrm{Var}_{X_T}(\Delta_\eta) \sim \int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 dx$$

and in distribution

$$\frac{\Delta_\eta - \mathbf{E}\,\Delta_\eta}{\sqrt{\mathrm{Var}\Delta_\eta}} \Rightarrow N(0,1)$$

as $T \to \infty$.

It is a straightforward computation to see that Theorem 3.1.1 follows from Theorem 3.1.2 by letting $\eta = \mathbf{1}_{[-1/2,1/2]}$.

Additionally, in the case of variances that converge:

Theorem 3.1.3. *For $n(T)$ and X_T as in Theorem 3.1.2, but η with compact support and bounded second derivative, the integral $\int |x||\hat{\eta}(x)|^2\,dx$ is necessarily finite, but the conclusion of Theorem 3.1.2 remains true even still.*

Remark: The condition that the test functions in Theorems 3.1.2 and 3.1.3 have compact support is not optimised; somewhat looser decay properties, along the lines of quadratic decay, are sufficient for the proof that follows.

We call Theorems 3.1.1, 3.1.2, and 3.1.3 'mesoscopic' central limit theorems as they concern collections of $n(T)$ zeros which grow to infinity, but intervals whose length $\frac{2\pi n(T)}{\log T}$ tends to 0 all the same.

On such mesoscopic intervals (averaged as in Theorems 3.1.1 and 3.1.2), all evidence points to the zeros resembling points in a determinantal point process with sine kernel. In the microscopic regime (when $n(T) = O(1)$) this is known to be the case, provided we restrict our attention to the statistics counted by sufficiently smooth test functions. (See Rudnick and Sarnak [63] or Hughes and Rudnick [42].) The techniques that follow allow us to recover these results for smooth test functions, as well as extend them to a mesoscopic regime, in a sense to be specified. These matters are discussed in the appendix.

For the moment, we may simply note the similarity of Theorems 3.1.1, 3.1.2, and 3.1.3 to certain results in the theories of random matrices and determinantal point processes (for an introduction to the latter, see the appendix B or alternatively [40] or the introduction of [71]):

Theorem 3.1.4 (Costin and Lebowitz)**.** *Let X be a determinantal point process on $\mathbb{R}$ with sine kernel $K(x,y) = \frac{\sin \pi(x-y)}{\pi(x-y)}$, and Δ a count of the number of points lying in the interval $[0, L]$. Then*

$$\mathop{\mathbf{E}}_{X} \Delta = L,$$

$$\mathrm{Var}_X(\Delta) \sim \frac{1}{\pi^2} \log L$$

and in distribution

$$\frac{\Delta - \mathbf{E}\,\Delta}{\sqrt{\mathrm{Var}\Delta}} \Rightarrow N(0,1)$$

as $L \to \infty$.

For more general test functions, the analogue for determinantal point processes of Theorem 3.1.3 is a corollary of both Theorem 3 or 4 of Soshnikov [71], who attributes this corollary to Spohn [74]. Proofs of closely related results can also be found in [72].

Theorem 3.1.5. *Let X be a determinantal point process on $\mathbb{R}$ with sine kernel $K(x,y) = \frac{\sin \pi(x-y)}{\pi(x-y)}$. For η a Schwartz function and L a positive number, define*

$$\Delta_\eta = \sum \eta(x_i/L)$$

where $((x_i))$ are the points of the point process. Note that for such η, $|x||\hat{\eta}(x)|^2$ is *always integrable. We have*

$$\mathop{\mathbf{E}}_{X} \Delta_\eta = L \int_{\mathbb{R}} \eta(\xi) d\xi,$$

$$\mathrm{Var}_X(\Delta_\eta) \sim \int_{-\infty}^{\infty} |x||\hat{\eta}(x)|^2 dx$$

and in distribution

$$\frac{\Delta_\eta - \mathbf{E}\,\Delta_\eta}{\sqrt{\mathrm{Var}\Delta_\eta}} \Rightarrow N(0,1)$$

as $L \to \infty$.

Note that in this case there is no limit on the growth rate of L, which is not surprising since the theorem is for a single point process X rather than a series of point processes.

Remark: The test functions for this theorem are Schwartz class. One expects the theorem to be true for any test functions η for which $|x||\hat{\eta}(x)|^2$ is integrable, a considerably larger class. A statement of this does not appear to be in the literature for the sine-kernel determinantal point process, but analogous results of this sort are known for the point processes induced by eigenvalues of random $n \times n$ unitary matrices, and test functions that count all n eigenvalues. (One may call such theorems 'macroscopic' as opposed to mesoscopic theorems for test functions that count only $o(n)$ eigenvalues.) Such results are known as strong Szegő theorems, and considerable literature surrounds the subject. (See [68], Ch. 6 for instance.)

Likewise, the analogue of Theorem 3.1.2 does not seem to be in the literature, but an analogue was proved by Diaconis and Evans [19] for counts of all n eigenvalues of unitary matrices, among other ensembles. The perspective of Diaconis and Evans is perhaps most similar to ours here.

We should therefore expect that Theorem 3.1.2 is true even in the case that $\int |x||\hat{\eta}(x)|^2\, dx$ converges with no more restrictions on η than a bound on variation

– this would encompass Theorem 3.1.3 – but in the latter theorem we require not only that this integral converge, but that it converge somewhat rapidly. Bounding an error term prevents us from accessing the results in between the two theorems, even though by analogy we should fully expect them to be true.

In fact, Fujii proved a more general result than Theorem 3.1.1, encompassing macroscopic intervals as well. In order to state Fujii's result succinctly, we recall the definition

$$S(t) := \arg \zeta(\tfrac{1}{2} + it),$$

where argument is defined by a continuous rectangular path from 2 to $2 + it$ to $\frac{1}{2} + it$, beginning with $\arg 2 = 0$, and by upper semicontinuity in case this path passes through a zero. $S(t)$, as it ends up, is small and oscillatory, and our interest in it derives from the fact that it appears as an error term in the zero counting function:

$$N(T) = \tfrac{1}{\pi} \arg \Gamma\big(\tfrac{1}{4} + i\tfrac{T}{2}\big) - \tfrac{T}{2\pi} \log \pi + 1 + S(T). \tag{3.1}$$

Theorem 3.1.6 (Fujii's macroscopic central limit theorem)**.** *Let X_T be as in Theorem 3.1.1, and $n(T)$ with $\log T \lesssim n(T) \lesssim T$. Define $\tilde{\Delta} = S(t + \frac{2\pi n(T)}{\log T}) - S(t)$. Then*

$$\mathop{\mathbf{E}}_{X_T} \tilde{\Delta} = o(1),$$

$$\mathrm{Var}_{X_T}(\tilde{\Delta}) \sim \frac{1}{\pi^2} \log \log T,$$

and in distribution

$$\frac{\tilde{\Delta}}{\sqrt{\mathrm{Var} \tilde{\Delta}}} \Rightarrow N(0,1)$$

as $T \to \infty$.

Note that in this case, if Δ is defined as before with respect to the function $N(t)$, $\mathbf{E}_{X_T} \Delta$ does not have quite as nice an expression owing to the growth of the logarithm function.

In fact, it will in general prove preferable to work with $S(t)$ in place of $N(t)$ in the computations that follow. Differentiating (3.1), we have

$$\left[\tilde{d}(\xi) - \tfrac{\Omega(\xi)}{2\pi}\right] d\xi = dS(\xi),$$

where

$$\tilde{d}(\xi) := \sum_{\gamma} \delta(\xi - \gamma),$$

with the sum over zeros counted with multiplicity, and

$$\Omega(\xi) := \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\left(\tfrac{1}{4} + i\tfrac{\xi}{2}\right) + \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\left(\tfrac{1}{4} - i\tfrac{\xi}{2}\right) - \log \pi.$$

Making use of the moment method[1] and Stirling's formula [2], we see that Theorem 3.1.2 will be implied by

Theorem 3.1.7. *For η a real-valued function with compact support and bounded variation, for $n(T) \to \infty$ as $T \to \infty$ in such a way that $n(T) = o(\log T)$,*

$$\frac{1}{T}\int_T^{2T} \left[\int_{\mathbb{R}} \eta\left(\tfrac{\log T}{2\pi n(T)}(\xi - t)\right) dS(\xi)\right]^k dt = (c_k + o(1))\left[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2\, dx\right]^{k/2},$$

provided the integral on the right diverges. Here $c_\ell := (\ell - 1)!!$ for even ℓ, and $c_\ell := 0$ for odd ℓ, are the moments of a standard normal variable.

Theorem 3.1.3, as well, will follow from the above statement where η is instead restricted as in Theorem 3.1.3.

In order to prove his results, Fujii made use of the moment method, and the

[1] An introduction to the moment method can be found in for example [77] section 2.2.3.

[2] Stirling's formula (proved in [1] section 1.4 for instance) implies that $\frac{\Omega(\xi)}{2\pi} = \frac{\log\left((|\xi|+2)/2\pi\right)}{2\pi} + O\left(\frac{1}{|\xi|+2}\right)$.

following (unconditional) approximation due to Selberg [65],[66],

$$\frac{1}{T}\int_T^{2T}\Big[S(t)+\tfrac{1}{\pi}\sum_{p\le T^{1/k}}\frac{\sin(t\log p)}{\sqrt{p}}\Big]^{2k}\,dt = O(1), \tag{3.2}$$

which Selberg had used earlier to derive a more global central limit theorem for $S(t)$,

$$\frac{1}{T}\int_T^{2T}|S(t)|^{2k}\,dt \sim \tfrac{(2k-1)!!}{(2\pi^2)^k}(\log\log T)^k.$$

These formulas are sufficient to prove Theorem 3.1.2 for test functions η which are sums of a finite number of indicator functions. They break down, however, in an attempt to prove the theorem for general η, since, although one can approximate η by simple functions, the error terms thus generated rapidly overwhelm the main terms of the moments.

We do not therefore make use of Selberg's approximation for $S(t)$, and indeed do not pass through his well known mollification for ζ'/ζ, which Selberg and many other authors typically make use of to prove statistical theorems about zeta zeros.

Our approach is outlined in the next section. Very roughly stated, it is a sort of weak analogue of the approach of Selberg and Fujii. In this, we follow the derivation [42] of Hughes and Rudnick of mock-gaussian behavior in the microscopic regime with respect to sufficiently smooth test functions. We extend these computations to the mesoscopic regime, still requiring smoothness, but the key point which allows us to obtain our central limit theorem is that any test function will become sufficiently smooth when dilated as they are in the central limit theorems 3.1.2 and 3.1.3. This is one clarifying feature of our proof. The proof of Fujii's theorem making use of Selberg's approximation for $S(t)$ leaves the link between this central limit theorem and the microscopic determinantal structure of the zeros somewhat mysterious.

This approach, with slightly more work, can be used to produce Fujii's Theorem 3.1.6 as well, although in this case an analogue of Theorem 3.1.2 is less satisfying. We shall not prove so in this note, but in the macroscopic case already if η is so much as absolutely continuous, the variance and higher moments of $\tilde{\Delta}_\eta$ (defined in the obvious way) tend to 0. This is a feature of the rigidity of the distribution of zeros at this regime, which while not quite as rigid as a clock distribution (see [50] for a definition), resemble at this level this distribution perhaps somewhat more than they do a sine kernel determinantal point process. One should compare this analogy with the classical theorems that for a fixed h, $N(t+h) - N(t) \asymp \log t$ for *all* sufficiently large t, with constants depending upon h. (See [79], Theorems 9.2 and 9.14.) In this regime, arithmetic factors play a heavy explicit role; this will be implicitly evident in the proof that follows. In this, we can recover the heuristic observations of Berry [8] regarding the origin for the variance terms in Fujii's theorems. Indeed, one can already discern, by comparing Fujii's central limit theorem to the central limit theorem of Costin and Lebowitz, that the statistics of the zeros in this regime cannot be modeled too closely by a sine-kernel determinantal process. Outside of the mesoscopic regime, these statistics demonstrate the important 'resurgence phenomenon' discovered heuristically by Bogomolny and Keating, and which has been discussed rigorously in chapter 2.

Zeev Rudnick pointed out to the author that he had used similar ideas with Faifman in [23] to prove a Fujii-type central limit theorem, for counting functions with a strict cutoff, in the finite field setting.

One can apply these ideas to get central limit theorems as well for the number of low-lying zeros of $L(s, \chi_d)$, where χ_d ranges over the family of primitive quadratic characters, by extending the microscopic statistics of Rubenstein [62].

After a note with these results had been posted online, Theorems 3.1.2 and 3.1.3 were independently proven, in a different manner, by Bourgade and Kuan [7]. They proceed by a method which makes use of the Helffer-Sjöstrand functional calculus and a mollification formula of Selberg. In brief summary, the dichotomy between the two treatments is that between the use of harmonic analysis (our approach) and complex analysis (the approach of Bourgrade and Kuan).

The conditions on admissible test functions in [7] differ very slightly from those in the theorems here, but not in an important way. Either method seems to include the class of test functions – lying within the class of test functions with converging variance – for which a central limit theorem, by analogy with random matrix theory and discussed in the remark above, ought to be true, but for which we have no proof. It would be interesting if other approaches could fill this small but pesky limitation in our knowledge.

3.2 A heuristic outline of the proof

Here we give a heuristic sketch of our approach before proceeding to a rigorous proof.

Instead of Selberg's metric $L^k[0,T]$ approximation for the function $S(t)$, we use the distributional formula,

$$dS(\xi) = -\int \frac{e^{ix\xi} + e^{-ix\xi}}{2} e^{-x/2}\, d\big(\psi(e^x) - e^x\big).$$

This formula is to be understood heuristically; some restrictions are entailed on the test functions in ξ against which it can be integrated, a precise statement of which are given by Theorem 3.3.1. The integral on the right has an arithmetic component,

$$-\frac{1}{\pi}\sum_n \frac{e^{i\xi\log n} + e^{-i\xi\log n}}{2}\frac{\Lambda(n)}{\sqrt{n}} \approx -\frac{1}{\pi}\sum_p \frac{e^{i\xi\log p} + e^{-i\xi\log p}}{2}\frac{\log p}{\sqrt{p}}$$

(compare with Selberg's (3.2)), and a continuous component

$$\int \frac{e^{ix\xi} + e^{-ix\xi}}{2} e^{x/2}\, dx.$$

It will emerge from computations that the measure in variables $\xi_1, ..., \xi_k$, given by

$$\frac{1}{T}\int_T^{2T} \prod_{\ell=1}^{k} dS(\xi_\ell + t)\, dt,$$

is extremely well approximated by substituting for dS in each variable only its arithmetic component:

$$\frac{1}{T}\int_T^{2T} \prod_{\ell=1}^{k} \Big(-\frac{1}{\pi}\sum_p \frac{e^{i\xi\log p} + e^{-i\xi\log p}}{2} \frac{\log p}{\sqrt{p}}\, d\xi_\ell \Big)\, dt,$$

so long as the measures are being integrated against functions $f(\xi_1, ..., \xi_k)$ that have their Fourier transform supported at a scale of $O(\log T)$. Said another way, this approximation is a good one so long as the test function f is sufficiently smooth, observed on intervals of size $1/\log T$.

The statistics in which we will be interested for our central limit theorem are

$$\frac{1}{T}\int_T^{2T} \int_{\xi\in\mathbb{R}^k} \prod_{\ell=1}^{k} \eta\Big(\tfrac{\log T}{2\pi n(T)}\xi_\ell\Big)\, dS(\xi_\ell + t)\, dt.$$

For any 'nice' function η, because $n(T)\to\infty$, for large enough T, the function

$$\eta\Big(\tfrac{\log T}{2\pi n(T)}\xi_\ell\Big)\cdots\eta\Big(\tfrac{\log T}{2\pi n(T)}\xi_\ell\Big)$$

will be basically smooth in ξ at a scale of $1/\log T$. (One would have to observe at the larger scale of $n(T)/\log T$ to see the variations in these test functions.)

Therefore the above integral can be approximated by

$$\frac{1}{T}\int_T^{2T}\prod_{\ell=1}^k \int_{-\infty}^{\infty} \eta\Big(\tfrac{\log T}{2\pi n(T)}\xi_\ell\Big)\Big(-\frac{1}{\pi}\sum_p \frac{e^{i\xi\log p}+e^{-i\xi\log p}}{2}\frac{\log p}{\sqrt{p}}\Big)\,d\xi_\ell\,dt$$
$$=\frac{1}{T}\int_T^{2T}\prod_{\ell=1}^k \frac{n(T)}{\log T}\sum_p \frac{\log p}{\sqrt{p}}\Big(\hat\eta\Big(\frac{\log p}{\log T/n(T)}\Big)e^{it\log p}+\hat\eta\Big(-\frac{\log p}{\log T/n(T)}\Big)e^{-it\log p}\Big)\,dt$$

For finite collections of primes p and large T, the quantities $e^{it\log p}$ behave like independent random variables as t ranges over $[T, 2T]$. We will be interested in collections of primes p that grow with T (with all primes p so that $\log p/(\log T/n(T)) = O(1)$ in fact, owing to the decay of the function $\hat\eta$), but by mimicking the analysis that leads to this observation, we are able to see that the above average contracts to a quantity close to

$$c_k\Big(\frac{n(T)}{\log T}\Big)^2 \sum_{\log p = O(\frac{\log T}{n(T)})} \frac{2\log^2 p}{p}\hat\eta\Big(\frac{\log p}{\log T/n(T)}\Big)\hat\eta\Big(-\frac{\log p}{\log T/n(T)}\Big)$$

where c_k are the moments of a standard normal variable. c_k is given also by the number of possible pairings among a set of k elements, and the only terms in this expression that have survived from the expression above it are pairings of equal primes in the expansion of the product inside the integral. Using the prime number theorem, we are able to show that this tends to the right hand side limit in Theorems 3.1.2 and 3.1.3.

There is one point of this proof which deserves further commentary, as it comprises a substantial part of the technical challenge ahead; this is the claim that for any nice function η, the rescaled function $\eta\big(\frac{\log T}{2\pi n(T)}\xi_\ell\big)\cdots\eta\big(\frac{\log T}{2\pi n(T)}\xi_\ell\big)$ will be sufficiently smooth at a scale of $1/\log T$. It is certainly not true for an arbitrary function of bounded variation η that this rescaling will be locally smooth in the sense we have used above: of having a Fourier transform supported at a scale of $\log T$. For instance, the rescaling of $\eta = \mathbf{1}_{[-1/2,1/2]}$ does not have this property, and indeed no function η will unless η has compact Fourier transform to begin with. What will be true, however, is that for any function η of the sort delimited in Theorems 3.1.2 and 3.1.3, this rescaling can be very well approximated by a function with Fourier transform supported at a scale of $\log T$. Making use of upper bounds for the average number of zeros in an interval of size $1/\log T$, we are able to show that the statistics of this approximation do not deviate much from the statistics of our original test function, and therefore obtain Theorems 3.1.2 and 3.1.3. (Indeed, it is because we must replace test functions with approximations that induce a small error term that we must restrict our attention to a slightly smaller domain of test functions in Theorem 3.1.3 than in 3.1.2.)

3.3 Local Limit Theorems for Smooth Test Functions

This section consists mainly in minor quantitative refinements in the argument of Hughes and Rudnick [42]. In turn, their argument is similar to Selberg's in making use of the fundamental theorem of arithmetic to evaluate certain integrals. We recall again the explicit formula relating zeros to primes:

Theorem 3.3.1. *[The explicit formula] For g a measurable function such that $g(x) = \frac{g(x+)+g(x-)}{2}$, and for some $\delta > 0$,*

$$(a) \quad \int_{-\infty}^{\infty} e^{(\frac{1}{2}+\delta)|x|}|g(x)|dx < +\infty,$$

$$(b) \quad \int_{-\infty}^{\infty} e^{(\frac{1}{2}+\delta)|x|}|dg(x)| < +\infty,$$

we have

$$\int_{-\infty}^{\infty} \hat{g}\big(\tfrac{\xi}{2\pi}\big)\, dS(\xi) = \int_{-\infty}^{\infty} [g(x)+g(-x)]e^{-x/2} d\big(e^x - \psi(e^x)\big),$$

where here $\psi(x) = \sum_{n\leq x} \Lambda(n)$, for the von Mangoldt function Λ.

The integral on the left denotes a principle value integral, $\lim_{L\to\infty}\int_{-L}^{L}$, and this limit necessarily converges when the conditions of the theorem for g are met. In what follows we will frequently work with test functions for which the distinction between this principle value integral and an ordinary integral disappears, and if this is the case we will cease to make one in notation.

Written in this way, the explicit formula is true only on the Riemann hypothesis. It is due in varying stages to Riemann [59], Guinand [31], and Weil [81], and expresses a Fourier duality between the error term in the prime number theorem and the error term for of the zero-counting function.

Without the Riemann hypothesis, we must write the left hand side as

$$\lim_{L\to\infty}\sum_{|\gamma|<L}\hat{g}\big(\tfrac{\gamma}{2\pi}\big)-\int_{-L}^{L}\frac{\Omega(\xi)}{2\pi}\hat{g}\big(\tfrac{\xi}{2\pi}\big)\,d\xi$$

where our sum is over γ (possibly complex) such that $\frac{1}{2}+i\gamma$ is a nontrivial zero of the zeta function, It is proven by a simple contour integration argument, making use of the the reflection formula to evaluate one-half of the contour. (For a proof, see [43] or [56].)

We will also need the following corollary of the prime number theorem.

Lemma 3.3.2 (A prime number asymptotic). *For f with compact support and bounded second derivative,*

$$\frac{1}{H^2}\sum_p\frac{\log^2 p}{p}f\left(\frac{\log p}{H}\right)=O\left(\frac{\|f\|_\infty+\|f'\|_\infty+\|f''\|_\infty}{H^2}\right)+\int_0^\infty xf(x)dx. \quad (3.3)$$

Proof. That something like this is true is evident from the prime number theorem (or even Chebyshev), but some formal care is required to get the desired error term. We will need that,

$$\sum_{p\le n} \frac{\log p}{p} = \log n + C + O(\tfrac{1}{\log^2 n})$$

for some constant C, which is a formula on the level of the prime number theorem (and can be proven from the prime number theorem with a strong error term using partial summation.)

We have then, using the abbreviation $F(x) = xf(x)$,

$$\begin{aligned}\frac{1}{H^2}\sum_p \frac{\log^2 p}{p} f\Big(\frac{\log p}{H}\Big) =& \frac{1}{H}\sum_n \Big[F\big(\tfrac{\log n}{H}\big) - F\big(\tfrac{\log(n+1)}{H}\big)\Big]\Big(\log n + C + O\big(\tfrac{1}{\log^2 n}\big)\Big)\\ =& O\Big(\frac{\|f\|_\infty + \|f'\|_\infty}{H^2}\Big) + \sum_n \frac{\log n - \log(n+1)}{H}\cdot\frac{\log n}{H} F'\big(\tfrac{\log n}{H}\big),\end{aligned}$$

by partial summation and the mean value theorem. Again using the mean value theorem, this time to approximate an integral, we have that this expression is

$$O\Big(\frac{\|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty}{H^2}\Big) + \int_0^\infty xF'(x)dx,$$

which upon integrating by parts is the right hand side of (3.3). □

In what follows instead of working with the average $\frac{1}{T}\int_T^{2T}$ we work with smooth averages $\int \sigma(t/T)/T$ for bump functions σ. What we will show is that

Theorem 3.3.3. *For η as in Theorem 3.1.7, and σ non-negative of mass 1 such that $\hat{\sigma}$ has compact support and $\sigma(t)\log^k(|t|+2)$ is integrable,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\Big[\int_{\mathbb{R}} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big)dS(\xi)\Big]^k dt = (c_k + o(1))\Big[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2\,dx\Big]^{k/2}.$$

We will show that this implies Theorem 3.1.7 at the end of Section 4. We have a computational lemma.

Lemma 3.3.4. *Suppose we are given non-negative integrable σ of mass 1 such that $\hat{\sigma}$ has compact support, and integrable functions $\eta_1, \eta_2, \ldots, \eta_k$ such that $\operatorname{supp}\hat{\eta}_\ell \subset [-\delta_\ell, \delta_\ell]$ with $\delta_1 + \delta_2 + \cdots + \delta_k = \Delta < 2$. There exists a T_0 depending on Δ and the the region in which $\hat{\sigma}$ is supported so that for $T \geq T_0$,*

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{k} \left(\int_{-\infty}^{\infty} \eta_\ell\big(\tfrac{\log T}{2\pi}(\xi_\ell - t)\big) dS(\xi_\ell) \right) dt \\
&= O_k\left(\frac{1}{T^{1-\Delta/2}} \prod_{\ell=1}^{k} \frac{\|\hat{\eta}_\ell\|_\infty}{\log T} \right) \\
&\quad + \left(\frac{-1}{\log T} \right)^k \sum_{n_1^{\epsilon_1} n_2^{\epsilon_2} \cdots n_k^{\epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\Lambda(n_\ell)}{\sqrt{n_\ell}} \hat{\eta}_\ell\big(\tfrac{\epsilon_\ell \log n_\ell}{\log T}\big),
\end{aligned} \tag{3.4}$$

where the sum is over all $n \in \mathbb{N}^k, \epsilon \in \{-1, 1\}^k$ such that $n_1^{\epsilon_1} n_2^{\epsilon_2} \cdots n_k^{\epsilon_k} = 1$.

Proof. By the explicit formula, the right hand side of (3.4) is

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left(\prod_{\ell=1}^{k} \int_{-\infty}^{\infty} \frac{1}{\log T} \Big[\hat{\eta}_\ell\big(-\tfrac{x_\ell}{\log T}\big) e^{-ix_\ell t} + \hat{\eta}_\ell\big(\tfrac{x_\ell}{\log T}\big) e^{ix_\ell t} \Big] e^{-x_\ell/2} d\big(e^{x_\ell} - \psi(e^{x_\ell})\big) \right) dt \\
&\quad = \sum_{\epsilon \in \{-1,1\}^k} \int_{\mathbb{R}^k} \frac{\hat{\sigma}\big(-\frac{T}{2\pi}(\epsilon_1 x_1 + \cdots + \epsilon_k x_k)\big)}{\log^k T} \prod_{\ell=1}^{k} \hat{\eta}_\ell\big(\tfrac{\epsilon_\ell x_\ell}{\log T}\big) e^{-x_\ell/2} d\big(e^{x_\ell} - \psi(e^{x_\ell})\big).
\end{aligned}$$

The second line follows by interchanging the order of integration, justified by the compact support of $\hat{\eta}_\ell$. We can expand the product $\prod e^{-x_\ell/2} d\big(e^{x_\ell} - \psi(e^{x_\ell})\big)$ into a sum of signed terms of the sort $d\beta_1(x_1) \cdots d\beta_k(x_k)$, where $d\beta_\ell(x)$ is either $e^{x/2}dx$ or $e^{-x/2}d\psi(e^x)$. In the case that at least one $d\beta_j$ in our product is $e^{x/2}dx$ we have

$$\left| \int_{\mathbb{R}} \frac{\hat{\sigma}\big(-\frac{T}{2\pi}(\epsilon_1 x_1 + \cdots + \epsilon_k x_k)\big)}{\log^k T} \hat{\eta}_j\big(\tfrac{\epsilon_j x_j}{\log T}\big) d\beta_j(x_j) \right| \lesssim \frac{\|\hat{\eta}_j\|_\infty}{T \log^k T} T^{\delta_j/2},$$

so that in this case

$$\begin{aligned}
&\left| \int_{\mathbb{R}^k} \frac{\hat{\sigma}\big(-\frac{T}{2\pi}(\epsilon_1 x_1 + \cdots + \epsilon_k x_k)\big)}{\log^k T} \prod_{\ell=1}^{k} \hat{\eta}\big(\tfrac{\epsilon_\ell x_\ell}{\log T}\big) d\beta_\ell(x_\ell) \right| \\
&\lesssim \frac{\|\hat{\eta}_j\|_\infty T^{\delta_j/2}}{T \log^k T} \int_{\mathbb{R}^{k-1}} \prod_{\ell \neq j} \hat{\eta}\big(\tfrac{\epsilon_\ell x_\ell}{\log T}\big) d\beta_\ell(x_\ell) \\
&\lesssim \frac{T^{\Delta/2}}{T} \prod_{\ell} \frac{\|\hat{\eta}_\ell\|_\infty}{\log T}
\end{aligned}$$

Into such error terms we can absorb all products $d\beta_1 \cdots d\beta_k$ except that product made exclusively of prime counting measures, namely $(-1)^k \prod e^{-x_\ell/2} d\psi(e^{x_\ell})$. Evaluating the integral of this product measure we have that the left hand side of (3.4) is

$$\begin{aligned}
&O_k\left(\frac{1}{T^{1-\Delta/2}} \prod_{\ell=1}^{k} \frac{\|\hat{\eta}_\ell\|_\infty}{\log T} \right) \\
&+ \left(\frac{-1}{\log T} \right)^k \sum_{\epsilon \in \{-1,1\}^k} \sum_{n \in \mathbb{N}^k} \hat{\sigma}\big(-\tfrac{T}{2\pi}(\epsilon_1 \log n_1 + \cdots + \epsilon_k \log n_k)\big) \prod_{\ell=1}^{k} \frac{\Lambda(n_\ell)}{\sqrt{n_\ell}} \hat{\eta}_\ell\big(\tfrac{\epsilon_\ell \log n_\ell}{\log T}\big).
\end{aligned}$$

Note that if $|\epsilon_1 \log n_1 + \cdots + \epsilon_k \log n_k|$ is not 0, it is greater than $|\log(1 - 1/\sqrt{n_1 \cdots n_k})| \geq \frac{\log 2}{\sqrt{n_1 \cdots n_k}}$ since n_i is always an integer. As $\sqrt{n_1 \cdots n_k} \leq T^{\Delta/2} = o(T)$ and $\hat{\sigma}$ has compact support, for large enough T our sum is over only those ϵ, n such that $\epsilon_1 \log n_1 + \cdots + \epsilon_k \log n_k = 0$. □

Finally, we can use our prime number asymptotic, Lemma 3.3.2, to obtain

Lemma 3.3.5. *For $u_1, ..., u_k$ with bounded second derivative*

$$\frac{1}{H^k} \sum_{n_1^{\epsilon_1} \cdots n_k^{\epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\Lambda(n_\ell)}{\sqrt{n_\ell}} u_\ell\big(\tfrac{\epsilon_\ell \log n_\ell}{H}\big) = S_{[k]} + \sum_{\emptyset \subseteq J \subsetneq [k]} S_J \cdot O_k\left(\prod_{\ell \notin J} \frac{\|u_\ell\|_\infty + \|u_\ell'\|_\infty + \|u_\ell''\|_\infty}{H} \right), \tag{3.5}$$

where $[k] = \{1, ..., k\}$ *and here for a set* J *we define*

$$S_J = \sum \prod_{\lambda} \int_{\mathbb{R}} |x| u_{i_\lambda}(x) u_{j_\lambda}(-x)\, dx$$

where the sum is over all partitions of J *into disjoint pairs* $\{i_\lambda, j_\lambda\}$*.*

Said another way,

$$S_J = \sum_{\pi \in C(J)} \prod_{\ell \in J} \left(\int_{\mathbb{R}} |x| u_\ell(x) u_{\pi(\ell)}(-x)\, dx \right)^{1/2}$$

where the set $C(J)$ is null for $|J|$ odd, and for $|J|$ even is the set of $(|J|-1)!!$ permutations of J whose cycle type is of $|J|/2$ disjoint 2-cycles.

Proof. By Lemma 3.3.2, for any i, j,

$$\begin{aligned}
&\frac{1}{H^2} \sum_{p_1^{\epsilon_1} p_2^{\epsilon_2} = 1} \frac{\log p_1 \log p_2}{\sqrt{p_1 p_2}} u_i\left(\tfrac{\epsilon_i \log p_i}{H}\right) u_j\left(\tfrac{\epsilon_2 \log p_2}{H}\right) \\
&= \int |x| u_i(x) u_j(-x)\, dx + O\left(\frac{\|u_i u_j\|_\infty + \|(u_i u_j)'\|_\infty + \|(u_i u_j)''\|_\infty}{H^2} \right) \\
&= \int |x| u_i(x) u_j(-x)\, dx \qquad (3.6)\\
&+ O\left(\left[\frac{\|u_i\|_\infty + \|u_i'\|_\infty + \|u_i''\|_\infty}{H} \right] \left[\frac{\|u_j\|_\infty + \|u_j'\|_\infty + \|u_j''\|_\infty}{H} \right] \right), \qquad (3.7)
\end{aligned}$$

where the initial sum is over all primes p_1, p_2 and signs ϵ_1, ϵ_2 with $p_1^{\epsilon_1} p_2^{\epsilon_2} = 1$.

It follows that

$$\frac{1}{H^k} \sum_{p^{\epsilon_1} \cdots p_k^{\epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\log p_\ell}{\sqrt{p_\ell}} u_\ell\left(\tfrac{\epsilon_\ell \log p_\ell}{H}\right) = S_{[k]} + \sum_{\emptyset \subseteq J \subsetneq [k]} S_J \cdot O_k\left(\prod_{\ell \notin J} \frac{\|u_\ell\|_\infty + \|u_\ell'\|_\infty + \|u_\ell''\|_\infty}{H} \right),$$

as, by the fundamental theorem of arithmetic, $p_1^{\epsilon_1} \cdots p_k^{\epsilon_k} = 1$ if and only if primes match up pairwise $p_i = p_j$ with $\epsilon_i = -\epsilon_j$. The error term listed accumulates by expanding those products in which terms of the sort (3.6) occur.

It remains to show that

$$\frac{1}{H^k}\sum_{p_1^{\epsilon_1\lambda_1}\cdots p_k^{\epsilon_k\lambda_k}=1}\prod_{\ell=1}^{k}\frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}}u_\ell\Big(\tfrac{\epsilon_\ell\lambda_\ell\log p_\ell}{H}\Big)=\sum_{\emptyset\subseteq J\subsetneq[k]}S_J\cdot O_k\bigg(\prod_{\ell\notin J}\frac{\|u_\ell\|_\infty+\|u_\ell'\|_\infty+\|u_\ell''\|_\infty}{H}\bigg), \tag{3.8}$$

where the sum is over primes $p_1,...,p_k$, signs $\epsilon_1,...,\epsilon_k$, and positive integers $(\lambda_1,...,\lambda_k)\in\mathbb{N}_+^k\setminus\{(1,1,...,1)\}$. But the left hand side sum of (3.8) restricted to λ with $\lambda_1\geq 3,...,\lambda_k\geq 3$ is plainly

$$O\bigg(\prod_{\ell=1}^{k}\frac{\|u_\ell\|}{H}\bigg).$$

On the other hand, for λ with λ_j fixed to equal 2 for some j, by the fundamental theorem of arithmetic $p_1^{\epsilon_1\lambda_1}\cdots p_k^{\epsilon_k\lambda_k}=1$ only in the case that $p_j=p_{j'}$ for some $j'\neq j$, so that thus restricted left hand side sum of (3.8) is

$$\sum_{\substack{I\subset[k]\\ j\notin I}}O\bigg(\sum_{p_j}\frac{\log^2 p_j}{p_j{}^{3/2}}\cdot\prod_{\ell'\notin I}\frac{\|u_{\ell'}\|_\infty}{H}\times\frac{1}{H^{|I|}}\sum_{p}\prod_{\ell\in I}\frac{\log p_\ell}{p^{\lambda_\ell/2}}u_\ell\Big(\tfrac{\epsilon_\ell\lambda_\ell\log p_\ell}{H}\Big)\bigg)$$

where the sum with index labeled p is over p,λ,ϵ such that $\prod_{\ell\in I}p_\ell^{\epsilon_\ell\lambda_\ell}=1$, and I has the function in this sum of collecting those p_i which are not equal to p_j. This expression is unpleasant, but our consolation is that it is only an error term. Applying it inductively, to bound the sums restricted to $\prod_{\ell\in I}p_\ell^{\epsilon_\ell\lambda_\ell}=1$, yields the Lemma. (We have here fixed $\lambda_j=2$, but of course to get an upper bound we need add at most k sums like this.) □

As a consequence of Lemmas 3.3.4 and 3.3.5, with $H=\frac{\log T}{n(T)}$,

Corollary 3.3.6. *For η and σ as in Lemma 3.3.4,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{k} \left(\int_{-\infty}^{\infty} \eta_\ell\big(\tfrac{\log T}{2\pi n(T)}(\xi_\ell - t)\big) dS(\xi_\ell) \right) dt$$
$$= S_{[k]} + O_k\left(\frac{1}{T^{1-\Delta/2}} \prod_{\ell=1}^{k} \frac{\|\hat{\eta}_\ell\|_\infty}{\log T/n(T)} \right)$$
$$+ \sum_{\emptyset \subseteq J \subsetneq [k]} S_J \cdot O_k\left(\prod_{\ell \notin J} \frac{\|\hat{\eta}_\ell\|_\infty + \|\hat{\eta}_\ell'\|_\infty + \|\hat{\eta}_\ell''\|_\infty}{\log T/n(T)} \right),$$

where S_J is defined as in Lemma 3.3.5 with $u_\ell = \hat{\eta}_\ell$.

Remark: As an aside, we note that by modifying the above analysis, making Δ small enough, one can obtain an asymptotic even in the case that $n(T)$ grows like $O(T^{1-\delta})$, for $\delta > 0$. In this case the result is less elegant, since the arithmetic factors present in Lemma 3.3.4 do not smooth out in the final asymptotic. We do not pursue these computations here, but they can be used to recover Fujii's macroscopic result, Theorem 3.1.6.

From Corollary 3.3.6 it is an easy computation to see that

Lemma 3.3.7. *For η, σ and $n(T)$ as in Theorem 3.1.7, with η, σ, and k fixed, and with K a fixed continuous function supported in $(-1/k, 1/k)$ such that $K(0) = 1$,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} \check{K}_{n(T)} * \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right]^k dt = (c_k + o(1)) \Big[\int_{-n(T)}^{n(T)} |x| |\hat{\eta}(x)|^2 \, dx \Big]^{k/2}. \tag{3.9}$$

Proof. Note that $[\check{K}_{n(T)} * \eta(\frac{\cdot}{n(T)})]\hat{}\,(\xi) = n(T) K(\xi) \hat{\eta}(n(T)\xi)$. By Corollary 3.3.6, for K chosen to be supported in $(-1/k, 1/k)$ we have the left hand side of (3.9) is

$$(c_k + o(1)) \Big[\int_{\mathbb{R}} K^2\Big(\frac{x}{n(T)}\Big) |x| \cdot |\hat{\eta}(x)|^2 \, dx \Big]^{k/2}.$$

Because η is of bounded variation, $\hat{\eta}(x) = O(1/x)$, and for any $c_1 > c_2 > 0$,

$$\int_{c_1 n(T)}^{c_2 n(T)} |x||\hat{\eta}(x)|^2 dx \lesssim \log(c_1/c_2) = o\Big(\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 dx\Big),$$

since this latter integral diverges. [3] As we have that when $x \to 0$, $K^2(x) = 1+o(1)$,

$$\int_{\mathbb{R}} K^2\Big(\frac{x}{n(T)}\Big)|x| \cdot |\hat{\eta}(x)|^2\, dx \sim \int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2\, dx.$$

□

3.4 An Upper Bound

We will be able to complete the proofs of our central limit theorems by showing that the left hand side of (3.9) is a good approximation to the left hand side of the equation in Theorem 3.1.7. We accomplish this mainly through the use of the following upper bound

Theorem 3.4.1. *For σ as in Lemma 3.3.4,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\Big[\int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi}(\xi - t)\big)\tilde{d}(\xi)d\xi\Big]^k dt$$
$$\lesssim_k \int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\Big[\int_{-\infty}^{\infty} M_k\eta\big(\tfrac{\log T}{2\pi}(\xi - t)\big)\log(|\xi|+2)\, d\xi\Big]^k dt, \qquad (3.10)$$

with

$$M_k\eta(\xi) = \sum_{\nu=-\infty}^{\infty} \sup_{I_k(\nu)} |\eta| \cdot \mathbf{1}_{I_k(\nu)}(\xi),$$

where for typographical reasons we have denoted the interval $[k\nu - k/2, k\nu + k/2)$ by $I_k(\nu)$, and the order of our bound depends upon k, $\|\hat{\sigma}\|$ and the region in which $\hat{\sigma}$ can be supported.

[3]Even in the case it converges this o-bound is true, albeit for a different reason.

Proof. We make use of the Fourier pair $V(\xi) = \left(\frac{\sin \pi\xi}{\pi\xi}\right)^2$ and $\hat{V}(x) = (1 - |x|)_+$. Note that

$$\eta(\xi) \lesssim \sum_{\nu} \sup_{I_k(\nu)} |\eta| \underbrace{V\left(\tfrac{\xi-\nu}{k}\right)}_{:=V_{\nu,k}(\xi)}.$$

The right hand side of this is similar to $M_k\eta$ and we denote it by $M_k'\eta$. What is important about the scaling is that $\hat{V}_{\nu,k}$ is supported in $(-1/k, 1/k)$. Note that the left hand side of (3.10) is bounded by

$$\begin{aligned} &\lesssim \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} M_k'\eta\left(\tfrac{\log T}{2\pi}(\xi - t)\right) \tilde{d}(\xi) d\xi \right]^k dt \\ &\lesssim [A^{1/k} + B^{1/k}]^k, \end{aligned}$$

where

$$\begin{aligned} A &= \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} M_k'\eta\left(\tfrac{\log T}{2\pi}(\xi - t)\right) dS(\xi) \right]^k dt, \\ B &= \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} M_k'\eta\left(\tfrac{\log T}{2\pi}(\xi - t)\right) \log(|\xi| + 2) d\xi \right]^k dt, \end{aligned}$$

by Minkowski, and the fact that $\Omega(\xi)/2\pi = O\big(\log(|\xi| + 2)\big)$.

By the restricted range of support for $\hat{V}_{\nu,l}$ and Lemmas 3.3.4 and 3.3.5, for integers $\nu_1, ..., \nu_k$

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{k} \left(\int_{-\infty}^{\infty} V_{\nu_\ell,k}\left(\tfrac{\log T}{2\pi}(\xi - t)\right) dS(\xi_\ell) \right) dt = O_k(1).$$

Whence, taking a multilinear sum,

$$\begin{aligned} A &\lesssim_k \prod_{\ell=1}^{k} \sum_{\nu} \sup_{I_k(\nu)} |\eta| \\ &\lesssim B \end{aligned}$$

as $\log(|\xi| + 2) \gtrsim 1$.

Finally,

$$M_k'\eta(\xi) \lesssim \sum_{\mu=-\infty}^{\infty} \frac{1}{1+\mu^2} M_k\eta(\xi+\mu),$$

so using $\log(|\xi+\mu|+2) \lesssim \log(|\xi|+2)\log(|\mu|+2)$,

$$B \lesssim \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \Big[\int_{-\infty}^{\infty} M_k\eta\big(\tfrac{\log T}{2\pi}(\xi-t)\big) \cdot \log(|\xi|+2)\, d\xi \Big]^k dt.$$

These estimates on A and B give us the result. □

This result should be viewed as a slight generalization of an $O_A(1)$ upper bound given by Fujii for the average number of zeros in an interval $[t, t+A/\log T]$ where t ranges from T to $2T$ [26].

3.5 Proof of Theorems 3.1.2 and 3.1.3

We are now finally in a position to prove our main results. We first prove Theorem 3.3.3, then pass to Theorem 3.1.7 (and hence to Theorem 3.1.2).

Proof of Theorem 3.3.3. We want to show that

$$E_T := \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \Big[\int_{\mathbb{R}} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi-t)\big) dS(\xi) \Big]^k - \Big[\int_{-\infty}^{\infty} \check{K}_{n(T)} * \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi-t)\big) dS(\xi) \Big]^k dt$$

is asymptotically negligible, where K is a fixed function that meets the conditions of Lemma 3.3.7. In part because k can be odd, we must use some care. To this end we have the following lemma.

Lemma 3.5.1. *For $(X, d\mu)$ a positive measure space, f, g real valued functions on X, and $k \geq 1$ an integer*

$$\Big| \int (f^k - g^k) d\mu \Big| \lesssim_k \|f-g\|_{L^k(d\mu)} \big(\|f\|_{L^k(d\mu)}^{k-1} + \|g\|_{L^k(d\mu)}^{k-1} \big).$$

Proof. If f^k and g^k are both almost everywhere the same sign, this is implied by Minkowski (with implied constant k). On the other hand, if f^k and g^k are almost always of opposite sign, the estimate is trivial. We can prove the lemma in general by breaking the integral over X into two integrals over these subcases, and combine our estimates by noting that for positive a and b, $a^\alpha + b^\alpha \leq 2\max(a^\alpha, b^\alpha) \lesssim (a+b)^\alpha$, where (in our case) $\alpha = (k-1)/k$. □

This leads us to consider

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} (\eta - \check{K}_{n(T)} * \eta)\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right]^k dt. \tag{3.11}$$

Trivally, this is bounded by

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} \Big| (\eta - \check{K}_{n(T)} * \eta)\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) \Big| |dS|(\xi) \right]^k dt, \tag{3.12}$$

which by Theorem 3.4.1 is bounded by

$$\begin{aligned}
&\lesssim \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\int_{-\infty}^{\infty} M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) \cdot \log(|\xi| + 2) d\xi \right]^k dt \\
&= \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left[\tfrac{2\pi n(T)}{\log T} \int_{-\infty}^{\infty} M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)(\xi) \log\left(\big|t + \tfrac{2\pi n(T)}{\log T}\xi\big| + 2\right) d\xi \right]^k dt \\
&\lesssim \left(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \frac{\log^k(|t|+2)}{\log^k T} dt \right) \left[n(T) \int_{-\infty}^{\infty} M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)\big(\xi\big) d\xi \right]^k \\
&\quad + \left[\frac{2\pi n(T)}{\log T} \int_{-\infty}^{\infty} M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)\big(\xi\big) \log(|\xi| + 2) d\xi \right]^k .
\end{aligned}$$

Note, if we label $L(\xi) = \log(|\xi|+2)$, we have $M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)\big(\xi\big) \log(|\xi|+2) \leq M_{k/n(T)}\big[(\eta - \check{K}_{n(T)} * \eta)L\big]\big(\xi\big)$.

At this point we make use of the fact that η is of bounded variation. Because

η has compact support,

$$\int \log(|\xi|+2)|d\eta(\xi)| < +\infty.$$

In addition, $\check{K}_{n(T)} * \eta$ is bounded in variation for the same reason that

$$\int \log(|\xi|+2)\big|d\check{K}_{n(T)} * \eta(\xi)\big| = K(0) \int \log(|\xi|+2)|d\eta(\xi)| < +\infty.$$

By the product rule then, $\mathrm{var}\big[(\eta - \check{K}_{n(T)} * \eta)L\big]$ is bounded, for $\mathrm{var}(\cdot)$ the total variation.

We have the following three lemmas:

Lemma 3.5.2. *For $f \in L^1(\mathbb{R})$ and of bounded variation $\mathrm{var}(f)$, and K as above,*

$$\|f - \check{K}_H * f\|_{L^1} \lesssim \frac{\mathrm{var}(f)}{H}.$$

The proof is utterly standard, but I was unable to find a reference. The key point is that K is smooth and compact, so that $|x||\check{K}(x)|$ is integrable.

Proof. Note that $\check{K}_H(x) = H\check{K}(Hx)$, so

$$\begin{aligned}
\|f - \check{K}_H * f\|_{L^1} &= \left\| \int H\check{K}(H\tau)f(t)\,d\tau - \int H\check{K}(H\tau)f(t-\tau)\,d\tau \right\|_{L^1(dt)} \\
&\le H \int \check{K}(H\tau)\|f(t) - f(t-\tau)\|_{L^1(dt)} d\tau \\
&\le H \int \check{K}(H\tau)\Big(\int_{\mathbb{R}} \int_{-\tau}^{0} |df(t+h)|\,dh\,dt \Big) d\tau \\
&= H \int \check{K}(H\tau)|\tau| d\tau \cdot \mathrm{var}(f) \\
&\lesssim \frac{\mathrm{var}(f)}{H}.
\end{aligned}$$

□

Likewise, because $|\check{K}(x)||x|^2$ is integrable, and $|\check{K}(x)||x|\log(|x|+2)$ is of order $|\check{K}(x)||x|$ around $x=0$ and is bounded up to a constant by $|\check{K}(x)||x|^2$ otherwise, we have similarly,

Lemma 3.5.3.

$$\|f - \check{K}_H * f\|_{L^1(\log(|t|+2)dt)} \lesssim \frac{1}{H}\int_{\mathbb{R}} \log(|t|+2)|df(t)|.$$

Finally,

Lemma 3.5.4. *For f of bounded variation, and any $\varepsilon > 0$,*

$$\sum_{k=-\infty}^{\infty} \varepsilon\|f\|_{L^\infty\left(\varepsilon[k-1/2,k+1/2)\right)} \lesssim \|f\|_{L^1} + \varepsilon \cdot \mathrm{var}(f).$$

Proof. For arbitrarily small ε', we can choose $x_k \in \varepsilon[k-1/2, k+1/2)$ so that $|f(x_k)|$ is sufficiently close to $\|f\|_{L^\infty\left(\varepsilon[k-1/2,k+1/2)\right)}$ that

$$\begin{aligned}\sum_{k=-\infty}^{\infty} \varepsilon\|f\|_{L^\infty\left(\varepsilon[k-1/2,k+1/2)\right)} &\le \varepsilon' + \varepsilon\sum_k |f(x_k)| \\ &\le \varepsilon' + \sum_j (x_{2j+2}-x_{2j})|f(x_{2j})| + \sum_{j'}(x_{2j'+1}-x_{2j'-1})|f(x_{2j-1})|.\end{aligned}$$

More,

$$\begin{aligned}\left|\int |f|dx - \sum_j (x_{2j+2}-x_{2j})|f(x_{2j})|\right| &\le \sum_j \int_{x_{2j}}^{x_{2j+2}} \Big||f(x)| - |f(x_{2j})|\Big| dx \\ &\le \sum_j (x_{2j+2}-x_{2j}) \int_{x_{2j}}^{x_{2j+2}} |df(x)| \\ &\le 3\varepsilon \cdot \mathrm{var}(f)\end{aligned}$$

as $(x_{2j+2} - x_{2j}) \le 3\varepsilon$ always. The same estimate holds for a sum over odd indices, and we have then

$$\sum_k \varepsilon \|f\|_{L^\infty\left(\varepsilon[k-1/2,k+1/2)\right)} \leq \varepsilon' + 6\varepsilon \cdot \mathrm{var}(f) + 2\int |f| dx.$$

As ε' was arbitrary, the lemma follows. □

Making use of these lemmas we have that

$$\int_{-\infty}^{\infty} M_{k/n(T)}(\eta - \check{K}_{n(T)} * \eta)\big(\xi\big) d\xi \lesssim_{\eta,k} \frac{1}{n(T)},$$

and

$$\int_{-\infty}^{\infty} M_{k/n(T)}\big[(\eta - \check{K}_{n(T)} * \eta) \cdot L\big]\left(\xi\right) d\xi \lesssim_{\eta,k} \frac{1}{n(T)}.$$

Hence (3.11) is bounded. By Lemma 3.5.1, with the averages over t with respect to σ playing the role of the positive measure μ,

$$\begin{aligned} E_T \lesssim_{\eta,k} & \left(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left| \int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right|^k dt \right)^{(k-1)/k} \\ & + \left(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left| \int_{-\infty}^{\infty} \check{K}_{n(T)} * \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right|^k dt \right)^{(k-1)/k}. \end{aligned} \tag{3.13}$$

For k even, this implies by Lemma 3.3.7 (our Fourier truncation central limit theorem), and the fact that $\int |x||\hat{\eta}|^2 dx = +\infty$,

$$\begin{aligned} & \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \Big[\int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \Big]^k dt \\ & = (c_k + o(1)) \Big[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 \, dx \Big]^{k/2} \\ & \quad + O\left[\left(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left| \int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right|^k dt \right)^{(k-1)/k} \right] \end{aligned}$$

This bound implies the left hand side diverges, and thus the conclusion of Theorem 3.3.3 for even k. For odd k, by Hölder (or Cauchy-Schwartz) and the result we have just proved for even k,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left| \int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right|^k dt \leq (\sqrt{c_{2k}} + o(1)) \Big[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 \, dx \Big]^{k/2}, \tag{3.14}$$

and hence, using (3.13) again, Theorem 3.3.3 for odd k as well. □

Proof of Theorem 3.1.7. To see that Theorem 3.3.3 implies Theorem 3.1.7, note that for any $\epsilon > 0$, we can find σ_1 of the sort delimited in Theorem 3.3.3, so that $\|\mathbf{1}_{[1,2]} - \sigma_1\|_{L^1} < \epsilon/2$. Further, we can find σ_2, a linear combination of translations and dilations of the function $\big(\frac{\sin \pi t}{\pi t}\big)^2$, so that σ_2 is non-negative and $|\mathbf{1}_{[1,2]}(t) - \sigma_1(t)| \leq \sigma_2(t)$ for all t, and $\|\sigma_2\|_{L^1} < \epsilon$. Note (for simplicity of notation) that (3.14) is true for even k as well, and by rescaling linearly, we have

$$\int_{\mathbb{R}} \frac{\sigma_2(t/T)}{T} \left| \int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \right|^k dt \leq \epsilon(\sqrt{c_{2k}} + o(1)) \Big[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 \, dx \Big]^{k/2}.$$

Then

$$\begin{aligned} &\int_{\mathbb{R}} \frac{\mathbf{1}_{[1,2]}(t/T)}{T} \Big[\int_{-\infty}^{\infty} \eta\big(\tfrac{\log T}{2\pi n(T)}(\xi - t)\big) dS(\xi) \Big]^k dt \\ &= [c_k + o(1) + \epsilon \cdot (O_k(1) + o(1))] \Big[\int_{-n(T)}^{n(T)} |x||\hat{\eta}(x)|^2 \, dx \Big]^{k/2}. \end{aligned}$$

(Note that here the $O_k(1)$ term is bounded absolutely by $\sqrt{c_{2k}}$.) As ϵ is arbitrary, the theorem follows. □

Proof of Theorem 3.1.3. A proof will follow almost exactly as before. We need only to show that Theorem 3.1.7 is true for η instead of the sort delimited in Theorem 3.1.3. The reader may check that the only part of the proof which requires modification is that the error term E_T, at the start of section 4, cannot be shown to be asymptotically negligible in the same way as before, since now asymptotically negligible means that $E_T = o(1)$. But using Lemma 3.5.1 in the same way as before, this will be the case, and therefore the theorem, so long as

$$\|\eta - \check{K}_H * \eta\|_{L^1} = o(1/H), \tag{3.15}$$

for some K as above. This is certainly the case for those η delimited in Theorem 3.1.2, using the fact that for such η, $\hat{\eta}(\xi) = o(1/(1+|\xi|)^2)$. □

Remark: (3.15) is true for a wider range of functions than $C_c^2(\mathbb{R})$; but it does not encompass the elegant criterion, "all functions which are of bounded variation and compactly supported." It is *not* the case for η a Cantor function, for instance. We expect the theorem to remain true in this case, but to prove this would seem to require upper bounds on correlation functions for zeta zeros with respect to oscillatory functions, extending outside the range of functions considered by Rudnick and Sarnak. Although here we require only upper bounds, not exact evaluations, this still goes beyond what we currently seem able to prove.

Let us conclude with this chapter on another similar point: that Selberg's approximation to $S(t)$, mentioned in the introduction, and therefore Fujii's Theorem's 3.1.1 and 3.1.6, are true unconditionally. The first of these claims was shown by Selberg, using a zero-density estimate to bound the number of zeroes lying off the critical line. I have been unable to extend this method to prove Theorem 3.1.2 unconditionally, where the points we are counting are the imaginary ordinates of non-trivial zeroes – zeroes which may in some instances lie off the critical line – and I leave it as a challenge for readers to do so.

CHAPTER 4

Mesoscopic and band-limited microscopic statistics: towards a more general mesoscopic theory

We include in this intercalary chapter a more general discussion of the statistics of the zeros of the zeta-function in the mesoscopic regime. Our discussion will culminate in Theorem 4.0.10, a statement from which one can deduce both the microscopic linear statistics of the sort considered by Rudnick and Sarnak and the central limit theorems discussed above, along with covariance statements for translated linear statistics separated by mesoscopic distances. Other theorems concerning the mesoscopic distribution of zeta zeros, which also depend upon the macroscopic statistics of the zeros, can be found in [5] and [47].

We show first that Corollary 3.3.6 implies the well-known result of Rudnick and Sarnak (building on work of Montgomery and concurrent with work of Hejhal) that, upon ordering the positive ordinates of the zeros $0 < \gamma_1 \leq \gamma_2 \leq ...$,

Theorem 4.0.5 (Rudnick-Sarnak)**.** *For* $\eta : \mathbb{R}^k \to \mathbb{R}$ *such that* $\operatorname{supp}\hat{\eta} \subseteq \{x \in \mathbb{R}^k : |x_1| + \cdots + |x_k| < 2\}$,

$$\lim_{T\to\infty} \frac{1}{T}\int_T^{2T} \sum_{\substack{i_1,...,i_k \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_{i_1}-t), ..., \tfrac{\log T}{2\pi}(\gamma_{i_k}-t)\Big)\, dt = \int_{\mathbb{R}^k} \eta(x) \det_{k\times k}[K(x_i-x_j)]\, d^k x,$$

where $K(\xi) = \frac{\sin \pi\xi}{\pi\xi}$.

That is to say, with respect to sufficiently smooth functions, the zeros of the zeta-function tend weakly to a determinantal point process with sine-kernel.

One may do this either through a combinatorial sieving procedure – effectively this is the proof of Rudnick and Sarnak – or alternatively one may use the combinatorics of Diaconis and Shahshahani. For us, it will be more enlightening to use the latter. Proceeding in this manner originated with Hughes and Rudnick, although our range of test functions will coincide with the slightly wider range used originally by Rudnick and Sarnak.

The theorem of Diaconis and Shahshahani we will need is

Theorem 4.0.6 (Diaconis-Shahshahani)**.** *Let $\mathcal{U}(n)$ be the set of $n \times n$ unitary matrices endowed with Haar measure. Consider $a = (a_1, ..., a_k)$ and $b = (b_1, ..., b_k)$ with $a_1, a_2, ..., b_1, b_2, ... \in \{0, 1, ...\}$.If $\sum_{j=1}^{k} ja_j \neq \sum_{j=1}^{k} jb_j$,*

$$\int_{\mathcal{U}(n)} \prod_{j=1}^{k} \mathrm{Tr}(g^j)^{a_j} \overline{\mathrm{Tr}(g^j)^{b_j}}\, dg = 0. \tag{4.1}$$

Furthermore, in the case that

$$\max\left(\sum_{j=1}^{k} ja_j, \sum_{j=1}^{k} jb_j\right) \leq n$$

we have

$$\int_{\mathcal{U}(n)} \prod_{j=1}^{k} \mathrm{Tr}(g^j)^{a_j} \overline{\mathrm{Tr}(g^j)^{b_j}}\, dg = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j! \tag{4.2}$$

In addition, for unrestricted a

$$\int_{\mathcal{U}(n)} \prod_{j=1}^{k} \left|\mathrm{Tr}(g^j)\right|^{2a_j} dg \leq \prod_{j=1}^{k} j^{a_j} a_j!$$

but we will not need this fact. In general, for products of traces outside of the restricted range of the theorem, no pattern emerges which is as nice as (4.2). Since our restricted range here corresponds – as we will show shortly – to the only range of test functions for which the statistics of the zeta-function's zeros can be rigorously evaluated, this fact must be seen as somewhat curious.

Here trace is defined in the standard way, so that $\mathrm{Tr}(I_{n\times n}) = n$. For a proof of Theorem 4.0.6, see [19] or [9].

It is a simple exercise in enumerative combinatorics to see that (4.1) and (4.2) imply that for $|j_1| + \cdots + |j_k| \leq 2n$

$$\int_{\mathcal{U}(n)} \prod_{\ell=1}^{k} \mathrm{Tr}(g^{j_\ell})\, dg = \sum \prod_{\lambda} |j_{\mu_\lambda}|\, \delta(j_{\mu_\lambda} = -j_{\nu_\lambda})$$

where once again the sum is over all partitions of $[k]$ into disjoint pairs $\{\mu_\lambda, \nu_\lambda\}$, and $\delta(j_{\mu_\lambda} = -j_{\nu_\lambda})$ is 1 or 0 depending upon whether $j_{\mu_\lambda} = -j_{\nu_\lambda}$ or not.

We are able to use this to study the determinantal point process with sine kernel because the eigenvalues of a random unitary matrix, properly spaced, are themselves a determinantal point process with kernel tending to that of the sine kernel. This is due, in effect, to Weyl, though Gaudin and Dyson deserve credit for the formulation in terms of correlation functions that follows.

Proposition 4.0.7. *Let $\{e(\theta_1), e(\theta_2), ..., e(\theta_n)\}$ be the eigenvalues of a random unitary matrix, distributed according to Haar measure, with $\theta_i \in [-1/2, 1/2)$ for all i. Then the points $\{n\theta_1, ..., n\theta_n\}$ comprise a determinantal point process $\mathcal{S}_n$ on $[-n/2, n/2)$ with kernel in x, y given by $K_n(x-y) = \frac{\sin \pi(x-y)}{n \sin(\pi(x-y)/n)}$. That is for any test function η,*

$$\begin{aligned}\mathop{\mathbf{E}}_{\mathcal{S}_n} \sum_{\substack{i_1,...,i_k \\ \text{distinct}}} \eta(\xi_{i_1}, ..., \xi_{i_k}) &= \int_{\mathcal{U}(n)} \sum_{\substack{i_1,...,i_k \\ \text{distinct}}} \eta(n\theta_{i_1}, ..., n\theta_{i_k})\, dg \\ &= \int_{[-n/2,n/2]^k} \eta(x_1, ..., x_k) \det_{k\times k}[K_n(x_i - x_j)]\, d^k x\end{aligned}$$

For further discussion see [12].

We use this to prove

Theorem 4.0.8. *If $\mathcal{S}$ is the determinantal point process with kernel in x, y given by $K(x-y)$ for $K(x) = \frac{\sin \pi x}{\pi x}$, then for functions $\eta_1, ..., \eta_k$ such that, as in Lemma 3.3.4,* $\operatorname{supp} \eta_\ell \in [-\delta_\ell, \delta_\ell]$ *with $\delta_1 + \cdots + \delta_k \leq 2$,*

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \left(\Delta_{\eta_\ell} - \mathop{\mathbf{E}}_{\mathcal{S}} \Delta_{\eta_\ell} \right) = S_{[k]} \tag{4.3}$$

where $S_{[k]}$ is defined as in Corollary 3.3.6, and here $\Delta_\eta = \sum \eta(\xi_i)$ as before, for $\{\xi_i\}$ the points of the process.

Note that here, by definition, $\mathbf{E}\, \Delta_\eta = \int \eta\, dx$.

Before we come to the proof, we note that as an easy consequence, upon expanding the product in 4.3 and applying induction,

Corollary 4.0.9. *A point process $\mathcal{P}$ satisfies* (4.3) *for all k over the range of test functions restricted as in Theorem 4.0.8 if and only if for all k and for any integrable η defined on $\mathbb{R}^k$ with* $\operatorname{supp} \hat{\eta} \subseteq \{y \in \mathbb{R}^k : |y_1| + \cdots + |y_k| \leq 2\}$,

$$\mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1, ..., i_k \\ \text{distinct}}} \eta(\xi_{i_1}, ..., \xi_{i_k}) = \int_{\mathbb{R}^k} \eta(x_1, ..., x_k) \det_{k \times k} [K(x_i - x_j)] d^k x.$$

Proof of Theorem 4.0.8. For a function η, define

$$\eta^{(n)}(\theta) = \sum_{k \in \mathbb{Z}} \eta(\theta + nk).$$

Note that for Schwartz η, $\eta^{(n)} \to \eta$ uniformly. We have then that for fixed

Schwartz $\eta_1, ..., \eta_k$,

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\Delta_{\eta_\ell} - \mathop{\mathbf{E}}_{\mathcal{S}} \Delta_{\eta_\ell} \Big) = \lim_{n\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}_n} \prod_{\ell=1}^{k} \Big(\Delta_{\eta_l^{(n)}} - \mathop{\mathbf{E}}_{\mathcal{S}_n} \Delta_{\eta_l^{(n)}} \Big)$$

$$= \lim_{n\to\infty} \int_{\mathcal{U}(n)} \prod_{\ell=1}^{k} \bigg(\sum_{\nu=1}^{n} \eta_\ell^{(n)}(n\theta_\nu) - n \int_{-1/2}^{1/2} \eta_\ell^{(n)}(n\theta)\, d\theta \bigg)\, dg$$

But by Poisson summation,

$$\eta_\ell^{(n)}(n\theta_\nu) - \int_{-1/2}^{1/2} \eta_\ell^{(n)}(n\theta)\, d\theta = \sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{n} \hat{\eta}_\ell\Big(\frac{j}{n}\Big) e(j\theta),$$

so that

$$\int_{\mathcal{U}(n)} \prod_{\ell=1}^{k} \bigg(\sum_{\nu=1}^{n} \eta_\ell^{(n)}(n\theta_\nu) - n \int_{-1/2}^{1/2} \eta_\ell^{(n)}(n\theta)\, d\theta \bigg)\, dg$$
$$= \int_{\mathcal{U}(n)} \prod_{\ell=1}^{k} \sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{n} \hat{\eta}_\ell\Big(\frac{j}{n}\Big) \mathrm{Tr}(g^j)\, dg$$
$$= \sum_{j_1,...,j_k\in\mathbb{Z}\setminus\{0\}} \prod_{\ell=1}^{k} \frac{1}{n} \hat{\eta}_\ell\Big(\frac{j_\ell}{n}\Big) \cdot \int_{\mathcal{U}(n)} \prod_{\ell=1}^{k} \mathrm{Tr}(g^{j_\ell})\, dg.$$

But for $\hat{\eta}_1, ..., \hat{\eta}_k$ restricted as in the Theorem, this sum is only over those j with $|\frac{j_1}{n}| + \cdots + |\frac{j_k}{n}| \leq 2$. In this case the above sum reduces to

$$\sum \prod_{\lambda} \Big(\sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{n} \frac{|j|}{n} \hat{\eta}_{\mu_\lambda}\Big(\frac{j}{n}\Big) \hat{\eta}_{\nu_\lambda}\Big(\frac{-j}{n}\Big) \Big).$$

Clearly this tends to $S_{[k]}$. □

Proof of Theorem 4.0.5. Using Corollary 3.3.6 for $n(T) = 1$, we have for $\eta_1, ..., \eta_\ell$ as in Theorem 4.0.8,

$$\lim_{T\to\infty} \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{k} \int_{-\infty}^{\infty} \eta\Big(\tfrac{\log T}{2\pi}(\xi_\ell - t)\Big)\, dS(\xi_\ell)\, dt = S_{[k]}.$$

But by Stirling's formula,

$$\int_{-\infty}^{\infty} \eta\Big(\tfrac{\log T}{2\pi}(\xi_\ell - t)\Big)\, dS(\xi) = \sum_{\gamma} \eta\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big) - \int \eta(x)\, dx + o(1).$$

Expanding the product as in Corollary 4.0.9, and passing from σ to $\mathbf{1}_{[1,2]}$ as before yeilds the claim for $\eta = \eta_1 \otimes \cdots \otimes \eta_k$. We can pass to general η by uniformly approximating such η and using Theorem 3.4.1 to bound the difference between the linear statistics of η and those of its approximation. □

The convergence here is microscopic, and therefore cannot, unless spread over a wider region as in Corollary 4.0.8, yield a mesoscopic central limit theorem like Fujii's or Theorem 3.1.2. In a general way, it does appear that in the mesoscopic regime, the zeros of the zeta function are spaced like the points of a sine-kernel determinantal point process – and that moreover we have knowledge of this fact as long as any test functions used remain microscopically band-limited. Stating this principle in a way which is both (i) precise, and (ii) satisfying, is a rough task however. We shall make an attempt below, but we should be forthright that it is only the first of these conditions and not the second that is really achieved. Before proceeding, it is worthwhile to discuss the matter heuristically somewhat further.

We say that a point process $\mathcal{P}$ is "mock-determinantal with sine-kernel" if its correlation functions agree with that of $\mathcal{S}$ with respect to sufficiently smooth test functions; that is

$$\mathop{\mathbf{E}}_{\mathcal{P}} \sum_{\substack{i_1,\ldots,i_k \\ \text{distinct}}} \eta(\xi_{i_1},\ldots,\xi_{i_k}) = \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1,\ldots,i_k \\ \text{distinct}}} \eta(\xi_{i_1},\ldots,\xi_{i_k})$$

with respect to – say for our purposes – η with Fourier transform $\hat{\eta}$ supported on $\{x \in \mathbb{R}^k : |x_1| + \cdots + |x_k| \leq 2\}$. Using the proof above for the zeros of the zeta function, one can show that Theorems 3.1.2 and 3.1.3 hold for any such $\mathcal{P}$.

That is for η restricted as in either theorem, a parameter L which grows, and $\Delta_\eta = \sum \eta(\xi_i/L)$,

$$\frac{\Delta_\eta - \mathbf{E}\,\Delta_\eta}{\sqrt{\mathrm{Var}\Delta_\eta}} \Rightarrow N(0,1),$$

as $L \to \infty$. (As here we are dealing with a single point process $\mathcal{P}$, 'mesoscopic' restrictions on the growth of L play no role.) We may ask whether there exists any such mock-determinantal point processes $\mathcal{P}$ for which η is of bounded variation, but $(\Delta_\eta - \mathbf{E}\,\Delta_\eta)/\sqrt{\mathrm{Var}\Delta_\eta}$ does not tend to the normal distribution. I do not know the answer to this, but I suspect that there does. This would imply that to fill the small gap between Theorems 3.1.2 and 3.1.3 and their random matrix analogues will require (a small amount of) statistical information about the zeros of the zeta function outside of that provided by test functions which are band-limited as in Rudnick-Sarnak.

We return to our goal of characterizing the zeros of the zeta-function in the mesoscopic regime in a way that retains microscopic statistics as well. We have:

Theorem 4.0.10. *Let σ be as in Theorem 3.3.3, and let $Z_T(\sigma)$ be the point process defined by the points $\{\frac{\log T}{2\pi}(\gamma - t)\}$ where γ runs through the ordinates of zeros of the zeta function, and t is a random variable in $\mathbb{R}$ with distribution given by $\sigma(t/T)/T$. For fixed $A < 2$, fixed r of compact Fourier support, and fixed $n(T)$ with $n(T) \to \infty$ but with $n(T) = o(\log T)$, we have that for $|\alpha_1| + \cdots + |\alpha_k| \leq A$,*

$$\mathop{\mathbf{E}}_{Z_T(\sigma)} \prod_{\ell=1}^{k} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big) = \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big) + \sum_{\emptyset \subseteq J \subsetneq [k]} \varepsilon([k]\backslash J) \cdot \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell \in J} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big),$$

where

$$\Delta_\ell = \sum r\Big(\frac{\xi_i}{n(T)}\Big) e(\alpha \xi_i)$$

for the terms $\varepsilon([k]\backslash J)$ having no dependence on α_i with $i \in J$, and tending to 0 uniformly as $T \to \infty$.

This may be proven by following exactly the proof of Corollary 3.3.6. By slightly modifying the proof, one may prove this theorem even for $\sigma = \mathbf{1}_{[1,2]}$ so that t is uniformly distributed between T and $2T$, but we do not pursue this matter here. By integrating in α, one can obtain microscopic and macroscopic statistics, and correlations thereof, uniformly for points separated by a distance asymptotically less than $m(T)$. One can, for instance, recover Corollary 3.3.6 for $n(T) = o(m(T))$ in this way. We are able to integrate in α without destroying error terms for the reason that $\varepsilon([k]\backslash J)$ has no dependence on α_i for $i \in J$.

In the same way, by modifying the proof of Theorem 4.0.8,

Theorem 4.0.11. *For fixed $A < 2$, fixed r of compact Fourier support, and fixed $n(N)$ with $n(N) \to \infty$ but with $n(N) = o(N)$, we have that for $|\alpha_1|+\cdots+|\alpha_k| \leq A$,*

$$\mathop{\mathbf{E}}_{\mathcal{S}_N} \prod_{\ell=1}^{k} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big) = \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big) + \sum_{\emptyset \subseteq J \subsetneq [k]} \varepsilon([k]\backslash J) \cdot \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell \in J} \Big(\Delta_\ell - \mathbf{E}\,\Delta_\ell\Big),$$

for Δ_ℓ (defined in the obvious way with respect to $n(N)$), and ε as in Theorem 4.0.10.

To have a more eloquent expression of the mesoscopic convergence expressed by these results would certainly be desirable.

CHAPTER 5

Microscopic statistics: Arithmetic consequences of the GUE conjecture

5.1 Background material

Recall once more the GUE Conjecture:

Conjecture 5.1.1 (GUE). *For any fixed n and any fixed $\eta \in C_c(\mathbb{R}^k)$,*

$$\lim_{T\to\infty} \frac{1}{T}\int_T^{2T} \sum_{\substack{\gamma_1,\ldots,\gamma_n\\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), \ldots, \tfrac{\log T}{2\pi}(\gamma_n - t)\Big)\, dt = \int_{\mathbb{R}^n} \eta(x) \det_{n\times n}\Big(K(x_i - x_j)\Big)\, d^n x \tag{5.1}$$

where the entries of the $n\times n$ determinant are formed from the function $K(x) = \frac{\sin \pi x}{\pi x}$.

The sum on the left is over all collections of distinctly labelled ordinates $\gamma_1, \ldots, \gamma_n$.

The compact support of test functions η means that our sums are effectively restricted to those γ's that are within $O(1/\log T)$ of the variable t. At such a scale, we have shown in the last chapter in Theorem 4.0.5 that we can rigorously verify equation (5.1) if η is restricted to a stringently smooth class of test functions.

It is a matter of longstanding interest, however, to see what can be said about the n-level correlation sums on the left hand side of (5.1) for functions not as smooth as those in Theorem 4.0.5 once additional assumptions have been made about the distribution of the primes. Even in the original paper of Montgomery, the $n = 2$ pair correlation conjecture for a wider class of test functions was supported on the assumption of a uniform version of a the Hardy-Littlewood conjecture about the likelihood that two primes are separated by a small distance h. (This argument appears in [54].)

An especially relevant result in this direction is the following:

Theorem 5.1.2 (Gallagher & Muller, and Goldston)**.** *(On RH.) The $n = 2$ pair correlation conjecture is equivalent to the statement that for fixed $\beta \geq 1$, as $T \to \infty$,*

$$\int_1^{T^\beta} \Big(\psi\big(x + \tfrac{x}{T}\big) - \psi(x) - \tfrac{x}{T}\Big)^2 \frac{dx}{x^2} \sim (\beta - \tfrac{1}{2}) \frac{\log^2 T}{T}. \tag{5.2}$$

The prime number theorem is a statement that the 'mean value' of $\psi(x)$ is x, so that this is a weighted estimate for the variance of the number of primes in short intervals $(x, x + x/T)$. That the pair correlation conjecture implies it is due to Gallagher and Mueller [30], the reverse implication to Goldston [32].

Unconditionally, for $\beta \leq 1$ the left hand side of (5.2) can be seen using the prime number theorem to be asymptotic to

$$\frac{\beta^2}{2} \frac{\log^2 T}{T}.$$

The somewhat unnatural weight dx/x^2 was removed in the work of Goldston and Montgomery [35], who showed that (on RH) a slightly stronger variant of the pair correlation conjecture is equivalent to a somewhat more naturally weighted estimate for the variance of primes in short intervals:

$$\frac{1}{X} \int_1^X \big(\psi(x + H) - \psi(x) - H\big)^2 \, dx \sim H \left(\log X - \log H\right) \tag{5.3}$$

uniformly for $X^\epsilon \leq H \leq X^{1-\epsilon}$ (for any fixed $\epsilon > 0$). The survey [33] is a nice introduction to this and other material.

We mention that the counts $(\psi(x+H) - \psi(x) - H)$ for x a random variable uniformly distributed between 1 and large X are widely expected to be normally distributed with variance given by (5.3) [55], though its higher moments are not directly related to the *local* statistics of zeros dealt with by Conjecture 5.1.1.

A computation reveals that neither (5.2) nor (5.3) are consistent with a heuristic model of Cramér [18] (see also [36], [73]) for the distribution of primes: that each number m has, roughly, an independent probability of $1/\log m$ of being prime. In these matters it is the predictions (5.2) and (5.3), rather than the Cramér model, that is widely expected to return the right answer. The Cramér model accurately predicts the Riemann hypothesis prediction that the error term in a count of primes in the interval $[1, x]$ is $O(x^{1/2+\epsilon})$, but quite apparently to accurately answer asymptotic questions about the distribution of primes in shorter intervals $[x, x+H]$ one must use a model of the primes that takes into account local arithmetic considerations.

Indeed, for higher correlations, Bogomolny and Keating [2], [3] argued heuristically that the m-level correlations correspond arithmetically to the likelihood that products of primes $p_1 \cdots p_\ell$ (each prime chosen from a specified region) are separated by a small distance from products of primes $p_{\ell+1} \cdots p_m$ (again with each prime drawn from a specified region) and that this likelihood – and therefore the GUE conjecture – can be understood as before by using Hardy-Littlewood conjectures. These predict the probability in terms of a, b, and h that both p_1 and p_2 are prime, given that $ap_1 - bp_2 = h$, where p_1 and p_2 are of order x. The prediction is not $1/\log^2 x$, as one might guess from a naïve use of the Cramér model.

It is thus a matter of longstanding interest to generalize the work mentioned above in for instance Theorem 5.1.2 from the pair correlation conjecture to higher order correlations, and this is the purpose of the present paper.

While this work was in progress, I learned that work on this same question had recently been undertaken by Farmer, Gonek, Lee and Lester [25]. Conditioned in addition to RH on technical hypotheses about the zeta zeros which they define and label Hypothesis AC and Hypothesis LC, the authors arrive at a solution in one direction, showing that knowing a Fourier-transformed evaluation of the n-point correlation sums in (5.1) (the *n-level form factor*) is sufficient to estimate the likelihood that products of primes in the fashion of Bogomolny and Keating are close to other products of primes.

Additionally motivated by the work of Goldston, Gonek, and Montgomery [34], the authors show conditioned on RH and Hypotheses AC and LC that knowing the n-level form factor for all n is sufficient to asymptotically evaluate

$$\frac{1}{T}\int_T^{2T} \prod_{\ell=1}^{j} \frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A_\ell}{\log T}+it\Big) \prod_{\ell'=1}^{k} \frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{B_{\ell'}}{\log T}-it\Big)\, dt \tag{5.4}$$

for positive constants $A_1, ..., A_j,\ B_1, ..., B_k$. Random matrix theory makes a prediction that this quantity will be asymptotic to a constant depending on the A_ℓ's and $B_{\ell'}$'s multiplied by $\log^{j+k} T$. The authors note that in the case $n = 2$ (considered in [34]) one can proceed in the converse direction, showing that the pair correlation conjecture follows form a conjectured asymptotic evaluation of (5.4) when $j = k = 1$, and write, "It would be interesting to know whether this generalizes to higher N."

Finally, Farmer, Gonek, Lee, and Lester show that by assuming Hypothesis AC and LC in addition to RH one can bound for any fixed $A > 0$,

$$\frac{1}{T}\int_T^{2T} \left|\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A}{\log T}+it\Big)\right|^k dt \lesssim \log^k T. \tag{5.5}$$

By Hölder this implies (5.4) is bounded by $O(\log^{j+k} T)$. On RH, the authors note, this is a correct lower bound.

We present in this chapter work that makes the advance of not requiring the hypotheses AC and LC for our analysis. We show in Theorem 5.2.4 that (on

RH) the GUE Conjecture, as stated in Conjecture 5.1.1 is logically equivalent to a statement about the likelihood that products of primes drawn from certain regions are close to other products of primes, and in Theorem 5.2.2 that the GUE Conjecture is equivalent to an asymptotic evaluation of (5.4). We thereby answer in the affirmative the question posed above by Farmer, Gonek, Lee, and Lester about whether these products are sufficient to characterize the local statistics of the zeta zeros.

Our techniques additionally yield (5.5) on the assumption of RH but no other hypothesis.

The work makes it possible to restate for instance the $k = 3, 4$ triple and quadruple correlation conjectures for the zeta zeros in terms of the distribution of prime numbers. Unfortunately the resulting statements about the primes are complicated algebraically. We draw out one elegant corollary of our work, however: that the GUE Conjecture implies a simple estimate for the covariance of almost primes in short intervals, where almost primes are weighted by the higher von Mangoldt functions famously used by Selberg and Erdős in an elementary proof of the prime number theorem [67],[21].

This is related to a conjecture first made by Farmer [24] concerning the average value of a ratio of products of the zeta function, and to arrive at our estimate we will make use of some elementary combinatorics that are in fact equivalent to an analogous ratio *theorem* for autocorrelations of characteristic polynomials over the unitary group.

5.2 A statement of main results

We obtain in the first place,

Theorem 5.2.1. *(On RH.) For fixed $k \geq 1$ and constant A with $\Re A > 0$,*

$$\frac{1}{T}\int_T^{2T} \left|\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A}{\log T} + it\Big)\right|^k dt \lesssim_{A,k} \log^k T.$$

Remark: As noted in [25] once can obtain this for $A \geq 4$ by using Lemma 3 of Selberg's paper [64]. In fact, using instead Lemma 2 of this paper of Selberg together with an upper bound due to Fujii (see Theorem 5.6.2), one can obtain exactly this theorem, for A arbitrarily close to 0 as above. We give a proof of Theorem 5.2.1 independent of Selberg's identity, since this will at any rate fit naturally into our framework, though we outline what the approach through Selberg's identity would look like. In some sense any possible proof must hinge upon the same ideas.

With sufficient effort one can trace through the implied constant in Theorem 5.2.1 in terms of A and k, obtaining a constant for positive real A of order

$$A^{-k}e^{O(k\log k)}$$

One should not expect this to be an optimal constant, or even necessarily the limit to which analysis on RH can be applied, though we do not pursue the matter further.

Indeed, for fixed $A > 0$ and positive integer λ, by assuming the GUE Conjecture one can show

$$\frac{1}{T}\int_T^{2T} \left|\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A}{\log T} + it\Big)\right|^{2\lambda} dt \sim C(A,2\lambda)\log^{2\lambda} T \tag{5.6}$$

where

$$C(A,2\lambda) := \lim_{N\to\infty} \frac{1}{N^{2\lambda}}\int_{U(N)} \left|\frac{Z'}{Z}\Big(\frac{A}{N}\Big)\right|^{2\lambda} du,$$

$U(N)$ is the group of $N \times N$ unitary matrices u with Haar probability measure

du, and

$$Z(\beta) := \det(1 - e^{-\beta}u).$$

Note that if $\omega_1, ..., \omega_n$ are the eigenvalues of the unitary matrix u,

$$\frac{Z'}{Z}(\beta) = \sum_i \frac{1}{1 - e^{-\beta}\omega_i} = \sum_{r=1}^{\infty} e^{-\beta r}\mathrm{Tr}(u^r). \tag{5.7}$$

That the limit $X(A, 2\lambda)$ exists can be seen from the proof of Theorem 5.2.2 to follow. By computation with correlation functions, not reproduced here, one can see that for fixed λ, $C(A, 2\lambda)$ is of order $A^{-2\lambda+1}$ which for small A is slightly better than what can be obtained without refining our methods. (Though note for $\lambda = 1$ this order of bound is achieved in [34].)

It is by only slightly extending (5.6) that one can obtain a statement equivalent to the GUE Conjecture.

Theorem 5.2.2. *(On RH.) The GUE Conjecture is equivalent to the statement that for all fixed* $j, k \geq 1$ *and all collections of fixed constants* $A_1, ..., A_j, B_1, ..., B_k$ *each with positive real part, the limit*

$$\lim_{T\to\infty} \frac{1}{\log^{j+k} T}\left(\frac{1}{T}\int_T^{2T} \prod_{\ell=1}^{j} \frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A_\ell}{\log T} + it\Big) \prod_{\ell'=1}^{k} \overline{\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{B_{\ell'}}{\log T} + it\Big)}\, dt\right) \tag{5.8}$$

exists and is equal to

$$\lim_{N\to\infty} \frac{1}{N^{j+k}}\left(\int_{\mathcal{U}(N)} \prod_{\ell=1}^{j} \frac{Z'}{Z}\Big(\frac{A_\ell}{N}\Big) \prod_{\ell'=1}^{k} \overline{\frac{Z'}{Z}\Big(\frac{B_{\ell'}}{N}\Big)}\, du\right) \tag{5.9}$$

Moreover, for each $n \geq 1$*, the claim that identity* (5.1) *holds for all* $k \leq n$ *(that is, the zeros* k*-level correlation functions tend to that of the sine-kernel determinantal point process), is equivalent to the claim that these limits are equal for all* $j+k \leq n$*.*

It has long been understood in a heuristic sense that the characteristic polynomial Z is statistically an analogue of the zeta-function. (See [49] for the first spectacular application of this philosophy). Theorem 5.2.2 shows that at a microscopic scale described by the GUE Conjecture this correspondence should be understood quite literally.

A theorem very much in the same spirt restates the GUE Conjecture in purely arithmetical terms.

To state the theorem more succinctly we require the notation

$$dz(x) := e^{-x/2}\, d\big(\psi(e^x) - e^x\big),$$

a measure which (because of its discrete part and growth as $|x| \to \infty$) we will only integrate against functions $\phi(x)$ that belong to a restricted class we call *admissible*:

Definition 5.2.3. *A function $\phi : \mathbb{R} \to \mathbb{R}$ is admissible if it is in $C^2(\mathbb{R})$, equal to 0 for sufficiently large x as $x \to \infty$, and bounded by $e^{\alpha|x|}$ for $\alpha < 1/2$ as $x \to -\infty$.*

If ϕ is admissible,

$$\int_{\mathbb{R}} \phi(x)\, dz(x) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\sqrt{n}} \phi(\log n) - \int_0^{\infty} \frac{1}{\sqrt{t}} \phi(\log t)\, dt,$$

which is a count of primes minus a regular approximation to that count.

Remark: By making use of improper integrals, in section 5.3 we will slightly extend the range of functions against which dz may be integrated, but any instance in which this extended definition is used will be made clear.

To reduce the length of formulas, we set

$$\upsilon_T(x, y) := \big(1 - T|x - y|\big)_+,$$

which plays the role of telling us when x and y are separated by a distance of $O(1/T)$.

Finally for bounded functions $f \in C^2(\mathbb{R}^j)$ and $g \in C^2(\mathbb{R}^k)$ such that $f \cdot \mathbf{1}_{\mathbb{R}^j_+}$ and $g \cdot \mathbf{1}_{\mathbb{R}^k_+}$ are compactly supported we define the arithmetical quantity

$$\Psi_T(f;g) = \Psi_T^{j,k}(f;g) \tag{5.10}$$
$$:= \frac{1}{\log^{j+k} T} \int_{\mathbb{R}^j} \int_{\mathbb{R}^k} f\big(\tfrac{x}{\log T}\big) g\big(\tfrac{y}{\log T}\big) \upsilon_T(x_1 + \cdots + x_k\,,\, y_1 + \cdots + y_k)\, dz^k(y)\, dz^j(x)$$

In the definition (5.10) for Ψ_T we will see later that the values f and g outside the quadrants $\mathbb{R}^j_+$ and $\mathbb{R}^k_+$ play no role asymptotically. Nonetheless, in (5.10) there is a certain algebraic significance to retaining integrals over all $\mathbb{R}^j \times \mathbb{R}^k$ rather than restricting to only this quadrant.

We likewise define the random matrix quantity

$$\Theta_N(f;g) = \Theta_N^{j,k}(f;g) \tag{5.11}$$
$$:= \frac{1}{N^{j+k}} \sum_{r \in \mathbb{N}^j_+} \sum_{s \in \mathbb{N}^k_+} f\Big(\frac{r}{N}\Big) g\Big(\frac{s}{N}\Big) \int_{U(N)} \prod_{\ell=1}^{j} (-\mathrm{Tr}\, u^{r_\ell}) \prod_{\ell'=1}^{k} \overline{(-\mathrm{Tr}\, u^{s_{\ell'}})}\, du,$$

As before, it is not immediately obvious that $\Theta_N(f;g)$ has a limiting value as $N \to \infty$ but we demonstrate this later.

Theorem 5.2.4. *(On RH.) The GUE Conjecture is equivalent to the statement that for all fixed $j, k \geq 1$, and all collections of fixed collections of admissible functions $f_1, ..., f_j, g_1, ..., g_k$, we have for $f = f_1 \otimes \cdots \otimes f_j$ and $g = g_1 \otimes \cdots \otimes g_k$*

$$\lim_{T \to \infty} \Psi_T(f\,;\,g) = \lim_{N \to \infty} \Theta_N(f\,;\,g). \tag{5.12}$$

Moreover, for each $n \geq 1$, the claim that identity (5.1) *holds for all $k \leq n$ (that is, the zeros k-level correlation functions tend to that of the sine-kernel determinantal point process), is equivalent to the claim that* (5.12) *holds for all $j + k \leq n$.*

Remark: $f_1 \otimes \cdots \otimes f_j$ is just the function $(x_1, ..., x_j) \mapsto f_1(x_1) \cdots f_j(x_j)$.

Though it is only a technical point, in our proof it is important that the functions in (5.12) are separable. To have a simple proof which extends to non-separable functions would be desirable. Morally, the reason that seperable functions by themselves are sufficient to recover the GUE Conjecture is that (5.12) is a linear relation, and linear combinations of such functions are sufficiently dense to approximate an arbitrary function. The same holds true for test functions $\exp(-A_1x_1 - \cdots - A_jx_j - B_1y_1 - \cdots - B_ky_k)$ in Theorem 5.2.2.

It is worthwhile to see that Theorem 5.2.4 generalizes Theorem 5.1.2, in particular that it implies identity (5.2). We do so heuristically for the moment, with a more rigorous development to follow later.

We know that the $n = 1$, 1-level density, case of the GUE Conjecture is true. It therefore follows from Theorem 5.2.4 that the pair correlation conjecture is equivalent to the claim that for all $f, g \in C_c^2(\mathbb{R})$,

$$\lim_{T\to\infty} \frac{1}{\log^2 T} \int_{\mathbb{R}} \int_{\mathbb{R}} f\Big(\tfrac{x}{\log T}\Big) g(\tfrac{y}{\log T}\Big) \upsilon_T(x,y)\, dz(x)\, dz(y) \tag{5.13}$$

is equal to

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} f\Big(\frac{r}{N}\Big) g\Big(\frac{s}{N}\Big) \int_{U(N)} \operatorname{Tr} u^r\, \overline{\operatorname{Tr} u^s}\, du. \tag{5.14}$$

We specialise to the case in which $f = g$ with both functions an arbitrarily close approximation to the characteristic function $\mathbf{1}_{[0,\beta)}$. In this way, choosing better and better approximations, one can see that the pair correlation conjecture implies that for all $\beta > 0$,

$$\lim_{T\to\infty} \frac{1}{\log^2 T} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{[0,\beta)}\Big(\tfrac{x}{\log T}\Big) \mathbf{1}_{[0,\beta)}(\tfrac{y}{\log T}\Big) \upsilon_T(x,y)\, dz(x)\, dz(y) \tag{5.15}$$

is equal to

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \mathbf{1}_{[0,\beta)}\Big(\frac{r}{N}\Big) \mathbf{1}_{[0,\beta)}\Big(\frac{s}{N}\Big) \int_{U(N)} \operatorname{Tr} u^r\, \overline{\operatorname{Tr} u^s}\, du. \tag{5.16}$$

In fact, with a little more work – using the fact that $\upsilon_T(x,y)$ constrains $x \approx y$ in (5.13) and (5.15) – one can see that (5.15) for all $\beta > 0$ is sufficient to recover (5.13) for general f and g; but we leave details of this argument to the reader.

To see that (5.15) provides the same information as (5.2) note that

$$\upsilon_T(x,y) = T \int \mathbf{1}_{[x-1/T,x]}(t)\mathbf{1}_{[y-1/T,y]}(t)\, dt$$

so that

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{[0,\beta)}\Big(\tfrac{x}{\log T}\Big)\mathbf{1}_{[0,\beta)}\Big(\tfrac{y}{\log T}\Big)\upsilon_T(x,y)\, dz(x)\, dz(y) \\
&= T\int_{\mathbb{R}}\int_0^{\beta\log T}\int_0^{\beta\log T} \mathbf{1}_{[x-1/T]}(t)\mathbf{1}_{[y-1/T,y]}(t)\, dz(x)dz(y)\, dt \\
&\sim T\int_0^{\beta\log T}\Big(\int_t^{t+1/T} dz(x)\Big)\Big(\int_t^{t+1/T} dz(y)\Big)\, dt \\
&\sim T\int_0^{\beta\log T} e^{-t}\Big(\int_t^{t+1/T} d\big(\psi(e^x)-e^x\big)\Big)^2\, dt \\
&\sim T\int_1^{T^\beta}\big(\psi(\tau e^{1/T})-\psi(\tau)-(e^{1/T}-1)\tau\big)^2\frac{d\tau}{\tau^2} \\
&\sim T\int_1^{T^\beta}\Big(\psi\Big(\tau+\tfrac{\tau}{T}\Big)-\psi(\tau)-\tfrac{\tau}{T}\Big)^2\frac{d\tau}{\tau^2}.
\end{aligned}$$

Our purpose at the moment is only to reassure the reader that the quantities we are working with are meaningful, so we have not made the effort to rigorously justify our passage from expression to expression. Rigorous justification is provided in a more general context in section 5.12. (None of the steps involve anything more involved than a straightforward bounding of error terms.)

On the other hand, to evaluate (5.16) we make use of the well known identity (see for instance [20]) that for $r \geq 1$,

$$\int_{U(N)} \operatorname{Tr} u^r\, \overline{\operatorname{Tr} u^s}\, du = \delta_{rs}\, r \wedge N. \tag{5.17}$$

(Here, recall the notation $r \wedge N$ to denote the minimum of r and N.) Hence (5.16) is given by

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{r \leq N^\beta} r \wedge N = \beta - 1/2$$

for $\beta \geq 1$. For $\beta < 1$ this limit is $\beta^2/2$.

The equality of (5.15) and (5.16) then, for $\beta \geq 1$ (the range of β for which we cannot simply evaluate (5.15) unconditionally from the prime number theorem), is exactly equation (5.2).

It is possible in this way to draw out arithmetical equivalences for the $k = 3, 4$ three and four point correlation conjectures for zeta zeros. The resulting arithmetical statements do not, however, have the simplicity of Theorem 5.1.2. We record them in Theorems 5.10.2 and 5.10.3.

At the same time, it is possible using Theorems 5.2.4 and 5.2.2 to generalize Theorem 5.1.2 in a way that is algebraically simple – though the compromise we suffer is that the result we shall now state is not equivalent to the GUE Conjecture, but is only a consequence of it.

We will require the higher-order von Mangoldt functions Λ_j, defined in the usual manner by

$$\Lambda_j(n) := \mu \star (\log^j)(n) = \sum_{d|n} \mu(d) \log^k(n/d) \tag{5.18}$$

or equivalently inductively by

$$\Lambda_j(n) = \Lambda \star \Lambda_{j-1}(n) + \log(n)\Lambda_{j-1}(n), \tag{5.19}$$

where $\Lambda_1 = \Lambda$, the usual von Mangoldt function, and we have used $\star$ to denote multiplicative convolution on the integers. This inductive definition makes clear that Λ_j is supported on integers with no more than j distinct prime factors. We likewise define

$$\psi_j(x) := \sum_{n \le x} \Lambda_j(n). \tag{5.20}$$

The properties of ψ_j are discussed in greater length in Appendix A. Unconditionally, from residue calculus and well-known zero free regions for the zeta function, we know that

$$\begin{aligned} \psi_j(x) &= \operatorname*{Res}_{s=1} \frac{\zeta^{(j)}(s)}{\zeta(s)} \frac{x^s}{s} + o(x) \\ &= x P_{j-1}(\log x) + o(x), \end{aligned} \tag{5.21}$$

where $P_{j-1}(x)$, defined by this expression, is a $j-1$ degree polynomial with

$$P_{j-1}(\log x) = j \log^{j-1} x + o(\log^{j-1} x).$$

The error term between ψ_j and its regular approximation,

$$\tilde{\psi}_j(x) := \psi_j(x) - x P_{j-1}(\log x), \tag{5.22}$$

on Riemann hypothesis has the better bound, $O_j(x^{1/2+\epsilon})$, and finally we define

$$\tilde{\psi}_j(x; H) = \tilde{\psi}_j(x + H) - \tilde{\psi}_j(x),$$

which is a count of almost primes in an interval of length H, minus its regular approximation. Its regular approximation should be thought of as its expected value.

We can arrive at counts of almost primes with the above von Mangoldt weights by repeatedly convolving the measures dz with one another, and in this way we will obtain

Theorem 5.2.5. *(On RH.) On the assumption of the GUE Conjecture, for fixed*

$\beta > 0$ and integers $j, k \geq 1$, let $X = T^\beta$ and $\delta = 1/T$. Then

$$\int_1^X \tilde{\psi}_j(x;\, \delta x)\tilde{\psi}_k(x,\, \delta x)\frac{dx}{x^2} \sim \tfrac{jk}{j+k-1}\frac{\log^{j+k} T}{T}\int_0^\beta y^{j+k-1} \wedge 1\, dy. \tag{5.23}$$

It is perhaps more instructive to write the right hand side of (5.23) as

$$\tfrac{jk}{j+k-1}\int_1^{T^\beta} \Big(\frac{x}{T}\Big)\Big(\log(x) - \log\big(\tfrac{x}{T} \vee 1\big)\Big)^{j+k-1}\frac{dx}{x^2}.$$

Recalling that $\delta x = x/T$ and $X = T^\beta$ above, it is reasonable therefore to make a conjecture in which the weight dx/x^2 has been replaced by the more natural weight dx.

Conjecture 5.2.6. *Fix any $\epsilon > 0$ and integers $j, k \geq 1$. Then as $X \to \infty$, uniformly for $X^\epsilon \leq H \leq X^{1-\epsilon}$,*

$$\frac{1}{X}\int_1^X \tilde{\psi}_j(x; H)\tilde{\psi}_k(x; H)\, dx \sim \tfrac{jk}{j+k-1}\, H\,(\log X - \log H)^{j+k-1}. \tag{5.24}$$

By a simple summability argument, this implies (5.23).

Note that for $j = k = 1$ (5.24) agrees with estimate (5.3).

An elementary combinatorial computation applied to the prime number theorem will reveal that

$$\frac{1}{X}\sum_{n \leq X} \Lambda_j(n)\Lambda_k(n) \sim \frac{jk}{j+k-1}\log^{j+k-1} X,$$

and from this one may see that (5.23) is true unconditionally for $\beta \leq 1$, or alternatively that

$$\frac{1}{X}\int_1^X \tilde{\psi}_j(x; H)\tilde{\psi}_k(x; H)\, dx \sim \tfrac{jk}{j+k-1}\, H\, \log^{j+k-1} X.$$

for $H \leq 1$.

It would be extremely interesting to know if there is an arithmetical reason, even one that is heuristic, that these formulas so strongly determine the form of equations (5.23) and (5.24). Analogous results for the counts $\Lambda \star \cdots \star \Lambda$, for instance, in place of Λ_j and Λ_k can be derived from the GUE Conjecture, but do not have nearly so simple a form as j and k grow.

On the other hand, the reason that equations (5.23) and (5.24) have such simple form is not obscure from a perspective prominent in random matrix theory. They are related to the ratio conjecture of Farmer [24], that with $s = 1/2 + it$,

$$\int_0^T \frac{\zeta(s+\alpha)\zeta(1-s+\beta)}{\zeta(s+\gamma)\zeta(1-s+\delta)} \sim T\frac{(\alpha+\delta)(\beta+\gamma)}{\alpha+\beta)(\gamma+\delta)} - T^{1-\alpha-\beta}\frac{(\delta-\beta)(\gamma-\alpha)}{(\alpha+\beta)(\gamma+\delta)} \quad (5.25)$$

for $\alpha, \beta, \gamma, \delta \lesssim 1/\log T$, with $\Re\gamma, \Re\delta > 0$.

The analogous result is known to hold for averages over the unitary group of ratios of characteristic polynomials. Different proofs are given in [10], [14], [15], each of which extends the result relevant here, of ratios of the product of 2 characteristic polynomials over the product of 2 characteristic polynomials, to an evaluation of ratios of products of any number of characteristic polynomials.

A result which is equivalent to the 2×2 ratio theorem and which we make use of in proving Theorem 5.2.5 is the following:

We define for a unitary matrix u the statistics $H_j(r)$ inductively as follows: for $r \geq 1$,

$$H_1(r) := -\mathrm{Tr}(u^r) \quad (5.26)$$

$$H_j(r) := \sum_{s=1}^{r-1} H_{j-1}(r-s)H_1(s) + rH_{j-1}(r). \quad (5.27)$$

The similarity to the inductive definition (5.19) of the higher von Mangoldt functions should be clear.

The result we prove is

Lemma 5.2.7.

$$\int_{U(N)} H_j(r)\overline{H_k(s)}\, du = \delta_{rs} \sum_{\nu=1}^{r\wedge N} \big(\nu^j - (\nu-1)^j\big)\big(\nu^k - (\nu-1)^k\big).$$

Note that the well-known identity (5.17) is the case $j = k = 1$. For $r > N$, this result extends beyond what can be derived making use of only the statistics of Diaconis and Shashahani (Theorem 5.7.1).

We give an simple proof of Lemma 5.2.7 in Appendix C, and use this to give an elementary and independent proof of the 2×2 ratio theorem for the unitary group.

5.3 Explicit formulae

In this section we write down the explicit formulae relating the zeros to prime and the primes to the zeta function in the critical strip. The first of these we have already discussed, and we only recall it again, using the new notation of the measure dz introduced above.

We recall the notation

$$S(t) := \frac{1}{\pi} \arg \zeta(\tfrac{1}{2} + it)$$

with argument defined by a continuous rectangular path from 2 to $2+it$ to $1/2+it$, starting with $\arg\zeta(2) = 0$. For us, the importance of the function $S(t)$ is that on the Riemann hypothesis,

$$dS(t) = \Big(\sum_{\gamma} \delta_\gamma(t) - \frac{\Omega(t)}{2\pi} \Big)\, dt,$$

where

$$\Omega(t) := \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\big(\tfrac{1}{4} + i\tfrac{t}{2}\big) + \tfrac{1}{2}\tfrac{\Gamma'}{\Gamma}\big(\tfrac{1}{4} - i\tfrac{t}{2}\big) - \log \pi.$$

By Stirling's formula

$$\frac{\Omega(t)}{2\pi} = \frac{\log\big((|t|+2)/2\pi\big)}{2\pi} + O\Big(\frac{1}{|t|+2}\Big),$$

and $\Omega(t)/2\pi$ is a regular approximation to the atomic mass at the γ's. $S(t)$ may therefore be thought of as an error term of a regular approximation to the zero counting function.

Theorem 5.3.1 (The explicit formula)**.** *(On RH.) For g a function in $C_c^2(\mathbb{R})$,*

$$\int_{-\infty}^{\infty} \hat{g}\Big(\frac{\xi}{2\pi}\Big)\, dS(\xi) = -\int_{-\infty}^{\infty} \big(g(x) + g(-x)\big)\, dz(x).$$

For g delimited as above, it follows from standard Fourier analysis that $\hat{g}$ decays quadratically or faster, so that the left hand integral converges absolutely (since the contribution of both the atomic mass of zeta zeros and the mass $\Omega(t)/2\pi\, dt$ on an interval $[\xi, \xi+1]$ is at most $O(\log(|\xi|+2))$.

A related identity we will make use of relates the measure dz to the values of the zeta function in the critical strip. As with Theorem 5.3.1 it is true only on the Riemann hypothesis.

Theorem 5.3.2. *(On RH.) For $\Re s \in (0, 1/2)$,*

$$-\frac{\zeta'}{\zeta}(1/2+s) = \int_{\to-\infty}^{\to\infty} e^{-sx}\, dz(x). \tag{5.28}$$

We have used the notation $\int_{\to-\infty}^{\to\infty}$ to denote an improper integral. Earlier to avoid any possible confusion we restricted the range of functions against which the measure dz can be integrated, and for this reason our improper integral must be defined in the following way:

We define the cutoff-function α_R by

$$\alpha(x) := \exp\Big(1 - \frac{1}{1-x^4}\Big)\mathbf{1}_{[-1,1]}(x), \tag{5.29}$$

$$\alpha_R(x) := \alpha(x/R). \tag{5.30}$$

For us the important features of α_R are that it is supported in $[-R, R]$, has continuous second derivative, and $\alpha(0) = 1$.

We thus define for a function $f \in C^2(\mathbb{R})$

$$\int_{\to-\infty}^{\to\infty} f(x)\,dz(x) = \lim_{R\to\infty} \int_{-\infty}^{\infty} \alpha_R(x) f(x)\,dz(x)$$

when the limit exists.

Note that we require the Riemann hypothesis to ensure that the integral in (5.28) converges; having assumed RH, that it does so follows from partial integration.

Proof of Theorem 5.3.2. Note that for $\Re s > 1$, (by dominated convergence for instance),

$$F(s) := \lim_{R\to\infty} \int_0^\infty \alpha_R(x) e^{-sx} d\big(\psi(e^x) - e^x\big) = -\frac{\zeta'}{\zeta}(s) - \frac{1}{s-1}. \tag{5.31}$$

But for any $\epsilon > 0$, it is easy to see by partial integration that the limit defining $F(s)$ converges uniformly for $\Re s \geq 1/2 + \epsilon$. Hence by analytic continuation (5.31) remains valid for $\Re s > 1/2$. Yet for $\Re s < 1$,

$$\int_{-\infty}^0 e^{(1-s)x}\,dx = -\frac{1}{s-1},$$

and so for $\Re s \in (1/2, 1)$,

$$\begin{aligned}
\lim_{R\to\infty} \int_{-\infty}^{\infty} \alpha_R(x) e^{-sx} d\big(\psi(e^x) - e^x\big) &= \lim_{R\to\infty} \int_{0}^{\infty} \alpha_R(x) e^{-sx} d\big(\psi(e^x) - e^x\big) + \int_{-\infty}^{0} e^{(1-s)x}\, dx \\
&= -\frac{\zeta'}{\zeta}(s)
\end{aligned}$$

by substituting (5.31). This is (5.28) with $s + 1/2$ replaced by s. $\square$

5.4 Notation: point processes and linear statistics

As in chapter 4, we recast the GUE Conjecture in the language of point processes. A short introduction to point processes is given in Appendix B; a more general introduction may be found in [70] or [40]. Those uncomfortable with the notion of a point process may be reassured that for us the processes defined below will just be an abbreviation allowing us to write formulas more succinctly and bring to mind the positivity of certain quantities. Even an intuitive understanding would suffice to translate these formulas into a more familiar form.

Definition 5.4.1. *Let T be a large real number, t a random variable uniformly distributed on $[T, 2T]$. We define Z_T to be the point process with point configurations*

$$\{\tfrac{\log T}{2\pi}(\gamma - t)\}$$

parameterized by t, where γ runs over all the ordinates of non-trivial zeros of the zeta function.

We will label the point configurations of Z_T by $\{\xi_i\}$. So for instance, this formalism has the consequence,

$$\mathop{\mathbf{P}}_{Z_T}\big(\#(i : \xi_i \in K) = n\big) = \frac{1}{T}\text{Meas}\Big\{t \in [T, 2T] : \#\{\gamma : \tfrac{\log T}{2\pi}(\gamma - t) \in K\} = n\Big\}$$

for a measurable set K and integer n, and for $\eta \in C_c(\mathbb{R}^2)$,

$$\mathop{\mathbf{E}}_{Z_T} \sum_{\substack{i,j \\ \text{distinct}}} \eta(\xi_i, \xi_j) = \frac{1}{T} \int_T^{2T} \sum_{\substack{\gamma_1, \gamma_2 \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), \tfrac{\log T}{2\pi}(\gamma_2 - t) \Big)\, dt.$$

Definition 5.4.2. *$\mathcal{S}$ is the determinantal point process with sine-kernel.*

As discussed in Appendix B, the process $\mathcal{S}$ is characterized by its correlation functions, which have the value,

$$\mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1, \dots, j_k \\ \text{distinct}}} \eta(x_{j_1}, \dots, x_{j_k}) = \int_{\mathbb{R}^k} \eta(x) \det_{k \times k} \Big(K(x_i - x_j) \Big)\, d^k x,$$

Remark: In this formalism, the GUE Conjecture is therefore just the statement that the processes Z_T tend in distribution as $T \to \infty$ to the process $\mathcal{S}$, which makes more clear and canonical what it is that the GUE Conjecture is describing. (See [48], ch. 16 for an account of the convergence of point processes.)

If the reader is unhappy with the concept of point processes he or she will lose none of the logical structure of the argument simply by substituting $\frac{1}{T}\int_T^{2T} \cdots dt$ and a sum over $\frac{\log T}{2\pi}(\gamma - t)$ anytime he or she sees $\mathbf{E}_{Z_T}$ and a sum over ξ_i, and likewise substituting determinantal integrals for the expected value of correlation sums over $\mathcal{S}$.

We quickly demonstrate the notational advantage of this device however: with it we can write

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{3} \sum_i \eta_\ell(x_i)$$

instead of

$$\mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1,i_2,i_3\\ \text{distinct}}} \eta_1(\xi_{i_1})\eta_2(\xi_{i_2})\eta_3(\xi_{i_3}) + \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1,i_2\\ \text{distinct}}} \eta_1(\xi_{i_1})\eta_2(\xi_{i_1})\eta_3(\xi_{i_2})$$

$$+ \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1,i_2\\ \text{distinct}}} \eta_1(\xi_{i_1})\eta_2(\xi_{i_2})\eta_3(\xi_{i_1}) + \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{i_1,i_2\\ \text{distinct}}} \eta_1(\xi_{i_1})\eta_2(\xi_{i_2})\eta_3(\xi_{i_2})$$

$$+ \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{i_1} \eta_1(x_{i_1})\eta_2(\xi_{i_1})\eta_3(\xi_{i_1})$$

$$= \int_{\mathbb{R}^3} \eta_1(x_1)\eta_2(x_2)\eta_3(x_3) \det_{3\times 3}\Big(K(x_i - x_j)\Big)\, d^3x + \int_{\mathbb{R}^2} \eta_1(x_1)\eta_2(x_1)\eta_3(x_2) \det_{2\times 2}\Big(K(x_i - x_j)\Big)\, d^2x$$

$$+ \int_{\mathbb{R}^2} \eta_1(x_1)\eta_2(x_2)\eta_3(x_1) \det_{2\times 2}\Big(K(x_i - x_j)\Big)\, d^2x + \int_{\mathbb{R}^2} \eta_1(x_1)\eta_2(x_2)\eta_3(x_2) \det_{2\times 2}\Big(K(x_i - x_j)\Big)\, d^2x$$

$$+ \int_{\mathbb{R}} \eta_1(x_1)\eta_2(x_1)\eta_3(x_1)\, dx_1.$$

The reader should check these expressions are the same.

In what follows, we will be using Fourier analysis in connection with the explicit formula, and for this reason it will be useful to replace the averages

$$\frac{1}{T}\int_T^{2T} \cdots\, dt = \int_{\mathbb{R}} \frac{\mathbf{1}_{[1,2]}(t/T)}{T} \cdots\, dt$$

with

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \cdots\, dt,$$

for σ a more general function. We define

Definition 5.4.3. *The point process $Z_T(\sigma)$ for σ a measurable function on $\mathbb{R}$ of mass 1 is defined by the point configurations*

$$\{\tfrac{\log T}{2\pi}(\gamma - t)\},$$

parameterized by a real valued random variable t with density $\sigma(t/T)/T$.

Note that under this definition, $Z_T = Z_T(\mathbf{1}_{[1,2]})$.

Definition 5.4.4. *For σ a measurable function on $\mathbb{R}$ of mass 1, we give the label* GUE(σ) *to the proposition that the processes $Z_T(\sigma)$ tend in distribution as $T \to \infty$ to the process $\mathcal{S}$.*

That is, in the language of correlation functions, GUE(σ) is the statement that for any $\eta \in C_c(\mathbb{R}^k)$,

$$\mathop{\mathbf{E}}_{Z_T(\sigma)} \sum_{\substack{j_1,\ldots,j_k \\ \text{distinct}}} \eta(\xi_{j_1}, \ldots, \xi_{j_k}) \sim \int_{\mathbb{R}^k} \eta(x) \cdot \det_{k\times k} \Big(K(x_i - x_j) \Big) \, d^k x$$

as $T \to \infty$.

In fact, there is nothing especially canonical about our use of $C_c(\mathbb{R}^k)$ test functions. Any class of test functions which are sufficiently dense and decay rapidly enough will do. We arrive at a more formal statement of this fact in section 5.8, where its proof will follow more easily.

We remind the reader that the eigenvalues of the unitary group, stretched out so as to have mean unit density, also converge to the process $\mathcal{S}$. This can be seen from the integration formula of Weyl, which gives an exact evaluation for the k-point correlation functions of the eigenvalues:

Theorem 5.4.5 (The Weyl-Dyson-Gaudin integration formula)**.** *Let $\{e(\theta_1), \ldots, e(\theta_N)\}$ be the eigenvalues of a random $N \times N$ unitary matrix u, distributed according to Haar measure du, with θ_j chosen to be in $[-1/2, 1/2)$ for all j, and define*

$$K_N(x) := \frac{\sin \pi x}{N \sin(\pi x / N)}.$$

Then for any $k \leq N$ and measurable $\eta : [-N/2, N/2)^k \mapsto \mathbb{C}$,

$$\int_{U(N)} \sum_{\substack{j_1,...,j_k \\ distinct}} \eta(N\theta_{j_1}, ..., N\theta_{j_k})\, du = \int_{[-N/2,N/2)^k} \eta(x_1, ..., x_k) \det_{k\times k} \Big(K_N(x_i - x_j)\Big)\, d^k x.$$

Note that $K_N(x) \to K(x)$ uniformly.

We can form a point process even closer to $\mathcal{S}$ by pulling back points of the process $\mathcal{S}_N$ defined in chapter 4, so that they are repeated with period N:

Definition 5.4.6. *The point process $\mathcal{S}'_N$ is defined by the point configurations*

$$\bigcup_{\nu\in\mathbb{Z}} \{N(\theta_1 + \nu), ..., N(\theta_N + \nu)\}$$

where $\theta_1, ..., \theta_N \in [-1/2, 1/2)$ are, as in the Weyl integration formula, such that $\{e(\theta_1), ..., e(\theta_N)\}$ are the eigenvalues of a random unitary matrix distributed according to Haar measure.

If we label the point configurations of $\mathcal{S}'_N$ by $\{x_j\}$, the the Weyl integration formula gives that for $\eta : \mathbb{R}^k \mapsto \mathbb{R}$ is integrable,

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \sum_{\substack{j_1,...,j_k \\ \text{distinct}}} \eta(x_{j_1}, ..., x_{j_k}) = \int_{\mathbb{R}^k} \eta(x) \cdot \det_{k\times k} \Big(K_N(x_i - x_j)\Big)\, d^k x.$$

In particular,

Proposition 5.4.7. *$\mathcal{S}'_N \to \mathcal{S}$ in distribution.*

Note that by Poisson summation for functions η that are, for instance, in $C^2_c(\mathbb{R})$,

$$\sum_{\nu\in\mathbb{Z}} \eta(N\theta + N\nu) = \sum_{r\in\mathbb{Z}} \frac{1}{N} \hat{\eta}\Big(\tfrac{r}{N}\Big) e(r\theta)$$

for all θ, so that for $\eta_1, ..., \eta_k$ of this sort,

$$\mathop{\mathbf{E}}_{S'_N} \prod_{\ell=1}^{k} \sum_i \eta_\ell(x_i) = \frac{1}{N^k} \sum_{r \in \mathbb{Z}^k} \hat{\eta}_1\Big(\tfrac{r_1}{N}\Big) \cdots \hat{\eta}_k\Big(\tfrac{r_k}{N}\Big) \int_{U(N)} \prod_{\ell=1}^{k} \mathrm{Tr}(u^{r_\ell})\, du. \tag{5.32}$$

Note that $\hat{\eta_\ell}$ for each ℓ will in this case decay quadratically, and for fixed N, $\mathrm{Tr}(u^r)$ remains bounded, so there is no difficulty in swapping the order of summation and integration.

It is therefore by passing through the processes S'_N that we will arrive at sums like (5.11).

Because the mapping $u \mapsto u^{-1}$ preserves Haar measure,

Proposition 5.4.8. *For $r \in \mathbb{Z}^k$,*

$$\int_{U(N)} \prod_{\ell=1}^{k} \mathrm{Tr}(u^{r_\ell})\, du = \int_{U(N)} \prod_{\ell=1}^{k} \mathrm{Tr}(u^{-r_\ell})\, du.$$

In particular these statistics are real valued.

Finally we introduce notation for linear statistics as they depend on the variable t. The mixed moments of these quantities carry the same information as the correlation sums (5.1) of the GUE Conjecture.

We define (for functions η that decay quadratically)

$$G_T(\eta, t) := \sum_\gamma \eta\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big). \tag{5.33}$$

An approximation to this count is given by substituting an integral against $\log(|\xi|+2)/2\pi$ for the sum over zeros:

$$L_T(\eta, t) := \int_{-\infty}^{\infty} \eta\Big(\tfrac{\log T}{2\pi}(\xi - t)\Big)\, \frac{\log(|\xi|+2)}{2\pi}\, d\xi \tag{5.34}$$

Note that, for η that decay quadratically,

$$L_T(\eta, t) = \tfrac{\log(|t|+2)}{\log T} \int_{-\infty}^{\infty} \eta(\alpha)\, d\alpha + O_\eta\Big(\frac{1}{\log T}\Big).$$

Finally, we define

$$\widetilde{G}_T(\eta, t) := \int_{-\infty}^{\infty} \eta\Big(\tfrac{\log T}{2\pi}(\xi - t)\Big)\, dS(\xi). \tag{5.35}$$

From Stirling's formula, for η that decay quadratically,

$$\begin{aligned}\widetilde{G}_T(\eta, t) &= G_T(\eta, t) - L_T(\eta, t) + O_\eta\Big(\frac{1}{\log T}\Big) \\ &= G_T(\eta, t) - \tfrac{\log(|t|+2)}{\log T} \int_{-\infty}^{\infty} \eta(\alpha)\, d\alpha + O_\eta\Big(\frac{1}{\log T}\Big),\end{aligned} \tag{5.36}$$

so that $\widetilde{G}_T(\eta, t)$ should be thought of as the linear statistic $G(\eta, t)$ minus its expected value.

Since we know unconditionally that the number of γ in any interval $[k, k+1)$ is at most $\log(|k| + 2)$, we have

$$G_T(\eta, t) \lesssim \sum_{k \in \mathbb{Z}} \log(|k| + 2) \max_{x \in [k, k+1)} \Big|\eta\Big(\tfrac{\log T}{2\pi}(x - t)\Big)\Big|, \tag{5.37}$$

with the same upper bound obviously holding for $L_T(\eta, t)$, and therefore $\widetilde{G}_T(\eta, t)$. A particular consequence of (5.37) that we will make use of later is that if

$$\eta(\xi) \lesssim 1/(1 + \xi^2),$$

then

$$G_T(\eta, t) \lesssim \log(|t| + 2),$$

and likewise for $L_T(\eta, t)$ and $\widetilde{G}_T(\eta, t)$.

The arithmetic significance of $\widetilde{G}_T(\eta, t)$ comes from the explicit formula:

Proposition 5.4.9. *For $g \in C^2(\mathbb{R})$,*

$$\widetilde{G}_T(\hat{g}, t) = \frac{-1}{\log T} \int_{-\infty}^{\infty} \Big(g\Big(\tfrac{x}{\log T}\Big) e^{ixt} + g\Big(\tfrac{-x}{\log T}\Big) e^{-ixt} \Big)\, dz(x). \tag{5.38}$$

It is worthwhile to see one example of the relation between the two notations introduced in this section. We have, for instance,

$$\mathop{\mathbf{E}}_{Z_T(\sigma)} \Big(\sum \eta(\xi_j) \Big)^k = \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} G_T(\eta, t)^k \, dt.$$

5.5 A plan of the proof

We are now in a position to outline our proofs. A tool we will find absolutely essential is an upper bound on the moments of point counts in the process Z_T first proved by Fujii [26] and which may be approximately stated in the following way, that for fixed k, as long as t has been averaged over a long enough interval (with length of order T), the k^{th} moment of the count

$$\#\{\tfrac{\log T}{2\pi}(\gamma - t) \in [A, A+k]\}$$

remains bounded, uniformly of the choice of A. In the language of point processes, this is to say the moments of counts of points inside course enough intervals can be bounded from above to the correct order.

This cannot be literally true as it has been stated, because for large enough A the density of γ around $t + 2\pi A/\log T$ will be larger than $\log T/2\pi$. A precise statement is that uniformly in a and for any $\epsilon > 0$,

$$\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T} \big| G_T(\eta, t) \big|^k \, dt \lesssim_k \int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T} \big| L_T(M_k \eta, t) \big|^k \, dt$$

for all $T \geq T_0$, where T_0 is a function only of ϵ. Here we have used the notation M_k, an upper bound on η through characteristic functions of size k:

$$M_k\eta(\xi) := \sum_{\nu=-\infty}^{\infty} \mathbf{1}_{I_k(\nu)}(\xi) \cdot \sup_{I_k(\nu)} |\eta|, \tag{5.39}$$

where for typographical reasons we denote the interval $[k\nu - k/2, k\nu + k/2)$ by $I_k(\nu)$. Recall that $L_T(\cdot, t)$ amounts to replacing the sum over γ in $G_T(\cdot, t)$ with a logarithmic mass that approximates this sum.

We also prove another upper bound which is considerably more subtle. This is that for a function g supported in an interval $[-X, X] \subset [-1/k, 1/k]$ and bounded in modulus by a constant A,

$$\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T} \left|\widetilde{G}_T(\hat{g}, t)\right|^k dt \lesssim_k A^k X^k, \tag{5.40}$$

for $T \geq T_0$, with T_0 a function of only ϵ and X.

This result should be surprising at first glance for the following reason: if $g(x) \in C^2(\mathbb{R})$ closely approximates the indicator function

$$\mathbf{1}_{[-\delta,\delta]}(x)$$

then $\hat{g}(\xi)$ will closely approximate the function

$$\frac{1}{\delta}\frac{\sin(\pi\xi/\delta)}{\pi\xi/\delta}.$$

In particular as g approaches $\mathbf{1}_{[-\delta,\delta]}$ (say uniformly), the $L^1(\mathbb{R})$ norm of $\hat{g}$ will grow arbitrarily large. Yet the naive approach to bounding the left hand side of

(5.40), namely

$$\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{T} \Big| \sum_\gamma \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big) - \int_{-\infty}^{\infty} \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big) \frac{\Omega(\xi)}{2\pi}\, d\xi \Big|^k dt$$
$$\lesssim \int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{T} \Big(\sum_\gamma \Big|\hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)\Big| + \int_{-\infty}^{\infty} \Big|\hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)\Big| \frac{\Omega(\xi)}{2\pi}\, d\xi \Big)^k dt$$

will be arbitrarily large as

$$\sum_\gamma \Big|\hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)\Big| \text{ and } \int_{-\infty}^{\infty} \Big|\hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)\Big| \frac{\Omega(\xi)}{2\pi}\, d\xi$$

will both be large for every t. Even ignoring this issue we do not see from this naive approach that the left hand side of (5.40) should become smaller as δ becomes smaller.[1] This can only be seen by exploiting the cancellation that arises by subtracting from

$$\sum_\gamma \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)$$

it's regular approximation

$$\int_{-\infty}^{\infty} \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big) \frac{\Omega(\xi)}{2\pi}\, d\xi.$$

The situation is analogous to estimating

$$\sum_{k\in\mathbb{Z}} f(k) - \int_{\mathbb{R}} f(x)\, dx = \sum_{k\in\mathbb{Z}} f(k) - \hat{f}(0).$$

A naive bound on this quantity is $2\|f\|_{L_1}$, but in fact for functions that do not

[1]Suppose for instance that we should exploit some additional cancellation in the oscillating $\hat{g}$ by looking at

$$\Big|\sum_\gamma \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big)\Big| \text{ and } \Big|\int_{-\infty}^{\infty} \hat{g}\Big(\tfrac{\log T}{2\pi}(\gamma - t)\Big) \frac{\Omega(\xi)}{2\pi}\, d\xi\Big|$$

instead. Even this refinement in insufficient to obtain a bound that decreases with δ.

oscillate much the sum over $\mathbb{Z}$ is close to the integral over $\mathbb{R}$: by Poisson summation if $\hat{f}$ is supported in $(-1,1)$, this quantity is exactly 0.

It is these two upper bounds that take the place in our proof of the Hypothesis AC and LC in [25]. They are proven in section 5.6 using the explicit formula.

Analogous upper bounds may be proven for the average distribution of eigenvalues of the unitary group under Haar measure. This is the content of section 5.7.

In section 5.8, we make use of the first of these upper bounds for the zeros of the zeta function to show that, for averages weighted by σ_1 and σ_2, the statements GUE(σ_1) and GUE(σ_2) are equivalent. This is a Tauberian theorem. We expand upon the ideas involved in its proof in section 5.8.

With this equivalence between weights we can give a first heuristic approximation to what lies behind our proof. By adding in lower correlations, GUE(σ) may be seen to be equivalent to the claim that

$$\mathop{\mathbf{E}}_{Z_T(\sigma)} \prod_{\ell=1}^{n} \Big(\sum_j \eta_\ell(\xi_j) - \int_{-\infty}^{\infty} \eta_\ell(\alpha)\, d\alpha \Big) \sim \mathop{\mathbf{E}}_{S} \prod_{\ell=1}^{n} \Big(\sum_j \eta_\ell(x_j) - \int_{-\infty}^{\infty} \eta_\ell(\alpha)\, d\alpha \Big) \tag{5.41}$$

for every $n \geq 1$ and every η_1, ..., and η_n belonging to a class of functions sufficiently dense in $C_c(\mathbb{R})$ and with suitable continuity and decay conditions. But the left hand side of (5.41) asymptotically is just

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{n} \widetilde{G}_T(\eta_\ell, t)\, dt \\
&= \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{n} \Big(\frac{1}{\log T} \int_{-\infty}^{\infty} \hat{\eta}_\ell\Big(\tfrac{x}{\log T}\Big) e^{ixt} + \hat{\eta}_\ell\Big(\tfrac{-x}{\log T}\Big) e^{-ixt}\, dz(x) \Big)\, dt \\
&= \frac{1}{\log^n T} \sum_{\varepsilon \in \{-1,1\}^n} \int_{\mathbb{R}^n} \hat{\eta}_1\Big(\tfrac{\varepsilon_1 x_1}{\log T}\Big) \cdots \hat{\eta}_n\Big(\tfrac{\varepsilon_n x_n}{\log T}\Big) \hat{\sigma}\big(\tfrac{T}{2\pi}(\varepsilon_1 x_1 + \cdots + \varepsilon_k x_k)\big)\, dz(x_1) \cdots dz(x_k),
\end{aligned} \tag{5.42}$$

with the last two equalities following simply from computation through the explicit formula and interchanging the order of integration. If we now select σ so that $\hat{\sigma}(y) = (1 - 2\pi|x|)_+$, the reader may check (5.42) is just a polarization of the quantities Ψ_T defined in (5.10).

On the other hand, the right hand side of (5.41) can be evaluated as the limiting case of random matrix statistics which end up being a polarized form of (5.11). Since by taking linear combinations of the identity (5.12) in Theorem 5.2.4 one can recover the polarized from above, it is comparatively easy in this way to see that (5.12) implies the GUE Conjecture.

To show that the GUE Conjecture implies (5.12) requires more work. To first approximation, the argument is nothing more than setting $\eta_1 \otimes \cdots \otimes \eta_n$ in (5.41) so that $\hat{\eta_1} \otimes \cdots \hat{\eta_n}$ is restricted to a given quadrant $\mathbb{R}_+^j \times \mathbb{R}_-^k$ of $\mathbb{R}^n$, and such that in these quadrants $\hat{\eta}_\ell$ has a sharp cutoff at the origin, say $\hat{\eta}_\ell(x) = \mathbf{1}_{R_{\varepsilon_\ell}} f_\ell(\varepsilon_\ell x)$ for functions f_ℓ admissible in the sense of Definition 5.2.3. In this way,

$$\begin{aligned}
&\frac{1}{\log^n T} \sum_{\varepsilon \in \{-1,1\}^n} \int_{\mathbb{R}^n} \hat{\eta}\Big(\tfrac{\epsilon_1 x_1}{\log T}\Big) \cdots \hat{\eta}_n\Big(\tfrac{\epsilon_n x_n}{\log T}\Big) \hat{\sigma}\big(\tfrac{T}{2\pi}(\varepsilon_1 x_1 + \cdots \varepsilon_n x_n)\big)\, d^n z(x) \\
&\sim \frac{1}{\log^n T} \int_{\mathbb{R}_+^j} \int_{\mathbb{R}_+^k} f_1\Big(\tfrac{x_1}{\log T}\Big) \cdots f_j\Big(\tfrac{x_j}{\log T}\Big) f_{j+1}\Big(\tfrac{x_{j+1}}{\log T}\Big) \cdots f_{j+k}\Big(\tfrac{x_{j+k}}{\log T}\Big) \\
&\qquad\qquad \times \hat{\sigma}\big(\tfrac{T}{2\pi}(x_1 + \cdots + x_j - x_{j+1} - \cdots - x_{j+k}\big)\, d^n z(x) \\
&\sim \frac{1}{\log^n T} \int_{\mathbb{R}^j} \int_{\mathbb{R}^k} f_1\Big(\tfrac{x_1}{\log T}\Big) \cdots f_j\Big(\tfrac{x_j}{\log T}\Big) f_{j+1}\Big(\tfrac{x_{j+1}}{\log T}\Big) \cdots f_{j+k}\Big(\tfrac{x_{j+k}}{\log T}\Big) \\
&\qquad\qquad \times \hat{\sigma}\big(\tfrac{T}{2\pi}(x_1 + \cdots + x_j - x_{j+1} - \cdots - x_{j+k}\big)\, d^n z(x)
\end{aligned}$$

The last line will follow by showing that the 'tails'

$$f\Big(\tfrac{x}{\log T}\Big)\, dz(x) = f\Big(\tfrac{x}{\log T}\Big) e^{x/2}\, dx, \qquad \text{for } x \leq 0$$

do not substantially contribute to these quantities asymptotically. We have thus recovered the terms Ψ_T.

This approach must be amended very substantially however, owing to the fact that for Fourier transforms $\hat{\eta}$ with a sharp cutoff at the origin, the original distribution η will in general not be integrable, and so the sums in (5.41) are not well defined.

We overcome the issue by choosing smooth functions $f_\ell|_{\epsilon_T}$ (depending upon T) that so closely approximate functions of sharp cutoff $f_\ell \cdot \mathbf{1}_{R_+}$ that we still have

$$\begin{aligned}
&\frac{1}{\log^n T}\int_{\mathbb{R}_+^j}\int_{\mathbb{R}_+^k} f_1\Big(\tfrac{x_1}{\log T}\Big)\cdots f_j\Big(\tfrac{x_j}{\log T}\Big) f_{j+1}\Big(\tfrac{x_{j+1}}{\log T}\Big)\cdots f_{j+k}\Big(\tfrac{x_{j+k}}{\log T}\Big)\\
&\qquad\qquad\qquad \times \hat{\sigma}\big(\tfrac{T}{2\pi}(x_1+\cdots+x_j-x_{j+1}-\cdots-x_{j+k}\big)\, d^n z(x)\\
&\sim \frac{1}{\log^n T}\int_{\mathbb{R}^j}\int_{\mathbb{R}^k} f_1|_{\epsilon_T}\Big(\tfrac{x_1}{\log T}\Big)\cdots f_j|_{\epsilon_T}\Big(\tfrac{x_j}{\log T}\Big) f_{j+1}|_{\epsilon_T}\Big(\tfrac{x_{j+1}}{\log T}\Big)\cdots f_{j+k}|_{\epsilon_T}\Big(\tfrac{x_{j+k}}{\log T}\Big)\\
&\qquad\qquad\qquad \times \hat{\sigma}\big(\tfrac{T}{2\pi}(x_1+\cdots+x_j-x_{j+1}-\cdots-x_{j+k}\big)\, d^n z(x).
\end{aligned}$$

It will indeed be the case that for this to be true, the closeness of our approximation of $f_\ell|_{\epsilon_T}$ to $f_\ell \cdot \mathbf{1}_{\mathbb{R}_+}$ must increase with T. All the same, for any $\delta > 0$, we show that there is a fixed approximation $f|_\epsilon$ so that

$$\begin{aligned}
\Bigg|&\frac{1}{\log^n T}\int_{\mathbb{R}^n} f_1|_\epsilon\big(\tfrac{x_1}{\log T}\Big)\cdots f_n|_\epsilon\Big(\tfrac{x_n}{\log T}\Big)\hat{\sigma}\big(\tfrac{T}{2\pi}(x_1+\cdots-x_n)\big)\, d^n z(x) \qquad (5.43)\\
&-\frac{1}{\log^n T}\int_{\mathbb{R}^n} f_1|_{\epsilon_T}\big(\tfrac{x_1}{\log T}\Big)\cdots f_n|_{\epsilon_T}\Big(\tfrac{x_n}{\log T}\Big)\hat{\sigma}\big(\tfrac{T}{2\pi}(x_1+\cdots-x_n)\big)\, d^n z(x)\Bigg| < \delta
\end{aligned}$$

Because the functions $f_\ell|_\epsilon$ closely approximate $f_\ell \cdot \mathbf{1}_{\mathbb{R}_\epsilon}$,

$$\frac{1}{\log^n T}\int_{\mathbb{R}^n} f_1|_\epsilon\big(\tfrac{x_1}{\log T}\Big)\cdots f_n|_\epsilon\Big(\tfrac{x_n}{\log T}\Big)\hat{\sigma}\big(\tfrac{T}{2\pi}(x_1+\cdots-x_n)\big)\, d^n z(x)$$

will be close to its polarization

$$\sum_{\varepsilon\in\{-1,1\}^n} \frac{1}{\log^n T}\int_{\mathbb{R}^n} f_1|_\epsilon\big(\tfrac{\varepsilon_1 x_1}{\log T}\big)\cdots f_n|_\epsilon\Big(\tfrac{\varepsilon_n x_n}{\log T}\Big)\hat{\sigma}\big(\tfrac{T}{2\pi}(\varepsilon_1 x_1+\cdots-\varepsilon_n x_n)\big)\,d^n z(x).$$

This last quantity, because the functions $f_\ell|_\epsilon$ are fixed and smooth, can be evaluated on the GUE Conjecture by identity (5.42). It is a straightforward matter finally to show that the resulting answer agrees with that of Theorem 5.2.4.

Although (5.43) is intuitive enough, we have not really fully justified it. It's proof in section 5.10 is technical and is accomplished only via the upper bound (5.40) and what is sometimes referred to as a tensorization trick. (This tensorization trick is the reason we work with seperable functions.) Note that it is natural to apply (5.40) here, as the functions $(f_\ell|_\epsilon - f_\ell|_{\epsilon_T})$ are supported in a small region around the origin.

It is through this same method, using (5.28) of Theorem (5.3.2) and the fact that linear combinations of function $\exp(-A_1x_1-\cdots-A_nx_n)$ are sufficiently dense in $C_c(\mathbb{R}^n_+)$, that we arrive at Theorem 5.2.2.

Theorem 5.2.1 is an application of the same method of decomposing test functions into parts with different Fourier support. Letting $f(x)=\exp(-Ax)$, Theorem 5.3.2 gives that

$$\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac12+\tfrac{A}{\log T}+it\Big) = \frac{1}{\log T}\int_{\to-\infty}^{\to\infty} f\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x).$$

We decompose this into

$$O\Big(\tfrac{1}{\log T}\Big)+\frac{1}{\log T}\int_{\mathbb{R}} f|_{\epsilon_T}^{1/k}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x)+\frac{1}{\log T}\int_{\mathbb{R}} f|_{1/k}^{R_T}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x),$$

where $f|_{\epsilon_T}^{1/k}$ is a function supported in the the interval $[0,1/k]$ and $f|_{1/k}^{R_T}$ is chosen so that

$$f|_{\epsilon_T}^{1/k} + f|_{1/k}^{R_T}$$

is a smooth compactly supported function (on an interval $[0, R_T]$ say) that closely approximates

$$f \cdot \mathbf{1}_{\mathbb{R}_+}.$$

Note that for fixed k, one should be able to (and indeed can) choose such functions $f|_{1/k}^{R_T}$ in a way that their second derivatives do not increase with T.

We have that

$$\begin{aligned}\frac{1}{T}\int_0^T \left|\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A}{\log T}+it\Big)\right|^k dt \lesssim_k & \frac{1}{T}\int_0^T \left|\frac{1}{\log T}\int_{\mathbb{R}} f|_{\epsilon_T}^{1/k}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x)\right|^k dt \\ &+\frac{1}{T}\int_0^T \left|\frac{1}{\log T}\int_{\mathbb{R}} f|_{1/k}^{R_T}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x)\right|^k dt \\ =&\frac{1}{T}\int_0^T \left|\widetilde{G}_T\big((f|_{\epsilon_T}^{1/k})\hat{},t\big)\right|^k dt \\ &+\frac{1}{T}\int_0^T \left|\widetilde{G}_T\big((f|_{1/k}^{R_T})\hat{},t\big)\right|^k dt\end{aligned}$$

The first of these terms can be bounded by (5.40). For the second, note that $f_{1/k}^{R_T}$ does not have increasing first or second derivative, even as T increases (because the cutoff from the origin to $1/k$ does not change with T). Therefore $(f_{1/k}^{R_T})\hat{}$ will decay quickly enough, for all T, that an appropriate bound can be gained from the Fujii upper bound.

The final part of this paper concerns evaluating the covariance of almost primes. We weight the almost primes in such a way as to produce an algebraically nice answer. The algebraic part involves random matrix statistics that closely parallel those studied by Diaconis and Shahshahani, but fall outside the domain of what can be deduced directly from their result. The evaluation of these statistics is included in Appendix C. On the other hand, we can quickly outline how it is that one arrives at counts of almost primes from Theorems 5.2.4 and 5.2.2 by convolving the measure dz with itself, so that for instance,

$$dz * dz(x) + x\, dz(x) = e^{-x/2} d\big(\psi_2(x) - xP_1(x)\big)$$

where ψ_2 and P_1 are defined by (5.20) and (5.21). In perhaps more familiar language, this is just that for $\Re s > 1$,

$$\frac{\zeta''}{\zeta}(s) = \sum_n \frac{\Lambda_2(n)}{n^s}.$$

To convolve the measure dz with itself, we must replace the test functions $f_1 \otimes \cdots \otimes f_n$ in Theorem 5.2.4 with test functions $f(x_1, ..., x_n)$ that are constant on level sets of $x_1+x_2+\cdots+x_n$, for instance. The fastest route to such a replacement is by appealing to Theorem 5.2.2, but because the test functions $\exp(-Ax)$ are not compactly supported, this route entaills a few technical challenges. These are discussed in more detail in section 5.12.

5.6 Upper bounds for counts of zeros

In this section we recall several lemmas first proved in chapter 3. Since it is now convenient to make use of slightly different notation, we restate them in full here.

Lemma 5.6.1. *Suppose we are given non-negative integrable σ of mass 1 such that $\hat{\sigma}$ has compact support, and suppose $g_1, ..., g_k$ are in $C_c^2(\mathbb{R})$ and satisfy* $\operatorname{supp} g_\ell \subset [-\delta_\ell, \delta_\ell]$ *with $\delta_1 + \cdots + \delta_k = \Delta \leq 2$. Then there exists a T_0 depending only on Δ and the region in which $\hat{\sigma}$ is supported so that for $T \geq T_0$,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^k \widetilde{G}_T(\,\hat{g}_\ell\,, t)\, dt = \Big(\frac{-1}{\log T}\Big)^k \sum_{n_1^{\epsilon_1} n_2^{\epsilon_2} \cdots n_k^{\epsilon_k} = 1} \prod_{\ell=1}^k \frac{\Lambda(n_\ell)}{\sqrt{n_\ell}} g_\ell\big(\tfrac{\epsilon_\ell \log n_\ell}{\log T}\big) \qquad (5.44)$$
$$+ O_k\Big(\frac{1}{T^{1-\Delta/2}} \prod_{\ell=1}^k \frac{\|g_\ell\|_\infty}{\log T}\Big),$$

where the sum is over all $n \in \mathbb{N}^k, \epsilon \in \{-1,1\}^k$ *such that* $n_1^{\epsilon_1} n_2^{\epsilon_2} \cdots n_k^{\epsilon_k} = 1$.

Proof. This is Lemma 3.3.4. □

As a consequence, we show that for coarse enough counts, linear statistics of zeta zeros can rigorously be bounded above to the correct order. This is the first upper bound outlined in section 5.5.

Lemma 5.6.2 (A Fujii-type upper bound)**.** *For* σ *non-negative and integrable such that* $\hat{\sigma}$ *is compactly supported, there exists a* T_0 *depending only on the region in which* $\hat{\sigma}$ *is supported, so that for all* $T \geq T_0$,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{j=1}^k G_t(\eta_j, t)\, dt = O_k\bigg(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{j=1}^k L_T(M_k \eta_j, t)\, dt \bigg),$$

where the implied constant depends only upon k.

The upper bound M_k is defined in (5.39).

Remark: Note that in the notation of point processes, the left hand side is

$$\mathop{\mathbf{E}}_{Z_t(\sigma)} \prod_{\ell=1}^k \sum_i \eta_\ell(\xi_i).$$

Proof. See that of Lemma 3.4.1. □

We can state the lemma in more intuitive terms.

Lemma 5.6.3 (A Fujii-type upper bound, restated)**.** *For* $\epsilon_0 > 0$*, there exists a* T_0 *depending only on* ϵ_0 *so that for all* $a \in \mathbb{R}$*, all* $\epsilon > \epsilon_0$ *and all* $T \geq T_0$,

$$\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T} \prod_{j=1}^k G_t(\eta_j, t)\, dt = O_k\bigg(\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T} \prod_{j=1}^k L_T(M_k \eta_j, t)\, dt \bigg),$$

where the implied constant depends only upon k.

Proof. Note that there is an absolute constant C so that

$$\frac{1}{\epsilon}\mathbf{1}_{[a,a+\epsilon]}(x) \leq CV_{a,\epsilon}(x)$$

for

$$V_{a,\epsilon}(x) := \frac{1}{\epsilon}V\Big(\frac{x-a}{\epsilon}\Big)$$

where

$$V(x) := \Big(\tfrac{\sin \pi x}{\pi x}\Big)^2.$$

Because

$$\hat{V}_{a,\epsilon}(x)(\xi) = e^{i2\pi a\xi}(1-\epsilon|x|)_+$$

is supported in $[-1/\epsilon_0, 1/\epsilon_0]$ for all $a \in \mathbb{R}$ and all $\epsilon > \epsilon_0$, we can apply Lemma 5.6.2 to bound the average in Lemma 5.6.3 from above. □

We now turn to the second upper bound outlined in section 5.5, for test functions with a narrowly supported Fourier transform. This is

Lemma 5.6.4. *For $\epsilon_0 > 0$, there exists a T_0 depending only on ϵ_0 so that for all $a \in \mathbb{R}$, all $\epsilon > \epsilon_0$, and all $g \in C_c^2(\mathbb{R})$ supported in $[-X, X]$ with $X \leq 1/k$, for all $T \geq T_0$*

$$\int_{\mathbb{R}} \frac{\mathbf{1}_{[a,a+\epsilon]}(t/T)}{\epsilon T}\big|\widetilde{G}_T(\,\hat{g}_\ell\,, t)\big|^k\,dt = O_k\Big(A^k\big(\tfrac{1}{\log^k T} + X^k\big)\Big),$$

where A is the maximum value of g.

To prove this bound we require another computational lemma that we will apply to Lemma 5.6.1.

Lemma 5.6.5. *For functions $g_1, ..., g_k$ each supported on the interval $[-X, X]$ and bounded in absolute value by a constant A, for $H \geq 1$ we have*

$$\frac{1}{H^k}\sum_{n_1^{\epsilon_1}n_2^{\epsilon_2}\cdots n_k^{\epsilon_k}=1}\prod_{\ell=1}^{k}\frac{\Lambda(n_\ell)}{\sqrt{n_\ell}}g\Big(\tfrac{\epsilon_\ell \log n_\ell}{H}\Big) = O_k(A^kX^k). \tag{5.45}$$

Remark: With control on the first and second derivatives of g_ℓ, a more exact evaluation can be made. See Lemma 12 of [61].

Proof of Lemma 5.6.5. We require from number theory only the Chebyshev estimate that

$$\sum_{p \le x} \log p = O(x).$$

As the von Mangoldt function Λ is supported on prime powers p^λ, the sum in (5.45) is just

$$\frac{1}{H^k} \sum_{p_1^{\lambda_1 \epsilon_1} \cdots p_k^{\lambda_k \epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}} g\Big(\tfrac{\epsilon_\ell \lambda_\ell \log p_\ell}{H}\Big) \le \frac{A^k}{H^k} \sum_{p_1^{\lambda_1 \epsilon_1} \cdots p_k^{\lambda_k \epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}} \mathbf{1}_{[0,x]}\Big(\tfrac{\lambda_\ell \log p_\ell}{H}\Big).$$

Here the sum ranges over all collections of k primes $\{p_1, ..., p_k\}$, k positive integers $\{\lambda_1, ..., \lambda_k\} \in \mathbb{N}_+^k$ and signs $\{\epsilon_1, ..., \epsilon_k\} \in \{-1, 1\}^k$ so that $p_1^{\lambda_1 \epsilon_1} \cdots p_k^{\lambda_k \epsilon_k} = 1$. Owing to the weights $p^{\lambda/2}$, our main contribution comes from terms in which $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 1$. By unique factorization, $p^{\epsilon_1} \cdots p^{\epsilon_k} = 1$ only when each p_i is equal to some pair, p_j. As there are c_k ways to form such pairs, where c_k is $(k-1)!!$ if k is even and 0 if k is odd,

$$\begin{aligned} \frac{A^k}{H^k} \sum_{p_1^{epsilon_1} \cdots p_k^{\epsilon_k} = 1} \prod_{\ell=1}^{k} \frac{\log p_\ell}{\sqrt{p_\ell}} \mathbf{1}_{[0,x]}\Big(\tfrac{\log p_\ell}{H}\Big) &= A^k c_k \cdot \Big(\frac{1}{H^2} \sum_{\log p \le XH} \frac{\log^2 p}{p}\Big)^{k/2} \\ &= O_k(A^k X^k) \end{aligned}$$

For the remaining terms in which one of $\lambda_1, ..., \lambda_k$ is greater than 1, note that if $\lambda_1, \lambda_2, ..., \lambda_k$ are each no less than 3,

$$\begin{aligned} \frac{A^k}{H^k} \sum_{\substack{p_1^{\lambda_1 \epsilon_1} \cdots p_k^{\lambda_k \epsilon_k} = 1 \\ \lambda_1, ..., \lambda_k \ge 3}} \prod_{\ell=1}^{k} \frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}} \mathbf{1}_{[0,x]}\Big(\tfrac{\lambda_\ell \log p_\ell}{H}\Big) &\le \frac{A^k}{H^k} \Big(\sum_{\lambda \ge 3, p} \frac{\log p}{p^{\lambda/2}} \mathbf{1}_{[0,X]}\Big(\tfrac{\lambda_\ell \log p_\ell}{H}\Big)\Big)^k \\ &= O_k\Big(\frac{A^k}{H^k}\Big). \end{aligned}$$

But because the sum is 0 if $\frac{3\log 2}{H} > X$, this is $O_k(A^k H^k)$ all the same. Finally, if some λ_j is fixed to be equal to 2 – suppose without generality $j = 1$ – then in our sum some p_i must equal p_1. If we with no loss of generality suppose the index i is 2, we have

$$\begin{aligned}
&\frac{A^k}{H^k} \sum_{p_1^{\lambda_1\epsilon_1}\cdots p_k^{\lambda_k\epsilon_k}=1} \prod_{\ell=1}^{k} \frac{\log p_\ell}{\sqrt{p_\ell}} \mathbf{1}_{[0,x]}\Big(\tfrac{\log p_\ell}{H}\Big) \\
&\leq \frac{A^k}{H^k}\Bigg(\sum_{p}\sum_{\lambda_2\geq 1} \frac{\log^2 p}{p^{1+\lambda_2/2}}\Bigg) \sum_{p_3^{\lambda_3\epsilon_3}\cdots p_k^{\lambda_k\epsilon_k}=1} \prod_{\ell=3}^{k} \frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}} \mathbf{1}_{[0,x]}\Big(\tfrac{\lambda_\ell \log p_\ell}{H}\Big) \\
&= O\Bigg(\frac{A^2}{H^2}\frac{A^{k-2}}{H^{k-2}} \sum_{p_3^{\lambda_3\epsilon_3}\cdots p_k^{\lambda_k\epsilon_k}=1} \prod_{\ell=3}^{k} \frac{\log p_\ell}{p_\ell^{\lambda_\ell/2}} \mathbf{1}_{[0,x]}\Big(\tfrac{\lambda_\ell \log p_\ell}{H}\Big)\Bigg).
\end{aligned}$$

An inductive argument shows this is $O_k(A^k/H^k)$, as again, for the sum to be nonzero we must have $1/H \lesssim X$. Since there are only k such cases that some λ_j may be fixed to be 2, we have shown that the sum (5.45) is $O_k(A^k X^k)$. □

From Lemmas 5.6.1 and 5.6.5,

Corollary 5.6.6. *Suppose we are given non-negative integrable σ of mass 1 such that $\hat{\sigma}$ has compact support, and suppose $g_1, ..., g_k$ are in $C_c^2(\mathbb{R})$ and each supported in a region $[-X, X]$ with $X < 1/k$ and each bounded in absolute value by a constant A. Then there exists a T_0 depending only on the region in which $\hat{\sigma}$ is supported so that for $T \geq T_0$,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{k} \widetilde{G}_T(\,\hat{g}_\ell\,, t)\, dt = O_k\Big(A^k\big(\tfrac{1}{\log^k T} + X^k\big)\Big).$$

With a little more work,

Corollary 5.6.7. *For σ as above in Corollary 5.6.6 and $g \in C_c^2(\mathbb{R})$ supported in $[-X, X]$ with $X \leq 1/k$ and bounded in absolute value by a constant A, there exists*

T_0 depending only on the region in which $\hat{\sigma}$ is supported so that for $T \geq T_0$,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left|\widetilde{G}_T(\,\hat{g}_\ell\,, t)\right|^k dt = O_k\Big(A^k\big(\tfrac{1}{\log^k T} + X^k\big)\Big).$$

Proof. By Cauchy-Schwarz,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left|\widetilde{G}_T(\,\hat{g}_\ell\,, t)\right|^k dt \leq \sqrt{\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \big(\widetilde{G}_T(\,\hat{g}_\ell\,, t)\big)^k \big(\widetilde{G}_T(\,\overline{\hat{g}_\ell}\,, t)\big)^k\, dt}.$$

But

$$\overline{\hat{g}(\xi)} = \int_{\mathbb{R}} e(-x\xi)\overline{g}(-x)\, dx$$

and $\overline{g}(-x)$ is also bounded in absolute value by A and supported in $[-X, X]$, so the corollary follows from Corollary 5.6.6. □

Lemma 5.6.4 then follows exactly in the same way as Lemma 5.6.3, by majorizing $\frac{1}{\epsilon}\mathbf{1}_{[a,a+\epsilon]}$ by $V_{a,\epsilon}$, exploiting the compactly supported Fourier tranform of the latter.

5.7 Upper bounds for counts of eigenvalues

In order to produce similar bounds for counts of eigenvalues, we need an analoge of Lemma 5.6.1. This is furnished by a result of Diaconis and Shahshahani [20], which we recall in full once again:

Theorem 5.7.1 (Diaconis-Shahshahani)**.** *Let $\mathcal{U}(n)$ be the set of $n \times n$ unitary matrices endowed with Haar measure. Consider $a = (a_1, ..., a_k)$ and $b = (b_1, ..., b_k)$ with $a_1, a_2, ..., b_1, b_2, ... \in \{0, 1, ...\}$.If $\sum_{j=1}^k ja_j \neq \sum_{j=1}^k jb_j$,*

$$\int_{\mathcal{U}(n)} \prod_{j=1}^k \mathrm{Tr}(g^j)^{a_j}\overline{\mathrm{Tr}(g^j)^{b_j}}\, dg = 0. \tag{5.46}$$

Furthermore, in the case that

$$\max\left(\sum_{j=1}^{k} ja_j, \sum_{j=1}^{k} jb_j\right) \leq n$$

we have

$$\int_{\mathcal{U}(n)} \prod_{j=1}^{k} \mathrm{Tr}(g^j)^{a_j} \overline{\mathrm{Tr}(g^j)^{b_j}}\, dg = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j! \tag{5.47}$$

Recall that a simple manipulation in enumerative combinatorics allows us to rephrase (5.47) as the statement that for integers $j_1, ..., j_k$ such that $|j_1| + \cdots + |j_k| \leq 2N$,

$$\int_{U(N)} \prod_{\ell=1}^{k} \mathrm{Tr}(u^{j_\ell})\, du = \sum \prod_{\lambda} |j_{\mu_\lambda}| \delta_{j_{\mu_\lambda} = -j_{\nu_\lambda}},$$

where the sum is over all partitions of $[k] = \{1, ..., k\}$ into disjoint pairs $\{\mu_\lambda, \nu_\lambda\}$ and $\delta_{j_{\mu_\lambda} = -j_{\nu_\lambda}}$ is 1 or 0 depending upon whether $j_{\mu_\lambda} = -j_{\nu_\lambda}$. For instance, the reader who skipped chapters 3 and 4 would be well to study the example that $\{1, 2, 3, 4\}$ can be partitioned into the disjoint pairs $\{\{1, 2\}; \{3, 4\}\}$, $\{\{1, 3\}; \{2, 3\}\}$, and $\{\{1, 4\}, \{2, 3\}\}$, and we have

$$\begin{aligned}\int_{U(N)} \mathrm{Tr}(u^{j_1})\mathrm{Tr}(u^{j_2})\mathrm{Tr}(u^{j_3})\mathrm{Tr}(u^{j_4})\, du =& |j_1|\delta_{j_1=-j_2}|j_3|\delta_{j_3=-j_4} \\ &+ |j_1|\delta_{j_1=-j_3}|j_2|\delta_{j_2=-j_4} \\ &+ |j_1|\delta_{j_1=-j_4}|j_2|\delta_{j_2=-j_3},\end{aligned}$$

when $|j_1| + |j_2| + |j_3| + |j_4| \leq 2N$.

For the point processes $\mathcal{S}'_N$, by using Poisson summation as in identity (5.32),

Corollary 5.7.2. *For $g_1, ..., g_k \in C_c^2(\mathbb{R})$ satisfying* supp $g_\ell \subset [-\delta_\ell, \delta_\ell]$ *with $\delta_1 + \cdots + \delta_k \leq 2$,*

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \prod_{\ell=1}^{k} \left(\sum_i \hat{g}_\ell(x_i) - \int_{-\infty}^{\infty} \hat{g}_\ell(\alpha)\, d\alpha\right) = \sum \prod_{\lambda} \left(\sum_{j \in \mathbb{Z}\setminus\{0\}} \frac{1}{N} \frac{|j|}{N} g_{\mu_\lambda}\left(\tfrac{j}{N}\right) g_{\nu_\lambda}\left(\tfrac{-j}{N}\right)\right),$$

where the first sum is, as above, over all partitions of $[k]$ *into disjoint parts* $\{\mu_\lambda, \nu_\lambda\}$.

With proofs proceeding exactly as in section 4, we obtain an analogue of Lemma 5.6.4,

Corollary 5.7.3. *For* $g_1, ..., g_k \in C_c^2(\mathbb{R})$ *each supported in a region* $[-X, X]$ *with* $X \leq 1/k$ *and each bounded in absolute value by a constant* A,

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \prod_{\ell=1}^{k} \Big(\sum_i \hat{g}_\ell(x_i) - \int_{-\infty}^{\infty} \hat{g}_\ell(\alpha)\, d\alpha \Big) = O_k(A^k X^k).$$

Corollary 5.7.4. *For* $g \in C_c^2(\mathbb{R})$ *supported in* $[-X, X]$ *with* $X \leq 1/k$, *and with maximum value* A,

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \Big| \sum_i \hat{g}(x_i) - \int_{-\infty}^{\infty} \hat{g}(\alpha)\, d\alpha \Big|^k = O_k(A^k X^k).$$

In the same way, we can produce an analogue of Fujii's bound:

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \Big| \sum_j \eta(x_j) \Big|^k \lesssim_k \Big| \int_{\mathbb{R}} M_k \eta(\alpha)\, d\alpha \Big|^k.$$

For our purposes this is rendered redundant by our ability to explicitly calculate the correlation functions of $\mathcal{S}'_N$, and in particular by knowing Proposition 5.4.7 – that $\mathcal{S}'_N \to S$.

5.8 A Tauberian interchange of averages

Recall that for a weight σ, GUE(σ) is an abbreviation for the proposition that the processes $Z_T(\sigma)$ tend in distribution to the sine-kernel determinantal process $\mathcal{S}$. In this section we show that for many σ, the proposition GUE(σ) is equivalent to GUE($\mathbf{1}_{[1,2]}$), that is to say the GUE Conjecture with which we began the paper.

We use the abbreviation

$$d\lambda_k(t) := \log^k(|t| + 2)\, dt.$$

Theorem 5.8.1. *Let $\sigma_1(t)$ and $\sigma_2(t)$ be non-negative piecewise continuous functions on $\mathbb{R}$ of mass 1 both dominated by a function $\varsigma(t)$ which decreases radially and is an element of $L^1(\mathbb{R}, d\lambda_k)$ for all $k \geq 1$. If for $f_1(x) = e^x \sigma_1(e^x)$ we have $\hat{f}_1(\xi) \neq 0$ for all ξ, then*

$$\mathrm{GUE}(\sigma_1) \Rightarrow \mathrm{GUE}(\sigma_2).$$

Our proof makes use of the first upper bound in section 5.6, the positivity of counts of zeros, and finally Weiner's Tauberian theorem to relate a specific σ to other weights.

We first develop an upper bound in terms of the weight σ. As a corollary of Lemma 5.6.3, making a change of variables $\tau = t/T$ and on the right, recalling the definition (5.34) of L_T, making the change of variables $x = \frac{\log T}{2\pi}(\xi - T\tau)$,

Corollary 5.8.2. *For $\epsilon > 0$ there exists T_0 such that for $T \geq T_0$,*

$$\int \mathbf{1}_{[a,a+\epsilon]}(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau$$
$$\lesssim_k \prod_{j=1}^{k} \|M_k \eta_j\|_{L^1(d\lambda_1)} \left(\int \mathbf{1}_{[a,a+\epsilon]}(\tau)\, d\tau + \int \mathbf{1}_{[a,a+\epsilon]}(\tau) \frac{\log^k(|\tau| + 2)}{\log^k T}\, d\tau \right).$$

for all $a \in \mathbb{R}$ and functions $\eta_1, ..., \eta_k$.

Remark: The importance of this bound is that it (and T_0) is independent of a and test functions η.

From this,

Corollary 5.8.3. *For σ_1 piecewise continuous and dominated by a function ς as* *in Theorem 5.8.1, for* $T \geq T_0$

$$\int_{\mathbb{R}} \sigma_1(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau \lesssim_k \prod_{j=1}^{k} \|M_k \eta_j\|_{L^1(d\lambda_1)} \Big(\|\sigma_1\|_{L^1(d\tau)} + \tfrac{1}{\log^k T} \|\sigma_1\|_{L^1(d\lambda_k)} \Big)$$

where T_0 *depends only on* ς *and* σ_1 *and the implied constant only on* k*.*

Proof of Corollary 5.8.3. Fix k. Let δ be an arbitrary positive number, and choose K so that

$$\int_{|t|>K} \varsigma(\tau) \log^k(|\tau| + 2)\, d\tau < \delta.$$

Likewise, choose ϵ positive but less than 1 so that

$$\int_{|\tau|<K+1} \big(M_\epsilon \sigma_1(\tau) - \sigma_1(\tau) \big) \log^K(|\tau| + 2)\, d\tau < \delta.$$

We have that

$$\begin{aligned} &\int \sigma_1(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau \\ &\lesssim \int M_\epsilon \sigma_1(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau \\ &\lesssim_k \prod_{j=1}^{k} \|M_k \eta_j\|_{L^1(d\lambda_1)} \bigg(\int M_\epsilon \sigma_1(\tau)\, d\tau + \int M_\epsilon \sigma_1(\tau) \tfrac{\log^k(|\tau|+2)}{\log^k T}\, d\tau \bigg) \end{aligned}$$

for $T \geq T_0$ depending only upon ϵ.

Because ς decays away from the origin and dominates σ_1,

$$\int_{|\tau|>K+1} M_\epsilon \sigma_1(\tau) \log^k(|\tau| + 2)\, d\tau < \delta,$$

and so for $T \geq T_0$,

$$\begin{aligned}
&\int \sigma_1(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau \\
&\lesssim_k \prod_{j=1}^{k} \|M_k \eta_j\|_{L^1(d\lambda_1)} \Big(\int_{|\tau < K+1} M_\epsilon \sigma_1(\tau) \Big(1 + \tfrac{\log^k(|\tau|+2)}{\log^k T}\Big)\, d\tau + \delta \cdot \Big(1 + \tfrac{1}{\log^k T}\Big)\Big) \\
&\lesssim_k \prod_{j=1}^{k} \|M_k \eta_j\|_{L^1(d\lambda_1)} \Big(\int \sigma_1(\tau) \Big(1 + \tfrac{\log^k(|\tau|+2)}{\log^k T}\Big)\, d\tau + 2\delta \cdot \Big(1 + \tfrac{1}{\log^k T}\Big)\Big).
\end{aligned}$$

As δ was arbitrary, we can let it be smaller for instance than $\|\sigma_1\|_{L^1(dt)}$ and obtain,

$$\int \sigma_1(\tau) \prod_{j=1}^{k} G_T(\eta_j, T\tau)\, d\tau \lesssim_k \prod_{j=1}^{k} \|M_k \eta\|_{L^1(d\lambda_1)} \Big(\|\sigma_1\|_{L^1(dt)} + \frac{1}{\log^k T} \|\sigma_1\|_{L^1(d\lambda_k)} \Big)$$

for sufficiently large T depending only upon ς and σ_1. □

Before proceeding to a proof of Theorem 5.8.1, we embark on a small digression. Corollary 5.8.3 yields a quick way to see that there is nothing special about using $C_c(\mathbb{R}^k)$ functions to test whether $Z_T(\sigma) \to \mathcal{S}$ in distribution.

Proposition 5.8.4. *For each $k \geq 1$, let $\mathcal{A}_k$ be a collection of functions $\eta : \mathbb{R}^k \to \mathbb{R}$ such that*

1. *For any $\eta \in \mathcal{A}_k$, η decays in each variable at a 3/2-power rate; that is, there is a constant A_η so that*

$$|\eta(x_1, ..., x_k)| \leq \frac{A_\eta}{(1 + |x_1|^{3/2}) \cdots (1 + |x_k|^{3/2})},$$

   and more

2. *For any $\rho \in C_c(\mathbb{R})$ any any $\epsilon > 0$, there exists $\eta \in \mathcal{A}_k$ so that for all $x \in \mathbb{R}^k$,*

$$|\rho(x) - \eta(x)| \leq \frac{\epsilon}{(1 + |x_1|^{3/2}) \cdots (1 + |x_k|^{3/2})}$$

Then for any $\sigma_1 : \mathbb{R} \to \mathbb{R}_+$ positive, piecewise continuous, and of mass 1, and dominated by a function ς as in Theorem 5.8.3, GUE(σ_1) *is equivalent to the statement that for all $k \geq 1$ and all $\eta \in \mathcal{A}_k$,*

$$\mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{\substack{j_1,...,j_k \\ \text{distinct}}} \eta(\xi_{j_1}, ..., \xi_{j_k}) \sim \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,...,j_k \\ \text{distinct}}} \eta(x_{j_1}, ..., x_{j_k}). \tag{5.48}$$

Remark: If not for the fact that the collections $\mathcal{A}_k$ may contain η which are not compactly supported, this proposition would be standard. The 3/2 power decay in (i) and (ii) is chosen for convenience rather than canonically. Some decay in the tails of functions η is necessary for the proposition to be true, and for technical reasons later on to have a proposition with for η whose tails decay more slowly than quadratically will be important.

Proof of Proposition 5.8.4. Recall that GUE(σ_1) is equivalent to the statement that for all $k \geq 1$ and all $\rho \in C_c(\mathbb{R}^k)$,

$$\mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{\substack{j_1,...,j_k \\ \text{distinct}}} \rho(\xi_{j_1}, .., \xi_{j_k}) \sim \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,...,j_k \\ \text{distinct}}} \rho(x_{j_1}, .., x_{j_k}).$$

By inductively including lower correlations, we see that this is equivalent to the statement that for all $k \geq 1$ and $\rho \in C_c(\mathbb{R}^k)$,

$$\mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{j_1,...,j_k} \rho(\xi_{j_1}, .., \xi_{j_k}) \sim \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{j_1,...,j_k} \rho(x_{j_1}, .., x_{j_k}).$$

The sums here are over indices which needn't be distinct. By applying Corollary 5.8.3 for sufficiently large T, for any η,

$$\Big| \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{j_1,...,j_k} \eta(\xi_{j_1}, ..., \xi_{j_k}) \Big| \lesssim_{k,\sigma_1} \int_{\mathbb{R}^k} M_k' \eta(x_1, ..., x_k)\, d\lambda_1(x_1) \cdots d\lambda(x_k),$$

where

$$M_k'\eta(x_1, ..., x_k) = \sum_{\nu\in\mathbb{Z}^k} \Big(\sup_{I_k'(\nu)} |\eta|\Big) \mathbf{1}_{I_k'(\nu)}(x),$$

where $I_k'(\nu)$ abbreviates the k-dimensional cube $k\nu + [-k/2, k/2)^k$.

Note that for any $\epsilon > 0$, any $\eta : \mathbb{R}^k \to \mathbb{R}^k$ which decays in each variable in the sense of condition (i) can be approximated by $\rho \in C_c(\mathbb{R}^k)$ so that both

$$\Big| \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{j_1,...,j_k} \big(\eta(x_{j_1}, ..., x_{j_k}) - \rho(x_{j_1}, ..., x_{j_k})\big)\Big| < \epsilon,$$

and

$$\int_{\mathbb{R}^k} M_k'(\eta - \rho)\, d\lambda(x_1) \cdots d\lambda(x_k) < \epsilon.$$

It therefore follows that for continuous $\eta : \mathbb{R}^k - \mathbb{R}$ decaying in each variable as in (i), GUE(σ_1) implies

$$\overline{\lim_{T\to\infty}} \Big| \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{j_1,...,j_k} \eta(\xi_{j_1}, ..., \xi_{j_k}) - \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{j_1,...,j_k} \eta(\xi_{j_1}, ..., \xi_{j_k})\Big| < 2\epsilon.$$

Because ϵ is arbitrary, this shows that GUE(σ_1) implies (5.48) for any $\eta \in \mathcal{A}_k$.

In the opposite direction, suppose that for all $k \geq 1$ and any $\eta \in \mathcal{A}_k$, (5.48) holds. Let ρ be an arbitrary element of $C_c(\mathbb{R}^k)$. For any $\epsilon > 0$, there exists an $\eta \in \mathcal{A}_k$ so that for all $x \in \mathbb{R}^k$,

$$|\eta(x) - \rho(x)| < \frac{\epsilon}{(1 + |x_1|^{3/2}) \cdots (1 + |x_k|^{3/2})}.$$

Thus it follows as before that

$$\begin{aligned}
&\Bigg| \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \rho(\xi_{j_1},\dots,\xi_{j_k}) - \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \eta(\xi_{j_1},\dots,\xi_{j_k}) \Bigg| \\
&\leq \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{j_1,\dots,j_k} \big| \rho(\xi_{j_1},\dots,\xi_{j_k}) - \eta(\xi_{j_1},\dots,\xi_{j_k}) \big| \\
&\lesssim_{k,\sigma_1} \epsilon
\end{aligned}$$

and

$$\Bigg| \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \rho(x_{j_1},\dots,x_{j_k}) - \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \eta(x_{j_1},\dots,x_{j_k}) \Bigg| \lesssim_k \epsilon.$$

As ϵ was arbitrary it follows that

$$\mathop{\mathbf{E}}_{Z_T(\sigma_1)} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \rho(\xi_{j_1},\dots,\xi_{j_k}) \sim \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} \rho(x_{j_1},\dots,x_{j_k}).$$

Because ρ was arbitrary, this is just GUE(σ_1). □

We return to the proof of Theorem 5.8.1. Recall Weiner's Tauberian Theorem:

Theorem 5.8.5 (Weiner)**.** *For $f_1, f_2 \in L^1(\mathbb{R}, dt)$ with $\hat{f}_1(\xi) \neq 0$, for any $\epsilon > 0$ there exists constants $w_1, ..., w_n$ and $a_1, ..., a_n$ so that*

$$\|f_2(t) - \sum a_i f_1(t - w_i)\|_{L^1(dt)} < \epsilon$$

That is, $\mathrm{span}_{w\in\mathbb{R}}\{f(t-w)\}$ is dense in $L^1(\mathbb{R}, dt)$. See for instance [48] for a proof.

With this we can proceed to a

Proof of Theorem 5.8.1. Choose $\epsilon > 0$. Weiner's Tauberian Theorem implies that there exist positive $h_1, ..., h_n$ and (possibly negative) $a_1, ..., a_n$ so that $a_1 + \cdots + a_n =$

1 and

$$\|\sigma_1(\tau) - \sum a_i h_i \sigma_1(\tau/h_i)\|_{L^1(dt)} < \epsilon.$$

Because σ_2 and σ_1 are both of mass 1, we can choose $a_1, ..., a_n$ so that $a_1 + \cdots + a_n = 1$. Because linear combinations of seperable and continously differentiable functions are dense in $C_c(\mathbb{R}^k)$, an expansion into lower order correlations shows that for σ either of σ_1 or σ_2, GUE(σ) is equivalent to the statement that for all k and continuously differentiable and compactly supported η_1, ..., η_k,

$$\lim_{T\to\infty} \mathop{\mathbf{E}}_{Z_T(\sigma)} \prod_{\ell=1}^{k} \sum_i \eta_\ell(\xi_i) = \prod_{\ell=1}^{k} \sum_i \eta_\ell(\xi_i).$$

Because any continuously differentiable η can be written as the difference of two radially non-increasing functions, e.g. for $x > 0$,

$$\eta(x) = \Big(\int_x^\infty \Big(\frac{d\eta}{dx}\Big)_+ dx\Big) - \Big(\int_x^\infty -\Big(\frac{d\eta}{dx}\Big)_- dx\Big),$$

GUE(σ) is equivalent to the statement that

$$\begin{aligned}\int \sigma(\tau) \prod_{\ell=1}^{k} G_T(\eta_\ell, T\tau)\, d\tau = & \mathop{\mathbf{E}}_{Z_T(\sigma)} \prod_{\ell=1}^{k} \sum_i \eta_j(\xi_i) \\ & \sim \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \sum_i \eta_j(\xi_i)\end{aligned}$$

for any collection $\eta_1, ..., \eta_j$ of radially non-increasing functions, continous and compactly supported.

We make use of a monotonicity argument to show that on the hypothesis of Theorem 5.8.1 for any $h > 0$,

$$h \int \sigma_1\Big(\frac{t}{h}\Big) \prod_{\ell=1}^{k} G_T(\eta_\ell, T\tau)\, d\tau \sim \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \sum_i \eta_\ell(\xi_i). \tag{5.49}$$

Clearly this is true for $h = 1$. For other h, the left hand side of (5.49) is equal to

$$\int \sigma_1(\tau) \prod_{\ell=1}^{k} G_T(\eta_\ell, Th\tau)\, d\tau.$$

If we define $\eta[\rho](x) := \eta(\rho^{-1}x)$, then for $\rho_1 < \rho_2$ (as long as η is non-increasing radially) $\eta[\rho_1] \leq \eta[\rho_2]$ pointwise. Also note

$$\begin{aligned} G_T(\eta_\ell, Th\tau) &= \sum_\gamma \eta_\ell\Big(\tfrac{\log T}{2\pi}(\gamma - Th\tau)\Big) \\ &= G_{Th}\Big(\eta_\ell\big[1 + \tfrac{\log h}{\log T}\big], Th\tau\Big). \end{aligned}$$

We consider first the case that $h < 1$. In this case, for $T > T'$ (because the quantity $1 + \frac{\log h}{\log T}$ decreases as T increases),

$$\begin{aligned} \int \sigma_1(\tau) \prod_{\ell=1}^{k} G_T(\eta_\ell, Th\tau)\, d\tau &\leq \int \sigma_1(\tau) \prod_{\ell=1}^{k} G_{Th}\Big(\eta_\ell\big[1 + \tfrac{\log h}{\log T'}\big], Th\tau\Big)\, d\tau \\ &\sim \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\sum_i \eta_\ell\big[1 + \tfrac{\log h}{\log T'}\big](\xi_i) \Big). \end{aligned} \qquad (5.50)$$

For the same reason,

$$\begin{aligned} \int \sigma_1(\tau) \prod_{\ell=1}^{k} G_T(\eta_\ell, Th\tau)\, d\tau &\geq \int \sigma_1(\tau) \prod_{\ell=1}^{k} G_{Th}(\eta_\ell, Th\tau)\, d\tau \\ &\sim \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\sum_i \eta_\ell(\xi_i) \Big). \end{aligned} \qquad (5.51)$$

As $T \to \infty$, we may choose T' arbitrarily large, and because the resulting limiting expression in (5.50) is continuous in $\frac{\log h}{\log T'}$, we have (5.49) as claimed.

In the case that $h < 1$, we may use the same argument, with the inequalities in both (5.50) and (5.51) reversed.

To complete the proof, note that by Corollary 5.8.3

$$\begin{aligned}
&\overline{\lim}_{T\to\infty}\left|\int_{-\infty}^{\infty}\big(\sigma_2(\tau)-\sum a_i h_i \sigma_1(\tau/h_i)\big)\prod_{\ell=1}^{k} G_T(\eta_\ell, T\tau)\,d\tau\right| \\
&\lesssim_{\eta,k} \overline{\lim}_{T\to\infty}\left(\Big\|\sigma_2(\tau)-\sum a_i h_i\sigma_1(\tau/h_i)\Big\|_{L^1(dt)}+\frac{1}{\log^k T}\Big\|\sigma_2(\tau)-\sum a_i h_i \sigma_1(\tau/h_i)\Big\|_{L^1(d\lambda_k(t))}\right) \\
&< \epsilon.
\end{aligned}$$

Because ϵ was arbitrary, (5.49) and the fact that $a_1+\cdots+a_n=1$ yield that

$$\begin{aligned}
\mathop{\mathbf{E}}_{Z_T(\sigma_2)}\prod_{\ell=1}^{k}\sum_i \eta_\ell(\xi_i) &= \int \sigma_2(\tau)\prod_{\ell=1}^{k} G_T(\eta_\ell, T\tau)\,d\tau \\
&\sim \mathop{\mathbf{E}}_{\mathcal{S}}\prod_{\ell=1}^{k}\sum_i \eta_\ell(x_i),
\end{aligned}$$

as claimed. □

We note two masses σ for which GUE(σ) reproduces itself to other masses.

Corollary 5.8.6. *The* GUE *Conjecture (*GUE($\mathbf{1}_{[1,2]}$)*, that is) implies* GUE(σ_2) *for any σ_2 which is piecewise continuous, in $L^1(d\lambda_k)$ for all k, and dominated by a decreasing function.*

Proof. It is apparent that $\sigma_1 := \mathbf{1}_{[1,2]}$ is itself non-negative, in $L^1(d\lambda_k)$ for all k, and non-increasing radially. In addition, the function $f_1(t) := e^t\mathbf{1}_{[1,2]}(e^t)$ satisfies

$$\hat{f}_1(\xi) = \frac{2^{1-i2\pi\xi}-1}{1-i2\pi\xi} \neq 0,$$

for all ξ. □

Likewise,

Corollary 5.8.7. *For*

$$\sigma_1(t) := \frac{1}{2\pi}\Big(\frac{\sin t/2}{t/2}\Big)^2, \tag{5.52}$$

GUE(σ_1) *implies* GUE(σ_2) *for any* σ_2 *which is piecewise continuous, in* $L^1(d\lambda_k)$ *for all* k*, and dominated by a decreasing function.*

Proof. Again it is apparent that σ_1 is non-negative and may be dominated by a function that is in $L^1(d\lambda_k)$ for all k and non-increasing radially. If $f_1(t) := e^t\sigma_1(e^t)$, then

$$\hat{f}_1(\xi) = \frac{\Gamma(-i2\pi\xi)\sin(-i\pi^2\xi)}{\pi(1-i2\pi\xi)} \neq 0$$

for all ξ. □

Corollary 5.8.8. *The GUE Conjecture is equivalent to* $GUE(\sigma_1)$ *where* σ_1 *is defined in* (5.52).

For us the significance of this particular σ_1 is that

$$\hat{\sigma}_1\left(\tfrac{x}{2\pi}\right) = (1-|x|)_+.$$

5.9 Approximating a principal value integral

We have come to the point to introduce the cutoff $f|_\epsilon$ of functions f mentioned in the outline in section 5.5. Recall (5.29) and (5.30), the definition of the bump function α and rescaled bump function α_ϵ of width 2ϵ. (Earlier our interest was a rescaling with large width, in the context of the present chapter, we rescale to small width.) The reader should check that $\alpha(0) = 1$ and $\alpha'(0) = \alpha''(0) = 0$. Using α_ϵ, we define

$$\omega_\epsilon(x) := 1 - \alpha_\epsilon(x)$$

$$\Omega_\epsilon(x) := \omega_\epsilon(x)\mathbf{1}_{\mathbb{R}_+}(x).$$

It is easy to verify that $\Omega_\epsilon \in C^2(\mathbb{R})$.

We define the cutoff function $f|_\epsilon$ for $f : \mathbb{R} \to \mathbb{R}$ by

$$f|_\epsilon(x) := f(x)\Omega_\epsilon(x)$$

For small ϵ this approximates $f \cdot \mathbf{1}_{\mathbb{R}_+}$. Further, for $b > a > 0$ we define

$$f|_a^b(x) := f|_a(x) - f|_b(x),$$

which is supported on the interval $[0, b]$ and morally acts as a restriction of f to the interval $[a, b]$.

The purpose of this section is to show that

Lemma 5.9.1. *For admissible g (see definition 5.2.3), and non-negative and integrable σ such that $\hat{\sigma}$ is compactly supported, there exists T_0 depending only on the region in which $\hat{\sigma}$ is supported so that for all $T > T_0$ and all $\epsilon > 0$,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left|\widetilde{G}_T\big((g|_\epsilon)\hat{}, t\big)\right|^k dt \lesssim_k \|g\|_{L^1(\mathbb{R})}^k + \|g'\|_{L^1(\mathbb{R})}^k + \|g''\|_{L^1(\mathbb{R})}^k$$

for $k \geq 1$.

This lemma may be at first surprising in the same way as the upper bound Lemma 5.6.4. In fact, it is true for much the same reason as Lemma 5.6.4. A partial explanation for the bound is that while $(g\mathbf{1}_{\mathbb{R}_+})\hat{}$ is not integrable for g smooth and $g(0) \neq 0$, for such g the principal value integral

$$\lim_{R\to\infty} \int_{-R}^{R} (g \cdot \mathbf{1}_{\mathbb{R}_+})\hat{}(\xi)\, d\xi$$

has the limit

$$= \tfrac{1}{2} g(0),$$

owing to the oscillatory nature of $g\mathbf{1}_{\mathbb{R}_+}$. For small ϵ, $g|_\epsilon$ resembles $g\mathbf{1}_{\mathbb{R}_+}$ and so in particular $\|g|_\epsilon\|_{L^1}$ will grow without bound. But at the same time $g|_\epsilon$ will capture the same oscillation as $g\mathbf{1}_{\mathbb{R}_+}$ and (much as in Lemma 5.6.4), this substantially reduces the size of $\widetilde{G}_T((g|_\epsilon)\hat{}, t)$.

In proving Lemma 5.9.1, it will be useful to have in mind some standard explicit bounds on the decay of $\hat{g}$ for $g \in C_c^2(\mathbb{R})$. Note that

$$\hat{g}(\xi) = -\frac{1}{4\pi^2\xi^2}\int_{\mathbb{R}} g''(x)e(-x\xi)\,dx$$

and because we have for all ξ (in particular for ξ close to the origin),

$$|\hat{g}(\xi)| \le \|g\|_{L^1(\mathbb{R})},$$

we have the estimate

$$\hat{g}(\xi) = O\Big(\frac{\|g\|_1 + \|g''\|_1}{\xi^2+1}\Big). \tag{5.53}$$

With this in mind,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} L_T(\,|\hat{g}|\,,t)^k\,dt = O_k\big((\|g\|_{L^1} + \|g''\|_{L^1})^k\big). \tag{5.54}$$

and so a trivial consequence then of Lemma 5.6.2 is

Lemma 5.9.2. *For σ non-negative and integrable such that $\hat{\sigma}$ is compactly supported, there exists a T_0 depending only on the region in which $\hat{\sigma}$ is supported, so that for all $T \ge T_0$,*

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T} G_T(\,|\hat{g}|\,,t)^k\,dt = O_k\big((\|g\|_{L^1} + \|g''\|_{L^1})^k\big).$$

From this, it is a short path to Lemma 5.9.1.

Proof of Lemma 5.9.1. From Minkowski's inequality,

$$\Big(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((g|_\epsilon)\hat{}, t\big)\big|^k\,dt\Big)^{1/k} \leq \Big(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((g|_\epsilon^{1/k})\hat{}, t\big)\big|^k\,dt\Big)^{1/k} + \Big(\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((g|_{1/k})\hat{}, t\big)\big|^k\,dt\Big)^{1/k}. \tag{5.55}$$

From Lemma 5.6.4, there is T_0 so that for $T \geq T_0$,

$$\begin{aligned}\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((g|_\epsilon^{1/k})\hat{}, t\big)\big|^k\,dt &\lesssim \|g\|_\infty^k\Big(\frac{1}{\log^k T} + \Big(\frac{1}{k}\Big)^k\Big)\\ &\lesssim \|g\|_\infty^k. \end{aligned}\tag{5.56}$$

On the other hand, applying equation (5.54) and its consequence, Lemma 5.9.2,

$$\begin{aligned}\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((g|_{1/k})\hat{}, t\big)\big|^k\,dt &\lesssim_k \int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\Big(\big|G_T\big((g|_{1/k})\hat{}, t\big)\big|^k + \big|L_T\big((g_{1/k}\hat{}, t\big)\big|^k\,dt\\ &\lesssim_k \|\Omega_{1/k}g\|_{L^1}^k + \|(\Omega_{1/k}g)''\|_{L^1}^k\\ &\lesssim_k \|g\|_{L^1} + \|g'\|_{L^1} + \|g''\|_{L^1}\end{aligned}\tag{5.57}$$

as $\Omega_{1/k}, \Omega'_{1/k}$, and $\Omega''_{1/k}$ are all bounded. (Here we have repeatedly used the inequality $(a+b)^k \lesssim_k a^k + b^k$.)

Substituting (5.56) and (5.57) into (5.55) gives the lemma. □

5.10 Zeros and arithmetic

From the Tauberian result, Corollary 5.8.8, the GUE Conjecture is equivalent to the claim GUE(σ_1), for σ_1 defined in (5.52). In this section we prove Theorem 5.2.4. Our proof is broken into two parts; we first show that the GUE Conjecture implies the identity (5.12) for admissible functions, and in a seperate second proof we demonstrate the converse.

Proof of Theorem 5.2.4: The GUE Conjecture implies (5.12). We begin by establishing that for fixed admissible f, there exists some positive ϵ_T (depending on T) so that

$$\widetilde{G}_T\big[(f|_{\epsilon_T})\hat{\ }, t\big] = \frac{-1}{\log T}\int_{-\infty}^{\infty} f\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x) + O_f\Big(\tfrac{1}{\log T}\Big), \tag{5.58}$$

and

$$\widetilde{G}_T\big[(f|_{\epsilon_T})\check{\ }, t\big] = \frac{-1}{\log T}\int_{-\infty}^{\infty} f\Big(\tfrac{x}{\log T}\Big)e^{-ixt}\,dz(x) + O_f\Big(\tfrac{1}{\log T}\Big), \tag{5.59}$$

For, for admissible f, there is some $\alpha < 1/2$ such that

$$\frac{1}{\log T}\int_{-\infty}^{0} f\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x) = O_f\bigg(\frac{1}{\log T}\int_{-\infty}^{0} e^{x(1/2-\alpha)}\,dx\bigg) = O_f\Big(\tfrac{1}{\log T}\Big).$$

and by continuity there exists some $\epsilon_T > 0$ so that

$$\frac{1}{\log T}\int_{0}^{\infty}(f - f|_{\epsilon_T})\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x) \le \frac{1}{\log T}$$

as $(f - f|_\epsilon)(x) \to 0$ pointwise for all $x > 0$ as $\epsilon \to 0^+$. (Of course, one could choose ϵ_T in a way that the left hand side is much smaller than $1/\log T$, if desired.)

On the other hand from Proposition 5.4.9,

$$\begin{aligned}\widetilde{G}_T\big((f|_{\epsilon_T})\hat{\ }, t\big) &= \frac{-1}{\log T}\int_{-\infty}^{\infty}(f|_{\epsilon_T})\Big(\tfrac{x}{\log T}\Big)e^{ixt} + (f|_{\epsilon_T})\Big(\tfrac{-x}{\log T}\Big)e^{-ixt}\,dz(x)\\ &= \frac{-1}{\log T}\int_{-\infty}^{\infty}(f|_{\epsilon_T})\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x) + O_f\Big(\tfrac{1}{\log T}\Big)\end{aligned}$$

Combining these equations gives (5.58), and (5.59) can be proved the same way (or alternatively, by conjugation). Note that we may suppose $\epsilon_T \to 0$, and if (5.58) and (5.59) hold true for some ϵ_T, they also hold true for any ϵ'_T with $\epsilon'_T \le \epsilon_T$.

We also have for admissible $f_1, ..., f_j$, $g_1, ..., g_k$, with $f := f_1 \otimes \cdots \otimes f_j$, $g :=$

$g_1 \otimes \cdots \otimes g_k$,

$$\Psi_T(f;g) = \frac{1}{\log^{j+k} T} \int_{\mathbb{R}^k} \int_{\mathbb{R}^j} f\Big(\tfrac{x}{\log T}\Big) g\Big(\tfrac{y}{\log T}\Big) \hat{\sigma_1}\big(\tfrac{T}{2\pi}(x_1 + \cdots + x_k - y_1 - \cdots - y_k)\big)$$
$$\times\ d^j z(x)\, d^k z(y)$$
$$= \frac{1}{\log^{j+k} T} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \Bigg(\prod_{\ell=1}^{j} \int_{-\infty}^{\infty} f\Big(\tfrac{x_\ell}{\log T}\Big) e^{ix_\ell t}\, dz(x_\ell)$$
$$\times \prod_{\ell'=1}^{k} \int_{-\infty}^{\infty} g_{\ell'}\Big(\tfrac{y_\ell}{\log T}\Big) e^{-iy_\ell}\, dz(y_\ell)\Bigg)\, dt$$

We are above able to interchange the order of integrations in the variable t or other variables as an application of Fubini's theorem because for fixed T and any admissible function f_ℓ (or $g_{\ell'}$) above,

$$\int_{\mathbb{R}} \Big| f_\ell\Big(\tfrac{x}{\log T}\Big)\Big|\, d\big(\psi(e^x) + e^x\big) < +\infty.$$

Hence from this representation of Ψ_T and (5.58) and (5.59), there is some $\epsilon_T \to \infty$ such that

$$\Psi_T(f;g) = (-1)^{j+k} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \Big(\widetilde{G}_T\big((f_\ell|_{\epsilon_T})\hat{}, t\big) + O_f\big(\tfrac{1}{\log T}\big)\Big)$$
$$\times \prod_{\ell'=1}^{k} \Big(\widetilde{G}_T\big((g_{\ell'}|_{\epsilon_T})\check{}, t\big) + O_g\big(\tfrac{1}{\log T}\big)\Big)\, dt$$
$$= (-1)^{j+k} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \widetilde{G}_T\big((f_\ell|_{\epsilon_T})\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_{\epsilon_T})\check{}, t\big)\, dt + O_{f,g}\Big(\tfrac{1}{\log T}\Big), \tag{5.60}$$

the second line following from expanding the product in the first, and using Hölder's inequality and Lemma 5.9.1 to bound those terms in which an error term appears.

We will show shortly that for all $\epsilon > \rho > 0$,

$$\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \widetilde{G}_T\big((f_\ell|_\rho)\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_\rho)\check{}, t\big)\, dt \tag{5.61}$$
$$= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{k} \widetilde{G}_T\big((f_\ell|_\epsilon)\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_\epsilon)\check{}, t\big)\, dt + O_{f,g}(\epsilon),$$

and likewise that

$$\mathop{\mathbf{E}}_{\mathcal{S}'_N} \Big(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\rho)\hat{}(x_i) \Big) \Big(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\rho)\check{}(x_i) \Big) \tag{5.62}$$
$$= \mathop{\mathbf{E}}_{\mathcal{S}'_N} \Big(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{}(x_i) \Big) \Big(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{}(x_i) \Big) + O_{f,g}(\epsilon),$$

Let us for the moment assume the truth of these bounds (5.61) and (5.62) to see that they allow us to derive identity (5.12) on the GUE Conjecture. From (5.60) and (5.61), with $\rho = \epsilon_T$, for any $\epsilon > 0$, for sufficiently large T,

$$\Psi_T(f; g) = (-1)^{j+k} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \widetilde{G}_T\big((f_\ell|_\epsilon)\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_\epsilon)\check{}, t\big)\, dt + O_{f,g}(\epsilon). \tag{5.63}$$

But from (5.36), because

$$\int_{\mathbb{R}} (f_\ell|_\epsilon)\hat{}(\alpha)\, d\alpha = \int_{\mathbb{R}} (g_{\ell'}|_\epsilon)\check{}(\alpha)\, d\alpha = 0$$

for all ℓ, ℓ', we have

$$\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \widetilde{G}_T\big((f_\ell|_\epsilon)\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_\epsilon)\check{}, t\big)\, dt$$
$$= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \Big(G_T\big((f_\ell|_\epsilon)\hat{}, t\big) + O_{f,\epsilon}\Big(\tfrac{1}{\log T}\Big)\Big) \prod_{\ell'=1}^{k} \Big(G_T\big((g_{\ell'}|_\epsilon)\check{}, t\big) + O_{g,\epsilon}\Big(\tfrac{1}{\log T}\Big)\Big)\, dt$$
$$= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} G_T\big((f_\ell|_\epsilon)\hat{}, t\big) \prod_{\ell'=1}^{k} G_T\big((g_{\ell'}|_\epsilon)\check{}, t\big)\, dt + O_{f,g,\epsilon}\Big(\tfrac{1}{\log T}\Big),$$

using in the last step the Fujii upper bound, Lemma 5.6.2. (The last line could also be obtained from the GUE Conjecture itself, since we are at this point assuming it.) We substitute this in equation (5.63).

In the language of point processes what we thus show is that

$$\Psi_T(f;g) = \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \left(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{\ }(\xi_i) \right) \left(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{\ }(\xi_i) \right) + O_{f,g}(\epsilon) + O_{f,g,\epsilon}\left(\frac{1}{\log T}\right). \tag{5.64}$$

GUE(σ_1) implies that

$$\begin{aligned} &\lim_{T\to\infty} \mathop{\mathbf{E}}_{Z_T(\sigma_1)} \left(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{\ }(\xi_i) \right) \left(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{\ }(\xi_i) \right) \\ &= \mathop{\mathbf{E}}_{\mathcal{S}} \left(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{\ }(x_i) \right) \left(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{\ }(x_i) \right). \end{aligned}$$

In particular, because ϵ is arbitrary in (5.64), $\Psi_T(f;g)$ has a limit as $T \to \infty$ for admissible f, g.

But in turn (from Proposition 5.4.7),

$$\begin{aligned} &\mathop{\mathbf{E}}_{\mathcal{S}} \left(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{\ }(x_i) \right) \left(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{\ }(x_i) \right) \\ &= \lim_{N\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}'_N} \left(\prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon)\hat{\ }(x_i) \right) \left(\prod_{\ell'=1}^{k} \sum_i (g_{\ell'}|_\epsilon)\check{\ }(x_i) \right) \end{aligned}$$

Note that, for any $\epsilon_N > 0$,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{\mathcal{S}'_N}\left(\prod_{\ell=1}^{j}\sum_i (f_\ell|_{\epsilon_N})\hat{}(x_i)\right)\left(\prod_{\ell'=1}^{k}\sum_i (g_{\ell'}|_{\epsilon_N})\check{}(x_i)\right)\\
&= \frac{1}{N^{j+k}}\sum_{r\in\mathbb{Z}^j}\sum_{s\in\mathbb{Z}^k}\prod_{\ell=1}^{j} f_\ell|_{\epsilon_N}\left(\tfrac{-r_\ell}{N}\right)\prod_{\ell'=1}^{k} g_{\ell'}|_{\epsilon_N}\left(\tfrac{s_{\ell'}}{N}\right)\int_{U(N)}\prod_{\ell=1}^{j}\mathrm{Tr}(u^{r_\ell})\prod_{\ell'=1}^{k}\mathrm{Tr}(u^{s_{\ell'}})\,du\\
&= \frac{1}{N^{j+k}}\sum_{r\in\mathbb{N}_+^j}\sum_{s\in\mathbb{N}_+^k}\prod_{\ell=1}^{j} f_\ell|_{\epsilon_N}\left(\tfrac{r_\ell}{N}\right)\prod_{\ell'=1}^{k} g_{\ell'}|_{\epsilon_N}\left(\tfrac{s_{\ell'}}{N}\right)\int_{U(N)}\prod_{\ell=1}^{j}\mathrm{Tr}(u^{-r_\ell})\prod_{\ell'=1}^{k}\mathrm{Tr}(u^{s_{\ell'}})\,du\\
&= \frac{1}{N^{j+k}}\sum_{r\in\mathbb{N}_+^j}\sum_{s\in\mathbb{N}_+^k}\prod_{\ell=1}^{j} f_\ell|_{\epsilon_N}\left(\tfrac{r_\ell}{N}\right)\prod_{\ell'=1}^{k} g_{\ell'}|_{\epsilon_N}\left(\tfrac{s_{\ell'}}{N}\right)\int_{U(N)}\prod_{\ell=1}^{j}\mathrm{Tr}(u^{r_\ell})\prod_{\ell'=1}^{k}\mathrm{Tr}(u^{-s_{\ell'}})\,du,
\end{aligned} \tag{5.65}$$

using Proposition 5.4.8 (that mixed moments of traces are real valued) in the last line.

For any function f, for $\epsilon_N \le 1/N$,

$$f|_{\epsilon_N}(r/N) = f(r/N)$$

for any positive integer r. Therefore for such ϵ_N, (5.65) is just

$$(-1)^{j+k}\Theta_N(f;g).$$

Letting ϵ be arbitrary, and using $\rho = \epsilon_N$ in (5.62), we see, in the same way as for Ψ_T, that $\Theta_N(f;g)$ has a limit as $N \to \infty$. But for any $\epsilon > 0$, both limits will be within $O_{f,g}(\epsilon)$ of

$$(-1)^{j+k}\mathop{\mathbf{E}}_{\mathcal{S}}\left(\prod_{\ell=1}^{j}\sum_i (f_\ell|_{\epsilon})\hat{}(x_i)\right)\left(\prod_{\ell'=1}^{k}\sum_i (g_{\ell'}|_{\epsilon})\check{}(x_i)\right)$$

and therefore $O_{f,g}(\epsilon)$ of each other. Because ϵ is arbitrary this is (5.12).

We therefore need only verify (5.61) and (5.62).

To verify (5.61), note that

$$\begin{aligned}\widetilde{G}_T\big((f_\ell|_\rho)\hat{}, t\big) &= \widetilde{G}_T\big((f_\ell|_\rho^\epsilon)\hat{}, t\big) + \widetilde{G}_T\big((f_\ell|_\epsilon)\hat{}, t\big) \\ &:= a_\ell + A_\ell.\end{aligned}$$

In addition to this shorthand, we also use

$$\begin{aligned}b_{\ell'} &:= \widetilde{G}_T\big((g_{\ell'}|_\rho^\epsilon)\hat{}, t\big) \\ B_{\ell'} &:= \widetilde{G}_T\big((g_{\ell'}|_\epsilon)\hat{}, t\big).\end{aligned}$$

Substituted into (5.61), we show that the terms $a_\ell, b_{\ell'}$ make a small contribution. More exactly,

$$\begin{aligned}&\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} \widetilde{G}_T\big((f_\ell|_\rho)\hat{}, t\big) \prod_{\ell'=1}^{k} \widetilde{G}_T\big((g_{\ell'}|_\rho)\check{}, t\big)\, dt \qquad (5.66)\\ &= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} (a_\ell + A_\ell) \prod_{\ell'=1}^{k} (b_{\ell'} + B_{\ell'})\, dt \\ &= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} A_\ell \prod_{\ell'=1}^{k} B_{\ell'}\, dt \\ &\quad + \sum_{\substack{\emptyset \subseteq J \subseteq [j] \\ \emptyset \subseteq K \subseteq [k] \\ J \cup K \neq \emptyset}} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell \in J} a_\ell \prod_{\lambda \notin J} A_\lambda \prod_{\ell' \in K} b_{\ell'} \prod_{\lambda' \notin K} B_{\lambda'}\, dt\end{aligned}$$

But for any of the terms in this last sum, by Hölder's inequality,

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell\in J} a_\ell \prod_{\lambda\notin J} A_\lambda \prod_{\ell'\in K} b_{\ell'} \prod_{\lambda'\notin K} B_{\lambda'}\, dt \\
&\le \prod_{\ell\in J} \left(\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} |a_\ell|^{j+k}\, dt \right)^{1/(j+k)} \prod_{\lambda\notin J} \left(\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} |A_\lambda|^{j+k}\, dt \right)^{1/(j+k)} \\
&\quad \times \prod_{\ell'\in K} \left(\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} |b_{\ell'}|^{j+k}\, dt \right)^{1/(j+k)} \prod_{\lambda'\notin K} \left(\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} |B_{\lambda'}|^{j+k}\, dt \right)^{1/(j+k)} \\
&= \prod_{\ell\in J} O_{f_\ell}(\epsilon) \prod_{\lambda\notin J} O_{f_\ell}(1) \prod_{\ell'\in K} O_{g_{\ell'}}(\epsilon) \prod_{\lambda'\notin K} O_{g_{\lambda'}}(1),
\end{aligned}$$

for sufficiently large T. Here we have used Lemma 5.6.2 (the Fujii bound) to bound those terms with A_λ or $B_{\lambda'}$, and Corollary 5.6.7 to bound those terms with a_ℓ or b_ℓ', recalling that $f_\ell|_\rho^\epsilon$ and $g_{\ell'}|_\rho^\epsilon$ are supported in the interval $[0, \epsilon]$.

In no term of the finite sum in the last line of (5.66) are both J and K empty, and so (5.66) is just

$$\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} A_\ell \prod_{\ell'=1}^{k} B_{\ell'}\, dt + O_{f,g}(\epsilon).$$

This demonstrates (5.61).

(5.62) is proven in the same way, substituting Proposition 5.4.7 for Fujii's upper bound, and Corollary 5.7.4 for Corollary 5.6.7. □

A proof in the opposite direction is less technically demanding.

Proof of Theorem 5.2.4: (5.12) *implies the GUE Conjecture.* Assume that (5.12) holds for all admissible functions. Let $f_1, ..., f_n$ be arbitrary $C_c^2(\mathbb{R})$ functions.

From (5.12), we have for any $\{\varepsilon_1, ..., \varepsilon_n\} \in \{-1, 1\}^n$,

$$\lim_{T\to\infty} \frac{(-1)^n}{\log^n T} \int_{\mathbb{R}^n} f_1\Big(\tfrac{\varepsilon_1 x_1}{\log T}\Big) \cdots f_n\Big(\tfrac{\varepsilon_n x_n}{\log T}\Big) \big(1 - T|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n|\big)_+ \, d^n z(x)$$
$$= \lim_{N\to\infty} \frac{1}{N^n} \sum_{r\in\mathbb{N}_+^n} f_1\Big(\tfrac{\varepsilon_1 r_1}{N}\Big) \cdots f_n\Big(\tfrac{\varepsilon_n r_n}{N}\Big) \int_{U(N)} \prod_{\ell=1}^n \mathrm{Tr}(u^{\epsilon_n r_n})\, du.$$

But by the explicit formula,

$$\sum_{\varepsilon\in\{-1,1\}^n} \frac{(-1)^n}{\log^n T} \int_{\mathbb{R}^n} f_1\Big(\tfrac{\varepsilon_1 x_1}{\log T}\Big) \cdots f_n\Big(\tfrac{\varepsilon_n x_n}{\log T}\Big) \big(1 - T|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n|\big)_+ \, d^n z(x)$$
$$= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^n \tilde{G}_T(\hat{f}_\ell, t)\, dt.$$

From Stirling's formula, in particular (5.36), this is equal as $T \to \infty$ to

$$\int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^n \Big(G_T(\hat{f}_\ell, t) - \frac{\log(|t|+2)}{\log T} \int_{\mathbb{R}} \hat{f}_\ell(\alpha)\, d\alpha + O_{f_\ell}\Big(\tfrac{1}{\log T}\Big)\Big)\, dt$$
$$= \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^n \Big(G_T(\hat{f}_\ell, t) - \frac{\log(|t|+2)}{\log T} \int_{\mathbb{R}} \hat{f}_\ell(\alpha)\, d\alpha\Big)\, dt + O_f\Big(\tfrac{1}{\log T}\Big), \quad (5.67)$$

the last line following from Lemma 5.6.2 (the Fujii bound) in the same manner we have used it previously. Because we have

$$\frac{\log(|t|+2)}{\log T} = 1 + O\Big(\frac{|\log\big(\frac{|t|}{T} + \frac{2}{T}\big)|}{\log T}\Big)$$

we may use Lemma 5.6.2 once again so see that the expression in (5.67) is equal to

$$\mathop{\mathbf{E}}_{Z_t(\sigma_1)} \prod_{\ell=1}^n \Big(\sum_i \hat{f}_\ell(\xi_i) - \int_{\mathbb{R}} \hat{f}_\ell(\alpha)\, d\alpha\Big) + O_f\Big(\tfrac{1}{\log T}\Big).$$

On the other hand,

$$\sum_{\varepsilon\in\{-1,1\}^n} \frac{1}{N^n} \sum_{r\in\mathbb{N}_+^n} f_1\Big(\tfrac{\varepsilon_1 r_1}{N}\Big)\cdots f_n\Big(\tfrac{\varepsilon_n r_n}{N}\Big) \int_{U(N)} \prod_{\ell=1}^n \mathrm{Tr}(u^{\epsilon_n r_n})\, du$$
$$= \mathop{\mathbf{E}}_{\mathcal{S}_N'} \prod_{\ell=1}^n \Big(\sum_i \hat{f}_\ell(x_i) - \int_{\mathbb{R}} \hat{f}_\ell(\alpha)\, d\alpha \Big)$$

by equation (5.32).

Thus, it inductively follows (by removing lower order correlations) that for $f_1, ..., f_n$ arbitrary $C_c^2(\mathbb{R})$ functions

$$\lim_{T\to\infty} \mathop{\mathbf{E}}_{Z_T(\sigma)} \sum_{\substack{j_1,\ldots,j_n \\ \text{distinct}}} \hat{f}_1(\xi_{j_1})\cdots \hat{f}_n(\xi_{j_n}) = \lim_{N\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}_N'} \sum_{\substack{j_1,\ldots,j_n \\ \text{distinct}}} \hat{f}_1(x_1)\cdots \hat{f}_n(x_n)$$
$$= \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,\ldots,j_n \\ \text{distinct}}} \hat{f}_1(x_1)\cdots \hat{f}_n(x_n). \tag{5.68}$$

Yet, any such $\hat{f}_1 \otimes \cdots \otimes \hat{f}_n$ will decay quadratically in each variable, and if $\mathcal{A}_n$ is the linear span of such functions:

$$\mathcal{A}_n := \mathrm{span}\{\eta : \mathbb{R}^n \to \mathbb{R} : \eta = \hat{f}_1 \otimes \cdots \otimes \hat{f}_n,\ f_1, ..., f_n \in C_c^2(\mathbb{R})\},$$

it is easy to see that for any $\rho \in C_c(\mathbb{R}^k)$ and any $\epsilon > 0$, there exists $\eta \in \mathcal{A}_n$ so that for all x,

$$\big|\rho(x) - \eta(x)\big| \le \frac{\epsilon}{(1+|x_1|^{3/2})\cdots(1+|x_n|^{3/2})}. \tag{5.69}$$

For, using (5.53), for any $\eta \in C_c(\mathbb{R})$ and $\epsilon > 0$, there exists $f \in C_c^2(\mathbb{R})$ such that for all x

$$|\eta(x) - \hat{f}(x)| \le \frac{\epsilon}{1+|x|^{3/2}}.$$

And quite generally if $\mathcal{B}$ is dense in $C_c(\mathbb{R})$, then the linear span of functions

$\{\eta_1 \otimes \cdots \otimes \eta_k : \eta_j \in \mathcal{B} \quad \forall j\}$ is dense in $C_c(\mathbb{R}^k)$. In the case that $\mathcal{B} = \{(1 + |x|^{3/2})\hat{f}(x) : f \in C_c^2(\mathbb{R})\}$ this yields (5.69).

Therefore, because $\mathcal{A}_n$ is in this sense sufficiently dense, by Proposition 5.8.4, (5.68) is sufficient to deduce GUE(σ), and therefore the GUE Conjecture proper. $\square$

Note that in the above proofs to pass from (5.12) to the GUE Conjecture and back, we did not require knowledge of correlation functions at all levels, but rather for any n, knowing the first n correlation functions of the zeta zeros was sufficient to pass to (5.12) for all $j + k \leq n$, and vice versa.

Because we know the $n = 1$ case of the GUE Conjecture is true unconditionally, we have as a corollary to Theorem 5.2.4 an arithmetic statement that is equivalent to the pair correlation conjecture.

Corollary 5.10.1. *The case $n = 2$ of the GUE Conjecture is equivalent to the claim that for all admissible $f, g : \mathbb{R} \to \mathbb{R}$,*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f\Big(\tfrac{x}{\log T}\Big) g\Big(\tfrac{y}{\log T}\Big) \upsilon_T(x,y)\, dz(x) dz(y) \sim \log^2 T \int_0^\infty f(\alpha) g(\alpha)(\alpha \wedge 1)\, d\alpha.$$

On the right hand side,

$$\int_0^\infty f(\alpha) g(\alpha)(\alpha \wedge 1)\, d\alpha = \frac{1}{N^2} \sum_{r,s=1}^{\infty} f\Big(\tfrac{r}{N}\Big) g\Big(\tfrac{s}{N}\Big) \int_{U(N)} \mathrm{Tr}(u^r)\overline{\mathrm{Tr}(u^s)}\, du,$$

which can be seen from either the Diaconis-Shashahani type identity (5.17) or the explicit calculation of correlation functions for eigenvalues of $U(N)$, Theorem 5.4.5. The latter approach is somewhat more tedious, involving as it does an inclusion-exclusion argument, but for us it will generalize.

We have outlined in section 5.1 how Corollary 5.10.1 reduces to Theorem 5.1.2, a weighted estimate for the variance of primes in short intervals, with an algebraically nice form. We record below the analogues of Corollary 5.10.1 for the cases $n = 3, 4$, but caution that there is no apparent way to put the resulting statements in a form that is of comparable simplicity to Theorem 5.1.2.

On the other hand, we do derive a generalization of Theorem 5.1.2 which is algebraically simple in section 5.12. This is the covariance of almost primes with higher order von Mangoldt weights. The estimates we consider there fall short however of implying in full that any n-level densities for the zeta zeros follow the GUE pattern, beyond $n = 2$.

Corollary 5.10.2 (The three point correlation conjecture)**.** *Assume the pair correlation conjecture, that* (5.1) *holds for* $n = 2$ *for all fixed* η*.*

Then the statement that (5.1) *holds for* $n = 3$ *for all* η *is equivalent to the statement that for all admissible* f_1, f_2, g

$$\int_{\mathbb{R}}\int_{\mathbb{R}^2} f_1\Big(\tfrac{x_1}{\log T}\Big) f_2\Big(\tfrac{x_2}{\log T}\Big) g\Big(\tfrac{y}{\log T}\Big) \upsilon_T(x_1+x_2, y)\, dz(x_1)dz(x_2)\, dz(y) \tag{5.70}$$
$$\sim \log^3 T \int_{\mathbb{R}^2_+} f_1(\alpha_1)f_2(\alpha_2)g(\alpha_1+\alpha_2)\big[(\alpha_1 \wedge 1) + (\alpha_2 \wedge 1) - ((\alpha_1+\alpha_2)\wedge 1)\big]\, d\alpha_1 d\alpha_2.$$

Corollary 5.10.3 (The four point correlation conjecture)**.** *Assume the pair correlation conjecture and the three point correlation conjectures, that is, that* (5.1) *holds for* $n = 2$ *and* 3 *for all fixed* η*.*

Then the statement that (5.1) *holds for* $n = 4$ *for all* η *is equivalent to the claim that both:*

(i) For all admissible f_1, f_2, f_3, g,

$$\int_{\mathbb{R}}\int_{\mathbb{R}^3} f_1\Big(\tfrac{x_1}{\log T}\Big) f_2\Big(\tfrac{x_2}{\log T}\Big) f_3\Big(\tfrac{x_3}{\log T}\Big) g\Big(\tfrac{y}{\log T}\Big) \upsilon_T(x_1+x_2+x_3, y)\, dz(x_1)dz(x_2)d(x_3)\, dz(y)$$

$$\sim \log^3 T \int_{\mathbb{R}_+^3} f_1(\alpha_1)f_2(\alpha_2)f_3(\alpha_3)g(\alpha_1+\alpha_2+\alpha_3)\big[(\alpha_1\wedge 1)+(\alpha_2\wedge 1)+(\alpha_3\wedge 1) \tag{5.71}$$

$$-((\alpha_1+\alpha_2)\wedge 1)-((\alpha_1+\alpha_3)\wedge 1)-((\alpha_2+\alpha_3)\wedge 1)+((\alpha_1+\alpha_2+\alpha_3)\wedge 1)\big]\, d\alpha_1 d\alpha_2 d\alpha_3.$$

and

(ii) For all admissible f_1, f_2, g_1, g_2,

$$\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f_1\Big(\tfrac{x_1}{\log T}\Big) f_2\Big(\tfrac{x_2}{\log T}\Big) g_1\Big(\tfrac{y_1}{\log T}\Big) g\Big(\tfrac{y_2}{\log T}\Big) \upsilon_T(x_1+x_2, y_1+y_2)\, dz(x_1)dz(x_2)\, dz(y_1)dz(y_2)$$

$$\sim \log^4 T \int_{\mathbb{R}_+^4} f_1(\alpha_1)f_2(\alpha_2)g_1(\beta_1)g_2(\beta_2)\Big[\delta(\alpha_1+\alpha_2-\beta_1-\beta_2)\Big(1+(1-\alpha_1)_+ +(1-\alpha_2)_+$$

$$+(1-\beta_1)_+ +(1-\beta_2)_+ -(1-\alpha_1-\alpha_2)_+ -(1-|\alpha_1-\beta_1|)_+ -(1-|\alpha_1-\beta_2|)_+$$

$$-2(1-\alpha_1\wedge\alpha_2\wedge\beta_1\wedge\beta_2)_+\Big)+\delta(\alpha_1-\beta_1)\delta(\alpha_2-\beta_2)(\alpha_1\wedge 1)(\alpha_2\wedge 1) \tag{5.72}$$

$$+\delta(\alpha_1-\beta_2)\delta(\alpha_2-\beta_1)(\alpha_1\wedge 1)(\alpha_2\wedge 1)\Big]\, d\alpha_1 d\alpha_2 d\beta_1 d\beta_2.$$

One can of course continue on in this way for even higher correlations.

5.11 Zeros and the zeta function

We turn to upper bounds for moments of ζ'/ζ.

Proof of Theorem 5.2.1. We recall Theorem 5.3.2, that for $f(x) = \exp(-Ax)$, and

$\log T \geq 2A$,

$$\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A}{\log T}-it\Big) = \frac{-1}{\log T}\int_{\to-\infty}^{\to\infty} f\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x)$$
$$= O_f\Big(\tfrac{1}{\log T}\Big) + \frac{-1}{\log T}\int_0^{\to\infty} f\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x).$$

There exists ϵ_T close enough to 0 so that this expression is

$$O_f\Big(\tfrac{1}{\log T}\Big) + \frac{-1}{\log T}\int_0^{\to\infty} f|_{\epsilon_T}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x)$$
$$= O_f\Big(\tfrac{1}{\log T}\Big) + \lim_{R\to\infty}\frac{-1}{\log T}\int_{\mathbb{R}} f|_{\epsilon_T}^{R}\Big(\tfrac{x}{\log T}\Big)e^{ixt}\,dz(x),$$

the second line being an easy exercise. Using Proposition 5.4.9, this is

$$O_f\Big(\tfrac{1}{\log T}\Big) + \lim_{R\infty}\widetilde{G}_T\big((f|_{\epsilon_T}^{R})\hat{\ },t\big).$$

Because

$$\lim_{R\to\infty}\sup_{R'>R}\|f|_R^{R'}\|_{L^1(\mathbb{R})} = 0$$
$$\lim_{R\to\infty}\sup_{R'>R}\|(f|_R^{R'})''\|_{L^1(\mathbb{R})} = 0$$

we have from (5.37) and (5.53) that for any $\delta > 0$, there exists R_δ so that

$$\lim_{R\to\infty}\big|\widetilde{G}_T\big((|_{R_\delta}^{R})\hat{\ },t\big)\big| \leq \delta\log(|t|+2).$$

In particular, setting $\delta = 1/T$, we see that

$$\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A}{\log T}-it\Big) = \widetilde{G}_T\big((f|_{\epsilon_T}^{R_{1/T}})\hat{\ },t\big) + O_f\Big(\tfrac{1}{\log T}\Big) + O\Big(\tfrac{\log(|t|+2)}{T}\Big)$$
$$= \widetilde{G}_T\big((f_{\epsilon_T}^{1/k})\hat{\ },t\big) + \widetilde{G}_T\big((f|_{1/k}^{R_{1/T}})\hat{\ },t\big) + O_f\Big(\tfrac{1}{\log T}\Big) + O\Big(\tfrac{\log(|t|+2)}{T}\Big),$$

so from Minkowski's inequality,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\Big|\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{A}{\log T}-it\Big)\Big|^k\,dt \lesssim_k \int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((f_{\epsilon_T}^{1/k})\hat{\ },t\big)\big|^k\,dt \qquad (5.73)$$
$$+\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((f|_{1/k}^{R_{1/T}})\hat{\ },t\big)\big|^k\,dt + o_f(1),$$

where in this case we define $\sigma(t)$ to be $\mathbf{1}_{[1,2]}(t)$ (though one could certainly use other weights). From Lemma 5.6.4 (or Lemma 5.6.7), because $f_{\epsilon_T}^{1/k}$ is supported in $[0,1/k]$,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((f_{\epsilon_T}^{1/k})\hat{\ },t\big)\big|^k\,dt = O_{f,k}(1).$$

Likewise, because

$$\|f_{1/k}^{R_{1/t}}\|_{L^1(\mathbb{R})} = O_{f,k}(1)$$

$$\|(f_{1/k}^{R_{1/t}})''\|_{L^1(\mathbb{R})} = O_{f,k}(1)$$

as $T\to\infty$, Lemma 5.9.1 implies that

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((f|_{1/k}^{R_{1/T}})\hat{\ },t\big)\big|^k\,dt = O_{f,k}(1)$$

Substituting these bounds in (5.73) gives the theorem. □

Remark: This analysis can be reproduced in a more classical way by using the famous Selberg mollification formula:

$$\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty}\frac{\Lambda_x(n)}{n^s} + \frac{1}{\log x}\sum_{\gamma}\frac{x^{1/2+i\gamma-s}-x^{2(1/2+i\gamma-s)}}{(1/2+i\gamma-s)^2}$$
$$+\frac{x^{2(1-s)}-x^{1-s}}{(1-s)^2\log x} + \frac{1}{\log x}\sum_{q=1}^{\infty}\frac{x^{-2q-s}-x^{-2(2q+s)}}{(2q+s)^2},$$

where

$$\Lambda_x(n) = \begin{cases} \Lambda(n) & \text{for } 1 \le n \le x \\ \Lambda(n)\frac{\log(x^2/n)}{\log x} & \text{for } x \le n \le x^2 \\ 0 & \text{otherwise} \end{cases} .$$

Letting $x = \frac{1}{2k}\log T$ and $s = \frac{1}{2} + \frac{A}{\log T} + it$, we can produce an upper bound for

$$\frac{1}{T}\int_T^{2T} \Big|\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A}{\log T} + it\Big)\Big|^k \, dt$$

in the same way as above, with

$$\frac{1}{\log T}\bigg(\sum \frac{\Lambda_x(n)}{n^s}\bigg) \quad \text{playing the role of} \quad \widetilde{G}_T\big((f_{\epsilon_T}^{1/k})\hat{\ }, t\big),$$

and

$$\frac{1}{\log T}\frac{1}{\log x}\sum_\gamma \frac{x^{1/2+i\gamma-s} - x^{2(1/2+i\gamma-s)}}{(1/2+i\gamma-s)^2} \quad \text{playing the role of} \quad \widetilde{G}_T\big((f|_{1/k}^{R_{1/T}})\hat{\ }, t\big)$$

and the remaining terms of order $O(1)$. Indeed, the latter sum over γ can be bounded from above by $G_T(\eta, t)$ for some η of quadratic decay, moments of which can be bounded with Fujii's theorem.

We can now turn to a proof of Theorem 5.2.2, our restatement of the GUE Conjecture in terms of the mixed moments of the zeta function. We shall demonstrate the asymptotic equality of (5.8) and (5.9) on the assumption of the GUE Conjecture. In fact, we show more:

Theorem 5.11.1. *For $1 \le j \le j$ and $1 \le \ell' \le k$, define*

$$f_\ell(x) := P_\ell(x)e^{-A_\ell x}$$

$$g_{\ell'}(x) := Q_{\ell'}(x)e^{-B_{\ell'} x}$$

where P_ℓ and $Q_{\ell'}$ are polynomials and $A_\ell, B_{\ell'}$ are constants with $\Re A_\ell, \Re B_{\ell'} > 0$. Let $\sigma(t)$ be either the function $\mathbf{1}_{[1,2]}$ or $\frac{1}{2\pi}(\sin(t/2)/(t/2))^2$ as in (5.52). Then the GUE Conjecture implies that

$$\lim_{T\to\infty} \frac{1}{\log^{j+k} T} \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \left(\prod_{\ell=1}^{k} \int_{\to\infty}^{\to\infty} f_\ell\big(\tfrac{x_\ell}{\log T}\big) e^{-ix_\ell t}\, dz(x_\ell) \prod_{\ell=1}^{k} \overline{\int_{\to-\infty}^{\to\infty} g_{\ell'}\big(\tfrac{x_{\ell'}}{\log T}\big) e^{-ix_{\ell'} t}\, dz(x_{\ell'})} \right) dt$$
$$= \lim_{N\to\infty} \Theta_N(f_1 \otimes \cdots \otimes f_j; \overline{g}_1 \otimes \cdots \otimes \overline{g}_k). \tag{5.74}$$

Remark: It follows from partial integration as before that the integrals

$$\int_{\to-\infty}^{\to\infty} f_\ell\left(\tfrac{x}{\log T}\right) e^{-ixt}\, dz(x)$$

converge.

In the case that $f_\ell = g_{\ell'} = 1$, we see from Theorem 5.3.2 that the left hand side of equation (5.74) is exactly $(-1)^{j+k}$ times the expression (5.8) in Theorem 5.2.2, while the right hand side is

$$\lim_{N\to\infty} \frac{1}{N^{j+k}} \sum_{r\in\mathbb{N}_+^j} \sum_{s\in\mathbb{N}_+^k} \left(\prod_{\ell=1}^{j} e^{-A_\ell r_\ell/N} \prod_{\ell'=1}^{k} \overline{e^{-B_{\ell'} s_{\ell'}/N}} \right)$$
$$\times \left(\int_{U(N)} \prod_{\ell=1}^{j} (-\mathrm{Tr}\, u^{r_\ell}) \prod_{\ell'=1}^{k} \overline{(-\mathrm{Tr}\, u^{s_{\ell'}})}\, du \right)$$
$$= \lim_{N\to\infty} (-1)^{j+k} \int_{U(N)} \prod_{\ell=1}^{j} \frac{Z'}{Z}\left(\frac{A_\ell}{N}\right) \prod_{\ell'=1}^{k} \overline{\frac{Z'}{Z}\left(\frac{B_{\ell'}}{N}\right)}\, du,$$

where we can swap the order of integration and summation because for fixed N, $\mathrm{Tr}(u^r)$ is bounded as $r \to \infty$. This is, of course $(-1)^{j+k}$ times expression (5.9).

More generally, if $f(x) = P(x)e^{-Ax}$ for a polynomial P, note that

$$\sum_{r=1}^{\infty} f\left(\tfrac{r}{N}\right) \mathrm{Tr}(u^r) = P\left(\frac{d}{dA}\right) \frac{Z'}{Z}\left(\frac{A}{N}\right),$$

and likewise for the zeta function,

$$\int_{\to -\infty}^{\to\infty} f\Big(\tfrac{x}{\log T}\Big) e^{-ixt}\, dz(x) = P\Big(\frac{d}{dA}\Big)\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A}{\log T} + it\Big).$$

We will use this more general framework when evaluating the covariance of almost primes.

Proof of Lemma 5.11.1. Let $\epsilon > 0$. We have as above that there exists ϵ_T so that

$$\frac{-1}{\log T}\int_{\to -\infty}^{\to\infty} f_\ell\Big(\tfrac{x}{\log T}\Big) e^{-ixt}\, dz(x) = O_f\Big(\tfrac{1}{\log T}\Big) + \widetilde{G}_T\big((f_\ell|_{\epsilon_T}^{\epsilon})\check{}, t\big) + \lim_{R\to\infty} \widetilde{G}_T\big((f_\ell|_{\epsilon}^{R})\check{}, t\big),$$

and likewise for $g_{\ell'}$. As before, there is some $R_{1/T}$ so that

$$\widetilde{G}_T\big((f_\ell|_{\epsilon}^{R_{1/T}})\check{}, t\big) = \lim_{R\to\infty} \widetilde{G}_T\big((f_\ell|_{\epsilon}^{R})\check{}, t\big) + O\Big(\tfrac{\log(|t|+2)}{T}\Big),$$

and one can find R', depending on ϵ, but not on T, so that as $T \to \infty$,

$$\begin{aligned} &\|f_\ell|_{R'}^{R_{1/T}}\|_{L^1} < \epsilon \\ &\|(f_\ell|_{R'}^{R_{1/T}})''\|_{L^1} < \epsilon, \end{aligned}$$

and therefore, using Lemma 5.9.2,

$$\int_{\mathbb{R}} \frac{\sigma(t/T)}{T}\big|\widetilde{G}_T\big((f_\ell|_{R'}^{R_{1/T}})\check{}, t\big)\big|^{j+k} \lesssim \epsilon^{j+k}$$

for sufficiently large T.

Therefore we may decompose our integral against dz:

$$\begin{aligned} \frac{-1}{\log T}\int_{\to -\infty}^{\to\infty} f_\ell\Big(\tfrac{x}{\log T}\Big) e^{-ixt}\, dz(x) =& O_f\Big(\tfrac{1}{\log T}\Big) + \widetilde{G}_T\big((f_\ell|_{\epsilon_T}^{\epsilon})\check{}, t\big) + \widetilde{G}_T\big((f_\ell|_{\epsilon}^{R'})\check{}, t\big) \\ &+ \widetilde{G}_T\big((f_\ell|_{R'}^{R_{1/T}})\check{}, t\big) + O\Big(\tfrac{\log(|t|+2)}{T}\Big), \end{aligned}$$

and likewise

$$\overline{\frac{-1}{\log T}\int_{\to-\infty}^{\to\infty} g_{\ell'}\Big(\tfrac{x}{\log T}\Big)e^{-ixt}\,dz(x)} = O_g\Big(\tfrac{1}{\log T}\Big) + \widetilde{G}_T\big((\overline{g}_{\ell'}|_{\epsilon_T}^{\epsilon})\hat{}, t\big) + \widetilde{G}_T\big((\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}, t\big)$$
$$+ \widetilde{G}_T\big((\overline{g}_{\ell'}|_{R'}^{R_{1/T}})\hat{}, t\big) + O\Big(\tfrac{\log(|t|+2)}{T}\Big).$$

Here the terms

$$\widetilde{G}_T\big((f_\ell|_{\epsilon}^{R'})\check{}, t\big) \text{ and } \widetilde{G}_T\big((\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}, t\big)$$

will be the main contributions. Note that in the second equation we have taken a Fourier transform $(\overline{g}\cdots)\hat{}$, as opposed to the inverse Fourier transform $(f\cdots)\check{}$ in the first equation; the reader should check that this is indeed what arises from conjugating the left hand side.

Applying Hölder's inequality to these decompositions as in section 5.10, as $T\to\infty$,

$$\frac{(-1)^{j+k}}{\log^{j+k} T}\int_{\mathbb{R}}\frac{\sigma(t/T)}{T}\left(\prod_{\ell=1}^{k}\int_{\to\infty}^{\to\infty} f_\ell\big(\tfrac{x_\ell}{\log T}\big)e^{-ix_\ell t}\,dz(x_\ell)\prod_{\ell=1}^{k}\overline{\int_{\to-\infty}^{\to\infty} g_{\ell'}\big(\tfrac{x_{\ell'}}{\log T}\big)e^{-ix_{\ell'}t}\,dz(x_{\ell'})}\right)dt$$
$$= \int_{\mathbb{R}}\frac{\sigma(t/T)}{T}\prod_{\ell=1}^{j}\widetilde{G}_T\big((f_\ell|_{\epsilon}^{R'})\check{}, t\big)\prod_{\ell'=1}^{k}\widetilde{G}_T\big((\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}, t\big)\,dt + O_{f,g}(\epsilon) + o_{f,g}(1). \quad (5.75)$$

But from the GUE Conjecture (which implies GUE(σ) for either choice of σ),

$$\int_{\mathbb{R}}\frac{\sigma(t/T)}{T}\prod_{\ell=1}^{j}\widetilde{G}_T\big((f_\ell|_{\epsilon}^{R'})\check{}, t\big)\prod_{\ell'=1}^{k}\widetilde{G}_T\big((\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}, t\big)\,dt$$
$$\sim \mathop{\mathbf{E}}_{\mathcal{S}}\prod_{\ell=1}^{j}\sum_i (f_\ell|_{\epsilon}^{R'})\check{}(\xi_i)\prod_{\ell'=1}^{k}\sum_i(\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}(\xi_i). \quad (5.76)$$

Here we have used that $f_\ell|_{\epsilon}^{R'}$ and $\overline{g}_{\ell'}|_{\epsilon}^{R'}$ are each smooth, implying that $(f_\ell|_{\epsilon}^{R'})\check{}$ and $(\overline{g}_{\ell'}|_{\epsilon}^{R'})\hat{}$ are guaranteed to have (much faster than) quadratic decay, so that Proposition 5.8.4 applies.

In turn, the right hand side of (5.76) is the limit as $N \to \infty$ of

$$\begin{aligned}
&\mathbf{E}_{\mathcal{S}'_N} \prod_{\ell=1}^{j} \sum_i (f_\ell|_\epsilon^{R'})\check{}(x_i) \prod_{\ell'=1}^{k} \sum_i (\overline{g}_{\ell'}|_\epsilon^{R'})\hat{}(x_i) \\
&= \mathbf{E}_{\mathcal{S}'_N} \prod_{\ell=1}^{j} \sum_i (f_\ell|_{\epsilon_N})\check{}(x_i) \prod_{\ell'=1}^{k} \sum_i (\overline{g}_{\ell'}|_{\epsilon_N})\hat{}(x_i) + O_{f,g}(\epsilon) + o_{f,g}(1),
\end{aligned}$$

for $\epsilon_N = 1/2N$, using the same estimates as above with the point processes $\mathcal{S}'_N$ in place of $Z_T(\sigma_2)$. But by applying Poisson summation (as in equation (5.32)),

$$\begin{aligned}
&\mathbf{E}_{\mathcal{S}'_N} \prod_{\ell=1}^{j} \sum_i (f_\ell|_{\epsilon_N})\check{}(x_i) \prod_{\ell'=1}^{k} \sum_i (\overline{g}_{\ell'}|_{\epsilon_N})\hat{}(x_i) \\
&= \int_{U(N)} \prod_{\ell=1}^{j} \Big(\sum_{r_\ell \in \mathbb{Z}} \frac{1}{N} f_\ell|_{\epsilon_N} \big(\tfrac{r_\ell}{N}\big) \mathrm{Tr}(u^{r_\ell}) \Big) \prod_{\ell'=1}^{k} \Big(\sum_{s'_\ell \in \mathbb{Z}} \frac{1}{N} \overline{g}_{\ell'}|_{\epsilon_N} \big(\tfrac{-s_{\ell'}}{N}\big) \mathrm{Tr}(u^{s_{\ell'}}) \Big)\, du \\
&= \frac{1}{N^{j+k}} \int_{U(N)} \prod_{\ell=1}^{j} \Big(\sum_{r=1}^{\infty} f_\ell|_{\epsilon_N} \big(\tfrac{r}{N}\big) \mathrm{Tr}(u^{r}) \Big) \prod_{\ell'=1}^{k} \Big(\sum_{s'=1}^{\infty} \overline{g}_{\ell'}|_{\epsilon_N} \big(\tfrac{s'}{N}\big) \mathrm{Tr}(u^{-s'}) \Big)\, du.
\end{aligned}$$

Interchanging integration and summation is plainly justified, and we see that this is

$$(-1)^{j+k} \Theta_N(f \otimes \cdots \otimes f_j \,;\, \overline{g}_1 \otimes \cdots \otimes \overline{g}_k)$$

Because for any $\epsilon > 0$ as $N \to \infty$ this is within $O(\epsilon) + o(1)$ of the right hand side of (5.76), we see that the right hand limit of (5.74) exists. But in the same way, for any $\epsilon > 0$ as $T \to \infty$ the left hand side of (5.75) is within $O(\epsilon)+o(1)$ of the right hand side of (5.76), so that the left hand limit of (5.74) exists. Therefore the two limits in (5.74) are within $O(\epsilon)$ of each other for any ϵ and are thus equal. □

In the converse direction,

Proof of Theorem 5.2.2: The equivalence of (5.8) *and* (5.9) *implies the GUE Conjecture.* Naturally, our proof will bear a similarity to the proof above of the second part of Theorem 5.2.4. We use the formula that for $L > 0$, τ real, and T sufficiently large,

$$\begin{aligned}
\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{2\pi(L-i\tau)}{\log T}-it\Big) &= -\frac{1}{\log T}\int_{\to-\infty}^{\to\infty}\exp\Big(-\tfrac{2\pi(L-i\tau)}{\log T}x\Big)e^{ixt}\,dz(x)\\
&= -\frac{1}{\log T}\int_{\to-\infty}^{\to\infty}\exp\Big(-\tfrac{2\pi L}{\log T}|x|+\tfrac{i2\pi\tau}{\log T}x\Big)e^{ixt}\,dz(x)\\
&\quad -\frac{1}{\log T}\int_{-\infty}^{0}\exp\Big(-\tfrac{2\pi(L-i\tau)}{\log T}x\Big)e^{ixt}e^{x/2}\,dx\\
&\quad +\frac{1}{\log T}\int_{-\infty}^{0}\exp\Big(\tfrac{2\pi(L+i\tau)}{\log T}x\Big)e^{ixt}e^{x/2}\,dx\\
&= -\frac{1}{\log T}\int_{\to-\infty}^{\to\infty}\exp\Big(-\tfrac{2\pi L}{\log T}|x|+\tfrac{i2\pi\tau}{\log T}x\Big)e^{ixt}\,dz(x)+O\Big(\frac{1}{\log T}\Big).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{2\pi(L-i\tau)}{\log T}-it\Big)+\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2}+\tfrac{2\pi(L+i\tau)}{\log T}+it\Big)\\
&= -\frac{1}{\log T}\int_{\to-\infty}^{\to\infty}\Big[\exp\Big(-\frac{2\pi L|x|-i2\pi\tau x}{\log T}\Big)e^{ixt}+\exp\Big(-\frac{2\pi L|x|+i2\pi\tau x}{\log T}\Big)e^{-ixt}\Big]\,dz(x)\\
&\quad +O\Big(\frac{1}{\log T}\Big).
\end{aligned}$$

If

$$f(x) := \exp(-2\pi L|x| + i2\pi\tau x)$$

then

$$\hat{f}(\xi) = h_{L,\tau}(\xi) := \frac{1}{L}h\Big(\frac{\xi-\tau}{L}\Big)$$

where

$$h(\xi) := \frac{1}{\pi(1+\xi^2)}.$$

Hence we have by the explicit formula

$$-\frac{1}{\log T}\int_{\to-\infty}^{\to\infty}\Big[\exp\Big(-\frac{2\pi L|x|-i2\pi\tau x}{\log T}\Big)e^{ixt}+\exp\Big(-\frac{2\pi L|x|+i2\pi\tau x}{\log T}\Big)e^{-ixt}\Big]\,dz(x)$$
$$=\lim_{R\to\infty}\int_{-\infty}^{\infty}\hat{\alpha}_R * h_{L,\tau}\big(\tfrac{\log T}{2\pi}(\xi-\tau)\big)\,dS(\xi)$$
$$=\int_{-\infty}^{\infty}h_{L,\tau}\big(\tfrac{\log T}{2\pi}(\xi-t)\big)\,dS(\xi),$$

the last line following from dominated convergence. Therefore, for positive constants $L_1, ..., L_k$ and real $\tau_1, ..., \tau_k$,

$$\frac{1}{T}\int_T^{2T}\prod_{\ell=1}^k \tilde{G}_T(h_{L_\ell,\tau_\ell},t)\,dt$$
$$=\frac{1}{T}\int_T^{2T}\prod_{\ell=1}^k\Big(\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac12+\tfrac{2\pi(L_\ell-i\tau_\ell)}{\log T}-it\Big)+\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac12+\tfrac{2\pi(L_\ell+i\tau_\ell)}{\log T}+it\Big)+O\Big(\frac{1}{\log T}\Big)\Big)\,dt$$
$$=\frac{1}{T}\int_T^{2T}\prod_{\ell=1}^k\Big(\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac12+\tfrac{2\pi(L_\ell-i\tau_\ell)}{\log T}-it\Big)+\frac{1}{\log T}\frac{\zeta'}{\zeta}\Big(\tfrac12+\tfrac{2\pi(L_\ell+i\tau_\ell)}{\log T}+it\Big)\Big)\,dt+O\Big(\frac{1}{\log T}\Big),$$

the last line following from Theorem 5.2.1.

On the assumption of condition (5.8) and (5.9), this is asymptotic to

$$Q:=\lim_{N\to\infty}\frac{1}{N^k}\int_{U(N)}\prod_{\ell=1}^k\Big(\frac{Z'}{Z}\Big(-\frac{2\pi L_\ell-i2\pi\tau_\ell}{N}\Big)+\overline{\frac{Z'}{Z}\Big(-\frac{2\pi L_\ell-i2\pi\tau_\ell}{N}\Big)}\Big)\,du$$

Using Poisson summation as before in (5.32)

$$\frac{1}{N}\Big(\frac{Z'}{Z}\Big(-\frac{2\pi L-i2\pi\tau}{N}\Big)+\overline{\frac{Z'}{Z}\Big(-\frac{2\pi L-i2\pi\tau}{N}\Big)}\Big)$$
$$=\Big(\sum_{j=1}^N\sum_{r\in\mathbb{Z}}\frac{1}{N}\exp\Big(-2\pi L\tfrac{|r|}{N}+i2\pi\tau\tfrac{r}{N}\Big)e^{i2\pi r\theta_j}\Big)-1$$
$$=\Big(\sum_{j=1}^N\sum_{\nu\in\mathbb{Z}}h_{L,\tau}\big(N(\theta_j+\nu)\big)\Big)-1,$$

so

$$Q = \lim_{N\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}'_N} \prod_{\ell=1}^{k} \Big(\sum_i h_{L_\ell,\tau_\ell}(x_i) - 1 \Big)$$
$$= \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\sum_i h_{L_\ell,\tau_\ell}(x_i) - 1 \Big).$$

By Stirling's formula,

$$\tilde{G}_T(h_{L_\ell,\tau_\ell}, t) = \sum_\gamma h_{L,\tau}\big(\tfrac{\log T}{2\pi}(\gamma - t)\big) - \frac{\log t}{\log T} \int_{\mathbb{R}} h_{L,\tau}(x)\, dx + O_{L,\tau}\Big(\frac{1}{\log T}\Big)$$
$$= \sum_\gamma h_{L,\tau}\big(\tfrac{\log T}{2\pi}(\gamma - t)\big) - 1 + O_{L,\tau}\Big(\frac{1}{\log T}\Big).$$

Using Corollary 5.8.3, we thus have

$$\lim_{T\to\infty} \mathop{\mathbf{E}}_{Z_T} \prod_{\ell=1}^{k} \Big(\sum_i h_{L_\ell,\tau_\ell}(\xi_i) - 1 \Big) = \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \Big(\sum_i h_{L_\ell,\tau_\ell}(x_i) - 1 \Big),$$

for all k and all sets of positive constants $L_1, ..., L_k$, and real constants $\tau_1, ..., \tau_k$. Inductively removing lower order correlations from the above sums, we obtain for any such series of constants that

$$\mathop{\mathbf{E}}_{Z_T} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} h_{L_1,\tau_1}(\xi_{j_1}) \cdots h_{L_k,\tau_k}(\xi_{j_k}) \sim \mathop{\mathbf{E}}_{\mathcal{S}} \sum_{\substack{j_1,\dots,j_k \\ \text{distinct}}} h_{L_1,\tau_1}(x_{j_1}) \cdots h_{L_k,\tau_k}(x_{j_k}). \quad (5.77)$$

But it is clear that if $\mathcal{A}_k := \mathrm{span}\{h_{L_1,\tau_1} \otimes \cdots \otimes h_{L_k,\tau_k} :\ L_1, ..., L_k > 0,\ \tau_1, ..., \tau_k \in \mathbb{R}\}$, then $\mathcal{A}_k$ satisfies the conditions of Proposition 5.8.4, so that (5.77) implies the GUE Conjecture. This proves the theorem. □

5.12 Counts of almost primes

We turn at last to the proof of Theorem 5.2.5. It is easy to give a heuristic outline of the main ideas involved, although the rigorous proof that follows will entail substantial modifications.

We note that if $d\mathcal{P}(x)$ is the measure given by $d\psi(e^x)$, then it is easy to verify that

$$d\mathcal{P} * d\mathcal{P}(x) + x\, d\mathcal{P}(x) = d\psi_2(e^x)$$

In the same way, preceding entirely formally, if we define

$$dz_2(x) = dz * dz(x) + x\, dz(x),$$

this measure is given by the above measure $d\psi_2(e^x)$ minus a regular approximation:

$$dz_2(x) = d\tilde{\psi}_2(e^x),$$

where, recall, $\tilde{\psi}_2$ was defined in section 5.2 in equation (5.22). If we have proved Theorem 5.2.4 for more general f, g than seperable functions, we could say that

$$\lim_{T\to\infty} \Psi_T^{2,1}(f\,;\,g) = \lim_{N\to\infty} \Theta_N^{2,1}(f\,;\,g), \tag{5.78}$$

where for $\beta > 0$,

$$f(x_1, x_2) := \mathbf{1}_{[0,\beta)}(x_1 + x_2)$$

$$g(y) := \mathbf{1}_{[0,\beta)}(y).$$

The advantage of this particular choice of f is that it allows us to convolve in the variables x_1 and x_2, and the left hand side of (5.78) reduces to

$$\frac{1}{\log^3 T} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y) \upsilon_T(x,y)\, dz * dz(x)\, dz(y),$$

while the right hand side reduces to

$$\frac{1}{N^3}\sum_{r,s\in\mathbb{N}_+}\mathbf{1}_{[0,\beta)}\left(\tfrac{r}{N}\right)\mathbf{1}_{[0,\beta)}\left(\tfrac{s}{N}\right)\int_{U(N)}\left(\sum_{r_1=1}^{r-1}[-\mathrm{Tr}(u^{r-r_1})][-\mathrm{Tr}(u^{r_1})]\right)\overline{[-\mathrm{Tr}(u^s)]}\,du$$
$$\sim -\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y)\delta(x-y)(x-1)_+\,dx\,dy$$

by explicit computation with correlation functions. (cf. Theorem 5.10.2).

On the other hand, setting

$$f_1(x) := x\mathbf{1}_{[0,\beta)}(x),$$

$$g_1(y) := \mathbf{1}_{[0,\beta)}(y)$$

in the identity

$$\lim_{T\to\infty}\Psi_T^{1,1}(f_1\,;\,g_1) = \lim_{N\to\infty}\Theta_N^{1,1}(f_1\,;\,g_1)$$

we obtain

$$\lim_{T\to\infty}\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y)\upsilon_T(x,y)\,x\,dz(x)\,dz(y)$$
$$= \lim_{N\to]\infty}\sum_{r,s\in\mathbb{N}_+}\mathbf{1}_{[0,\beta)}\left(\tfrac{r}{N}\right)\mathbf{1}_{[0,\beta)}\left(\tfrac{s}{N}\right)\int_{U(N)}[-r\mathrm{Tr}(u^r)]\overline{[-\mathrm{Tr}(u^s)]}\,du.$$

This left hand limit as $N\to\infty$ tends to

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y)\delta(x,y)\,x(x\wedge 1)\,dx\,dy.$$

By adding the results, we obtain

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y)\upsilon_T(x,y)\,dz_2(x)\,dz(y) \sim \log^3 T\int_{\mathbb{R}}\mathbf{1}_{[0,\beta)}(x)(x\wedge 1)^2\,dx.$$

The right hand side above can also be written in the form

$$\log^3 T \Bigg(\lim_{N\to\infty} \frac{1}{N^3} \sum_{r,s} \mathbf{1}_{[0,\beta)}\Big(\tfrac{r}{N}\Big)\mathbf{1}_{[0,\beta)}\Big(\tfrac{s}{N}\Big) \int_{U(N)} H_2(r)\overline{H_1(s)}\, du \Bigg),$$

where $H_j(r)$ was defined in (5.26) and (5.27).

We can generalize this argument. Letting $dz_j(x) := d\tilde{\psi}_j(x)$, we obtain

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{\log^{j+k} T} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{[0,\beta)}(x)\mathbf{1}_{[0,\beta)}(y)\upsilon_T(x,y)\, dz_j(x)\, dz_k(y) \\ &= \lim_{N\to\infty} \sum_{r,s\in\mathbb{N}_+} \mathbf{1}_{[0,\beta)}\Big(\tfrac{r}{N}\Big)\mathbf{1}_{[0,\beta)}\Big(\tfrac{s}{N}\Big) \int_{U(N)} H_j(r)\overline{H_k(s)}\, du. \end{aligned}$$

It is from Lemma 5.2.7 that we can simplify the random matrix part of this identity. On the other hand, as in section 5.1, the arithmetic side is given by

$$\lim_{T\to\infty} \frac{T}{\log^{j+k} T} \int_1^{T^\beta} \Big(\tilde{\psi}_j\big(\tau+\tfrac{\tau}{T}\big) - \tilde{\psi}_j(\tau)\Big)\Big(\tilde{\psi}_k\big(\tau+\tfrac{\tau}{T}\big) - \tilde{\psi}_k(\tau)\Big)\,\frac{d\tau}{\tau^2}.$$

In this manner we have arrived at a (purely formal) derivation of Theorem 5.23. We are prevented from making this argument rigorous in the above form in that we have proved Theorem 5.2.4 only for functions f, g that are separable. In particular, we cannot approximate $f(x_1, x_2) = \mathbf{1}_{[0,\beta)}(x_1 + x_2)$ with a single separable function. Even to approximate this function with a linear combination of separable functions will not do, as we have proved no continuity properties for Ψ_T (an integral against signed measures). Equation (5.78) is therefore unjustified for the test functions we have made use of. We are therefore left with two routes to make the above sketch rigorous. In the first we could reprove Theorem 5.2.4 for test functions f and g that are not separable. This should certainly be possible, but will entail making the proof of the theorem more complicated. (The reader is encouraged to try to come up with a simple argument!) In the second possible approach, we make use of seperable functions that allow for convolution – these are exactly the exponential functions, and therefore the case we have considered in Theorems 5.2.2 and 5.11.1. This is the route we shall take. It involves the additional complication that exponential functions are not compactly supported, and this fact entails a sort of gymnastics that we must go through in the proof that follows.

Proof of Theorem 5.2.5. We note that Theorem 5.11.1 may be rewritten in the form that, conditioned on the GUE Conjecture

$$\lim_{T\to\infty} \frac{1}{\log^{j+k} T} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{j} P_\ell\Big(\frac{d}{dA_\ell}\Big)\Big(\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A_\ell}{\log T} + it\Big)\Big) \times \overline{\prod_{\ell'=1}^{k} Q_{\ell'}\Big(\frac{d}{dB_{\ell'}}\Big)\Big(\frac{\zeta'}{\zeta}\Big(\tfrac{1}{2} + \tfrac{B_{\ell'}}{\log T} + it\Big)\Big)}\, dt$$

$$= \lim_{N\to\infty} \frac{1}{N^{j+k}} \int_{U(N)} \prod_{\ell=1}^{j} P_\ell\Big(\frac{d}{dA_\ell}\Big)\Big(\frac{Z'}{Z}\Big(\tfrac{A_\ell}{N}\Big)\Big) \overline{\prod_{\ell'=1}^{k} Q_{\ell'}\Big(\frac{d}{dB_{\ell'}}\Big)\Big(\frac{Z'}{Z}\Big(\tfrac{B_{\ell'}}{N}\Big)\Big)}\, du,$$

for any polynomials $P_1,..,P_j, Q_1,...,Q_k$, where $\sigma_1(t) := \frac{1}{2\pi}\Big(\frac{\sin t/2}{t/2}\Big)^2$ as in (5.52). We will use this definition of σ_1 throughout this proof.

Because

$$\frac{\zeta^{(j)}}{\zeta}(s) = \Big(\frac{\zeta'}{\zeta} + \frac{d}{ds}\Big)\frac{\zeta^{(j-1)}}{\zeta}(s),$$

and likewise for $Z^{(j)}/Z$, we can inductively show from Theorem 5.11.1,

$$\lim_{T\to\infty} \frac{1}{\log^{J+K} T} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \prod_{\ell=1}^{J} P_\ell\Big(\frac{d}{dA_\ell}\Big)\Big(\frac{\zeta^{(j_\ell)}}{\zeta}\Big(\tfrac{1}{2} + \tfrac{A_\ell}{\log T} + it\Big)\Big) \times \overline{\prod_{\ell'=1}^{K} Q_{\ell'}\Big(\frac{d}{dB_{\ell'}}\Big)\Big(\frac{\zeta^{(k_{\ell'})}}{\zeta}\Big(\tfrac{1}{2} + \tfrac{B_{\ell'}}{\log T} + it\Big)\Big)}\, dt$$

$$= \lim_{N\to\infty} \frac{1}{N^{J+K}} \int_{U(N)} \prod_{\ell=1}^{J} P_\ell\Big(\frac{d}{dA_\ell}\Big)\Big(\frac{Z^{(j_\ell)}}{Z}\Big(\tfrac{A_\ell}{N}\Big)\Big) \overline{\prod_{\ell'=1}^{k} Q_{\ell'}\Big(\frac{d}{dB_{\ell'}}\Big)\Big(\frac{Z^{(k_{\ell'})}}{Z}\Big(\tfrac{B_{\ell'}}{N}\Big)\Big)}\, du,$$

We specialize to the case $J = K = 1$ and $A_1 = B_1$ real to obtain

$$\lim_{T\to\infty} \frac{1}{\log^{j+k} T} \int_{\mathbb{R}} \frac{\sigma_1(t/T)}{T} \left((-1)^j \frac{\zeta^{(j)}}{\zeta} \left(\tfrac{1}{2} + \tfrac{A}{\log T} + it \right) \right) \left((-1)^k \frac{\zeta^{(k)}}{\zeta} \left(\tfrac{1}{2} + \tfrac{A}{\log T} - it \right) \right) dt$$
$$= \lim_{N\to\infty} \frac{1}{N^{j+k}} \int_{U(N)} \left((-1)^j \frac{Z^{(j)}}{Z} \left(\tfrac{A}{N} \right) \right) \left((-1)^k \frac{Z^{(k)}}{Z} \left(\tfrac{A}{N} \right) \right) du \qquad (5.79)$$

This is in fact the identity we need, albeit in a somewhat veiled form. We now prove the theorem in four steps. In the first two steps, our development mimics the elegant approach in [34], which in turn draws from Selberg [64].

Step 1: We show for positive A and

$$f_\kappa(s) := \frac{e^{\kappa s} - 1}{s}$$

that for $\alpha := \frac{1}{2} + \frac{A}{\log T}$,

$$\int_0^\infty \frac{1}{r^{2\alpha}} \big(\tilde{\psi}_j(e^\kappa r) - \tilde{\psi}_j(r) \big) \big(\tilde{\psi}_k(e^\kappa r) - \tilde{\psi}_k(r) \big) \, dr \qquad (5.80)$$
$$= \int_{\mathbb{R}} \left((-1)^j \frac{\zeta^{(j)}}{\zeta}(\alpha + it) \right) \left((-1)^k \frac{\zeta^{(k)}}{\zeta}(\alpha - it) \right) \frac{|f_\kappa(\alpha + it)|^2}{2\pi} \, dt.$$

Step 2: We show that for κ_1 such that $e^{\kappa_1} - 1 = 1/T$, and α and f defined as in step 1,

$$\int_{\mathbb{R}} \left((-1)^j \frac{\zeta^{(j)}}{\zeta}(\alpha + it) \right) \left((-1)^k \frac{\zeta^{(k)}}{\zeta}(\alpha - it) \right) \frac{|f_\kappa(\alpha + it)|^2}{2\pi} \, dt$$
$$- \frac{1}{T} \int_{\mathbb{R}} \left((-1)^j \frac{\zeta^{(j)}}{\zeta}(\alpha + it) \right) \left((-1)^k \frac{\zeta^{(k)}}{\zeta}(\alpha - it) \right) \frac{\sigma(t/T)}{T} \, dt \qquad (5.81)$$
$$= O_A\left(\frac{\log^{2(j+k)+1} T}{T} \right),$$

Step 3: We combine these steps with (5.79) and the random matrix statistic

Lemma 5.2.7. We obtain that for any positive constant A

$$\int_0^\infty \frac{1}{r^{2+2A/\log T}} \tilde{\psi}_j\big(r; \tfrac{r}{T}\big)\tilde{\psi}_k\big(r; \tfrac{r}{T}\big)\, dr \sim \frac{\log^{j+k} T}{T} \frac{jk}{j+k-1} \int_0^\infty e^{-2Ay} (y \wedge 1)^{j+k-1}\, dy. \tag{5.82}$$

Step 4: We use a Tauberian argument to pass between the weights $e^{-\beta x}$ and $\mathbf{1}_{[0,\beta)}(x)$, thereby showing that (5.82) implies the covariance asymptotic (5.23) for any constant $\beta > 0$.

Having verified these steps, our proof will be complete.

Step 1: It follows from a standard argument in residue calculus (using the bound of Appendix A for $\zeta^{(j)}/\zeta$ at large heights) that when $x > 0$ is not an integer, for $\alpha \in (1/2, 1)$,

$$\tilde{\psi}_j(x) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \left((-1)^j \frac{\zeta^{(j)}}{\zeta}(s) \right) \frac{x^s}{s}\, ds.$$

Continuing the mimick the arguments [34], differencing the values when $x = e^{\tau+\kappa}$ and $x = e^\tau$ gives for almost all τ,

$$\frac{\tilde{\psi}_j(e^\kappa e^\tau) - \tilde{\psi}_j(e^\tau)}{e^{\tau\alpha}} = \frac{1}{2\pi} \int_{-\infty}^\infty (-1)^j \frac{\zeta^{(j)}}{\zeta}(\alpha + it) \Big(\frac{e^{\kappa(\alpha+it)} - 1}{\alpha + it} \Big) e^{it\tau}\, dt.$$

The right hand side is the inverse Fourier transform of $(-1)^j \zeta^{(j)}/\zeta(\alpha + i2\pi t) f_\kappa(\alpha + i2\pi t)$, while the left hand side is obviously real valued. It is moreover easy to see from the elementary estimates in Appendix A that the left hand side is square integrable in τ and so by an application of Plancherel

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\big(\tilde{\psi}_j(e^\kappa e^\tau) - \tilde{\psi}_j(e^\tau)\big)\big(\tilde{\psi}_k(e^\kappa e^\tau) - \tilde{\psi}_k(e^\tau)\big)}{e^{2\tau\alpha}}\, d\tau \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \left((-1)^j \frac{\zeta^{(j)}}{\zeta}(\alpha + it) \right) \left((-1)^k \frac{\zeta^{(k)}}{\zeta}(\alpha - it) \right) |f_\kappa(\alpha + it)|^2\, dt.
\end{aligned}$$

Making the change of variables $r = e^\tau$ and setting $\alpha = 1/2 + A/\log T$, this is

(5.80).

Step 2: We quote the estimate from [34], that for $\alpha \le 1$, $0 < \kappa \le 1$,

$$|f_\kappa(\alpha+it)|^2 - |f_\kappa(it)|^2 = O\Big(\tfrac{\kappa}{|t|^2} \wedge \kappa^2\Big).$$

Likewise, because $|f_\kappa(it)|^2 = \Big(\frac{\sin \kappa t/2}{t/2}\Big)^2$ and

$$\sin^2 x - \sin^2 y = O\big(|x-y| \wedge 1\big),$$

we have for real κ_1, κ_2 and $t \ge 1$

$$|f_{\kappa_1}(it)|^2 - |f_{\kappa_2}(it)|^2 = O\Big(\tfrac{|\kappa_1-\kappa_2|}{|t|} \wedge \tfrac{1}{|t|^2}\Big),$$

while for $t \le 1$ and $\kappa \le 1$, clearly

$$|f_\kappa(it)|^2 = O(\kappa^2).$$

We also make use of the basic pointwise bound proved in Appendix A,

$$\frac{\zeta^{(j)}}{\zeta}(\alpha+it) = O\Big(\frac{\log^j(|t|+2)}{(\alpha-1/2)^j}\Big),$$

for $|\sigma+it-1| \ge 1/4$, say.

We let κ_1 be such that $e^{\kappa_1}-1 = 1/T$ and $\kappa_2 = 1/T$. Note that $f_{\kappa_2}(it) = \sigma_1(t/T)/T^2$, and $\kappa_1-\kappa_2 = O(1/T^2)$. Hence, the left hand side of (5.81) has the bound

$$
\begin{aligned}
&\lesssim \int_{|t|\geq 1} \frac{\log^{j+k}(|t|+2)}{A^{j+k}/\log^{j+k} T} \Big(O\Big(\tfrac{1}{T|t^2|} \wedge \tfrac{1}{T^3}\Big) + O\Big(\tfrac{1}{T^2|t|} \wedge \tfrac{1}{|t|^2}\Big)\Big)\, dt + \int_{|t|<1} O\Big(\tfrac{1}{T^2}\Big)\, dt \\
&\lesssim_A \log^{j+k} T \Big(\int_1^T \frac{1}{T^3} + \int_T^\infty \frac{1}{T|t|^3} + \int_1^{T^2} \frac{1}{T^2|t|} + \int_{T^2}^\infty \frac{1}{|t|^2}\, dt \Big) + O\Big(\tfrac{1}{T^2}\Big) \\
&\lesssim_A \frac{\log^{2(j+k)+1} T}{T^2},
\end{aligned}
$$

as claimed.

Step 3: We turn to evaluating the right hand side of (5.79). We first demonstrate inductively that

$$
(-1)^j \frac{Z^{(j)}}{Z}(\beta) = \sum_{r=1}^\infty e^{-\beta r} H_j(r). \tag{5.83}
$$

For, the identity (5.7) says just that

$$
-\frac{Z'}{Z}(\beta) = \sum_{r=1}^\infty e^{-\beta r} H_1(r),
$$

while the fact that

$$
(-1)^j \frac{Z^{(j)}}{Z}(\beta) = \Big(-\frac{Z'}{Z}(\beta) - \frac{d}{d\beta}\Big) \Big((-1)^{j-1} \frac{Z^{(j-1)}}{Z}(\beta)\Big)
$$

and the definition (5.26) and (5.27) of $H_j(r)$ completes the induction to $j > 1$.

From Lemma 5.2.7 therefore,

$$
\int_{U(N)} \Big((-1)^j \frac{Z^{(j)}}{Z}(\beta)\Big)\Big((-1)^k \frac{Z^{(k)}}{Z}(\beta)\Big)\, du = \sum_{r=1}^\infty e^{-2\beta r} \sum_{\nu=1}^{r\wedge N} (\nu^j - (\nu-1)^j)(\nu^k - (\nu-1)^k).
$$

(The interchange of integration and summation is easy to justify, as for fixed N,

$H_j(r)$ is bounded.) Hence,

$$\frac{1}{N^{j+k}}\int_{U(N)}\left((-1)^j\frac{Z^{(j)}}{Z}\Big(\frac{A}{N}\Big)\right)\left((-1)^k\frac{Z^{(k)}}{Z}\Big(\frac{A}{N}\Big)\right)$$
$$=\frac{1}{N}\sum_{r=1}^{\infty}e^{-2Ar/N}\frac{1}{N^{j+k-1}}\Big(\tfrac{jk}{j+k-1}(r\wedge N)^{j+k-1}+O((r\wedge N)^{j+k-2})\Big)$$
$$\sim \tfrac{jk}{j+k-1}\int_0^\infty e^{-2Ay}(y\,\wedge\,1)^{j+k-1}\,dy,$$

since the sum over r is just a Riemann sum.

Using (5.80), (5.81), and (5.79) in succession, we arrive at (5.82).

Step 4: In the first place note that

$$\int_0^1\frac{1}{r^{2+2A/\log T}}\tilde\psi_j\big(r;\,\tfrac{r}{T}\big)\tilde\psi_k\big(r;\,\tfrac{r}{T}\big)\,dr=O\bigg(\int_0^1\frac{1}{r^{2A/\log T}}\frac{|\log r|^{j+k-2}}{T}\,dr\bigg)=O\Big(\frac{\log^{j+k-1}T}{T}\Big),$$

so (5.82) is equivalent to

$$\int_1^\infty\frac{1}{r^{2+2A/\log T}}\tilde\psi_j\big(r;\,\tfrac{r}{T}\big)\tilde\psi_k\big(r;\,\tfrac{r}{T}\big)\,dr\sim\frac{\log^{j+k}T}{T}\frac{jk}{j+k-1}\int_0^\infty e^{-2Ay}\,(y\wedge 1)^{j+k-1}\,dy. \tag{5.84}$$

With this simplification, we move on to the Tauberian part of the proof, assuming the truth of (5.84) and using it to demonstrate (5.23).

Note that for any continuous function ϕ of compact support, and for any $\epsilon>0$, there is a polynomial P so that

$$\big|P\big(\tfrac{1}{w}\big)-\phi(w)\big|\le\epsilon/w\qquad\text{for } w\ge 1. \tag{5.85}$$

For, note that by its compact support, $\phi(1/x)/x$ is continuous on the interval $[0,1]$ (defined by continuity to take the value 0 at $x=0$). Hence by Weierstrass's approximation theorem, there is some polynomial Q so that

$$|Q(x)-\phi(1/x)/x|<\epsilon\qquad\text{for all } x\in[0,1].$$

$P(x) := xQ(x)$ thus satisfies (5.85).

We use this to show that for any continuous f of compact support,

$$\int_1^\infty f\Big(\tfrac{\log r}{\log T}\Big)\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)\, dr \sim \frac{\log^{j+k} T}{T}\frac{jk}{j+k-1}\int_0^\infty f(y)(y\,\wedge\,1)^{j+k-1}\, dy. \tag{5.86}$$

For if $\phi(x) = f(\log x)$, then for P as in (5.85), using Cauchy-Schwarz,

$$\begin{aligned}
&\int_1^\infty \Big(f\Big(\tfrac{\log r}{\log T}\Big) - P\Big(\frac{1}{r^{1/\log T}}\Big)\Big)\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)\,\frac{dr}{r^2}\\
&\le \Big(\int_1^\infty \frac{\epsilon}{r^{1/\log T}}\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)^2\,\frac{dr}{r^2}\Big)^{1/2}\Big(\int_1^\infty \frac{\epsilon}{r^{1/\log T}}\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)^2\,\frac{dr}{r^2}\Big)^{1/2}\\
&\lesssim_{j,k} \frac{\log^{j+k} T}{T}\epsilon
\end{aligned}$$

by an applications of (5.84) in the case $j = k$.

On the other hand, from (5.84) again,

$$\begin{aligned}
&\int_1^\infty P\Big(\frac{1}{r^{1/\log T}}\Big)\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)\,\frac{dr}{r^2}\\
&\sim \frac{\log^{j+k} T}{T}\frac{jk}{j+k-1}\int_0^\infty P(e^{-y})\,(y\,\wedge\,1)^{j+k-1}\, dy\\
&= \frac{\log^{j+k} T}{T}\frac{jk}{j+k-1}\Big(\int_0^\infty f(y)\,(y\,\wedge\,1)^{j+k-1}\, dy + O(\epsilon)\Big).
\end{aligned}$$

As ϵ was arbitrary, this proves (5.86) for continuous and compactly supported f.

We want finally to show that (5.86) remains true when $f = \mathbf{1}_{[0,\beta)}$. This function is not continuous, but for any $\epsilon > 0$, plainly there exist continuous functions of compact support, f_1 and h, so that

$$f(x) = f_1(x) \text{ for } x \in [0, \beta)$$

$$|f(x) - f_1(x)| \le h(x) \text{ for all } x, \text{ and } \int_0^\infty h(x)\, dx < \epsilon.$$

Hence,

$$
\begin{aligned}
&\int_1^\infty \Big(f\Big(\tfrac{\log r}{\log T}\Big) - f_1\Big(\tfrac{\log r}{\log T}\Big)\Big)\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)\,\frac{dr}{r^2}\\
&\leq \bigg(\int_1^\infty h\Big(\tfrac{\log r}{\log T}\Big)\tilde{\psi}_j\big(r;\, \tfrac{r}{T}\big)^2\,\frac{dr}{r^2}\bigg)^{1/2}\bigg(\int_1^\infty h\Big(\tfrac{\log r}{\log T}\Big)\tilde{\psi}_k\big(r;\, \tfrac{r}{T}\big)^2\,\frac{dr}{r^2}\bigg)^{1/2}\\
&\lesssim_{j,k} \frac{\log^{j+k} T}{T}\epsilon.
\end{aligned}
$$

In the second line we used the positivity of $\tilde{\psi}_j^2$ and $\tilde{\psi}_k^2$ to replace $|f - f_1|$ by its majorant.

Clearly

$$\left|\int_0^\infty \big(f(y) - f_1(y)\big)\,(y \wedge 1)^{j+k-1}\,dy\right| < \epsilon$$

as well. Because ϵ is arbitrary, this proves that (5.86) is true even when $f = \mathbf{1}_{[0,\beta)}$, which is what we sought to show.

This completes step 4, and therefore the proof of Theorem 5.2.5. □

CHAPTER 6

Applying the Hardy-Littlewood conjectures

6.1 A conditional analysis

In this final chapter we show that conditioned on a uniform version of certain famous conjectures of Hardy & Littlewood [37], one may verify, with respect to a slightly wider range of test functions than those considered by Montgomery-Hejhal-Rudnick & Sarnak in Theorem 1.1.6, that the joint intensities of the zeta zeros tend to those of the sine-determinantal / GUE point process. We consider two, three, and four point joint intensities in particular, the analysis of the two point pair correlation case being due to Montgomery. What can be said rigorously – by assuming Hardy-Littlewood conjecture – beyond the range of Theorem 1.1.6 in these cases is the content of Theorems 6.1.4-6.1.6. We also record in Theorem 6.3.1 a conditional extension for *all* higher correlations, but with respect to a somewhat esoteric class of test functions. In principle our analysis could be extended to a more natural class of test functions, but seemingly only by passing through a series of somewhat indecent computations.

In our arguments we follow in outline the well-known heuristic account in Bogomolny & Keating [2], [3], though in making this implication rigorous we will deviate from it in some ways. The conjectures of Hardy & Littlewood that we will make use of, in order of generality are

Conjecture 6.1.1 (A uniform version of Hardy-Littlewood 1)**.** *For any* $\epsilon > 0$*,* *and all* $1 \le h \le x^{1-\epsilon}$

$$\sum_{n \le x} \Lambda(n)\Lambda(n+h) = \mathfrak{S}(h)x + O_\epsilon(x^{1/2+\epsilon}), \tag{6.1}$$

where

$$\begin{aligned} \mathfrak{S}(h) &:= \prod_p \mathbf{P}\big(n+h \not\equiv 0(p) \big| n \not\equiv 0(p)\big) \\ &= \prod_{p \mid h} \frac{1}{1-p^{-1}} \prod_{p \nmid h} \frac{1-(p-1)^{-1}}{1-p^{-1}}. \end{aligned}$$

Really it is the second line above that defines $\mathfrak{S}(k)$; the probabilistic notation in the first line may be taken intuitively.

More generally,

Conjecture 6.1.2 (A uniform version of Hardy-Littlewood 2)**.** *For any* $\epsilon > 0$*, and all* $1 \le a, h \le x^{1-\epsilon}$

$$\sum_{n \le x} \Lambda(n)\Lambda(an+h) = \big[(a,h)=1\big]\mathfrak{S}(ah)x + O_\epsilon(x^{1/2+\epsilon}). \tag{6.2}$$

Here we have used the notation that for a statement $\mathfrak{A}$, $\big[\mathfrak{A}\big]$ is 0 or 1 depending on whether $\mathfrak{A}$ is true or false. Note that

$$\prod_p \mathbf{P}\big(an+h \not\equiv 0(p) \big| n \not\equiv 0(p)\big) = \big[(a,h)=1\big]\mathfrak{S}(ah).$$

More generally still,

Conjecture 6.1.3 (A uniform version of Hardy-Littlewood 3)**.** *For any* $\epsilon > 0$*,* *and all* $1 \le a, b, h \le x^{1-\epsilon}$

$$\sum_{n \le x} \Lambda(n)\Lambda\Big(\frac{an+h}{b}\Big)\big[b|an+h\big] \tag{6.3}$$
$$= \big[(a,b)|h\big]\big[(a,b) = (a,h)\big]\big[(a,b) = (b,h)\big]\frac{(a,b)}{b}\mathfrak{S}(abh)x + O_\epsilon\Big(\frac{(a,b)}{b}x^{1/2+\epsilon}\Big).$$

To have less faith in these conjectures as they become more general is perhaps the right attitude. For *fixed* a, b and h there is every reason to believe that all are correct, however. These conjectures were made by Hardy and Littlewood for fixed a, b and h with the error term $O_\epsilon(x^{1/2})$ replaced by the more modest $o(x)$.

We have written these conjectures seperately in increasing order of generality because one may treat pair, triple, and quadruple correlation respectively using only Conjecture 6.1.1, 6.1.2, and 6.1.3.

The results we will prove are

Theorem 6.1.4 (Montgomery)**.** *Conditioned on Conjecture 6.1.1 and RH, for all Schwartz* $\eta : \mathbb{R} \to \mathbb{R}$ *such that* $\operatorname{supp} \hat{\eta} \subset \{(x_1, x_2) : |x_1| + |x_2| < 4\}$*,*

$$\lim_{T\to\infty} \frac{1}{T}\int_T^{2T} \sum_{\substack{\gamma_1,\gamma_2 \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), \tfrac{\log T}{2\pi}(\gamma_2 - t)\Big)\, dt$$
$$= \int_{\mathbb{R}^2} \eta(x_1, x_2)\big(1 - K(x_1 - x_2)^2\big)\, dx_1\, dx_2.$$

Remark: Using a 1-level density estimate, this is equivalent to the claim that

$$\frac{2\pi}{T\log T} \sum_{\substack{T \le \gamma_1,\gamma_2 \le 2T \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - \gamma_2)\Big) \sim \int_{\mathbb{R}} \eta(x)\big(1 - K(x)^2\big)\, dx$$

for all fixed Schwart $\eta : \mathbb{R} \to \mathbb{R}$ such that $\operatorname{supp} \hat{\eta} \subset (-2, 2)$.

This dates back to Mongtomery's original paper on the pair correlation conjecture [53], though the argument did not actually appear in print until [54]. We generalize the result to three and four point correlations:

Theorem 6.1.5. *Conditioned on Conjecture 6.1.2 and RH, for all Schwartz* $\eta : \mathbb{R}^3 \to \mathbb{R}$ *such that* $\mathrm{supp}\, \hat{\eta} \subset \{(x_1, x_2, x_3) : |x_1| + |x_2| + |x_3| < 8/3\}$,

$$\lim_{T\to\infty} \frac{1}{T} \int_T^{2T} \sum_{\substack{\gamma_1,\gamma_2,\gamma_3 \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), \tfrac{\log T}{2\pi}(\gamma_2 - t), \tfrac{\log T}{2\pi}(\gamma_3 - t)\Big)\, dt$$
$$= \int_{\mathbb{R}^3} \eta(x) \det_{3\times 3} \big(K(x_i - x_j)\big)\, d^3x.$$

Theorem 6.1.6. *Conditioned on Conjecture 6.1.3 and RH, for all Schwartz* $\eta : \mathbb{R}^4 \to \mathbb{R}$ *such that* $\mathrm{supp}\, \hat{\eta} \subset \{(x_1, x_2, x_3, x_3) : |x_1| + |x_2| + |x_3| + |x_4| < 12/5\}$,

$$\lim_{T\to\infty} \frac{1}{T} \int_T^{2T} \sum_{\substack{\gamma_1,\gamma_2,\gamma_3,\gamma_3 \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), ..., \tfrac{\log T}{2\pi}(\gamma_4 - t)\Big)\, dt$$
$$= \int_{\mathbb{R}}^4 \eta(x) \det_{4\times 4} \Big(K(x_i - x_j)\Big)\, d^4x.$$

Remark: Note that 8/3 and 12/5 are each greater that 2, which corresponds to the region in Theorem 1.1.6. In fact, one can prove slightly more than these theorems state; in Theorem 6.1.5 for instance, it is trivial to verify the above equation when η is supported in the region $x_1 + x_2 + x_3 \neq 0$. On the other hand, if η is supported in a region in which x_i, x_j are positive (x_1 and x_2 both positive for instance), one may verify the above equation when η is supported in the region $x_i \wedge x_j + x_i \vee x_j/2 < 1$. By permuting the coordinates of $\mathbb{R}^3$, one can verify (on Conjecture 6.1.2 !) that the triple correlation function of zeta zeros agrees with that of a sine-determinantal point process for a region slightly larger than even that of Theorem 6.1.5. Even still, this region will miss some interesting features of the sine-determinant. We are prevented from enlarging it, in the first place, because of the error term in (6.2). But even ignoring these error terms altogether – on for instance a somewhat heuristic assumption that a great deal of cancelation occurs when adding them – what falls out of our analysis is not really the sine-determinant itself, only a proxy for it that agrees against sufficiently band limited test functions. Similar considerations hold for Theorem 6.1.6. The reason for this will be made more clear in the proofs that follow, but corresponds to the fact that Conjectures 6.1.2-6.1.3 are quite false when a, b, or h are larger than x.

6.2 Oscillations of the singular series

In order to prove Theorems 6.1.4-6.3.1, we will need information about the size of the singular series $\mathfrak{S}(h)$ on average, and its interaction on average with prime sums. In proving Theorem 6.1.4, Montgomery made use of the fact that

Theorem 6.2.1 (Montgomery)**.**

$$\sum_{h\leq H}\Big(1-\frac{h}{H}\Big)\big(\mathfrak{S}(h)-1\big) = -\tfrac{1}{2}\log H + O(1).$$

In other words, on average $\mathfrak{S}(h)-1$ looks like $-1/2h$, which agrees at any rate with our intution that to first order $\mathfrak{S}(h)$ will oscillate around 1. It is worth noting that aside aesthetic considerations there is no special signicance to the weight $(1-h/H)_+$, except that it is slightly easier analytically to work with than $\mathbf{1}_{[0,H]}(h)$. The theorem is no harder to prove when $(1-h/H)_+$ is replaced by $A(h/H)$ where A is compactly supported with bounded second derivative such that $A(0) = 1$.

It is convenient to extend Montgomery's result in the following way:

Theorem 6.2.2. *For A with compact support and bounded second derivative such that $A(0) = 1$,*

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big)\big([(h,p^\lambda)=1]\mathfrak{S}(hp^\lambda)-1\big) = -\tfrac{1}{2}\big(\log(H)-\log_+(H/p)\big)+O(1)+O\Big(\frac{\log H}{p}\Big),$$

with the implied constants of error terms depending only on the function A.

More generally,

Theorem 6.2.3. *For k fixed and $N = p_1^{\lambda_1} \cdots p_k^{\lambda_k}$ and A as above,*

$$\sum_{h \geq 1} A\Big(\frac{h}{H}\Big)\Big(\big[(h,N)=1\big]\mathfrak{S}(hN)-1\Big) = -\tfrac{1}{2}\sum_{\delta | N}\mu(\delta)\log_+(H/\delta)+O(1)+O\Big(\frac{\log H}{p_1 \wedge \cdots \wedge p_k}\Big),$$

with the implied constants of error terms depending only on A and k. If $N=1$ the second error term $O(\log H/p_1 \wedge \cdots \wedge p_k)$ may be deleted.

Remark: In fact, what we require of A for such a result is even looser than a bounded second derivative; in our proof we require only that $\hat{A}$ decay quadratically.

At least two related methods exist to prove results of this sort. The first involves the Dirichlet series

$$\sum_{h} \frac{\big[(h,N)=1\big]\mathfrak{S}(hN)-1}{h^s}, \tag{6.4}$$

which can also be written

$$\zeta(s)\bigg(2cT(s)\prod_{p|N}\Big(1-\tfrac{1}{p-1}+\tfrac{1}{p^s-1}\Big)^{-1}-1\bigg), \tag{6.5}$$

where

$$c := \prod_{p>2}\Big(1-\tfrac{1}{(p-1)^2}\Big)$$

$$T(s) := \zeta(s+1)\Big(1-\tfrac{1}{2^{s+1}}\Big)\prod_{p>2}\Big(1+\frac{2}{(p-2)p^{s+1}}-\frac{1}{(p-2)p^{2s+1}}\Big),$$

so that T(1)=1/2c, with (6.5) defined at $s=1$ by continuity. This shows that (6.3) converges (though not absolutely of course) for $\Re s > 0$. In addition $2cT(\epsilon) = 1/\epsilon + O(1)$ for small ϵ. With a certain additional amount of calculation these formulas allow us to prove Theorems 6.2.1 - 6.2.3. The equivalence of (6.3) and

(6.5) for $N = 1$ is used in [54] in this way to demonstrate Theorem 6.2.1 and therefore 6.1.4. (In fact, by an elementary sieving procedure, one can deduce Theorems 6.2.2 and 6.2.3 directly from Theorem 6.2.1, at least for the weight $A(x) = (1-x)_+$.)

That the computations thus outlined yield Theorem 6.1.4, the authors of [54] note, seems "fortuitous... but miracles do not happen by accident, so it seems there is something going on here that remains to be understood." We will take a different approach to prove these theorems that – while still leaving the parsimony of Theorems 6.2.1-6.2.3 slightly opaque – at least makes more transparent the oscillations of $\mathfrak{S}(h)$ or $\big[(h,N)=1\big]\mathfrak{S}(hN)$ around the value 1. We will expand these functions into a linear combination of Ramanujan sums, and then apply Fourier analysis to the latter.

In fact we will really only prove in full detail Theorem 6.2.2, with the proofs of Theorems 6.2.1 and 6.2.3 being almost identical (the latter case just involving more book-keeping).

Proof of Theorem 6.2.2. We note that

$$\begin{aligned}\big[(h,N)=1\big]\mathfrak{S}(hN) &= \prod_p \Big(1+\frac{1-p[p|N]}{(p-1)^2}c_p(h)\Big)\\ &= \sum_{q=1}^{\infty}\frac{\mu(q)^2}{\phi(q)^2}c_q(h)\prod_{p|q}(1-p[p|N]),\end{aligned}\tag{6.6}$$

where

$$c_q(h) := \sum_{(a,q)=1} e\big(\tfrac{ah}{q}\big)$$

are Ramanujan sums, satisfying for a prime p,

$$c_p(h) = \begin{cases} p-1 & \text{if } p|h \\ -1, & \text{if } p\nmid h\end{cases}$$

and for $(q, q') = 1$,

$$c_q(h)c_{q'}(h) = c_{qq'}(h).$$

None of these properties lies deep, the last following from the Chinese remainder theorem.

It is the expression (6.6) that will allow us to perform Fourier analysis on sums of singular series. Note that

$$\big[(h, N) = 1\big]\mathfrak{S}(hN) - 1 = \sum_{q>1} \frac{\mu(q)^2}{\phi(q)^2} c_q(h)\phi\big((q, N)\big)\mu\big((q, N)\big).$$

We use this to establish

Lemma 6.2.4. *For a fixed function A meeting the condition of the theorem,*

$$\begin{aligned}\sum_{h\geq 1} A\Big(\frac{h}{H}\Big)\big(\big[(h, N) = 1\big]\mathfrak{S}(hN) - 1\big) = &-\tfrac{1}{2}\sum_{q\leq H} \frac{\mu(q)^2}{\phi(q)^2}\mu\big((q, N)\big)\phi\big((q, N)\big) \\ &+ O\bigg(\frac{1}{H}\sum_{q\leq H} \mu(q)^2\Big(\frac{q}{\phi(q)}\Big)^2 \phi\big((q, N)\big)\bigg) \\ &+ O\bigg(\sum_{H<q\leq H'} \frac{\mu(q)^2 q}{\phi(q)^2}\phi\big((q, N)\big)\bigg) \\ &+ O\bigg(H\sum_{q>H'} \frac{\mu(q)^2 d(q)}{\phi(q)^2}\phi\big((q, N)\big)\bigg), \quad (6.7)\end{aligned}$$

uniformly for all N and $H \geq H'$.

Remark: In the instances we will use this theorem (that N is a product of a fixed number of primes), the error terms will all be $O(1)$ with a proper choice of H'.

Proof. To prove (6.7), we rely upon a trick whose usefulness can only be made clear in the course of our proof. We extend the definition of A from $\mathbb{R}_+$ to $\mathbb{R}$ by symmetry across 0, so that $A(-x) = A(x)$. The left hand side of (6.7) is thus

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big) \sum_{q>1} \frac{\mu(q)^2}{\phi(q)^2} c_q(h)\mu\big((q,N)\big)\phi\big((q,N)\big) \tag{6.8}$$
$$= \tfrac{1}{2} \sum_{q>1} \frac{\mu(q)^2}{\phi(q)^2} \Big(\sum_{h\in\mathbb{Z}} A\Big(\frac{h}{H}\Big) c_q(h) - \phi(q) \Big) \mu\big((q,N)\big)\phi\big((q,N)\big),$$

as $c_q(0) = \phi(q)$ and $c_q(h) = c_q(-h)$.

We shall show that each of the following bounds hold:

$$\sum_{h\in\mathbb{Z}} A\Big(\frac{h}{H}\Big) c_q(h) = \begin{cases} O(q^2/H) \\ O(q) \\ \phi(q) + O(Hd(q)). \end{cases} \tag{6.9}$$

By applying in (6.8) the first of these relations for $q \leq H$, the second for $H \leq g < H'$, and the third for $q > H'$, we obtain (6.7).

We therefore turn to (6.9). Note that the left hand side is, by Poisson summation,

$$\sum_{(a,q)=1} \sum_{h\in\mathbb{Z}} e\big(\tfrac{ah}{q}\big) A\Big(\frac{h}{H}\Big) = \sum_{(a,q)=1} \sum_{\ell\in\mathbb{Z}} H\hat{A}\big(A(\tfrac{a}{q} + \ell)\big). \tag{6.10}$$

This is turn is no more than

$$\sum_{\lambda\in\mathbb{Z}\setminus\{0\}} H\big|\hat{A}(H\tfrac{\lambda}{q})\big| \lesssim_A H \sum_{\lambda\neq 0} \frac{1}{1 + (H/q)^2\lambda^2},$$

as $A'' \in L^1(\mathbb{R})$. (Here one sees that we really only need $\hat{A}$ to have quadratic decay, for instance.) Clearly this sum is both

$$O\Big(H \cdot \frac{q^2}{H^2}\Big) = O\Big(\frac{q^2}{H}\Big)$$

and (a better bound when H/q is small)

$$O(q).$$

On the other hand, making the definition

$$S_n := \sum_{a=0}^{n-1} \sum_{h \in \mathbb{Z}} e\big(\tfrac{ah}{q}\big) A\Big(\frac{h}{H}\Big) = n \sum_{j \in Z} A\Big(\frac{nj}{H}\Big),$$

we have

$$\sum_{(a,q)=1} \sum_{h \in \mathbb{Z}} e\big(\tfrac{ah}{q}\big) A\Big(\frac{h}{H}\Big) = \sum_{\delta | q} \mu\big(\tfrac{q}{\delta}\big) S_\delta.$$

Yet

$$S_n = nA(0) + n \sum_{j \neq 0} A\Big(\frac{nj}{H}\Big) = n + O(H),$$

and the third relationship in (6.9) follows immediately. □

This lemma established, for $N = p^\lambda$ a prime power, it is easy to verify that

$$\begin{aligned}
&\frac{1}{H} \sum_{q \leq H} \mu(q)^2 \Big(\frac{q}{\phi(q)}\Big)^2 \phi\big((q, p^\lambda)\big) \\
&= \frac{1}{H} \sum_{\substack{(q,p)=1 \\ q \leq H}} \mu(q)^2 \Big(\frac{q}{\phi(q)}\Big)^2 + \frac{1}{H} \sum_{\substack{p|q \\ q \leq H}} \mu(q)^2 \Big(\frac{q}{\phi(q)}\Big)^2 (p-1) \\
&= O(1)
\end{aligned}$$

and using the same decomposition,

$$\sum_{H < q \leq H'} \frac{\mu(q)^2 q}{\phi(q)^2} \phi\big((q, p^\lambda)\big) = O\big(1 + \log(H'/H)\big),$$

$$H \sum_{q > H'} \frac{\mu(q)^2 d(q)}{\phi(q)^2} \phi\big((q, p^\lambda)\big) = O\Big(\frac{H}{H'} \log H'\Big).$$

Choosing $H' = H \log H$, for instance, we see that (6.7) implies

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big)\big(\big[(h,p^\lambda)=1\big]\mathfrak{S}(hp^\lambda)-1\big) = -\tfrac{1}{2}\sum_{q\leq H}\frac{\mu(q)^2}{\phi(q)}\mu\big((q,p)\big)\phi\big((q,p)\big)+O(1). \tag{6.11}$$

From the well known formula first proved in [80] (a by now standard exercise with Dirichlet series):

$$\sum_{q\leq H}\frac{\mu(q)^2}{\phi(q)} = \log H + O(1),$$

we obtain

$$\begin{aligned}\sum_{q\leq H}\frac{\mu(q)^2}{\phi(q)}\mu\big((q,p)\big)\phi\big((q,p)\big) &= \sum_{q\leq H}\frac{\mu(q)^2}{\phi(q)} - \sum_{\substack{q'\leq H/p\\(q',p)=1}}\frac{\mu(q')^2}{\phi(q'p)}p\\ &= \sum_{q\leq H}\frac{\mu(q)^2}{\phi(q)} - \frac{p}{\phi(p)}\sum_{q'\leq H/p}\frac{\mu(q')^2}{\phi(q')} + \frac{p}{\phi(p)^2}\sum_{\substack{q''\leq H/p^2\\(q'',p)=1}}\frac{(\mu(q'')^2}{\phi(q'')}\\ &= \log H - \log_+(H/p) + O(1) + O\Big(\tfrac{\log H}{p}\Big).\end{aligned}$$

□

We have therefore proved in full detail Theorem 6.2.2. Theorem 6.2.1 follows rather more directly from Lemma 6.2.4, while Theorem 6.2.3 follows from a straightforward (but in notation somewhat tiresome) extension of the above argument.

A simple consequence of these results that we will require later is

Corollary 6.2.5.

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big)\big(\big[p|h\big]\mathfrak{S}(ph)-1\big) = -\tfrac{1}{2}\log_+(H/p)+O(1).$$

We note corollaries of these results that have a slightly different form from what we will require later to compute correlation functions, but which are worth recording anyway:

Theorem 6.2.6. *For A as in Theorem 6.2.2, and $f \in C_c(\mathbb{R})$,*

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big) \sum_{n\geq 1} f\Big(\tfrac{\log n}{\log H}\Big) \frac{\Lambda(n)}{n} \big([(h,n)=1]\mathfrak{S}(hn)-1\big)$$
$$= -\tfrac{1}{2}\log^2 H \int_0^\infty f(x)(1-(1-x)_+)\,dx + o(\log H).$$

More generally,

Theorem 6.2.7. *For A as in Theorem 6.2.2, and $f \in C_c(\mathbb{R}^k)$,*

$$\sum_{h\geq 1} A\Big(\frac{h}{H}\Big) \sum_{n_1,\ldots,n_k\geq 1} f\Big(\tfrac{\log n_1}{\log H},\ldots,\tfrac{\log n_k}{\log H}\Big) \frac{\Lambda(n_1)}{n_1}\cdots\frac{\Lambda(n_k)}{n_k}$$
$$\times \big([(h,n_1\cdots n_k)=1]\mathfrak{S}(hn_1\cdots n_k)-1\big)$$
$$= -\tfrac{1}{2}\log^{k+1} H \int_{\mathbb{R}_+^k} f(x) \sum_{\emptyset\subseteq S\subseteq[k]} (-1)^{|S|}(1-x_S)_+\,d^k x + o(\log^k H),$$

where for a set $S \subseteq [k]$,

$$x_S := \sum_{\ell\in S} x_\ell.$$

6.3 An evaluation of correlation functions

Finally we are in a position to prove Theorem 6.1.5 and Theorem 6.1.6. As a warm-up, we will give Montgomery's proof of Theorem 6.1.4 in the notation that will prove convenient in the proofs that follow.

Proof of Theorem 6.1.4. We assume the truth of the Hardy-Littlewood conjecture 6.1.2. We will show that for any Schwartz σ of mass 1 with $\hat{\sigma}$ compactly supported and $\eta = \eta_1 \otimes \eta_2$ with supp $\hat{\eta} \subset \{(x_1, x_2) : |x_1| + |x_2| < 4\}$,

$$\lim_{T\to\infty} \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \sum_{\substack{\gamma_1,\gamma_2 \\ \text{distinct}}} \eta_1\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t)\Big)\eta_2\Big(\tfrac{\log T}{2\pi}(\gamma_2 - t)\Big)\, dt$$
$$= \int_{\mathbb{R}^2} \eta_1(x_1)\eta_2(x_2)\big(1 - K(x_1 - x_2)^2\big)\, dx_1\, dx_2.$$

It is then easy, exactly as in chapter 4 by using Theorem 3.4.1, to prove this result when $\sigma = \mathbf{1}_{[1,2]}$ and η is not necessarily separable, by approximating $\mathbf{1}_{[1,2]}$ in the $L^1(\mathbb{R})$ metric by sufficiently smooth σ, and general Schwartz η uniformly by linear combinations of seperable η.

We use once again the measure dS, first discussed in Chapter 2 following the explicit formula. From a 1-level density estimate, the above equation is equivalent to

$$\lim_{T\to\infty} \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \int_{-\infty}^{\infty} \eta_1\Big(\tfrac{\log T}{2\pi}(\xi_1 - t)\Big)\, dS(\xi_1) \int_{-\infty}^{\infty} \eta_2\Big(\tfrac{\log T}{2\pi}(\xi_2 - t)\Big)\, dS(\xi_2)\, dt$$
$$= \int_{\mathbb{R}^2} \eta_1(x_1)\eta_2(x_2)\big(\delta(x_1 - x_2) - K(x_1 - x_2)^2\big)\, dx_1\, dx_2.$$

To at this point briefly repeat some of the analysis of Chapter 5: by using the explicit formula to write each side in terms of $\hat{\eta_1}$ and $\hat{\eta_2}$ rather than η_1 and η_2, we see that this is equivalent to

$$\lim_{T\to\infty} \frac{1}{\log^2 T} \sum_{\varepsilon \in \{-1,1\}^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{\sigma}\big(\tfrac{-T}{2\pi}(\epsilon_1 x_1 + \epsilon_2 x_2)\big)\hat{\eta_1}\Big(\tfrac{\epsilon_1 x_1}{\log T}\Big)\hat{\eta_2}\Big(\tfrac{\epsilon_2 x_2}{\log T}\Big)\, dz(x_1)\, dz(x_2)$$
$$= \int_{\mathbb{R}} \hat{\eta_1}(y)\hat{\eta_2}(-y)\big(1 - (1 - |y|)_+\big)\, dy. \tag{6.12}$$

The summand of the left hand side is plainly asymptotically 0 unless $\epsilon_1 = -\epsilon_2$.

In this remaining case, by symmetry, if we can show that

$$\begin{aligned}&\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{\sigma}\big(\tfrac{-T}{2\pi}(x_1-x_2)\big)\hat{\eta_1}\Big(\tfrac{x_1}{\log T}\Big)\hat{\eta_2}\Big(\tfrac{-x_2}{\log T}\Big)\,dz(x_1)\,dz(x_2)\\ &\sim \log^2 T\int_0^\infty \hat{\eta_1}(y)\hat{\eta_2}(-y)\big(1-(1-|y|)_+\big)\,dy. \end{aligned} \tag{6.13}$$

we will recover (6.12). Letting $\hat{\eta_1}(x) = f_1(x)$ and $\hat{\eta_2}(-x) = f_2(x)$ to ease notation, note that the left hand side of (6.13) is

$$\begin{aligned}&\sum_{n,m}\frac{\Lambda(n)\Lambda(m)}{\sqrt{nm}}\hat{\sigma}\Big(\tfrac{-T}{2\pi}\log\big(\tfrac{n}{m}\big)\Big)f_1\Big(\tfrac{\log n}{\log T}\Big)f_2\Big(\tfrac{\log m}{\log T}\Big)\\ &-\sum_n\frac{\Lambda(n)}{\sqrt{n}}\int_{\mathbb{R}}\hat{\sigma}\big(\tfrac{-T}{2\pi}(x_2-\log n)\big)e^{x_2/2}f_1\Big(\tfrac{\log n}{\log T}\Big)f_2\Big(\tfrac{x_2}{\log T}\Big)\,dx_2\\ &-\sum_m\frac{\Lambda(m)}{\sqrt{m}}\int_{\mathbb{R}}\hat{\sigma}\big(\tfrac{-T}{2\pi}(x_1-\log m)\big)e^{x_1/2}f_1\Big(\tfrac{x_1}{\log T}\Big)f_2\Big(\tfrac{\log m}{\log T}\Big)\,dx_1\\ &+\int_{\mathbb{R}}\int_{\mathbb{R}} e^{(x_1+x_2)/2}\hat{\sigma}\big(\tfrac{-T}{2\pi}(x_1-x_2)\big)f_1\Big(\tfrac{x_1}{\log T}\Big)f_2\Big(\tfrac{x_2}{\log T}\Big)\,dx_1\,dx_2.\end{aligned}$$

Note that for a Schwartz function g, if $\hat{\sigma}$ is supported in $[-\kappa,\kappa]$,

$$\begin{aligned}\int_{\mathbb{R}}\hat{\sigma}\big(\tfrac{-T}{2\pi}(y-Y)\big)e^{y/2}g\Big(\tfrac{y}{\log T}\Big)\,dy &= \sigma(0)e^{Y/2}g\Big(\tfrac{Y}{\log T}\Big)+O\Big(\frac{1}{T^2}\sup\frac{d}{dy}e^{y/2}g\Big(\tfrac{y}{\log T}\Big)\Big)\\ &= \sigma(0)e^{Y/2}g\Big(\tfrac{Y}{\log T}\Big)+O\Big(\frac{e^{Y/2}}{T^2}\Big),\end{aligned} \tag{6.14}$$

where the supremum is taken over all y within $2\pi\kappa/T$ of Y, and the implied constants will depend on σ and g.

The left hand side of (6.13) is thus

$$\begin{aligned}&\sum_{n,m}\frac{\Lambda(n)\Lambda(m)}{\sqrt{nm}}\hat{\sigma}\Big(\tfrac{-T}{2\pi}\log\big(\tfrac{n}{m}\big)\Big)f_1\Big(\tfrac{\log n}{\log T}\Big)f_2\Big(\tfrac{\log m}{\log T}\Big)\\ &-\frac{2\sigma(0)}{T}\sum_n\Lambda(n)f_1\Big(\tfrac{\log n}{\log T}\Big)f_2\Big(\tfrac{\log n}{\log T}\Big)\\ &+\frac{\sigma(0)}{T}\int_1^\infty f_1\Big(\tfrac{\log\nu}{\log T}\Big)f_2\Big(\tfrac{\log\nu}{\log T}\Big)\,d\nu+O\Big(\frac{T^{\beta_1\vee\beta_2}}{T^2}\Big).\end{aligned} \tag{6.15}$$

Suppose f_1 is supported in $(-\beta_1, \beta_1)$ and f_2 is supported in $(-\beta_2, \beta_2)$. As supp $\hat{\eta} \subset \{(x_1, x_2) : |x_1| + |x_2| < 4\}$, it is the case that $\beta_1 + \beta_2 < 4$. As in (6.13) our integrand is nonzero only when x_1 is close to x_2, we lose no generality in assuming $\beta_1 = \beta_2$, both less than 2.

It is the first of the terms in (6.15) that is most interesting. We have

$$\begin{aligned}
&\sum_{n,m} \frac{\Lambda(n)\Lambda(m)}{\sqrt{nm}} \hat{\sigma}\Big(\tfrac{-T}{2\pi} \log\big(\tfrac{n}{m}\big)\Big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log m}{\log T}\Big) \\
&= \sum_{n} \frac{\Lambda^2(n)}{n} f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) + \sum_{n \neq m} \frac{\Lambda(n)\Lambda(m)}{\sqrt{nm}} \hat{\sigma}\Big(\tfrac{-T}{2\pi} \log\big(\tfrac{n}{m}\big)\Big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log m}{\log T}\Big) \\
&= \sum_{n} \frac{\Lambda^2(n)}{n} f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) \\
&\quad + \sum_{n} \sum_{\substack{h \neq 0 \\ n+h>0}} \frac{\Lambda(n)\Lambda(n+h)}{n} \hat{\sigma}\big(\tfrac{T}{2\pi} \log\big(1 + \tfrac{h}{n}\big)\big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) + O\Big(\frac{\log^2 T}{T}\Big).
\end{aligned}$$

By partial summation to access the sums of Conjecture 6.1.2, and rearranging back, for any $\epsilon > 0$,

$$\begin{aligned}
&\sum_{n} \sum_{\substack{h \neq 0 \\ n+h>0}} \frac{\Lambda(n)\Lambda(n+h)}{n} \hat{\sigma}\big(\tfrac{T}{2\pi} \log\big(1 + \tfrac{h}{n}\big)\big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) \\
&= \sum_{n \geq 1} \sum_{\substack{h \neq 0 \\ n+h>0}} \frac{\mathfrak{S}(h)}{n} \hat{\sigma}\big(\tfrac{T}{2\pi} \log\big(1 + \tfrac{h}{n}\big)\big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) \\
&\quad + O_\epsilon\Bigg(\sum_{n \leq T^{\beta_1 \wedge \beta_2}} \sum_{\substack{h \neq 0 \\ n+h>0}} \frac{1}{n^{3/2-\epsilon}} \Bigg)
\end{aligned}$$

Hence (6.15) is equal to

$$\sum_n \frac{\Lambda(n)^2}{n} f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) + \sum_n \sum_{h \neq 0} \frac{\mathfrak{S}(h)-1}{n} \hat{\sigma}\big(\tfrac{T}{2\pi}\big(1+\tfrac{h}{n}\big)\big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big)$$
$$+ \sum_n \sum_{h \neq 0} \frac{1}{n} \hat{\sigma}\big(\tfrac{T}{2\pi}\big(1+\tfrac{h}{n}\big)\big) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big) - \frac{2\sigma(0)}{T} \sum_n \Lambda(n) f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big)$$
$$+ \frac{\sigma(0)}{T} \int_1^\infty f_1\Big(\tfrac{\log \nu}{\log T}\Big) f_2\Big(\tfrac{\log \nu}{\log T}\Big)\, d\nu + O_\epsilon\Big(\frac{T^{(1/2+\epsilon)\beta_1 \wedge \beta_2}}{T} + \frac{T^{(\beta_1+\beta_2)/2}}{T}\Big)$$

We note that by Theorem 6.1.4

$$\sum_{h \neq 0} (\mathfrak{S}(h)-1) \hat{\sigma}\big(\tfrac{T}{2\pi} \log\big(1+\tfrac{h}{n}\big)\big) = \sum_{0<|h| \lesssim n/T} (\mathfrak{S}(h)-1)\Big(\hat{\sigma}\big(\tfrac{T}{2\pi}\tfrac{h}{n} + O\big(T\big(\tfrac{h}{n}\big)^2\big)\Big)$$
$$= -\log_+(n/T) + O(1+n/T^2).$$

We use the Riemann hypothesis and Theorem 6.1.4 to evaluate the above expression. The expression (6.15) – and therefore the left hand side of (6.13) – are equal to

$$\sum_n \frac{\Lambda^2(n)}{n} f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big)$$
$$- \sum_n \frac{\log_+(n/T) + O(1+n/T^2)}{n} f_1\Big(\tfrac{\log n}{\log T}\Big) f_2\Big(\tfrac{\log n}{\log T}\Big)$$
$$+ O_\epsilon\Big(\log T + \frac{T^{(1/2+\epsilon)\beta_1 \wedge \beta_2}}{T} + \frac{T^{(\beta_1+\beta_2)/2}}{T}\Big).$$

As long as $\beta_1 + \beta_2 < 4$, this expression is asymptotic to

$$\sim \log^T \int_0^\infty \big(x - (x-1)_+\big) f(x)\, dx,$$

which is the right hand side of (6.13). □

Proof of Theorem 6.1.5. Our proof runs along the same lines as above. We we will ultimately seek to show is that for a certain class of functions f_1, f_2 and f_3 (cf. Corollary 5.10.2.)

$$\begin{aligned}
&\int_{\mathbb{R}^3} \hat{\sigma}\Big(\tfrac{\log T}{2\pi}(x_3 - x_1 - x_2)\Big) f_1\Big(\tfrac{x_1}{\log T}\Big) f_2\Big(\tfrac{x_2}{\log T}\Big) f_3\Big(\tfrac{x_3}{\log T}\Big)\, dz(x_1) dz(x_2) dz(x_3) \\
&\sim -\log^3 T \int_0^\infty \int_0^\infty f_1(x_1) f_2(x_2) f_3(x_1+x_2) \qquad (6.16)\\
&\qquad\qquad \times (1 - (1-x_1)_+ - (1-x_2)_+ + (1 - x_1 - x_2)_+)\, dx_1 dx_2.
\end{aligned}$$

Let us see that this is sufficient for our proof. We borrow from chapter 5 the notation of point processes, not an essential step, but one which greatly simplifies computations. We define for Schwartz η_1, η_2, η_3 the quantity

$$\begin{aligned}
Q &:= \int_{\mathbb{R}} \frac{\sigma(t/T)}{T} \prod_{\ell=1}^{3} \int_{-\infty}^{\infty} \eta_\ell\Big(\tfrac{\log T}{2\pi}(\xi_\ell - t)\Big)\, dS(\xi_\ell)\, dt \\
&= \mathop{\mathbf{E}}_{Z_T(\sigma)} \prod_{\ell=1}^{3} \Big(\sum_i \eta_\ell(\xi_i) - \int \eta_\ell \Big) + o(1).
\end{aligned}$$

If $\eta = \eta_1 \otimes \eta_2 \otimes \eta_3$ is a test function of the sort prescribed by the theorem, then supp $\hat{\eta}_\ell \subset [-\beta_\ell, \beta_\ell]$ where $\beta_1 + \beta_2 + \beta_3 < 8/3$. Hence $\beta_i + \beta_j$ is certainly less than 4 for any pair i, j, and thus expanding the product above and using a 1-level density estimate and Theorem 6.1.4 to evaluate lower level correlation functions, we see that the theorem will be proved if we can show that Q is equal to

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell}^{3} \Big(\sum_i \eta_\ell(x_i) - \int \eta_\ell \Big).$$

The expansion of the right hand side above is somewhat complicated, but fortunately simplifies in the end; we let $\varphi_1 = \hat{\eta}_1, \varphi_2 = \hat{\eta}_2, \varphi_3 = \hat{\eta}_3$ and define, in analogy with $\phi|_\epsilon$,

$$\varphi||_\epsilon(x) := (1 - \alpha_\epsilon(x))\varphi(x),$$

where $\alpha_\epsilon(x) = \alpha(x/\epsilon)$ for α some even bump function centered at 0 with $\alpha(0) = 1$

as in previous chapters. Note that

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{3} \Big(\sum_i \eta_\ell(x_i) - \int \eta_\ell \Big) = \lim_{N\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}'_N} \prod_{\ell=1}^{3} \Big(\sum_i \eta_\ell(x_i) - \int \eta_\ell \Big)$$

$$= \lim_{N\to\infty} \mathop{\mathbf{E}}_{\mathcal{S}'_N} \prod_{\ell=1}^{3} \Big(\sum_i (\varphi_\ell ||_\epsilon)\check{}(x_i) - \int (\varphi ||_\epsilon)\check{} \Big) + O(\epsilon)$$

$$= \lim_{\epsilon\to\epsilon^+} \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{j} \sum_i (\varphi ||_\epsilon)\check{}(x_i).$$

We expand this last expression in terms of a Fourier integral, broken into quadrants:

$$\sum_{\varepsilon\in\{-1,1\}^3} \lim_{\epsilon\to 0^+} \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{3} \sum_i \Big(\int_{\mathbb{R}} \phi_\ell(y_\ell) \Omega_\epsilon(\varepsilon_\ell y_\ell) e(x_i y_\ell)\, dy_\ell \Big),$$

where, recall $\Omega_\epsilon = (1-\alpha_\epsilon)\mathbf{1}_{[0,\infty)}$. (The utility of this will be to more accurately match our sine-determinantal expression to the arithmetic expression we will obtain by applying the explicit formula to Q.)

We now must resort to a few unfortunately somewhat tedious calculations. Note that by an expansion into correlation sums over distinct indices (cf. the expansion in 5.4),

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{3} \sum_i \eta_\ell(x_i) = \int_{\mathbb{R}^3} \eta_1(x_1)\eta_2(x_2)\eta_3(x_3) J_3(x_1,x_2,x_3)\, d^3x \qquad (6.17)$$

where

$$\begin{aligned}
J_3(x_1,x_2,x_3) :=& \det_{3\times 3}[K(x_i-x_j)] + \delta(x_1-x_2)\det_{\{2,3\}^2}[K(x_i-x_j)] \\
&+ \delta(x_2-x_3)\det_{\{3,1\}^2}[K(x_i-x_j)] + \delta(x_3-x_1)\det_{\{1,2\}^2}[K(x_i-x_j)] \\
&+ \delta(x_1-x_2)\delta(x_2-x_3) \\
=& 1 - K^2(x_2-x_3) - K^2(x_1-x_2) - K^2(x_1-x_3) \\
&+ 2K(x_1-x_2)K(x_2-x_3)K(x_3-x_1) \\
&+ \delta(x_1-x_2)\big(1-K^2(x_2-x_3)\big) + \delta(x_2-x_3)\big(1-K^2(x_1-x_3)\big) \\
&+ \delta(x_1-x_3)\big(1-K^2(x_1-x_2)\big) + \delta(x_1-x_2)\delta(x_2-x_3).
\end{aligned}$$

so that

$$\begin{aligned}
\check{J}_3(y_1,y_2,y_3) =& \delta(y_1)\delta(y_2)\delta(y_3) - \delta(y_1)\delta(y_2+y_3)(1-|y_2|)_+ \\
&- \delta(y_2)\delta(y_3+y_1)(1-|y_3|)_+ - \delta(y_3)\delta(y_1+y_2)(1-|y_1|)_+ \\
&+ 2\delta(y_1+y_2+y_3)\int_{\mathbb{R}} \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v)\mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v+y_1)\mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v+y_1+y_2)\,dv \\
&+ \delta(y_1+y_2+y_3)\Big(\delta(y_3) - (1-|y_3|)_+ + \delta(y_1) - (1-|y_1|)_+ + \delta(y_2) \\
&- (1-|y_2|)_+\Big) + \delta(y_1+y_2+y_3)
\end{aligned}$$

This distribution is supported on the plane $y_1+y_2+y_3=0$, and by inspection one can see that on this plane,

$$\int_{\mathbb{R}} \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v)\mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v+y_1)\mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}(v+y_1+y_2)\,dv = (1-|y_1|\vee|y_2|\vee|y_3|)_+.$$

Hence it follows, as $\Omega_\epsilon(0) = 0$,

$$\lim_{\epsilon \to 0^+} \mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{3} \sum_i \left(\int_{\mathbb{R}} \phi_\ell(y_\ell) \Omega_\epsilon(\varepsilon_\ell y_\ell) e(x_i y_\ell)\, dy_\ell \right)$$

$$= \int_{\mathbb{R}_+^3} \delta(\varepsilon_1 y_1 + \varepsilon_2 y_2 + \varepsilon_3 y_3) \varphi(\varepsilon_1 y_1)\varphi(\varepsilon_2 y_2)\varphi(\varepsilon_3 y_3) \Big(1 - (1 - |y_1|)_+ - (1 - |y_2|)_+$$

$$- (1 - |y_3|)_+ + 2(1 - |y_1| \vee |y_2| \vee |y_3|)_+ \Big)\, d^3 y.$$

$$= S_\varepsilon \qquad \text{say.}$$

On the other hand, from the explicit formula,

$$\begin{aligned} Q =& \frac{-1}{\log^3 T} \sum_{\varepsilon \in \{-1,1\}^3} \int_{\mathbb{R}^3} \hat{\sigma}\Big(\tfrac{-T}{2\pi} (\varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_3 x_3) \Big) \\ & \times \varphi_1\Big(\tfrac{\varepsilon_1 x_1}{\log T} \Big) \varphi_2\Big(\tfrac{\varepsilon_2 x_2}{\log T} \Big) \varphi_3\Big(\tfrac{\varepsilon_3 x_3}{\log T} \Big)\, dz(x_1) dz(x_2) dz(x_3) \\ =& Z_\varepsilon \qquad \text{say.} \end{aligned}$$

Thus if we can show for all ε that

$$S_\varepsilon \sim Z_\varepsilon$$

we will have proved the theorem. This is obvious when $\varepsilon = (1,1,1)$ or $(-1,-1,-1)$, and the reader may check that all remaining cases boil down to verifying (6.16) for all functions f_1, f_2, f_3 with supp $f_\ell \subset (-\infty, \beta_\ell)$ and $\beta_1 + \beta_2 + \beta_3 < 8/3$.[1] Clearly we lose no generality in assuming $\beta_3 = \beta_1 + \beta_2$, so that $\beta_1 + \beta_2 < 4/3$. Likewise, with no loss of generality, by symmetry we may suppose $\beta_1 \geq \beta_2$.

By repeating the analysis of our proof of Theorem 6.1.4, using in particular

[1] Clearly the values f_ℓ take on the negative reals play no role in evaluating (6.16). We support f_ℓ in $(-\infty, \beta_\ell)$ rather than $[0, \beta_\ell)$ only so that $f_\ell(0)$ is not forced to be 0.

(6.14), we see that the left hand side of (6.16) is

$$\sum_{n_1,n_2,m} \frac{\Lambda(n_1)\Lambda(n_2)\Lambda(m)}{\sqrt{n_1 n_2 m}} \hat{\sigma}\Big(\tfrac{T}{2\pi}\log\big(\tfrac{m}{n_1 n_2}\big)\Big) f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) f_3\Big(\tfrac{\log m}{\log T}\Big)$$
$$- \frac{\sigma(0)}{T}\int_{\mathbb{R}} \sum_{n_2} \Lambda(n_2) f_1\Big(\tfrac{\log \nu_1}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) f_3\Big(\tfrac{\log \nu_1 n_2}{\log T}\Big) d\big(\psi(\nu_1)-\nu_1\big)$$
$$- \frac{\sigma(0)}{T}\int_{\mathbb{R}} \sum_{n_2} \frac{\Lambda(n_2)}{n_2} f_1\Big(\tfrac{\log(\nu_3/n_2)}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) f_3\Big(\tfrac{\log n_3}{\log T}\Big)\, d\big(\psi(\nu_3)-\nu_3\big)$$
$$- \frac{\sigma(0)}{T}\int_{\mathbb{R}} \sum_{n_1} \frac{\Lambda(n_1)}{n_1} f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log(\nu_3/n_1)}{\log T}\Big) f_3\Big(\tfrac{\log \nu_3}{\log T}\Big)\, d\big(\psi(\nu_3)-\nu_3)\big)$$
$$- \frac{\sigma(0)}{T}\int_{\mathbb{R}} f_1\Big(\tfrac{\log \nu_1}{\log T}\Big) f_2\Big(\tfrac{\log \nu_2}{\log T}\Big) f_3\Big(\tfrac{\log(\nu_1\nu_2)}{\log T}\Big)\, d\nu_1 d\nu_2 + O\Big(\frac{T^{\beta_1+\beta_2}\log T}{T^2}\Big)$$
$$= \mathrm{I} - \mathrm{II} - \mathrm{III} - \mathrm{IV} - \mathrm{V} + O\Big(\frac{T^{\beta_1+\beta_2}\log T}{T^2}\Big) \qquad \text{say.}$$

Each of $\mathrm{II}, \mathrm{III}, \mathrm{IV}$ are plainly no more than $O(T^{\beta_1/2+\beta_2}/T)$.[2] On the other hand, there are no values n_1, n_2, and m which are prime such that $n_1 n_2 = m$. Therefore

$$I = \sum_{n_1,n_2} \sum_{\substack{h\neq 0 \\ n_1n_2+h>0}} \frac{\Lambda(n_1)\Lambda(n_2)\Lambda(n_1n_2+h)}{\sqrt{n_1n_2(n_1n_2+h)}} \hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\tfrac{h}{n_1n_2}\Big)\Big)$$
$$\times f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) f_3\Big(\tfrac{\log(n_1n_2+h)}{\log T}\Big)$$
$$= \sum_{n_1,n_2} \sum_{\substack{h\neq 0 \\ n_1n_2+h>0}} \frac{\Lambda(n_1)\Lambda(n_2)\Lambda(n_1n_2+h)}{n_1n_2} \hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\tfrac{h}{n_1n_2}\Big)\Big)$$
$$\times f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) f_3\Big(\tfrac{\log(n_1n_2)}{\log T}\Big) + O\Big(\frac{\log^3 T}{T}\Big).$$

[2] There is an alternative way to eliminate these terms, which is to use a Tauberian argument to show that σ such that $\sigma(0) = 0$ is a dense enough class of averages to recover the average $\mathbf{1}_{[1,2]}$ – so that in this case $\mathrm{II}, \mathrm{III}, \mathrm{IV}$ and V are all $O(T^{\beta_1+\beta_2}\log T/T^2)$, and the right hand side of (6.16) is in fact $\mathrm{I} + O(T^{\beta_1+\beta_2/2}\log T/T^2)$.

By using Conjecture 6.1.2 to sum in the variable n_1, this expression is

$$\sum_{n_1',n_2}\sum_{h\neq 0}\frac{\mathfrak{S}(n_2h)\big[(n_2,h)=1\big]\Lambda(n_2)}{n_1'n_2}\hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\frac{h}{n_1'n_2}\Big)\Big)$$
$$\times f_1\Big(\tfrac{\log n_1'}{\log T}\Big)f_2\Big(\tfrac{\log n_2}{\log T}\Big)f_3\Big(\tfrac{\log(n_1'n_2)}{\log T}\Big)$$
$$+O\Big(\sum_{n_1\leq T^{\beta_1}}\sum_{n_2\leq T^{\beta_2}}\sum_{|h|\lesssim n_1n_2/T}\frac{1}{n_1^{3/2}n_2}\Big).$$

Note that we can apply Conjecture 6.1.2 because $n_1n_2/T\leq (T^{\beta_2})^{1-\epsilon}$ for some ϵ in this range.

Of all the error terms we have encountered in this proof, this last one is the most fundamental to curtailing our knowledge of correlation functions. It is $O_\epsilon(T^{\beta_2+(1/2+\epsilon)\beta_1}/T)$ for any $\epsilon>0$. Writing the sum over n_2 as a sum over p^λ, making use of Theorem 6.2.2 and otherwise imitating the proof of Theorem 6.1.4,

$$\begin{aligned}\mathrm{I}-\mathrm{V}=&-\sum_{n_1,p^\lambda}\frac{\log p}{n_1p^\lambda}\Big(\log_+(n_1p^\lambda/T)-\log_+(n_1p^{\lambda-1}/T)+O(1)\\&+O\big(\tfrac{\log_+(n_1p^\lambda/T)}{p}\big)\Big)f_1\Big(\tfrac{\log n_1}{\log T}\Big)f_2\Big(\tfrac{\log n_2}{\log T}\Big)f_2\Big(\tfrac{\log(n_1n_2)}{\log T}\Big)+O(\log^2 T)\\=&-\log^3 T\int_0^\infty\int_0^\infty\big((x_1+x_2-1)_+-(x_1-1)_+\big)\\&\times f_1(x_1)f_2(x_2)f_3(x_1+x_2)\,dx_1dx_2+O(\log^2 T).\end{aligned}$$

At first glance this expression seems not to agree with the right hand (random matrix) side of (6.16), where $f_1(x_1)f_2(x_2)f_3(x_1+x_2)\,dx_1dx_2$ is integrated against

$$\begin{aligned}&1-(1-x_1)_+-(1-x_2)_++(1-x_1-x_2)_+\\&=(x_1+x_2-1)_+-(x_1-1)_+-(x_2-1)_+.\end{aligned}$$

So long as $x_2\leq 1$, however, this *does* agree with $(x_1+x_2-1)_+-(x_1-1)_+$, and we have indeed delimited $f_2(x_2)$ to be supported in a region in which $x_2\leq 1$.

Moreover, for β_1, β_2 in the range under consideration, $O(T^{\beta_2+(1/2+\epsilon)\beta_1}/T) = o(1)$. This verifies (6.16) and therefore proves the theorem. $\square$

It is interesting that $(x_1+x_2-1)_+ - (x_1-1)_+$ does not equal $(x_1+x_2-1)_+ - (x_1-1)_+ - (x_2-1)_+$ in general. Though Conjecture 6.1.2 allows us to verify that triple correlation sums of zeta zeros tend to a GUE limit for a wider class of functions Theorem 1.1.6, even an application of Conjecture 6.1.2 is not enough to see all the 'interesting' phenomena of the triple correlation form-factor, on the right hand side of (6.16). This is in contrast to pair correlation. There does not seem to be a rigorously spelled out way to get around this limitation.

In this connection, it is perhaps worth noting at least in passing that

$$\sum_{n_1,n_2}\sum_{h\neq 0}\Big(\mathfrak{S}(n_1n_2h)\big[(n_1n_2,h)=1\big]-1\Big)\frac{\Lambda(n_1)\Lambda(n_2)}{n_1n_2}\hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\frac{h}{n_1n_2}\Big)\Big)$$
$$\times f_1\Big(\tfrac{\log n_1}{\log T}\Big)f_2\Big(\tfrac{\log n_2}{\log T}\Big)f_3\Big(\tfrac{\log(n_1n_2)}{\log T}\Big)$$
$$\sim -\log^3 T\int_{\mathbb{R}_+^2}(1-(1-x_1)_+-(1-x_2)_++(1-x_1-x_2)_+)$$
$$\times f_1(x_1)f_2(x_2)f_3(x_1+x_2)\,dx_1dx_2.$$

Proof of Theorem 6.1.6. Because of the similarity of this proof to the preceding we will not be discursive. The reader should check that to prove the theorem, we must investigate the asymptotic value of two arithmetic quantities:

$$Z_A := \int_{\mathbb{R}^4} f_1\Big(\tfrac{x_1}{\log T}\Big)f_2\Big(\tfrac{x_2}{\log T}\Big)f_3\Big(\tfrac{x_3}{\log T}\Big)g_1\Big(\tfrac{y_1}{\log T}\Big)\hat{\sigma}\Big(\tfrac{T}{2\pi}(y_1-x_1-x_2-x_3)\Big)$$
$$\times\ dz(x_1)dz(x_2)dz(x_3)\,dz(y_1),$$

and

$$Z_B := \int_{\mathbb{R}^4} f_1\Big(\tfrac{x_1}{\log T}\Big) f_2\Big(\tfrac{x_2}{\log T}\Big) g_1\Big(\tfrac{y_1}{\log T}\Big) g_2\Big(\tfrac{y_2}{\log T}\Big) \hat{\sigma}\Big(\tfrac{T}{2\pi}(y_1+y_2-x_1-x_2)\Big)$$
$$\times\ dz(x_1)dz(x_2)\, dz(y_1)dz(y_2).$$

After a computation, random matrix theory predicts (we cite this computation in Corollary 5.10.3 as well, but leave writing it to the reader) that Z_A and Z_B will be respectively

$$\begin{aligned}\sim \log^4 T \int_{\mathbb{R}^3_+} & f_1(x_1)f_2(x_2)f_3(x_3)g_1(x_1+x_2+x_3)\big[(x_1\wedge 1)+(x_2\wedge 1)+(x_3\wedge 1)\\ &-((x_1+x_2)\wedge 1)-((x_1+x_3)\wedge 1)-((x_2+x_3)\wedge 1)\\ &+((x_1+x_2+x_3)\wedge 1)\big]\, dx_1dx_2dx_3.\end{aligned} \tag{6.18}$$

and

$$\begin{aligned}\sim \log^4 T \int_{\mathbb{R}^4_+} & f_1(x_1)f_2(x_2)g_1(y_1)g_2(y_2)\Big[\delta(x_1+x_2-y_1-y_2)\Big(1+(1-x_1)_+\\ &+(1-x_2)_+ + (1-y_1)_+ + (1-y_2)_+ - (1-x_1-x_2)_+ - (1-|x_1-y_1|)_+\\ &-(1-|x_1-y_2|)_+ - 2(1-x_1\wedge x_2\wedge y_1\wedge y_2)_+\Big)\\ &+\delta(x_1-y_1)\delta(x_2-y_2)(x_1\wedge 1)(x_2\wedge 1)\\ &+\delta(x_1-y_2)\delta(x_2-y_1)(x_1\wedge 1)(x_2\wedge 1)\Big]\, dx_1dx_2dy_1dy_2.\end{aligned} \tag{6.19}$$

To prove the theorem it will be sufficient to verify these relationships for functions with support in the regions

- for Z_A, when supp $f_1 \subset (-\infty,\alpha_1)$, supp $f_2 \subset (-\infty,\alpha_2)$, supp $f_3 \subset (-\infty,\alpha_3)$, (and supp $g_1 \subset (-\infty,\beta_1)$ with $\beta_1 = \alpha_1+\alpha_2+\alpha_3$), for constants $\alpha_1 \le \alpha_2 \le \alpha_3$ with $\alpha_1+\alpha_2+\alpha_3/2 \le 1$. In this way we can verify it for $\alpha_1+\alpha_2+\alpha_3 \le 6/5$, and hence $\alpha_1+\alpha_2+\alpha_3+\beta_1 \le 12/5$.

- for Z_B, when supp $f_1 \subset (-\infty, \alpha_1)$, supp $f_2 \subset (-\infty, \alpha_2)$, supp $g_1 \subset (-\infty, \beta_1)$, supp $g_2 \subset (\infty, \beta_2)$, with $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$, $\beta_1 \le \alpha_1 \le \alpha_2 \le \beta_2$, and $\alpha_1 + \alpha_2 + \beta_1 + \beta_2/2 < 2$. (Note that this contains the region of α, β ordered in this way such that $\alpha_1 + \alpha_2 + \beta_1 + \beta_2 < 12/5$. We also will use later that this forces the relationship $\beta_1 + \beta_2/2 < 1$.)

In applying 6.1.2 to verify this for Z_A, our computations proceed exactly as before, and we do not record the details.

On the other hand, to verify this relationship for Z_B, we must apply Conjecture 6.1.3. In this way, one may – with patience and fortitude, but no real cleverness required – see that

$$\begin{aligned} Z_B = &\sum_{n_1,m_1,m_2} \sum_{h\in\mathbb{Z}} \frac{1}{m_1m_2}\Big(\Lambda(m_2)\Lambda\Big(\frac{m_1m_2+h}{n_1}\Big)\big[n_1|(m_1m_2+h)\big]-1\Big) \\ &f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log(m_1m_2/n_1)}{\log T}\Big) g_1\Big(\tfrac{\log m_1}{\log T}\Big) g_2\Big(\tfrac{\log m_2}{\log T}\Big)\hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\frac{h}{m_1m_2}\Big)\Big) \\ &+ O\Big(\frac{T^{\beta_1+\beta_2/2}\log^2 T}{T}\Big) + O\Big(\frac{T^{\beta_2/2+\beta_1+\alpha_1+\alpha_2}}{T^2}\Big). \end{aligned}$$

By summing in the variable m_2 when $h \neq 0$, this sum is equal to, for all $\epsilon > 0$,

$$\begin{aligned} &\sum_{n_1n_2=m_1m_2} \frac{\Lambda(n_1)\Lambda(n_2)\Lambda(m_1)\Lambda(m_2)}{m_1m_2} f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log n_2}{\log T}\Big) g_1\Big(\tfrac{\log m_1}{\log T}\Big) g_2\Big(\tfrac{\log m_2}{\log T}\Big) \\ &+ \sum_{n_1,m_1,m_2}\sum_{h\neq 0} \frac{\Lambda(n_1)\Lambda(m_1)}{m_1m_2}\Big[\big[(n_1,m_1)|h\big]\big[(n_1,m_2)=(n_1,h)\big]\big[(n_1,m_1)=(m_1,h)\big] \\ &\qquad\times \frac{(n_1,m_1)}{n_1}\mathfrak{S}(n_1m_1h) - 1\Big] f_1\Big(\tfrac{\log n_1}{\log T}\Big) f_2\Big(\tfrac{\log(m_1m_2/n_1)}{\log T}\Big) g_1\Big(\tfrac{\log m_1}{\log T}\Big) \\ &\qquad\times g_2\Big(\tfrac{\log m_2}{\log T}\Big)\hat{\sigma}\Big(\tfrac{T}{2\pi}\log\Big(1+\frac{h}{m_1m_2}\Big)\Big) \\ &+ O_\epsilon\Big(\frac{T^{\beta_1+\beta_2(1/2+\epsilon)}}{T}\Big). \end{aligned}$$

We label the two multi-index sums here Σ_1 and Σ_2, respectively in the order they appear above.

The sum Σ_1 is labelled 'diagonal' in much of the literature, and we have evaluated sums of this sort before in this thesis. From the prime number theorem and unique factorization, it takes the asymptotic value

$$\sim \log^4 T \int_{\mathbb{R}_+^2} x_1 x_2 f_1(x_1) f_2(x_2) g_1(x_1) g_2(x_2) \\ + x_1 x_2 f_1(x_1) f_2(x_2) g_1(x_2) g_2(x_1)\, dx_1 dx_2. \tag{6.20}$$

In the sum Σ_2, the summand will be zero unless n_1 and m_1 are prime powers, say p^r and q^s. Breaking the sum up into the cases that one of r or s is more than 1, r and s are both the same prime, and r and s are both distinct primes,

$$\sum_{n_1,m_1,m_2} \sum_{h\neq 0} = \sum_{\substack{p^r,q^s,m_2 \\ r\vee s\geq 2}} \sum_{h\neq 0} + \sum_{\substack{p,q,m_2 \\ p=q}} \sum_{h\neq 0} + \sum_{\substack{p^r,q^s,m_2 \\ p\neq q}} \sum_{h\neq 0},$$

it is easy to see that the first of these terms is

$$O(\log^3 T),$$

while the second is

$$\sum_{p,m_2} \sum_{h\neq 0} \frac{\log^2 p}{pm_2} \big([p|h] \mathfrak{S}(ph) - 1 \big) \hat{\sigma} \Big(\tfrac{T}{2\pi} \log \Big(1 + \frac{h}{pm_2} \Big) \Big) \\ \times f_1\Big(\tfrac{\log p}{\log T}\Big) g_1\Big(\tfrac{\log p}{\log T}\Big) f_2\Big(\tfrac{\log m_2}{\log T}\Big) g_2\Big(\tfrac{\log m_2}{\log T}\Big) \\ \sim -\log^4 T \int_{\mathbb{R}_+^2} x_1 (x_2 - 1)_+ f_1(x_1) f_2(x_2) g_1(x_1) g_2(x_2)\, dx_1 dx_2, \tag{6.21}$$

using Corollary 6.2.5. The third term is

$$\sum_{\substack{p,q\\ distinct}} \sum_{m_2} \sum_{h \neq 0} \frac{\log p \log q}{pqm_2} \big(\big[(pq,h) = 1 \big] \mathfrak{S}(pqh) - 1 \big) \hat{\sigma} \Big(\tfrac{T}{2\pi} \log \Big(1 + \frac{h}{qm_2} \Big) \Big)$$

$$\times f_1\Big(\tfrac{\log p}{\log T}\Big) f_2\Big(\tfrac{\log(qm_2/p)}{\log T}\Big) g_1\Big(\tfrac{\log q}{\log T}\Big) g_2\Big(\tfrac{\log m_2}{\log T}\Big)$$

$$\sim -\log^4 T \int_{\mathbb{R}_+^3} \big((y_1+y_2-1)_+ - (y_1+y_2-x_1-1)_+ - (y_2-1)_+ + (y_2-x_1-1)_+ \big)$$

$$\times f_1(x_1) f_2(y_1+y_2-x_1) g_1(y_1) g_2(y_2)\, dx_1\, dy_1 dy_2, \tag{6.22}$$

using Theorem 6.2.3.

Putting (6.20), (6.21), and (6.22) together, we see that conditionally for f_1, f_2, g_1, and g_2 with the regions of support delimited above,

$$\begin{aligned} Z_B \sim \log^4 T \int_{\mathbb{R}_+^4} & f_1(x_1) f_2(x_2) g_1(y_1) g_2(y_2) \Big[-\delta(x_1+x_2-y_1-y_2) \Big((y_1+y_2-1)_+ \\ & - (x_2-1)_+ - (y_2-1)_+ + (y_2-x_1-1)_+ \Big) \\ & + \delta(x_1-y_1)\delta(x_2-y_2) x_1 \big(x_2 - (x_2-1)_+ \big) \\ & + \delta(x_1-y_2)\delta(x_2-y_1) x_1 x_2 \Big]\, dx_1 dx_2 dy_1 dy_2. \end{aligned} \tag{6.23}$$

This does not seem at first glance to agree with (6.19), but we will show that for functions f, g with the support of f_1 and g_1 lying in $(-\infty, 1)$, the right hand sides of (6.19) and (6.23) agree. This collection of test functions is restricted, but the reader may check that it contains the class of test functions for which we set out to verify (6.19).

Plainly if $x_1, y_1 \leq 1$ the terms

$$\delta(x_1-y_1)\delta(x_2-y_2)x_1\big(x_2-(x_2-1)_+\big) \quad \text{and} \quad \delta(x_1-y_1)\delta(x_2-y_2)(x_1 \wedge 1)(x_2 \wedge 1)$$

agree, and the same is true of

$$\delta(x_1 - y_2)\delta(x_2 - y_1)x_1x_2 \quad \text{and} \quad \delta(x_1 - y_2)\delta(x_2 - y_1)(x_1 \wedge 1)(x_2 \wedge 1).$$

To verify that the coefficients of $\delta(x_1 + x_2 - y_1 - y_2)$ agree, we need to see that

$$\begin{aligned}
\mathcal{E} :=& \Big(1 + (1 - x_1)_+ + (1 - x_2)_+ + (1 - y_1)_+ + (1 - y_2)_+ - (1 - x_1 - x_2)_+ \\
&- (1 - |x_1 - y_1|)_+ - (1 - |x_1 - y_2|)_+ - 2(1 - x_1 \wedge x_2 \wedge y_1 \wedge y_2)_+\Big) \\
&+ \Big((y_1 + y_2 - 1)_+ - (x_2 - 1)_+ - (y_2 - 1)_+ + (y_2 - x_1 - 1)_+\Big)
\end{aligned}$$

is 0 when $x_1 + x_2 = y_1 + y_2$ and $x_1, y_1 \leq 1$. Using that $(1 - x_1)_+ = 1 - x_1$ and $(1 - y_1)_+ = 1 - y_1$ in this region, and the identity

$$(w - 1)_+ - (1 - w)_+ = w - 1,$$

true for all non-negative w, one sees that in this region,

$$\mathcal{E} = -y_1 - y_2 + |y_2 - x_1| + |y_1 - x_1| + 2x_1 \wedge x_2 \wedge y_1 \wedge y_2.$$

By inspection one can see that when $x_1 + x_2 = y_1 + y_2$, the right hand side is 0.

This verifies (6.19) for the desired class of test functions, and completes the proof. □

Again, one does not see all the interesting phenomena in (6.19) and (6.20), and one cannot hope to by using Conjecture 6.1.3 alone in this manner. At present, it is only from the introduction of what seem to the author to be somewhat ad hoc devices that one can recover the full part of the right hand sides of (6.19) and (6.20).

One can continue on in this fashion, but perhaps to somewhat diminishing returns. The combinatorial part of the analysis becomes increasingly cumbersome. There is one class of test functions lying outside the range of Theorem 1.1.6, but for which the Hardy-Littlewood conjectures can be applied and the combinatorics remains relatively simple. That in higher correlations the combinatorics of this case is nicer was noticed by Bogomolny & Keating also.

Theorem 6.3.1. *Conditioned on Conjecture 6.1.2 and RH, for all* $j \geq 1$ *and Schwartz* $\eta : \mathbb{R}^j \to \mathbb{R}$ *such that* $\operatorname{supp} \hat{\eta} \subset \{(x_1, ..., x_j) \in \mathbb{R}_+^{j-1} \times R_- : x_1 + \cdots + x_{j-1} - x_1 \vee \cdots \vee x_{j-1}/2 < 1\}$

$$\lim_{T\to\infty} \frac{1}{T} \int_T^{2T} \sum_{\substack{\gamma_1,...,\gamma_j \\ \text{distinct}}} \eta\Big(\tfrac{\log T}{2\pi}(\gamma_1 - t), \cdots, \tfrac{\log T}{2\pi}(\gamma_j - t)\Big)\, dt$$
$$= \int_{\mathbb{R}^j} \eta(x) \det_{j\times j} \big(K(x_i - x_j)\big)\, d^j x.$$

Note that in this case our result is unsatisfying owing to the peculiar region to which our test function's support has been restricted. We give only a sketch of a proof.

To prove Theorem 6.3.1, on the arithmetic side we make use of the same sort of analysis used to prove Theorem 6.1.5 or evaluate Z_A in the proof of Theorem 6.1.6. We are concerned with the likelihood that a product of $k-1$ primes is close to a k^{th} prime.

On the other hand, one must also make use of an identity belonging to random matrix theory: that for test functions $\eta_1, ..., \eta_k$ with $\operatorname{supp} \hat{\eta}_1, ..., \operatorname{supp} \hat{\eta}_{k-1} \subset \mathbb{R}_+$,

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell=1}^{k} \sum_i \eta_\ell(x_i) = \int_{\mathbb{R}^{k-1}} \hat{\eta}_1(\xi_1) \cdots \hat{\eta}_{k-1}(\xi_{k-1}) \hat{\eta}_k(-x_1 - \cdots - x_{k-1})$$
$$\times \sum_{\emptyset \subseteq S \subseteq [k-1]} (-1)^{|S|} (1 - x_S)_+ \, d^k x. \tag{6.24}$$

We outline a proof of this identity. We note the formula that for $u_1, .., u_r$ with supp $\hat{u}_1, .., \hat{u}_{r-1} \subset \mathbb{R}_+$,

$$\int_{\mathbb{R}^r} u_1(x_1) \cdots u_r(x_r) \det_{r\times r} \big(K(x_i - x_j)\big)\, d^r x \tag{6.25}$$
$$= (-1)^{r-1}(r-1)! \int_{\mathbb{R}^{r-1}} \hat{u}_1(\xi_1) \cdots \hat{u}_{r-1}(\xi_{r-1}) \hat{u}_r(-\xi_{[r-1]})(1 - \xi_{[r-1]})_+ \, d^{r-1}\xi.$$

This formula is perhaps most easily proven by making use of the identity (used extensively in [63]),

$$\int_{\mathbb{R}^r} K(x_1 - x_2)K(x_2 - x_3) \cdots K(x_{r-1} - x_j)K(x_r - x_1)\, d^r x$$
$$= \delta(\xi_{[r]}) \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(t) \prod_{j=1}^{r-1} \mathbf{1}_{[0,1]}(t + \xi_1 + \cdots + \xi_j)\, dt.$$

From (6.25), by expanding the multi-linear statistic into a sum of determinantal integrals, one may see that

$$\mathop{\mathbf{E}}_{\mathcal{S}} \prod_{\ell}^{k} \sum_{i} \eta_\ell(x_i) = \sum_{\substack{\pi\in\Pi_{[k]} \\ k\in\pi_1}} (-1)^{|\pi|-1}(|\pi|-1)!$$
$$\times \int_{\mathbb{R}^{k-1}} \hat{\eta}_1(\xi_1) \cdots \hat{\eta}_{k-1}(\xi_{k-1}) \hat{\eta}_k(-x_{[k-1]})(1 - x_{[k]\setminus\pi_1})_+ \, d^{k-1}x$$
$$= \sum_{\emptyset\subseteq S\subseteq[k-1]} \bigg(\sum_{\pi'\in\Pi_S} (-1)^{|\pi'|}\big(|\pi'|\big)! \bigg)$$
$$\times \int_{\mathbb{R}^{k-1}} \hat{\eta}_1(\xi_1) \cdots \hat{\eta}_{k-1}(\xi_{k-1}) \hat{\eta}_k(-x_{[k-1]})(1 - x_S)_+ \, d^{k-1}x$$

where Π_S is the collection of all partitions $\pi = \{\pi_1, ...\}$ of a set S.

By applying the elementary identity

$$\sum_{\pi\in\Pi_{[n]}} (-1)^{|\pi|}\big(|\pi|\big)! = (-1)^n,$$

we obtain (6.24).

APPENDIX A

Counts of almost primes in long intervals

We made us of the following estimates earlier; as in the rest of the document, we require the Riemann hypothesis for their proof.

Theorem A.0.2. *For fixed j with $\alpha \in (1/2, 1)$ and $|\sigma + it - 1| \geq 1/4$*

$$\frac{\zeta^{(j)}}{\zeta}\Big(\sigma + it\Big) = O\big(\frac{\log^j(|t| + 2)}{(\sigma - 1/2)^j}\big)$$

The region above are chosen so that they do not include the singularity at $\alpha + it = 1$.

Theorem A.0.3. *For fixed j,*

$$\psi_j(x) = \int_0^x j \log^{j-1} y \, dy + O_j\big(x^{1/2} \log^{2j+1} x\big).$$

We prove Theorem A.0.3 on the basis of Theorem A.0.2. The error term bound $O(x^{1/2} \log^{2j+1} x)$ is not optimal; by refining our technique (by using the mean value estimates in this preprint for instance), one can obtain an error term of $O(x^{1/2} \log^{j+1} x)$, an estimate on the level of the classical von Koch estimate for $j = 1$. It is likely that even this estimate is not optimal (for $j = 1$ for instance Montgomery has conjectured the error term is of order $x^{1/2}(\log \log \log x)^2 / \log x)$ but either estimate will at any rate be sufficient for our purposes.

Proof of Theorem A.0.2. We have for $|t| \geq 1$ and $\sigma > 1/2$

$$\begin{aligned}\frac{\zeta'}{\zeta}\Big(\sigma + it\Big) &= \sum_{|\gamma - t| \leq 1} \frac{1}{\sigma + it - (1/2 + i\gamma)} + O\big(\log(|t|+2)\big)\\ &= O\Big(\frac{\log(|t|+2)}{\sigma - 1/2}\Big) + O\big(\log(|t|+2)\big),\end{aligned}$$

with the first line following from Lemma 12.1 of [56] (essentially taking a logarithmic derivative of a Hadamard product), and the second from bounding the number of zeros that lie in a unit interval at height t. We show inductively that

$$\frac{\zeta^{(j)}}{\zeta}\Big(\sigma + it\Big) = O_j\big(\big((\sigma - 1/2)^{-1} \vee 1\big)^j \log^j(|t|+2)\big)$$

for $|t| \geq 1$; we have just demonstrated it for $j = 1$. Suppose we have the estimate for $\zeta^{(j-1)}/\zeta$. Then for $s = \sigma + it$, $|t| \geq 2$ and $\delta = (\sigma - 1/2)^{-1} \wedge 1$,

$$\begin{aligned}\frac{\zeta^{(j)}}{\zeta}(s) &= \Big(\frac{\zeta^{(j-1)}}{\zeta}\Big)'(s) + \frac{\zeta'}{\zeta}(s)\frac{\zeta^{(j-1)}}{\zeta}(s)\\ &= \frac{1}{2\pi i}\int_{|z-s|=\delta} \frac{\zeta^{(j)}}{\zeta}(z)\frac{dz}{(z-s)^2} + O_j\big(\big((\sigma - 1/2)^{-1} \vee 1\big)^j \log^j(|t|+2)\big)\\ &= O_j\big(\big((\sigma - 1/2)^{-1} \vee 1\big)^j \log^j(|t|+2)\big).\end{aligned}$$

For $t \in (1,2)$ clearly

$$\frac{\zeta^{(j)}}{\zeta}(s) = O_j(1),$$

which completes our induction.

As moreover for $|t| \in (0,1)$ but $|\sigma + it - 1| \geq 1/4$,

$$\frac{\zeta^{|j|}}{\zeta}\Big(\sigma + it\Big) = O_j(1),$$

we have proved the theorem. ☐

To prove Theorem A.0.3 we reference Lemma 3.12 from Titchmarsh's tract

[79]:

Lemma A.0.4 (Lemma 3.12 of [79])**.** *Let*

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \quad (\sigma > 1)$$

where $a_n = O(\rho(n))$*,* $\rho(n)$ *non-decreasing, and*

$$\sum_{n=1}^{\infty} \frac{|a_n|}{n^\sigma} = O\Big(\frac{1}{(\sigma-1)^\alpha}\Big),$$

as $\sigma \to 1$*. Then if* $c > 0$*,* $\sigma + c > 1$*,* x *not an integer, and* N *is the integer nearest to* x*,*

$$\begin{aligned}\sum_{n \le x} \frac{a_n}{n^s} =& \frac{1}{2\pi i} \int_{c-iT}^{c+iT} f(s+w) \frac{x^w}{w}\, dw + O\Big(\frac{x^c}{T(\sigma+c-1)^\alpha}\Big) \\ &+ O\Big(\frac{\rho(2x) x^{1-\sigma} \log x}{T}\Big) + O\Big(\frac{\rho(N) x^{1-\sigma}}{T|x-N|}\Big).\end{aligned}$$

Proof of Theorem A.0.3. Using the lemma with $a_n = \Lambda_j(n)$, we have

$$f(s) = (-1)^j \frac{\zeta^{(j)}}{\zeta}(s),$$

$$a_n = O(\log^j n)$$

$$\sum_{n=1}^{\infty} \frac{|a_n|}{n^\sigma} = (-1)^j \frac{\zeta^{(j)}}{\zeta}(\sigma) \sim \frac{j!}{(\sigma-1)^j}.$$

Setting $s = 1/2$, $c = 3/4$, and $T = x^2$ for $x = N + 1/2$, we have

$$\begin{aligned}
\sum_{n\le x}\frac{\Lambda_j(n)}{\sqrt{n}} =& \frac{1}{2\pi i}\int_{3/4-iT}^{3/4+iT}(-1)^j\frac{\zeta^{(j)}}{\zeta}\big(\tfrac{1}{2}+w\big)\frac{x^w}{w}\,dw + o(1)\\
&= \operatorname*{Res}_{w=1/2}\Big(\frac{j!}{(w-1/2)^j}\frac{x^w}{w}\Big)\\
&\quad + \frac{1}{2\pi i}\left(\int_{1/\log T+iT}^{3/4+iT}+\int_{1/\log T-iT}^{1/\log T+iT}+\int_{3/4-iT}^{1/\log T-iT}\right)(-1)^j\frac{\zeta^{(j)}}{\zeta}\big(\tfrac{1}{2}+w\big)\frac{x^w}{w}\,dw + o(1)\\
&= \int_0^x \frac{j\log^{j-1}y}{\sqrt{y}}\,dy + O\left(\int_{-T}^{T}\frac{\log^j T\log^j(|t|+2)}{\big|\frac{1}{\log T}+it\big|}\,dt\right)+o(1)\\
&= \int_0^x \frac{j\log^{j-1}y}{\sqrt{y}}\,dy + O(\log^{2j+1}x).
\end{aligned}$$

Because $\log^{2j+1} x$ is a slowly growing function we obtain this for all x, not only $x = N + 1/2$. The theorem then follows from partial integration. $\square$

APPENDIX B

The sine-kernel determinantal point process

A point process $(X, \mathfrak{F}, \mathbb{P})$ on $\mathbb{R}$ is a probability measure $\mathbb{P}$ on the σ-algebra $(X, \mathfrak{F})$, where X is the set of locally finite configurations of sequences of real numbers:

$$X := \{\xi = ((..., \xi_{-1}, \xi_0, \xi_1, ...)) : \xi_i \in \mathbb{R}\ \forall i \in \mathbb{Z},$$
$$\text{and for any compact } K \subset \mathbb{R},\ \#_K(\xi) = \#\{i : \xi_i \in K\} \leq \infty\}$$

and $\mathfrak{F}$ is the σ-algebra with a basis consisting of the cylinder sets

$$C_n^B := \{\xi \in X :\ \#_B(\xi) = n\}$$

where $n = 0, 1, 2, ...$ and B is any Borel subset of $\mathbb{R}$ with compact closure.

In this way for any Borel $B_1, ..., B_K$, the expectation

$$\mathbf{E} \sum_{j_1,...,j_k} \mathbf{1}_{B_1}(\xi_{j_1}) \cdots \mathbf{1}_{B_k}(\xi_{j_k}) = \mathop{\mathbf{E}}_{(X,\mathfrak{F},\mathbb{P})} \#_{B_1}(\xi) \cdots \#_{B_k}(\xi)$$

can be evaluated and, from approximation by simple functions, for any measurable $\eta : \mathbb{R}^k \to \mathbb{R}$, the expectation

$$\mathbf{E} \sum_{j_1,...,j_k} \eta(\xi_{j_1}, ..., \xi_{j_k})$$

can be evaluated as well. By a combinatorial sieving procedure, so too can

$$\mathbf{E} \sum_{\substack{j_1,\ldots,j_k \\ \text{distinct}}} \eta(\xi_{j_1}, ..., \xi_{j_k}).$$

For instance,

$$\mathbf{E} \sum_{j_1 \neq j_2} \eta(\xi_{j_1}, \xi_{j_2}) = \mathbf{E} \sum_{j_1, j_2} \eta(\xi_{j_1}, \xi_{j_2}) - \mathbf{E} \sum_{j} \eta(\xi_j, \xi_j).$$

This defines a measure $d\mu_k$ on $\mathbb{R}^k$, called the k-level joint intensity measure:

$$\mathbf{E} \sum_{\substack{j_1,\ldots,j_k \\ \text{distinct}}} \eta(\xi_{j_1}, ..., \xi_{j_k}) = \int_{\mathbb{R}^k} \eta(x_1, ..., x_k)\, d\mu_k(x_1, ..., x_k).$$

These should be thought of as a generalization of moments for random variables. By no means do all collections of measures $\{d\mu_1, d\mu_2, ...\}$ on $\mathbb{R}^1, \mathbb{R}^2, ...$ correspond to the joint intensity of a point process, but in the case that

$$d\mu_k(x_1, ..., x_k) = \det_{k \times k} \big(K(x_i - x_j)\big)\, dx_1 dx_2 \cdots dx_k$$

it is known that there exists a unique point process, labeled 'the sine-kernel determinantal point process', with these joint intensities. Details of its construction and a more general account of the theory of determinantal point processes can be found in [70].

APPENDIX C

Some random matrix statistics

We recall without proof some essential facts from symmetric function theory. A more complete introduction with proofs of the facts cited below is found in [9]. The references [75] and [29] are also useful, the latter being a streamlined introduction from the perspective of random matrices and analytic number theory.

In the variables $\omega_1, ..., \omega_N$, recall the definitions that for $k = 0, 1..., N$

$$e_k = e_k(\omega_1, ..., \omega_N) := \sum_{j_1 < \cdots < j_k} \omega_{j_1} \cdots \omega_{j_k}$$

with $e_o := 1$, while for $k = 0, 1, ...,$

$$h_k = h_k(\omega_1, ..., \omega_N) := \sum_{j_1 \leq \cdots \leq h_k} \omega_{j_1} \cdots \omega_{j_k}.$$

with $h_0 := 1$.

A partition $\lambda = (\lambda_1, \lambda_2, ...)$ is a sequence of non-negative integers $\lambda_1 \geq \lambda_2 \geq ...$ such that for large enough n, $\lambda_{n+1} = 0$. λ may then be thought of as just $(\lambda_1, ..., \lambda_n)$, and the largest n such that $\lambda_n \neq 0$ is called the *length* of λ.

If the length of λ is no more than N, we define the Schur function s_λ by

$$s_\lambda = s_\lambda(\omega_1, ..., \omega_N) := \frac{\det_{N\times N}\left(\omega_i^{\lambda_j + N - j}\right)}{\det_{N\times N}\left(\omega_i^{N-j}\right)}.$$

The functions also satisfy

$$s_\lambda = \sum_T (\omega_{t[1,1]}\omega_{t[1,2]} \cdots \omega_{t[1,\lambda_1]}) \cdots (\omega_{t[2,1]} \cdots \omega_{t[2,\lambda_2]})(\omega_{t[n,1]} \cdots \omega_{t[n,\lambda_n]}),$$

where n is the length of λ and the sum is over all so-called *semi-standard Young tableau* of shape λ, numbers

$$\begin{array}{lllll} t[1,1] & t[1,2] & \dots & \dots & t[1,\lambda_1] \\ t[2,1] & t[2,2] & \dots & t[2,\lambda_2] & \\ \vdots & \vdots & \ddots & & \\ t[n,\lambda_n] & \dots & t[n,\lambda_n] & & \end{array}$$

with $t[i,j] \in \{1, 2, ..., N\}$ for all i, j, so that in rows numbers from left-to-right are non-decreasing:

$$t[i,1] \leq t[i,2] \leq \cdots \leq t[i,\lambda_i],$$

while in columns

$$\begin{array}{l} t[1,j] \\ < t[2,j] \\ \quad \vdots \\ < t[\cdot,j] \end{array}$$

numbers are strictly increasing. For instance, when $N = 3$, the semi-standard Young tableaux of the partition $(2, 1)$ are

$$\begin{array}{cc} 1 & 1 \\ 2 & \end{array} \quad \begin{array}{cc} 1 & 2 \\ 2 & \end{array} \quad \begin{array}{cc} 1 & 3 \\ 2 & \end{array} \quad \begin{array}{cc} 1 & 1 \\ 3 & \end{array} \quad \begin{array}{cc} 1 & 2 \\ 3 & \end{array} \quad \begin{array}{cc} 1 & 3 \\ 3 & \end{array} \quad \begin{array}{cc} 2 & 2 \\ 3 & \end{array} \quad \begin{array}{cc} 2 & 3 \\ 3 & \end{array}$$

and

$$s_{(2,1)}(\omega_1,\omega_2,\omega_3) = \omega_1^2\omega_2 + \omega_1\omega_2^2 + \omega_1^2\omega_3 + \omega_1\omega_3^2 + \omega_2^2\omega_3 + \omega_2\omega_3^2 + 2\omega_1\omega_2\omega_3.$$

For us the importance of Schur functions is that

$$\int_{U(N)} s_{\lambda_1}(\omega_1,...,\omega_N)\overline{s_{\lambda_2}(\omega_1,...,\omega_N)}\, du = \delta_{\lambda_1=\lambda_2}, \tag{C.1}$$

for all partitions λ_1, λ_2 of length no more than N, where $\omega_1,...,\omega_N$ are the eigenvalues of $U(N)$; a proof of this fact can be found in [9] or [29].

Finally, let us introduce the abbreviation

$$\lambda = (\lambda_1,...,\lambda_j,\underbrace{1,...,1}_{k}) = (\lambda_1,...,\lambda_j,1^k).$$

This generalizes in the obvious way, but the above usage is all that we will make use of.

Note that in the case that the variables $\omega_1,...,\omega_N$ are eigenvalues of a random unitary matrix,

$$Z(\beta) = \sum_{n=0}^{N}(-1)^n e_n e^{-\beta n} \tag{C.2}$$

$$Z^{(j)}(\beta) = (-1)^j \sum_{n=0}^{N}(-1)^n n^j e_n e^{-\beta n} \tag{C.3}$$

$$\frac{1}{Z(\beta)} = \sum_{m=0}^{\infty} h_m e^{-\beta m}. \tag{C.4}$$

Proof of Lemma 5.2.7. All symmetric functions in this section are in the variables $\omega_1,...,\omega_N$, eigenvalues of a random unitary matrix. We show that

$$H_j(r) = \sum_{\nu=1}^{r\wedge N}(-1)^\nu(\nu^j - (\nu-1)^j)s_{(r-\nu+1,1^{\nu-1})}. \tag{C.5}$$

Lemma 5.2.7 then follows from the Schur orthogonality relation (C.1).

In the first place, from (5.83) and (C.3) and (C.4), we have by pairing coefficients,

$$H_j(r) = \sum_{\nu=1}^{r\wedge N} (-1)^\nu \nu^j e_\nu h_{r-\nu}. \tag{C.6}$$

But note that for $n \geq 1$, $m \geq 0$,

$$e_n h_m = \sum_{\alpha_1 < \cdots < \alpha_n} \sum_{\beta_1 \leq \cdots \leq \beta_m} \omega_{\alpha_1} \cdots \omega_{\alpha_n} \omega_{\beta_1} \cdots \omega_{\beta_m},$$

where if $m = 0$ the sum over β is understood to be empty.

In the case that $m \neq 0$, breaking the sum into two parts depending on whether $\alpha_1 \leq \beta_1$ or $\beta_1 < \alpha_1$, this sum is

$$\begin{aligned} e_n h_m = & \sum_{\alpha_1 \leq \beta_1 \cdots \leq \beta_m} \omega_{\alpha_1} \omega_{\beta_1} \cdots \omega_{\beta_n} \sum_{\substack{\alpha_2, \ldots, \alpha_n \\ \text{such that} \\ \alpha_1 < \alpha_2 < \cdots < \alpha_n}} \omega_{\alpha_2} \cdots \omega_{\alpha_n} \\ & + \sum_{\beta_1 \leq \cdots \leq \beta_m} \omega_{\beta_1} \cdots \omega_{\beta_m} \sum_{\substack{\alpha_1, \ldots, \alpha_n \\ \text{such that} \\ \beta_1 < \alpha_1 < \cdots < \alpha_n}} \omega_{\alpha_1} \cdots \omega_{\alpha_n} \\ = & s_{(m+1, 1^{n-1})} + s_{(m, 1^n)} \end{aligned}$$

Provided we adopt the convention that $s_{(0,1^n)} = 0$, this remains true when $m = 0$.

On the other hand, if $n = N$,

$$\begin{aligned} e_n h_m &= \omega_1 \cdots \omega_N \sum_{\beta_1 \leq \cdots \leq \beta_m} \omega_{\beta_1} \cdots \omega_{\beta_m} \\ &= s_{(m+1, 1^{N-1})} \end{aligned}$$

since in this case, for any indices $\beta_1, ..., \beta_m$ of the sum, $1 \leq \beta_1$.

Therefore, for all $j \geq 1$ and $r \geq 1$,

$$
\begin{aligned}
H_j(r) &= \sum_{\nu=1}^{r\wedge N} (-1)^{\nu} \nu^j \big(s_{(r-\nu+1,1^{\nu-1})} + \delta_{\nu \neq r,N} s_{(r-\nu,1^{\nu})} \big) \\
&= \sum_{\nu=1}^{r\wedge N} (-1)^{\nu} (\nu^j - (\nu-1)^j) s_{(r-\nu+1,1^{\nu-1})},
\end{aligned}
$$

as claimed. Applying (C.1) to this proves the lemma. □

We mentioned earlier that this result is equivalent to the 2×2 ratio theorem for the unitary group.

Theorem C.0.5. *For A, B, C, D complex numbers with $|C|, |D| \leq 1$, for $N \geq 1$,*

$$
\begin{aligned}
&\int_{U(N)} \frac{\det(1-Au)\det(1-Bu^{-1})}{\det(1-Cu)\det(1-Du^{-1})}\, du \\
&= \frac{(1-BC)(1-AD)}{(1-AB)(1-CD)} + (AB)^N \frac{(1-CA^{-1})(1-DB^{-1})}{(1-(AB)^{-1})(1-CD)}. \qquad \text{(C.7)}
\end{aligned}
$$

Proof. We let

$$A = e^{-\beta_1 + s_1}$$

$$B = e^{-\beta_2 + s_2}$$

$$C = e^{-\beta_1}$$

$$D = e^{-\beta_2}$$

with $\Re\beta_1, \Re\beta_2 > 0$. (There is no real loss of generality to assume A, B, C and D are real, and this would make less to keep track of in the argument that follows if the reader desires.)

Then the left hand side of (C.7) is

$$\int_{U(N)} \frac{Z(\beta_1 - s_1)}{Z(\beta_1)} \overline{\left(\frac{Z(\overline{\beta_2} - \overline{s_2})}{Z(\overline{\beta_2})} \right)} \, du$$
$$= \sum_{j,k=0}^{\infty} \frac{s_1^j}{j!} \frac{s_2^k}{k!} \int_{U(N)} \left((-1)^j \frac{Z^{(j)}}{Z}(\beta_1) \right) \overline{\left((-1)^k \frac{Z^{(k)}}{Z}(\overline{\beta_2}) \right)} \, du$$
$$= 1 + \sum_{j,k=1}^{\infty} \sum_{r,s=1}^{\infty} \frac{s_1^j}{j!} \frac{s_2^k}{k!} e^{-r\beta_1 - s\beta_2} \int_{U(N)} H_j(r) \overline{H_j(s)} \, du.$$

Here we have used that,

$$\int_{U(N)} \frac{Z^{(0)}}{Z}(\beta_1) \overline{\frac{Z^{(0)}}{Z}(\overline{\beta_2})} \, du = 1$$

and slightly less trivially that

$$\int_{U(N)} \frac{Z^{(j)}}{Z}(\beta) \, du = 0$$

for $j \geq 1$, which follows from, for instance, equations (C.3), (C.4) (with all exponents of ω being positive in both identities) and that for any θ, $u \mapsto e^{i2\pi\theta} u$ preserves the Haar measure of the unitary group.

On the other hand, after rearranging the right hand side of (C.7), it is just

$$\begin{aligned}
&1+\frac{(1-e^{-s_1})(1-e^{-s_2})}{1-e^{-\beta_1-\beta_2}}e^{s_1+s_2-\beta_1-\beta_2}\left(\frac{1-(e^{s_1+s_2-\beta_1-\beta_2})^N}{1-e^{s_1+s_2-\beta_1-\beta_2}}\right)\\
&=1+\frac{(1-e^{-s_1})(1-e^{-s_2})}{1-e^{-\beta_1-\beta_2}}\sum_{\nu=1}^{N}e^{\nu(s_1+s_2)}e^{-\nu(\beta_1+\beta_2)}\\
&=1+(1-e^{-s_1})(1-e^{-s_2})\sum_{\nu=1}^{N}e^{\nu(s_1+s_2)}\sum_{r=\nu}^{\infty}e^{-(\beta_1+\beta_2)r}\\
&=1+\sum_{r=1}^{\infty}e^{-(\beta_1+\beta_2)r}\sum_{\substack{\nu\le r\\ \nu\le N}}\left[e^{\nu s_1}-e^{(\nu-1)s_1}\right]\left[e^{\nu s_2}-e^{(\nu-1)s_2}\right]\\
&=1+\sum_{j,k=1}^{\infty}\frac{s_1^j}{j!}\frac{s_2^k}{k!}\sum_{r=1}^{\infty}e^{-(\beta_1+\beta_2)r}\sum_{\nu=1}^{r\wedge N}\big(\nu^j-(\nu-1)^j\big)\big(\nu^k-(\nu-1)^k\big).
\end{aligned}$$

By pairing coefficients, Lemma 5.2.7 is therefore equivalent to the right and left hand sides of (C.7) being equal. □

One may wonder whether this same analysis can be used at this point to show that the GUE Conjecture implies the ratio conjecture (5.25) for the zeta function. The answer is not quite; one would require some uniformity in j and k for our covariance asymptotics. In particular, an estimate which would be sufficient to deduce Farmer's ratio conjecture from the GUE Conjecture is

Conjecture C.0.6. *For all fixed positive constants A,*

$$\left(\frac{1}{T}\int_0^T\left|\frac{\zeta^{(j)}}{\zeta}\left(\tfrac{1}{2}+\tfrac{A}{\log T}+it\right)\right|^2dt\right)^{1/2j}\lesssim_A\log T \tag{C.8}$$

uniformly for all $j\geq 1$ and $T\geq 1$.

Of course one could only hope to prove such an estimate on RH. On RH, a weaker estimate in which the right hand side grows in j such a bound can be obtained from the methods of section 5.11. It would be interesting if these methods can be refined to produce the conjectured uniformity in j.

References

[1] G.E. Andrews, R. Askey and R. Roy. *Special Functions,* Encyclopedia of Mathematics and its Applications, vol. 71, Cambridge University Press (1999).

[2] Bogomolny, E. B., and J. P. Keating. "Random matrix theory and the Riemann zeros. I. Three-and four-point correlations." *Nonlinearity* 8.6 (1999): 1115.

[3] Bogomolny, E. B., and J. P. Keating. "Random matrix theory and the Riemann zeros II: n-point correlations." *Nonlinearity* 9.4 (1999): 911.

[4] Bogomolny, E.B. and J. P. Keating. "Gutzwiller's trace formula and spectral statistics: beyond the diagonal approximation", *Physical Review Letters* 77, no. 8 (1996): 1472-1475.

[5] Bourgade, P., "Mesoscopic fluctuations of the zeta zeros". *Probab. Theory and Related Fields* 148 (2010), no. 3-4, 479-500.

[6] Bourgade, P., and J.P. Keating. "Quantum chaos, random matrix theory, and the Riemann ζ-function." *Seminaire Poincare* XIV (2010): 115-153.

[7] Bourgade, P., and J. Kuan. "Strong Szego asymptotics and zeros of L-functions", *http://arxiv.org/abs/1203.5328*, to appear in *Communications in Pure and Applied Math.*

[8] Berry, M. "Semiclassical formula for the number variance of the Riemann zeros", *Nonlinearity* 1, 399-407 (1988)

[9] Bump, D., *Lie groups.* Vol. 225 Graduate Texts in Mathematics. Springer, 2004.

[10] Bump, D., and A. Gamburd. "On the averages of characteristic polynomials from classical groups." *Communications in mathematical physics* 265, no. 1 (2006): 227-274.

[11] Chan, T.H. "More precise pair correlation of zeros and primes in short intervals", *J. London Math. Soc.* 68 (2) (2003) 579 -598

[12] Conrey, J.B. *Notes on eigenvalue distributions for the classical compact groups.* in "Recent Perspectives in Random Matrix Theory and Number Theory", LMS, Lecture Note Series 322 (2005) 111 - 145.

[13] Conrey, J.B., D.W. Farmer, and M.R. Zirnbauer. "Autocorrelation of ratios of L-functions", *Comm. Number Theory and Physics*, 2 no. 3 (2008): 593-636.

[14] Conrey, J. B., D.W. Farmer, and M.R. Zirnbauer, "Howe pairs, supersymmetry, and ratios of random characteristic polynomials for the unitary groups U (N)." *arXiv preprint arXiv:math-ph/0511024: 0511024.* (2005)

[15] Conrey, J. B., P. J. Forrester, and N. C. Snaith. "Averages of ratios of characteristic polynomials for the compact classical groups." *International Mathematics Research Notices* 2005, no. 7 (2005): 397-431.

[16] Conrey, J. B., and N. C. Snaith. "Applications of the L-functions ratios conjectures." *Proceedings of the London Mathematical Society* 94.3 (2007): 594-646.

[17] J.B. Conrey and N.C. Snaith. "Correlations of eigenvalues and Riemann zeros", *Comm. Number Theory and Phyics*, 2(3) (2008), 477-536.

[18] Cramér, H. "On the order of magnitude of the difference between consecutive prime numbers." *Acta Arithmetica* 2, no. 1 (1936): 23-46.

[19] Diaconis, P., and S. Evans. "Linear Functionals of Eigenvalues of Random Matrices", *Trans. Amer. Math. Soc.*, 353 (2001), 2615-2633.

[20] Diaconis, P., and M. Shahshahani. "On the eigenvalues of random matrices." *Journal of Applied Probability* (1994): 49-62.

[21] Erdös, P. "On a new method in elementary number theory which leads to an elementary proof of the prime number theorem." *Proceedings of the National Academy of Sciences of the United States of America* 35, no. 7 (1949): 374-384.

[22] Erdös, L., and H.T. Yau. "A comment on the Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices." *arXiv preprint arXiv:1201.5619* (2012).

[23] Faifman, D., and Z. Rudnick. *Statistics of the zeros of zeta functions in families of hyperelliptic curves over a finite field.* Compositio Mathematica 146 (1), (2010) 81-101.

[24] Farmer, D. W. "Long mollifiers of the Riemann zeta-function." *Mathematika* 40.01 (1993): 71-87.

[25] Farmer, D. W., S. M. Gonek, Y. Lee, and S.J. Lester. "Mean values of $\zeta'/\zeta(s)$, correlations of zeros and the distribution of almost primes." to appear, *The Quarterly Journal of Mathematics* (2013).

[26] Fujii, A. "On the Berry conjecture." *Kyoto Journal of Mathematics* 37.1 (1997): 55-98.

[27] Fujii, A. "On the zeros of Dirichlet L-functions I", *Trans. Amer. Math. Soc.*, 196 (1974), 225-235.

[28] A. Fujii, "Explicit formulas and oscillations", *Emerging Applications of Number Theory*, Springer (1999), 219-267.

[29] Gamburd, A. "Some applications of symmetric functions theory in random matrix theory." *LMS Lecture Note Series* 341 (2007): 143-170.

[30] Gallagher, P. X., and J. H. Mueller. "Primes and zeros in short intervals." *J. reine angew. Math* 303.304 (1978): 205-220.

[31] Guinand, A. P. "A summation formula in the theory of prime numbers." *Proc. London Math. Soc* 2.50 (1948): 107-119.

[32] Goldston, D. A. "On the pair correlation conjecture for zeros of the Riemann zeta-function." *J. reine angew. Math* 385 (1988): 24-40.

[33] Goldston, D.A. "Notes on pair correlation of zeros and prime numbers." in *Recent Perspectives in Random Matrix Theory and Number Theory*, LMS Lecture Note Series 322, Edited by F. Mezzadri and N. C. Snaith, Cambridge University Press (2005): 79-110.

[34] Goldston, D. A., S. M. Gonek, and H. L. Montgomery. "Mean values of the logarithmic derivative of the Riemann zeta-function with applications to primes in short intervals." *J. reine angewa. Math* (2001): 105-126.

[35] Goldston, D. A., and H. L. Montgomery. "Pair correlation of zeros and primes in short intervals." *Progress in Math* 70 (1987): 183-203.

[36] Granville, A. "Harald Cramér and the distribution of prime numbers." *Scandinavian Actuarial Journal* 1995, no. 1 (1995): 12-28.

[37] Hardy, G.H., and J.E. Littlewood. "On some problems of "partitio numerorum" III: On the expression of a number as a sum of primes". *Acta Math.* 44 (1923): 1-70.

[38] Hejhal, D. A. "On the triple correlation of zeros of the zeta function." *International Mathematics Research Notices* 1994.7 (1994): 293-302.

[39] Ghaith, A.H. and Odlyzko, A. M. "The zeta function on the critical line: Numerical evidence for moments and random matrix theory models." *Mathematics of Computation* 81.279 (2011): 1723.

[40] Hough J.B., M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian analytic functions and determinantal point processes.* Vol. 51. Amer Mathematical Society, 2009.

[41] Hughes, C. "Mock-Gaussian Behavior" in *Recent Perspectives in Random Matrix Theory and Number Theory*, LMS Lecture Note Series 322, Edited by F. Mezzadri and N. C. Snaith, Cambridge University Press (2005) 337-355.

[42] Hughes, C., and Z. Rudnick. "Linear statistics for zeros of Riemann's zeta function." *CR Math. Acad. Sci. Paris* (2002): 667-670.

[43] Iwaniec, Henryk, and Emmanuel Kowalski. *Analytic number theory.* Vol. 53. American Mathematical Society, 2004.

[44] Kallenberg, Olav. *Foundations of modern probability.* Springer, 2002.

[45] Katz, Nicholas, and Peter Sarnak. "Zeroes of zeta functions and symmetry." *Bulletin of the American Mathematical Society* 36, no. 1 (1999): 1-26.

[46] Katz, N., and P. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy.* Vol. 45. Amer Mathematical Society, 1999.

[47] V. Kargin, "On fluctuations of Riemann's Zeta zeros", *http://arxiv.org/abs/1203.5309*, to appear *Probability Theory and Related Fields.*

[48] Katznelson, Y.*An introduction to harmonic analysis.* Cambridge University Press, 2004.

[49] Keating, J. P., and N. C. Snaith. "Random matrix theory and $\zeta(1/2+it)$." *Communications in Mathematical Physics* 214, no. 1 (2000): 57-89.

[50] Killip, R., and M. Stoiciu, "Eigenvalue statistics for CMV matrices: from Poisson to Clock via CβE." *Duke Math. J.* 146 (3) (2009), 361-399.

[51] von Koch, H. "Sur la distribution des nombres premiers." *Acta Mathematica* 24.1 (1901): 159-182.

[52] Laurincikas, A. *Limit theorems for the Riemann zeta-function,* Kluwer Academic, 1996.

[53] Montgomery, H. L. "The pair correlation of zeros of the zeta function." *Proc. Symp. Pure Math. Vol.* 24. (1973): 181-193.

[54] Montgomery, H. L., and K. Soundararajan. H. Montgomery and K. Soundararajan. "Beyond pair correlation", *Paul Erdős and his mathematics. I*, Bolyai Soc. Math. Stud. 11, (1999): 507-514.

[55] Montgomery, H. L., and Soundararajan, K. "Primes in short intervals." *Communications in mathematical physics* 252.1 (2004): 589-617.

[56] Montgomery, H. L., and R. C. Vaughan. *Multiplicative number theory I: Classical theory.* Vol. 97. Cambridge University Press, 2006.

[57] Odlyzko, A. M. "On the distribution of spacings between zeros of the zeta function." *Math. Comp* 48. 177 (1987): 273-308.

[58] Pérez-Marco, R. *Statistics on Riemann Zeros*, *arxiv preprint http://arxiv.org/abs/1112.0346*.

[59] Riemann, B. "Ueber die Anzahl der Primzahlen unter einer gegebenen Grosse." *Ges. Math. Werke und Wissenschaftlicher Nachlass* 2 (1859): 145-155.

[60] Rodgers, B. "Macroscopic Pair Correlation of the Riemann Zeroes for Smooth Test Functions", to appear in *Quart. J. of Math.*.

[61] Rodgers, B. "A Central Limit Theorem for the Zeroes of the Zeta Function." *arXiv preprint arXiv:1205.0303* (2012).

[62] Rubinstein, M. *Low-lying zeros of L-functions and random matrix theory*, Duke Math. J. 109 (2001) 147-181.

[63] Rudnick, Z., and P. Sarnak. "Zeros of principal L-functions and random matrix theory." *Duke Mathematical Journal* 81.2 (1996): 269-322.

[64] Selberg, A. "On the normal density of primes in small intervals, and the difference between consecutive primes." *Arch. Math. Naturvid.* 47 (1943): 87-105.

[65] Selberg, A. "On the remainder in the formula for $N(T)$, the number of zeroes of $\zeta(s)$ in the strip $0 < t < T$", *Avh. Norske Vid. Akad. Oslo.* I. (1944), no. 1, 27.

[66] Selberg, A. "Contributions to the theory of the Riemann zeta-function", *Arch. Mat. Naturvid.* 48 no. 5(1946): 89-155.

[67] Selberg, A. "An elementary proof of the prime-number theorem." *Annals of Math.* 50, no. 2 (1949): 305-313.

[68] Simon, B. *Orthogonal polynomials on the unit circle.* American Mathematical Soc., 2009.

[69] Snaith, N.C. "Riemann zeros and random matrix theory", *Milan J. Math.* 78 (2010), 135-152

[70] Soshnikov, A. "Determinantal random point fields." *Russian Mathematical Surveys* 55.5 (2007): 923.

[71] Soshnikov, A. "Gaussian Limit for Determinantal Random point Fields", *Ann. Probab.* 30 (2002) 171-187.

[72] Soshnikov, A. “The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities”. *Ann. Probab.* 28 (2000) 1353-1370.

[73] Soundararajan, K. ”The distribution of prime numbers.” *In Equidistribution in number theory, an Introduction*, pp. 59-83. Springer, 2007.

[74] Spohn, H. “Interacting Brownian particles: a study of Dyson’s model,” in *Hydrodynamic Behavior and Interacting Particle Systems,* Springer, IMA Volumes in Mathematics and its Applications, 9 (1987) 151-179.

[75] Stanley, R. P. *Enumerative Combinatorics, Vol. 2.* Cambridge Studies in Advanced Mathematics 62, 2001.

[76] Steuding, J. *Value-distribution of L-functions,* Vol. 1877. Springer, 2007.

[77] Tao, T. *Topics in random matrix theory,* Vol. 132. American Mathematical Soc., 2012.

[78] Tao, T., and V. Vu. “The Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices.” *Electronic Journal of Probability* 16 (2011): 2104-2121.

[79] Titchmarsh, E. C., and D.R. Heath-Brown. *The theory of the Riemann zeta-function.* Oxford University Press, 1986.

[80] Ward, D. R. “Some series involving Euler’s function”, *J. London Math. Soc.* 2 (1927) 210-214.

[81] Weil, A. “Sur les “formules explicites” de la theorie des nombres premiers”, *Comm. Sem. Math. Univ. Lund* (1952): 252-265.