CRITICAL ZEROS OF HECKE L-FUNCTIONS

BY JORGE CANTILLO

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Henryk Iwaniec

and approved by

New Brunswick, New Jersey

October, 2014

ABSTRACT OF THE DISSERTATION

Critical Zeros of Hecke L-Functions

by Jorge Cantillo

Dissertation Director: Henryk Iwaniec

In this dissertation, we established that in average taken over the family of all Hecke L-functions of weight $k \asymp K$ associated with the full modular group, at least 35% of their zeros lie on the critical line as $K \to \infty$. We used Levinson's method employing a mollifier of length $K^{2\theta}$ with θ sufficiently close to $\frac{1}{2}$. To handle such a long mollifier, it was necessary to develop an Asymptotic Large Sieve that evaluated a bilinear form by taking advantage of sum cancellations resulting from the quasi-orthogonality property of Hecke eigenvalues for a sufficiently large number of weights k.

Acknowledgements

I would like to specially thank my thesis adviser Professor Henryk Iwaniec for the unvaluable help he provided throughout the entire process of creating this work. His insightful suggestions carried me through at times the problems seemed to difficult to overcome. Also, I would like to thank the Mathematics Department of Rutgers University for their support.

Dedication

To my wife Lorna, my son Matthew, and my daugther Abigail for all the time with me they have sacrificed so that this work could be completed.

Table of Contents

Chapter 1

Introduction

1.1 A Brief History

Given the level of difficulty associated with proving Riemann Hypothesis (RH), efforts have been directed at solving problems about zeros of the zeta function that are more approchable with current technology. One such problem consists of giving the best lower bound possible for the number of zeros $N_0(T)$ on the critical line up to height T in comparison with $N(T)$ which is the total number of nontrivial zeros of height less than T. The latter have been shown to be asymptotically

$$N(T) \sim \frac{2T}{2\pi} \log T.$$

The first such result was given by G. H. Hardy and J. E. Littlewood who proved

$$N_0(T) \gg T.$$

However, it was A. Selberg who first established a lower bound of the right order of magnitud, namely

$$N_0(T) > cT \log T.$$

On the other hand, Selberg's method produced a very small constant c. In 1976, employing a new method, N. Levinson [9] proved

$$N_0(T) > .34N(T).$$

In other words, more than 34 % of zeros lie on the critical line. As described in more precise terms later in this thesis, Levinson's method counts critical zeros $\rho = \frac{1}{2} + i\gamma$ using the detector

$$\arg G(\rho) = \frac{\pi}{2} \mod \pi$$

where $G(s)$ is a linear combination of $\zeta(s)$ and its first derivative,

$$G(s) = \zeta(s) + \lambda\zeta'(s).$$

After a few steps culminating in the use of Littlewood's lemma [7], the problem of finding a lower bound for $N_0(T)$ is transformed into that of giving an upper bound for

$$\frac{1}{2T}\int_{-T}^{T} \log|G(a+it)|dt$$

where the integration takes place on a vertical line $\Re s = a$ slightly shifted left away from the critical line. Unfortunately, if applied directly this way, the method fails to produce the right order of magnitude due to large values of $G(s)$. To overcome this difficulty, Levinson introduces the old idea of mollification. More precisely, $G(s)$ is multiplied times a Dirichlet polynomial $M(s)$ of length T^{θ}. The mollifier is built so that it mimics the inverse of $\zeta^{-1}(s)$ with the hope of taming large values of $G(s)$, but at the expense of also counting possible zeros of $M(s)$. Then, by concavity of log and Cauchy's inequality, we are left with the problem of evaluating the following mean

$$\frac{1}{2T}\int_{-T}^{T} |M(a+it)G(a+it)|^2 dt$$

for which there are analytic techniques available. Levinson obtained his result employing a mollifier of length T^{θ}, with $\theta = \frac{1}{2}$.

In 1986, after further refinements of Levinson's method, B. Conrey [3] produced a large inprovement; namely that more than two fifth of the zeros lie on the critical line,

$$N_0(T) > \frac{2}{5}N(T). \tag{1.1}$$

His improvement depended upon been able to handle a longer mollifier with $\theta = \frac{4}{7}$. Key to this was the application of results by H. Iwaniec and J. M. Deshoullier on averages of Kloosterman sums. Since Conrey's result there have only been slight improvements of the percentage of critical zeros. See the work of M. Young, H. Bui, and B. Conrey in [1] or independent work by S. Feng on the subject. It should be noticed that Levinson's method could produce a percentage of critical zeros above the 50% threshold if it were possible to handle a mollifier of length $\theta = 1$. Unfortunately, current technology is insufficient to accomplish that for $\zeta(s)$ alone.

1.2 Generalizing Levinson's Method

It is naturally possible to generalize the problem about the density of critical zeros as well as Levinson's method to other L-functions in the context of the Generalized Riemann Hypothesis (GRH). However, significant progress in terms of proving a larger percentage of critical zeros is made by tweaking the original problem into a statement about the average number of critical zeros in a family of L-functions. In [4], B. Conrey, H. Iwaniec, and K. Soundararajan consider the family of L-functions $L(s,\chi)$ obtained by twisting a single L-function $L(s)$ by every primitive Dirichlet character $\chi \mod q$ for every positive integer q. With $N(T,\chi)$ and $N_0(T,\chi)$ being the number of non-trivial and critical zeros of $L(s,\chi)$ with absolute imaginary part less than T, respectively, the authors of [4] define the following averages

$$\mathcal{N}(T,Q) = \sum_q \psi\left(\frac{q}{Q}\right) \frac{1}{\phi(q)} \sideset{}{^*}\sum_{\chi \mod q} N(T,\chi)$$

and

$$\mathcal{N}_0(T,Q) = \sum_q \psi\left(\frac{q}{Q}\right) \frac{1}{\phi(q)} \sideset{}{^*}\sum_{\chi \mod q} N_0(T,\chi)$$

where $\psi \in C_0^\infty(\mathbb{R}^+)$ is a smooth function of compact support and the asterisk indicates that the inner sum is restricted to primitive characters only. Then, they prove

$$\mathcal{N}_0(T,Q) > \kappa \mathcal{N}(T,Q)$$

for $(\log Q)^6 \leq T \leq (\log Q)^A$ with $A \geq 6$, as $Q \to \infty$. The authors of [4] give the following explicit formula for κ,

$$\kappa = \kappa(\theta, r, R) = 1 - \frac{1}{R}\log c(\theta, r, R) \tag{1.2}$$

where the constant $c(\theta, r, R)$ is the mean value in the average

$$\sum_q \psi\left(\frac{q}{Q}\right) \frac{1}{\phi(q)} \sideset{}{^*}\sum_{\chi \mod q} I_\chi \sim c(\theta, r, R) \sum_q \psi\left(\frac{q}{Q}\right) \frac{\phi^*(q)}{\phi(q)}$$

as $Q \to \infty$, with

$$I_\chi = \frac{1}{2T}\int_{-T}^{T} \left| M\left(\frac{1}{2}+it,\chi\right) G(a+it,\chi)\right|^2 dt$$

for T in the range $(\log Q)^6 \leq T \leq (\log Q)^A$. Here, $G(s,\chi)$ is given as the linear combination

$$G(s,\chi) = L(s,\chi) + \lambda L'(s,\chi)$$

with $\lambda = (r \log \frac{q}{2\pi})^{-1}$ and $a = \frac{1}{2} - \frac{R}{\log \frac{q}{2\pi}}$. In addition, the mollifier $M(s,\chi)$ is chosen as a Dirichlet polynomial of length $\left(\frac{q}{2\pi}\right)^\theta$ that mimics the inverse of $L(s,\chi)$. Finally, the mean value $c(\theta,r,R)$ is given by the following formulas

$$r^2 c(\theta,r,R) = C(\theta,r,R) + C(\theta,1-r,-R)\exp\{2R\} \tag{1.3}$$

with

$$C(\theta,r,R) = -\left(\frac{r^2}{2} + \frac{1}{4R^2}\right)\left(\frac{1}{R\theta} + \frac{R\theta}{3}\right) - \frac{r}{2R}\left(\frac{1}{R\theta} - \frac{R\theta}{3}\right) + \frac{r^2}{2}. \tag{1.4}$$

When $L(s) = \zeta(s)$, the family of L-functions consist of the classic Dirichlet L-functions. Then, choosing $\theta = 1$, $r = 1.06$, and $R = 0.75$, formulas (1.2), (1.3), and (1.4) produce the following value of κ,

$$\kappa(1, 1.06, 0.75) \approx 0.586.$$

The authors of [4] succeed in overcoming the 50% threshold by employing a mollifier of length $\left(\frac{q}{2\pi}\right)^\theta$, with $\theta = 1$. They are able to handle a mollifier of such length by exploiting cancellation produced by the orthogonality property satisfied by the Dirichlet characters modulo q. They also take advantage of having a sufficiently large number of q's relative to the effective length of $G(s,\chi)$ and the mollifier. In the process, they have to develop the Asymptotic Large Sieve in [5] which implements the above ideas.

If $L(s)$ is chosen to be a degree two L-function, then the length of the mollifier in this case would correspond to the length of a mollifier in the case of a degree one L-function with $\theta = \frac{1}{2}$. Then, choosing $r = 0.96$ and $R = 1.24$, the authors obtained the value of κ,

$$\kappa\,(0.5, 0.96, 1.24) \approx 0.356. \tag{1.5}$$

In the present work, we follow the recipe provided by [4] and obtain a density theorem for a family of Hecke L-functions. These are degree two L-functions so we give a result with the same value of κ as in (1.5). We are able to this because the fomulas that we arrive at in the current work match (1.2), (1.3), and (1.4). We will show this in chapter 3 while a precise statement of our main result is given at the end of the next chapter.

Chapter 2

Hecke L-functions

2.1 The Hecke Basis of Cusp Forms and their Hecke Eigenvalues

Let Γ be the modular group $\mathrm{Sl}_2(\mathbb{Z})$, then for every even $k \geq 12$ we denote the space of cusp forms of weight k by $S_k(\Gamma)$. This is a finite dimension linear space of holomorphic functions $f(z)$ of the upper half-plane $\mathbb{H}$ that satisfy the modularity relation

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \tag{2.1}$$

for all $z \in \mathbb{H}$ and vanish at the cusp of the fundamental region $\Gamma\backslash\mathbb{H}$ located at infinity. From these defining properties, it follows that every $f(z)$ in $S_k(\Gamma)$ has a Fourier series expansion around the point at infinity given by

$$f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{\frac{k-1}{2}} e(nz) \tag{2.2}$$

where $e(z) = \exp\{2\pi i z\}$. If we introduce the Petersson inner product

$$\langle f, g\rangle = \int_{\Gamma\backslash\mathbb{H}} f(z)\bar{g}(z) y^{k-2} dx dy \quad \text{for all } f \text{ and } g \in S_k(\Gamma)$$

then, the theory by E. Hecke constructs a family of commuting self-adjoint operators T_n, $n \in \mathbb{N}$ over the space $S_k(\Gamma)$ that satisfy

$$T_{mn} = T_m T_n \quad \text{if} \quad (m,n) = 1 \tag{2.3}$$

(see [6] for details.) This construction is done such that the members f of the natural orthonormal basis $H_k(1)$ of common eigenfunctions to all Hecke operators T_n have as their respective eigenvalue their n^{th} Fourier coefficient $\lambda_f(n)$. Namely, if $f \in H_k(1)$

$$T_n f = \lambda_f(n) f \quad \text{for all } n \in \mathbb{N}. \tag{2.4}$$

Thus, (2.3) and (2.4) imply that the Hecke eigenvalues $\lambda_f(n)$ for every $f \in H_k(1)$ are multiplicative arithmetic functions, i.e. they satisfy

$$\lambda_f(mn) = \lambda_f(m)\lambda_f(n) \quad \text{if} \quad (m,n) = 1.$$

In general, we have the formula

$$\lambda_f(m)\,\lambda_f(n) = \sum_{d|(m,n)} \lambda_f\left(\frac{mn}{d^2}\right). \tag{2.5}$$

The lack of complete multiplicativity of the Hecke eigenvalues is a fact that will make our arguments a bit more complicated than the treatment in [4]. Finally, since the Hecke operators T_n are self-adjoint, we have that the Hecke eigenvalues $\lambda_f(n)$ are real numbers for all $n \in \mathbb{N}$ and $f \in H_k(1)$.

2.1.1 The Petersson Formula

Here, we will establish an analogy between the set of Hecke eigenvalues and the set of Dirichlet characters. This analogy is made clear by the Petersson formula which will resemble the orthogonality property satisfied by the Dirichlet characters. Let $\Delta_k(m,n)$ be defined by the expression

$$\sum_{f\in H_k(1)}^{h} \lambda_f(m)\lambda_f(n) \tag{2.6}$$

where $\sum^h$ is the weighted sum

$$\sum_{f\in H_k(1)}^{h} \alpha_f = \frac{\Gamma(k-1)}{(4\pi)^{k-1}} \sum_{f\in H_k(1)} \frac{\alpha_f}{\|f\|^2} \tag{2.7}$$

with $\|f\|^2 = \langle f, f\rangle$. Then, the Petersson formula establishes the quasi-orthogonality property

$$\Delta_k(m,n) = \delta(m,n) + i^k J_{k-1}(m,n) \tag{2.8}$$

where $\delta(m,n)$ is the Kronecker symbol

$$\delta(m,n) = \begin{cases} 1 & \text{if} \quad m = n \\ 0 & \text{otherwise} \end{cases}$$

and $J_{k-1}(m,n)$ is the series

$$J_{k-1}(m,n) = \sum_{c>0} \frac{S(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right). \tag{2.9}$$

In this series,

$$S(m,n;c) = \sum_{ab=1 \mod c} e\left(\frac{am+bn}{c}\right)$$

are the Kloosterman sums and $J_{k-1}(x)$ are the classical Bessel function of integral order $k-1$. Just as the the orthogonality property satisfied by Dirichlet characters played a crucial role in [4], the Petersson formula will play the same role in our arguments.

2.1.2 The Ramanujan-Petersson Conjecture

One last property about the Hecke eigenvalues which will be useful to us is that they satisfy the Ramanujan-Petersson conjecture, which was proved by P. Deligne. This property states that

$$\lambda_f(n) \leq \tau(n). \tag{2.10}$$

Then, estimating trivially the Fourier series (2.2) of $f(z)$ using the Ramanujan conjecture, we have that $f(z)$ decays exponentially in vertical lines

$$|f(x+iy)| \ll \exp\{-2\pi y\} \quad \text{for } y \geq Y \tag{2.11}$$

with the implicit constant depending on k and Y.

2.2 The Hecke L-functions

Based on all the properties satisfied by the Hecke eigenvalues, it is natural to consider the following Dirichlet series with Euler product

$$L(s,f) = \sum_{n=1}^{\infty} \frac{\lambda_f(n)}{n^s} = \prod_p (1 - \lambda_f(p)p^{-s} + p^{-2s})^{-1} \tag{2.12}$$

for each $f \in H_k(1)$. They are an analogue of the classical Dirichlet L-functions and which were first introduced by E. Hecke. In addition, other relevant Dirichlet series derived from $L(s,f)$ will be referenced in this work. These include its inverse

$$\frac{1}{L(f,s)} = \prod_p (1 - \lambda_f(p)p^{-s} + p^{-2s}) = \sum \frac{\mu_f(m)}{m^s} \tag{2.13}$$

where

$$\mu_f(m) = \sum_{m=nr^2} \lambda_f(n)\mu(nr)\mu(r)$$

and its logarithmic derivative

$$-\frac{L'}{L}(s,f) = \sum \frac{\Lambda_f(m)}{m^s} \tag{2.14}$$

with

$$\Lambda_f(m) = \sum_{m=nr^2} \lambda_f(n)a(n,r)$$

and

$$a(n,r) = \begin{cases} \Lambda(n) & \text{if } r = 1. \\ \Lambda(r)\sum_{c|n}\mu(cr) & \text{if } r \neq 1. \end{cases}$$

The coefficients for the logarithmic derivative are obtained by multiplying the inverse of $L(s,f)$ times its derivative and using formula (2.5). From the Ramanujan-Petersson conjecture, it follows inmediately that the three Dirichlet series (2.12), (2.13), and (2.14) are absolutely convergent for $\Re s > 1$. Also, they have bounds independent of f and the weight k in this region.

2.2.1 The Complete L-function

Let $f \in H_k(1)$. Then, the function

$$\Lambda(s,f) = \gamma(s)L(s,f)$$

is known as the complete L-function with

$$\gamma(s) = (2\pi)^{-s}\Gamma\left(s + \frac{k-1}{2}\right)$$

denoting the gamma factor. Using the integral representation of $\Gamma(s)$, we have

$$\begin{aligned} \Lambda(s,f) &= (2\pi)^{-s}\int_0^\infty e^{-y}y^{s+\frac{k-1}{2}}\frac{dy}{y}\sum_{n=1}\lambda_f(n)n^{-s} \\ &= \sum_{n=1}\lambda_f(n)\int_0^\infty y^{\frac{k-1}{2}}e^{-y}\left(\frac{y}{2\pi n}\right)^s\frac{dy}{y} \\ &= \sum_{n=1}\lambda_f(n)\int_0^\infty (2\pi yn)^{\frac{k-1}{2}}e^{-2\pi ny}y^s\frac{dy}{y}. \end{aligned}$$

Thus, we obtain the formula

$$\Lambda(s,f) = \int_0^\infty (2\pi y)^{\frac{k-1}{2}} f(iy) y^s \frac{dy}{y}.$$

From the modularity relations (2.1), we have

$$f(iy) = i^k y^{-k} f\left(\frac{i}{y}\right).$$

Then,

$$\begin{aligned}\Lambda(s,f) &= \int_0^1 (2\pi y)^{\frac{k-1}{2}} f(iy) y^s \frac{dy}{y} + \int_1^\infty (2\pi y)^{\frac{k-1}{2}} f(iy) y^s \frac{dy}{y} \\ &= \int_0^1 \left(\frac{2\pi}{y}\right)^{\frac{k-1}{2}} f\left(\frac{i}{y}\right) i^k y^{s-1} \frac{dy}{y} + \int_1^\infty (2\pi y)^{\frac{k-1}{2}} f(iy) y^s \frac{dy}{y}.\end{aligned}$$

If we apply the change of variables $y \to y^{-1}$ to the first integral above, then the following integral representation of $\Lambda(s,f)$ is obtained

$$\Lambda(s,f) = \int_1^\infty (2\pi y)^{\frac{k-1}{2}} f(iy) \left(y^s + i^k y^{1-s}\right) \frac{dy}{y}. \tag{2.15}$$

Hence, two immediate important consequences follow from this formula. First, the integral representation (2.15) makes it possible to extend $\Lambda(s,f)$ to an entire function of order 1 because of the exponential decay (2.11) satisfied by $f(iy)$. Second, the following functional equation is established

$$\Lambda(s,f) = i^k \Lambda(1-s,f). \tag{2.16}$$

2.2.2 The Zeros of $L(s,f)$

If $f \in H_k(1)$, then the Euler product formula (2.12) implies that $L(s,f)$ does not have zeros ρ with $\Re\rho > 1$. Thus, it follows that $\Lambda(s,f)$ have no zeros ρ with $\Re\rho > 1$ because $\frac{1}{\gamma(s)}$ is an entire function with simple zeros

$$s = -n - \frac{k-1}{2} \quad , n \in \mathbb{N}. \tag{2.17}$$

Hence, by the functional equation (2.16), $\Lambda(s,f)$ also has no zeros ρ with $\Re\rho < 0$. This implies that the first order poles (2.17) of $\gamma(s)$ are simple zeros of $L(s,f)$. They are commonly known as the trivial zeros of $L(s,f)$.

From the theory of entire functions of a finite order developed by Hadamard, it follows that each $\Lambda(s,f)$ has an infinite number of zeros ρ in the strip $0 \le \Re s \le 1$ such that the series

$$\sum_{\rho} \frac{1}{|\rho|^{\alpha}} \tag{2.18}$$

converges for every $\alpha > 1$. These set of zeros of $L(s,f)$ in the critical strip are at the center of the theory of classical L-functions. From Hadamard's theory, we also have the following product formula

$$\Lambda(s,f) = e^{-B_f s + A_f} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{\frac{s}{\rho}} \tag{2.19}$$

with A_f and B_f real numbers. Taking the logarithmic derivative of this formula, we obtain

$$\frac{\Lambda'}{\Lambda}(s,f) = -B_f + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \tag{2.20}$$

Also, taking the logarithmic derivative of the functional equation (2.16), we have the following functional equation of the logarithmic derivative of $\Lambda(s,f)$

$$\frac{\Lambda'}{\Lambda}(s,f) = -\frac{\Lambda'}{\Lambda}(1-s,f). \tag{2.21}$$

Hence, formulas (2.20) and (2.21) imply

$$2B_f = \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{1-s-\rho} + \frac{2}{\rho}\right).$$

In particular, for $s = 0$ we have

$$2B_f = \sum_{\rho} \left(\frac{1}{1-\rho} + \frac{1}{\rho}\right).$$

Taking the real part in the previous formula combined with the fact that

$$\sum_{\rho} \Re\frac{1}{\rho} \le \sum_{\rho} \frac{1}{|\rho|^2} \ll 1$$

we obtain

$$2B_f = \sum_{\rho} \Re\frac{1}{1-\rho} + \sum_{\rho} \Re\frac{1}{\rho}.$$

By the functional equation (2.16), if ρ is a zero of $\Lambda(s,f)$, then $1-\rho$ is also a zero of $\Lambda(s,f)$. Thus, we have

$$\begin{aligned} B_f &= \sum_{\rho} \Re \frac{1}{\rho} \\ &= \sum_{\rho} \frac{\Re \rho}{|\rho|^2}. \end{aligned} \tag{2.22}$$

The following useful standard lemma evaluates the logarithmic derivative of $L(s,f)$ in the critical strip $0 \leq \Re s \leq 1$.

Lemma 2.2.1. *For every $f \in H_k(1)$ with $k \asymp K$, let $\rho = \beta + i\gamma$ be a generic element of the set of non-trivial zeros of $L(s,f)$. If $T = o(K)$, then, as $K \to \infty$*

i)

$$\sum_{\rho} \frac{1}{1+(t-\gamma)^2} \ll \log K$$

uniformly for all $f \in H_k(1)$ with $k \asymp K$ and $|t| \leq T$.

ii) The number of zeros ρ of $L(s,f)$ with $t-1 \leq \gamma \leq t+1$ is $O(\log K)$ uniformly for all $f \in H_k(1)$ with $k \asymp K$ and $|t| \leq T$.

iii)

$$-\frac{L'}{L}(s,f) = -\sum_{\substack{\rho \\ |t-\gamma| \leq 1}} \frac{1}{s-\rho} + O(\log K)$$

uniformly for all $f \in H_k(1)$ with $k \asymp K$ and all $s = \sigma + it$ with $0 \leq \sigma \leq 1$ and $|t| \leq T$.

This lemma is stated and proven in greater generality in [8], chapter 5.

Proof. If we evaluate formula (2.20) at $s = 2+it$ and take real parts, we obtain

$$\Re \frac{L'}{L}(2+it,f) + \Re \frac{\gamma'}{\gamma}(2+it) = \sum_{\rho} \frac{2-\beta}{(2-\beta)^2 + (t-\gamma)^2}.$$

Then, from the Ramanujan-Petersson conjecture, we have

$$-\frac{L'}{L}(2+it,f) \ll 1$$

with an absolute constant. Also, the asymptotic formula

$$\frac{\Gamma'}{\Gamma}(s) = \log s + O\left(\frac{1}{|s|}\right) \tag{2.23}$$

as $|s| \to \infty$, uniformly for $-\pi+\delta < \arg(s) < \pi-\delta$ with the implicit constant depending on δ, gives

$$\frac{\gamma'}{\gamma}(s) \sim \log K \tag{2.24}$$

as $K \to \infty$ uniformly for all $k \asymp K$ and all $s = \sigma + it$ with $0 \le \sigma \le 1$ and $|t| \le T$. Hence, part i) of the lemma follows. Part ii) follows by dropping from the expression given in part i) those terms which do not satisfy $t-1 \le \gamma \le t+1$. To obtain part iii) we substract formula (2.20) evaluated at $s = \sigma + it$ with $0 \le \sigma \le 1$ from the same formula evaluated at $2+it$. Thus, we have

$$-\frac{L'}{L}(s,f) = -\frac{L'}{L}(2+it,f) - \frac{\gamma'}{\gamma}(2+it) + \frac{\gamma'}{\gamma}(s) - \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right).$$

By the same arguments used to prove part i), we have

$$-\frac{L'}{L}(s,f) = -\sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log K).$$

If $|t-\gamma| > 1$, then

$$\begin{aligned}\left|\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right| &= \frac{2-\beta}{|s-\rho||2+it-\rho|}\\ &= \frac{2-\beta}{|2+it-\rho|^2}\sqrt{\frac{(2-\beta)^2+(t-\gamma)^2}{(\sigma-\beta)^2+(t-\gamma)^2}}\\ &\ll \frac{1}{1+(t-\gamma)^2}.\end{aligned}$$

Thus, we obtain

$$-\frac{L'}{L}(s,f) = -\sum_{\substack{\rho\\ |t-\gamma|\le 1}}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log K).$$

Since

$$\left|\frac{1}{2+it-\rho}\right| = \frac{1}{(2-\beta)^2+(t-\gamma)^2} \le 1$$

part iii) of the lemma follows by part ii). □

2.2.3 Counting Zeros in the Critical Strip

The distribution of the zeros of $L(s,f)$ in the critical strip is a key problem in the theory of classical L-functions. A formula that asymptotically evaluates the number of zeros $\rho = \beta + i\gamma$ of $L(s, f)$ with $0 \le \beta \le 1$ such that $|\gamma| \le T$ is a standard feature in the theory. In the case of the Riemann zeta function, this is traditionally known as the Riemann-von Mangoldt formula. A version of this formula for Hecke L-functions is given in the following proposition.

Proposition 2.2.2. *For every $f \in H_k(1)$ with $k \asymp K$, let $N(T, f)$ be the number of zeros $\rho = \beta + it$ of $L(s, f)$ counted with multiplicity such that $0 \le \beta \le 1$ and $|\gamma| \le T$. If*

$$T = o(K) \text{ and } T^{-1} = o(1)$$

as $K \to \infty$, then

$$N(T, f) \sim \frac{2T}{2\pi} \log\left(K^2\right) \tag{2.25}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$.

Proof. Let

$$R = \Gamma_1 - \Gamma_2 - \Gamma_3 + \Gamma_4$$

be a rectangle in the complex plane traveled counter-clockwise where Γ_1, Γ_2, Γ_3, and Γ_4 are the segments $[2 - iT, 2 + iT]$, $[-1 + iT, 2 + iT]$, $[-1 - iT, -1 + iT]$, and $[-1 - iT, 2 - iT]$, respectively. Also, individual vertical and horizontal segments are assumed to be traveled in an upward and horizontal direction, respectively. Then, we have

$$N(T, f) = \frac{1}{2\pi} \Delta_R \arg \Lambda(s, f) = \Im \frac{1}{2\pi} \int_R \frac{\Lambda'}{\Lambda}(s, f) ds$$

where Δ_R is the variation of the argument over R. Thus, it follows that

$$\begin{aligned} N(T, f) = &\frac{1}{2\pi} \Delta_{\Gamma_1} \arg \Lambda(s, f) - \frac{1}{2\pi} \Delta_{\Gamma_2} \arg \Lambda(s, f) \\ &- \frac{1}{2\pi} \Delta_{\Gamma_3} \arg \Lambda(s, f) + \frac{1}{2\pi} \Delta_{\Gamma_4} \arg \Lambda(s, f) \end{aligned}$$

with

$$\frac{1}{2\pi} \Delta_{\Gamma_i} \arg \Lambda(s, f) = \Im \frac{1}{2\pi} \int_{\Gamma_i} \frac{\Lambda'}{\Lambda}(s, f) ds \quad \text{for } i = 1, 2, 3, 4.$$

By the functional equation,

$$\frac{\Lambda'}{\Lambda}(s, f) = -\frac{\Lambda'}{\Lambda}(1 - s, f)$$

and the appropriate change of variables, we have

$$\begin{aligned}\frac{1}{2\pi}\Delta_{\Gamma_1}\arg\Lambda(s,f) &= \Im\frac{1}{2\pi}\int_{\Gamma_1}\frac{\Lambda'}{\Lambda}(s,f)ds\\ &= \Im\frac{1}{2\pi}\int_{\Gamma_1}\frac{\Lambda'}{\Lambda}(1-s,f)ds\\ &= -\Im\frac{1}{2\pi}\int_{\Gamma_3}\frac{\Lambda'}{\Lambda}(s,f)ds\\ &= -\frac{1}{2\pi}\Delta_{\Gamma_3}\arg\Lambda(s,f).\end{aligned}$$

Similarly, we have

$$-\frac{1}{2\pi}\Delta_{\Gamma_2}\arg\Lambda(s,f) = \frac{1}{2\pi}\Delta_{\Gamma_4}\arg\Lambda(s,f).$$

Hence, we obtain the equation

$$N(T,f) = \frac{1}{\pi}\Delta_{\Gamma_1}\arg\Lambda(s,f) - \frac{1}{\pi}\Delta_{\Gamma_2}\arg\Lambda(s,f).$$

Then, by the formula

$$\frac{\Lambda'}{\Lambda}(s,f) = \frac{L'}{L}(s,f) + \frac{\gamma'}{\gamma}(s,f),$$

we get

$$\begin{aligned}N(T,f) = &\frac{1}{\pi}\Delta_{\Gamma_1}\arg\gamma(s,f) + \frac{1}{\pi}\Delta_{\Gamma_1}\arg L(s,f)\\ &-\frac{1}{\pi}\Delta_{\Gamma_2}\arg\gamma(s,f) - \frac{1}{\pi}\Delta_{\Gamma_2}\arg L(s,f).\end{aligned}$$

By the already established asymptotic formula

$$\frac{\gamma'}{\gamma}(\sigma+it) \sim \log K$$

as $K\to\infty$, uniformly for all $k \asymp K$ and all $s=\sigma+it$ with $0\le\sigma\le 1$ and $|t|\le T$, we have that

$$\begin{aligned}\frac{1}{\pi}\Delta_{\Gamma_1}\arg\gamma(s) &\sim \frac{2T}{\pi}\log K\\ &= \frac{2T}{2\pi}\log\left(K^2\right)\end{aligned} \tag{2.26}$$

and

$$\frac{1}{\pi}\Delta_{\Gamma_2}\arg\gamma(s) \ll \log K \tag{2.27}$$

uniformly for all $k\asymp K$. In the case of the variation of the argument of $L(s,f)$, by the bound

$$\frac{L'}{L}(2+it,f) \ll 1,$$

we have

$$\frac{1}{\pi}\Delta_{\Gamma_1} \arg L(s,f) \ll T \tag{2.28}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ and $k \asymp K$. On the other hand, by part iii) of Lemma 2.2.1, we obtain

$$-\frac{1}{\pi}\Delta_{\Gamma_2} \arg L(s,f) = -\sum_{\substack{\rho \\ |t-\gamma|\leq 1}} \frac{1}{\pi}\Delta_{\Gamma_1} \arg(s-\rho) + O(\log K)$$

Since

$$\frac{1}{\pi}\Delta_{\Gamma_1} \arg(s-\rho) \leq 1$$

for every non-trivial zero ρ of $L(s,f)$, we obtain from part ii) of Lemma 2.2.1 that

$$\frac{1}{\pi}\Delta_{\Gamma_2} \arg L(s,f) \ll \log K \tag{2.29}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ and $k \asymp K$. Hence, the lemma follows after gathering the results from (2.26), (2.27), (2.28), (2.29), and the condition

$$T^{-1} = o(1)$$

as $K \to \infty$. □

2.3 The Main Result

The Generalized Riemann Hypothesis (GRH) asserts that all non-trivial zeros ρ of classical L-functions lie in the critical line, i.e. they satisfy $\Re\rho = \frac{1}{2}$. Given that this is a very difficult problem, we could try to answer a relatively simpler or rather more approchable question. Namely, we would like to determine what fraction of the non-trivial zeros ρ counted with multiplicity, i.e. those satifying $0 \leq \Re\rho \leq 1$ and with the additional restriction that $|\Im\rho| \leq T$ lie in the critical line for any given large enough T. For this purpose, we define for every $f \in H_k(1)$, the quantity $N_0(T,f)$ which counts with multiplicity the number of critical zeros $\rho = \frac{1}{2} + i\gamma$ of $L(s,f)$ with $|\gamma| \leq T$. Thus, the precise problem consists of finding a constant κ with $0 < \kappa < 1$ such that

$$N_0(T,f) > \kappa N(T,f)$$

as $T \to \infty$. In a recent paper by D. Bernard , [2], the author proves that at least 2.97% of zeros of a single L-function associated with a holomorphic primitive cusp form lie on the critical line, i.e. $\kappa = .0297$. In contrast, by tweaking this problem like in [4] and using Levinson's method, we will establish a lower bound for the average number of critical zeros in terms of a fraction of the average number of non-trivial zeros with the averaging taking place over a large family of Hecke L-functions. The prize will be a significantly better constant κ than the one that can be produced for a single L-function. The specific family $\mathcal{F}$ that concerns us is the totality of Hecke L-functions. Hence, we have formally

$$\mathcal{F} = \{f \in H_k(1) \quad | \quad \text{with } k \geq 12 \text{ even}\} . \tag{2.30}$$

As we will see in later chapters, the idea behind the solution to the tweaked problem is to exploit the quasi-orthogonality property of the Hecke eigenvalues featured in the Petersson formula. In practice, when we average over the family $\mathcal{F}$, this will produce sufficient cancellation in some sums arising in the implementation of Levinson's method with their lengths directly linked to the resulting constant κ. Hence, by averaging the counting of zeros over the family $\mathcal{F}$ of L-functions we are able to effectively handle a longer "mollifier" which ultimately results in a better constant κ. We mentioned the mollifier already in the introduction although its formal definition will be given in the next chapter. Not surprisingly, since the Petersson formula plays such a prominent role, the appropriate and natural weights used when counting the zeros is derived from (2.7) which are the same weights featured in the Petersson formula when summing over the members of the Hecke basis. Hence, we introduce the following definitions

$$\mathcal{N}(T, \mathcal{F}) = \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} N(T, f) \tag{2.31}$$

and

$$\mathcal{N}_0(T, \mathcal{F}) = \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} N_0(T, f) \tag{2.32}$$

where $\psi(x) \in C_0^\infty(\mathbb{R}^+)$ is a smooth positive function of compact support in the positive reals. From the result (2.25) in Proposition 2.2.2, it follows inmediately that

$$\mathcal{N}(T, \mathcal{F}) \sim \frac{2T}{2\pi} \left(\log\left(K^2\right)\right) \mathcal{N}_{\mathcal{F}}(\psi) \tag{2.33}$$

where $\mathcal{N}_{\mathcal{F}}(\psi)$ is defined by

$$\mathcal{N}_{\mathcal{F}}(\psi) = \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right)\Delta_k(1,1) \tag{2.34}$$

with $\Delta_k(m,n)$ given by (2.6). Finally, this thesis will be devoted to the proof of the following theorem

Theorem 2.3.1. *If*

$$(\log\log K)^A \leq T = o(K)$$

as $K \to \infty$ *with* $A > 0$ *sufficiently large, then*

$$\mathcal{N}_0(T, \mathcal{F}) > .35\mathcal{N}(T, \mathcal{F}) \tag{2.35}$$

as $K \to \infty$.

We believe that with some small extra effort this result can be extended to one about simple zeros. The problem of establishing positive density of simple critical zeros for a single degree two L-function is open which makes our result more interesting.

Chapter 3

Levinson's Method

3.1 A Linear Combination of $L(s, f)$ and $L'(s, f)$

Levinson's method, as we explained in the introduction, produces a lower bound for the number of critical zeros $\rho = \frac{1}{2} + i\gamma$ with $|\gamma| \le T$. In the case of Hecke L-functions, we denote this quantity by $N_0(T, f)$ for $f \in H_k(1)$. Traditional theorems in the theory of functions of a complex variable are effective at determining the number of zeros of an analytic function in a region but cannot determine how these zeros are spatially distributed with respect to each other. For example, they do not determine whether the zeros are located in a straight line. Surprisingly, Levinson's idea to catch zeros in the critical line is extremely simple at its outset. The current implementation of the method follows the outline given in the appendix of [4].

Given $f \in H_k(1)$ with $k \asymp K$, the method begins by considering the following linear combination of $L(s, f)$ and its derivative

$$G(s, f) = L(s, f) + \lambda L'(s, f) \tag{3.1}$$

where the coefficient λ is chosen such that

$$\lambda = \left(r \log\left(K^2\right)\right)^{-1} \tag{3.2}$$

with a fixed parameter $r > 0$. Behind this choice of λ is the idea that if $L(s, f)$ were assumed to be a Dirichlet polynomial of length K, then taking its derivative would produce a loss of $\log K$ in order of magnitude. Heuristically, λ attempts to correct this.

3.1.1 A Functional Equation for $G(s,f)$

By the definition (3.1), we have

$$\gamma(s)G(s,f) = \Lambda(s,f)\left(1+\lambda\frac{L'(s,f)}{L(s,f)}\right).$$

Similarly,

$$\begin{aligned} i^k\gamma(1-s)G(1-s,f) &= i^k\Lambda(1-s,f)\left(1+\lambda\frac{L'}{L}(1-s,f)\right) \\ &= \Lambda(s,f)\left(1+\lambda\frac{L'}{L}(1-s,f)\right). \end{aligned}$$

Adding the two last equations, we obtain

$$\Lambda(s,f)\left(2+\lambda\frac{L'}{L}(s,f)+\lambda\frac{L'}{L}(1-s,f)\right) = \gamma(s)G(s,f)+i^k\gamma(1-s)G(1-s,f).$$

Now, we define the function

$$Y(s) = 2-\lambda\frac{\gamma'}{\gamma}(s)-\lambda\frac{\gamma'}{\gamma}(1-s). \tag{3.3}$$

Then, by the functional equation (2.21), we have

$$\begin{aligned} 0 &= \frac{\Lambda'}{\Lambda}(s,f)+\frac{\Lambda'}{\Lambda}(1-s,f) \\ &= \frac{L'}{L}(s,f)+\frac{L'}{L}(1-s,f)+\frac{\gamma'}{\gamma}(s)+\frac{\gamma'}{\gamma}(1-s) \end{aligned}$$

from which we obtain that $Y(s)$ satisfies

$$Y(s) = 2+\lambda\frac{L'}{L}(s,f)+\lambda\frac{L'}{L}(1-s,f).$$

Hence, we have the following functional equation of $G(s,f)$

$$Y(s)\Lambda(s,f) = \gamma(s)G(s,f)+i^k\gamma(1-s)G(1-s,f). \tag{3.4}$$

If we multiply equation (3.4) by

$$\eta = e^{-\frac{k\pi}{4}i},$$

then we obtain the even more symmetric expression

$$\eta Y(s)\Lambda(s,f) = \eta\gamma(s)G(s,f)+\bar{\eta}\gamma(1-s)G(1-s,f).$$

Thus, we make the important observation that on the critical line $\Re s = \frac{1}{2}$ the previous functional equation can be written as

$$\eta Y(s)\Lambda(s,f) = 2(\Re\eta)\gamma(s)G(s,f). \tag{3.5}$$

3.1.2 Detecting Critical Zeros of $L(s, f)$

Equation (3.5) is the reason for introducing the function $G(s, f)$. The following lemma expresses the ability of $G(s, f)$ of detecting critical zeros of $L(s, f)$.

Lemma 3.1.1. *Let $f \in H_k(1)$ with $k \asymp K$ even. If $\rho = \frac{1}{2} + i\delta$ with*

$$|\delta| = o(K)$$

as $K \to \infty$, then $L(\rho, f) = 0$ iff either $G(\rho, f) = 0$ or

$$\arg \gamma(\rho) G(\rho, f) \equiv \begin{cases} \frac{\pi}{2} & \mod \pi \quad \textit{if } k \equiv 0 \mod 4 \\ 0 & \mod \pi \quad \textit{if } k \equiv 2 \mod 4 \end{cases} \tag{3.6}$$

Proof. If $L(\rho, f) = 0$, then equation (3.5) implies that either $G(\rho, f) = 0$ or

$$\arg \eta\gamma(\rho) G(\rho, f) \equiv \frac{\pi}{2} \mod \pi \tag{3.7}$$

because $\gamma(s)$ does not vanish anywhere in the complex plane and its poles which are given by expression (2.17) are far away from the critical line when K is large. If $G(\rho, f) \neq 0$, then the result of condition (3.6) follows because

$$\arg \eta = \begin{cases} \frac{\pi}{2} & \mod \pi \quad \text{if } k \equiv 2 \mod 4. \\ 0 & \mod \pi \quad \text{if } k \equiv 0 \mod 4. \end{cases}$$

On the other hand, when either condition (3.6) or $G(\rho, f) = 0$ is satisfied then

$$Y(\rho)\Lambda(\rho, f) = 0.$$

From the previous chapter, we have that if $\rho = \frac{1}{2} + i\gamma$ with $|\gamma| = o(K)$, then

$$\frac{\gamma'}{\gamma}(\rho) \sim \log K$$

as $K \to \infty$, uniformly for all $k \asymp K$. Thus, we have

$$Y(\rho) = 2 - \lambda\frac{\gamma'}{\gamma}(\rho) - \lambda\frac{\gamma'}{\gamma}(1-\rho) \sim 2\left(1 - \frac{1}{r}\right)$$

as $K \to \infty$, uniformly for all $k \asymp K$. Hence, if $r \neq 1$, then $Y(\rho)\Lambda(\rho, f) = 0$ implies $L(\rho, f) = 0$ when K is sufficiently large. We can effectively still have $r = 1$ by choosing r as close to 1 as necessary. □

Armed with the Lemma 3.1.1, we can detect critical zeros of $L(s, f)$ by essentially counting the number of π-size variations of the argument of $G(s, f)$ along the critical line. When the argument of $G(s, f)$ is not well defined because of the presence of a zero of $G(s, f)$, we will see that asymptotically these zeros will also account for π-size variations of the argument. The end result will be establishing a lower bound for the number of critical zeros $N_0(T, f)$. Before we give such lower bound, several standard lemmas need to be proven first.

3.1.3 The Variation of the Argument of $G(s, f)$

At this point, Levinson's methods requires establishing an asymptotic expression that evaluates the variation of the argument of $G(s, f)$ along a segment of the critical line in terms of the number of zeros $\rho = \beta + i\gamma$ in the interior of an adjecent rectangular region R to the right of the critical line. However, it will be necessary to create small dents in the side of the region which coinsides with the critical line to avoid posible critical zeros of $G(s, f)$. We do this in the following lemma.

Lemma 3.1.2. *For every $f \in H_k(1)$ with $k \asymp K$, let the quantity $N_G(T, f)$ count with multiplicity the zeros $\rho = \beta + i\gamma$ of $G(s, f)$ such that $\frac{1}{2} < \beta < 4$ and $|\gamma| \leq T$. If*

$$T = o(K)$$

as $K \to \infty$, then

$$N_G(T, f) = -\lim_{\delta \to 0} \frac{1}{2\pi} \Delta_{\Gamma_\delta} \arg G(s, f) + O(T + \log K) \tag{3.8}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$. Here, Γ_δ is the vertical interval $\left[\frac{1}{2} - iT, \frac{1}{2} + iT\right]$ with small right-hand side semicircle dents of radius δ centered at the zeros of $G(s, f)$ in the interval $\left[\frac{1}{2} - iT, \frac{1}{2} + iT\right]$.

Proof. Let R_δ be the rectangular region bounded by $[4 - iT, 4 + iT]$, $\left[\frac{1}{2} + iT, 4 + iT\right]$, Γ_δ, and $\left[\frac{1}{2} - iT, 4 - iT\right]$. Then, we can assume without loss of generality that there are no zeros of $G(s, f)$ in these segments. Thus, we have

$$\begin{aligned} N_G(T, f) &= \lim_{\delta \to 0} \frac{1}{2\pi} \Delta_{\partial R_\delta} \arg G(s, f) \\ &= -\lim_{\delta \to 0} \frac{1}{2\pi} \Delta_{\Gamma_\delta} \arg G(s, f) + \frac{1}{2\pi} \Delta_{\partial R_\delta \setminus \Gamma_\delta} \arg G(s, f) \end{aligned}$$

where

$$\partial R_\delta \backslash \Gamma_\delta = \left[\frac{1}{2} - iT, 4 - iT\right] + [4 - iT, 4 + iT] - \left[\frac{1}{2} + iT, 4 + iT\right].$$

By the definition of $G(s, f)$, we have

$$G(s, f) = L(s, f)\left(1 + \lambda \frac{L'}{L}(s, f)\right).$$

Then, taking the logarithmic derivative of $G(s, f)$, we obtain

$$\frac{G'}{G}(s, f) = \frac{L'}{L}(s, f) + \frac{\lambda \left(\frac{L'}{L}\right)'(s, f)}{1 + \lambda \frac{L'}{L}(s, f)}.$$

By absolute convergence and the Ramanujan-Petersson conjecture, we have

$$\frac{L'}{L}(4 + it, f) \ll 1$$

and

$$\left(\frac{L'}{L}\right)'(4 + it, f) \ll 1$$

with absolute constants. Thus, it follows that

$$\begin{aligned}\frac{G'}{G}(4 + it, f) &= \frac{L'}{L}(4 + it, f) + O((\log K)^{-1}) \\ &\ll 1.\end{aligned}$$

Hence, the variation of the argument of $G(s, f)$ along the interval $[4 - iT, 4 + iT]$ satisfies

$$\Im \frac{1}{2\pi} \int_{-T}^{T} \frac{G'}{G}(4 + it, f) i dt \ll T$$

uniformly for all $f \in H_k(1)$ with $k \asymp K$. On other hand, the variation of the argument along the horizontal segments $\left[\frac{1}{2} - iT, 4 - iT\right]$ and $\left[\frac{1}{2} + iT, 4 + iT\right]$ is bounded by $O(\log K)$ as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$. This statement is not proven here but we assume its validity based on the analogy of this situation with the case of the Riemann zeta function. □

3.1.4 A Lower Bound for $N_0(T, f)$

Now, we are ready to give a lower bound for the number of critical zeros $\rho = \frac{1}{2} + i\gamma$ of $L(s, f)$ with $|\gamma| \leq T$. The following lemma establishes such bound.

Lemma 3.1.3. *If*

$$T = o(K)$$

as $K \to \infty$, *then*

$$N_0(T, f) \geq \frac{2T}{2\pi}\left[\log\left(K^2\right)\right](1 - o(1)) - 2N_G(T, f) \tag{3.9}$$

as $K \to \infty$, *uniformly for all* $f \in H_k(1)$ *with* $k \asymp K$.

Proof. Let $\rho = \frac{1}{2} + i\gamma$ be a zero of $G(s, f)$ of order m. Then, the variation of the argument of $G(s, f)$ around a right-hand side semicircle of radius δ approaches $m\pi$ as $\delta \to 0$. Hence, from Lemma 3.1.1., it follows that

$$\begin{aligned} N_0(T, f) &\geq \lim_{\delta \to 0} \frac{1}{\pi}\Delta_{\Gamma_\delta} \arg \gamma(s) G(s, f) \\ &= \frac{1}{\pi}\Delta_\Gamma \arg \gamma(s) + \lim_{\delta \to 0} \frac{1}{\pi}\Delta_{\Gamma_\delta} \arg G(s, f) \end{aligned}$$

where Γ_δ is as in Lemma 3.1.2. and Γ is the segment of the critical line $\left[\frac{1}{2} - iT, \frac{1}{2} + iT\right]$, or equivalently, Γ_δ without the dents. Replacing Γ_δ with Γ when we compute the variation of the argument of $\gamma(s)$ is possible due to the analyticity of $\frac{\gamma'}{\gamma}(s)$. Then, by the asymptotic formula (2.24),

$$\begin{aligned} \frac{1}{\pi}\Delta_\Gamma \arg \gamma(s) &= \frac{1}{\pi}\Im \int_\Gamma \frac{\gamma'}{\gamma}(s) ds \\ &\sim \frac{2T}{\pi}\log K \\ &= \frac{2T}{2\pi}\log\left(K^2\right) \end{aligned}$$

as $K \to \infty$, uniformly for all $k \asymp K$. Hence, the lemma follows from formula (3.8). □

3.2 The Mollifier

Up until this point, Levinson's method has transformed the problem of giving a lower bound for $N_0(T, f)$ to that of finding an upper bound for the number of zeros $N_G(T, f)$ of the function $G(s, f)$ inside the rectangular region R determined by $\frac{1}{2} \leq \sigma \leq 4$ and $|t| \leq T$. As mentioned in the introduction, we rely on Littlewood's lemma to estimate this number by an integral of the logarithm of $G(s, f)$. Unfortunately, large values of

$G(s,f)$ cause the loss of the right order of magnitude when estimating this integral which should be bounded by

$$\frac{2T}{2\pi}\log\left(K^2\right).$$

This situation prevents us from producing a lower bound for $N_0(T,f)$ in terms of $N(T,f)$. The solution to this dilema is to multiply $G(s,f)$ by another function, that effectively truncates its size at very large values and does not introduce too many new zeros. Thus, we consider the inverse of $L(s,f)$ which we introduced in the previous chapter and is given by the series

$$\frac{1}{L(f,s)} = \prod_p (1-\lambda_f(p)p^{-s}+p^{-2s}) = \sum \frac{\mu_f(m)}{m^s}$$

where

$$\mu_f(m) = \sum_{m=nr^2} \lambda_f(n)\mu(nr)\mu(r). \tag{3.10}$$

Then, we define the mollifier so that it mimics $L(s,f)^{-1}$. Heuristically, something resembling the inverse of $L(s,f)$ should truncate large values of $L(s,f)$. This is our inspiration when choosing the mollifier even if it is actually large values of $G(s,f)$ we want to mollify. Let the function $P(u)$ be smooth on the interval $[0,1]$. Hence, we define the mollifier by the following Dirichlet polynomial

$$M_f(s) = \sum c_f(m)m^{-s} \tag{3.11}$$

with coefficients

$$c_f(m) = \sum_{\substack{m=nr^2\\ r\leq\Delta}} \lambda_f(n)P\left(\gamma_n\right)\mu(nr)\mu(r) \tag{3.12}$$

with γ_n defined as

$$\gamma_n = \frac{\log n}{\log\left(K^2\right)}. \tag{3.13}$$

We would also like $P(u)$ to closely mimic the shape of the function

$$\begin{cases} 1-\frac{u}{\theta} & \text{if } u\in[0,\theta]. \\ 0 & \text{if } u\in[\theta,1]. \end{cases} \tag{3.14}$$

To make this construction more precise, we let δ be a small fixed parameter kept at our disposal. Thus, we choose a smooth function $P(u)$ that satisfies

$$P(0)=1 \text{ and } P(u)\equiv 0 \text{ for } u\in[\theta-\delta,1]$$

and such that

$$P'(u) = -\frac{1}{\theta} + o(1) \quad \text{if } u \in [0, \theta - 2\delta]$$

as $\delta \to 0$. In the small transition segment $(\theta - 2\delta, \theta - \delta)$ we only need $P'(u)$ to be absolutely bounded. This is as precise as we need the choice of $P(u)$ to be in order to conduct computations and obtain the percentage featured in the statement (2.36) of the main result. Finally, the support of $P(u)$ ensures that we have a mollifier of effective length $K^{2(\theta-\delta)}$. In practice, as we will see later, the largest possible value of θ for which we can handle a mollifier of such length in our implementation of Levinson's method will be $\frac{1}{2}$.

The sum in the variable r that appears in the definition of the mollifier is a nuissance that we would gladly do away with. However, we keep it in a small range Δ also for technical reasons. It turns out that this sum helps dealing with the lack of complete multiplicativity of the Hecke eigenvalues. The sum in r will create a small gap when convoluted with another sum originating from formula (2.5). Then a small integration in the t aspect effectively kills their constribution. All the details are given when time comes but for now we let the reader know our perhaps puzzling choice of Δ.

$$\Delta = \Delta(K) = \exp\left(\log\log K\right)^3 \tag{3.15}$$

3.3 Applying Littlewood's Lemma

In this section, we follow the presentation of this material given in chapters 21 and 22 of course notes [7] by H. Iwaniec on the analogous situation of the Riemann zeta function. Let

$$F(s, f) = M(s, f)G(s, f)$$

for $f \in H_k(1)$ and weight $k \asymp K$. Hence

$$N_G(T, f) \leq N_F(T, f)$$

where $N_F(T,f)$ is defined exactly as $N_G(T,f)$ is in Lemma 3.1.2. Let $\mathcal{D}$ be the rectangular region determined by $a \le \sigma \le 4$ and $|t| \le T$ where

$$a = \frac{1}{2} - \frac{R}{\log(K^2)}$$

with $R > 0$ a fixed parameter at our disposal and T satisfying

$$T = o(K)$$

as $K \to \infty$. Then, notice that although we actually want to count the zeros of $F(s,f)$ in the rectangle R determined by $\frac{1}{2} \le \sigma \le 4$ and $|t| \le T$, it is essential for Levinson's method to create a gap by moving the left side of R a bit to the left as the region $\mathcal{D}$ shows. Without loss of generality, we can always assume that $\partial\mathcal{D}$ does not contain any zeros of $F(s,f)$. Thus, we can define the logarithm of $F(s,f)$ by

$$\log F(s,f) = \log|F(s,f)| + i \arg F(s,f)$$

as a continous branch of logarithm on $\partial\mathcal{D}$ where the argument is defined by continuous variation clockwise. Then Littlewood's lemma, in chapter 21 of [7], states that

$$\sum_{\rho \in \mathcal{D}} \mathrm{dist}(\rho) = -\Re\left(\frac{1}{2\pi i}\int_{\partial\mathcal{D}} \log F(s,f)ds\right)$$

where ρ are the zeros of $F(s,f)$ in $\mathcal{D}$ and $\mathrm{dist}(\rho)$ are their respective distances to the left side of $\mathcal{D}$. Taking the real part of the integral in Littlewood's lemma gives

$$\begin{aligned}\sum_{\rho \in \mathcal{D}} \mathrm{dist}(\rho) = &\frac{1}{2\pi}\int_{-T}^{T} (\log|F(a+it,f)| - \log|F(4+it,f)|)\, dt \\ &+ \frac{1}{2\pi}\int_{a}^{4} (\arg F(\sigma - iT,f) - \arg F(\sigma+iT,f))\, dt.\end{aligned} \tag{3.16}$$

By the Ramanujan-Petersson conjecture, we obtain

$$1 < |G(4+it,f)| \ll 1 \text{ and } 1 < |M(4+it,f)| \ll 1. \tag{3.17}$$

Thus, we have

$$\log|F(4+iT,f)| = \log|G(4+iT,f)| + \log|M(4+iT,f)| \ll 1$$

from which we get the estimate

$$\frac{1}{2\pi}\int_{-T}^{T} \log|F(4+it,f)|dt \ll T. \tag{3.18}$$

Now, we show how to estimate

$$\frac{1}{2\pi}\int_a^4 (\arg F(\sigma - iT, f) - \arg F(\sigma + iT, f))\, d\sigma. \tag{3.19}$$

Since we defined $\arg F(s)$ as a continuous variation of the argument along $\partial\mathcal{D}$, we have

$$\arg F(4+iT, f) - \arg F(\sigma + iT, f) = \int_\sigma^4 \Im \frac{F'}{F}(u+iT, f)du.$$

Then, integrating this equation over $[a, 4]$, we have

$$\begin{aligned}(4-a)\arg F(4+iT,f) - \int_a^4 \arg F(\sigma+iT,f)d\sigma &= \int_a^4\int_\sigma^4 \Im\frac{F'}{F}(u+iT,f)dud\sigma \\ &= \int_a^4 (u-a)\Im\frac{F'}{F}(u+iT,f)du.\end{aligned}$$

This last formula is also valid with T replaced by $-T$. Then, substracting the formula with $-T$ from the one with T, we obtain the following expression for (3.19)

$$(4-a)\Delta_{\Gamma_0}\arg F(s,f) - \int_a^4 (u-a)\Im\left(\frac{F'}{F}(u+iT,f) - \frac{F'}{F}(u-iT,f)\right) du$$

or equivalently

$$(4-a)\Delta_{\Gamma_0}\arg F(s,f) - 2\int_a^4 (u-a)\Im\frac{F'}{F}(u+iT,f)du$$

where Γ_0 is the right side of $\partial\mathcal{D}$. By the Ramanujan-Petersson conjecture and the lower bounds given in (3.17) , we have that

$$\begin{aligned}\Delta_{\Gamma_0}\arg F(s,f) &= \Im\int_{\Gamma_0}\frac{F'}{F}(s,f)ds \\ &= \Im\int_{\Gamma_0}\left(\frac{G'}{G}(s,f)+\frac{M'}{M}(s,f)\right)ds \ll T.\end{aligned} \tag{3.20}$$

On the other hand,

$$\int_a^4 (u-a)\Im\frac{F'}{F}(u+iT,f)du = \int_a^4 (u-a)\Im\frac{G'}{G}(u+iT,f)du + \int_a^4 (u-a)\Im\frac{M'}{M}(u+iT,f)du.$$

Again, although we do not give a proof here, we state the following estimates by analogy with the case of the Riemann zeta function

$$\int_a^4 (u-a)\Im\frac{G'}{G}(u+iT,f)du \ll \log K \tag{3.21}$$

and

$$\int_a^4 (u-a)\Im\frac{M'}{M}(u+iT,f)du \ll \log K. \tag{3.22}$$

The second of these two integrals is treated using a lemma about $\frac{M'}{M}(s)$ for general Dirichlet polynomials $M(s)$ appearing in chapter 8 of [7]. Now, we can put estimates (3.18), (3.20), (3.21) and (3.22) back to expression (3.16). The result is the following bound for $N_G(T,f)$

$$\begin{aligned}\left(\frac{1}{2}-a\right)N_G(T,f) &\le \left(\frac{1}{2}-a\right)N_F(T,f)\\ &\le \frac{1}{2\pi}\int_{-T}^{T}\log|F(a+it,f)|dt + O\left(\log K + T\right)\end{aligned}$$

where the coefficient $\left(\frac{1}{2}-\sigma\right)$ follows from dropping all zeros $\rho=\beta+i\gamma$ of $F(s,f)$ in the region $\mathcal{D}$ with $a<\beta<\frac{1}{2}$. If we replace $N_G(T,f)$ in formula (3.9) of Lemma 3.1.3. with the estimate above, then we have

$$\begin{aligned}N_0(T,f) &\ge \frac{2T}{2\pi}\left[\log\left(K^2\right)\right](1-o(1)) - \frac{1}{\pi\left(\frac{1}{2}-a\right)}\int_{-T}^{T}\log|F(a+it,f)|dt + O\left(\log K + T\right)\\ &\ge \frac{2T}{2\pi}\left[\log\left(K^2\right)\right](1-o(1)) - \frac{\log\left(K^2\right)}{\pi R}\int_{-T}^{T}\log|F(a+it,f)|dt + O\left(\log K + T\right)\end{aligned}$$

as $K\to\infty$. Hence, the previous estimate and the asymptotic formula (2.25) for $N(T,f)$ produce a lower bound for $N_0(T,f)$ which we state in the form of the following lemma.

Lemma 3.3.1. *If*

$$a = \frac{1}{2} - \frac{R}{\log\left(K^2\right)}$$

with $R>0$, and

$$T = o(K) \text{ and } T^{-1} = o(1)$$

as $K\to\infty$, then

$$N_0(T,f) \ge N(T,f)\left(1-\left(\frac{R}{2}\right)^{-1}\frac{1}{2T}\int_{-T}^{T}\log|F(a+it,f)|dt + o(1)\right) \tag{3.23}$$

as $K\to\infty$, uniformly for all $f\in H_k(1)$ with $k\asymp K$.

3.4 A Lower Bound for $\mathcal{N}_0(T,\mathcal{F})$

After establishing the lower bound (3.23) for $N_0(T,f)$, Levinson's method continues using the convexity property of $\log(x)$ and Cauchy-Schwarz inequality to give a lower bound that depends on asymptotically evaluating the integral

$$\frac{1}{2T}\int_{-T}^{T}|F(a+it,f)|^2dt.$$

In our case however, we digress slightly from this path and follow [4] in order to produce a lower bound for $\mathcal{N}_0(T,\mathcal{F})$ instead. Recall that $\mathcal{N}_0(T,\mathcal{F})$ and $\mathcal{N}(T,\mathcal{F})$ were the averages defined in (2.31) and (2.32), respectively. Thus, we give the following proposition.

Proposition 3.4.1. *Let $\mathcal{F}$ be the family given in (2.30) and let $\mathcal{N}_{\mathcal{F}}(\psi)$ be defined as in (2.34). If*

$$a = \frac{1}{2} - \frac{R}{\log(K^2)}$$

with $R > 0$, and T satisfies

$$T = o(K) \text{ and } T^{-1} = o(1)$$

as $K \to \infty$, then

$$\mathcal{N}_0(T,\mathcal{F}) \geq \mathcal{N}(T,\mathcal{F})\left(1 - R^{-1}\log\mathcal{L}(T,\mathcal{F}) + o(1)\right) \tag{3.24}$$

as $K \to \infty$, where $\mathcal{L}(T,\mathcal{F})$ is given by

$$\mathcal{L}(T,\mathcal{F}) = \frac{1}{\mathcal{N}_{\mathcal{F}}(\psi)} \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} \frac{1}{2T}\int_{-T}^{T} |F(a+it,f)|^2 dt. \tag{3.25}$$

Proof. Since the asymptotic formula (2.25) of $N(T,f)$ is uniform for all $f \in H_k(1)$ with $k \asymp K$, we have

$$\begin{aligned}
&\sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} N(T,f)\frac{1}{2T}\int_{-T}^{T} \log|F(a+it,f)|dt \\
&\sim \frac{2T}{2\pi}\left[\log\left(K^2\right)\right]\mathcal{N}_{\mathcal{F}}(\psi)\mathcal{J}(T,\mathcal{F}) \sim \mathcal{N}(T,\mathcal{F})\mathcal{J}(T,\mathcal{F})
\end{aligned}$$

as $K \to \infty$ where $\mathcal{J}(T,\mathcal{F})$ is given by

$$\mathcal{J}(T,\mathcal{F}) = \mathcal{N}_{\mathcal{F}}^{-1}(\psi) \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} \frac{1}{2T}\int_{-T}^{T} \log|F(a+it,f)|dt.$$

Thus, by the convexity of the function $\log(x)$, we obtain

$$\mathcal{J}(T,\mathcal{F}) \leq \log\mathcal{H}(T,\mathcal{F})$$

where $\mathcal{H}(T,\mathcal{F})$ is defined by

$$\mathcal{H}(T,\mathcal{F}) = \mathcal{N}_{\mathcal{F}}^{-1}(\psi) \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} \frac{1}{2T}\int_{-T}^{T} |F(a+it,f)|dt.$$

Finally, applying Cauchy-Schwarz inequality to the previous expression, we have

$$\mathcal{J}(T, \mathcal{F}) \leq \log \mathcal{H}(T, \mathcal{F}) \leq \frac{1}{2} \log \mathcal{L}(T, \mathcal{F}).$$

Then the Proposition follows after applying the weighted average (2.7) to both sides of the inequality (3.23). $\square$

Chapter 4

The Mean Value Theorem

4.1 The Statement of the Mean Value Theorem

In this section, we give the statement of a mean value theorem from which we derive a corollary for evaluating $\mathcal{L}(T, \mathcal{F})$ asymptotically as $K \to \infty$. Thus, for all $f \in \mathcal{F}$, we define the following integral

$$I_f(\alpha, \beta) = \frac{1}{2T} \int_{-T}^{T} \left| M\left(\frac{1}{2} + it, f\right) \right|^2 L\left(\frac{1}{2} + \alpha + it, f\right) L\left(\frac{1}{2} + \beta - it, f\right) dt. \quad (4.1)$$

where $M(s, f)$ is the moolifier we defined in (3.11). Then, we can state the following theorem.

Theorem 4.1.1. *If $\theta = \frac{1}{2}$ in the definition of the mollifier $M(s, f)$, and*

$$(\log \log K)^A \leq T = o(K)$$

as $K \to \infty$ for a sufficiently large $A > 0$, then

$$\sum_{k \ even} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} I_f(\alpha, \beta) \sim j\left(\alpha \log\left(K^2\right), \beta \log\left(K^2\right)\right) \mathcal{N}_{\mathcal{F}}(\psi) \quad (4.2)$$

as $K \to \infty$, uniformly for all complex numbers α and β satisfying

$$|\alpha|, |\beta| \ll \left(\log\left(K^2\right)\right)^{-1}$$

and

$$|\alpha + \beta| \asymp \left(\log\left(K^2\right)\right)^{-1}$$

where

$$j(x, y) = J(x, y) + \exp\{-(x + y)\} J(-x, -y) \quad (4.3)$$

with

$$J(x, y) = \frac{1}{x + y} \int_0^1 (P'(u) - xP(u))(P'(u) - yP(u)) du \quad \textit{for } x \neq y \quad (4.4)$$

and $\mathcal{N}_{\mathcal{F}}(\psi)$ defined as in (2.34).

Opening the product in the integrand of (4.4), we obtain another formula for $J(x,y)$. Namely,

$$J(x,y) = \frac{A(\theta)}{x+y} - B(\theta) + \frac{xy}{x+y}C(\theta) \quad \text{for } x \neq y \tag{4.5}$$

where

$$A(\theta) = \int_0^1 (P'(u))^2 du, \tag{4.6}$$

$$B(\theta) = \int_0^1 P'(u)P(u)du = -\frac{P^2(0)}{2}, \text{ and} \tag{4.7}$$

$$C(\theta) = \int_0^1 P^2(u)du. \tag{4.8}$$

The rest of the chapters in this thesis will be devoted to proving the Theorem 4.1.1. However, in the remainder of this chapter we complete the proof of the main result (2.35), Theorem 2.3.1.

4.2 Evaluating $\mathcal{L}(T,\mathcal{F})$ Asymptotically

In the following corollary, we employ the mean value theorem from the previous section to evaluate $\mathcal{L}(T,\mathcal{F})$.

Corollary 4.2.1. *If $\theta = \frac{1}{2}$,*

$$a = \frac{1}{2} - \frac{R}{\log(K^2)} \tag{4.9}$$

with $R > 0$, and T satisfies

$$(\log\log K)^{-A} \leq T = o(K)$$

as $K \to \infty$, for a sufficiently large $A > 0$, then

$$\mathcal{L}(T,\mathcal{F}) \sim c(\theta,r,R) \tag{4.10}$$

as $K \to \infty$, where $c(\theta,r,R)$ satisfies

$$r^2 c(\theta,r,R) = C(\theta,r,R) + C(\theta,1-r,-R)\exp(2R) \tag{4.11}$$

as in (1.3) and with $C(\theta,r,R)$ given by the formula

$$C(\theta,r,R) = -\left(\frac{r^2}{2} + \frac{1}{4R^2}\right)\left(\frac{1}{\theta R} + \frac{\theta R}{3}\right) - \frac{r}{2R}\left(\frac{1}{\theta R} - \frac{\theta R}{3}\right) + \frac{r^2}{2} \tag{4.12}$$

as in (1.4).

Proof. This proof follows the work in [4] closely. Let

$$I_f = \frac{1}{2T}\int_{-T}^{T}\left|M\left(\frac{1}{2}+it,f\right)G(a+it,f)\right|^2 dt \tag{4.13}$$

where

$$G(s,f) = L(s,f) + \lambda L'(s,f)$$

and

$$\lambda = (r\log\left(K^2\right))^{-1}.$$

The reader might have noticed already that integral (4.13) is not exactly

$$\frac{1}{2T}\int_{-T}^{T}|F(a+it,f)|^2\,dt$$

which is the one appearing in the definition of $\mathcal{L}(T,\mathcal{F})$. However, we can account for the discrepancy by changing the definition of the coefficients $c(m)$ of the mollifier by the factor

$$\exp\{-R\gamma_m\} \leq 1$$

where $\gamma_m = \frac{\log m}{\log(K^2)}$. This changes do not affect any of the estimates involving $M(s,f)$ in the previous chapter since in those cases the only thing assumed about $M(s,f)$ is that it is a Dirichlet polynomial of length K with relatively small coefficients. The reason for having the mollifier on the critical line is because it produces optimal results. Having settled this issue, we proceed by opening the square of the absolute value of $G(a+it,f)$ which gives the formula

$$\begin{aligned}|G(a+it,f)|^2 &= L(a+it,f)L(a-it,f) + \lambda L(a+it,f)L'(a-it,f)\\ &\quad + \lambda L'(a+it,f)L(a-it,f) + \lambda^2 L'(a+it,f)L'(a-it,f).\end{aligned}$$

By Cauchy's formula

$$f(s) = \frac{1}{2\pi i}\int_{|z|=\epsilon} f(s+z)z^{-2}dz = \oint f(s+z)z^{-2}dz,$$

we can replace the derivatives in the previous expression of $|G(a+it,f)|^2$ by complex integrals. Hence, we have the formula

$$\begin{aligned}|G(a+it,f)|^2 = {} & L(a+it,f)L(a-it,f)\\ & + \lambda \oint L(a+it,f)L(a-it+\eta,f)\eta^{-2}d\eta\\ & + \lambda \oint L(a+it+\xi,f)L(a-it,f)\xi^{-2}d\xi\\ & + \lambda^2 \oint \oint L(a+it+\xi,f)L(a-it+\eta,f)\xi^{-2}\eta^{-2}d\xi d\eta\end{aligned}$$

which, by (4.9), also have the following equivalent expression

$$\begin{gathered}L\left(\tfrac{1}{2}+it-\tfrac{R}{\log(K^2)},f\right)L\left(\tfrac{1}{2}-it-\tfrac{R}{\log(K^2)},f\right)\\ +\lambda\oint L\left(\tfrac{1}{2}+it-\tfrac{R}{\log(K^2)},f\right)L\left(\tfrac{1}{2}-it-\tfrac{R}{\log(K^2)}+\eta,f\right)\eta^{-2}d\eta\\ +\lambda\oint L\left(\tfrac{1}{2}+it-\tfrac{R}{\log(K^2)}+\xi,f\right)L\left(\tfrac{1}{2}-it-\tfrac{R}{\log(K^2)},f\right)\xi^{-2}d\xi\\ +\lambda^2\oint\oint L\left(\tfrac{1}{2}+it-\tfrac{R}{\log(K^2)}+\xi,f\right)L\left(\tfrac{1}{2}-it-\tfrac{R}{\log(K^2)}+\eta,f\right)\xi^{-2}\eta^{-2}d\xi d\eta.\end{gathered}$$

Inserting the above expression into (4.13), we obtain

$$\begin{aligned}I_f = {} & I_f\left(-\frac{R}{\log(K^2)},-\frac{R}{\log(K^2)}\right)\\ & + \lambda\oint I_f\left(-\frac{R}{\log(K^2)},-\frac{R}{\log(K^2)}+\eta\right)\eta^{-2}d\eta\\ & + \lambda\oint I_f\left(-\frac{R}{\log(K^2)}+\xi,-\frac{R}{\log(K^2)}\right)\xi^{-2}d\xi\\ & + \lambda^2\oint\oint I_f\left(-\frac{R}{\log(K^2)}+\xi,-\frac{R}{\log(K^2)}+\eta\right)\xi^{-2}\eta^{-2}d\xi d\eta\end{aligned} \tag{4.14}$$

with $I_f(\alpha,\beta)$ defined by formula (4.1). If the complex integration $\oint$ is performed on circles of sufficiently small radius, i.e. with ξ and η satisfying

$$|\xi|,|\eta| = o\left(\left(\log(K^2)\right)^{-1}\right)$$

as $K\to\infty$, then we can apply Theorem 4.1.1 with $\theta=\frac{1}{2}$ to each one of the four means that appear when replacing I_f by the equation (4.14) in the expression

$$\sum_k \psi\left(\frac{k}{K}\right)\sum_{f\in H_k(1)}^{h} I_f.$$

After normalization by $\mathcal{N}_{\mathcal{F}}(\psi)$, the statement (4.10) of the corollary is obtained with

$c(r, R)$ given by the formula

$$\begin{aligned} c(r,R) &= j(-R,-R) \\ &+ \lambda \oint j(-R, -R + \eta \log\left(K^2\right))\eta^{-2} d\eta \\ &+ \lambda \oint j(-R + \xi \log\left(K^2\right), -R)\xi^{-2} d\xi \\ &+ \lambda^2 \oint \oint j(-R + \xi \log\left(K^2\right), -R + \eta \log\left(K^2\right))\xi^{-2}\eta^{-2} d\xi d\eta \end{aligned}$$

which, by the change of variables $\xi \to \frac{\xi}{\log\left(K^2\right)}$ and $\eta \to \frac{\eta}{\log\left(K^2\right)}$, also has the equivalent expression

$$j(-R,-R) + r^{-1} \oint j(-R, -R + \eta)\eta^{-2} d\eta + r^{-1} \oint j(-R + \xi, -R)\xi^{-2} d\xi \\ + r^{-2} \oint \oint j(-R + \xi, -R + \eta)\xi^{-2}\eta^{-2} d\xi d\eta.$$

Thus, using Cauchy's formula, we obtain

$$r^2 c(\theta, r, R) = [\mathcal{D}(r)j](-R, -R) \tag{4.15}$$

where the differential operator $\mathcal{D}(r)$ is defined by the following formula

$$r^2 + r\frac{\partial}{\partial x} + r\frac{\partial}{\partial y} + \frac{\partial^2}{\partial x \partial y}. \tag{4.16}$$

For the purpose of easing computations we decompose the operator $\mathcal{D}(r)$ as follows

$$\mathcal{D}(r) = \mathcal{D}_y(r)\mathcal{D}_x(r)$$

with operators $\mathcal{D}_x(r)$ and $\mathcal{D}_u(r)$ defined by

$$r + \frac{\partial}{\partial x} \quad \text{and} \quad r + \frac{\partial}{\partial y}$$

respectively. Hence, applying the operator $\mathcal{D}_x(r)$ to formula (4.3), we have

$$[\mathcal{D}_x(r)j](x,y) = [\mathcal{D}_x(r)J](x,y) - \exp\{-(x+y)\}[\mathcal{D}_x(1-r)J](-x,-y)$$

and applying the operator $\mathcal{D}_y(r)$ to the above expression, we get

$$[\mathcal{D}(r)j](x,y) = [\mathcal{D}(r)J](x,y) + \exp\{-(x+y)\}[\mathcal{D}(1-r)J](-x,-y). \tag{4.17}$$

If we define $C(\theta, r, R)$ by the following formula

$$C(\theta, r, R) = [\mathcal{D}(r)J](-R, -R) \tag{4.18}$$

then the formula (4.11) in the statement of the corollary follows from formulas (4.15), (4.17) and (4.18). Then, from the formulas

$$\frac{\partial J}{\partial x}(x,y) = -\frac{A(\theta)}{(x+y)^2} + \frac{y^2}{(x+y)^2}C(\theta),$$
$$\frac{\partial J}{\partial y}(x,y) = -\frac{A(\theta)}{(x+y)^2} + \frac{x^2}{(x+y)^2}C(\theta),$$
$$\frac{\partial^2 J}{\partial x \partial y}(x,y) = \frac{2A(\theta)}{(x+y)^3} + \frac{2xy}{(x+y)^3}C(\theta),$$

and the definition (4.16) of operator $\mathcal{D}(r)$, it follows that

$$[\mathcal{D}(r)J](x,y) =$$
$$-\left(\tfrac{r^2}{2} + \tfrac{1}{(x+y)^2}\right)\left(-\tfrac{2A(\theta)}{x+y} - \tfrac{2xy}{x+y}C(\theta)\right) + \tfrac{r}{(x+y)}\left(-\tfrac{2A(\theta)}{(x+y)} + \tfrac{x^2+y^2}{(x+y)}C(\theta)\right) - B(\theta)r^2.$$

Thus, by the definition of $C(\theta,r,R)$ given by equation (4.18), we have

$$C(\theta,r,R) = -\left(\frac{r^2}{2} + \frac{1}{4R^2}\right)\left(\frac{A(\theta)}{R} + RC(\theta)\right) - \frac{r}{2R}\left(\frac{A(\theta)}{R} - RC(\theta)\right) - B(\theta)r^2. \tag{4.19}$$

Now, we determine the constants $A(\theta)$, $B(\theta)$, and $C(\theta)$ using their defining equations (4.6), (4.7), and (4.8), respectively. Recall that when we defined the mollifier $M(s,f)$ in chapter 3, we chose $P(u)$ so that it satisfies

$$P(0) = 1 \text{ and } P(u) \equiv 0 \text{ for } u \in [\theta - \delta, 1]$$

and such that

$$P'(u) = -\frac{1}{\theta} + o(1) \quad \text{if } u \in [0, \theta - 2\delta]$$

as $\delta \to 0$. Hence, we have

$$A(\theta) = \int_0^1 (P'(u))^2 du \to \frac{1}{\theta}$$
$$B(\theta) = \int_0^1 P'(u)P(u)du = -\frac{P^2(0)}{2} = -\frac{1}{2}$$
$$C(\theta) = \int_0^1 P^2(u)du \to \frac{\theta}{3}$$

as $\delta \to 0$. Then, formula (4.12) follows immediately from (4.19). □

With the results from Corollary 4.2.1, we can finalize the proof of the main result (2.35). Notice that by the asymptotic equation (4.10), Proposition 3.4.1. is equivalent to

$$\mathcal{N}_0(T,f) \geq (\kappa(\theta,r,R) + o(1))\mathcal{N}(T,f)$$

as $K \to \infty$, with $\kappa(\theta, r, R)$ given by

$$\kappa(\theta, r, R) = 1 - \frac{1}{R} \log c(\theta, r, R).$$

Since Corollary 4.2.1 produced the same formulas (1.3) and (1.4) as given by the authors of [4], we can take advantage of their computations (1.5). Thus

$$\kappa\,(0.5, .96, 1.24) \approx .356.$$

This proves our main result, Theorem 2.3.1.

Chapter 5

The Approximate Functional Equation

5.1 The Approximate Functional Equation

The mean value theorem from the previous chapter evaluates asymptotically an average of the integrals

$$I_f(\alpha,\beta) = \frac{1}{2T}\int_{-T}^{T} \left| M\left(\frac{1}{2}+it,f\right)\right|^2 L\left(\frac{1}{2}+\alpha+it,f\right) L\left(\frac{1}{2}+\beta-it,f\right) dt$$

over the family $\mathcal{F}$. Since α, β are very small,

$$\alpha,\beta \ll \left(\log\left(K^2\right)\right)^{-1},$$

then, as $K\to\infty$, the integral $I_f(\alpha,\beta)$ resembles the mean

$$\frac{1}{2T}\int_{-T}^{T} \left| M\left(\frac{1}{2}+it,f\right) L\left(\frac{1}{2}+it,f\right)\right|^2 dt.$$

The standard procedure for evaluating the above integral entails using an approximate functional equation of $L\left(\frac{1}{2}+it,f\right)$ to essentially replace the L-function with a Dirichlet polynomial. As a result, we can make use of harmonic anlysis techniques to evaluate such integral. However, instead of establishing approximate functional equations for $L\left(\frac{1}{2}+\alpha+it,f\right)$ and $L\left(\frac{1}{2}+\beta-it,f\right)$ separately and computing their product, the approach taken in [4] involves giving a single approximate functional equation for the product $L\left(\frac{1}{2}+\alpha+it,f\right)L\left(\frac{1}{2}+\beta-it,f\right)$. This has the benefit of sparing us from having to deal with annoying cross terms that inevitably appear when we pursue the other approach. Before we establish our own version of the approximate functional equation, we give some definitions. Thus, consider the following double Dirichlet series

$$D_f(s_1,s_2) = \sum_{l_1}\sum_{l_2} \lambda_f(l_1)\lambda_f(l_2) l_1^{-s_1} l_2^{-s_2} H\left(\frac{k}{K},\frac{l_1 l_2}{K^2}\right) \tag{5.1}$$

where

$$H(y,Y) = \frac{1}{2\pi i}\int_{(c)} \frac{\gamma(y,s_1+z)}{\gamma(y,s_1)}\frac{\gamma(y,s_2+z)}{\gamma(y,s_2)}(K^2Y)^{-z}\frac{w(z)}{z}dz \tag{5.2}$$

where the function $\gamma(y,s)$ is given by the formula

$$\gamma(y,s) = (2\pi)^{-s}\Gamma\left(s+\frac{yK-1}{2}\right) \tag{5.3}$$

and the test function $w(z)$ is chosen to be

$$w(z) = \left(1-\left(\frac{2z}{s_1+s_2-1}\right)^2\right)e^{z^2} \tag{5.4}$$

Lemma 5.1.1. *If s_1 and s_2 are complex numbers satisfying*

$$-B < \Re s_i < B \quad \text{for } i = 1,2$$

with $B > 0$ being an absolute constant, and such that

$$s_1 + s_2 \neq 1$$

then for $f \in H_k(1)$, we have

$$L(s_1,f)L(s_2,f) = D_f(s_1,s_2) + \Theta\left(\frac{k}{K},s_1,s_2\right)D_f(1-s_1,1-s_2) \tag{5.5}$$

where, $\Theta(y,s_1,s_2)$ denotes the expression

$$\Theta(y,s_1,s_2) = \frac{\gamma(y,1-s_1)}{\gamma(y,s_1)}\frac{\gamma(y,1-s_2)}{\gamma(y,s_2)}. \tag{5.6}$$

Proof. Here, we follow closely the proof of the the approximate functional equation given in [4]. Consider the complex integral

$$\frac{1}{2\pi i}\int_{(c)} \Lambda(s_1+z,f)\Lambda(s_2+z,f)\frac{w(z)}{z}dz \tag{5.7}$$

where $\Lambda(s,f)$ is the complete L-function defined in chapter 2. Let $c > 0$ be chosen temporarily such that the Dirichlet series of $L(s_1+z,f)$ and $L(s_2+z,f)$ converge absolutely. The definition (5.4) of $w(z)$ implies that for any fixed c

$$|w(z)| \ll \exp\{-(\Im z)^2\} \tag{5.8}$$

with the implicit constant depending on s_1, s_2 and c. In the second chapter, it was mentioned that $\Lambda(s,f)$ is an entire function of order 1. Hence, the complex integral

(5.7) is absolutely convergent for every fixed $c \neq 0$. Since w(0)=1, if we move the integral to the line $\Re z = -c$ by contour integration, then we pick up a pole at $z = 0$ with residue $\Lambda(s_1, f)\Lambda(s_2, f)$. Thus, we have

$$\begin{aligned}&\frac{1}{2\pi i}\int_{(c)} \Lambda(s_1 + z, f)\Lambda(s_2 + z, f)\frac{w(z)}{z}dz\\ &= \Lambda(s_1, f)\Lambda(s_2, f) + \frac{1}{2\pi i}\int_{(-c)} \Lambda(s_1 + z, f)\Lambda(s_2 + z, f)\frac{w(z)}{z}dz.\end{aligned}$$

By the functional equation (2.16) of $\Lambda(s, f)$ and the property

$$w(-z) = w(z),$$

we have, after the change of variables $z \to -z$, the following identity

$$\begin{aligned}&\frac{1}{2\pi i}\int_{(-c)} \Lambda(s_1 + z, f)\Lambda(s_2 + z, f)\frac{w(z)}{z}dz\\ &= -\frac{1}{2\pi i}\int_{(c)} \Lambda(1 - s_1 + z, f)\Lambda(1 - s_2 + z, f)\frac{w(z)}{z}dz.\end{aligned}$$

Thus, we obtain the formula

$$\begin{aligned}&\frac{1}{2\pi i}\int_{(c)} \Lambda(s_1 + z, f)\Lambda(s_2 + z, f)\frac{w(z)}{z}dz\\ &= \Lambda(s_1, f)\Lambda(s_2, f) - \frac{1}{2\pi i}\int_{(c)} \Lambda(1 - s_1 + z, f)\Lambda(1 - s_2 + z, f)\frac{w(z)}{z}dz.\end{aligned} \tag{5.9}$$

Since $\Lambda(s, f) = \gamma(s)L(s, f)$, the statement (5.5) of the lemma follows, with $D_f(s_1, s_2)$ given by the integral

$$D_f(s_1, s_2) = \frac{1}{2\pi i}\int_{(c)} \frac{\gamma(s_1 + z)}{\gamma(s_1)}\frac{\gamma(s_2 + z)}{\gamma(s_2)}L(s_1 + z, f)L(s_2 + z, f)\frac{w(z)}{z}dz$$

after dividing formula (5.9) by $\gamma(s_1)\gamma(s_2)$ and also multiplying and dividing the integral in the right hand side of the same formula by $\gamma(1-s_1)\gamma(1-s_2)$. Also, the reader should notice that

$$\gamma(s) = \gamma\left(\frac{k}{K}, s\right).$$

Finally, the expression (5.1) for $D_f(s_1, s_2)$ is obtained when we replace $L(s_1 + z, f)$ and $L(s_2 + z, f)$ in the previous formula by their respective Dirichlet series and switching the order of integration and summation. These last two steps are possible because we had appropriately chosen c such that the Dirichlet series of $L(s_1 + z, f)$ and $L(s_2 + z, f)$ were absolutely convergent on the line $\Re z = c$. □

It is not difficult to prove absolute convergence of the integral (5.2) when $c \neq 0$ using Stirling's formula and the estimate (5.8) for a fixed K. Thus, by Cauchy's theorem, $H(y, Y)$ is independent of any $c > 0$ chosen to define the integral. In addition, the reader might be puzzled by the introduction of the polynomial factor

$$1 - \left(\frac{2z}{s_1 + s_2 - 1}\right)^2$$

in the definition of the test fuction $w(z)$. It seemed unnecesary for any of the arguments in the previous proof. Like the material discussed in the next section, this is a technicality the purpose of which will become clear later in the last chapter.

5.2 Some Technicalities

It will be useful to us to know the behavior of $H(y, Y)$, $\Theta(y; s_1, s_2)$, and all their derivatives with respect to $y > 0$ as $K \to \infty$. In preparation for this study, we first prove the following lemma.

Lemma 5.2.1. *Let $\psi(z)$ be the logarithmic derivative of $\Gamma(z)$. Then,*

$$\log \Gamma(u + v) - \log \Gamma(u) = v \log K + O\left(|v| + (|v|^2 + 1)K^{-1}\right) \tag{5.10}$$

and

$$\psi^{(l)}(u + v) - \psi^{(l)}(u) \ll (|v| + 1)K^{-l-1} \tag{5.11}$$

for $l = 0, 1, \ldots,$ as $K \to \infty$, uniformly for all complex numbers u and v satisfying

$$\Re u \asymp K \quad \textit{and} \quad \Im u = o(K) \tag{5.12}$$

and

$$-A < \Re v < A \tag{5.13}$$

with $A > 0$, where the implicit constants depend on A and l.

Proof. Let $0 < \theta_0 < \frac{\pi}{2}$ be a fixed angle and assume that $0 \leq |\arg v| < \pi - \theta_0$. If $w = \frac{v}{u}$, then $0 \leq |\arg w| < \pi - \theta_0$ for a sufficiently large K because condition (5.12) implies

$\arg u \to 0$ as $K \to \infty$. Thus

$$|\log(1+w)| = \left|\int_0^{|w|} \frac{e^{i\arg w}}{1+re^{i\arg w}} dr\right|$$
$$\leq \frac{|w|}{\sin\theta_0}.$$

On the other hand, if $\pi - \theta_0 \leq |\arg v| \leq \pi$, then $|w| \to 0$ as $K \to 0$. Thus,

$$\log(1+w) \ll |w|.$$

Hence, we have proven that

$$\log\left(1+\frac{v}{u}\right) \ll \left|\frac{v}{u}\right| \tag{5.14}$$

as $K \to \infty$, uniformly for all u and v satisfying conditions (5.12) and (5.13), respectively. Stirling's formula states that

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log(z) - z + \log\sqrt{2\pi} + O\left(|z|^{-1}\right)$$

as $|z| \to \infty$, uniformly for all z satisfying $|\arg z| < \pi - \delta$ with implicit constant depending on δ. From it, we deduce that

$$\log\Gamma(u+v) - \log\Gamma(u) = v\log u + \left(u + v - \frac{1}{2}\right)\log\left(1+\frac{v}{u}\right) - v + O\left(K^{-1}\right) \tag{5.15}$$

as $K \to \infty$, uniformly for all u and v satisfying conditions (5.12) and (5.13), respectively. This is true because conditions (5.12) and (5.13) imply

$$|u| \geq |\Re u| \gg K \quad \text{and} \quad |u+v| \geq |\Re(u+v)| \gg K \tag{5.16}$$

Morover, applying the estimate (5.14) to equation (5.15) above gives

$$\log\Gamma(u+v) - \log\Gamma(u) = v\log u + O\left(|v| + (|v|^2+1)K^{-1}\right)$$

as $K \to \infty$, uniformly for all u and v satisfying conditions (5.12) and (5.13), respectively. Also, condition (5.12) of the lemma implies

$$v\log u = v\log K + O(|v|) \tag{5.17}$$

as $K \to \infty$. Hence, the statement (5.10) of the lemma follows. An asymptotic formula similar to Stirling's, gives the following evaluation of the logarithmic derivative of $\Gamma(z)$

$$\psi(z) = \log(z) + O\left(|z|^{-1}\right)$$

as $|z| \to \infty$, uniformly for all z satisfying $|\arg z| < \pi - \delta$ with the implicit constant depending on δ. Thus, (5.16) and the previous formula imply

$$\psi(u+v) - \psi(u) = \log\left(1 + \frac{v}{u}\right) + O\left(K^{-1}\right)$$

as $K \to \infty$, uniformly for all u and v satisfying conditions (5.12) and (5.13), respectively. Hence, the statement (5.11) of the lemma follows for $l = 0$ after applying (5.14) to the previous equation. Finally, there are also the following asymptotic formulas

$$\psi^{(l)}(z) = \frac{l!(-1)^{l-1}}{z^l} + O\left(|z|^{-l-1}\right)$$

for $l = 1, 2, \ldots$ as $|z| \to \infty$, uniformly for all z satisfying $|\arg z| < \pi - \delta$ with the implicit constant depending on δ. Thus, (5.16) and the previous formula imply

$$\psi^{(l)}(u+v) - \psi^{(l)}(u) = (l-1)!(-1)^{l-1}\left(\frac{1}{(u+v)^l} - \frac{1}{u^l}\right) + O\left(K^{-l-1}\right)$$

for $l = 1, 2, \ldots$ as $K \to \infty$, uniformly for all u and v satisfying conditions (5.12) and (5.13), respectively. Since

$$\left(\frac{1}{(u+v)^l} - \frac{1}{u^l}\right) = \left(\frac{v}{u(u+v)}\right)\left(\sum_{i+j=l-1} \frac{1}{(u+v)^i u^j}\right)$$

the statement (5.11) of the lemma for the cases $l = 1, \ldots$ is deduced after trivially estimating the above expression using the lower bounds in (5.16). $\square$

5.2.1 Estimating $\frac{\partial^j H}{\partial y^j}(y, Y)$

Now, we are ready to state a lemma about the asymptotic behaviour of $H(y, Y)$ and its derivatives with respect to $y > 0$ as $K \to \infty$.

Lemma 5.2.2. *Let $H(y, Y)$ be as in (5.2). Then, for any $a > 1$, $H(y, Y)$ is a smooth function of $y \in (a^{-1}, a)$ when K is sufficiently large. Also, for any $c > 0$, it satisfies*

$$\frac{\partial^l H}{\partial y^l}(y, Y) \ll Y^{-c}(\log K)^2 \quad \text{for } l = 0, 1, 2, \ldots \tag{5.18}$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and any complex numbers s_1 and s_2 satisfying

$$-B < \Re s_i < B \quad \text{for } i = 1, 2 \tag{5.19}$$

and

$$\Im s_i = o(K) \quad \textit{for } i = 1, 2 \tag{5.20}$$

with the implicit constant depending on a, c, B, and l.

Proof. Let

$$F_i(y, z) = \log \frac{\gamma(y, s_i + z)}{\gamma(y, s_i)}$$

for $i = 1, 2$. Then, for any $a > 1$, the definition of $\gamma(y, s)$ implies that $F_i(y, z)$ is smooth for $y \in (a^{-1}, a)$ when K is sufficiently large. Now, we can apply Lemma 5.2.1. with $u = s_i + \frac{yK-1}{2}$ and $v = z$ for all $y \in (a^{-1}, a)$, all z satisfying $\Re z = c$, and all complex numbers s_1 and s_2 satisfying (5.19) and (5.20). Thus, we have

$$F_i(y, z) = z \log K + O\left(|z| + (|z|^2 + 1)K^{-1}\right)$$

and

$$\frac{\partial^l F_i}{\partial y^l}(y, z) \ll (|z| + 1) \quad \text{for } l = 0, 1, 2, \ldots$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$, all z with $\Re z = c$, and all complex numbers s_1 and s_2 satisfying (5.19) and (5.20) with the implicit constants depending on a, B, c, and l. Now, we define

$$G_i(y, z) = \exp\{F_i(y, z)\}$$

and, for $l = 1, \ldots,$ we apply Faà di Bruno's formula which states that

$$\frac{d^l}{dx^l}\left[g(f(x))\right] = \sum_{n=0}^{l} g^{(n)}(f(x)) B_{l,n}\left(f'(x), ..., f^{(l-n+1)}(x)\right) \tag{5.21}$$

where $B_{l,n}(x_1, ..., x_{l-n+1})$ are the Bell polynomials

$$B_{l,n}(x_1, ..., x_{l-n+1}) = \sum \frac{l!}{j_1! j_2! ... j_{l-n+1}!} \left(\frac{x_1}{j_1!}\right)^{j_1} ... \left(\frac{x_{l-n+1}}{j_{l-n+1}!}\right)^{j_{l-n+1}} \tag{5.22}$$

with the sum defined over the integers $j_1 \geq 0$,..., $j_{l-n+1} \geq 0$ satisfying $j_1 + ... + j_{l-n+1} = l$ and $j_1 + 2j_2 + ... + (l - n + 1)j_{l-n+1} = l$. Thus, we obtain the following formula for $\frac{\partial^l G_i}{\partial y^l}(y, z)$

$$\frac{\partial^l G_i}{\partial y^l}(y, z) = \exp\{F_i(y, z)\} \sum_{n=1}^{l} B_{l,n}\left(\frac{\partial F_i}{\partial y}(y, z), ..., \frac{\partial^{l-n+1} F_i}{\partial y^{l-n+1}}(y, z)\right).$$

Then, estimating trivially the above expression by using the estimates for $\frac{\partial^l F_i}{\partial y^l}(y,z)$, it follows that

$$\frac{\partial^l G_i}{\partial y^l}(y,z) \ll \exp\{c\log K + O\left(|z| + (1+|z|^2)K^{-1}\right)\}(|z|+1)^l \quad \text{for } l = 0,1,2,\ldots$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$, all z with $\Re z = c$, and all complex numbers s_1 and s_2 satisfying (5.19) and (5.20) with the implicit constants depending on a, B, c, and l. Moreover, if we define $G(y,z) = G_1(y,z)G_2(y,z)$ and compute its l^{th} order partial derivative with respect to y, we obtain

$$\frac{\partial^l G}{\partial y^l}(y,z) = \sum_{i+j=l} \frac{l!}{i!j!}\frac{\partial^i G_1}{\partial y^i}(y,z)\frac{\partial^j G_2}{\partial y^j}(y,z).$$

Then, we can estimate trivially the above expression using the estimates for $\frac{\partial^l G_i}{\partial y^l}(y,z)$. Thus, we have

$$\frac{\partial^l G}{\partial y^l}(y,z) \ll \exp\{2c\log K + O\left(|z| + (|z|^2+1)K^{-1}\right)\}(|z|+1)^l \quad \text{for } l = 0,1,2,\ldots$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$, all z with $\Re z = c$, and all complex numbers s_1 and s_2 satisfying (5.19) and (5.20) with the implicit constants depending on a, B, c and l. Hence, the statement of the lemma follows by estimating trivially the l^{th} order partial derivative of $H(y,Y)$ with respect to y

$$\frac{\partial^l H}{\partial y^l}(y,Y) = \frac{1}{2\pi i}\int_{(c)} \frac{\partial^l G}{\partial y^l}(y,z)(K^2Y)^{-z}\frac{w(z)}{z}dz.$$

□

5.2.2 Estimating $\frac{\partial^j \Theta}{\partial y^j}(y, s_1, s_2)$

Similarly, we state a lemma describing the behaviour of $\Theta(y, s_1, s_2)$ and its derivatives with respect to y as $K \to \infty$.

Lemma 5.2.3. *Let $\Theta(y, s_1, s_2)$ be given by formula (5.6). Then, for any $a > 1$, $\Theta(y, s_1, s_2)$ is a smooth function of $y \in (a^{-1}, a)$ when K sufficiently large. Also, it satisfies*

$$\frac{\partial^l \Theta}{\partial y^l}(y, s_1, s_2) \ll 1 \quad \textit{for } l = 0,1,\ldots \tag{5.23}$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying

$$-B < \Re s_i < B \quad \textit{for } i = 1, 2 \tag{5.24}$$

$$\Im s_i = o(K) \quad \textit{for } i = 1, 2 \tag{5.25}$$

and

$$1 - s_1 - s_2 \ll \left(\log\left(K^2\right)\right)^{-1} \tag{5.26}$$

with implicit constant depending on a, B, and l. Moreover, for the case $l = 0$ we have the precise asymptotic formula

$$\Theta(y, s_1, s_2) \sim K^{2(1-s1-s2)} \tag{5.27}$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying conditions (5.24), (5.25), and (5.26) with implicit constant depending on a and B.

Proof. In this case, we consider the quotient

$$F(y, s_1, s_2) = \frac{\gamma(y, 1 - s_2)}{\gamma(y, s_1)}.$$

For any $a > 1$, the definition of $\gamma(y, s)$ again implies that the function $F(y, s_1, s_2)$ is smooth for $y \in (a^{-1}, a)$ when K sufficiently large. Then, we apply Lemma 5.2.1 to $F(y, s_1, s_2)$ with $u = s_2 + \frac{yK-1}{2}$ and $v = 1 - s_1 - s_2$ for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying conditions (5.24), (5.25), and (5.26). Thus, we obtain

$$F(y, s_1, s_2) = (1 - s_1 - s_2) \log K + O\left((\log K)^{-1}\right)$$

and

$$\frac{\partial^l F}{\partial y^l}(y, s_1 s_2) \ll 1 \quad \text{for } l = 0, 1, 2, \ldots$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying conditions (5.24), (5.25), and (5.26) with the implicit constant depending on a, B, and l. The same two equations hold for $F(y, s_2, s_1)$ and under the same uniformity conditions if we apply Lemma 5.2.1 with $u = s_1 + \frac{yK-1}{2}$ instead. Hence, if we define $G(y, s_1, s_2)$ by

$$G(y, s_1, s_2) = \exp\left\{F(y, s_1, s_2)\right\}$$

then, we have

$$\Theta(y, s_1, s_2) = G(y, s_1, s_2)G(y, s_2, s_1) \sim K^{2(1-s_1-s_2)}$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying conditions (5.24), (5.25), and (5.26) with the implicit constant depending on a and B. By the same arguments used in the proof of Lemma 5.2.2., it also follows that

$$\frac{\partial^l G}{\partial y^l}(y, s_1, s_2) \ll 1 \quad \text{for } l = 0, 1, 2, \ldots$$

as $K \to \infty$, uniformly for all $y \in (a^{-1}, a)$ and all complex numbers s_1 and s_2 satisfying conditions (5.24), (5.25), and (5.26) with the implicit constant depending on a, B, and l. Then, estimating trivially the l^{th} order derivative of $G(y, s_1, s_2)G(y, s_2, s_1)$, the statement (5.23) of the lemma follows. □

5.2.3 The Effective Length of $D_f(s_1, s_2)$

One immediate application of Lemma 5.2.2 is establishing that for any $\epsilon > 0$, the sums over variables l_1 and l_2 in $D_f(s_1, s_2)$ are effectively constrained by the condition

$$l_1 l_2 \leq K^{2+\epsilon} \tag{5.28}$$

as $K \to \infty$. More precisely, we have the following Corollary of Lemma 5.2.2.

Corollary 5.2.4. *Let $D_f(s_1, s_2)$ be defined as in (5.1). For any $\epsilon > 0$, the tail $\mathcal{T}_f$ of $D_f(s_1, s_2)$ defined by*

$$\mathcal{T}_f = \sum \sum_{l_1 l_2 > K^{2+\epsilon}} \lambda_f(l_1)\lambda_f(l_2) l_1^{-s_1} l_2^{-s_2} H\left(\frac{k}{K}, \frac{l_1 l_2}{K^2}\right)$$

satisfies, for every $D > 0$

$$\mathcal{T}_f \ll_D K^{-D}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$ and all complex numbers s_1 and s_2 satisfying

$$-B < \Re s_i < B \quad \text{for } i = 1, 2 \tag{5.29}$$

and

$$\Im s_i = o(K) \quad \text{for } i = 1, 2. \tag{5.30}$$

Proof. To show this, we choose $a > 1$ such that $\frac{k}{K} \in (a^{-1}, a)$ for all k in the range determined by the condition $k \asymp K$. Then, we estimate trivially the tail $\mathcal{T}_f$ of $D_f(s_1, s_2)$ using the estimate (5.18) from Lemma 5.2.2. Thus, we have

$$\mathcal{T}_f \ll K^{2c}(\log K)^2 \sum_{l > K^{2+\epsilon}} c(l) l^{-c}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$ and all complex numbers s_1 and s_2 satisfying conditions (5.29) and (5.30). In the previous equation, the coefficients $c(l)$ are given by

$$c(l) = \left| \sum_{l = l_1 l_2} \lambda_f(l_1)\lambda_f(l_2) l_1^{-s_1} l_2^{-s_2} \right| .$$

By the Ramanujan-Petersson Conjecture (2.10), we have

$$c(l) \ll \tau_4(l) l^{2B} \ll_\epsilon l^{2B+\epsilon}$$

which implies

$$\begin{aligned} \mathcal{T}_f \ll & K^{2c+(1+2B+\epsilon-c)(2+\delta)}(\log K)^2 \\ \ll & K^{(1+2B+\epsilon)(2+\epsilon)-c\epsilon+\epsilon} \end{aligned}$$

as $K \to \infty$, uniformly for all $f \in H_k(1)$ with $k \asymp K$ and all complex numbers s_1 and s_2 satisfying conditions (5.29) and (5.30). Hence, the result of the Corollary follows after choosing

$$c = [D + \epsilon + (1 + 2B + \epsilon)(2 + \epsilon)]\, \epsilon^{-1}.$$

□

Chapter 6

The Large Sieve

6.1 An Introduction to the Large Sieve

At the end of the previous chapter, we showed that the double series $D_f(s_1, s_2)$ in the functional equation (5.5) of the product $L(s_1, f)L(s_2, f)$ has an effective length determined by the condition

$$l_1 l_2 \leq K^{2+\epsilon}$$

as $K \to \infty$. If we replace s_1 by $\frac{1}{2} + \alpha + it$ and s_2 by $\frac{1}{2} + \beta - it$ with α and β satisfying

$$\alpha,\, \beta \ll \left(\log\left(K^2\right)\right)^{-1}$$

then we could think of the double series $D_f(s_1, s_2)$ as essentially the square of the absolute value of the following Dirichlet polynomial of length K

$$\sum_{l \leq K} \frac{\lambda_f(l)}{\sqrt{l}} l^{it}.$$

Hence, applying the approximate functional equation reduces the mean value theorem to essentially evaluating the average over the family $\mathcal{F}$ of the integrals

$$\frac{1}{2T} \int_{-T}^{T} \left| M\left(\frac{1}{2} + it, f\right) \left(\sum_{l \leq K} \frac{\lambda_f(l)}{\sqrt{l}} l^{it}\right) \right|^2 dt$$

Since $T = o(K)$ as $K \to \infty$, we can postpone the integration in the t aspect and focus on the major source of cancellation which is provided by averaging over the family $\mathcal{F}$. Recall the definition (3.12) of the mollifier $M\left(\frac{1}{2} + it, f\right)$ given in the third chapter; it was essentially the following Dirichlet polynomial of length slightly smaller than K

$$\sum_{n \leq K^{1-\epsilon}} \frac{\lambda_f(n)c(n)}{\sqrt{n}} n^{it}.$$

Let $\psi(x)$ be a smooth function of compact support in the positive reals. If we momentarily disregard the lack of complete multiplicativity of the Hecke eigenvalues $\lambda_f(n)$, $f \in \mathcal{F}$, then all these simplifications show that we are essentially trying to evaluate the expression

$$\sum_k \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} \left| \sum_{m \leq K^{2-\epsilon}} a_m \lambda_f(m) \right|^2 \tag{6.1}$$

where the linear form inside the absolute value is the result of the product between the mollifier and the Dirichlet polynomial from the approximate functional equation.

A Large Sieve Inequality provides an upper bound for expressions like (6.1) with arbitrary coefficients a_m in terms of the norm

$$\|a\|^2 = |a_1|^2 + |a_2|^2 + ...$$

This bound produces a non-trivial saving if the number of "harmonics" λ_f, for $f \in \mathcal{F}$, over which we take the average is larger than the length of the linear form

$$\sum_{m \leq K^{2-\epsilon}} a_m \lambda_f(m)$$

Obviously, this is the case for us as the initial heuristic has shown, given that we average over approximately K^2 members of the family $\mathcal{F}$. One of the reasons Large Sieve results are so powerful is because they do not assume anything about the coefficients a_n which in practice could be quite complicated.

Unfortunately, for our purposes it will be necessary to obtain the sharpest result possible or we risk the disastrous result of proving a negative percentage of critical zeros due to the delicate nature of Levinson's method. This means it will be necessary to evaluate the expression (6.1) instead. It turns out we will pay a price for this in the form of giving up on the arbitrariness of the coefficients a_m. In [5], the authors develop the Asymptotic Large Sieve (ALS) which accomplishes these objectives in the case in which the harmonics associated with the family $\mathcal{F}$ are Dirichlet characters modulo q. In this chapter, we develop a version of ALS where the harmonics are the Hecke eigenvalues of each $f \in H_k(1)$ for every even weight k.

6.2 The Statement of the Asymptotic Large Sieve

In this section, we state a version of the Asymptotic Large Sieve which is suitable for use in the proof of the Mean Value Theorem. First, we define the billinear form

$$\mathcal{S}(\mathcal{A}\times\mathcal{B}) = \sum_{n_1}\sum_{l_1}\sum_{n_2}\sum_{l_2}\alpha_{n_1l_1}\beta_{n_2l_2}\sum_{k \text{ even}} F\left(\frac{k}{K},\frac{l_1l_2}{L}\right)\Delta_k(n_1l_1,n_2l_2), \tag{6.2}$$

where $\Delta_k(m,n)$ is given by the formula (2.6). Also, we assume that the test function $F(y,Y) \in C_0^\infty(\mathbb{R}^+)$ has compact support inside some positive open interval where, for every $A \geq 0$ and $\epsilon > 0$, it satisfies

$$\frac{\partial^j F}{\partial y^j}(y,Y) \ll Y^{-A}K^\epsilon \quad \text{for } j = 0,1,2,\ldots \tag{6.3}$$

for all y with the implied constant depending on j, ϵ, and A only. If we replace $\Delta_k(m,n)$ in (6.2) with the Kronecker symbol $\delta(m,n)$ as the result of applying the Petersson formula (2.8), then we obtain

$$\mathcal{S}_{\text{diag}}(\mathcal{A}\times\mathcal{B}) = \sum_{n_1}\sum_{l_1}\sum_{n_2}\sum_{l_2}\alpha_{m_1l_1}\beta_{m_2l_2}\delta(m_1l_1,m_2l_2)\sum_{k \text{ even}} F\left(\frac{k}{K},\frac{l_1l_2}{L}\right). \tag{6.4}$$

This last definition is intended to extract the "diagonal" terms of the bilinear form (6.2). The following theorem states that if we were to evaluate the bilinear form (6.2), then the main term would necessarily come from evaluating the diagonal. However, as we mentioned in the previous section, it will be necessary to make some assumptions about the coefficients.

Theorem 6.2.1. *(The Asymptotic Large Sieve) Let the coefficients α_{nl} and β_{nl} satisfy*

$$\alpha_{nl}, \beta_{nl} \ll_\epsilon (nl)^{-\frac{1}{2}+\epsilon} \tag{6.5}$$

for all $\epsilon > 0$ with the variables n_1, n_2 supported on the interval $[1,N]$. If

$$K^\epsilon \leq L \leq K^2$$

and

$$N \leq K^{1-\delta}$$

then

$$\mathcal{S}(\mathcal{A}\times\mathcal{B}) - \mathcal{S}_{diag}(\mathcal{A}\times\mathcal{B}) \ll K^\phi \tag{6.6}$$

as $K \to \infty$ for some $\phi < 1$ with the implicit constant depending on ϕ.

6.3 The Main Lemma

The main ingredient in the proof of the Asymptotic Large Sieve is the following lemma of H. Iwaniec for evaluating Neumann series of the form

$$G_a(x) = \sum_{l \equiv a \mod 4} g(l) J_l(x) \tag{6.7}$$

for $a = \pm 1$ where the functions $J_l(x)$ are classical Bessel function of integral order and $g \in C_0^\infty(\mathbb{R}^+)$ is a smooth function with compact support in some positive open interval.

Lemma 6.3.1. *If $G_a(x)$ is defined as in (6.7) for $a = \pm 1$, then*

$$4G_a(x) = g(x) + i^{1-a} h(x) + O\left(x c_3(g)\right) \tag{6.8}$$

with

$$h(x) = \int_0^\infty g(\sqrt{2xu}) \sin\left(x + u - \frac{\pi}{4}\right) (\pi u)^{-\frac{1}{2}} du \tag{6.9}$$

and

$$c_j(g) = \int_0^\infty |\hat{g}(t)| \, |t|^j dt \tag{6.10}$$

where $\hat{g}(t)$ is the Fourier transform of $g(x)$

$$\hat{g}(t) = \int g(x) e(x) dx.$$

We do not give the proof of this lemma here and instead remit the reader to [6] where a detailed proof is given. However, we hint that the proof is essentially an application of the Poisson summation formula.

6.4 The Proof of ALS

In this section, we prove ALS. As mentioned earlier, the Petersson formula (2.8) which was introduced in chapter two, is used to isolate the main term $\mathcal{S}_{\text{diag}}(\mathcal{A} \times \mathcal{B})$ in the evaluation of the bilinear form $\mathcal{S}(\mathcal{A} \times \mathcal{B})$. Thus, we have

$$\mathcal{S}(\mathcal{A} \times \mathcal{B}) = \mathcal{S}_{\text{diag}}(\mathcal{A} \times \mathcal{B}) + \mathcal{R}(\mathcal{A} \times \mathcal{B}) \tag{6.11}$$

where the remainder $\mathcal{R}(\mathcal{A} \times \mathcal{B})$ collects the effect of those other terms that, because of the quasi-orthogonality property satisfied by the Hecke eigenvalues, end up producing sufficient cancellation and do not constribute to the main term. According to the Petersson formula, we define the remainder by

$$\mathcal{R}(\mathcal{A} \times \mathcal{B}) = \sum\sum\sum\sum \alpha_{n_1 l_1} \beta_{n_2 l_2} G(n_1 l_1, n_2 l_2) \tag{6.12}$$

with $G(m, n)$ given by the series

$$G(m, n) = \sum_{c>0} \frac{S(m, n; c)}{c} G\left(\frac{4\pi\sqrt{mn}}{c}\right) \tag{6.13}$$

and where $G(x)$ is a Neumann series of the form

$$G(x) = \sum_{k \text{ even}} i^k F\left(\frac{k}{K}, Y\right) J_{k-1}(x). \tag{6.14}$$

It turns out that the necessary cancellation can only be obtained if the Petersson formula is applied for a sufficiently large number of weights k. This last imput is technically accomplished by the main lemma. If we apply the change of variable $k \to l + 1$, then the above series can be split into two Neumann series

$$G(x) = G_+(x) + G_-(x)$$

where

$$G_a(x) = -a \sum_{l \equiv a \mod 4} F\left(\frac{l+1}{K}, Y\right) J_l(x)$$

for $a = \pm 1$. Thus, we can apply the main lemma to these two series with $g(x) = F\left(\frac{x}{K}, Y\right)$ and obtain the following evaluation of $G(x)$

$$4G(x) = -h(x) + O\left(x c_3(g)\right). \tag{6.15}$$

To proceed any further, it will be necessary to produce estimates for $c_3(g)$ and $h(x)$.

6.4.1 Estimating $c_j(g)$

Since $g(x) = F\left(\frac{x}{K}, Y\right)$, the conditions imposed on the test function $F(y, Y)$ imply that the Fourier transform of $g(x)$ satisfies

$$\begin{aligned}
\hat{g}(t) &= \int_{-\infty}^{\infty} g(x) e(xt) dx \\
&= \int_{-\infty}^{\infty} F\left(\frac{x}{K}, Y\right) e(xt) dx \\
&= K \int_{-\infty}^{\infty} F(y, Y) e(yKt) dy
\end{aligned}$$

which proves that for all $A > 0$ and $\epsilon > 0$

$$|\hat{g}(t)| \ll Y^{-A} K^{1+\epsilon} \tag{6.16}$$

for all t with the implicit constant depending on A and ϵ. Also, by partial integration, for each $j \geq 1$ we have

$$\begin{aligned}(-t2\pi i)^j \hat{g}(t) &= \int_{-\infty}^{\infty} g^j(x) e(xt) dx \\ &= K^{-j} \int_{-\infty}^{\infty} \frac{\partial^j}{\partial x^j} F\left(\frac{x}{K}, Y\right) e(xt) dx \\ &= K^{1-j} \int_{-\infty}^{\infty} \frac{\partial^j F}{\partial y^j} (y, Y)\, e(yKt) dy\end{aligned}$$

from which we obtain that for all $A > 0$ and $\epsilon > 0$

$$(K|t|)^j |\hat{g}(t)| \ll Y^{-A} K^{1+\epsilon} \quad \text{for } j = 1, 2, \ldots \tag{6.17}$$

for all t with the implicit constant depending on A, j, and ϵ. Combining the estimates (6.16) and (6.17), we have for all $A > 0$ and $\epsilon > 0$

$$(1 + K|t|)^j |\hat{g}(t)| \ll Y^{-A} K^{1+\epsilon} \quad \text{for } j = 0, 1, 2, \ldots$$

for all t. Hence, from the definition of $c_j(g)$, it follows that

$$\begin{aligned}c_j(g) &= \int_0^{\infty} |\hat{g}(t)|\, t^j dt \\ &\ll \left[\int_0^{\infty} (1 + Kt)^{-2-j} t^j K dt\right] Y^{-A} K^{\epsilon} \\ &\ll \left[\int_0^{\infty} (1 + v)^{-2} dv\right] Y^{-A} K^{-j+\epsilon}.\end{aligned}$$

Thus, for all $A > 0$ and $\epsilon > 0$ we have

$$c_j(g) \ll Y^{-A} K^{-j+\epsilon} \quad \text{for } j = 0, 1, 2, \ldots \tag{6.18}$$

with the implicit constant depending on A, j, and ϵ.

6.4.2 Estimating $h(x)$

In (6.9), we defined $h(x)$ by

$$h(x) = \int_0^{\infty} g(\sqrt{2xu}) \sin\left(x + u - \frac{\pi}{4}\right) (\pi u)^{-\frac{1}{2}} du.$$

By partial integration, for each $j \geq 0$, we obtain

$$\begin{aligned}|h(x)| &\leq \pi^{-\frac{1}{2}} \int_0^\infty \left| \frac{\partial^i}{\partial u^j} \left[g(\sqrt{2xu}) u^{-\frac{1}{2}} \right] \right| du \\ &= \pi^{-\frac{1}{2}} \int_0^\infty \left| \sum_{i+k=j} \frac{j!}{i!k!} \frac{\partial^i}{\partial u^i} \left[g(\sqrt{2xu}) \right] \left(-\frac{1}{2} \right) \dots \left(-\frac{1}{2} - k + 1 \right) u^{-\frac{1}{2}-k} \right| du \\ &\ll \int_0^\infty \sum_{i+k=j} \left| \frac{\partial^i}{\partial u^i} \left[g(\sqrt{2xu}) \right] \right| u^{-\frac{1}{2}-k} du\end{aligned}$$

with the implicit constant depending on j. By Faà di Bruno's formula (5.21) which was introduced in the previous chapter, we have

$$\left| \frac{\partial^i}{\partial u^i} \left[g(\sqrt{2xu}) \right] \right| \leq \sum_{n=0}^{i} \left| g^{(n)}(\sqrt{2xu}) B_{i,n} \left(\left(\sqrt{2xu} \right)', \left(\sqrt{2xu} \right)'', \dots \right) \right|$$

where the Bell polynomials (5.22) satisfy

$$\left| B_{i,n} \left(\left(\sqrt{2xu} \right)', \left(\sqrt{2xu} \right)'', \dots \right) \right| \ll \left(\sqrt{2xu} \right)^i u^{-i}$$

with the implicit constant depending on i. Thus, we have

$$\begin{aligned}|h(x)| &\ll \int_0^\infty u^{-j} \sum_{i+k=j} \left(\sqrt{2xu} \right)^i \left\{ \sum_{n=0}^{i} \left| g^{(n)}(\sqrt{2xu}) \right| \right\} u^{-\frac{1}{2}} du \\ &= \int_0^\infty u^{-j} \sum_{i+k=j} \left(\sqrt{2xu} K^{-1} \right)^i \left\{ \sum_{n=0}^{i} \left| \frac{\partial^n F}{\partial y^n} \left(\sqrt{2xu} K^{-1}, Y \right) \right| \right\} u^{-\frac{1}{2}} du \\ &\ll \left(\int_0^\infty y^{-j} \sum_{i+k=j} y^{-k} \sum_{n=0}^{i} \left| \frac{\partial^n F}{\partial y^n} (y, Y) \right| dy \right) \left(xK^{-2} \right)^{-j-\frac{1}{2}}.\end{aligned}$$

This proves that when $x \leq K^2$, for all $A > 0$ and $\epsilon > 0$ we have

$$|h(x)| \ll \left(xK^{-2} \right)^{-j} Y^{-A} K^\epsilon \quad \text{for } j = 0, 1, 2, \dots \tag{6.19}$$

with the implicit constant depending on j, A, and ϵ.

6.4.3 Estimating $G(m, n)$

Now that we have estimates (6.18) and (6.19) for $c_3(g)$ and $h(x)$, respectively, we obtain for all $A > 0$ and $\epsilon > 0$ the following bound for $G(x)$

$$G(x) \ll \left[\left(xK^{-2} \right)^j + xK^{-3} \right] Y^{-A} K^\epsilon \quad \text{for } j = 0, 1, 2, \dots$$

with the implicit constant depending only on j, A, and ϵ. Thus, from the definition (6.13) of $G(m,n)$, it follows that

$$G(m,n) \ll \left[\left(\sqrt{mn}K^{-2}\right)^j \sum_{c>0} \frac{|S(m,n;c)|}{c^{1+j}} + \sqrt{mn}K^{-3} \sum_{c>0} \frac{|S(m,n;c)|}{c^2} \right] Y^{-A} K^{\epsilon}.$$

By the Weil bound for Kloosterman sums

$$|S(m,n;c)| \ll (m,n;c)^{\frac{1}{2}} c^{\frac{1}{2}} \tau(c),$$

if $j \geq 1$, then we have

$$\begin{aligned}
\sum_{c>0} \frac{|S(m,n;c)|}{c^{1+j}} &\leq \sum_{c>0} \frac{|S(m,n;c)|}{c^2} \\
&\ll \sum_{c>0} (m,n;c)^{\frac{1}{2}} \tau(c) c^{-\frac{3}{2}} \\
&= \sum_{d|(m,n)} \frac{\tau(d)}{d} \sum_{\substack{a>0 \\ (m,n,a)=1}} \tau(a) a^{-\frac{3}{2}} \\
&\ll (mn)^{\epsilon}.
\end{aligned}$$

Hence, we obtain the following estimate for $G(m,n)$. For all $A > 0$ and $\epsilon > 0$, we have

$$G(m,n) \ll \left[\left(\sqrt{mn}K^{-2}\right)^j + \sqrt{mn}K^{-3} \right] Y^{-A} (mnK)^{\epsilon} \quad \text{for } j = 1, 2, \ldots \tag{6.20}$$

with the implicit constant depending only on j, A, and ϵ.

6.4.4 Estimating $\mathcal{R}(\mathcal{A} \times \mathcal{B})$

To estimate $\mathcal{R}(\mathcal{A} \times \mathcal{B})$, we will first estimate the summation in the variables n_1 and n_2. Thus, we write

$$\mathcal{R}(\mathcal{A} \times \mathcal{B}) = \sum_{l_1} \sum_{l_2} \mathcal{R}(l_1, l_2)$$

where

$$\mathcal{R}(l_1, l_2) = \sum_{n_1} \sum_{n_2} \alpha_{n_1 l_1} \beta_{n_2 l_2} G(n_1 l_1, n_2 l_2).$$

Then, we estimate $\mathcal{R}(\mathcal{A} \times \mathcal{B})$ trivially using the assumptions (6.5) of the theorem about the coefficients

$$\alpha_{n,l},\ \beta_{n,l} \ll_{\epsilon} (nl)^{-\frac{1}{2}+\epsilon}$$

and the bound (6.20) for $G(m,n)$. Thus, we have

$$\mathcal{R}(l_1,l_2) \ll \left[\left(\sqrt{l_1l_2}NK^{-2}\right)^j (l_1l_2)^{-\frac{1}{2}}N + K^{-3}N^2\right] \left(l_1l_2L^{-1}\right)^{-A} (l_1l_2NK)^{\epsilon}$$

for all $j \geq 1$, $A > 0$, and $\epsilon > 0$ with the implicit constant depending only on j, A, and ϵ. Now, if we let

$$\mathcal{R}(l) = \sum_{l=l_1l_2} \mathcal{R}(l_1,l_2)$$

then the previous estimate gives

$$\mathcal{R}(l) \ll \left[\left(\sqrt{l}NK^{-2}\right)^j l^{-\frac{1}{2}}N + K^{-3}N^2\right] \left(lL^{-1}\right)^{-A} (lNK)^{\epsilon} \tag{6.21}$$

for all $j \geq 1$, $A > 0$, and $\epsilon > 0$ with the implicit constant depending only on j, A, and ϵ. Let $\eta > 0$ be relatively small compared to δ, then we will split $\mathcal{R}(\mathcal{A}\times\mathcal{B})$ into two sums

$$\mathcal{R}(\mathcal{A}\times\mathcal{B}) = \mathcal{R}_0 + \mathcal{R}_1,$$

where

$$\mathcal{R}_0 = \sum_{l \leq L^{1+\eta}} \mathcal{R}(l),$$

and

$$\mathcal{R}_1 = \sum_{l > L^{1+\eta}} \mathcal{R}(l).$$

First, we estimate $\mathcal{R}_0$. Here, we will take advantage of the arbitrary size of A, ϵ, and j. Thus, by estimate (6.21) with $N \leq K^{1-\delta}$ and $L \leq K^2$, we have

$$\begin{aligned}\mathcal{R}_0 &\ll \left[\left(L^{\frac{1}{2}(1+\eta)}NK^{-2}\right)^j L^{\frac{1}{2}(1+\eta)}N + L^{1+\eta}K^{-3}N^2\right] L^A(L^{1+\eta}NK)^{\epsilon} \\ &\leq \left[K^{j(\eta-\delta)+2+\eta-\delta} + K^{1+2(\eta-\delta)}\right] K^{2A+\epsilon(4+2\eta-\delta)}.\end{aligned}$$

Hence, if we choose A and ϵ sufficiently small and j sufficiently large, then there exists $0 < \phi < 1$ such that

$$\mathcal{R}_0 \ll K^{\phi}$$

as $K \to \infty$ with the implicit constant depending on β_0. In the case of $\mathcal{R}_1$, we have

$$\begin{aligned}\mathcal{R}_1 &\ll \left[L^{(1+\eta)(-A+\frac{j}{2}+\frac{1}{2}+\epsilon)}\left(NK^{-2}\right)^j N + L^{(1+\eta)(-A+1+\epsilon)}K^{-3}N^2\right] L^A(NK)^{\epsilon} \\ &\leq L^{-A\eta}\left[L^{(1+\eta)(\frac{j}{2}+\frac{1}{2}+\epsilon)}\left(NK^{-2}\right)^j N + L^{(1+\eta)(1+\epsilon)}K^{-3}N^2\right](NK)^{\epsilon}.\end{aligned}$$

Since $L \geq K^\epsilon$, choosing A sufficiently large gives

$$\mathcal{R}_1 \to 0$$

as $K \to \infty$. This finishes the proof of Theorem 6.2.1 (ALS.)

6.4.5 Evaluating $\mathcal{N}_\mathcal{F}(\psi)$

Here, we evaluate the normalizing factor $\mathcal{N}_\mathcal{F}(\psi)$. This is obtained at no extra cost from the procedure we used to give the estimate (6.20) of $G(m,n)$. Thus, we have the following lemma

Lemma 6.4.1. *Let $\psi \in C_0^\infty(\mathbb{R}^+)$ is a smooth function with compact support in some positive open interval and let $\mathcal{N}_\mathcal{F}(\psi)$ be defined as in (2.34). Then, we have*

$$\mathcal{N}_\mathcal{F}(\psi) \sim \sum_{k \ even} \psi\left(\frac{k}{K}\right) \tag{6.22}$$

as $K \to \infty$

Proof. From the definition of $\mathcal{N}_\mathcal{F}(\psi)$ and the Petersson formula, it follows that

$$\sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \Delta_k(1,1) = \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) + G(1,1)$$

where $G(m,n)$ is given by (6.13). If we take $g(x) = \psi\left(\frac{x}{K}\right)$, then, proceeding in the same fashion as before, we obtain

$$c_3(g) \ll xK^{-3}$$

and

$$h(x) \ll (xK^{-2})^j \quad \text{for } j = 0, 1, 2, \ldots$$

and from this estimate it follows that

$$G(1,1) \ll K^{-2j} + K^{-3} \quad \text{for } j = 1, 2, \ldots$$

Hence, the lemma follows. □

Chapter 7

The Proof of the Mean Value Theorem

7.1 Averaging over the Family $\mathcal{F}$

Finally, all necessary tools have been developed to prove the mean value theorem. As the reader may recall, we will be evaluating the average over the family $\mathcal{F}$ of the integrals

$$I_f(\alpha,\beta) = \frac{1}{2T}\int_{-T}^{T}\left|M\left(\frac{1}{2}+it,f\right)\right|^2 L\left(\frac{1}{2}+\alpha+it,f\right) L\left(\frac{1}{2}+\beta-it,f\right)dt$$

with

$$|\alpha|,|\beta| \ll \left(\log\left(K^2\right)\right)^{-1}$$

and

$$|\alpha+\beta| \asymp \left(\log\left(K^2\right)\right)^{-1}$$

as $K\to\infty$. However, since

$$T = o(K)$$

we are going to postpone the integration in t due to its secondary role and we are going to concentrate temporarily on the k aspect. Let

$$J_f = \left|M\left(\frac{1}{2}+it,f\right)\right|^2 L\left(s_1,f\right) L\left(s_2,f\right)$$

with s_1 and s_2 equal to $\frac{1}{2}+\alpha+it$ and $\frac{1}{2}+\beta+it$, respectively. Then, we will evaluate the following expression

$$\sum_k \psi\left(\frac{k}{K}\right)\sum_{f\in H_k(1)}^{h} J_f$$

uniformly for all $|t|\le T$ as $K\to\infty$.

7.1.1 Applying the Functional Equation

The first step would be to use the approximate functional equation (5.5) so that we can replace $L(s_1, f)\, L(s_2, f)$ with Dirichlet series. Thus, we have

$$\sum_k \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} J_f = \sum_k \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} J_f(s_1, s_2)$$
$$+ \sum_k \psi\left(\frac{k}{K}\right) \Theta\left(\frac{k}{K}, s_1, s_2\right) \sum_{f \in H_k(1)}^{h} J_f(1-s_1, 1-s_2)$$

where

$$J_f(s_1, s_2) = \left| M\left(\frac{1}{2} + it, f\right)\right|^2 D_f(s_1, s_2).$$

Then, we will focus on evaluating

$$\sum_k \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} J_f(s_1, s_2) \tag{7.1}$$

as $K \to \infty$. The treatment of

$$\sum_k \psi\left(\frac{k}{K}\right) \Theta\left(\frac{k}{K}, s_1, s_2\right) \sum_{f \in H_k(1)}^{h} J_f(1-s_1, 1-s2)$$

will be essentially the same since the smoothness, compact support, and decay properties satisfied by $\psi(y)$ are also satisfied by $\psi(y)\Theta(y, s_1, s_2)$ because of Lemma 5.2.3. Before stating a lemma that evaluates (7.1), we give some definitions. Let the function $W(u; x, y)$ be defined by

$$W(u; x, y) = \frac{1}{x+y}\big(P'(u) - xP(u)\big)\big(P'(u) - yP(u)\big) \tag{7.2}$$

where $P(u)$ is the function used in the definition of the mollifier. Also, let $\Delta(n, t)$ be the Dirichlet polynomial

$$\Delta(n, t) = \sum_{(m,n)=1} \frac{\mu(m)}{m} g(m)\omega(m) m^{2it} \tag{7.3}$$

where the coefficients $g(m)$ are given by the convolution

$$g(m) = \sum_{\substack{m=rd \\ r,d \le \Delta}} \mu(r) \tag{7.4}$$

and

$$\omega(m) = \prod_{p|m} \left(1 - p^{-1}\right)^{-1}.$$

Finally, consider the expression

$$J(t; x, y) = \sum\sum_{(n,h)=1} \frac{\mu(h)}{h^2} \frac{\mu^2(n)}{n} \omega^2(nh) |\Delta(nh, t)|^2 W(\gamma_n; x, y). \tag{7.5}$$

Now, we are ready to state the following lemma

Lemma 7.1.1. *Let $s_1 = \frac{1}{2} + \alpha + it$ and $s_2 = \frac{1}{2} + \beta + it$. If $\theta = \frac{1}{2}$ and*

$$T = o(K)$$

as $K \to \infty$, then, we have

$$\sum_k \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} J_f(s_1, s_2) \sim \frac{J(t; \alpha \log\left(K^2\right), \beta \log\left(K^2\right))}{\log\left(K^2\right)} \sum_k \psi\left(\frac{k}{K}\right) \tag{7.6}$$

as $K \to \infty$, uniformly for all $|t| \leq T$ and all complex numbers α and β satisfying

$$|\alpha|, |\beta| \ll \left(\log\left(K^2\right)\right)^{-1}$$

and

$$|\alpha + \beta| \asymp \left(\log\left(K^2\right)\right)^{-1}$$

with Δ in the definition (7.4) given by

$$\Delta = \Delta(K) = \exp\left(\log\log K\right)^3.$$

The prove of the lemma will be given in the rest of the section, and the next one.

7.1.2 Applying the Asymptotic Large Sieve

The first step in the proof of Lemma 7.1.1 consists of applying the asymptotic large sieve. Thus, we begin by opening the sums in $J_f(s_1, s_2)$. The square of the absolute value of the mollifier is

$$\left|M\left(\frac{1}{2} + it, f\right)\right|^2 = \sum\sum \frac{c_f(m_1)}{\sqrt{m_1}} \frac{c_f(m_2)}{\sqrt{m_2}} \left(\frac{m_2}{m_1}\right)^{it}$$

with $c_f(m)$ defined by the expression (3.12)

$$c_f(m) = \sum_{\substack{m=nr^2 \\ r\le\Delta}} \lambda_f(n) P\left(\gamma_m\right) \mu(nr)\mu(r).$$

Then, we combine the sum above with the double series

$$D_f(s_1, s_2) = \sum_{l_1}\sum_{l_2} \lambda_f(l_1)\lambda_f(l_2) l_1^{-s_1} l_2^{-s_2} H\left(\frac{k}{K}, \frac{l_1 l_2}{K^2}\right).$$

For this, we use the multiplication formula

$$\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \lambda_f\left(\frac{mn}{d^2}\right)$$

and make a change of variables that replaces n and l with nd and ld, repectively. Hence, we obtain the following formula for expression (7.1)

$$\sum_{r_1\le\Delta}\sum_{d_1}\sum_{r_2\le\Delta}\sum_{d_2} \frac{\mu(r_1)}{r_1 d_1}\frac{\mu(r_2)}{r_2 d_2}\left(\frac{r_2 d_2}{r_1 d_1}\right)^{2it} \mathcal{S}(\mathcal{A}\times\mathcal{B})$$

where $\mathcal{S}(\mathcal{A}\times\mathcal{B})$ is the bilinear form

$$\mathcal{S}(\mathcal{A}\times\mathcal{B}) = \sum_{n_1}\sum_{l_1}\sum_{n_2}\sum_{l_2} \alpha_{n_1 l_1}\beta_{n_2 l_2} \sum_{k \text{ even}} F\left(\frac{k}{K}, \frac{l_1 l_2 d_1 d_2}{K^2}\right) \Delta_k(n_1 l_1, n_2 l_2)$$

with the coefficients α_{n_1,l_1} and β_{n_2,l_2} defined by the following expressions

$$\alpha_{n_1,l_1} = \frac{\mu(n_1 d_1 r_1)}{\sqrt{n_1 l_1}} P(\gamma_{n_1 d_1})(d_1 l_1)^{-\alpha}(n_1 l_1)^{-it} \tag{7.7}$$

and

$$\beta_{n_2,l_2} = \frac{\mu(n_2 d_2 r_2)}{\sqrt{n_2 l_2}} P(\gamma_{n_2 d_2})(d_2 l_2)^{-\beta}(n_2 l_2)^{it}. \tag{7.8}$$

Here, the test function $F(y, Y)$ is given by

$$F(y, Y) = \psi(y) H(y, Y)$$

Also, by the definition of the mollifier $M(s, f)$ and more specifically since $P(x) \equiv 0$ for $x \in (\theta - \delta, 1)$ with $\theta = \frac{1}{2}$, we have that n_1, n_2, d_1, and d_2 all have support $\le K^{1-2\delta}$. This implies that

$$\frac{K^2}{d_1 d_2} \ge K^{4\delta}$$

Hence, we can apply the Asymptotic Large Sieve theorem with $N \leq K^{1-2\delta}$ and $L = \frac{K^2}{d_1 d_2}$ obtaining

$$\mathcal{S}(\mathcal{A} \times \mathcal{B}) = \mathcal{S}_{\text{diag}}(\mathcal{A} \times \mathcal{B}) + \mathcal{R}(\mathcal{A} \times \mathcal{B})$$

with

$$\mathcal{R}(\mathcal{A} \times \mathcal{B}) \ll_{\phi} K^{\phi}$$

as $K \to \infty$ for some $0 < \phi < 1$. From this, it follows by trivial estimation that

$$\sum_{r_1 \leq \Delta} \sum_{d_1 \leq K} \sum_{r_2 \leq \Delta} \sum_{d_2 \leq K} \frac{\mu(r_1)}{r_1 d_1} \frac{\mu(r_2)}{r_2 d_2} \left(\frac{r_2 d_2}{r_1 d_1} \right)^{2it} \mathcal{R}(\mathcal{A} \times \mathcal{B}) \ll_{\phi_1} K^{\phi_1} \tag{7.9}$$

as $K \to \infty$ for some $0 < \phi_1 < 1$.

7.1.3 Removing the Kronecker Symbol

Having dealt with the remainder terms in the previous section, we switch our focus to evaluating the "diagonal" terms which we denote by $\mathcal{D}(\mathcal{A} \times \mathcal{B})$ and are given by

$$\mathcal{D}(\mathcal{A} \times \mathcal{B}) = \sum_{r_1 \leq \Delta} \sum_{d_1} \sum_{r_2 \leq \Delta} \sum_{d_2} \frac{\mu(r_1)}{r_1 d_1} \frac{\mu(r_2)}{r_2 d_2} \left(\frac{r_2 d_2}{r_1 d_1} \right)^{2it} \mathcal{S}_{\text{diag}}(\mathcal{A} \times \mathcal{B}).$$

Also, since

$$\sum_{k} \psi \left(\frac{k}{K} \right) \sim \hat{\psi}(0) K$$

it follows from formula (7.9) that

$$\sum_{k} \psi \left(\frac{k}{K} \right) \sum_{f \in H_k(1)}^{h} J_f(s_1, s_2) \sim \mathcal{D}(\mathcal{A} \times \mathcal{B}).$$

Now, it is necessary to evaluate $\mathcal{D}(\mathcal{A} \times \mathcal{B})$ asymptotically. For that purpose, we decompose the test function back into its components $\psi(y)$ and $H(y, Y)$

$$F(y, Y) = \psi(y) H(y, Y).$$

Then, we give the following expression for $\mathcal{D}(\mathcal{A} \times \mathcal{B})$

$$\mathcal{D}(\mathcal{A} \times \mathcal{B}) = \mathcal{D} \sum_{k \text{ even}} \psi \left(\frac{k}{K} \right)$$

where

$$\mathcal{D} = \sum_{r_1 \leq \Delta} \sum_{r_2 \leq \Delta} \sum_{d_1} \sum_{d_2} \frac{\mu(r_1)}{r_1 d_1} \frac{\mu(r_2)}{r_2 d_2} \left(\frac{r_2 d_2}{r_1 d_1}\right)^{2it}$$
$$\times \sum_{n_1} \sum_{l_1} \sum_{n_2} \sum_{l_2}{}' \alpha_{m_1 l_1} \beta_{m_2 l_2} H\left(\frac{k}{K}, \frac{l_1 l_2 d_1 d_2}{K^2}\right)$$

with the notation $\sum\sum\sum\sum'$ indicating that the variables n_1, n_2, l_1, and l_2 are subject to the restriction

$$n_1 l_1 = n_2 l_2$$

which is imposed by the Kronecker symbol. We will remove this restriction by solving the above arithmetic equation. The solution is given by $n_1 = na_1$, $n_2 = na_2$, $l_1 = la_2$, and $l_2 = la_1$ where a_1 and a_2 satisfy $(a_1, a_2) = 1$ and with the variables n and l completely free. Thus, by the definition of the coefficients α_{n_1,l_1} and β_{n_2,l_2} given in (7.7) and (7.8), respectively, we have

$$\alpha_{na_1,la_2} = \frac{\mu(na_1 d_1 r_1)}{\sqrt{nla_1a_2}} P(\gamma_{na_1d_1})(d_1 a_2 l)^{-\alpha}(nla_1a_2)^{-it}$$

and

$$\beta_{na_2,la_1} = \frac{\mu(na_2 d_2 r_2)}{\sqrt{nla_2a_1}} P(\gamma_{na_2d_2})(d_2 a_1 l)^{-\beta}(nla_2a_1)^{it}$$

. Hence, we can write $\mathcal{D}$ as

$$\mathcal{D} = \sum_n \Omega(n,t) n^{-1} \tag{7.10}$$

with the coefficient $\Omega(n,t)$ is given by the following sums

$$\Omega(n,t) = \sum_{r_1 \leq \Delta} \sum_{r_2 \leq \Delta} \frac{\mu(r_1)}{r_1} \frac{\mu(r_2)}{r_2} \left(\frac{r_2}{r_1}\right)^{2it} \mathcal{A}(r_1, r_2; t) \tag{7.11}$$

where $\mathcal{A}(r_1, r_2; t)$ is defined by

$$\begin{aligned} \mathcal{A}(r_1, r_2; t) = \sum_{d_1} \sum_{d_2} \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} & \frac{\mu(na_1 d_1 r_1)}{a_1 d_1} \frac{\mu(na_2 d_2 r_2)}{a_2 d_2} \\ & \times \left(\frac{d_2}{d_1}\right)^{2it} Z(a_1, a_2, d_1, d_2) \end{aligned} \tag{7.12}$$

and the function $Z(x_1, x_2, y_1, y_2)$ is given by the expression

$$Z(x_1, x_2, y_1, y_2) = (x_2 y_1)^{-\alpha}(x_1 y_2)^{-\beta} P(\gamma_{nx_1y_1}) P(\gamma_{nx_2y_2}) Z\left(\frac{k}{K}, \frac{x_1 x_2 y_1 y_2}{K^2}\right) \tag{7.13}$$

where

$$Z\left(y,Y\right)=\sum_{l}H\left(\frac{k}{K},l^2Y\right)l^{-1-\alpha-\beta}.$$

Because of the definition of $H(y,Y)$ in (5.2), we also have

$$Z\left(y,Y\right)=\frac{1}{2\pi i}\int_{(c)}\zeta(1+\alpha+\beta+2z)\frac{\gamma(y,s_1+z)}{\gamma(y,s_1)}\frac{\gamma(y,s_2+z)}{\gamma(y,s_2)}(K^2Y)^{-z}\frac{w(z)}{z}dz.$$

7.1.4 Evaluating $\mathcal{D}$

We would like to replace $\mathcal{D}$ with a simpler expression. Thus, we give the following general lemma about the function $Z(y,Y)$ defined by

$$Z\left(y,Y\right)=\frac{1}{2\pi i}\int_{(c)}\zeta(s_1+s_2+2z)\frac{\gamma(y,s_1+z)}{\gamma(y,s_1)}\frac{\gamma(y,s_2+z)}{\gamma(y,s_2)}(K^2Y)^{-z}\frac{w(z)}{z}dz. \quad (7.14)$$

From this formula, we notice that $w(z)$ was defined in (5.4) with a zero at $z=\frac{1-s_1-s_2}{2}$ so that it would kill the simple pole of $\zeta(s_1+s_2+2z)$.

Lemma 7.1.2. *For any $a>1$ and all $y\in(a^{-1},a)$, $Z(y,Y)$ is a smooth function of Y over the positive reals which, for any $c>0$, satisfies*

$$\frac{\partial^l Z}{\partial Y^l}(y,Y)=\begin{cases}O\left(Y^{-c}(\log K)^2\right) & \text{if} \quad Y\geq 1\\ \zeta(s_1+s_2)+O\left(Y^{c}(\log K)^2\right) & \text{if} \quad Y<1\end{cases} \quad (7.15)$$

as $K\to\infty$, uniformly for all $y\in(a^{-1},a)$ and all complex numbers s_1 and s_2 satisfying

$$-B<\Re s_i<B \qquad \text{for } i=1,2 \quad (7.16)$$

and

$$\Im s_i=o(K) \qquad \text{for } i=1,2 \quad (7.17)$$

with the implicit constant depending on a, c, B, and l.

Proof. We already established in the proof of Lemma 5.2.2 that

$$\frac{\gamma(y,s_1+z)}{\gamma(y,s_1)}\frac{\gamma(y,s_2+z)}{\gamma(y,s_2)}\ll\exp\left\{2c\log K+O\left(|z|+(|z|^2+1)K^{-1}\right)\right\}$$

as $K\to\infty$, uniformly for all $y\in(a^{-1},a)$ and all complex numbers s_1 and s_2 satisfying conditions (7.16) and (7.17). Also, we know from the theory of the Riemann zeta function that $\zeta(s_1+s_2+2z)$ is polynomially bounded on $\Re z = c \ll 1$ with the implicit constant depending on B and c. Then, the complex integral (7.14) that defines $Z(y,Y)$ is absolutely convergent for any c. Thus,

$$\frac{\partial^l Z}{\partial Y^l}(y,Y) = \frac{1}{2\pi i}\int_{(c)} \zeta(s_1+s_2+2z)\frac{\gamma(y,s_1+z)}{\gamma(y,s_1)}\frac{\gamma(y,s_2+z)}{\gamma(y,s_2)}(K^2Y)^{-z}p(-z)\frac{w(z)}{z}dz$$

with $p(-z) = -z(-z-1)\ldots(-z-l+1)$. By the same argument, the above integral is also absolutely convergent for any c. Starting with $c>0$, if $Y>1$, then the result follows by estimating the integral trivially. If $Y<1$, then the result follows moving the line of integration left to $\Re z = -c$, picking up the residue of the pole at $z=0$, and estimating the integral trivially. □

Now, we define $Z_0(x_1,x_2,y_1,y_2)$ by the expression

$$Z_0(x_1,x_2,y_1,y_2) = (x_2y_1)^{-\alpha}(x_1y_2)^{-\beta}P(\gamma_{nx_1y_1})P(\gamma_{nx_2y_2}) \tag{7.18}$$

Then, we define $\mathcal{A}_0(r_1,r_2;t)$ replacing $Z(x_1,x_2,y_1,y_2)$ with $Z_0(x_1,x_2,y_1,y_2)$ in equation (7.12). Thus, we have

$$\mathcal{A}_0(r_1,r_2;t) = \sum_{d_1}\sum_{d_2}\sum_{a_1}\sum_{\substack{a_2\\(a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1d_1}\frac{\mu(na_2d_2r_2)}{a_2d_2}\left(\frac{d_2}{d_1}\right)^{2it} Z_0(a_1,a_2,d_1,d_2). \tag{7.19}$$

We do the same in equations (7.11) and (7.10) obtaining

$$\Omega_0(n,t) = \sum_{r_1\le\Delta}\sum_{r_2\le\Delta}\frac{\mu(r_1)}{r_1}\frac{\mu(r_2)}{r_2}\left(\frac{r_2}{r_1}\right)^{2it}\mathcal{A}_0(r_1,r_2;t) \tag{7.20}$$

and

$$\mathcal{D}_0 = \sum_n \Omega_0(n,t)n^{-1}. \tag{7.21}$$

Then, with these definitions, we give the following corollary of Lemma 7.1.2.

Corollary 7.1.3. *Let $s_1 = \frac{1}{2}+\alpha+it$ and $s_2 = \frac{1}{2}+\beta+it$. If T=o(K) as $K\to\infty$, then for all $\eta>0$*

$$\mathcal{D} = \zeta(1+\alpha+\beta)\mathcal{D}_0 + O\left(K^{-\eta}\right) \tag{7.22}$$

as $K\to\infty$, uniformly for all $|t|\le T$ and complex numbers α and β satisfying

$$|\alpha|,|\beta| \ll \left(\log\left(K^2\right)\right)^{-1} \tag{7.23}$$

and

$$|\alpha+\beta| \asymp \left(\log\left(K^2\right)\right)^{-1}. \tag{7.24}$$

Proof. From the definition (7.13) of $Z(x_1, x_2, y_1, y_2)$

$$Z(x_1,x_2,y_1,y_2) = (x_2y_1)^{-\alpha}(x_1y_2)^{-\beta}P(\gamma_{nx_1y_1})P(\gamma_{nx_2y_2})Z\left(\frac{k}{K}, \frac{x_1x_2y_1y_2}{K^2}\right)$$

we notice that x_1y_1 and x_2y_2 have both support less than $K^{1-2\delta}$ because of the definition of the function $P(u)$. Thus, by Lemma 7.1.2.,

$$\frac{x_1x_2y_1y_2}{K^2} \le K^{-4\delta} < 1$$

implies that for any $c > 0$

$$Z(x_1,x_2,y_1,y_2) = \zeta(1+\alpha+\beta)Z_0(x_1,x_2,y_1,y_2) + O\left(K^{-4\delta c}(\log K)^2\right)$$

as $K \to \infty$, uniformly for all $|t| \le T$ and all complex numbers α and β satisfying (7.23) and (7.24). Then, we have

$$\mathcal{A}(r_1,r_2;t) = \zeta(1+\alpha+\beta)\mathcal{A}_0(r_1,r_2;t) + O\left(K^{-4\delta c}(\log K)^3\right)$$

which implies

$$\Omega(n,t) = \zeta(1+\alpha+\beta)\Omega_0(n,t) + O\left(K^{-4\delta c}(\log K)^5\right)$$

and

$$\mathcal{D} = \zeta(1+\alpha+\beta)\mathcal{D}_0 + O\left(K^{-4\delta c}(\log K)^6\right)$$

as $K \to \infty$, uniformly for all $|t| \le T$ and all complex numbers α and β satisfying (7.23) and (7.24). Hence, the lemma follows choosing $c = \eta(4\delta)^{-1}$ □

From (7.22), it follows that

$$\mathcal{D}(\mathcal{A}\times\mathcal{B}) \sim \zeta(1+\alpha+\beta)\mathcal{D}_0 \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right). \tag{7.25}$$

7.2 Möbius cancellation

In this step of the proof of Lemma 7.1.1., we dramatically reduce the range size of the variables a_1, a_2, d_1, and d_2 in $\mathcal{A}_0(r_1, r_2; t)$. The idea is drawn from the approach taken in [7] when dealing with the same problem but in the case of the Riemann zeta function. One thing that these four variables have in common is that they determine Dirichlet polynomials with the Möbius function appearing in the coefficients. This produces cancellation in the sums due to sign change of the Möbius function which is a result that follows from the Prime Number Thorem. More precisely, we have that if

$$\Delta = \Delta(K) = \exp\left(\log\log K\right)^3 \tag{7.26}$$

then, for all $N > 0$

$$\sum_{\Delta < n \leq X} \frac{\mu(n)}{n} \ll \exp\left(-\sqrt{\log \Delta}\right) \ll_N (\log K)^{-N} \tag{7.27}$$

for any $X \geq \Delta$ as $K \to \infty$. Perhaps now, the reader might not find so strange that we chose Δ as the range of the variable r in the definition of the mollifier. More clarity on that will be provided later. Since our Möbius functions are attached to $Z_0(a_1, a_2, d_1, d_2)$, we combine the Möbius cancellation property with partial summation to truncate the sums in a_1, a_2, d_1, and d_2. Hence, we have

$$\mathcal{A}_0(r_1, r_2; t) = \sum_{d_1 \leq \Delta} \sum_{d_2 \leq \Delta} \sum_{a_1 \leq \Delta} \sum_{\substack{a_2 \leq \Delta \\ (a_1, a_2) = 1}} \frac{\mu(na_1d_1r_1)}{a_1d_1} \frac{\mu(na_2d_2r_2)}{a_2d_2} \left(\frac{d_2}{d_1}\right)^{2it} Z_0(a_1, a_2, d_1, d_2)$$
$$+ O\left((\log K)^{-N_0}\right)$$

for a very large N_0 as $K \to \infty$. Define $\mathcal{A}_1(r_1, r_2; t)$ as $\mathcal{A}_0(r_1, r_2; t)$ with the variables a_1, a_2, d_1, and d_2 truncated at Δ,

$$\mathcal{A}_1(r_1, r_2; t) = \sum_{d_1 \leq \Delta} \sum_{d_2 \leq \Delta} \sum_{a_1 \leq \Delta} \sum_{\substack{a_2 \leq \Delta \\ (a_1, a_2) = 1}} \frac{\mu(na_1d_1r_1)}{a_1d_1} \frac{\mu(na_2d_2r_2)}{a_2d_2} \left(\frac{d_2}{d_1}\right)^{2it} \times Z_0(a_1, a_2, d_1, d_2). \tag{7.28}$$

Then, we have

$$\mathcal{A}_0(r_1, r_2; t) = \mathcal{A}_1(r_1, r_2; t) + O\left((\log K)^{-N_0}\right).$$

Now, we naturally define $\Omega_1(n,t)$ and $\mathcal{D}_1$ by replacing $\mathcal{A}_0(r_1, r_2; t)$ and $\Omega_0(n,t)$ with $\mathcal{A}_1(r_1, r_2; t)$ and $\Omega_1(n,t)$ in the definitions (7.20) and (7.21) of $\Omega_0(n,t)$ and $\mathcal{D}_0$, respectively. Hence,

$$\Omega_1(n,t) = \sum_{r_1 \leq \Delta} \sum_{r_2 \leq \Delta} \frac{\mu(r_1)}{r_1} \frac{\mu(r_2)}{r_2} \left(\frac{r_2}{r_1}\right)^{2it} \mathcal{A}_1(r_1, r_2; t) \tag{7.29}$$

and

$$\mathcal{D}_1 = \sum_n \Omega_1(n,t) n^{-1}. \tag{7.30}$$

The reader may recall that the range of r_1 and r_2 is less than Δ and the range of n less than $K^{1-2\delta}$ so we only lose less than $(\log K)^{1+\epsilon}$ in order when estimating trivially the sums over those variables. Hence, we obtain trivially

$$\Omega_0(n,t) = \Omega_1(n,t) + O\left((\log K)^{-N_1}\right)$$

and

$$\mathcal{D}_0 = \mathcal{D}_1 + O\left((\log K)^{-N_1}\right)$$

for large N_1. Finally, this implies that the asymptotic formula (7.25) remains true with $\mathcal{D}_0$ replaced by $\mathcal{D}_1$,

$$\mathcal{D}(\mathcal{A} \times \mathcal{B}) \sim \zeta(1+\alpha+\beta)\mathcal{D}_1 \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right). \tag{7.31}$$

7.2.1 Taylor Approximation

A first application of truncating the variables a_1, a_2, d_1, and d_2 will be to free the variables a_1 and a_2 so that they can be extended to an unrestricted sum by reversing our earlier application of Möbius cancellation. This way the sums in a_1 and a_2 can be evaluated using zeta function theory. To accomplish this, we begin rewriting $Z_0(a_1, a_2, d_1, d_2)$ so all variables appear in the logarithmic scale. Consider the change of variables

$$x = \alpha \log\left(K^2\right) \quad \text{and} \quad y = \beta \log\left(K^2\right).$$

Then,

$$\begin{aligned} Z_0(x_1, x_2, y_1, y_2) &= (x_2 y_1)^{-\alpha} (x_1 y_2)^{-\beta} P(\gamma_{n x_1 y_1}) P(\gamma_{n x_2 y_2}) \\ &= P(\gamma_{n x_1 y_1}) \exp\{-x \gamma_{x_2 y_1}\} P(\gamma_{n x_2 y_2}) \exp\{-y \gamma_{x_1 y_2}\} \\ &= Q_x\left(\gamma_n + \gamma_{x_1 y_1}, \gamma_{x_2 y_1}\right) Q_y\left(\gamma_n + \gamma_{x_2 y_2}, \gamma_{x_1 y_2}\right) \end{aligned}$$

where

$$Q_\eta(u,v) = P(u)e^{-\eta v}$$

Let $\gamma_{11} = \gamma_{x_1y_1}$, $\gamma_{21} = \gamma_{x_2y_1}$, $\gamma_{22} = \gamma_{x_2y_2}$, and $\gamma_{12} = \gamma_{x_1y_2}$. Then,

$$\begin{aligned} \gamma_{11}\ ,\ \gamma_{21}\ ,\ \gamma_{22}\ ,\ \gamma_{11} &\le \frac{\log \Delta}{\log (K^2)} \\ &\ll_\epsilon (\log K)^{-1+\epsilon} \end{aligned}$$

with Δ as in (7.26). Since $Q_\eta(u,v)$ is smooth, we can use a Taylor approximation to give

$$\begin{aligned} Q_x(\gamma_n + \gamma_{11}, \gamma_{21}) = Q_x &+ \frac{\partial Q_x}{\partial u}\gamma_{11} + \frac{\partial Q_x}{\partial v}\gamma_{21} \\ &+ \frac{\partial^2 Q_x}{\partial u^2}\gamma_{11}^2 + 2\frac{\partial^2 Q_x}{\partial u \partial v}\gamma_{11}\gamma_{21} + \frac{\partial^2 Q_x}{\partial v^2}\gamma_{21}^2 + (\text{"higher order terms"}) \end{aligned}$$

where

$$\begin{aligned} Q_x = P(\gamma_n) &\quad \frac{\partial Q_x}{\partial u} = P'(\gamma_n) \\ \frac{\partial Q_x}{\partial v} = -xP(\gamma_n) &\quad \frac{\partial^2 Q_x}{\partial u^2} = P''(\gamma_n) \\ \frac{\partial Q_x}{\partial u \partial v} = -xP'(\gamma_n) &\quad \frac{\partial^2 Q_x}{\partial v^2} = x^2 P(\gamma_n) \end{aligned}$$

Also, we have

$$(\text{"higher order terms"}) \ll_\epsilon (\log K)^{-3+\epsilon}.$$

The same equations hold for $Q_y(\gamma_n + \gamma_{22}, \gamma_{12})$ with x, γ_{11}, and γ_{21} replaced with y, γ_{22}, and γ_{12}, respectively. Hence, we have

$$Q_x(\gamma_n + \gamma_{x_1y_1}, \gamma_{x_2y_1})\, Q_y(\gamma_n + \gamma_{x_2y_2}, \gamma_{x_1y_2}) = T_0 + T_1 + T_2 + (\text{"higher order terms"})$$

$$\begin{aligned} T_0 &= Q_x Q_y \\ T_1 &= Q_x\frac{\partial Q_y}{\partial u}\gamma_{22} + Q_x\frac{\partial Q_y}{\partial v}\gamma_{12} + \frac{\partial Q_x}{\partial u}Q_y\gamma_{11} + \frac{\partial Q_x}{\partial v}Q_y\gamma_{21} \\ T_2 &= \frac{1}{2}Q_x\frac{\partial^2 Q_y}{\partial u^2}\gamma_{22}^2 + Q_x\frac{\partial^2 Q_y}{\partial u \partial v}\gamma_{22}\gamma_{12} + \frac{1}{2}Q_x\frac{\partial^2 Q_y}{\partial v^2}\gamma_{12}^2 \\ &+ \frac{1}{2}\frac{\partial^2 Q_x}{\partial u^2}Q_y\gamma_{11}^2 + \frac{\partial^2 Q_x}{\partial u \partial v}Q_y\gamma_{11}\gamma_{21} + \frac{1}{2}\frac{\partial^2 Q_x}{\partial v^2}Q_y\gamma_{21}^2 \\ &+ \frac{\partial Q_x}{\partial u}\frac{\partial Q_y}{\partial u}\gamma_{11}\gamma_{22} + \frac{\partial Q_x}{\partial v}\frac{\partial Q_y}{\partial v}\gamma_{21}\gamma_{12} + \frac{\partial Q_x}{\partial u}\frac{\partial Q_y}{\partial v}\gamma_{11}\gamma_{12} + \frac{\partial Q_x}{\partial v}\frac{\partial Q_y}{\partial u}\gamma_{21}\gamma_{22}. \end{aligned}$$

Now, define $\mathcal{A}_2(r_1, r_2; t)$ as $\mathcal{A}_1(r_1, r_2; t)$ with $Z_0(a_1, a_2, d_1, d_2)$ replaced by $T_0 + T_1 + T_2$,

$$\mathcal{A}_2(r_1, r_2; t) = \sum_{d_1 \le \Delta} \sum_{d_2 \le \Delta} \sum_{a_1 \le \Delta} \sum_{\substack{a_2 \le \Delta \\ (a_1, a_2) = 1}} \frac{\mu(na_1 d_1 r_1)}{a_1 d_1} \frac{\mu(na_2 d_2 r_2)}{a_2 d_2} \left(\frac{d_2}{d_1}\right)^{2it} \times (T_0 + T_1 + T_2). \tag{7.32}$$

Then, we have

$$\mathcal{A}_1(r_1, r_2; t) = \mathcal{A}_2(r_1, r_2; t) + O\left((\log K)^{-3+\epsilon}\right)$$

with the implicit constant depending on ϵ. Again, we naturally define $\Omega_2(n, t)$ and $\mathcal{D}_2$ by replacing $\mathcal{A}_1(r_1, r_2; t)$ and $\Omega_1(n, t)$ with $\mathcal{A}_2(r_1, r_2; t)$ and $\Omega_2(n, t)$ in the definitions (7.29) and (7.30) of $\Omega_1(n, t)$ and $\mathcal{D}_1$, respectively. Hence,

$$\Omega_2(n, t) = \sum_{r_1 \le \Delta} \sum_{r_2 \le \Delta} \frac{\mu(r_1)}{r_1} \frac{\mu(r_2)}{r_2} \left(\frac{r_2}{r_1}\right)^{2it} \mathcal{A}_2(r_1, r_2; t) \tag{7.33}$$

and

$$\mathcal{D}_2 = \sum_n \Omega_2(n, t) n^{-1}. \tag{7.34}$$

Hence, it follows that

$$\Omega_1(n, t) = \Omega_2(n, t) + O\left((\log K)^{-3+\epsilon}\right)$$

and

$$\mathcal{D}_1 = \mathcal{D}_2 + O\left((\log K)^{-2+\epsilon}\right).$$

Thus, since

$$\zeta(1 + \alpha + \beta) \asymp \log\left(K^2\right)$$

as $K \to \infty$, then (7.31) is again true with $\mathcal{D}_1$ replaced by $\mathcal{D}_2$,

$$\mathcal{D}(\mathcal{A} \times \mathcal{B}) \sim \zeta(1 + \alpha + \beta) \mathcal{D}_2 \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right). \tag{7.35}$$

7.2.2 Evaluating the Sums over the Variables a_1 and a_2

As indicated earlier, we would like to evaluate the sums over a_1 and a_2 by extending them to unrestricted sums. Thus, we define the following series

$$\Gamma = \sum_{a_1} \sum_{\substack{a_2 \\ (a_1, a_2) = 1}} \frac{\mu(na_1 d_1 r_1)}{a_1} \frac{\mu(na_2 d_2 r_2)}{a_2} \tag{7.36}$$

and

$$\Gamma_{11} = \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1} \frac{\mu(na_2d_2r_2)}{a_2} \gamma_{11}. \tag{7.37}$$

Then, we define the series Γ_{22}, Γ_{12}, and Γ_{21} by replacing γ_{11} with γ_{22}, γ_{12}, and γ_{21}, respectively in equation (7.37). Also, we define

$$\Gamma_{11,11} = \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1} \frac{\mu(na_2d_2r_2)}{a_2}. \tag{7.38}$$

Then, again we define series $\Gamma_{22,22}$, $\Gamma_{22,12}$, $\Gamma_{12,12}$, $\Gamma_{11,21}$, $\Gamma_{21,21}$, $\Gamma_{11,22}$, $\Gamma_{21,12}$, $\Gamma_{11,12}$, and $\Gamma_{21,22}$ replacing γ_{11}^2 with γ_{22}^2, $\gamma_{22}\gamma_{12}$, γ_{12}^2, $\gamma_{11}\gamma_{21}$, γ_{21}^2, $\gamma_{11}\gamma_{22}$, $\gamma_{21}\gamma_{12}$, $\gamma_{11}\gamma_{12}$, and $\gamma_{21}\gamma_{22}$, respectively in equation (7.38). Before we carry out the evalution of these series, we justify extending the sums in a_1 and a_2 variables to infinity. If we assume momentarily that the series defined above are actually finite sums with $a_1 \leq \Delta$ and $a_2 \leq \Delta$, then we can write

$$\Omega_2(n,t) = \sum_{r_1 \leq \Delta} \sum_{d_1 \leq \Delta} \sum_{r_2 \leq \Delta} \sum_{d_2 \leq \Delta} \frac{\mu(r_1)}{r_1d_1} \frac{\mu(r_2)}{r_2d_2} \left(\frac{r_2d_2}{r_1d_1} \right)^{2it} (\mathcal{T}_0 + \mathcal{T}_1 + \mathcal{T}_2)$$

where $\mathcal{T}_0$, $\mathcal{T}_1$, and $\mathcal{T}_2$ are given by the formulas

$$\begin{aligned} \mathcal{T}_0 &= \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1} \frac{\mu(na_2d_2r_2)}{a_2} T_0 \\ &= Q_xQ_y\Gamma, \end{aligned}$$

$$\begin{aligned} \mathcal{T}_1 &= \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1} \frac{\mu(na_2d_2r_2)}{a_2} T_1 \\ &= Q_x \frac{\partial Q_y}{\partial u} \Gamma_{22} + Q_x \frac{\partial Q_y}{\partial v} \Gamma_{12} + \frac{\partial Q_x}{\partial u} Q_y \Gamma_{11} + \frac{\partial Q_x}{\partial v} Q_y \Gamma_{21}, \end{aligned}$$

$$\begin{aligned} \mathcal{T}_2 &= \sum_{a_1} \sum_{\substack{a_2 \\ (a_1,a_2)=1}} \frac{\mu(na_1d_1r_1)}{a_1} \frac{\mu(na_2d_2r_2)}{a_2} T_2 \\ &= \frac{1}{2} Q_x \frac{\partial^2 Q_y}{\partial u^2} \Gamma_{22,22} + Q_x \frac{\partial^2 Q_y}{\partial u \partial v} \Gamma_{22,12} + \frac{1}{2} Q_x \frac{\partial^2 Q_y}{\partial v^2} \Gamma_{12,12} \\ &+ \frac{1}{2} \frac{\partial^2 Q_x}{\partial u^2} Q_y \Gamma_{11,11} + \frac{\partial^2 Q_x}{\partial u \partial v} Q_y \Gamma_{11,21} + \frac{1}{2} \frac{\partial^2 Q_x}{\partial v^2} Q_y \Gamma_{21,21} \\ &+ \frac{\partial Q_x}{\partial u} \frac{\partial Q_y}{\partial u} \Gamma_{11,22} + \frac{\partial Q_x}{\partial v} \frac{\partial Q_y}{\partial v} \Gamma_{21,12} + \frac{\partial Q_x}{\partial u} \frac{\partial Q_y}{\partial v} \Gamma_{11,12} + \frac{\partial Q_x}{\partial v} \frac{\partial Q_y}{\partial u} \Gamma_{21,22}. \end{aligned}$$

By Möbius cancellation, extending the sums in a_1 and a_2 to infinity happens at no cost since these tails will be bounded by $(\log K)^{N_2}$ with N_2 large. On the other hand, summing over variables r_1, r_2, d_1, d_2 only produces a loss of the order $(\log K)^{\epsilon}$ and summing over n produces a loss of $\log K$. To avoid introducing new notation, we redefine $\Omega_2(n,t)$ and $\mathcal{D}_2$ by considering the sums in a_1 and a_2 to be unrestricted while still preserving the validity of the asymptotic formula (7.35),

$$\mathcal{D}(\mathcal{A}\times\mathcal{B}) \sim \zeta(1+\alpha+\beta)\mathcal{D}_2 \sum_{k \text{ even}} \psi\left(\frac{k}{K}\right).$$

The following lemma and its corollary will help us evaluate the series that define $\mathcal{T}_0$, $\mathcal{T}_1$, and $\mathcal{T}_2$.

Lemma 7.2.1.

$$\sum_{(n,a)=1} \frac{\mu(n)}{n} = 0 \tag{7.39}$$

and

$$\sum_{(n,a)=1} \frac{\mu(n)}{n} \log n = -\omega(a) \tag{7.40}$$

where

$$\omega(a) = \prod_{p|a} \left(1-p^{-1}\right)^{-1} \tag{7.41}$$

Proof. The Prime Number Theorem establishes the following formula

$$\zeta^{-1}(s) \prod_{p|a} \left(1-p^{-s}\right)^{-1} = \sum_{(n,a)=1} \mu(n) n^{-s} \tag{7.42}$$

for all s in a small neighborhood of 1 within the zero free region. Thus, evaluating (7.42) at $s=1$, we have

$$\sum_{(n,a)=1} \frac{\mu(n)}{n} = 0.$$

Also, (7.41) follows when we take the derivative of equation (7.42) and evaluate it at $s=1$. □

From this last lemma, we deduce the following corollary

Corollary 7.2.2.

$$\sum_{(a_1,b)} \sum_{\substack{(a_2,c)=1 \\ (a_1,a_2)=1}} \frac{\mu(a_1)}{a_1} \frac{\mu(a_2)}{a_2} (\log a_2)^N = 0 \quad \textit{for} \quad N = 0, \dots \tag{7.43}$$

and

$$\sum_{(a_1,b)} \sum_{\substack{(a_2,c)=1 \\ (a_1,a_2)=1}} \frac{\mu(a_1)}{a_1} \frac{\mu(a_2)}{a_2} (\log a_1)(\log a_2) = \omega(b)\omega(c) \sum_{(h,bc)=1} \frac{\mu(h)}{h^2} \omega^2(h). \tag{7.44}$$

Proof. By Mobius inversion, the constrain $(a_1, a_2) = 1$ is eliminated. Thus, expressions (7.43) and (7.44) become

$$\sum_{(h,bc)=1} \frac{\mu(h)}{h^2} \left(\sum_{(a_1,bh)=1} \frac{\mu(a_1)}{a_1} \right) \left(\sum_{(a_2,ch)=1} \frac{\mu(a_2)}{a_2} (\log a_2 h)^N \right) \tag{7.45}$$

and

$$\sum_{(h,bc)=1} \frac{\mu(h)}{h^2} \left(\sum_{(a_1,bh)=1} \frac{\mu(a_1)}{a_1} \log a_1 h \right) \left(\sum_{(a_2,ch)=1} \frac{\mu(a_2)}{a_2} \log a_2 h \right) \tag{7.46}$$

respectively. Then, the corollary follows from the previous lemma. □

With the previous corollary, we immediately determine that all the series in the definition of $\mathcal{T}_0$, $\mathcal{T}_1$, and $\mathcal{T}_2$ vanish except for $\Gamma_{22,12}$, $\Gamma_{11,21}$, $\Gamma_{11,22}$, and $\Gamma_{21,12}$. These non-vanishing four series satisfy

$$\Gamma_{22,12} = \Gamma_{11,21} = \Gamma_{11,22} = \Gamma_{21,12} = \mathcal{G}(nr_1d_1, nr_2d_2)$$

with $\mathcal{G}(b,c)$ defined by

$$\mathcal{G}(b,c) = \left(\log\left(K^2\right)\right)^{-2} \mu(b)\mu(c)\omega(b)\omega(c) \sum_{(h,bc)=1} \frac{\mu(h)}{h^2} \omega^2(h).$$

Thus, we have

$$\begin{aligned} \mathcal{T}_0 + \mathcal{T}_1 + \mathcal{T}_2 &= \mathcal{G}(nr_1d_1, nr_2d_2) \left(Q_x \frac{\partial^2 Q_y}{\partial u \partial v} + \frac{\partial^2 Q_x}{\partial u \partial v} Q_y + \frac{\partial Q_x}{\partial u} \frac{\partial Q_y}{\partial u} + \frac{\partial Q_x}{\partial v} \frac{\partial Q_y}{\partial v} \right) \\ &= \mathcal{G}(nr_1d_1, nr_2d_2) \left(P'(\gamma_n) - xP(\gamma_n)\right) \left(P'(\gamma_n) - yP(\gamma_n)\right). \end{aligned}$$

Since

$$x = \alpha \log\left(K^2\right) \quad \text{and} \quad y = \beta \log\left(K^2\right)$$

we have

$$\Omega_2(n,t) = \left(\log\left(K^2\right)\right)^{-1}(\alpha+\beta)\mu^2(n)\sum_{(n,h)=1}\frac{\mu(h)}{h^2}\omega^2(hn)\left|\Delta(n,t)\right|^2 W(\gamma_n;x,y)$$

and

$$\mathcal{D}_2 = \left(\log\left(K^2\right)\right)^{-1}(\alpha+\beta)\sum\sum_{(n,h)=1}\frac{\mu(h)}{h^2}\frac{\mu^2(n)}{n}\omega^2(nh)|\Delta(nh,t)|^2 W(\gamma_n;x,y)$$

where $\Delta(n,t)$ and $W(\gamma_n;x,y)$ are given by (7.3) and (7.2), respectively. Hence, we have

$$\mathcal{D}_2 = \left(\log\left(K^2\right)\right)^{-1}(\alpha+\beta)J(t;x,y).$$

The last equation and (7.35) imply

$$\mathcal{D}(\mathcal{A}\times\mathcal{B}) \sim \left(\log\left(K^2\right)\right)^{-1}\zeta(1+\alpha+\beta)(\alpha+\beta)J(t;x,y)\sum_{k \text{ even}}\psi\left(\frac{k}{K}\right)$$

as $K\to\infty$, uniformly for all $|t|\leq T$ and all complex numbers α and β satisfying

$$|\alpha|,|\beta| \ll \left(\log\left(K^2\right)\right)^{-1}$$

and

$$|\alpha+\beta| \asymp \left(\log\left(K^2\right)\right)^{-1}.$$

Then, Lemma 7.1.1 follows from

$$\zeta(1+\alpha+\beta)(\alpha+\beta)\sim 1$$

as $K\to\infty$.

7.3 Small Averaging in the t Aspect

In the last two sections, we carried out the averaging in the k aspect. As argued before, Lemma 7.1.1 can also be applied to evaluate the dual expression

$$\sum_{k \text{ even}}\psi\left(\frac{k}{K}\right)\Theta\left(\frac{k}{K},s_1,s_2\right)\sum_{f\in H_k(1)}^{h} J_f(1-s_1,1-s2).$$

In this case, $1-s_1=\frac{1}{2}-\alpha-it$ and $1-s_2=\frac{1}{2}-\beta+it$. Thus, if $T=o(K)$, then

$$\begin{aligned}\sum_{k \text{ even}}\psi\left(\frac{k}{K}\right)\Theta\left(\frac{k}{K},s_1,s_2\right)\sum_{f\in H_k(1)}^{h} J_f(1-s_1,1-s2)\\ \sim J(-t;-x,-y)\left(\log\left(K^2\right)\right)^{-1}\sum_{k \text{ even}}\psi\left(\frac{k}{K}\right)\Theta\left(\frac{k}{K},s_1,s_2\right)\end{aligned} \tag{7.47}$$

as $K \to \infty$, uniformly for all $|t| \leq T$ and all complex numbers x and y such that x,$y \ll 1$. Recall that x and y were defined as the following change of variables.

$$x = \alpha \log\left(K^2\right) \quad \text{and} \quad y = \beta \log\left(K^2\right).$$

By Lemma 5.2.3., if $T = o(K)$, it follows that

$$\begin{aligned}\Theta(y, s_1, s_2) &\sim K^{2(1-s1-s2)} \\ &= \exp\{-(x+y)\}\end{aligned}$$

as $K \to \infty$, uniformly for all $k \asymp K$ and for all $|t| \leq T$ and all complex numbers x and y such that x,$y \ll 1$. This implies that

$$\sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \Theta\left(\frac{k}{K}, s_1, s_2\right) \sim \exp\{-(x+y)\} \sum_{k} \psi\left(\frac{k}{K}\right)$$

as $K \to \infty$, uniformly for all $|t| \leq T = o(K)$ and all complex numbers x and y such that $|x|$,$|y| \asymp 1$. Hence, we obtain

$$\sum_{k \text{ even}} \psi\left(\frac{k}{K}\right) \sum_{f \in H_k(1)}^{h} I_f(\alpha, \beta) \sim \mathcal{J}(K) \sum_{k} \psi\left(\frac{k}{K}\right)$$

as $K \to \infty$, for all complex numbers x and y such that $|x|$,$|y| \asymp 1$ where

$$\mathcal{J}(K) = \left(\log\left(K^2\right)\right)^{-1} \frac{1}{2T} \int_{-T}^{T} \left(J(t; x, y) + \exp\{-(x+y)\} J(-t; -x, -y)\right) dt. \tag{7.48}$$

7.3.1 The Gap Principle

The next step in the proof of the Mean Value Theorem, Theorem 4.1.1, requires evaluating $\mathcal{J}(K)$ asymptotically as $K \to \infty$. This involves a small averaging over the t aspect that was inherited from Levinson's method. From the definition (7.5) of $J(t, x, y)$, we have that

$$\frac{1}{2T} \int_{-T}^{T} J(t; x, y) dt = \sum \sum_{(n,h)=1} \frac{\mu(h)}{h^2} \frac{\mu^2(n)}{n} \omega^2(nh) \frac{1}{2T} \int_{-T}^{T} |\Delta(nh, t)|^2 W(\gamma_n; x, y).$$

The following lemma will be an application of the mean value theorem for general Dirichlet polynomials [8], chapter 9. We will be using it on $\Delta(m, t)$ so when $\Delta \geq T$, the application is direct and trivial. However, if we want the widest possible range for T in our main result (2.35), Theorem 2.3.1, then we have to contemplate the case when

$T \leq \Delta$ which would make a trivial application of the mean value theorem for general Dirichlet polynomials impossible. Fortunately, we can take advantage of the precise shape of the coefficients $g(m)$ which have a gap due to Möbius cancellation. As we will show in the next lemma, the existence of this gap is key to evaluating the mean value of $\Delta(m, t)$. We hope the reader can now fully grasp the reason why we did not get rid of the variable r completely when we were defining the coefficients of the mollifier (3.12) earlier in the third chapter.

lemma 7.3.1. *Let $\Delta(t)$ be the following Dirichlet polynomial*

$$\Delta(t) = \sum_m g(m) \frac{a_m}{m} m^{it}$$

with

$$g(m) = \sum_{\substack{m=rd \\ r,d \leq \Delta}} \mu(r),$$

$$a_1 = 1 \text{ and } a_m \ll_B (\log m)^B$$

for some $B > 0$. If

$$(\log \Delta)^{2B+5} < T \leq \Delta$$

then

$$\frac{1}{2T} \int_{-T}^{T} |\Delta(n, t)|^2 dt = 1 + O\left(T^{-\frac{1}{2}} (\log \Delta)^{B+\frac{5}{2}}\right). \tag{7.49}$$

Proof. By Mobius cancellation,

$$g(m) = \begin{cases} 1 & \text{if} \quad m = 1. \\ 0 & \text{if} \quad 1 < m \leq \Delta. \end{cases}$$

Thus $m \neq 1$ is supported over the interval $[\Delta, \Delta^2]$. Then, this interval is divided into diadic intervals which is an standard technique to deal with shorter sums instead,

$$\Delta(t) = 1 + \sum_i \sum_{m \sim \Delta_i} \frac{b(m)}{m} m^{2it}$$

where the Δ_i's are defined by

$$\Delta_i = \Delta 2^i \quad \text{for} \quad i = 0, 1, 2, ...$$

and we define $b(m)$ by

$$b(m) = \begin{cases} g(m)a_m & \text{if} \quad m \leq \Delta^2 \\ 0 & \text{otherwise} \end{cases}$$

with $i \ll \log\Delta$. Since $T \ll \Delta$, the diadic intervals are further subdivided into subintervals of length T

$$\Delta(t) = 1 + \sum_i \sum_j S_{ij}(t)$$

where

$$S_{ij}(t) = \sum_{\Delta_{ij} < m \leq \Delta_{ij}+T} \frac{b(m)}{m} m^{it}$$

and $j \ll \Delta T^{-1}$. By the mean value theorem for Dirichlet Polynomials, we have

$$\begin{aligned} \frac{1}{2T}\int_{-T}^{T} |S_{ij}(t)|^2 dt &\ll \sum_{\Delta_{ij} < m \leq \Delta_{ij}+T} \frac{|b(m)|^2}{m^2} \\ &\ll \sum_{\Delta_{ij} < m \leq \Delta_{ij}+T} \tau^2(m) m^{-2} (\log m)^{2B} \\ &\ll T\Delta^{-2}(\log\Delta)^{2B+3}. \end{aligned}$$

Then, it follows that

$$\begin{aligned} \sum_i \sum_j \left(\frac{1}{T}\int_{-T}^{T} |S_{ij}(t)|^2 dt\right)^{\frac{1}{2}} &\ll \sum_i \sum_j T^{\frac{1}{2}} \Delta^{-1} (\log\Delta)^{B+\frac{3}{2}} \\ &\ll T^{-\frac{1}{2}} (\log\Delta)^{B+\frac{5}{2}}. \end{aligned}$$

Thus, by the Cauchy-Schwarz's inequality, we have

$$\frac{1}{2T}\int_{-T}^{T} |\Delta(t)|^2 dt = 1 + O\left(\sum_i \sum_j \left(\frac{1}{T}\int_{-T}^{T} |S_{ij}(t)|^2 dt\right)^{\frac{1}{2}}\right).$$

Hence, the lemma follows. □

Since $\Delta(n,t)$ satisfies the conditions of the previous lemma, then even when

$$T \leq \Delta = e^{(\log\log K)^3}$$

we would still have

$$\frac{1}{2T}\int_{-T}^{T} |\Delta(nh,t)|^2 \sim 1.$$

This implies that

$$\frac{1}{2T}\int_{-T}^{T} J(t;x,y) dt \sim \sum \sum_{(n,h)=1} \frac{\mu(h)}{h^2} \frac{\mu^2(n)}{n} \omega^2(nh) W(\gamma_n; x, y).$$

7.4 Evaluating the Series $\mathcal{J}(K)$

To evaluate $\mathcal{J}(K)$, the sum of the convergent series in h is first computed

$$\sum_{(n,h)=1} \frac{\mu(h)}{h^2}\omega^2(h) = \prod_{p|n}\left(1-\frac{\omega^2(p)}{p^2}\right)^{-1}\prod\left(1-\frac{\omega^2(p)}{p^2}\right)$$

where by convention when n even, the factors associated with 2 are not present in neither of the products. Thus

$$\sum \frac{\mu^2(n)}{n^s}\omega^2(n)\prod_{p|n}\left(1-\frac{\omega^2(p)}{p^2}\right)^{-1} = \prod\left(1-\frac{\omega^2(p)}{p^2}+\frac{\omega^2(p)}{p^s}\right)\prod\left(1-\frac{\omega^2(p)}{p^2}\right)^{-1}.$$

Then, since

$$\prod\left(1-\frac{\omega^2(p)}{p^2}+\frac{\omega^2(p)}{p^s}\right)\zeta^{-1}(s)\sim 1 \quad \text{when} \quad s\to 1$$

it follows that

$$\sum_{n\le N}\frac{\mu^2(n)}{n}\omega^2(n)\prod_{p|n}\left(1-\frac{\omega^2(p)}{p^2}\right)^{-1} \sim \log N\prod\left(1-\frac{\omega^2(p)}{p^2}\right)^{-1}.$$

Hence, by Abel's summation

$$\sum\sum_{(n,h)=1}\frac{\mu(h)}{h^2}\frac{\mu^2(n)}{n}\omega^2(nh)W(\gamma_n;x,y)\sim\left(\log\left(K^2\right)\right)\int_0^1 W(u;x,y)du.$$

Thus, we have obtained that

$$\mathcal{J}(K)\sim j\left(\alpha\log\left(K^2\right),\beta\log\left(K^2\right)\right)$$

where $j(x,y)$ is given by the formula

$$j(x,y)=J(x,y)+\exp\{-(x+y)\}J(-x,-y)$$

with $J(x,y)$ given by

$$J(x,y)=\int_0^1 W(u;x,y)du.$$

This culminates the proof of our mean value theorem (Theorem 4.1.1).

References

[1] H.M. Bui, J.B. Conrey, M.P. Young, *More than 41% of the zeros of the zeta function are on the critical line*, Acta Arithmetica, Vol: 150, Issue: 1 (2011), 35 - 64

[2] D. Bernard, *Modular case of Levinson's theorem*, Preprint 2014, http://arxiv.org/abs/1404.6432

[3] J.B. Conrey, *More than two fifths of the zeros of the Riemann zeta function are on the critical line*, J. Reine Angew, Math. 399 (1989), 1-26

[4] J.B. Conrey, H. Iwaniec, and K. Soundararajan, *Critical zeros of Dirichlet L-functions*, J. Reine Angew. Math. 681 (2013), 175198

[5] J.B. Conrey, H. Iwaniec, and K. Soundararajan, *Asymptotic large sieve*, Preprint 2011, http://arxiv.org/ abs/1105.1176.

[6] H. Iwaniec, *Topics in Classical Automorphic Forms*, Graduate Studies in Mathematics, Vol 17, AMS, Providence, RI, 1997

[7] H. Iwaniec, *Lectures on the Riemann zeta function*, Rutgers 2012, pedrohpontes.files.wordpress.com/2014/01/riemannzeta.pdf

[8] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications, vol 53, AMS, Providence, RI, 2004

[9] N. Levinson, *More than one third of zeros of the Riemann's zeta-function are on* $\sigma = 1/2$, Adv. in Math. 13 (1974), 383-436