n-LEVEL DENSITIES OF THE LOW-LYING ZEROS OF QUADRATIC DIRICHLET L-FUNCTIONS

by

JAKE LEVINSON

STEVEN J. MILLER, ADVISOR

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Mathematics

WILLIAMS COLLEGE
Williamstown, Massachusetts

May 20, 2011

ABSTRACT

The statistical distributions of zeros of L-functions can be used to study prime numbers, elliptic curves and even the ideal class groups of number fields. L-functions have been studied in connection with random matrix theory, which provides easier methods of computing these distributions. One statistic, the n-level density of low-lying zeros for a family of L-functions, measures the distribution of zeros near the central point $s = 1/2$. The Density Conjecture of Katz and Sarnak states that the n-level density for an L-function family depends on a classical compact group associated to the family. We extend previous work by Gao [4] on the n-level densities of quadratic Dirichlet L-functions. Our main result is to confirm up to $n = 6$ that, for test functions of suitable support, the density is as predicted by random matrix theory. We also consider a (conjectural) combinatorial identity for certain Fourier transforms of the test functions which, if true, would help in extending the result to all n.

ACKNOWLEDGEMENTS

First and foremost, I thank my wonderful advisor, Steven Miller, for his constant support and encouragement this past year, and for many late-night and early-morning email exchanges on a wide variety of subjects, including (but not limited to) the actual math research. His guidance made this work possible, and opened my eyes to the depth and beauty of analytic number theory. I am also very grateful to Mihai Stoiciu for being a dedicated second reader and for offering many helpful suggestions to improve both the thesis and thesis defense presentation.

I thank my friends, classmates and professors in the Williams math community for making this place so welcoming and dynamic, and Nick, Patricia and Rob for endless math conversations, because a steady influx of new ideas is the surest route towards mathematical progress.

Finally, I thank my family for their love and support; Aaron, Ellen and Casey for being there since the beginning; and Dave, MJ, and all the residents of Milham House '11, for an excellent year together.

Contents

1. INTRODUCTION

1.1. **The Theory of L-functions and Random Matrix Theory.** The distributions of zeros of L-functions $L(s, f)$ have implications throughout analytic number theory. The leading example is the Riemann zeta function $\zeta(s)$, defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{1.1}$$

when $\Re(s) > 1$, whose zeros give information about numerous problems, including the size of the error term to the approximation

$$\pi(x) := \#\{\text{primes } p \leq x\} \approx \frac{x}{\log x}. \tag{1.2}$$

The widely-studied Riemann Hypothesis (RH) states that all the nontrivial zeros of the zeta function lie on the line with $\Re(s) = \frac{1}{2}$, which is called the critical line; in this case, the error term is essentially $O(x^{1/2} \log x)$. Other examples of L-functions include the Dirichlet L-functions $L(s, \chi)$, the elliptic curve L-functions $L(s, E)$ and the most general class (conjecturally), the L-functions $L(s, f)$ attached to automorphic forms. In each case, the zeros contain important information about the object of study, and the Generalized Riemann Hypothesis (or GRH) states that the zeros of these L-functions again lie on the line with real part $\frac{1}{2}$.

An L-series attached to an object of study f is defined by a coefficient function $a_f : \mathbb{N} \to \mathbb{C}$ related to f. Usually $a_f(n)$ is multiplicative and does not grow too fast – say, $a_f(n) = O(n^c)$ for some $c \geq 0$. We study the series

$$\sum_{n=1}^{\infty} \frac{a_f(n)}{n^s}, \tag{1.3}$$

which converges absolutely when $\Re(s) > 1 + c$. The L-function $L(s, f)$ is the analytic continuation of the above series to an entire function on $\mathbb{C}$ (possibly meromorphic: $\zeta(s)$ has a simple pole at $s = 1$). According to the Langlands Program, all the L-functions arising in number theory can be written as products of L-functions $L(s, \pi)$ attached to automorphic representations π of GL_m over $\mathbb{Q}$, so we will use the notation $L(s, \pi)$ for a 'general' L-function, with coefficients $a_\pi(n)$. For more information on the Langlands Program, see [1].

By the multiplicativity of a_π it can be shown that the series has an Euler product (a product over the primes): if a_π is completely multiplicative, this is just

$$L(s,\pi) = \sum_{n=1}^{\infty} \frac{a_\pi(n)}{n^s} = \prod_{p \text{ prime}} \frac{a_\pi(p)}{1 - \frac{1}{p^s}}. \tag{1.4}$$

By taking the logarithmic derivative and using contour integration, we get an expression which connects the zeros and poles of $L(s,\pi)$ to a sum over the primes, giving us the 'explicit formula': for the Riemann zeta function, this leads to the approximation (1.2) above, whose error term comes from a sum over the zeros. Knowledge about the distributions of zeros is therefore useful for understanding and improving these formulae (among myriad other uses).

The zeta function satisfies the functional equation

$$\zeta(s) = 2^s \pi^{s-1} \sin(\pi s/2)\Gamma(1-s)\zeta(1-s), \tag{1.5}$$

which relates values in the region $\Re(s) > \frac{1}{2}$ to values in the region $\Re(s) < \frac{1}{2}$. This functional equation is an important tool for studying $\zeta(s)$: for example, it tells us that if there is a zero with real part $0 < \Re(s) < \frac{1}{2}$, then there is a corresponding zero at $1-s$ with real part $1 - \Re(s)$. The Riemann Hypothesis is the statement that the zeros of $\zeta(s)$ with real part $0 < \Re(s) < 1$ lie at the center of the symmetry, the line $\Re(s) = \frac{1}{2}$.

Other L-functions satisfy similar functional equations, and we always normalize if necessary, so that the axis of symmetry is again the line $\Re(s) = \frac{1}{2}$. To find the functional equation for a general L-function $L(s,\pi)$, we associate to L a product of Gamma factors $L_\infty(s,\pi)$ (see [20], 2.2 for details) and define the completed (or normalized) L-function

$$\Lambda(s,\pi) = L_\infty(s,\pi)L(s,\pi). \tag{1.6}$$

We call $L_\infty(s,\pi)$ the Gamma factor because it includes factors of the form $\Gamma(s+\mu_\pi(j))$, where the $\mu_\pi(j) \in \mathbb{C}$ are complex numbers which depend on π. The poles of the Gamma function will remove some of the zeros of $L(s,\pi)$ (for the Riemann zeta function: all the negative even integers). Those are called the trivial zeros, and the ones that remain in $\Lambda(s,\pi)$ are the nontrivial zeros. The completed L-function satisfies a functional equation

$$\Lambda(s,\pi) = \varepsilon(\pi)Q_\pi^{(1-s)/2}\Lambda(1-s,\tilde{\pi}), \tag{1.7}$$

where $Q_\pi \in \mathbb{R}$ is called the L-function's conductor, $\varepsilon(\pi) \in \mathbb{C}$ is the sign of the functional equation, $|\varepsilon(\pi)| = 1$, and $L(s,\tilde{\pi})$ is another L-function, called the contragredient of π. It is defined by

$$L(s,\tilde{\pi}) = \overline{L(\overline{s},\pi)}, \tag{1.8}$$

i.e. it has coefficients $a_{\tilde{\pi}}(n) = \overline{a_\pi(n)}$, and its completed form is $\Lambda(s,\tilde{\pi})$. We note that when the coefficients $a_\pi(n)$ are real, the contragredient is the same as the original L-function, so the functional equation allows us to study $L(s,\pi)$, particularly its zeros, in every region of the plane. For the family of L-functions under consideration in this thesis, this will be the case as the coefficients are all 1 or -1. The symmetry in the functional equation $s \leftrightarrow 1-s$ leads to the Generalized Riemann Hypothesis, which states that the nontrivial zeros of $L(s,\pi)$ have real part $\frac{1}{2}$.

Random matrix ensembles have been used to model the distributions of zeros of L-functions since the 1970s, when Montgomery [15] proved that the local pair correlation of zeros of the zeta function is the same as the eigenvalue pair correlation of the Gaussian Unitary Ensemble (GUE), under certain conditions (the Schwartz function $f \in \mathcal{S}(\mathbb{R})$ used to measure the correlation had to have Fourier transform $\hat{f}$ supported in $(-1,1)$). In that case, it was shown that

$$\lim_{N\to\infty} \sum_{1\le j\ne k\le N} f(\tilde{\theta}_j - \tilde{\theta}_k) = \int_{\mathbb{R}} f(x) C^{(2)}_{\mathrm{GUE}}(x)dx, \tag{1.9}$$

where $\tilde{\theta}_j$ ranges over the (normalized) nontrivial zeros of $\zeta(s)$, and $C^{(2)}_{\mathrm{GUE}}(x)$ is the eigenvalue pair correlation distribution of the GUE, given by

$$C^{(2)}_{\mathrm{GUE}}(x) = 1 - \left(\frac{\sin \pi x}{\pi x}\right)^2. \tag{1.10}$$

This work inspired the so-called GUE Conjecture: that all the statistics of the zeros of $\zeta(s)$ are the same as the statistics of eigenvalues of matrices in the GUE, and that this is true regardless of the support of the function f. In the 1980s, Odlyzko [16] performed extensive numerical computations on the local spacing statistics of zeros of the zeta function, providing additional evidence that the zeros of the zeta function have the same statistics as the GUE.

Further advances were made by Hejhal [5], who proved that (under similar conditions to Montgomery) the triple correlation for the zeros of $\zeta(s)$ is the same as the GUE, and Rudnick and Sarnak [20], who confirmed the GUE Conjecture for all the n-level correlations, $n \ge 2$, and for all L-functions attached to automorphic forms (again for suitably restricted test functions).

We study the local statistics of the zeros as follows: we assume GRH and write the nontrivial zeros of $L(s, f)$ as $\frac{1}{2} + i\gamma_L^{(j)}$, where $\gamma_L^{(j)} \in \mathbb{R}$ and

$$\cdots \leq \gamma_L^{(-2)} \leq \gamma_L^{(-1)} < 0 \leq \gamma_L^{(1)} \leq \gamma_L^{(2)} \leq \cdots .$$

It is known (see, for example, [8]) that the number $N_L(T)$ of zeros with $|\gamma_L^{(j)}| < T$ satisfies, as $T \to \infty$,

$$N_L(T) \sim \frac{M}{\pi} T \log T; \tag{1.11}$$

here M is a constant depending on L. In particular, since $N_L(T)$ increases faster than T, the spacings between zeros tend to 0 as $T \to \infty$. To study the local spacing distributions on the correct scale, we normalize (rescale) the zeros:

$$\tilde{\gamma}_L^{(j)} = \gamma_L^{(j)} \frac{M}{2\pi} \log |\gamma_L^{(j)}|. \tag{1.12}$$

We call the $\tilde{\gamma}$'s the normalized zeros, and the mean spacing between adjacent normalized zeros is 1. We can thus study the distribution of zeros on the local scale.

For the n-level correlations, we wish to study the density of 'clusters' of n zeros close to each other. We select Schwartz functions $f_1, \dots, f_{n-1}$, and define the n-level correlation using $f_1, \dots, f_{n-1}$ to be

$$C(L, n; f) = \lim_{N \to \infty} \frac{n!}{N} \sum_{\substack{1 \leq j_1, \dots, j_n \leq N \\ j_i \neq j_k}} f_1(\tilde{\gamma}_{j_2} - \tilde{\gamma}_{j_1}) \cdots f_{n-1}(\tilde{\gamma}_{j_n} - \tilde{\gamma}_{j_{n-1}}). \tag{1.13}$$

Since the f_i's decay rapidly, an n-tuple of zeros $\tilde{\gamma}_{j_1}, \dots, \tilde{\gamma}_{j_n}$ will only contribute to the sum if the successive differences are small. We can think of the f_i's as approximating a compact box $B \subset \mathbb{R}^{n-1}$; the above sum effectively counts clusters of n zeros whose successive differences are in the range indicated by B.

We can define the analogous statistic for correlations of eigenvalues of Hermitian matrices and obtain a density function $C_{\mathrm{GUE}}^{(n)}(x)$; the result of Rudnick and Sarnak in [20] is that, if $\hat{f}_1, \dots, \hat{f}_{n-1}$ are supported in $\sum_{i=1}^{n-1} |u_i| < 1$, then the n-level correlations are the same:

$$C(L, n; f) = \int_{\mathbb{R}^{n-1}} C_{\mathrm{GUE}}^{(n)}(x) f_1(x_1) \cdots f_{n-1}(x_{n-1}) dx_1 \cdots dx_{n-1}. \tag{1.14}$$

The n-level correlation describes the limiting behavior of zeros of the L-functions under study; it is, by definition, insensitive to the locations of any finite collection of zeros. In particular, it cannot describe the behavior of zeros near the central point. In order to understand these low-lying zeros, we use a different statistic, the n-level density.

Unlike for the correlations, where we have infinitely-many zeros to average over (in the limit), if we restrict our attention to low-lying zeros (say, on a compact set containing the point $s = \frac{1}{2}$), a single L-function has only finitely-many zeros there, not enough to average over. As Steve Miller might say, wherever the zeros are – that's where they are! Instead, we define the n-level density by averaging over an (infinite) family $\mathcal{F}$ of L-functions. We order the L-functions by conductor X, and study the statistic for finite slices of the family (parameterized by X). Then we take the limit as $X \to \infty$.

We define the n-level density as follows. Given a family $\mathcal{F}$ of L-functions, we write the normalized zeros as $\frac{1}{2} + i\tilde{\gamma}_L^{(j)}$ as above. Then we let $\mathcal{F}(X)$ denote the set of L-functions from $\mathcal{F}$ having conductor $d \in [X, 2X)$. Given a bounded, symmetric, rapidly decaying measurable function $f : \mathbb{R}^n \to \mathbb{C}$, we define

$$D^{(n)}(\mathcal{F}; f, X) = \frac{1}{|\mathcal{F}(X)|} \sum_{L \in \mathcal{F}(X)} \sum_{\substack{j_1, \ldots, j_n \\ j_i \neq \pm j_k}} f(\tilde{\gamma}_L^{(j_1)}, \ldots, \tilde{\gamma}_L^{(j_n)}). \tag{1.15}$$

In practice, we will choose even Schwartz functions $f_1, \ldots, f_n$ and let $f(x_1, \ldots, x_n) = \prod f_i(x_i)$.

Since f decays rapidly, only the low-lying zeros, those near the central point $s = \frac{1}{2}$, contribute to the sum in the limit. We wish to study the limit

$$D^{(n)}(\mathcal{F}; f) = \lim_{X \to \infty} D^{(n)}(\mathcal{F}; f, X). \tag{1.16}$$

We can study the analogous statistic for eigenvalues on the random matrix theory side. We first look at the n-level density of eigenvalues of random $N \times N$ matrices from a classical compact group $\mathcal{G}(N)$ (chosen with Haar measure), then take the limit as $N \to \infty$. All the classical compact groups consist of unitary matrices, so the eigenvalues can be written as $\lambda_j = e^{i\theta_j}$ for some $-\pi < \theta_j \leq \pi$; we study the spacings between the (normalized) eigenangles θ_j. Given a matrix $A \in \mathcal{G}$ with normalized eigenangles $\tilde{\theta}_1 \leq \cdots \leq \tilde{\theta}_N$, we define

$$W_{\mathcal{G}(N)}^{(n)}(A, f) = \sum_{\substack{1 \leq j_1, \ldots, j_n \leq N \\ j_i \neq j_k}} f(\tilde{\theta}_{j_1}, \ldots, \tilde{\theta}_{j_n}) \tag{1.17}$$

and study

$$\lim_{N \to \infty} \int_{\mathcal{G}(N)} W_{\mathcal{G}(N)}^{(n)}(A, f) dA. \tag{1.18}$$

Whereas the n-level correlations of all the classical compact groups have the same limit as $N \to \infty$ as the GUE, the n-level densities are distinct for each group, allowing us to distinguish between different types of symmetry. Katz and Sarnak [10, 11] showed that for each classical compact group, there exist measures $W_{\mathcal{G}}^{(n)}$ with values in $\mathbb{R}_{\geq 0}^n$ that depend on $\mathcal{G}$ and n, such that, if f has Fourier transform supported in $(-1, 1)$,

$$\lim_{N\to\infty} \int_{\mathcal{G}(N)} W_{\mathcal{G}(N)}^{(n)}(A, f)dA = \int_{\mathbb{R}_{\geq 0}^n} W_{\mathcal{G}}^{(n)}(x)f(x)dx. \tag{1.19}$$

In particular, we have the following densities:

Classical compact groups and their n-level densities.

$\mathcal{G}$	$W_{\mathcal{G}}^{(n)}$
U	$\det(K_0(x_i, x_j))_{1\leq i,j\leq n}$
USp	$\det(K_{-1}(x_i, x_j))_{1\leq i,j\leq n}$
SO(even)	$\det(K_1(x_i, x_j))_{1\leq i,j\leq n}$
SO(odd)	$\det(K_{-1}(x_i, x_j))_{1\leq i,j\leq n} + \sum_{\ell=1}^n \delta(x_\ell)\det(K_{-1}(x_i, x_j))_{i,j\neq\ell}$

where

$$K_\varepsilon(x, y) = \frac{\sin\pi(x-y)}{\pi(x-y)} + \varepsilon\frac{\sin\pi(x+y)}{\pi(x+y)}.$$

Since these densities are all different, the n-level density distinguishes the particular group symmetry. The Density Conjecture states that, for each family $\mathcal{F}$ of L-functions, there is an associated classical compact group $\mathcal{G}(\mathcal{F})$, and

$$D^{(n)}(\mathcal{F}; f) = \int_{\mathbb{R}_{\geq 0}^n} f(x)W_{\mathcal{G}(\mathcal{F})}^{(n)}(x)dx. \tag{1.20}$$

In this way we can articulate more precisely the symmetries in families of L-functions and their connections to random matrix theory. Determining the type of symmetry exhibited by a family of L-functions is especially helpful because the random matrix ensemble, once pinpointed, can be used thereafter to obtain predictions on other statistical questions about the family. The Density Conjecture has been verified for a number of different L-function families (up to small support of $\hat{f}$), including elliptic curves, weight k level N cuspidal newforms, certain families of $GL(n)$ L-functions, and Dirichlet and quadratic Dirichlet L-functions, with different classes of symmetry (unitary, orthogonal, and symplectic); see, for example, [2, 3, 6, 7, 9, 12, 14, 18, 21] and for quadratic Dirichlet L-functions, [4, 17, 19]. In this thesis we study quadratic Dirichlet L-functions, adding to the evidence from [4] that the symmetry type is symplectic.

1.2. **Quadratic Dirichlet L-functions.** We recall that a Dirichlet character is a group homomorphism $\chi : (\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}^*$, extended to the integers by setting

$$\chi(n) = \begin{cases} \chi(n \bmod m) & \text{if } (n,m) = 1, \\ 0 & \text{if } (n,m) \neq 1. \end{cases} \tag{1.21}$$

If $m \mid m'$, then a character χ on m gives rise to a character χ' on m' by setting $\chi'(n) = \chi(n)$ if $(n, m') = 1$ and $\chi'(n) = 0$ otherwise. In this case we say the character χ' is induced by χ; we say that a character χ is primitive if it is not induced by any other character. Finally, the conductor d of a character χ of modulus m is the smallest divisor d of m such that χ is induced by a character of modulus d. If χ is primitive, the conductor is the same as the modulus. It is a fact [8] that the conductor of a real primitive character is 1 or of the form $d, 4d$ or $8d$, for some $d \in \mathbb{N}$ odd and squarefree.

We say that χ_d is a quadratic character (of modulus d) if

$$\chi_d(n) = \left(\frac{d}{n}\right), \tag{1.22}$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. Note that χ_d takes values in $\{-1, 0, 1\}$.

The L-function family under consideration in this thesis is the collection of quadratic Dirichlet L-functions $L(s, \chi_d)$, given by

$$L(s, \chi_d) = \sum_{n=1}^{\infty} \frac{\chi_d(n)}{n^s}, \text{ if } \Re(s) > 1, \tag{1.23}$$

analytically continued to $\mathbb{C}$, and χ_d is a primitive quadratic character. Note that since the coefficients of the series are real, the functional equation only involves values of $L(s, \chi_d)$.

This family is conjectured to have symplectic symmetry, i.e. $\mathcal{G}(\mathcal{F}) = \mathrm{USp}$. We will extend results by Gao [4] showing that the n-level densities of zeros of quadratic Dirichlet L-functions are the same as the n-level densities of eigenvalues of unitary symplectic matrices. Following Gao, we restrict to the subfamily $\{L(s, \chi_{8d})\}$ containing only the L-functions with conductor $8d$, where $d \in \mathbb{N}$ is odd and square free. This simplifies the analysis by excluding χ_2, and facilitates applications of Poisson summation in the analysis in [4]. Note that χ_{8d} is a real primitive character with even sign (i.e. $\chi_{8d}(-1) = 1$.)

1.3. **The n-level density and thesis results.** The long-term goal of the work in this thesis (following Gao [4] and Rubinstein [19]) is to show, under the most general conditions possible, that the n-level density of eigenvalues of unitary symplectic matrices (with Haar measure) equals the n-level density of zeros of quadratic Dirichlet L-functions $L(s, \chi_{8d})$ of conductor $8d$.

Let $D(X)$ denote the set of all $d \in [X, 2X)$ of this form. We average over finite subsets of the family, then take the limit as $X \to \infty$. Note that

$$|D(X)| \sim \frac{4X}{\pi^2} \tag{1.24}$$

asymptotically [13], so we sum over the L-functions $L(s, \chi_{8d})$ with $d \in D(X)$ and multiply by $\frac{\pi^2}{4X}$. We wish to show the Density Conjecture:

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \sum_{\substack{j_1,\ldots,j_n \\ j_i \neq \pm j_k}} \prod_{i=1}^{n} f_i(L\gamma_{8d}^{(j_i)}) = \int_{\mathbb{R}^n} \prod_{i=1}^{n} f_i(x_i) W_{\mathrm{USp}}^{(n)}(x)dx, \tag{1.25}$$

where $\gamma_{8d}^{(j_1)}, \ldots, \gamma_{8d}^{(j_n)}$ ranges over the (imaginary parts of) n-tuples of zeros along the critical line with $j_i \neq \pm j_k$, $L = \frac{\log X}{\pi^2}$ is the normalizing factor, and $W_{USp}^{(n)}(x)$ is the n-level scaling density of the compact group USp of unitary symplectic matrices (see previous page).

The Density Conjecture for our L-function family is that equality holds in (1.25) for test functions f_i of arbitrary support. The more f_i is concentrated near 0, the larger the support of $\hat{f}_i$, so ideally we wish to show equality for $\hat{f}_i$'s having as large support as possible. (Indeed, if we could take f to be the Dirac delta $\delta(x)$, then the one-level density with f counts the rank of the central point zero. As useful as this might be, we would first have to prove equality in (1.25) for $\hat{f}(u) = \widehat{\delta(x)}(u) = 1(u)$, which has unbounded support!)

Let $f_1, \ldots, f_n$ be even Schwartz functions. Equality in (1.25) was first shown by Rubinstein [19], assuming that $\hat{f}_1, \ldots, \hat{f}_n$ are supported in $\sum_{i=1}^{n} |u_i| < 1$. Later, Gao [4] improved this result, computing the left-hand side of (1.25) explicitly for $\hat{f}_1, \ldots, \hat{f}_n$ supported in the larger region $\sum_{i=1}^{n} |u_i| < 2$. However, Gao could only show equality up to $n = 3$.

The main result of this thesis is showing that equality holds for $n \leq 6$ under the same conditions of support:

Theorem 1.1. *Assume GRH. Then equality holds in* (1.25) *if* $n \leq 6$, $f_1, \ldots, f_n$ *are even Schwartz functions and* $\hat{f}_1, \ldots, \hat{f}_n$ *are supported in* $\sum_{i=1}^{n} |u_i| < 2$.

The key difference between our approach and Gao's is as follows. Gao verified the cases $n = 1, 2, 3$ by using various Fourier Transform identities, and explicitly computing formulas for (sums of) integrals over certain regions in $\mathbb{R}^n$ ($n \leq 3$), such as (equation 5.11 from [4]) :

$$\int_{\substack{\mathbb{R}^3_{\geq 0} \\ u_1 > 1+u_2+u_3}} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i = \int_1^\infty \int_0^{u_1-1} \int_0^{u_1-u_2-1} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i,$$

$$\int_{\substack{\mathbb{R}^3_{\geq 0} \\ u_2 > 1+u_1+u_3}} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i = \int_0^\infty \int_{1+u_1}^{\infty} \int_0^{u_2-u_1-1} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i,$$

and showed that these sums yielded zero over various sub-regions of the support region $|u_1| + |u_2| + |u_3| < 2$. In contrast, we will write

$$\int_{\substack{\mathbb{R}^3_{\geq 0} \\ u_1 > 1+u_2+u_3}} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i = \int_{\mathbb{R}^3_{\geq 0}} \tilde{\chi}(u_1 - u_2 - u_3) \prod_{i=1}^{3} \hat{f}_i(u_i) du_i,$$

$$\int_{\substack{\mathbb{R}^3_{\geq 0} \\ u_1 > 1+u_2+u_3}} \prod_{i=1}^{3} \hat{f}_i(u_i) du_i = \int_{\mathbb{R}^3_{\geq 0}} \tilde{\chi}(-u_1 + u_2 - u_3) \prod_{i=1}^{3} \hat{f}_i(u_i) du_i,$$

where $\tilde{\chi}$ is the characteristic function of the interval $[1, \infty)$, and show equality by analyzing the combinatorics of various sums of products of characteristic functions. In Appendix A, we show how this approach simplifies the last step for the case $n = 3$.

1.4. **Outline of the Thesis.** The bulk of the thesis, Sections 3 through 5, is dedicated to the proof of the Density Conjecture for the 4-level density. The results for $n = 5, 6$, which use the same method of proof, are discussed afterward, in Section 6.1, and in Section 6.2 we discuss a possible approach towards proving the Density Conjecture for all other n.

Note: some of the expressions in the 4-level density are very long, so the symbolic manipulations for some steps of the proof were carried out in Mathematica. All the computations can be checked in the Mathematica notebooks, which are available at: http://www-personal.umich.edu/~jakelev/index.html.

1.4.1. *Outline of the Proof of Theorem 1.1,* $n = 4$. In the next section, we quote various expressions from Gao [4] for the n-level densities and terms thereof, to be used later on. Three important identities involve terms of the form

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d \in D(X)} \frac{1}{\log^n X} \prod_{i=1}^{n} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p} \right) \widehat{f}_i \left(\frac{\log p}{\log X} \right) \right),$$

for $n \leq 3$, which occur in the number theoretic density expression.

In Section 3, we reduce the (very long) 4-level density equation to a much shorter form by matching identical terms, discarding terms that are identically zero, and using the identities (2.14), (2.15), and (2.16) for all but one of the $\sum_p$ terms. Since the expressions are very long, we carry out this step using symbolic manipulation in Mathematica; we describe the code in Section 3.1.1. (The results can be verified in the Mathematica notebooks.) We reduce to an equation involving only 'new' 4-level terms, each of which is an integral with a distinct integrand or region of integration, so no further cancellation is immediately possible.

In Section 4, we find a 'canonical' form for these remaining unmatched terms. These terms involve integrals of Fourier transforms of products of the four test functions and the characteristic function χ, such as:

$$\int_{\mathbb{R}^2} \chi(u+v)\chi(u-v)\widehat{f_1 f_2}(u)\widehat{f_3 f_4}(v)dudv. \tag{1.26}$$

We recall here the definition of the convolution $\star$ of two functions,

$$\hat{f} \star \hat{g}(u) = \int_{\mathbb{R}} \hat{f}(u)\hat{g}(t-u)dt. \tag{1.27}$$

The Fourier transform converts products $\widehat{fg}$ to convolutions:

$$\widehat{fg}\ (u) = \hat{f} \star \hat{g}\ (u). \tag{1.28}$$

We apply this conversion, rewriting the Fourier transforms as convolutions, then as multiple integrals. This puts all of the integrals over $\mathbb{R}^4$. We then change variables, so that each integrand is a product of characteristic functions χ with the product $\prod_{i=1}^4 \hat{f}_i(u_i)du_i$.

Finally, since the test functions are even, we break each integral over $\mathbb{R}^4$ into a sum of integrals over $\mathbb{R}^4_{\geq 0}$. This form is easier to work with, since the remaining terms arising from the number theory side are also over $\mathbb{R}^4_{\geq 0}$. Moreover, certain simplifying arguments about the characteristic functions are only possible to make when all the u_i's are positive.

We finish the proof for $n = 4$ in Section 5. First, we consider the special case where each of the $\hat{f}_i$'s is supported in $(-\frac{1}{2}, \frac{1}{2})$. This case arises in practice: for instance, in computing the centered moments, we use the same test function n times. It is also often convenient to use the same test function when computing the n-level density numerically. In this case, the above steps are sufficient to show agreement between the random matrix theory and number theory. In Section 5.2, we treat the general case, where $\hat{f}_1(u_1), \ldots, \hat{f}_4(u_4)$ are supported in $\sum |u_i| < 2$. We rewrite the leftover terms to use the characteristic function of $[1, \infty)$ instead of $[-1, 1]$ (this simplifies several combinatorial arguments related to the support of the test functions), and show that the remaining terms cancel. We do not provide a line-by-line verification of this step, since the details are tedious to verify by hand; instead, we again carry out the computations in Mathematica. These steps can again be checked in the Mathematica notebooks.

2. Density Preliminaries

In this section, we present the general expressions for the n-level density of the eigenvalues of the Gaussian Symplectic Ensemble (equation (2.3)), the zeros of quadratic Dirichlet L-functions (equations (2.13) to (2.16)), and the explicit expressions representing the densities in the case $n = 4$ (equations (2.18) and (2.19)).

Note: Throughout this section, we use notation to describe sums and products indexed by various combinatorial objects. For clarity, examples of each notation are provided in Appendix B.

2.1. Random Matrix Theory Eigenvalue Densities.

Katz and Sarnak showed that the n-level density of eigenvalues for the group USp of unitary sympletic matrices is given by

$$W^{(n)}_{USp}(x)dx = \det\big(K_{-1}(x_i, x_j)\big)_{\substack{1\leq i\leq n,\\ 1\leq j\leq n}} \tag{2.1}$$

where

$$K_\varepsilon(x,y) = \frac{\sin\pi(x-y)}{\pi(x-y)} + \varepsilon\frac{\sin\pi(x+y)}{\pi(x+y)}. \tag{2.2}$$

To facilitate matching with the terms from the L-function densities, we instead use the following expression (developed by Gao) for the n-level eigenvalue density for USp (this is equation 4.12 in [4]):

$$\int_{\mathbb{R}^n}\prod_{i=1}^n f_i(x)W^{(n)}_{USp}(x)dx = \sum_{\underline{F}}(-2)^{n-\nu(\underline{F})}\prod_{l=1}^{\nu(\underline{F})}(P_l + Q_l + R_l + S_l), \tag{2.3}$$

where

$$\begin{aligned}
P_l &= (|F_l|-1)!(\frac{-1}{2})\int_{\mathbb{R}}\hat{F}_l(u)du,\\
Q_l &= (|F_l|-1)!\int_{\mathbb{R}}F_l(x)dx,\\
R_l &= -\sum_{[H,H^c]}(|H|-1)!(|H^c|-1)!\int_{\mathbb{R}}|u|\widehat{\prod_{i\in H}f_i}(u)\widehat{\prod_{i\in H^c}f_i}(u)du,\\
S_l &= \frac{1}{2}\int_{\mathbb{R}^{|F_l|}}\left((|F_l|-1)! - \chi^*_{F_l}(u_{i_1},\dots,u_{i_{|F_l|}})\right)\prod_{i\in F_l}\hat{f}_i(u_i)du_i,
\end{aligned}$$

where $\underline{F}$ ranges over the ways of partitioning $\{1,\dots,n\}$ into disjoint subsets $\{F_1,\dots,F_{\nu(\underline{F})}\}$, and we let

$$F_l(x)=\prod_{i\in F_l} f_i(x).$$

Also, the sum $\sum_{[H,H^c]}$ ranges over the ways of decomposing F_l into two proper disjoint subsets H and H^c (so $\varnothing\neq H,H^c\subsetneq F_l$; $H\cap H^c=\varnothing$; and $H\cup H^c=F_l$), and $\chi^*_{F_l}$ is defined as follows:

$$\begin{aligned}\chi(u)&=\begin{cases}0 & |u|>1\\ 1 & |u|\le 1\end{cases}\\ \chi^*_{F_l}(u_{i_1},\dots,u_{i_{|F_l|}})&=\sum_{\{i\,::\,F_l\}}\left(\prod_{k=1}^{|F_l|}\chi(u_{i_1}+\cdots+u_{i_k}-u_{i_{k+1}}-\cdots-u_{i_{|F_l|}})\right),\end{aligned}\tag{2.4}$$

where the notation $\sum_{\{i\,::\,F_l\}}$ means summation over all $(|F_l|-1)!$ cyclic permutations $(i_1=1,i_2\dots,i_{|F_l|})$ of the elements of F_l.

In particular, for $|F_l|=k\le 4$, the χ^*_k terms are as follows:

$$\chi^*_2(u_1,u_2)=\chi(u_1+u_2)\chi(u_1-u_2)=\begin{cases}0 & |u_1|+|u_2|>1\\ 1 & |u_1|+|u_2|\le 1,\end{cases}\tag{2.5}$$

$$\begin{aligned}\chi^*_3(u_1,u_2,u_3)&=\chi(u_1+u_2+u_3)\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3)\\ &\quad+\chi(u_1+u_2+u_3)\chi(u_1+u_2-u_3)\chi(u_1-u_2-u_3),\end{aligned}\tag{2.6}$$

$$\begin{aligned}&\chi^*_4(u_1,u_2,u_3,u_4)=\\ &\quad\chi(u_1+u_2+u_3+u_4)\chi(u_1-u_2+u_3+u_4)\chi(u_1-u_2-u_3+u_4)\chi(u_1-u_2-u_3-u_4)\\ &+\chi(u_1+u_2+u_3+u_4)\chi(u_1-u_2+u_3+u_4)\chi(u_1-u_2+u_3-u_4)\chi(u_1-u_2-u_3-u_4)\\ &+\chi(u_1+u_2+u_3+u_4)\chi(u_1+u_2-u_3+u_4)\chi(u_1+u_2-u_3-u_4)\chi(u_1-u_2-u_3-u_4)\\ &+\chi(u_1+u_2+u_3+u_4)\chi(u_1+u_2-u_3+u_4)\chi(u_1-u_2-u_3+u_4)\chi(u_1-u_2-u_3-u_4)\\ &+\chi(u_1+u_2+u_3+u_4)\chi(u_1+u_2+u_3-u_4)\chi(u_1-u_2+u_3-u_4)\chi(u_1-u_2-u_3-u_4)\\ &+\chi(u_1+u_2+u_3+u_4)\chi(u_1+u_2+u_3-u_4)\chi(u_1+u_2-u_3-u_4)\chi(u_1-u_2-u_3-u_4).\end{aligned}\tag{2.7}$$

2.2. **Number Theory densities.** Gao's expression for the n-level density of zeros of quadratic Dirichlet L-functions (equation (2.16) in [4]) is:

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \sum_{\substack{j_1,\ldots,j_n \\ j_i \neq \pm j_k}} \prod_{i=1}^{n} f_i(L\gamma_{8d}^{(j_i)}) = \tag{2.8}$$

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \sum_{\underline{F}} (-2)^{n-\nu(\underline{F})} \prod_{l=1}^{\nu(\underline{F})} (|F_l|-1)!\big(A_l + B_l\big),$$

where

$$A_l = \int_{\mathbb{R}} F_l(x)dx - \frac{1}{2}\int_{\mathbb{R}} \hat{F}_l(u)du, \tag{2.9}$$

$$B_l = -\frac{2}{\log X} \sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{F_l}\left(\frac{\log p}{\log X}\right). \tag{2.10}$$

Here $\underline{F}$ ranges over the ways of partitioning $\{1,\ldots,n\}$ into disjoint subsets, $\underline{F} = \{F_1,\ldots,F_{\nu(\underline{F})}\}$, and

$$F_l(x) = \prod_{i\in F_l} f_i(x),$$

$\sum_p$ is over the primes and $\left(\frac{8d}{p}\right)$ is the Legendre symbol.

Note that the A_l terms are independent of d and X. Hence, if we expand the products, the A_l terms can be pulled past the $\lim_{X\to\infty}\sum_{d\in D(X)}$ operator, making their contributions easy to compute:

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \left(\prod_l A_l\right) = \left(\prod_l A_l\right) \cdot \lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} 1 = \prod_l A_l, \tag{2.11}$$

since $|D(X)| \sim \frac{4X}{\pi^2}$ asymptotically. The main difficulty comes from the expressions of the form

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \prod_{l\in S} B_l,$$

where $S \subseteq \{1,\ldots,\nu(\underline{F})\}$.

For these, Gao computes the following:

Lemma 2.1. *Let $\hat{g}_i$, $i \in S$, be even test functions supported in $\sum_{i\in S}|u_i| < 2$. Then*

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{1}{\log^{|S|} X} \prod_{i=1}^{|S|} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{g_i}\left(\frac{\log p}{\log X}\right)\right) = \tag{2.12}$$

$$\left(\frac{1+(-1)^{|S|}}{2}\right) \sum_{(A;B)} \prod_{i=1}^{|S|/2} \int_0^\infty u_i \hat{g}_{a_i}(u_i)\hat{g}_{b_i}(u_i)du_i - \frac{1}{2} \sum_{\substack{S_2 \subsetneq S \\ |S_2| \text{ even}}} \left(\sum_{(C;D)} \prod_{i=1}^{|S_2|/2} \right.$$

$$\left. \int_0^\infty u_i \hat{g}_{c_i}(u_i)\hat{g}_{d_i}(u_i)du_i\right) \cdot \left(\sum_{I \subsetneq S_2^c} (-1)^{|I|} \int_{\mathbb{R}_{\geq 0}^{|S_2^c|}} \tilde{\chi}(\sum_{I^c} u_i - \sum_I u_i) \prod_{S_2^c} \hat{g}_i(u_i)du_i \right),$$

where S_2^c denotes the complement of S_2, the notations $\sum_{(A;B)}$ and $\sum_{(C;D)}$ indicate sums over the ways of pairing up the elements of S and S_2, respectively, and $\tilde{\chi}$ is the characteristic function

$$\tilde{\chi}(u) = \begin{cases} 0 & u \leq 1 \\ 1 & u > 1. \end{cases}$$

Empty products are 1.

By combining expressions (2.8) and (2.12), we can write down a general formula for the n-level density. Using

$$\prod_{i=1}^{\nu(\underline{F})} (A_l + B_l) = \sum_{S \subseteq \{1,\ldots,\nu(\underline{F})\}} \prod_S A_l \prod_{S^c} B_l,$$

we get the following expression (equation (2.3) in [4]):

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \sum_{\substack{j_1,\ldots,j_n \\ j_i \neq \pm j_k}} \prod_{i=1}^n f_i(L\gamma_{8d}^{(j_i)}) = \tag{2.13}$$

$$\sum_{\underline{F}} (-2)^{n-\nu(\underline{F})} \left(\prod_{l=1}^{\nu(\underline{F})} (|F_l|-1)! \right) \sum_S \left(\prod_{l\in S^c} \int_{\mathbb{R}} F_l(x)dx \right) \cdot \sum_{S_2 \subseteq S} \left(\frac{-1}{2}\right)^{|S_2^c|} \prod_{l \in S_2^c} \int_{\mathbb{R}} \hat{F}_l(u)du \cdot$$

$$\cdot \left(\frac{1+(-1)^{|S_2|}}{2} \cdot 2^{\frac{|S_2|}{2}} \sum_{(A;B)} \prod_{i=1}^{\frac{|S_2|}{2}} \int_{\mathbb{R}} |u_i| \hat{F}_{a_i}(u_i)\hat{F}_{b_i}(u_i)du_i - \frac{1}{2} \sum_{\substack{S_3 \subsetneq S_2 \\ |S_3| \text{ even}}} 2^{\frac{|S_3|}{2}} \right.$$

$$\left. \left(\sum_{(C;D)} \prod_{i=1}^{|S_3|/2} \int_{\mathbb{R}} |u_i| \hat{F}_{c_i}(u_i)\hat{F}_{d_i}(u_i)du_i \right) \cdot \sum_{I \subsetneq S_3^c} (-1)^{|I|} (-2)^{|S_3^c|} \int_{\mathbb{R}_{\geq 0}^{|S_3^c|}} \tilde{\chi}(\sum_{I^c} u_i - \sum_I u_i) \prod_{i=1}^{|S_3^c|} \hat{F}_i(u_i)du_i \right),$$

where S ranges over the subsets of $\{1,\ldots,\nu(\underline{F})\}$, and the rest of the notation is the same as in equations (2.8) and (2.12). Empty products are 1.

We note, however, that the integrals in (2.12) are desymmetrized into the region $\mathbb{R}^m_{\geq 0}$, whereas the terms of the random matrix theory density (2.3) are over all of $\mathbb{R}^m$. Where possible, we would like to avoid repeating the work involved in determining how the terms from (2.12) match with terms on the other side.

To this end, we note that the only partition $\underline{F}$ of $\{1,\ldots,n\}$ that gives rise to a sum with n factors,

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\prod^{n}B_l,$$

is the partition $\{\{1\},\ldots,\{n\}\}$. All other partitions yield $\prod B_l$ terms with at most $n-1$ factors. When we study the n-level density, we can use existing results on the lower densities to immediately substitute terms with $n-1$ or fewer B_l's with the appropriate terms from the random matrix theory side. We will only need to make explicit use of the formula (2.12) once, for this last term, and will be able to avoid using the longer expression (2.13) directly.

2.2.1. *The cases n=1,2,3.* We quote the identities obtained by Gao for the terms

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{1}{\log^n X}\prod_{i=1}^{n}\left(\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{g_i}\left(\frac{\log p}{\log X}\right)\right),$$

in the cases $n=1,2,3$.

Let χ and χ^* be defined as in equation (2.4). Let $\hat{g}_1,\hat{g}_2,\hat{g}_3$ be even test functions. If $\hat{g}_1$ is supported in $(-2,2)$, then

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{-2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{g_1}\left(\frac{\log p}{\log X}\right)=\frac{1}{2}\int_{\mathbb{R}}(1-\chi(u))\hat{g}_1(u)du. \tag{2.14}$$

If $\hat{g}_1,\hat{g}_2$ are supported in $|u_1|+|u_2|<2$, then

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{4}{\log^2 X}\prod_{i=1}^{2}\left(\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{g_i}\left(\frac{\log p}{\log X}\right)\right)= \tag{2.15}$$
$$\int_{\mathbb{R}}(1-\chi(u))\widehat{f_1f_2}(u)du-\int_{\mathbb{R}^2}(1-\chi(u_1+u_2)\chi(u_1-u_2))\hat{f}_1(u_1)du_i\hat{f}_2(u_2)du_2$$
$$+2\int_{\mathbb{R}}|u|\hat{f}_1(u)\hat{f}_2(u)du,$$

and if $\hat{g}_1, \hat{g}_2, \hat{g}_3$ are supported in $|u_1|+|u_2|+|u_3|<2$, then

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{-8}{\log^3 X}\prod_{i=1}^{3}\left(\sum_{p}\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{g_i}\left(\frac{\log p}{\log X}\right)\right)= \tag{2.16}$$
$$\sum_{a=1}^{3}\left(\int_{\mathbb{R}}(1-\chi(u))\hat{f}_a(u)du\cdot\int_{\mathbb{R}}|u|\hat{f}_b(u)\hat{f}_c(u)du\right)+2\int_{\mathbb{R}}(1-\chi(u))\widehat{f_1f_2f_3}(u)du$$
$$-2\sum_{a=1}^{3}\int_{\mathbb{R}^2}(1-\chi(u+v)\chi(u-v))\hat{f}_a(u)\widehat{f_bf_c}(v)dudv$$
$$+2\int_{\mathbb{R}^3}\big(2-\chi(u_1+u_2+u_3)\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3)$$
$$-\chi(u_1+u_2+u_3)\chi(u_1+u_2-u_3)\chi(u_1-u_2-u_3)\big)\prod_{i=1}^{3}\hat{f}_i(u_i)du_i.$$

Later, we use expressions (2.14), (2.15) and (2.16) to match up many of the terms in the 4-level density with corresponding terms from the random matrix theory expression.

2.2.2. *The case n=4.* In this section we present the random matrix theory and number theory expressions in the case $n=4$. We wish to show

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\sum_{\substack{j_1,\dots,j_4\\ j_i\neq\pm j_k}}\prod_{i=1}^{4}f_i(L\gamma_{8d}^{(j_i)})=\int_{\mathbb{R}^4}\prod_{i=1}^{4}f_i(x)W_{USp}^{(4)}(x)dx. \tag{2.17}$$

For $\{1,2,3,4\}$ there are 15 partitions, hence 15 summands on each side:

{{1},{2},{3},{4}}	{{1,4},{2},{3}}	{{1,2,3},{4}}
{{1,2},{3,4}}	{{1,3},{2},{4}}	{{1,3,4},{2}}
{{1,3},{2,4}}	{{1},{2,3},{4}}	{{1},{2,3,4}}
{{1,4},{2,3}}	{{1,2},{3},{4}}	{{1,2,4},{3}}
{{1},{2},{3,4}}	{{1},{2,4},{3}}	{{1,2,3,4}}.

Partitions of {1,2,3,4}.

Using expression (2.8), we can write the left-hand side of the 4-level density as

$$\begin{aligned}
&\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\Bigg(\prod_{i=1}^{4}\left(\int_{\mathbb{R}} f_i(x)dx-\frac{1}{2}\int_{\mathbb{R}}\hat{f}_i(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_i\left(\frac{\log p}{\log X}\right)\right)\\
&-2\sum_{\{\{a\},\{b\},\{c,d\}\}}^{6}\prod_{a,b}\left(\int_{\mathbb{R}} f_i(x)dx-\frac{1}{2}\int_{\mathbb{R}}\hat{f}_i(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_i\left(\frac{\log p}{\log X}\right)\right)\cdot\\
&\qquad\cdot\left(\int_{\mathbb{R}} f_c(x)f_d(x)dx-\frac{1}{2}\int_{\mathbb{R}}\widehat{f_cf_d}(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f_cf_d}\left(\frac{\log p}{\log X}\right)\right)\\
&+4\sum_{\{\{a,b\},\{c,d\}\}}^{3}\prod^{2}\left(\int_{\mathbb{R}} f_i(x)f_j(x)dx-\frac{1}{2}\int_{\mathbb{R}}\widehat{f_if_j}(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f_if_j}\left(\frac{\log p}{\log X}\right)\right)\\
&+8\sum_{\{\{a\},\{b,c,d\}\}}^{4}\left(\int_{\mathbb{R}} f_a(x)dx-\frac{1}{2}\int_{\mathbb{R}}\hat{f}_a(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_a\left(\frac{\log p}{\log X}\right)\right)\cdot\\
&\qquad\cdot\left(\int_{\mathbb{R}} f_b(x)f_c(x)f_d(x)dx-\frac{1}{2}\int_{\mathbb{R}}\widehat{f_bf_cf_d}(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f_bf_cf_d}\left(\frac{\log p}{\log X}\right)\right)\\
&-48\left(\int_{\mathbb{R}} f_1f_2f_3f_4(x)dx-\frac{1}{2}\int_{\mathbb{R}}\widehat{f_1f_2f_3f_4}(u)du-\frac{2}{\log X}\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f_1f_2f_3f_4}\left(\frac{\log p}{\log X}\right)\right)\Bigg),
\end{aligned}\tag{2.18}$$

where $\sum_{\{\{a\},\{b,c,d\}\}}^{4}$ is over the (four) distinct partitions of the form {{a},{b,c,d}}, and so on. (Note that we cannot use (2.12) for the $\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_i\left(\frac{\log p}{\log X}\right)$ terms until we expand out the products.)

On the other hand, using (2.3), we can write the right-hand side as

$$
\begin{aligned}
&\prod_{i=1}^{4}\left(\int_{\mathbb{R}} f_i(x)dx - \frac{1}{2}\int_{\mathbb{R}} \hat{f}_i(u)du + \frac{1}{2}\int_{\mathbb{R}}(1-\chi(u))\hat{f}_i(u)du\right) \qquad (2.19)\\
&-2\sum_{\{\{a\},\{b\},\{c,d\}\}}^{6}\prod_{a,b}\left(\int_{\mathbb{R}} f_i(x)dx - \frac{1}{2}\int_{\mathbb{R}} \hat{f}_i(u)du + \frac{1}{2}\int_{\mathbb{R}}(1-\chi(u))\hat{f}_i(u)du\right)\cdot\\
&\quad\cdot\left(\int_{\mathbb{R}} f_c(x)f_d(x)dx - \frac{1}{2}\int_{\mathbb{R}}\widehat{f_c f_d}(u)du - \int_{\mathbb{R}}|u|\widehat{f_c}(u)\widehat{f_d}(u)du\right.\\
&\quad\left.+\frac{1}{2}\int_{\mathbb{R}^2}(1-\chi(u_c+u_d)\chi(u_c-u_d))\hat{f}_c(u_c)\hat{f}_d(u_d)du_c du_d\right)\\
&+4\sum_{\{\{a,b\},\{c,d\}\}}^{3}\prod^{2}\left(\int_{\mathbb{R}} f_i(x)f_j(x)dx - \frac{1}{2}\int_{\mathbb{R}}\widehat{f_i f_j}(u)du - \int_{\mathbb{R}}|u|\hat{f}_i(u)\hat{f}_j(u)du\right.\\
&\quad\left.+\frac{1}{2}\int_{\mathbb{R}^2}(1-\chi(u_i+u_j)\chi(u_i-u_j))\hat{f}_i(u_i)\hat{f}_j(u_j)du_i du_j\right)\\
&+4\sum_{\{\{a\},\{b,c,d\}\}}^{4}\left(\int_{\mathbb{R}} f_a(x)dx - \frac{1}{2}\int_{\mathbb{R}}\hat{f}_a(u)du + \frac{1}{2}\int_{\mathbb{R}}(1-\chi(u))\hat{f}_a(u)du\right)\cdot\\
&\quad\cdot\left(2\int_{\mathbb{R}} f_b(x)f_c(x)f_d(x)dx - \int_{\mathbb{R}}\widehat{f_b f_c f_d}(u)du - \sum_{\{i,j\}\subset\{b,c,d\}}^{3}\int_{\mathbb{R}}|u|\widehat{f_i f_j}(u)\hat{f}_k(u)du\right.\\
&\quad+\frac{1}{2}\int_{\mathbb{R}^3}\big(2-\chi(u_b+u_c+u_d)\chi(u_b-u_c+u_d)\chi(u_b-u_c-u_d)\\
&\quad\left.-\chi(u_b+u_c+u_d)\chi(u_b+u_c-u_d))\chi(u_b-u_c-u_d)\big)\prod_{b,c,d}\hat{f}_i(u_i)du_i\right)\\
&-8\left(6\int_{\mathbb{R}} f_1f_2f_3f_4(x)dx - 3\int_{\mathbb{R}}\widehat{f_1f_2f_3}f_4(u)du - \sum_{\{a,b\},\{c,d\}}^{3}\int_{R}|u|\hat{f}_a(u)\hat{f}_b(u)du\right.\\
&\quad\left.-2\sum_{a=1}^{4}\int_{\mathbb{R}}|u|\hat{f}_a(u)\widehat{f_b f_c f_d}(u)du + \frac{1}{2}\int_{\mathbb{R}^4}(6-\chi_4^*(u_1,u_2,u_3,u_4))\prod_{i=1}^{4}\hat{f}_i(u_i)du_i\right),
\end{aligned}
$$

where the notations $\sum_{\{\{a,b\},\{c,d\}\}}^{3}$ are the same as for the left-hand side, and similarly for the other sums; and χ and χ_k^* are defined as in (2.4).

3. CANCELING LOWER TERMS

Writing out the expressions (2.18) and (2.19) for the 4-level densities in full detail (expanding the sums and products) takes several hundred lines. Many terms occur identically on both sides, and the first step of the verification is to find and remove these matching terms. Moreover, we can use Gao's previous results on expressions

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{1}{\log^k X}\prod_{i=1}^{k}\left(\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{F}_i\left(\frac{\log p}{\log X}\right)\right), (k\le 3),$$

to match these terms without having to use formula (2.12).

The main goal is to show the following:

Lemma 3.1. *Let $f_1,\dots,f_4$ be even test functions and $\hat{f}_1,\dots,\hat{f}_4$ be supported in $\sum_{i=1}^4|u_i|<2$. Then equality holds in equation* (1.25) *if and only if the following equation is true:*

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{16}{\log^4 X}\prod_{i=1}^{4}\left(\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_i\left(\frac{\log p}{\log X}\right)\right)=S_0+S_\star, \tag{3.1}$$

where

$$S_0=2\sum_{\{a,b\}}^{6}\int_{\mathbb{R}}|u|\hat{f}_a(u)\hat{f}_b(u)du\bigg(\int_{\mathbb{R}}(1-\chi(u))\widehat{f_cf_d}(u)du \tag{3.2}$$
$$-\int_{\mathbb{R}^2}(1-\chi_2^*(u_1,u_2))\hat{f}_c(u_1)\hat{f}_d(u_2)du_1du_2+\int_{\mathbb{R}}|u|\hat{f}_c(u)\hat{f}_d(u)du\bigg),$$

$$S_\star=\ -4\sum_{a=1}^{4}\left(\int_{\mathbb{R}^2}(1-\chi_2^*(u_1,u_2))\hat{f}_a(u_1)\widehat{f_bf_cf_d}(u_2)du_1du_2\right) \tag{3.3}$$
$$-\ 4\sum_{\{a,b\},\{c,d\}}^{3}\left(\int_{\mathbb{R}^2}(1-\chi_2^*(u_1,u_2))\widehat{f_af_b}(u_1)\widehat{f_cf_d}(u_2)du_1du_2\right)$$
$$+\ 4\sum_{\{a,b\}}^{6}\int_{\mathbb{R}^3}(2-\chi_3^*(u_1,u_2,u_3))\widehat{f_af_b}(u_1)\hat{f}_c(u_2)\hat{f}_d(u_3)du_1du_2du_3$$
$$-4\int_{\mathbb{R}^4}(6-\chi_4^*(u_1,u_2,u_3,u_4))\prod_{i=1}^{4}\hat{f}_i(u_i)du_i+4\int_{\mathbb{R}}(1-\chi(u))\widehat{f_1f_2f_3f_4}(u)du,$$

where the notations $\sum_{\{a,b\}}^6$ and $\sum_{\{a,b\},\{c,d\}}^3$ mean, respectively, summing over the (six) two-element subsets of $\{1,2,3,4\}$ and summing over the (three) ways of pairing up the elements of $\{1,2,3,4\}$, and each χ_i^* is a sum of products of characteristic functions, as in equation (2.4).

In particular, equation (3.1) is equivalent to Theorem 1.1. We describe the steps used to reduce equation (1.25) to equation (3.1) in more detail below. In the end, we are left with the above expressions, and we note that each of the integrals in (3.1) has a distinct integrand or region of integration. Since there are no matching counterparts, no further cancellation is immediately possible, and we complete the analysis in Section 4.

3.1. **Summary of Steps.** First, we expand out the expressions (2.18) and (2.19). Immediately, we can cancel terms such as

$$\prod_{i=1}^{\nu(\underline{F})} \int_{\mathbb{R}} \hat{F}_i(u)du, \tag{3.4}$$

but integrals involving $|u|$ and χ from the random matrix theory side remain, as do all the $\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{F}\left(\frac{\log p}{\log X}\right)$ terms from the number theory side. The resulting expression is roughly 800 lines long and contains around 550 terms.

We next use the substitutions (2.14), (2.15) and (2.16) for terms of the form

$$\prod_{i=1}^{k} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{F_i}\left(\frac{\log p}{\log X}\right)\right),$$

$k = 1, 2, 3$. This allows many of the terms from the random matrix theory side to cancel, but the result is still very long (200 lines of output involving 119 terms).

The last step is to use the support assumption,

$$\sum_{i=1}^{4} \operatorname{supp}(\hat{f}_i) < 2,$$

to discard integrals located outside this region. In particular, we have the following lemma:

Lemma 3.2. *Let F_l and F_k be disjoint subsets of $\{1, \ldots, n\}$. Then*

$$\int_{\mathbb{R}^{|F_l|}} \left((|F_l|-1)! - \chi^*_{F_l}(u)\right) \prod_{i\in F_l} \hat{f}_i(u_i)du_i \cdot \int_{\mathbb{R}^{|F_k|}} \left((|F_k|-1)! - \chi^*_{F_k}(u)\right) \prod_{i\in F_k} \hat{f}_i(u_i)du_i = 0, \tag{3.5}$$

where $\chi^*_{F_l}(u)$ is shorthand for $\chi^*_{F_l}(u_{i_1}, \ldots, u_{i_{|F_l|}})$, as defined in equation (2.4).

Proof. Since F_l and F_k are disjoint, we must have either

$$\sum_{i\in F_l} \operatorname{supp}(\hat{f}_i) < 1 \text{ or } \sum_{j\in F_k} \operatorname{supp}(\hat{f}_j) < 1,$$

since the total support is less than 2. Without loss of generality, assume F_l's total support is less than 1. Then

$$|\underbrace{\varepsilon_{i_1}u_{i_1} + \cdots + \varepsilon_{i_k}u_{i_k}}_{i_j \in F_l}| \leq \sum_{F_l} |u_i| < 1$$

in the region of support, so $\chi(\sum_{F_l} \varepsilon_i u_i) = 1$ for any $\varepsilon_i = \pm 1$. So, since $\chi^*_{F_l}$ is a sum of $(|F_l| - 1)!$ products of χ's, the F_l integrand is identically 0. □

Note that, using the notation from the expression for the random matrix density (2.3), Lemma 3.2 says that $S_l \cdot S_k = 0$, where S_l and S_k are the terms from the product

$$\prod_{i=1}^{\nu(\underline{F})} (P_l + Q_l + R_l + S_l)$$

corresponding to a partition $\underline{F} = \{F_1, \ldots, F_{\nu(\underline{F})}\}$ of $\{1, \ldots, n\}$.

For $n = 4$, Lemma 3.2 reduces to the cases

$$\int_{\mathbb{R}} (1 - \chi(u))\hat{f}_a(u)du \cdot \int_{\mathbb{R}^{|F_l|}} ((|F_l| - 1)! - \chi^*_{F_l}(u)) \prod_{i \in F_l} \hat{f}_i(u_i)du_i, \text{ where } 1 \leq |F_l| \leq 3, \tag{3.6}$$

$$\int_{\mathbb{R}^2} (1 - \chi(u+v)\chi(u-v)) \prod_{a,b} \hat{f}_i(u_i)du_i \cdot \int_{\mathbb{R}^2} (1 - \chi(u+v)\chi(u-v)) \prod_{c,d} \hat{f}_i(u_i)du_i. \tag{3.7}$$

These steps reduce the expression to the form in Lemma 3.1. No further cancellation is possible without finding a way of rewriting the integrals to resemble one another. In Section 4, we show how we can convert each term into a 'canonical' form, leading to the desired cancellation.

3.1.1. *A technical note: representing the expressions in Mathematica.* Given the length and complexity of the expressions (2.18) and (2.19), these first steps were carried out symbolically in Mathematica. We subtract the left-hand side from the right-hand side and use symbolic manipulation to simplify.

Rather than reproducing the expressions line-by-line (which would not be enlightening), we discuss the code used to reduce the expressions to a more manageable form to be verified by hand.

Mathematica attempts to evaluate the $\int$ symbol even if the integrand contains undefined functions, and is occasionally successful:

$$\text{In}[1] := \int_{-\infty}^{\infty} \text{DiracDelta}[x] f[x] dx$$
$$\text{Out}[1] := f[0]$$

However, our expressions are long enough that using the $\int$ symbol makes the computations very slow, as Mathematica attempts to evaluate each term before displaying it symbolically. Moreover, Mathematica does not simplify integrals with different formal variables:

$$\text{In}[3] := \int_{-\infty}^{\infty} f[u_1] du_1 - \int_{-\infty}^{\infty} f[u_1] du_1$$
$$\text{Out}[3] := 0$$

but

$$\text{In}[2] := \int_{-\infty}^{\infty} f[u_1] du_1 - \int_{-\infty}^{\infty} f[u_2] du_2$$
$$\text{Out}[2] := \int_{-\infty}^{\infty} f[u_1] du_1 - \int_{-\infty}^{\infty} f[u_2] du_2.$$

The choice of variable is not always 'natural': for example, although it might seem reasonable to always assign u_1 to $\hat{f}_1$, it is not necessarily clear what variable should go with $\widehat{f_1 f_3}$. Consequently the expressions may include terms that are equal but might not be recognized as such.

To avoid the first problem, we use undefined symbols to represent integrals, Fourier transforms and the characteristic functions χ. We have terms that look like

$$\text{IntR}[f_3[x], x]\text{IntR}[(1 - \chi[u])\widehat{f_1 f_2}[u], u],$$

where the IntR and χ functions, and the $\widehat{\cdot}$ notation, are not given any definition as Mathematica symbols, but stand for $\int_{\mathbb{R}}$, the characteristic function χ, and the Fourier transform, respectively. This approach makes the manipulations fast and provides precise control over pattern matching and replacements. (Alternately, we could have used the Hold[ ] or Unevaluated[ ] operators and kept the $\int$ symbol.)

The full list of symbols used is:

Formal Symbol	Meaning
IntR[f[x],x]	$\int_{\mathbb{R}} f(x)dx$
IntR0[f[x],x]	$\int_{\mathbb{R}_{\geq 0}} f(x)dx$
IntRN[$\{i_1, \ldots, i_n\}$, f]	$\int_{\mathbb{R}^n} f dx_{i_1} \cdots dx_{i_n}$
IntR0N[$\{i_1, \ldots, i_n\}$, f]	$\int_{\mathbb{R}^n_{\geq 0}} f dx_{i_1} \cdots dx_{i_n}$
$\widehat{f_1 \cdots f_n}$	Fourier transform
$\chi[x], \tilde{\chi}[x]$	The characteristic functions

To deal with the problem of recognizing integrals with different formal variables, we use replacement rules to standardize subscripts and ambiguous notation as necessary. We also supply replacement rules to effectively make the IntR and other such functions linear, so that terms like

$$\text{IntR}[f[x], x] + \text{IntR}[-f[x], x]$$

are eventually simplified and canceled.

Finally, we include patterns to match the $\sum_p$ and χ products in equations (3.4) and (3.6), and replace them with the appropriate expressions (or 0).

4. THE UNMATCHED TERMS

After cancelling as many terms as possible, we are left with the unmatched terms shown in Lemma 3.1. We wish to show that equation (3.1) holds:

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{16}{\log^4 X} \prod_{i=1}^{4} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{f}_i \left(\frac{\log p}{\log X}\right)\right) = S_0 + S_\star,$$

where S_0 and $S_\star$ are as defined in equation (3.1). As noted earlier, each of the remaining terms has a distinct integrand or region of integration, so no further cancellation is possible without finding another way to rewrite the terms. The goal of this section is to put the unmatched terms into a canonical form, allowing them to be added together and simplified.

First, in Section 4.1, we make a slight modification to Gao's formula for the left-hand side of equation (3.1), to make it easier to match with S_0 and $S_\star$. Then, in Section 4.2, we consider the terms $S_\star$ and S_0. We prove the following lemma:

Lemma 4.1. *Each of the terms of $S_\star$ and S_0 from equation* (3.1) *can be written as sums of integrals of the following form:*

$$\int_{\mathbb{R}^k_{\geq 0}} \chi^{**}(u_1, \ldots, u_k) \prod_{i=1}^{k} \hat{f}_i(u_i) du_i, \tag{4.1}$$

where $\chi^{**}(u_1, \ldots, u_k)$ is either 1 or a product of terms of the form

$$\chi(\varepsilon_1 u_1 + \cdots + \varepsilon_k u_k),$$

with $\varepsilon_i = \pm 1$ varying from term to term. For $S_\star$, $k = 4$, and for S_0, $k = 2$.

The identity for S_0 is

$$\begin{aligned} S_0 = 4 &\sum_{\{a,b\},\{c,d\}}^{3} \int_{\mathbb{R}} |u| \hat{f}_a(u) \hat{f}_b(u) du \int_{\mathbb{R}} |u| \hat{f}_c(u) \hat{f}_d(u) du \\ &+ 4 \sum_{\{a,b\}}^{6} \int_{\mathbb{R}} |u| \hat{f}_a(u) \hat{f}_b(u) du \int_{\mathbb{R}^2_{\geq 0}} \big(\chi(u_c + u_d) - \chi(u_c - u_d)\big) \prod_{c,d} \hat{f}_i(u_i) du_i, \end{aligned} \tag{4.2}$$

and in the case where each $\hat{f}_i$ is supported in $(-\frac{1}{2}, \frac{1}{2})$, the identity for $S_\star$ is

$$\begin{aligned} S_\star = 24 &\int_{\mathbb{R}^4_{\geq 0}} \prod_{i=1}^{4} \hat{f}_i(u_i) du_i - 8 \sum_{\ell=1}^{4} \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_\ell + \sum_{j\neq \ell} u_j) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i \\ &+ 8 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i. \end{aligned} \tag{4.3}$$

Since the Fourier transforms convert products of the test functions into convolutions, we unfold each of the convolutions to get integrals over $\mathbb{R}^4$. We then change variables to put all the integrands in the form used by Lemma 4.1, and desymmetrize into the region $\mathbb{R}^4_{\geq 0}$.

4.1. **The Left-Hand Side of** (3.1)**.** We first examine the left-hand side of (3.1). All the terms that occur there are desymmetrized into the region $\mathbb{R}^n_{\geq 0}$, and use the characteristic function $\tilde{\chi}$ instead of χ. We make a slight change to Gao's formula, to convert these terms into a form closer to that of the terms on the right-hand side of (3.1).

We first recall Gao's expression to be used for the LHS of equation (3.1) (note that this is equation (3.13) from [4]) :

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{1}{\log^n X} \prod_{i=1}^{n} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{g_i}\left(\frac{\log p}{\log X}\right)\right) = \tag{4.4}$$

$$\left(\frac{1+(-1)^n}{2}\right) \sum_{(A;B)} \prod_{i=1}^{n/2} \int_0^\infty u_i \hat{g}_{a_i}(u_i)\hat{g}_{b_i}(u_i)du_i - \frac{1}{2} \sum_{\substack{S\subsetneq\{1,\ldots,n\}\\ |S| \text{ even}}} \left(\sum_{(A_S;B_S)} \prod_{i=1}^{|S|/2} \right.$$

$$\left. \int_0^\infty u_i \hat{g}_{a_i}(u_i)\hat{g}_{b_i}(u_i)du_i \right) \cdot \left(\sum_{I\subsetneq S^c} (-1)^{|I|} \int_{\mathbb{R}^{|S^c|}_{\geq 0}} \tilde{\chi}(\sum_{I^c} u_i - \sum_I u_i) \prod_{S^c} \hat{g}_i(u_i)du_i \right),$$

where $\sum_{(A;B)}$ and $\sum_{(A_S;B_S)}$ mean summing over the distinct ways of pairing up the elements of $\{1,\ldots,n\}$ and S, respectively, and the function $\tilde{\chi}$ is given by

$$\tilde{\chi}(u) = \begin{cases} 0 & u \leq 1 \\ 1 & u > 1. \end{cases}$$

Note that

$$\tilde{\chi}(u) + \tilde{\chi}(-u) = 1 - \chi(u).$$

We rewrite this formula in a way that makes it easier to match with the right-hand side. First, since all test functions are even, we have

$$\int_0^\infty u\hat{f}_i(u)\hat{f}_j(u)du = \frac{1}{2}\int_{\mathbb{R}} |u|\hat{f}_i(u)\hat{f}_j(u)du. \tag{4.5}$$

Assuming n is even, we can also rewrite the last part of the expression. Consider $S \subsetneq \{1,\ldots,n\}$ with $|S|$ even (so $|S^c|$ is also even). Let $I \subsetneq S^c$. If I is nonempty, then both I and $S^c \setminus I = I^c$ occur in the sum over all $I \subsetneq S^c$. Moreover, $|I|$ and $|I^c|$ have the same parity, so

$$(-1)^{|I|} = (-1)^{|I^c|}.$$

Thus the $\tilde{\chi}$ terms for I and I^c will reinforce each other, so we can replace them with χ terms. In particular, if we let $\sum_{[I,I^c]:S^c}$ denote a sum over the distinct ways of decomposing S^c into two disjoint proper subsets, then

$$\begin{aligned}
&\sum_{I\subsetneq S^c}(-1)^{|I|}\cdot\tilde{\chi}(\sum_{I^c}u_i-\sum_{I}u_i)\\
&\quad=\tilde{\chi}(u_{s_1}+\cdots+u_{s_k})+\sum_{[I,I^c]:S^c}(-1)^{|I|}\bigg(\tilde{\chi}(\sum_{I^c}u_i-\sum_{I}u_i)+\tilde{\chi}(\sum_{I}u_i-\sum_{I^c}u_i)\bigg)\\
&\quad=(1-\chi(u_{s_1}+\cdots+u_{s_k}))+\sum_{[I,I^c]:S^c}(-1)^{|I|}(1-\chi(\sum_{I}u_i-\sum_{I^c}u_i))\\
&\quad=\,-\chi(u_{s_1}+\cdots+u_{s_k})-\sum_{[I,I^c]:S^c}(-1)^{|I|}\chi(\sum_{I}u_i-\sum_{I^c}u_i).
\end{aligned}$$

Note that the first term is the summand corresponding to $I=\varnothing, I^c=S^c$, (and $\tilde{\chi}(-u_1-\cdots-u_n)=0$ since all u_i are positive), and that the 1's cancel because

$$1+\sum_{[I,I^c]:S^c}(-1)^{|I|}=\frac{1}{2}\bigg(\#\{\text{even subsets of } S^c\}-\#\{\text{odd subsets of } S^c\}\bigg)=0$$

is a well-known combinatorial identity.

These substitutions yield, for n even, the following expression:

$$\begin{aligned}
&\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{1}{\log^n X}\prod_{i=1}^{n}\left(\sum_p\frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{g}_i\left(\frac{\log p}{\log X}\right)\right)= \qquad (4.6)\\
&\frac{1}{2^{n/2}}\sum_{(A;B)}\prod_{i=1}^{n/2}\int_{\mathbb{R}}|u_i|\hat{g}_{a_i}(u_i)\hat{g}_{b_i}(u_i)du_i+\frac{1}{2}\sum_{\substack{S\subsetneq\{1,\ldots,n\}\\|S|\text{ even}}}\frac{1}{2^{|S|/2}}\cdot\\
&\bigg(\sum_{(A_S;B_S)}\prod_{i=1}^{|S|/2}\int_{\mathbb{R}}|u_i|\hat{g}_{a_i}(u_i)\hat{g}_{b_i}(u_i)du_i\bigg)\cdot\bigg(\int_{\mathbb{R}_{\geq 0}^{|S^c|}}\chi(u_{s_1}+\cdots+u_{s_k})\prod_{S^c}\hat{g}_i(u_i)du_i\\
&+\sum_{[I,I^c]:S^c}(-1)^{|I|}\int_{\mathbb{R}_{\geq 0}^{|S^c|}}\chi(\sum_{I}u_i-\sum_{I^c}u_i)\prod_{S^c}\hat{g}_i(u_i)du_i\bigg).
\end{aligned}$$

The advantage of this expression is that Gao's expressions for the n-level density on the random matrix theory side already involve sums of the form $\sum_{[I,I^c]}$ over the ways of decomposing a set into two disjoint proper subsets, and all the integrals involve χ, not $\tilde{\chi}$.

Applying equation (4.6) gives the following identity:

$$\lim_{X\to\infty}\frac{\pi^2}{4X}\sum_{d\in D(X)}\frac{16}{\log^4 X}\prod_{i=1}^{4}\left(\sum_p \frac{\log p}{\sqrt{p}}\left(\frac{8d}{p}\right)\widehat{f}_i\left(\frac{\log p}{\log X}\right)\right)=\tilde{S}_0+\tilde{S}_\star, \tag{4.7}$$

where

$$\tilde{S}_0 = 4\sum_{\{a,b\},\{c,d\}}^{3}\left(\int_{\mathbb{R}}|u|\hat{f}_a(u)\hat{f}_b(u)du\int_{\mathbb{R}}u\hat{f}_c(u)\hat{f}_d(u)du\right) \tag{4.8}$$

$$+4\sum_{\{a,b\}}^{6}\int_{\mathbb{R}}|u|\hat{f}_a(u)\hat{f}_b(u)du\cdot\int_{\mathbb{R}^2_{\geq 0}}(\chi(u_c+u_d)-\chi(u_c-u_d))\prod_{c,d}\hat{f}_i(u_i)du_i,$$

$$\tilde{S}_\star = 8\sum_{\{a,b\},\{c,d\}}^{3}\int_{\mathbb{R}^4_{\geq 0}}\chi(u_a+u_b-u_c-u_d)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i \tag{4.9}$$

$$+8\int_{\mathbb{R}^4_{\geq 0}}\chi(\sum_{i=1}^{4}u_i)\prod_{i=1}^{4}\hat{f}_i(u_i)du_i-8\sum_{\ell=1}^{4}\int_{\mathbb{R}^4_{\geq 0}}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\prod_{i=1}^{4}\hat{f}_i(u_i)du_i.$$

We will see that this expression allows us to match the remaining terms of the 4-level density. In particular, we will show that $S_0=\tilde{S}_0$ and $S_\star=\tilde{S}_\star$.

4.2. **The Right-Hand Side of** (3.1)**.** We now prove Lemma 4.1 by analyzing the terms from the right-hand side of (3.1).

Proof of Lemma 4.1. We first consider the term $S_\star$.

Each term from $S_\star$ is an integral containing all four test functions, multiplied together and Fourier-transformed according to some partition $\{F_1,\dots,F_k\}$ of $\{1,2,3,4\}$. In particular, each has the form

$$\int_{\mathbb{R}^k}\left((k-1)!-\chi_k^*(u_1,\dots,u_k)\right)\big(\widehat{\prod_{j\in F_1}f_j}\big)(u_1)du_1\cdots\big(\widehat{\prod_{j\in F_k}f_j}\big)(u_k)du_k, \tag{4.10}$$

There are two difficulties in simplifying these terms and matching them with terms from the left-hand-side of (3.1). First of all, each term of $S_\star$ arises from a unique partition of $\{1,2,3,4\}$, hence involves a unique combination of Fourier transforms of the test functions. Second, Gao's formula for the left-hand side involves integrals in the region $\mathbb{R}^n_{\geq 0}$, not all of $\mathbb{R}^n$. Hence, in order to show equality, we convert each term into a sum of 4-dimensional integrals of $\prod_{i=1}^4\hat{f}_i(u_i)du_i$ (with characteristic functions) in the region $\mathbb{R}^4_{\geq 0}$.

Since the Fourier transform converts multiplication into convolution $\star$, the term (4.10) is the same as

$$\int_{\mathbb{R}^k} \big((k-1)!-\chi_k^*(u_1,\dots,u_k)\big)\big(\underbrace{\hat{f}_{j_{1,1}}\star\cdots\star\hat{f}_{j_{1,|F_1|}}}_{j\in F_1}\big)(u_1)du_1\cdots\big(\underbrace{\hat{f}_{j_{k,1}}\star\cdots\star\hat{f}_{j_{k,|F_k|}}}_{j\in F_k}\big)(u_k)du_k. \tag{4.11}$$

Since an n-fold convolution is just an $(n-1)$-dimensional integral, and since each test function occurs once, if we 'unfold' each convolution, we always get an integral over $\mathbb{R}^4$. Then we desymmetrize into the region $\mathbb{R}^4_{\geq 0}$. This will allow us first to simplify many of the products of characteristic functions, and, second, to match all of the remaining terms.

4.2.1. *Reaching a canonical form.* We rewrite the convolutions in two steps. First, we repeatedly use the definition $(f\star g)(u)=\int_{\mathbb{R}} f(t-u)g(t)dt$ to obtain an integral over the region $\mathbb{R}^4$. Since all the test functions are even, for notational simplicity we use

$$(f\star g)(u)=\int_{\mathbb{R}} f(u+t)g(t)dt \tag{4.12}$$

instead (we don't have to keep track of signs). Applying (4.12) repeatedly gives the identity (for even functions)

$$\widehat{f_1\cdots f_k}(u_1)=\int_{\mathbb{R}^{k-1}}\hat{f}_1(u_1+\cdots+u_k)\prod_{i=2}^k\hat{f}_i(u_i)du_i. \tag{4.13}$$

We apply this identity to each of the convolutions $\widehat{\prod_{j\in F_\ell} f_j}$ in (4.10); for clarity, we renumber the variables so that $u_{\ell,j}$ is the j-th variable associated to the ℓ-th convolution (after expanding into an integral over $\mathbb{R}^{|F_\ell|}$). The result is an integrand of the following form:

$$\big((k-1)!-\chi_k^*(u_{1,1},\dots,u_{k,1})\big)\prod_{i=1}^k\Bigg(\underbrace{\hat{f}_{j_1}(u_{i,1}+\cdots+u_{i,|F_i|})\hat{f}_{j_2}(u_{i,2})\cdots\hat{f}_{j_{|F_i|}}(u_{i,|F_i|})}_{j\in F_i}\Bigg). \tag{4.14}$$

We then make the change of variables

$$
\begin{aligned}
u_{1,1}+u_{1,2}+\cdots+u_{1,|F_1|} &\mapsto u_1 \\
u_{1,j} &\mapsto u_j, j=2,\ldots,|F_1| \\
u_{2,1}+u_{2,2}+\cdots+u_{2,|F_2|} &\mapsto u_{|F_1|+1} \\
u_{2,j} &\mapsto u_{|F_1|+j}, j=2,\ldots,|F_2| \\
&\vdots \\
u_{k,1}+\cdots u_{k,|F_k|} &\mapsto u_{|F_1|+\cdots+|F_{k-1}|+1} \\
u_{k,j} &\mapsto u_{|F_1|+\cdots+|F_{k-1}|+j}, j=2,\ldots,|F_k|.
\end{aligned} \tag{4.15}
$$

Note that the matrix A sending $(u_{1,1},\ldots,u_{k,|F_k|})$ to $(u_1,\ldots,u_n)$ is a block diagonal matrix of the form $A=A_1\oplus\cdots\oplus A_k$, where each block corresponds to one of the F_i's, and is an $|F_i|\times|F_i|$ matrix of the form

$$
A_i=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, \qquad A_i^{-1}=\begin{pmatrix} 1 & -1 & \cdots & -1 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},
$$

with 1's on the first row and main diagonal, and zeros elsewhere. Hence the Jacobian determinant for this change of coordinates is always 1, so the value of the integral is unchanged.

After renaming the variables if necessary, the integrand is now in the form

$$
\big((k-1)!-\chi_k^*(*,\ldots,*)\big)\prod_{i=1}^{k}\hat{f}_i(u_i)du_i,
$$

where the χ_k^* term is now being evaluated at sums of the u_i's.

In particular, in the case $n=4$, we get the following identities for even functions $\hat{f}_i$:

$$
\int_{\mathbb{R}}(1-\chi(u))\widehat{f_1f_2f_3f_4}(u)du=\int_{\mathbb{R}^4}\big(1-\chi(u_1-u_2-u_3-u_4)\big)\prod_{i=1}^{4}\hat{f}_i(u_i)du_i, \tag{4.16}
$$

$$
\int_{\mathbb{R}^2}\big(1-\chi_2^*(u_a,u_b)\big)\hat{f}_a(u_a)\widehat{f_bf_cf_d}(u_b)du_adu_b=\int_{\mathbb{R}^4}\big(1-\chi_2^*(u_a,u_b-u_c-u_d)\big)\prod_{i=1}^{4}\hat{f}_i(u_i)du_i, \tag{4.17}
$$

$$\int_{\mathbb{R}^2}(1-\chi_2^*(u_a,u_c))\widehat{f_af_b}(u_a)\widehat{f_cf_d}(u_c)du_adu_c = \int_{\mathbb{R}^4}\left(1-\chi_2^*(u_a-u_b,u_c-u_d)\right)\prod_{i=1}^4 \hat{f}_i(u_i)du_i, \tag{4.18}$$

$$\int_{\mathbb{R}^3}(2-\chi_3^*(u_a,u_c,u_d))\widehat{f_af_b}(u_a)\hat{f}_c(u_c)\hat{f}_d(u_d)du_adu_cdu_d = \int_{\mathbb{R}^4}\left(2-\chi_3^*(u_a-u_b,u_c,u_d)\right)\prod_{i=1}^4 \hat{f}_i(u_i)du_i. \tag{4.19}$$

(The term with χ_4^* in the integrand is already an integral over $\mathbb{R}^4$, so we leave it unchanged.)

Finally, we rewrite each term as an integral over the region $\mathbb{R}^4_{\geq 0}$. Letting $\chi^{**}(u_1,\ldots,u_4)$ denote the $(k-1)!-\chi_k^*$ term after the change of variables, we break the integral into 16 pieces corresponding to the 'quadrants' of $\mathbb{R}^4$:

$$\int_{\mathbb{R}^4}\chi_k^{**}(u_1,\ldots,u_4)\prod_{i=1}^4\hat{f}_i(u_i)du_i = \sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4=\pm1}\int_{\mathbb{R}^4_{\geq 0}}\chi_k^{**}(\varepsilon_1u_1,\varepsilon_2u_2,\varepsilon_3u_3,\varepsilon_4u_4)\prod_{i=1}^4\hat{f}_i(u_i)du_i. \tag{4.20}$$

(Note that $\hat{f}_i(-u_i)=\hat{f}_i(u_i)$ since the test functions are even.)

4.2.2. *Simplifying the products* $\chi(\sum_i\varepsilon_iu_i)$. Now all of our terms have been desymmetrized into the region $\mathbb{R}^4_{\geq 0}$, so we can simplify the products of characteristic functions, using the fact that $0\leq u_i$ for each i. This assumption causes certain terms to be identically 1, allowing for some cancellation.

Most importantly, we have the following: in the region $\mathbb{R}^4_{\geq 0}$,

$$\chi(\varepsilon_1u_1+\cdots\varepsilon_nu_n)\cdot\chi(u_1+\cdots+u_n)=\chi(u_1+\cdots+u_n) \tag{4.21}$$

for all choices of signs $\varepsilon_i=\pm1$. To see this, we observe that

$$\left|\sum_{i=1}^4\varepsilon_iu_i\right|\leq\sum_{i=1}^4u_i, \tag{4.22}$$

so if $\chi(\sum_{i=1}^4u_i)=1$, it follows that $\chi(\sum_{i=1}^4\varepsilon_iu_i)=1$ for any choice of signs. Thus, in this case any product containing $\chi(\sum_{i=1}^4u_i)$ is 1. Of course, if instead $\chi(\sum_{i=1}^4u_i)=0$, then any product containing $\chi(\sum_{i=1}^4u_i)$ is 0.

After the desymmetrizing step, a large number of terms remain. Rather than reproducing the terms in full generality (there are hundreds of terms), we invoke the following stronger support assumption:

$$\mathrm{supp}(\hat{f}_i)\subset\left(-\frac{1}{2},\frac{1}{2}\right)$$

for each i. This enables the following additional simplification:

$$\chi(u_a + u_b - u_c - u_d) = 1, \tag{4.23}$$

since

$$|u_a + u_b - u_c - u_d| \leq \max(u_a + u_b, u_c + u_d) \leq 1.$$

This step greatly increases the amount of cancellation and simplification. We note that often we use test functions with the same support when computing n-level densities in practice (for instance, in computing the n-th moment, we use the same test function n times.) Later we show how to proceed without this simplifying assumption.

4.2.3. *An example computation.* We simplify the sums by applying the above simplification steps, then collecting like terms and counting. As an example computation, we have, after breaking the integral over $\mathbb{R}^4$ into 16 pieces,

$$\int_{\mathbb{R}^4} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i = \int_{\mathbb{R}^4_{\geq 0}} \Bigg(\chi(u_1 + u_2 + u_3 + u_4) + \chi(-u_1 - u_2 - u_3 - u_4)$$
$$+ \sum_{\{a,b,c\} \subset \{1,2,3,4\}}^{4} \chi(-u_a - u_b - u_c + u_d) + \sum_{\{a,b\} \subset \{1,2,3,4\}}^{6} \chi(-u_a - u_b + u_c + u_d)$$
$$+ \sum_{a=1}^{4} \chi(-u_a + u_b + u_c + u_d) \Bigg) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i.$$

The six $\chi(-u_a - u_b + u_c + u_d)$ terms are all identically 1. Also, since χ is an even function, the first two χ terms (with all $+$ or all $-$) are the same, and likewise

$$\chi(u_a - u_b - u_c - u_d) = \chi(-u_a + u_b + u_c + u_d).$$

So, by counting, we have

$$\int_{\mathbb{R}^4} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i = 2 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i$$
$$+ 2 \sum_{a=1}^{4} \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_a + u_b + u_c + u_d) \prod_{i=1}^{4} \hat{f}_i(u_i) du_i + 6 \int_{\mathbb{R}^4_{\geq 0}} \prod_{i=1}^{4} \hat{f}_i(u_i) du_i.$$

Applying these steps to each of the terms produce the following identities (assuming symmetric support):

$$\int_{\mathbb{R}}(1-\chi(u))\widehat{f_1\widehat{f_2 f_3} f_4}(u)du \tag{4.24}$$

$$= 10\int_{\mathbb{R}^4_{\geq 0}}\prod_{i=1}^4 \hat{f}_i(u_i)du_i - 2\int_{\mathbb{R}^4_{\geq 0}}\chi(u_1+u_2+u_3+u_4)\prod_{i=1}^4 \hat{f}_i(u_i)du_i$$

$$-2\sum_{\ell=1}^4\int_{\mathbb{R}^4_{\geq 0}}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i,$$

$$\int_{\mathbb{R}^2}(1-\chi_2^*(u_a,u_b))\hat{f}_a(u_a)\widehat{f_b f_c f_d}(u_b)du_a du_b = \tag{4.25}$$

$$= 16\int_{\mathbb{R}^4_{\geq 0}}\prod_{i=1}^4 \hat{f}_i(u_i)du_i - 4\int_{\mathbb{R}^4_{\geq 0}}\chi(u_1+u_2+u_3+u_4)\prod_{i=1}^4 \hat{f}_i(u_i)du_i$$

$$-4\sum_{\ell=b,c,d}\int_{\mathbb{R}^4_{\geq 0}}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i,$$

$$\int_{\mathbb{R}^2}(1-\chi_2^*(u_a,u_c))\widehat{f_a f_b}(u_a)\widehat{f_c f_d}(u_c)du_a du_c \tag{4.26}$$

$$= 12\int_{\mathbb{R}^4_{\geq 0}}\prod_{i=1}^4 \hat{f}_i(u_i)du_i - 4\int_{\mathbb{R}^4_{\geq 0}}\chi(u_1+u_2+u_3+u_4)\prod_{i=1}^4 \hat{f}_i(u_i)du_i$$

$$-4\int_{\mathbb{R}^4_{\geq 0}}\left(\prod_{\ell=a,b}\chi(-u_\ell+\sum_{j\neq\ell}u_j)+\prod_{\ell=c,d}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\right)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i,$$

$$\int_{\mathbb{R}^3}(2-\chi_3^*(u_a,u_c,u_d))\widehat{f_a f_b}(u_a)\hat{f}_c(u_c)\hat{f}_d(u_d)du_a du_c du_d \tag{4.27}$$

$$= 32\int_{\mathbb{R}^4_{\geq 0}}\prod_{i=1}^4 \hat{f}_i(u_i)du_i - 12\int_{\mathbb{R}^4_{\geq 0}}\chi(u_1+u_2+u_3+u_4)\prod_{i=1}^4 \hat{f}_i(u_i)du_i$$

$$-4\int_{\mathbb{R}^4_{\geq 0}}\Big(2\cdot\prod_{\ell=a,b}\chi(-u_\ell+\sum_{j\neq\ell}u_j)+\sum_{\ell=a,b}\chi(-u_\ell+\sum_{j\neq\ell}u_j)$$

$$+\prod_{\ell=c,d}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\Big)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i,$$

$$\int_{\mathbb{R}^4}(6-\chi_4^*(u_1,u_2,u_3,u_4))\prod_{i=1}^4 \hat{f}_i(u_i)du_i \tag{4.28}$$

$$= 96\int_{\mathbb{R}^4_{\geq 0}}\prod_{i=1}^4 \hat{f}_i(u_i)du_i - 48\int_{\mathbb{R}^4_{\geq 0}}\chi(u_1+u_2+u_3+u_4)\prod_{i=1}^4 \hat{f}_i(u_i)du_i$$

$$-8\sum_{\{a,b\}\subset S}^{6}\int_{\mathbb{R}^4_{\geq 0}}\prod_{\ell=a,b}\chi(-u_\ell+\sum_{j\neq\ell}u_j)\prod_{a,b,c,d}\hat{f}_i(u_i)du_i.$$

Substituting identities (4.24)-(4.28) into the expression (3.3) for $S_\star$ yields the following, considerably simpler, expression (assuming symmetric support):

$$\begin{aligned} S_\star = 24 \int_{\mathbb{R}^4_{\geq 0}} \prod_{i=1}^4 \hat{f}_i(u_i) du_i - 8 \sum_{\ell=1}^4 \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_\ell + \sum_{j \neq \ell} u_j) \prod_{i=1}^4 \hat{f}_i(u_i) du_i \qquad (4.29) \\ + 8 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^4 \hat{f}_i(u_i) du_i. \end{aligned}$$

This is the desired expression for $S_\star$ from equation (4.3). We now consider S_0.

4.2.4. *Desymmetrizing the term* S_0. We can show that the term S_0 matches with terms from the left-hand-side of (3.1) under the general support condition $\sum_{i=1}^4 |u_i| < 2$. We perform the same sequence of steps for the terms from S_0 that involve convolutions, but we do not assume $\operatorname{supp}(\hat{f}_i) \subset (-\frac{1}{2}, \frac{1}{2})$. (Indeed, under the stricter condition, the convolution terms of S_0 are all identically 0.) We rewrite Fourier transforms of products as multiple integrals (in this case, over $\mathbb{R}^2$). We then break each integral over $\mathbb{R}^2$ into 4 pieces corresponding to the four quadrants, simplify the sums if necessary, and collect terms. The result is the following:

$$\begin{aligned} S_0 = 4 \sum_{\{a,b\},\{c,d\}}^{3} \int_{\mathbb{R}} |u| \hat{f}_a(u) \hat{f}_b(u) du \int_{\mathbb{R}} |u| \hat{f}_c(u) \hat{f}_d(u) du \qquad (4.30) \\ + 4 \sum_{\{a,b\}}^{6} \int_{\mathbb{R}} |u| \hat{f}_a(u) \hat{f}_b(u) du \int_{\mathbb{R}^2_{\geq 0}} \big(\chi(u_c + u_d) - \chi(u_c - u_d)\big) \prod_{c,d} \hat{f}_i(u_i) du_i. \end{aligned}$$

This is the desired expression for S_0 from equation (4.2), and completes the proof of Lemma 4.1. $\square$

Now that S_0 and $S_\star$ have been desymmetrized into the region $\mathbb{R}^n_{\geq 0}$, we can match them with terms on the left-hand side of (3.1).

5. THE PROOF OF THEOREM 1.1, $n = 4$

We now complete the proof of Theorem 1.1 for the case $n = 4$ by showing that the terms $S_0, S_\star, \tilde{S}_0$ and $\tilde{S}_\star$ from equations (3.1) and (4.7) satisfy

$$S_0 = \tilde{S}_0$$
$$S_\star = \tilde{S}_\star.$$

First, by comparing expressions (4.2) and (4.8), we immediately see that $S_0 = \tilde{S}_0$ for $\hat{f}_1, \ldots, \hat{f}_4$ supported in $\sum_{i=1}^4 |u_i| < 2$. It remains to show that $S_\star = \tilde{S}_\star$.

For $S_\star$, the expressions obtained after desymmetrizing into the $\mathbb{R}^4_{\geq 0}$ region are very long, so for simplicity we first present the proof for the case where each $\hat{f}_i$ is supported in $(-\frac{1}{2}, \frac{1}{2})$. For the general support condition, we include a sketch of the argument in Section 5.2, without including line-by-line details.

5.1. **Proof of Theorem 1.1 under Symmetric Support.**

After desymmetrizing and simplifying the term $S_\star$, we have, from equation (4.3),

$$S_\star = 24 \int_{\mathbb{R}^4_{\geq 0}} \prod_{i=1}^4 \hat{f}_i(u_i) du_i - 8 \sum_{\ell=1}^4 \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_\ell + \sum_{j \neq \ell} u_j) \prod_{i=1}^4 \hat{f}_i(u_i) du_i \tag{5.1}$$
$$+ 8 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^4 \hat{f}_i(u_i) du_i.$$

On the other hand, we had reduced the left hand side to the following (equation (4.9)):

$$\tilde{S}_\star = 8 \sum_{\{a,b\},\{c,d\}}^3 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_a + u_b - u_c - u_d) \prod_{a,b,c,d} \hat{f}_i(u_i) du_i \tag{5.2}$$
$$+ 8 \int_{\mathbb{R}^4_{\geq 0}} \chi(\sum_{i=1}^4 u_i) \prod_{i=1}^4 \hat{f}_i(u_i) du_i - 8 \sum_{\ell=1}^4 \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_\ell + \sum_{j \neq \ell} u_j) \prod_{i=1}^4 \hat{f}_i(u_i) du_i.$$

We note that the first term in (5.2) has the characteristic function

$$\chi(u_a + u_b - u_c - u_d),$$

which is identically 1 in the $\mathbb{R}^4_{\geq 0}$ region when $\operatorname{supp}(\hat{f}_i) \leq \frac{1}{2}$ ($i = 1, \ldots, 4$), by the argument in Section 4.2. Consequently

$$\begin{aligned}\tilde{S}_\star = 24 \int_{\mathbb{R}^4_{\geq 0}} \prod_{i=1}^4 \hat{f}_i(u_i) du_i + 8 \int_{\mathbb{R}^4_{\geq 0}} \chi(u_1 + u_2 + u_3 + u_4) \prod_{i=1}^4 \hat{f}_i(u_i) du_i \\ - 8 \sum_{\ell=1}^4 \int_{\mathbb{R}^4_{\geq 0}} \chi(-u_\ell + \sum_{j \neq \ell} u_j) \prod_{i=1}^4 \hat{f}_i(u_i) du_i,\end{aligned} \tag{5.3}$$

which is identical to equation (4.3), so $S_\star = \tilde{S}_\star$. □

We have shown:

Theorem 5.1 (Symmetric Support)**.** *Assume GRH. Then for $n = 4$, equality holds in equation* (1.25) *if each f_i is an even Schwartz function and* $\operatorname{supp}(\hat{f}_i) \subset (-\frac{1}{2}, \frac{1}{2})$ *for each i.*

5.1.1. *Automating the remaining steps.* Although we have computed the remaining cancellations by hand, it is possible to automate these steps as well. While not necessary for the case where $\operatorname{supp}(\hat{f}_i) = \frac{1}{2}$ for all i, it becomes very helpful when we remove this assumption (see below). We briefly discuss the Mathematica code used to desymmetrize and simplify the integrals and products of χ terms.

The code proceeds in six steps:

(1) We repeatedly replace convolutions with integrals:

$$\widehat{f_1 \cdots f_n}[x] \mapsto \text{IntRN}[\{u_n\}, \widehat{f_1 \cdots f_{n-1}}[x + u_n] \hat{f}_n[u_n]].$$

(2) Then we repeatedly replace nested integrals with multiple integrals:

$$\text{IntRN}[vars, \text{IntRN}[vars2, f] \cdot g] \mapsto \text{IntRN}[\text{Join}[vars, vars2], f \cdot g].$$

(3) We perform the change of variables (4.15) dynamically: given the arguments of the functions $\hat{f}_i$, we solve the linear system that sends the i-th argument to u_i, and perform the corresponding replacements on the entire integrand. (Note that by the argument in Section 4, this change of variables always has Jacobian determinant 1.)

(4) We break each integral over $\mathbb{R}^n$ (represented by the symbol IntRN) into a sum of 2^n integrals, each over $\mathbb{R}^n_{\geq 0}$ (represented by IntR0N), using a replacement rule to change the signs of the appropriate u_i's for each summand.

(5) We expand each integral linearly:

$$\text{IntR0N}[vars, f + g] \mapsto \text{IntR0N}[vars, f] + \text{IntR0N}[vars, g]$$

$$\text{IntR0N}[vars, k\_\_Number f] \mapsto k \cdot \text{IntR0N}[vars, f].$$

(6) We apply the simplifications to the χ terms:

$$\chi[\ ]\cdot\chi[u_1+u_2+u_3+u_4]\mapsto\chi[u_1+u_2+u_3+u_4]$$
$$\chi[u_a+u_b-u_c-u_d]\mapsto 1 \text{ (if } \operatorname{supp}(\hat{f}_i)\le 1/2 \text{ for each } i.)$$

At this point, under the symmetric support assumption, we are left with 0. Otherwise we proceed as described below.

5.2. **Removing Symmetric Support.** Since Gao's result holds for the more general support condition of $\sum|u_i|<2$, we would like to show equality without having to assume that all the test functions have the same support. In fact, we are able to show this with some additional work.

Without the simplifying support assumption, however, the equation

$$\chi(u_a+u_b-u_c-u_d)=1 \text{ in } \mathbb{R}^4_{\ge 0}$$

no longer holds, since any u_i may be ≥ 1. Consequently, much of the cancellation no longer occurs right away, the resulting expressions are more unwieldy, and the last steps of the proof involve additional simplification steps. Rather than reproducing the expressions line-by-line (there are hundreds of lines after desymmetrizing the integrals), we include a sketch of the key steps in the proof. (The full computations can be reproduced and verified in the Mathematica notebook.)

Theorem 5.2 (General Support)**.** *Assume GRH. Then for $n\le 4$, equality holds in equation* (1.25) *if each f_i is an even Schwartz function and $\hat{f}_i$ is supported in $\sum_{i=1}^4|u_i|<2$.*

Sketch of proof. First, note that our proof of $S_0=\tilde{S}_0$ still holds; we need only verify the equation

$$S_\star-\tilde{S}_\star=0.$$

Initially we apply the steps described in section 4 to desymmetrize the terms of $S_\star$. We are left with sums of products of terms $\chi(\sum_{i=1}^4\varepsilon_iu_i)$. We proceed as follows:

Step 1. Rather than using the altered formula for the left-hand side of (3.1), we keep the original formula and use

$$\chi(u)=1-\tilde{\chi}(u)-\tilde{\chi}(-u) \tag{5.4}$$

to rewrite the χ terms from the right-hand side as $\tilde{\chi}$, where $\tilde{\chi}$ is the characteristic function of the interval $[1,\infty)$. The advantage of using $\tilde{\chi}$ again comes from not being even: to know when $\chi(\sum\varepsilon_iu_i)=0$, we need to consider both $\sum\varepsilon_iu_i>1$ and $\sum\varepsilon_iu_i<-1$, but with $\tilde{\chi}$ only the first case matters. This will facilitate several simplifications.

Step 2. We now adopt some notation. Letting A be the set of indices for which $\varepsilon_i = +1$ and B the set for which $\varepsilon_i = -1$, we separate by signs the sum

$$\sum_i \varepsilon_i u_i = \sum_A u_{a_i} - \sum_B u_{b_i}$$

and write

$$\tilde{\chi}_{A|B} := \tilde{\chi}(\sum_A u_{a_i} - \sum_B u_{b_i}). \tag{5.5}$$

For example, we write

$$\tilde{\chi}(u_1 + u_2 - u_3 - u_4) = \tilde{\chi}_{\{1,2\}|\{3,4\}}.$$

This notation allows us to reduce arguments about products of $\tilde{\chi}(\sum_i \varepsilon_i u_i)$ to set-theoretic and combinatorial arguments relating to the subsets A, B of $\{1, \ldots, n\}$.

Step 3. Finally, we order the test functions by their support: without loss of generality, we assume

$$\mathrm{supp}(\hat{f}_1) < \cdots < \mathrm{supp}(\hat{f}_n). \tag{5.6}$$

This assumption allows us to obtain bounds on the u_i's (for example, $\mathrm{supp}(\hat{f}_1) \leq \frac{2}{n}$), which we can use to eliminate certain $\tilde{\chi}$ terms.

The simplifications. Let A, B and A', B' be two decompositions of the set $\{1, \ldots, n\}$. The first simplification is as follows:

$$\text{If } A \subset A', \text{ then } \tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = \tilde{\chi}_{A|B}. \tag{5.7}$$

To see this, note that if $A \subset A'$, then $B' \subset B$, and in particular

$$\sum_{A'} u_{a'_i} \geq \sum_A u_{a_i} \text{ and } \sum_B u_{b_i} \geq \sum_{B'} u_{b'_i}.$$

Hence, $\sum_A u_{a_i} - \sum_B u_{b_i} > 1$ implies

$$\sum_{A'} u_{a'_i} - \sum_{B'} u_{b'_i} \geq \sum_A u_{a_i} - \sum_B u_{b_i} \geq 1.$$

So if $\tilde{\chi}_{A|B} = 1$, it follows that $\tilde{\chi}_{A'|B'} = 1$. Thus, given products of $\tilde{\chi}$ terms, we need only keep the $\tilde{\chi}$'s whose A's are minimal with respect to containment (i.e., the fewest positive signs).

The second simplification is:

$$\text{if } A \cap A' = \varnothing, \text{ then } \tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = 0. \tag{5.8}$$

To see this, observe that

$$\sum_A u_{a_i} + \sum_{A'} u_{a'_i} \leq \sum_{i=1}^{n} u_i < 2,$$

so we cannot have both $\sum_A u_{a_i} > 1$ and $\sum_{A'} u_{a'_i} > 1$.

In the case $n = 4$, the steps (5.7) and (5.8) reduce the hundreds of lines of sums of products of $\tilde{\chi}$ terms to a sum of 18 remaining terms.

The remaining simplifications take advantage of the assumption that the test functions are ordered by support.

For $A \subset \{1, \ldots, n\}$, by the pigeonhole principle

$$\max_A(\mathrm{supp}(\hat{f}_{a_i})) \geq \frac{1}{|A|} \sum_A \mathrm{supp}(\hat{f}_{a_i}). \tag{5.9}$$

Applying this to the support regions,

$$\begin{aligned} 2 > \sum_{i=1}^{n} \mathrm{supp}(\hat{f}_i) &= \sum_{i=1}^{\max\{k \in A\}} \mathrm{supp}(\hat{f}_i) + \sum_{\max\{k \in A\}+1}^{n} \mathrm{supp}(\hat{f}_i) \\ &\geq \sum_A \mathrm{supp}(\hat{f}_{a_i}) \ + \ (n - \max\{k \in A\}) \cdot \max_A(\mathrm{supp}(\hat{f}_{a_i})), \end{aligned}$$

which yields a general bound

$$2 > \left(1 + \frac{n - \max\{k \in A\}}{|A|}\right) \cdot \sum_A \mathrm{supp}(\hat{f}_{a_i}). \tag{5.10}$$

In particular, we have

$$\sum_A \mathrm{supp}(\hat{f}_{a_i}) < 1 \text{ whenever } |A| + \max\{k \in A\} \leq n. \tag{5.11}$$

This condition implies that many of the terms we encounter are identically zero. In particular, we can detect when certain $\tilde{\chi}$ terms or products of $\tilde{\chi}$ terms are zero. We state this as a lemma.

Lemma 5.3. *Let A, B and A', B' be decompositions of $\{1, \dots, n\}$, that is, $A \cup B = \{1, \dots, n\}$ and $A \cap B = \varnothing$, and similarly for A', B'. Then:*

$$\int_{\mathbb{R}^n_{\geq 0}} \tilde{\chi}_{A|B} \prod_{i=1}^{n} \hat{f}_i(u_i) du_i = 0 \textit{ whenever } |A| + \max\{k \in A\} \leq n. \tag{5.12}$$

$$\int_{\mathbb{R}^n_{\geq 0}} \tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} \prod_{i=1}^{n} \hat{f}_i(u_i) du_i = 0 \textit{ whenever } |A \cap A'| + \max\{k \in A \cap A'\} \leq n. \tag{5.13}$$

Proof. For (5.12), note that $\sum_A \operatorname{supp}(\hat{f}_a) < 1$ by the bound (5.11). Hence

$$\sum_A u_a - \sum_B u_b < 1$$

everywhere in the region of integration, so $\tilde{\chi}_{A|B}$ is always zero.

Similarly, for (5.13), note that if $\tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = 1$, then $\sum_A u_a - \sum_B u_b > 1$ and $\sum_{A'} u_{a'} - \sum_{B'} u_{b'} > 1$. Adding these inequalities together, we have

$$\begin{aligned} 2 &< \sum_A u_{a_i} - \sum_B u_{b_i} + \sum_{A'} u_{a'_i} - \sum_{B'} u_{b'_i} \\ &= 2 \cdot \left(\sum_{A \cap A'} u_a - \sum_{B \cap B'} u_b \right) \\ &< 2 \cdot \sum_{A \cap A'} u_a, \end{aligned}$$

so we must have $\sum_{A \cap A'} u_a > 1$. By condition (5.11), this cannot occur. $\square$

Conditions (5.12) and (5.13) allow us to show that the remaining terms are identically zero. For example, these conditions imply the following:

(1) $\tilde{\chi}_{A|B} = 0$ for all $A \subseteq \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$,
(2) $\tilde{\chi}_{A|B} = 0$ for all $A = \{u_i\}$, $i \leq n-1$.
(3) $\tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = 0$ whenever $A \cap A'$ is one of the above subsets.

These arguments are sufficient to show that $S_\star - \tilde{S}_\star = 0$ in the case $n = 4$, concluding the proof of Theorem 1.1. $\square$

6. Where to go from here

6.1. **The Cases $n = 5, 6$.** We have shown that the Density Conjecture (1.25) holds for $n = 4$. In fact, we can show equality for $n = 5, 6$ using the same method, by using the earlier results to reduce the equation (for $n = 5$) to the form

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{-32}{\log^5 X} \prod_{i=1}^{5} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{f_i}\left(\frac{\log p}{\log X}\right)\right) = S_0 + S_\star,$$

where S_0 is the collection of terms having a factor $\int_{\mathbb{R}} |u| \hat{f}_a(u)\hat{f}_b(u)du$, and $S_\star$ consists of the remaining terms, which involve all 5 test functions, convolved together every possible way:

$$\begin{aligned} S_\star = & -8 \int_{\mathbb{R}} (1-\chi(u))\widehat{f_1\cdots f_5}(u)du \\ & + 8 \sum_{\{\{a\},\{b,c,d,e\}\}}^{5} \int_{\mathbb{R}^2} (1-\chi_2^*(u_1,u_2))\hat{f}_a(u_1)\widehat{f_b\cdots f_e}(u_2)du_1du_2 \\ & + 8 \sum_{\{\{a,b\},\{c,d,e\}\}}^{10} \int_{\mathbb{R}^2} (1-\chi_2^*(u_1,u_2))\widehat{f_af_b}(u_1)\widehat{f_cf_df_e}(u_2)du_1du_2 \\ & - 8 \sum_{\{\{a\},\{b\},\{c,d,e\}\}}^{10} \int_{\mathbb{R}^3} (1-\chi_3^*(u_1,u_2,u_3))\hat{f}_a(u_1)\hat{f}_b(u_2)\widehat{f_cf_df_e}(u_3)du_1du_2du_3 \\ & - 8 \sum_{\{\{a\},\{b,c\},\{d,e\}\}}^{15} \int_{\mathbb{R}^3} (1-\chi_3^*(u_1,u_2,u_3))\hat{f}_a(u_1)\widehat{f_bf_c}(u_2)\widehat{f_df_e}(u_3)du_1du_2du_3 \\ & + 8 \sum_{\{\{a\},\{b\},\{c\},\{d,e\}\}}^{10} \int_{\mathbb{R}^4} (1-\chi_4^*(u_1,u_2,u_3,u_4))\hat{f}_a(u_1)\hat{f}_b(u_2)\hat{f}_c(u_3)\widehat{f_df_e}(u_4)du_1du_2du_3du_4 \\ & - 8 \int_{\mathbb{R}^5} (1-\chi_5^*(u_1,u_2,u_3,u_4,u_5)) \prod_{i=1}^{5} \hat{f}_i(u_i)du_i. \end{aligned} \tag{6.1}$$

If we then expand the $\sum_p$ term using Gao's formula, we can use our work on the cases $n \le 4$ to show directly that the $\int_{\mathbb{R}} |u|\hat{f}_a(u)\hat{f}_b(u)du$ terms cancel with S_0, and we are left with the term $\tilde{S}_\star$ consisting of integrals over $\mathbb{R}^5_{\ge 0}$:

$$\tilde{S}_\star = 16 \sum_{I \subsetneq \{1,2,3,4,5\}} (-1)^{|I|} \int_{\mathbb{R}^5_{\ge 0}} \tilde{\chi}(\sum_{I^c} u_i - \sum_I u_i) \prod_{i=1}^{5} \hat{f}_i(u_i)du_i, \tag{6.2}$$

where the sum ranges over the proper subsets $I \subsetneq \{1,2,3,4,5\}$ and $I^c = \{1,2,3,4,5\} \setminus I$.

We use the same steps as in Section 4.2 to convert the convolution terms of $S_\star$ into integrals over $\mathbb{R}^5$, then break them up into sums of integrals over $\mathbb{R}^5_{\geq 0}$. We then replace the characteristic functions $\chi = \mathbb{I}_{[-1,1]}$ with $\tilde{\chi} = \mathbb{I}_{[1,\infty)}$ via

$$\chi(u) = 1 - \tilde{\chi}(u) - \tilde{\chi}(-u). \tag{6.3}$$

Finally, we use the combinatorial arguments from equations (5.7) - (5.12) to deal with the resulting sums of products of $\tilde{\chi}$ terms.

Running these computations in Mathematica confirms that equality holds in the case $n = 5$. Applying this procedure to $n = 6$ also works, with two caveats. First, the final steps become very slow. We are rewriting each integral over $\mathbb{R}^6$ as a sum of $2^6 = 64$ integrals over $\mathbb{R}^6_{\geq 0}$, and we have products of up to six χ terms before we apply the substitution for $\tilde{\chi}$. In the final step, before applying the combinatorial arguments, there were (tens of) thousands of terms, and the computations had to be divided into substeps in order to run successfully. The computational complexity of this method is more than exponential, and so is highly prohibitive for larger n.

The second problem in the case $n = 6$ is that three terms remain after the final step: we end up with $S_\star - \tilde{S}_\star = R$, where

$$\begin{aligned}
R = & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.4)\\
& \tilde{\chi}(u_1 + u_2 + u_3 + u_4 + u_5 - u_6)\tilde{\chi}(-u_1 + u_2 + u_3 + u_4 - u_5 + u_6)\tilde{\chi}(u_1 - u_2 - u_3 - u_4 + u_5 + u_6)\\
& + \tilde{\chi}(u_1 + u_2 + u_3 + u_4 + u_5 - u_6)\tilde{\chi}(u_1 - u_2 + u_3 + u_4 - u_5 + u_6)\tilde{\chi}(-u_1 + u_2 - u_3 - u_4 + u_5 + u_6)\\
& + \tilde{\chi}(u_1 + u_2 + u_3 + u_4 + u_5 - u_6)\tilde{\chi}(u_1 + u_2 - u_3 + u_4 - u_5 + u_6)\tilde{\chi}(-u_1 - u_2 + u_3 - u_4 + u_5 + u_6)\\
= & \tilde{\chi}_{\{1,2,3,4,5\}|\{6\}} \cdot \tilde{\chi}_{\{2,3,4,6\}|\{1,5\}} \cdot \tilde{\chi}_{\{1,5,6\}|\{2,3,4\}}\\
& + \tilde{\chi}_{\{1,2,3,4,5\}|\{6\}} \cdot \tilde{\chi}_{\{1,3,4,6\}|\{2,5\}} \cdot \tilde{\chi}_{\{2,5,6\}|\{1,3,4\}}\\
& + \tilde{\chi}_{\{1,2,3,4,5\}|\{6\}} \cdot \tilde{\chi}_{\{1,2,4,6\}|\{3,5\}} \cdot \tilde{\chi}_{\{3,5,6\}|\{1,2,4\}}.
\end{aligned}$$

Note that the arguments from equations (5.7) - (5.12) do not simplify these leftover terms: in terms of the notation $\chi_{A|B}$, for each of the three products, there are no subset containment relations between the $A's$; none of the A's individually satisfy the bound (5.12); and none of the pairs A, A' are such that $A \cap A'$ satisfies the bound (5.11).

Luckily, we have the following simple argument:

Lemma 6.1. *Each summand of the term R in equation* (6.4) *is identically 0.*

Proof. Consider the second and third factors in the first term. In order for to have $\tilde{\chi}_{\{2,3,4,6\}|\{1,5\}} \cdot \tilde{\chi}_{\{1,5,6\}|\{2,3,4\}} \neq 0$, we must have

$$-u_1 + u_2 + u_3 + u_4 - u_5 + u_6 > 1$$
$$u_1 - u_2 - u_3 - u_4 + u_5 + u_6 > 1,$$

so by adding the inequalities we have $2u_6 > 2$, hence $u_6 > 1$. Similarly, by examining the other terms, we observe that $\{1,3,4,6\} \cap \{2,5,6\} = \{6\}$ and $\{1,2,4,6\} \cap \{3,5,6\} = \{6\}$, so those terms also only contribute when $u_6 > 1$. However, all three terms include the factor $\tilde{\chi}_{\{1,2,3,4,5\}|\{6\}}$, which contributes only when

$$u_1 + u_2 + u_3 + u_4 + u_5 - u_6 > 1,$$

which is impossible when $u_6 > 1$ and the $\hat{f}_i$ have a total support of 2. $\square$

We are thus still able to confirm the density conjecture for $n = 6$. Still, the arguments detailed in Section 5.2 are insufficient to generalize to $n = 6$ directly, and we certainly expect that the necessary combinatorics would become more complex as n increases.

6.2. **Towards a General Argument.** In each of the cases $n \leq 6$, the final step was to show an identity between integrals over $\mathbb{R}^k$ (for various $k \leq n$) of convolutions of the n test functions (multiplied by characteristic functions), and integrals over $\mathbb{R}^n_{\geq 0}$ of $\prod_{i=1}^n \hat{f}_i(u_i) du_i$ (multiplied by a different characteristic function). For these cases, we were able to show equality by altering the $\mathbb{R}^k$ terms via the steps outlined in Section 4:

(1) Rewrite the convolutions as multiple integrals, turning integrals over $\mathbb{R}^k$, $k < n$, into integrals over $\mathbb{R}^n$.
(2) Change variables so that every integrand has $\prod_{i=1}^n \hat{f}_i(u_i) du_i$.
(3) Break each integral over $\mathbb{R}^n$ into 2^n pieces corresponding to the orthants of $\mathbb{R}^n$.
(4) Replace each characteristic function $\chi = \mathbb{I}_{[-1,1]}$ with $\tilde{\chi} = \mathbb{I}_{[1,\infty]}$ via $\chi(u) = 1 - \tilde{\chi}(u) - \tilde{\chi}(-u)$.
(5) Simplify the expressions using combinatorics and support arguments.

Unfortunately, the third and fourth steps increase the number of terms exponentially, making this method intractable to compute explicitly for larger n. In order to prove the Density Conjecture for all n, we wish to find another approach to this argument.

In this section, we discuss how to generalize the above argument in two steps (though note that we do not currently have proofs for either step). First, we state a conjecture for the general form of the identity, involving the integrals of the test functions, that must be shown for each n. Second, we consider how to show that this identity actually implies the Density Conjecture in general.

The two conjectures are as follows:

Conjecture 6.2. *Let $f_1, \ldots, f_n$ be even Schwartz functions such that $\hat{f}_1, \ldots, \hat{f}_n$ are supported in $\sum_{i=1}^{n} |u_i| < 2$. Then*

$$S_\star(f_1, \ldots, f_n) = 2\, \tilde{S}_\star(f_1, \ldots, f_n), \tag{6.5}$$

where

$$S_\star(f_1, \ldots, f_n) = \tag{6.6}$$

$$\sum_{\underline{F}} (-1)^{n-\nu(\underline{F})} \int_{\mathbb{R}^{\nu(\underline{F})}} \left((\nu(\underline{F}) - 1)! - \chi^*_{\nu(\underline{F})}(u_1, \ldots, u_{\nu(\underline{F})})\right) \prod_{\ell=1}^{\nu(\underline{F})} \hat{F}_\ell(u_\ell) du_\ell$$

and

$$\tilde{S}_\star(f_1, \ldots, f_n) = \sum_{I \subsetneq \{1,\ldots,n\}} (-1)^{|I|} \int_{\mathbb{R}^n_{\geq 0}} \tilde{\chi}(\sum_{I^c} u_i - \sum_{I} u_i) \prod_{i=1}^{n} \hat{f}_i(u_i) du_i. \tag{6.7}$$

Here $\underline{F}$ ranges over the partitions of $\{1, \ldots, n\}$ into disjoint subsets, numbered as $\underline{F} = \{F_1, \ldots, F_{\nu(\underline{F})}\}$, and we write $F_\ell(x) = \prod_{i \in F_\ell} f_i(x)$. The functions χ^*_k are defined as in equation (2.4)*, and $\tilde{\chi}$ is the characteristic function of $[1, \infty)$. Finally, I ranges over the proper subsets of $\{1, 2, 3, 4, 5\}$ and $I^c = \{1, 2, 3, 4, 5\} \setminus I$.*

Conjecture 6.3. *Let $f_1, \ldots, f_n$ be even Schwartz functions such that $\hat{f}_1, \ldots, \hat{f}_n$ are supported in $\sum_{i=1}^{n} |u_i| < 2$. If Conjecture 6.2 holds for $1 \leq k \leq n$, then Theorem 1.1 is true for n.*

We do not have a proof for Conjecture 6.2 for $n > 6$, due to the complexity of the $\chi^*_{\nu(\underline{F})}$ terms and the rapid growth in the number and form of the partitions $\underline{F}$. We note that the terms of S^* greatly outnumber the terms of $\tilde{S}^*$, so we expect a lot of cancellation on the left-hand-side of (6.5). It may be the case that equality in (6.5) should somehow hold term-by-term (after grouping terms from S^* together in some way) – if so, it might be possible to break down the conjecture into simpler steps.

For the cases $n = 1, 2, 3$, Conjecture 6.2 holds by [4], pages 57-58 and equation 5.7. (We quoted related identities in section 2.2.1.)

In the case $n = 4$, the terms $S_\star$ and $\tilde{S}_\star$ defined in Lemma 3.1 and equation (4.9) are identical to $4\ S_\star(f_1, f_2, f_3, f_4)$ and $8\ \tilde{S}_\star(f_1, f_2, f_3, f_4)$ as defined above, and we showed equality in section 5.2. Equality up to $n = 6$ is discussed in Section 6.1. For each of these cases, the details can be checked in the Mathematica notebooks.

Conjecture 6.3 is easier to deal with in principle, though we do not have a proof for $n > 6$. All the terms of $S_\star$ arise from the random matrix theory expression, while all the terms of $\tilde{S}_\star$ arise from the number theory side, in the last part of Gao's formula (2.12) for the $\sum_p$ terms. In fact, with the above notation, we can write the formula for the $\sum_p$ terms as:

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{1}{\log^n X} \prod_{i=1}^{n} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{g}_i \left(\frac{\log p}{\log X}\right)\right) = \tag{6.8}$$

$$\left(\frac{1+(-1)^n}{2}\right) \frac{1}{2^{n/2}} \sum_{(A;B)} \prod_{i=1}^{n/2} \int_{\mathbb{R}} |u_i| \hat{g}_{a_i}(u_i) \hat{g}_{b_i}(u_i) du_i$$

$$-\frac{1}{2} \sum_{\substack{S\subsetneq\{1,\dots,n\} \\ |S| \text{ even}}} \frac{1}{2^{|S|/2}} \left(\sum_{(A_S;B_S)} \prod_{i=1}^{|S|/2} \int_{\mathbb{R}} u_i \hat{g}_{a_i}(u_i) \hat{g}_{b_i}(u_i) du_i \right) \cdot \tilde{S}_\star(g_i : i \in S^c),$$

where $\sum_{(A;B)}$ ranges over the ways of pairing up the elements of the set $\{1, \dots, n\}$, and similarly for $\sum_{(A_S;B_S)}$ and the set S.

Thus, by substituting $\frac{1}{2}S_\star$ for $\tilde{S}_\star$ in the above formula, we can easily check Conjecture 6.3 computationally, though it is more difficult to check Conjecture 6.2.

Appendix A. The 3-level Density

We show how our approach systematizes the final step in proving the Density Conjecture for $n = 3$. We follow the approach described in Sections 3 through 5.2.

Note that our approach is the same as in Section 5.1 of [4] for the first few steps (equations (5.5) through (5.7)): we cancel the terms that appear identically in the number theory and random matrix theory expressions, and use the results from $n = 1, 2$ to match the terms

$$\lim_{X\to\infty} \frac{\pi^2}{4X} \sum_{d\in D(X)} \frac{1}{\log^k X} \prod_{i=1}^{k} \left(\sum_p \frac{\log p}{\sqrt{p}} \left(\frac{8d}{p}\right) \widehat{F}_i \left(\frac{\log p}{\log X}\right)\right)$$

for $k = 1, 2$. We then use Gao's formula (2.12) to compute the last term of this form (where $k = 3$). All the terms containing factors $\int_{\mathbb{R}} \hat{f}_a(u)\hat{f}_b(u)du$ cancel immediately, and we are left with the following equation (5.7 from [4]):

$$\begin{aligned} 2\int_{\mathbb{R}^3_{\geq 0}} \Big(&\tilde{\chi}(u_1 - u_2 - u_3) + \tilde{\chi}(-u_1 + u_2 - u_3) + \tilde{\chi}(-u_1 - u_2 + u_3) \\ &- \tilde{\chi}(u_1 + u_2 - u_3) - \tilde{\chi}(u_1 - u_2 + u_3) - \tilde{\chi}(-u_1 + u_2 + u_3) \\ &+ \tilde{\chi}(u_1 + u_2 + u_3)\Big) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i \\ = &\int_{\mathbb{R}} (1-\chi(u))\widehat{f_1 f_2 f_3}(u)du - \int_{\mathbb{R}^2} (1 - \chi_2^*(u, v))\hat{f}_1(u)\widehat{f_2 f_3}(v)dudv \\ - &\int_{\mathbb{R}^2} (1-\chi_2^*(u, v))\hat{f}_2(u)\widehat{f_1 f_3}(v)dudv - \int_{\mathbb{R}^2} (1 - \chi_2^*(u, v))\hat{f}_3(u)\widehat{f_1 f_2}(v)dudv \\ + &\int_{\mathbb{R}^3} (2-\chi_3^*(u_1, u_2, u_3)) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i. \end{aligned} \tag{A.1}$$

Here χ_k^* is defined as in equation (2.4), so

$$\begin{aligned} \chi_2^*(u, v) &= \chi(u + v)\chi(u - v), \\ \chi_3^*(u_1, u_2, u_3) &= \chi(u_1 - u_2 - u_3)\chi(u_1 + u_2 - u_3)\chi(u_1 + u_2 + u_3) \\ &\quad + \chi(u_1 - u_2 - u_3)\chi(u_1 - u_2 + u_3)\chi(u_1 + u_2 + u_3). \end{aligned}$$

The remaining steps in our proof are different from those used in [4]. Also note that, with the notation of Section 6, equation (A.1) is just the $n = 3$ version of Conjecture 6.2, i.e. the equation $2\,\tilde{S}_\star(f_1, f_2, f_3) = S_\star(f_1, f_2, f_3)$.

The first step is to expand each term on the right-hand side into an integral over $\mathbb{R}^3$, and apply the change of variables from Section 4. This makes the right-hand side equivalent to

$$\int_{\mathbb{R}^3}(1-\chi(u_1-u_2-u_3))\prod_{i=1}^{3}\hat{f}_i(u_i)du_i+\int_{\mathbb{R}^3}(2-\chi_3^*(u_1,u_2,u_3))\prod_{i=1}^{3}\hat{f}_i(u_i)du_i \quad \text{(A.2)}$$
$$-\int_{\mathbb{R}^3}(1-\chi_2^*(u_2,u_1-u_3))\prod_{i=1}^{3}\hat{f}_i(u_i)du_i-\int_{\mathbb{R}^3}(1-\chi_2^*(u_3,u_1-u_2))\prod_{i=1}^{3}\hat{f}_i(u_i)du_i$$
$$-\int_{\mathbb{R}^3}(1-\chi_2^*(u_1,u_2-u_3))\prod_{i=1}^{3}\hat{f}_i(u_i)du_i.$$

We can collect the integrals together (the 1's cancel) to get

$$\int_{\mathbb{R}^3}\Big(-\chi(u_1-u_2-u_3)+\chi_2^*(u_1,u_2-u_3)+\chi_2^*(u_2,u_1-u_3) \quad \text{(A.3)}$$
$$+\chi_2^*(u_3,u_1-u_2)-\chi_3^*(u_1,u_2,u_3)\Big)\prod_{i=1}^{3}\hat{f}_i(u_i)du_i.$$

Now we break up the integral into a sum of 8 integrals corresponding to the distinct octants of $\mathbb{R}^3$. The right-hand side becomes the following:

$$2\int_{\mathbb{R}^3_{\geq 0}}\Big(-\chi(u_1-u_2-u_3)-\chi(u_1+u_2-u_3)-\chi(u_1-u_2+u_3)-\chi(u_1+u_2+u_3)$$
$$+2\chi(u_1+u_2-u_3)\chi(u_1-u_2-u_3)+2\chi(u_1-u_2+u_3)\chi(u_1+u_2+u_3)$$
$$+2\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3)+2\chi(u_1+u_2+u_3)\chi(u_1-u_2-u_3)$$
$$+2\chi(u_1+u_2-u_3)\chi(u_1-u_2+u_3)+2\chi(u_1+u_2-u_3)\chi(u_1+u_2+u_3)$$
$$-2\chi(u_1+u_2-u_3)\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3)$$
$$-2\chi(u_1+u_2-u_3)\chi(u_1+u_2+u_3)\chi(u_1-u_2-u_3)$$
$$-2\chi(u_1-u_2+u_3)\chi(u_1+u_2+u_3)\chi(u_1-u_2-u_3)$$
$$-2\chi(u_1+u_2-u_3)\chi(u_1-u_2+u_3)\chi(u_1+u_2+u_3)\Big)\prod_{i=1}^{3}\hat{f}_i(u_i)du_i.$$
(A.4)

Recall that, by equation (4.21), in the region $\mathbb{R}^3_{\geq 0}$, we have

$$\chi(\varepsilon_1u_1+\varepsilon_2u_2+\varepsilon_3u_3)\chi(u_1+u_2+u_3)=\chi(u_1+u_2+u_3),$$

for any signs $\varepsilon_i = \pm 1$. Applying this simplification reduces expression (A.5) to

$$
\begin{aligned}
2\int_{\mathbb{R}^3_{\geq 0}} \Big(& -\chi(u_1-u_2+u_3) - \chi(u_1+u_2+u_3) - \chi(u_1-u_2-u_3) \\
& -\chi(u_1+u_2-u_3) + 2\chi(u_1+u_2-u_3)\chi(u_1-u_2-u_3) \\
& + 2\chi(u_1+u_2-u_3)\chi(u_1-u_2+u_3) + 2\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3) \\
& - 2\chi(u_1+u_2-u_3)\chi(u_1-u_2+u_3)\chi(u_1-u_2-u_3)\Big) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i.
\end{aligned} \tag{A.5}
$$

We now use the substitution $\chi(u) = 1 - \tilde{\chi}(u) - \tilde{\chi}(-u)$. To shorten the notation, we use the notation from Section 5.2 and write $\tilde{\chi}(\sum_A u_a - \sum_B u_b) = \tilde{\chi}_{A|B}$. The expression becomes the following:

$$
\begin{aligned}
2\int_{\mathbb{R}^3} \Big(& -\tilde{\chi}_{\{1\}|\{2,3\}} - \tilde{\chi}_{\{2\}|\{1,3\}} - \tilde{\chi}_{\{3\}|\{1,2\}} - \tilde{\chi}_{\{1,2\}|\{3\}} - \tilde{\chi}_{\{1,3\}|\{2\}} - \tilde{\chi}_{\{2,3\}|\{1\}} \\
& + \tilde{\chi}_{\{1,2,3\}|\{\}} + 2\tilde{\chi}_{\{1\}|\{2,3\}}\tilde{\chi}_{\{2\}|\{1,3\}}\tilde{\chi}_{\{3\}|\{1,2\}} + 2\tilde{\chi}_{\{1\}|\{2,3\}}\tilde{\chi}_{\{2\}|\{1,3\}}\tilde{\chi}_{\{1,2\}|\{3\}} \\
& + 2\tilde{\chi}_{\{1\}|\{2,3\}}\tilde{\chi}_{\{3\}|\{1,2\}}\tilde{\chi}_{\{1,3\}|\{2\}} + 2\tilde{\chi}_{\{1\}|\{2,3\}}\tilde{\chi}_{\{1,2\}|\{3\}}\tilde{\chi}_{\{1,3\}|\{2\}} \\
& + 2\tilde{\chi}_{\{2\}|\{1,3\}}\tilde{\chi}_{\{3\}|\{1,2\}}\tilde{\chi}_{\{2,3\}|\{1\}} + 2\tilde{\chi}_{\{2\}|\{1,3\}}\tilde{\chi}_{\{1,2\}|\{3\}}\tilde{\chi}_{\{2,3\}|\{1\}} \\
& + 2\tilde{\chi}_{\{3\}|\{1,2\}}\tilde{\chi}_{\{1,3\}|\{2\}}\tilde{\chi}_{\{2,3\}|\{1\}} + 2\tilde{\chi}_{\{1,2\}|\{3\}}\tilde{\chi}_{\{1,3\}|\{2\}}\tilde{\chi}_{\{2,3\}|\{1\}}\Big) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i.
\end{aligned} \tag{A.6}
$$

Finally, we apply the arguments from Section 5.2 to argue that many of the terms are identically zero. We assume without loss of generality that

$$\operatorname{supp}(\hat{f}_1) \leq \operatorname{supp}(\hat{f}_2) \leq \operatorname{supp}(\hat{f}_3).$$

First we apply condition (5.12), which states that $\chi_{A|B} = 0$ if $\max\{k \in A\} + |A| \leq 3$. In particular, this means that $\tilde{\chi}_{\{1\}|\{2,3\}} = \tilde{\chi}_{\{2\}|\{1,3\}} = 0$. We remove those terms and are left with the following:

$$
\begin{aligned}
2\int_{\mathbb{R}^3} \Big(& -\tilde{\chi}_{\{3\}|\{1,2\}} - \tilde{\chi}_{\{1,2\}|\{3\}} - \tilde{\chi}_{\{1,3\}|\{2\}} - \tilde{\chi}_{\{2,3\}|\{1\}} + \tilde{\chi}_{\{1,2,3\}|\{\}} \\
& + 2\tilde{\chi}_{\{3\}|\{1,2\}}\tilde{\chi}_{\{1,3\}|\{2\}}\tilde{\chi}_{\{2,3\}|\{1\}} + 2\tilde{\chi}_{\{1,2\}|\{3\}}\tilde{\chi}_{\{1,3\}|\{2\}}\tilde{\chi}_{\{2,3\}|\{1\}}\Big) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i.
\end{aligned} \tag{A.7}
$$

Next, we apply condition (5.7), which states that

$$\tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = \tilde{\chi}_{A|B} \text{ whenever } A \subset A'.$$

This reduces the second-to-last term in (A.7) to just $+2\tilde{\chi}_{\{3\}|\{1,2\}}$.

Finally, we apply condition (5.13), which states that

$$\tilde{\chi}_{A|B} \cdot \tilde{\chi}_{A'|B'} = 0 \text{ if } \max\{A \cap A'\} + |A \cap A'| \leq 3.$$

The last term of (A.7) includes the product $\tilde{\chi}_{\{1,2\}|\{3\}}\tilde{\chi}_{\{1,3\}|\{2\}}$; since $\{1,2\} \cap \{1,3\} = \{1\}$, the condition applies, so this term is 0. Hence the right-hand side is just

$$2\int_{\mathbb{R}^3} \left(\tilde{\chi}_{\{3\}|\{1,2\}} - \tilde{\chi}_{\{1,2\}|\{3\}} - \tilde{\chi}_{\{1,3\}|\{2\}} - \tilde{\chi}_{\{2,3\}|\{1\}} + \tilde{\chi}_{\{1,2,3\}|\{\}}\right) \prod_{i=1}^{3} \hat{f}_i(u_i)du_i. \quad \text{(A.8)}$$

Except for the terms $\tilde{\chi}_{\{1\}|\{2,3\}}$ and $\tilde{\chi}_{\{2\}|\{1,3\}}$, this is the same as the left-hand side in (A.1). But both of these ‘missing’ terms are identically 0. Hence (A.1) holds, which proves the Density Conjecture for $n = 3$.

Appendix B. Notation Glossary

The expressions for the n-level densities include sums indexed by combinatorial objects, involving notation that may be confusing. We include here an example of each of the notations used. This section can (and should) be skipped by those who are already comfortable with the notation.

B.1. Sums over Pairings.

The notation

$$\sum_{(A;B)} \prod_{i=1}^{n/2} \int_{\mathbb{R}} \hat{f}_{a_i}(u)\hat{f}_{b_i}(u)du \quad \text{(B.1)}$$

is a sum over the ways of pairing up the elements of $\{1,\ldots,n\}$. The sets $A = \{a_1,\ldots,a_t\}$ and $B = \{b_1,\ldots,b_t\}$ are meant to list (in order) the first and second elements from each pair. Note that there are $(n-1)!!$ ways to do so.

For example, one way of pairing up the elements $\{1,2,3,4,5,6\}$ is

$$\{1,4\}, \{6,2\}, \{3,5\}.$$

The sets A and B for this pairing are $A = \{1,6,3\}$, $B = \{4,2,5\}$. Hence the term corresponding to A and B is the product

$$\int_{\mathbb{R}} \hat{f}_1(u)\hat{f}_4(u)du \cdot \int_{\mathbb{R}} \hat{f}_6(u)\hat{f}_2(u)du \cdot \int_{\mathbb{R}} \hat{f}_3(u)\hat{f}_5(u)du. \quad \text{(B.2)}$$

B.2. Sums and Products over Permutations.

The notation

$$\sum_{\{i::\{1,\dots,n\}\}} \left(\prod_{k=1}^{n} \chi(u_{i_1} + \cdots + u_{i_k} - u_{i_{k+1}} - \cdots - u_{i_n}) \right) \tag{B.3}$$

is a sum over the cyclic permutations of the elements of $\{1,\dots,n\}$. (Here χ is the characteristic function of the interval $[-1,1]$.) By a cyclic permutation we mean a permutation of the form $(i_1 \to i_2 \to \cdots \to i_n)$, where the i_j are a reordering of $1,\dots,n$. Note that there are $(n-1)!$ such permutations.

For example, one such permutation of $\{1,2,3,4\}$ is the cycle $(1 \to 3 \to 4 \to 2)$. The term corresponding to this cycle is

$$\chi(u_1-u_2-u_3-u_4)\chi(u_1-u_2+u_3-u_4)\chi(u_1-u_2+u_3+u_4)\chi(u_1+u_2+u_3+u_4), \tag{B.4}$$

where the signs are changed from -1 to $+1$ in the order specified by the cycle (starting with u_1).

B.3. Sums over Decompositions

The notation

$$\sum_{[H,H^c]} (|H|-1)!(|H^c|-1)! \int_{\mathbb{R}} |u| \widehat{\prod_{i\in H} f_i}(u) \widehat{\prod_{i\in H^c} f_i}(u)du \tag{B.5}$$

is a sum over the ways of decomposing the set $\{1,\dots,n\}$ into two disjoint subsets H, H^c, such that $H \cup H^c = \{1,\dots,n\}$ and $H, H^c \neq \varnothing$. Note that there are $2^{n-1}-1$ such decompositions.

For example, one decomposition of $\{1,2,3,4\}$ is given by $H = \{3\}$, $H^c = \{1,2,4\}$. The term corresponding to $[H,H^c]$ is

$$2\int_{\mathbb{R}} \hat{f}_3(u)\widehat{f_1 f_2 f_4}(u)du. \tag{B.6}$$

B.4. Sums over Subsets.

The notation

$$\sum_{I\subseteq\{1,\dots,n\}} (-1)^{|I|} \int_{\mathbb{R}^n_{\geq 0}} \tilde{\chi}(\sum_{I^c} u_i - \sum_{I} u_j) \prod_{i=1}^{n} \hat{f}_i(u_i)du_i \tag{B.7}$$

is a sum over the subsets $I \subseteq \{1,\dots,n\}$, and $I^c = \{1,\dots,n\} \setminus I$. (Here $\tilde{\chi}$ is the characteristic function of the interval $[1,\infty)$.)

For example, the term corresponding to the subset $\{2,3,5\} \subset \{1,2,3,4,5,6\}$ is

$$-\int_{\mathbb{R}^6_{\geq 0}} \tilde{\chi}(u_1 - u_2 - u_3 + u_4 - u_5 + u_6) \prod_{i=1}^{6} \hat{f}_i(u_i) du_i. \tag{B.8}$$

References

[1] D. Bump, J. W. Cogdell, E. de Shalit, D. Gaitsgory, E. Kowalski, and S. S. Kudla. *An introduction to the Langlands program*. Birkhäuser Boston Inc., Boston, MA, 2003. Lectures presented at the Hebrew University of Jerusalem, Jerusalem, March 12–16, 2001, Edited by Joseph Bernstein and Stephen Gelbart.

[2] Eduardo Dueñez and Steven J. Miller. The low lying zeros of a $\mathrm{GL}(4)$ and a $\mathrm{GL}(6)$ family of L-functions. *Compos. Math.*, 142(6):1403–1425, 2006.

[3] Eduardo Dueñez and Steven J. Miller. The effect of convolving families of L-functions on the underlying group symmetries. *Proc. Lond. Math. Soc. (3)*, 99(3):787–820, 2009.

[4] Peng Gao. *n-level Density of the Low-Lying Zeros of Quadratic Dirichlet L-functions*. PhD thesis, The University of Michigan, 2005.

[5] Dennis A. Hejhal. On the triple correlation of zeros of the zeta function. *Internat. Math. Res. Notices*, (7):293ff., approx. 10 pp. (electronic), 1994.

[6] C. P. Hughes and Steven J. Miller. Low-lying zeros of L-functions with orthogonal symmetry. *Duke Math. J.*, 136(1):115–172, 2007.

[7] C. P. Hughes and Z. Rudnick. Linear statistics of low-lying zeros of L-functions. *Q. J. Math.*, 54(3):309–333, 2003.

[8] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[9] Henryk Iwaniec, Wenzhi Luo, and Peter Sarnak. Low lying zeros of families of L-functions. *Inst. Hautes Études Sci. Publ. Math.*, (91):55–131 (2001), 2000.

[10] Nicholas M. Katz and Peter Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[11] Nicholas M. Katz and Peter Sarnak. Zeroes of zeta functions and symmetry. *Bull. Amer. Math. Soc. (N.S.)*, 36(1):1–26, 1999.

[12] Steven J. Miller. One- and two-level densities for rational families of elliptic curves: evidence for the underlying group symmetries. *Compos. Math.*, 140(4):952–992, 2004.

[13] Steven J. Miller. A symplectic test of the L-functions ratios conjecture. *Int. Math. Res. Not. IMRN*, (3):Art. ID rnm146, 36, 2008.

[14] Steven J. Miller. Lower order terms in the 1-level density for families of holomorphic cuspidal newforms. *Acta Arith.*, 137(1):51–98, 2009.

[15] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., 24, St. Louis, Mo., 1972)*, pages 181–193. Amer. Math. Soc., Providence, R.I., 1973.

[16] A. M. Odlyzko. On the distribution of spacings between zeros of the zeta function. *Math. Comp.*, 48(177):273–308, 1987.

[17] A. E. Özlük and C. Snyder. On the distribution of the nontrivial zeros of quadratic L-functions close to the real axis. *Acta Arith.*, 91(3):209–228, 1999.

[18] Emmanuel Royer. Petits zéros de fonctions L de formes modulaires. *Acta Arith.*, 99(2):147–172, 2001.

[19] M. Rubinstein. *Evidence for a Spectral Interpretation of the Zeros of L-Functions*. PhD thesis, Princeton University, 1998.

[20] Zeév Rudnick and Peter Sarnak. Zeros of principal L-functions and random matrix theory. *Duke Math. J.*, 81(2):269–322, 1996. A celebration of John F. Nash, Jr.

[21] Matthew P. Young. Low-lying zeros of families of elliptic curves. *J. Amer. Math. Soc.*, 19(1):205–250, 2006.