LIMITING DISTRIBUTIONS AND ZEROS OF ARTIN L-FUNCTIONS

by

NATHAN CHRISTOPHER NG

B.Sc., University of British Columbia, 1994
M.Sc., University of Toronto, 1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Mathematics

We accept this thesis as conforming
to the required standard

. . . .

. . . .

. . . .

. . . .

. . . .

THE UNIVERSITY OF BRITISH COLUMBIA

November 2000

Abstract

This thesis is concerned with behaviour of some famous arithmetic functions. The first part of the thesis deals with prime number races. Rubinstein-Sarnak [62] developed a technique to study primes in arithmetic progressions. This thesis studies prime numbers that are described by Chebotarev's density theorem using the techniques developed in [62]. Let L/K be a normal extension with Galois group G. Consider conjugacy classes $C_1, \ldots, C_r$ and the set

$$P_{L/K;1,2,\ldots,r} = \{x \geq 2 \mid \frac{|G|}{|C_1|}\pi_{C_1}(x) \geq \frac{|G|}{|C_2|}\pi_{C_2}(x) \geq \cdots \geq \frac{|G|}{|C_r|}\pi_{C_r}(x)\} \ .$$

Following Rubinstein-Sarnak and by applying effective versions of Chebotarev's density theorem we do the following:

1. A limiting distribution $\mu_{L/K;1,2,\ldots,r}$ attached to the set $P_{L/K;1,2,\ldots,r}$ is constructed.

2. The Fourier transform of $\mu_{L/K;1,2,\ldots,r}$ is calculated. It can be expressed as an infinite product of $J_0(x)$ Bessel functions evaluated at zeros of the corresponding Artin L-functions.

3. Logarithmic densities of some specific examples of the sets $P_{L/K;1,2,\ldots,r}$ are computed. This computation requires many zeros of Artin L-functions. Some of these were computed using programs written in Fortran and C. Others were provided by Robert Rumely.

4. An explanation of Chebyshev's bias in the Galois group setting is given. In addition, the algebraic bias coming from possible zeros of an L-function at the centre of the critical is considered. Two examples of quaternion Galois groups were studied. The Dedekind zeta function of one of these fields has a zero at $s = \frac{1}{2}$.

5. The analogous problems for class groups are also considered. A simple explanation of when a bias occurs in two-way races for complex quadratic fields is presented. We also compute some logarithmic densities and derive a central limit theorem in this setting.

The second part of the thesis studies the summatory function of the Möbius function

$$M(x) = \sum_{n \leq x} \mu(n) \ .$$

Assuming conjectures due to Gonek and Hejhal concerning the reciprocal of the zeta function, the following results are shown:

1. The weak Mertens conjecture is true. Precisely,

$$\int_2^X \left(\frac{M(x)}{x}\right)^2 dx \ll \log X \ .$$

2. $M(x) = O(x(\log\log x)^{\frac{3}{2}})$ except on a set of finite logarithmic measure.

3. The function $e^{-\frac{y}{2}}M(e^y)$ has a limiting distribution $\nu(t)$.

4. Assuming the zeros of the Riemann zeta function are linearly independent over the rationals leads to bounds on the tails of $\nu(t)$. For V large let $B_V = [V, \infty)$ or $(-\infty, -V]$. It is shown that

$$\exp(-c_1 V^{\frac{8}{5}} \exp(c_2 V^{\frac{4}{5}})) \leq \nu(B_V) \leq \exp(-c_3 V^2 \exp(c_4 V^{\frac{4}{5}}))$$

for some effective constants $c_1, c_2, c_3, c_4 > 0$.

5. The true order of $M(x)$ is investigated via the above bounds. It appears that

$$M(x) = \Omega_{\pm}(x^{\frac{1}{2}}(\log\log\log x)^{\frac{5}{4}})$$

is the true lower bound.

Table of Contents

Acknowledgements

I would like to thank my supervisor, Professor David Boyd, who diligently guided me, generously gave of his time, and shared of his extensive knowledge of mathematics. I am very grateful for the financial and academic support that he has provided throughout my graduate studies.

Thank-you also, to Professor Rajiv Gupta, my second reader, whose helpful comments and encouraging remarks assisted me during the final sprint towards home.

Many thanks to Professor Robert Rumely for kindly providing me with his list of zeros of Dirichlet L-functions. I am also grateful to fellow grad student, Glen Pugh, who provided me with his Riemann-Siegel programs.

I must mention my two childhood friends, Wayne Tran and Steven Tieu. Thank-you Wayne for the late night programming tips and many thanks to Steven who set up my computer for me (visit his website at `www.sfu.ca/~stieu/`).

I would also like to thank my many friends and relatives who continually enquired about the progress of my work and encouraged me along the way.

I would like to thank my family. Thank-you to my sisters who faithfully provided meals-on-wheels at crucial times. Thank-you to my cousin Corinne who helped me maintain my sanity by forcing me to work out at the fitness centre at least twice a week. Finally, I would like to thank my father and my mother for their love, support, and guidance. I am particularly grateful that they dissuaded me during third year university from leaving math to study history.

Dedicated to my parents Wally and Lenny

Chapter 1
Introduction

1.1 Prime number races

The main topic of this thesis is the irregularity in the distribution of prime numbers. In the final chapter of Davenport's Multiplicative Number Theory [11], the author writes, "The principal omission in these lectures has been the lack of any account of work on irregularities of distributions, both of the primes as a whole and of primes in the various progessions to the same modulus q." In the past century, many articles in analytic number theory have been written on this topic. This thesis will employ the techniques of the recent article by Rubinstein and Sarnak [62] to study prime numbers that are described by Chebotarev's density theorem and to study the summatory function of the Möbius function.

We define the prime counting function

$$\pi(x) = \#\{p \leq x \mid p \text{ prime } \}$$

and the logarithmic integral

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t} \ .$$

The famous prime number theorem states that

$$\pi(x) \sim \mathrm{Li}(x) \ .$$

This was proven by Hadamard and de la Vallee Poussin in 1896. They proved the theorem in the form

$$\pi(x) = \mathrm{Li}(x) + O\left(x \exp(-c\sqrt{\log x})\right) \ .$$

where c is some effective constant. For details of this proof see Davenport's book [11] pp. 115-124. The true size of the error term is something that is still not known.

However, the Riemann Hypothesis implies a better bound on the error. Recall that the Riemann zeta function is defined to be

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

valid for $\mathrm{Re}(s) > 1$. This function is holomorphic in the complex plane except at $s = 1$ where it has a simple pole. Riemann conjectured that:

Riemann Hypothesis (RH). If $\zeta(\rho) = 0$ and $0 < \mathrm{Re}(\rho) < 1$ then $\mathrm{Re}(\rho) = \frac{1}{2}$.

This is one of the most famous unsolved problem in mathematics. If RH is true then

$$\pi(x) - \mathrm{Li}(x) \ll x^{\frac{1}{2}} \log x \ .$$

The proof of the prime number theorem was the culmination of a century's work by many talented mathematicians, including Legendre, Gauss, Chebyshev, and Riemann. The most profound contribution to the solution was the groundbreaking work of Bernhard Riemann. The key techniques used in Hadamard and de la Vallee Poussin's proof are essentially due to Riemann. Despite the proof of the prime number theorem, there are many things we still do not know about the function $\pi(x) - \mathrm{Li}(x)$. For instance, it was observed by a young Gauss that $\pi(x) < \mathrm{Li}(x)$ for small values of x. Here is a small table of values of these functions computed by Deléglise and Rivat [12].

x	$\pi(x)$	$\mathrm{Li}(x) - \pi(x)$
$1 \cdot 10^{15}$	29 844 570 422 669	1 052 619
$2 \cdot 10^{15}$	58 478 215 681 891	1 317 791
$3 \cdot 10^{15}$	86 688 602 810 119	1 872 580
$4 \cdot 10^{15}$	114 630 988 904 000	1 364 039
$5 \cdot 10^{15}$	142 377 417 196 364	2 277 608
$6 \cdot 10^{15}$	169 969 662 554 551	1 886 041
$7 \cdot 10^{15}$	197 434 994 078 331	2 297 328
$8 \cdot 10^{15}$	224 792 606 318 600	2 727 671
$9 \cdot 10^{15}$	252 056 733 453 928	1 956 031
$1 \cdot 10^{16}$	279 238 341 033 925	3 214 632
$2 \cdot 10^{16}$	547 863 431 950 008	3 776 488
$3 \cdot 10^{16}$	812 760 276 789 503	4 651 601
$4 \cdot 10^{16}$	1 075 292 778 753 150	5 538 861
$5 \cdot 10^{16}$	1 336 094 767 763 971	6 977 890
$6 \cdot 10^{16}$	1 595 534 099 589 274	5 572 837
$7 \cdot 10^{16}$	1 853 851 099 626 620	8 225 687
$8 \cdot 10^{16}$	2 111 215 026 220 444	6 208 817
$9 \cdot 10^{16}$	2 367 751 438 410 550	9 034 988

10^{17}	2 623 557 157 654 233	7 956 589
$2 \cdot 10^{17}$	5 153 329 362 645 908	10 857 072
$3 \cdot 10^{17}$	7 650 011 911 220 803	14 592 271
$4 \cdot 10^{17}$	10 125 681 208 311 322	19 808 695
$5 \cdot 10^{17}$	12 585 956 566 571 620	19 070 319
$6 \cdot 10^{17}$	15 034 102 021 263 820	20 585 416
$7 \cdot 10^{17}$	17 472 251 499 627 256	18 395 468
$8 \cdot 10^{17}$	19 901 908 567 967 065	16 763 001
$9 \cdot 10^{17}$	22 324 189 231 374 849	26 287 786
$1 \cdot 10^{18}$	24 739 954 287 740 860	21 949 555

It would be tempting to conjecture that the difference $\mathrm{Li}(x) - \pi(x)$ is always positive and perhaps goes to infinity also. It is not known whether Gauss made this conjecture. However, Littlewood [47], with remarkable insight, proved that

$$\pi(x) - \mathrm{Li}(x) = \Omega_{\pm}\left(\frac{x^{\frac{1}{2}} \log\log\log x}{\log x}\right).$$

(Note that the notation $f(x) = \Omega_+(g(x))$ means that there exists a positive constant c and an increasing infinite sequence of numbers x_n for $n = 1, 2, \ldots$ such that

$$f(x_n) > cg(x_n) \text{ for } n \geq 1 .$$

Likewise, $f(x) = \Omega_-(g(x))$ means that there exists a positive constant c and an increasing infinite sequence of numbers x_n for $n = 1, 2, \ldots$ such that

$$f(x_n) < -cg(x_n) \text{ for } n \geq 1.$$

The notation $f(x) = \Omega_{\pm}(g(x))$ means that both Ω_+ and Ω_- are true.) Littlewood's result not only implies that this function has an infinite number of sign changes, but it demonstrates how large the function can become. Amazingly, to this day no sign change of $\pi(x) - \mathrm{Li}(x)$ has ever been found. However, large upper bounds for the first sign change have been computed. It is now known that the first sign change of this function is less than 10^{370} [59] . Despite Littlewood's proof, there was still no adequate explanation for why $\mathrm{Li}(x)$ was larger than $\pi(x)$ for small values of x. In addition, Littlewood's result teaches us that no matter how convincing numerical evidence seems, we cannot always believe our intuition.

Similarly, Chebyshev investigated prime numbers modulo four. He noticed that there seem to be more primes congruent to three modulo four than to one modulo four. Define

$$\pi(x; 4, 1) = \#\{p \leq x \mid p \equiv 1 \bmod 4\ \} \text{ and } \pi(x; 4, 3) = \#\{p \leq x \mid p \equiv 3 \bmod 4\ \} .$$

Likewise, we could ask whether $\pi(x;4,3) > \pi(x;4,1)$ for all real numbers. Littlewood also showed that the difference of these functions has an infinite number of sign changes. The first sign change for $\pi(x;4,3) - \pi(x;4,1)$ is easily found to be at $x = 26861$. The notion of comparing primes in different arithmetic progressions is called a prime number race. In the above example, we think of 3 and 1 racing each other. The one that is leading the race depends on whether $\pi(x;4,3) > \pi(x;4,1)$ or $\pi(x;4,1) > \pi(x;4,3)$. We are interested in knowing which residue class is leading the race most of the time.

The prime number races can be generalized to an arbitrary modulus and to more than two residue classes. Set q to be a fixed modulus and $(a,q) = 1$. We use the notation

$$\pi(x;q,a) = \#\{p \le x \mid p \equiv a \bmod q \} \ .$$

We consider a fixed set of reduced residue classes mod q. Suppose these classes are labelled $a_1, a_2, \ldots, a_r$ with $r \le \phi(q)$. Consider the set,

$$P_{q;a_1,a_2,\ldots,a_r} = \{x \ge 2 \mid \pi(x;q,a_1) > \pi(x;q,a_2) > \cdots > \pi(x;q,a_r)\} \ .$$

We can ask if there are infinitely many integers x that belong to $P_{q;a_1,a_2,\ldots,a_r}$. In fact, the set $P_{q;a_1,a_2,\ldots,a_r}$ describes how often the race between the reduced residue classes $a_1, a_2, \ldots a_r$ has the order of a_1 leading a_2 leading a_3 etc...

In order to study prime number races of the above type it is necessary to work with Dirichlet L-functions. For a fixed modulus q we consider $(\mathbb{Z}/q\mathbb{Z})^*$, the group of reduced residue class mod q. A characters χ is a group homomorphism

$$\chi : (\mathbb{Z}/q\mathbb{Z})^* \longrightarrow \mathbb{C}^* \ .$$

Attached to each character χ, is the Dirchlet L-function $L(s,\chi)$. This function is defined by

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n \bmod q)}{n^s}$$

valid for $\mathrm{Re}(s) > 1$. It has a holomorphic continuation to the whole complex plane, except possibly $s = 1$. For more details on this function see [11].

In the 1950's, Turán and Knapowski investigated versions of the prime number race problem by making various assumptions about the location of the zeros of Dirichlet L-functions mod q. A major breakthrough was made on this problem in a 1994 article by Rubinstein and Sarnak [62]. Their work depends on the assumption of the Riemann Hypothesis for Dirichlet L-functions. The Riemann Hypothesis for all Dirichlet L-functions has traditionally been referred to as the Generalized Riemann Hypothesis (GRH). Throughout this thesis, we will loosely refer to the GRH in the context of other L-functions. For example, GRH may refer to the Riemann Hypothesis holding for a set of Artin L-functions. The other conjecture Rubinstein and Sarnak assumed is the Linear Independence conjecture.

Linear Independence conjecture (LI). Consider all Dirichlet characters χ mod q. Assume all $L(s,\chi)$ satisfy the Riemann Hypothesis. If all zeros are written as $\frac{1}{2}+i\gamma$ then the imaginary ordinates of the zeros $\gamma \geq 0$ are linearly independent over $\mathbb{Q}$.

Unfortunately, the method employed by Rubinstein and Sarnak makes it very difficult to remove either of these hypotheses. The assumption of the GRH is not too troubling as this is a well-accepted conjecture in mathematics. However, eminent mathematicians such as Littlewood and Turan did not believe the GRH is true. On the other hand, LI is not very well known outside of analytic number theory. In addition, the assumption of LI has only been used very recently. Due to the nature of LI, numerical evidence is very limited.

We also need to define logarithmic density of a set P of positive real numbers. We define

$$\delta(P) = \lim_{X\to\infty} \frac{1}{\log X} \int_{P\bigcap[2,X]} \frac{dt}{t} \ ,$$

if this limit exists. Note that if a set has a natural density, then it will also have a logarithmic density. This follows from partial integration. On the other hand, a logarithmic density does not always guarantee the existence of a natural density. Rubinstein and Sarnak showed, assuming GRH and LI, that $\delta(P_{q;a_1,a_2,\ldots,a_r})$ exists and is non-zero. The existence of a logarithmic density implies that there are infinitely many members of $P_{q;a_1,a_2,\ldots,a_r}$. This gives a conditional solution to the Shanks-Renyi race game. Another interesting aspect of the Rubinstein-Sarnak work is the calculation of a variety of $\delta(P_{q;a_1,a_2})$. For example, they found that $\delta(P_{q;a_1,a_2}) = \frac{1}{2}$ if a_1 and a_2 are both squares or non-squares mod q. If one of the residue classes is a square and the other a non-square, then $\delta(P_{q;a_1,a_2}) \neq \frac{1}{2}$. Using many zeros of certain Dirichlet L-functions, Rubinstein and Sarnak computed,

$$\delta(P_{3;2,1}) = 0.9990\ldots \text{ and } \delta(P_{4;3,1}) = 0.9959\ldots$$

The computations of these densities is one of the more amusing aspects of their article. These computations demonstrate what Rubinstein and Sarnak have aptly referred to as Chebyshev's Bias. The high percentages indicate that 3 leads the race modulo 4 most of the time. Although 1 may lead the race, this happens for a much smaller logarithmic percentage of the time. Hence, Chebyshev's original intuition that there seem to be more primes 3 modulo 4 than 1 modulo 4 is in some sense correct. In Rubinstein-Sarnak, only the densities of two-way races are computed. Recently, Andrey Feuerverger and Greg Martin [20] have developed a formula that enables the computation of prime number races with three or more residue classes. Feuerverger and Martin also use large lists of zeros of Dirichlet L-functions to compute logarithmic densities of the $P_{q;a_1,,a_2,\ldots,a_r}$ for $r = 2,3,4$. The primary reason for computing densities of races of greater than two residue classes, is that Rubinstein and Sarnak observed that there is an unexpected asymmetry among these sets. For example, it was expected that if all of $a_1, a_2, \ldots, a_r$ are squares or non-squares then $\delta_{q;a_1,a_2,\ldots,a_r} = \frac{1}{r!}$. However, this only occurs for $r = 2$ and one special case for $r = 3$. This asymmetry phenomenon is still not completely understood. Feuerverger and Martin can explain why certain examples of races have either equal or unequal densities.

The purpose of this thesis is to study prime number races, where the primes under consideration cannot be simply described as lying in an arithmetic progression. In particular, we will be considering the following situation. Let L/K be a normal extension of number fields. Let $G = \text{Gal}(L/K)$. In the following notation, $\mathfrak{p}$ refers to a prime ideal of $\mathcal{O}_K$. Furthermore, the symbol $\sigma_{\mathfrak{p}}$ refers to a canonical conjugacy class of G associated to $\mathfrak{p}$. The exact definition of $\sigma_{\mathfrak{p}}$ is explained at the beginning of Chapter Two. Denote a subset of the conjugacy classes of G as $C_1, C_2, \ldots C_r$. For each conjugacy class C_i set

$$\pi_i(x) = \frac{|G|}{|C_i|}\pi_{C_i}(x) = \frac{|G|}{|C_i|}\#\{\mathfrak{p} \subset \mathcal{O}_K \mid \mathfrak{p} \text{ unramified, } \mathbb{N}\mathfrak{p} \leq x \ \& \ \sigma_{\mathfrak{p}} = C_i\}$$

for $1 \leq i \leq r$. By Chebotarev's density theorem, observe that $\pi_{C_i}(x) \sim \frac{|G|}{|C_i|}\text{Li}(x)$. Thus, $\pi_i(x) \sim \text{Li}(x)$. Define the subset $P_{L/K;1,2,\ldots,r}$ of $\mathbb{R}$ as

$$P_{L/K;1,2,\ldots,r} = \{x \geq 2 \mid \pi_1(x) > \pi_2(x) > \cdots > \pi_r(x)\}.$$

Under the assumption of RH for Artin L-functions and a modification of LI for Artin L-functions, we will show that the logarithmic density of these sets exist. Here is the modified version of LI.

Linear Independence conjecture (LI). The set of $\gamma > 0$ such that $L(\frac{1}{2} + i\gamma, \chi, L/K) = 0$, for any χ running over irreducible characters of $\text{Gal}(L/K)$, is linearly independent over $\mathbb{Q}$.

This formulation of LI takes into account the vanishing at $s = \frac{1}{2}$ of some Artin L-functions. If we included all $\gamma \geq 0$, then LI would be trivially false if $\gamma = 0$ were in the set. In the next section we will see examples of zeros at the central point. It seems plausible that the imaginary ordinates of the zeros above the real axis are linearly independent.

We will use the following abbreviation

$$\delta_{L/K;1,2,\ldots,r} = \delta(P_{L/K;1,2,\ldots,r}).$$

In showing the existence of $\delta_{L/K;1,2,\ldots,r}$, we work with the zeros of Artin L-functions. The major difference between Artin L-functions and Dirichlet L-functions is that the holomorphy of the Artin L-functions is not yet known. In addition, some Artin L-functions have a zero at $s = \frac{1}{2}$. On the other hand, it is widely believed that Dirichlet L-functions never vanish at $s = \frac{1}{2}$.

The other type of prime number race considered will occur in the setting of class groups. Let K denote some number field with class group $\mathcal{H}_K$ and class number $h_K = h$. Denote a subset of its ideal classes as $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_r$. Set

$$\pi_{\mathfrak{a}_i}(x) = \sum_{N\mathfrak{p} \leq x,\ \mathfrak{p} \in \mathfrak{a}_i} 1.$$

Note that we have $\pi_{\mathfrak{a}_i}(x) \sim \frac{1}{h}\mathrm{Li}(x)$. Define the subset of $\mathbb{R}^+$ as

$$P_{K;1,2,\ldots,r} = \{x \geq 2 \mid \pi_{\mathfrak{a}_1}(x) > \pi_{\mathfrak{a}_2}(x) > \cdots > \pi_{\mathfrak{a}_r}(x)\ \}$$

In this thesis, we will consider a number of specific examples and compute the logarithmic densities for certain prime number races. These examples will attempt to explain when there are biases in the Chebotarev density type prime number races and prime ideal races in the class group setting. The calculation of the densities required the calculation of many zeros of Dirichlet L-functions and weight one modular form L-functions. These zeros were computed with programs written in C, Fortran, and Maple. The methods of computing the zeros are due to Rumely [63] and Rubinstein [60].

Also, in the class group setting we explain the Chebyshev bias term. When the field under consideration is complex quadratic, we can explain when a certain ideal class wins or loses a race. In the complex quadratic case, we also show that the limiting behaviour of the prime ideals under consideration becomes unbiased as the discriminants of the fields gets larger.

1.2 The summatory function of the Möbius function

The final chapter of this thesis is concerned with the average value of the Möbius function. The Möbius function is defined for positive integers n by

$$\begin{aligned} &\mu(1) = 1, \\ &\mu(n) = 0 \text{ if } n \text{ is not squarefree}, \\ &\mu(n) = (-1)^k \text{ if } n \text{ is squarefree and } n = p_1 \ldots p_k\ . \end{aligned} \tag{1.1}$$

The summatory function of the Möbius function is defined to be

$$M(x) = \sum_{n \leq x} \mu(n)\ .$$

Based on numerical evidence, it was believed by some early number theorists that the bound

$$|M(x)| = \left|\sum_{n\le x} \mu(n)\right| \le \sqrt{x}$$

holds for all $x \ge 2$. Mertens verifed this for all integers $x \le 10000$ and made the above conjecture. Later Von Sterneck verified the inequality up to 500,000. Number theorists were particularly interested in the Mertens conjecture since it implies the famous Riemann Hypothesis. Here is a table of values of $M(x)$ computed by Deléglise and Rivat [15]. Note that the Mertens bound is satisfied in all instances.

n	10	11	12	13	14	15
$M(1\cdot 10^n)$	−33722	−87856	62366	599582	−875575	−3216373
$M(2\cdot 10^n)$	48723	−19075	−308413	127543	2639241	1011871
$M(3\cdot 10^n)$	42411	133609	190563	−759205	−2344314	5334755
$M(4\cdot 10^n)$	−25295	202631	174209	−403700	−3810264	−6036592
$M(5\cdot 10^n)$	54591	56804	−435920	−320046	4865646	11792892
$M(6\cdot 10^n)$	−56841	−43099	268107	1101442	−4004298	−14685733
$M(7\cdot 10^n)$	7971	111011	−4252	−2877017	−2605256	4195668
$M(8\cdot 10^n)$	−1428	−268434	−438208	−99222	3425855	6528429
$M(9\cdot 10^n)$	−5554	10991	290186	1164981	7542952	−12589671

However, an interesting paper by Ingham [34] showed that Mertens hypothesis implies that the imaginary ordinates of the zeta function satisfy some linear relations. This result led the experts to believe that it was more likely for the imaginary parts of the zeros of the Riemann zeta function to be linearly independent than for Mertens hypothesis to be true. In 1985, Odlyzko and te Riele [56] proved that the Mertens hypothesis is false. Their techniques used zeros of the zeta function computed to many decimal points. In fact they showed that

$$\liminf_{x\to\infty} \frac{M(x)}{\sqrt{x}} < -1.009 \text{ and } \limsup_{x\to\infty} \frac{M(x)}{\sqrt{x}} > 1.06\ .$$

In hindsight, we see that it is not plausible to believe the Mertens conjecture. Note that $M(x)$ is closely related to the function $\pi(x) - \mathrm{Li}(x)$ and it also has an explicit formula. If we could prove a result analagous to Littlewood's omega result, then the Mertens conjecture would be proven false. We should expect $M(x)$ to exhibit similar behaviour to $\pi(x) - \mathrm{Li}(x)$.

Although we now know that the Mertens conjecture is false, we still do not have an example of a number x for which $M(x) > \sqrt{x}$. Pintz has shown that there is an $x < \exp(3.21\times 10^{64})$ for which the Mertens conjecture fails. This problem is related to the prime number race problem as we know a certain inequality fails infinitely often, yet we cannot produce a single example of a number demonstrating this falsehood. These types of inconsistencies demonstrate that we are unable to compute primes "large enough" to observe their true behaviour. For computational mathematicians, it still is considered an intriguing problem to find the first counterexample in these type of problems.

A related question to the Mertens conjecture is Pólya's conjecture. Let $\lambda(n) = (-1)^{\Omega(n)}$ where $\Omega(n)$ is the total number of prime factors of n. Like the Mertens conjecture, early numerical evidence suggested that

$$L(x) = \sum_{n \leq x} \lambda(n) \leq 0$$

for all integers n. This conjecture would also imply the Riemann Hypothesis. However, Haselgrove gave a numerical disproof of this conjecture in 1958 [29]. In the same article by Ingham it was also shown that Pólya's conjecture implies that LI is false. It would seem that the disproofs of these conjectures would have laid to rest the mention of these conjectures. However, the behaviour of these types of functions is still not completely understood. The final chapter of this thesis shows that the true nature of these functions depends on the behaviour of the zeros of the zeta function. Another reason to study these conjectures is because of an interesting development concerning automorphic versions of these conjectures made by Ram Murty. His work concerns generalizations of the Pólya conjecture to modular forms. Specifically, consider a cusp form $f(z)$ of weight k for $\Gamma_0(N)$. Assume $f(z)$ is a Hecke eigenform with Fourier expansion at $i\infty$, $f(z) = \sum_{n=1}^{\infty} a_f(n) q^n$. Let $L(s, f)$ be the corresponding normalized L-function. In addition, assume $L(s, f)$ has a zero at $s = \frac{1}{2}$ of multiplicity greater or equal to two. Consider sums of the form

$$L_f(x) = \sum_{n \leq x} a_f(n) \lambda(n) \ .$$

Murty shows that these functions go to infinity as x gets large. In fact, he derives an asymptotic formula for this function assuming certain conditions on the zeros of the corresponding L-functions. When the zero vanishes to order less than two, Murty proves that the sum oscillates infinitely often. In this case, it behaves the same as its natural analogue $\sum_{n \leq x} \lambda(n)$. To prove the asymptotic formula, Murty applies a version of a conjecture made by Steve Gonek and Dennis Hejhal on the zeros of the zeta function. It is this same conjecture that will be required in analyzing $M(x)$. The interesting feature of Murty's result is that it demonstrates that L-functions of modular forms do not always behave in the same manner as their classical analogues. Naively, one would have assumed that the sum $\sum_{n \leq x} a_f(n) \lambda(n)$ oscillates infinitely often. Similarly, we discover, in our studies of Chebyshev's bias in Galois group, that zeros of Artin L-functions at the center of the critical strip can also change the "expected" behaviour.

In Ingham's article concerning the Mertens and Pólya conjecture he showed that LI implies that

$$\liminf_{x\to\infty} \frac{M(x)}{\sqrt{x}} = -\infty \text{ and } \limsup_{x\to\infty} \frac{M(x)}{\sqrt{x}} = \infty \ .$$

One of the key topics of the final chapter of this thesis is to show that the assumption of LI suggests that the behaviour of the above functions can be refined. Using techniques from Rubinstein-Sarnak [62] and Montgomery [48] we will study this function in an attempt to find its true size of magnitude. The analysis differs from Rubinstein-Sarnak in that we need to assume deep conjectures about the zeros of the Riemann zeta function. These conjectures appeared in the literature fairly recently and were made independently by Gonek and Hejhal. Without assuming these conjectures, it would appear to be hopeless to investigate $M(x)$. It will be shown that the Gonek-Hejhal conjecture in conjuction with the Riemann Hypothesis implies that the function $e^{-\frac{y}{2}}M(e^y)$ has a limiting distribution. This gives a conditional solution to a question posed by Heath-Brown [30]. Surprisingly, the same assumption implies the weak Mertens conjecture. Lastly, we will show that the Gonek-Hejhal conjectures suggest the true lower order of $M(x)$ is

$$M(x) = \Omega_{\pm}\left(x^{\frac{1}{2}}(\log\log\log x)^{\frac{5}{4}}\right).$$

This is explained more fully in the final chapter.

Chapter 2
Artin L-Functions

2.1 Introduction

In this section, we will define Artin L-functions and describe some of their crucial properties. The Artin L-functions are the natural analogue of Dirichlet L-functions to number fields. Originally, Artin struggled with giving the correct definition of the Artin L-function. His original definition avoided defining the local factors of the ramified primes. Interestingly, his original definition led him to conjecture and prove the reciprocity law of Class Field Theory. One of the main purposes of Artin L-functions is to describe the distribution of the Frobenius symbol. The Frobenius substitution is significant because it describes how prime ideals split when lifted from a smaller ring of integers to a larger ring of integers.

2.2 The function $L(s, \rho)$

Let L/K be a normal extension of number fields with $G = \text{Gal}(L/K)$. Let $\rho : G \to \mathbb{GL}_n(\mathbb{C})$ be a group representation. Attached to this representation is a meromorphic L-function originally defined by Artin. We now give the details of the definition.

Let $\mathcal{O}_L$ and $\mathcal{O}_K$ be the corresponding ring of integers. To each unramified prime ideal $\mathfrak{p} \subset \mathcal{O}_K$, let $\mathfrak{q}$ be a prime ideal in $\mathcal{O}_L$ such that $\mathfrak{q}$ lies over $\mathfrak{p}$. Define the decomposition group as $D_{\mathfrak{q}} = \{\sigma \in G | \sigma\mathfrak{q} = \mathfrak{q}\}$. There exists a canonical map from $D_{\mathfrak{q}}$ to $\text{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$. Note that if we set $q = \mathbb{N}\mathfrak{p} = p^m$ then $\mathcal{O}_K/\mathfrak{p}$ can be identified with $\mathbb{F}_q$ the field of q elements. Furthermore, if $[\mathcal{O}_L/\mathfrak{q} : \mathcal{O}_K/\mathfrak{p}] = f$, then we can think of $\mathcal{O}_L/\mathfrak{q}$ as $\mathbb{F}_{q^f}$. Therefore, we can regard $\text{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$ as $\text{Gal}(\mathbb{F}_{q^f}/\mathbb{F}_q)$. By Galois theory we know that this group is cyclic of order f. In fact, it is generated by the element $\tau_q : x \to x^q$ for $x \in \mathbb{F}_{q^f}$. The canonical map from $D_{\mathfrak{q}}$ to $\text{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$ is defined by sending $\sigma \to \overline{\sigma}$ where $\overline{\sigma}(x + \mathfrak{q}) = \sigma(x) + \mathfrak{q}$. This is a well-defined map and it can be shown that it is surjective (see Lang pp. 15-16 [45]) . Define the inertia group to be $I_{\mathfrak{q}} = \ker(D_{\mathfrak{q}} \to \text{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p}))$. $I_{\mathfrak{q}}$ can be described as

$I_{\mathfrak{q}} = \{\sigma \in G \mid \sigma(x) \equiv x \bmod \mathfrak{q}, \forall \mathrm{x} \in \mathcal{O}_{\mathrm{L}}\}$. If $\mathfrak{p}$ is unramified, then $I_{\mathfrak{q}} = 1$ and we have $D_{\mathfrak{q}} \cong \mathrm{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$. The Frobenius element $\sigma_{\mathfrak{q}}$ is defined to be the element of $D_{\mathfrak{q}}$ that maps to τ_q under this isomorphism. Now define the Frobenius substitution attached to $\mathfrak{p}$ to be the conjugacy class

$$\sigma_{\mathfrak{p}} = \{\sigma_{\mathfrak{q}} \mid \mathfrak{q} \text{ divides } \mathfrak{p}\ \}.$$

The Frobenius substitution is of fundamental importance in algebraic number theory. It gives information on how the prime $\mathfrak{p}$ factors in the larger ring $\mathcal{O}_L$. We can now define the unramified factors of $L(s,\rho)$.

$$L_{ur}(s,\rho,L/K) = \prod_{\mathfrak{p} \text{ unramified}} (\det(\mathrm{I}_n - \rho(\sigma_{\mathfrak{p}})(\mathbb{N}\mathfrak{p})^{-s})^{-1} \tag{2.1}$$

Observe that this definition is well defined. In fact, it suffices to replace the term $\rho(\sigma_{\mathfrak{p}})$ by $\rho(\sigma_{\mathfrak{q}})$ for any $\mathfrak{q}$ dividing $\mathfrak{p}$. This term is well defined, since any other prime $\mathfrak{q}^{'}$ would produce a conjugate element $\sigma_{\mathfrak{q}'}$. However, by the elementary determinant property, $\det(X) = \det(YXY^{-1})$, the local factors are the same. The most subtle part in defining the Artin L-function is at the local factors of the ramified primes. The problem is that the inertia group is non-trivial in this situation. In fact, we have for $\mathfrak{q}$ lying above $\mathfrak{p}$

$$D_{\mathfrak{q}}/I_{\mathfrak{q}} \cong \mathrm{Gal}(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p}).$$

We can now define $\sigma_{\mathfrak{q}}$ to be the element of the quotient $D_{\mathfrak{q}}/I_{\mathfrak{q}}$ that maps to the generator τ_q. Notice that $\sigma_{\mathfrak{q}}$ is no longer an element of the Galois group G, but a coset of $I_{\mathfrak{q}}$. For this reason we no longer work with the vector space $V = \mathbb{C}^n$. Define a new vector space

$$V^{I_{\mathfrak{q}}} = \{x \in V \mid \rho(\sigma)x = x\ \forall \sigma \in I_{\mathfrak{q}}\ \}.$$

This is the subspace of $I_{\mathfrak{q}}$ invariants. We now define the ramified part as

$$L_{ram}(s,\rho,L/K) = \prod_{\mathfrak{p} \text{ ramified}} (\det\nolimits_{V^{I_{\mathfrak{q}}}}(\mathrm{I} - \rho(\sigma_{\mathfrak{q}})|V^{\mathrm{I}_{\mathfrak{q}}}(\mathbb{N}\mathfrak{p})^{-s})^{-1}. \tag{2.2}$$

In the above equation, I is the identity map on $V^{I_{\mathfrak{q}}}$. Also, $\rho(\sigma_{\mathfrak{q}})|V^{I_{\mathfrak{q}}}$ does not depend on the element in the coset of $\sigma_{\mathfrak{q}}$. This is because we are now working in the vector space $V^{I_{\mathfrak{q}}}$ which consists of elements fixed by $I_{\mathfrak{q}}$. Consequently, nothing is changed if you shift by an element of $I_{\mathfrak{q}}$.

Finally, we can define the Artin L-function $L(s,\rho,L/K)$ as

$$L(s,\rho,L/K) = L_{ur}(s,\rho,L/K)L_{ram}(s,\rho,L/K).$$

Note: If we fix the fields L and K, we abbreviate $L(s,\rho,L/K)$ to $L(s,\rho)$. Furthermore, if we denote the character attached to ρ as $\chi = \mathrm{Tr}\rho$, we also write $L(s,\chi,L/K)$

and $L(s, \chi)$ in place of the above functions.

In analytic number theory, it is convenient to "complete" the L-function under consideration. This means you include gamma-factors and a conductor term of the form Q^s for some positive real number Q. The completed function will then satisfy a nice functional equation relating the Artin L-function's value at the points s with the contragradient Artin L-function at the point $1-s$. We will now define the completed function $\Lambda(s, \chi)$. It will have the form

$$\Lambda(s, \chi) = A(\chi)^{\frac{s}{2}} \gamma_\chi(s) L(s, \chi)$$

where $A(\chi)$ is some real number and $\gamma_\chi(s)$ is the product of gamma factors.

Recall that the finite primes of K are the prime ideals of $\mathcal{O}_K$. The infinite primes correspond to embeddings of K into $\mathbb{C}$. A real infinite prime is a real embedding $\sigma_\nu : K \to \mathbb{R}$ and a complex infinite prime is a pair of complex conjugate embeddings $\sigma_\nu, \sigma_{\nu'} : K \to \mathbb{C}$ and $\sigma_\nu \neq \sigma_{\nu'}$. To each real infinite prime σ_ν we associate a real Archimedean valuation ν of K and to each pair of complex infinite primes σ_ν and $\sigma_{\nu'}$ we associate a complex Archimedean valuation. Let $\gamma(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})$. The gamma factor γ_χ is defined as a product $\gamma_\chi(s) = \prod_\nu \gamma_\chi^\nu(s)$ where ν runs over the archimedean valuations corresponding to the infinite primes. If ν is complex, set $\gamma_\chi^\nu(s) = (\gamma(s)\gamma(s+1))^{\chi(1)}$. On the other hand, if ν is real, let w be a valuation of L extending ν. Attached to w is the decomposition group $G(w) = \{\sigma \in G \mid \sigma w = w\ \}$. In fact, it is shown in [39] that $G(w) \cong \mathrm{Gal}(L_w/K_\nu)$. Since $K_\nu = \mathbb{R}$ and $L_w = \mathbb{R}$ *or* $\mathbb{C}$, $G(w)$ has order 1 or 2. Let σ_w be the generator of $G(w)$. This is analogous to the Frobenius substitution at the finite primes. Now $\rho(\sigma_w)$ acts on V with eigenvalues $+1$ and -1. Thus V has a decomposition into eigenspaces $V = V_\nu^+ \oplus V_\nu^-$. Now set $\gamma_\chi^\nu(s) = \gamma(s)^{\mathrm{dim} V_\nu^+}\gamma(s+1)^{\mathrm{dim} V_\nu^-}$. This completes the definition of the gamma factors. Let $\gamma_\chi(s)$ denote the gamma factors in the completed Artin L-function. Tate expresses this in the following convenient form. Let r_2 denote the number of complex places of K and set

$$\begin{aligned} a_1 = a_1(\chi) &= \sum_{\nu \text{ real}} \mathrm{dim} V^{D_\nu} \\ a_2 = a_2(\chi) &= \sum_{\nu \mid \infty} \mathrm{codim} V^{D_\nu} = \sum_{\nu \text{ real}} \mathrm{codim} V^{D_\nu} \\ n_K = [K : \mathbb{Q}] &= \frac{1}{\chi(1)}(a_1(\chi) + a_2(\chi) + 2r_2\chi(1)) \end{aligned} \tag{2.3}$$

then

$$\gamma_\chi(s) = 2^{r_2\chi(1)(1-s)} \cdot \pi^{-\frac{a_2}{2} - \frac{s}{2}n_K\chi(1)} \cdot \Gamma(s)^{r_2\chi(1)} \cdot \Gamma\left(\frac{s}{2}\right)^{a_1} \cdot \Gamma\left(\frac{1+s}{2}\right)^{a_2}.$$

This form of the gamma factors term makes it convenient to observe where the trival zeros of the Artin L-function $L(s,\chi)$ are. Note that $\Gamma(s)$ has simple poles at $s = 0, -1, -2, -3..$, $\Gamma(\frac{s}{2})$ has simple poles at $s = 0, -2, -4, -6, ..$, and $\Gamma(\frac{s+1}{2})$ has simple poles at $s = -1, -3, -5, ..$. Combining these facts and applying the functional equation shows that $L(s, \rho, L/K)$ has trivial zeros at $s = -2k$ for $k \geq 0$ of order $a_1 + r_2\chi(1)$. Likewise, $L(s, \rho, L/K)$ has a zero at $s = -(2k+1)$ for $k \geq 0$ of order $a_2 + r_2\chi(1)$.

Finally, we need to define the term A_χ. Let $\mathfrak{p}$ be a prime ideal in $\mathcal{O}_K$. Let $\mathfrak{q}$ be a prime ideal of $\mathcal{O}_L$ dividing $\mathfrak{p}$. Consider $\{G_i(\mathfrak{q})\}$ $, i \geq 0$, the ramification groups of G relative to $\mathfrak{q}$. They are defined as $G_i(\mathfrak{q}) = \{\sigma \in G \mid \sigma(x) \equiv x \bmod \mathfrak{q}^{i+1}, \forall x \in \mathcal{O}_L\}$ for each $i \geq 0$. Note that they form a decreasing sequence of normal subgroups

$$G_0 \supseteq G_1 \supseteq G_2 \supseteq G_3 \ldots$$

In addition, $G_0(\mathfrak{q})$ equals the inertia group $I_\mathfrak{q}$ and it also known that for i large enough $G_i = 1$. Let $g_i = |G_i|$. Define the rational number $n(\chi, \mathfrak{p})$ attached to $\mathfrak{p}$ by

$$n(\chi, \mathfrak{p}) = \sum_{i=0}^{\infty} \frac{g_i}{g_0} \operatorname{codim} V^{G_i}.$$

In fact, it is a theorem due to Artin, that $n(\chi, \mathfrak{p})$ is an integer. Also, if $\mathfrak{p}$ is an unramified prime ideal, $n(\chi, \mathfrak{p}) = 0$. We now define an ideal in $\mathcal{O}_K$ called the Artin conductor $\mathfrak{f}(\chi, L/K)$. It is defined as

$$\mathfrak{f}(\chi, L/K) = \mathfrak{f}(\chi) = \prod_{\mathfrak{p}} \mathfrak{p}^{n(\chi,\mathfrak{p})}.$$

Observe that the above product is only over the ramified prime ideals. We now define $A(\chi)$ as follows:

$$A(\chi) = |d_K|^{\chi(1)} \mathbb{N}_{K/\mathbb{Q}}(\mathfrak{f}(\chi)).$$

Here d_K denotes the absolute discriminant of K and $\mathbb{N}_{K/\mathbb{Q}}$ is the norm map. This completes the definition of the Artin L-function.

Brauer, using representation theoretic arguments, proved the Artin L-function is meromorphic and has a functional equation.

Brauer's Induction Theorem 2.2.1 *Let $\Lambda(s,\chi) = A(\chi)^{\frac{s}{2}} \gamma_\chi(s) L(s,\chi)$ for $Re(s) > 1$. Then $\Lambda(s,\chi)$ has a meromorphic continuation to all of $\mathbb{C}$. Also, it satifies the functional equation*

$$\Lambda(1-s,\chi) = W(\chi)\Lambda(s,\overline{\chi})$$

for all $s \in \mathbb{C}$ and $W(\chi)$ is a number of absolute value one. $W(\chi)$ is known as the root number.

Proof See [31] pp. 223-225.

2.3 Properties of Artin L-functions

Although the Dirichlet series expansion of the Artin L-function is quite mysterious, it is possible to give a more concrete expression for $\log L(s,\rho)$.

Proposition 2.3.1

$$\log L(s,\rho) = \sum_{\mathfrak{p}\in\mathcal{O}_K}\sum_{m=1}^{\infty}\frac{\chi(\sigma_{\mathfrak{p}}^m)}{m\mathbb{N}(\mathfrak{p})^m}$$

Note: In the above sum, $\chi(\sigma_{\mathfrak{p}}^m)$ is well-defined for $\mathfrak{p}$ unramified. If $\mathfrak{p}$ is ramified, then let $\mathfrak{q}|\mathfrak{p}$ and $\chi(\sigma_{\mathfrak{p}}^m) = \frac{1}{|I_{\mathfrak{q}}|}\sum_{\tau\in I_{\mathfrak{q}}}\chi(\sigma_{\mathfrak{q}}^m\tau)$.

Proof
Let $n = \dim(\rho)$. Then for an unramified prime $\mathfrak{p}$, consider the matrix $\rho(\sigma_{\mathfrak{p}})$. Let its eigenvalues be $\lambda_i(\mathfrak{p})$ for $i = 1 \leq n$. Then,

$$\det(\mathrm{I}_n - \mathbb{N}(\mathfrak{p})^{-s}\rho(\sigma_{\mathfrak{p}})) = \prod_{i=1}^{n}(1-\lambda_i(\mathfrak{p})\mathbb{N}(\mathfrak{p})^{-s})$$

Therefore, taking logarithms we obtain,

$$\log(\det(\mathrm{I}_n - \mathbb{N}(\mathfrak{p})^{-s}\rho(\sigma_{\mathfrak{p}}))^{-1}) = \sum_{i=1}^{n}\sum_{m=1}^{\infty}\frac{\lambda_i(\mathfrak{p})^m}{m\mathbb{N}(\mathfrak{p})^{ms}} = \sum_{m=1}^{\infty}\frac{\chi(\sigma_{\mathfrak{p}}^m)}{m\mathbb{N}(\mathfrak{p})^{ms}}\ .$$

The argument for the ramified primes is analogous. □

Here are some of the key properties of Artin L-functions.

Theorem 2.3.2
(a) $L(s,\chi,L/K)$ is regular for $Re(s) > 1$.
(b) $L(s,1,L/K) = \zeta_K(s)$ where 1 is the trivial representation.
(c) $L(s,\chi_1+\chi_2,L/K) = L(s,\chi_1,L/K)L(s,\chi_2,L/K)$
(d) $L(s,Ind_H^G\chi,L/K) = L(s,\chi,L/L^H)$ where H is a subgroup of G and L^H is the corresponding fixed field.
(e) Let H be a normal subgroup of G. ρ' is a representation of the factor group G/H and ρ is the corresponding representation of G given by composition with projection. If χ and χ' are the corresponding characters, we have $L(s,\chi',L^H/K) = L(s,\chi,L/K)$.
(f) $\zeta_L(s) = \prod_{\chi\in Irr(G)} L(s,\chi,L/K)^{\chi(1)}$.

Proof
(a) Follows from noticing that

$$|L(s,\chi)| \ll \zeta_K(\sigma)^{\chi(1)}.$$

(b) Follows directly from the definition of an Artin L-function.
(c) By the Dirichlet series expansion of $\log L(s, \chi)$, it is clear that

$$\log L(s, \chi_1 + \chi_2) = \log L(s, \chi_1) + \log L(s, \chi_2).$$

Taking exponentials gives the expression.
(d) See Heilbronn [31] p. 222.
(e) See Heilbronn [31] p. 221.
(f) Consider the regular representation $\mathrm{reg}_G : \mathrm{Gal}(L/K) \to \mathbb{GL}_n(\mathbb{C})$ where $n = |G|$. From representation theory, there is the decomposition

$$\mathrm{reg}_G = \sum_{\chi} \chi(1)\chi \ .$$

Also, by definition of induction of a representation, $\mathrm{reg}_G = \mathrm{Ind}_e^G 1$ where e is the identity element of G and 1 is the trivial representation of e. Hence we obtain by properties (d) and (e) above,

$$\prod_{\chi} L(s, \chi, L/K)^{\chi(1)} = L(s, \mathrm{reg}_G, L/K) = L(s, \mathrm{Ind}_e^G 1, L/K) = L(s, 1, L/L) = \zeta_L(s) \ .$$

In the 1920's Takagi and Artin made innovations in algebraic number theory by proving the main theorems in class field theory. Class field theory is particularly significant as it connects the theory of abelian Galois extensions and generalized ideal class groups. In particular, class field theory helps explain what a one dimensional Artin L-function is. In fact, a one dimensional Artin L-function can be interpreted as a Hecke L-function attached to some ray class group. This is significant since Hecke and Tate gave proofs of the holomorphy for these L-functions. In class field theory, the Artin map plays a significant role. Given a normal abelian extension L/K consider a prime $\mathfrak{p} \subset \mathcal{O}_K$ relatively prime to some modulus $\mathfrak{m}$ of K. We get a map from the set of primes in K to $\mathrm{Gal}(L/K)$ by considering the Frobenius element $\sigma_{\mathfrak{p}}$. This induces by multiplicativity a map

$$\Phi_{\mathfrak{m}} : I_K(\mathfrak{m}) \to \mathrm{Gal}(L/K)$$

from the fractional ideals of K prime to $\mathfrak{m}$ to the Galois group. The first theorem of class field theory shows that $\Phi_{\mathfrak{m}}$ is a surjective map. In addition, if the finite primes dividing $\mathfrak{m}$ are sufficiently large, then $\ker(\Phi_{\mathfrak{m}})$ is a congruence subgroup for $\mathfrak{m}$. That is,

Theorem 2.3.3 Abelian Reciprocity Law *Suppose L/K is an abelian extension of number fields. Let χ be a one-dimensional representation of $G = \mathrm{Gal}(L/K)$. There exists a modulus $\mathfrak{f} = \mathfrak{f}(L/K)$ divisible by all ramified primes (finite and infinite) such that $I_K(\mathfrak{f})/\ker(\Phi(\mathfrak{f})) \cong \mathrm{Gal}(L/K)$.*

Under the above isomorphism, we define a character ψ of the class group by $\psi(\mathfrak{p}) = \chi(\sigma_{\mathfrak{p}})$ for $\mathfrak{p}$ relatively prime to $\mathfrak{f}$. Then

$$L(s, \chi) = L(s, \psi)$$

where the latter function is the Hecke L-function attached to the generalized ideal class group. The abelian reciprocity law is a very deep theorem in number theory. As shown above, it gives a description of all one-dimensional Artin L-functions. It is called the abelian reciprocity law because it is known to reduce to the classical reciprocity laws in many cases. For example, quadratic and cubic reciprocity are consequences of the above theorem.

Abelian reciprocity describes one-dimensional Artin L-functions. In the last century, one of the major trends in representation theory and number theory has been to understand two and larger dimensional Artin L-functions. In general, little is known about these higher dimensional Artin L-functions. However, for irreducible odd two-dimensional representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ there is an almost complete description of all such Artin L-functions This is mostly due to groundbreaking work of Robert Langlands. Consider a continuous, odd, irreducible representation

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{GL}_2(\mathbb{C})$$

Such a representation is odd when $\det\rho(c) = -1$ where c is a complex conjugation element of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Consider the group $G = \mathrm{Im}(\rho)$. This is a finite subgroup of $\mathbb{GL}_2(\mathbb{C})$. Denote the image of G in $\mathbb{PGL}_2(\mathbb{C}) = \mathbb{GL}_2(\mathbb{C})/\mathbb{C}^*$ as PG. All subgroups of $\mathbb{PGL}_2(\mathbb{C})$ have been classified and G is isomorphic to one of the following groups:

$$PG \cong \begin{cases} \text{Dihedral group } D_n \text{ of order } 2n \ (n \geq 2), \\ \text{Alternating group } A_4, \\ \text{Symmetric group } S_4, \\ \text{Alternating group } A_5. \end{cases}$$

(There are also cyclic subgroups of $\mathbb{PGL}_2(\mathbb{C})$, however an irreducible representation excludes this possibility). Although the above extension is infinite, we are actually considering the Artin L-function $L(s, \rho)$ of the finite extension $L = \overline{\mathbb{Q}}^{\ker\rho}/\mathbb{Q}$. In this case, the completed L-function has the form

$$\Lambda(s, \rho) = A(\rho)^{s/2}(2\pi)^{-s}\Gamma(s)L(s, \rho)$$

where $A(\rho)$ is the Artin conductor. Hecke showed that if $PG \cong D_n$, then $L(s, \rho)$ is holomorphic. In the dihedral case, it can be written as a linear combination of theta series (see [1] for a very precise description). When $PG \cong A_4$, Langlands showed that $L(s, \rho)$ is holomorphic. This is a very deep result and used powerful techniques from representation theory. Tunnell extended Langlands techniques to the S_4 case.

However, the holomorphy of $L(s,\rho)$ in the A_5 case remains an open problem. Joe Buhler [5] gave a proof of one A_5 example of Artin conductor 800 being holomorphic in his Ph.D. dissertation. There are now more examples of holomorphic Artin L-functions in the A_5 case (see [24]). Recently, Kevin Buzzard and Richard Taylor have made some advances on this problem. Although we know that in most cases the two-dimensional irreducible Artin L-functions are most likely holomorphic, this does not describe their exact nature. Weil introduced the idea of twisting L-functions by a linear character. Suppose a representation satisfies the following:

Condition A

There exists a positive integer M such that, for all one-dimensional linear representations χ of G with conductor prime to M, $\Lambda(s,\rho\otimes\chi)$ is holomorphic function of s for $s\neq 0,1$

If an Artin L-function satisfies the above condition then the function is related to a certain weight one modular form. This is descibed precisely:

Weil-Langlands Theorem 2.3.4 *Let ρ be an irreducible two-dimensional complex linear representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with conductor N and $\epsilon=\det(\rho)$ odd. Assume that ρ satisfies Condition A. Suppose $L(s,\rho)=\sum_{n=1}^{\infty}a_n n^{-s}$, and let $f(z)=\sum_{n=1}^{\infty}a_n q^n$. Then f is a normalized newform of weight one on $\Gamma_0(N)$ and character ϵ.*

This theorem applies to those representations of type D_n, S_4, and A_4. On the other hand, Serre and Deligne proved a sort of converse to the above theorem.

Theorem 2.3.5 *(Serre-Deligne) Let f be a normalised newform on $\Gamma_0(N)$ of type $(1,\epsilon)$. Then there exists an irreducible two-dimensional complex linear representation ρ of $G_{\mathbb{Q}}$ such that $L_f(s)=L(s,\rho)$. Further, the conductor of ρ is N, and $det(\rho)=\epsilon$.*

Abelian reciprocity and Langland's visionary work on two-dimensional Artin L-functions suggest that Artin L-functions are in fact holomorphic. This conjecture was made in Artin's original paper on these L-functions.

Artin's Holomorphy Conjecture (AC) Let L/K be normal with Galois group G. If ρ is a non-trivial irreducible representation of G, then $L(s,\rho)$ is a holomorphic function.

The Artin conjecture is one of the famous unsolved problems of number theory. This conjecture has many significant applications. For example, it gives an improvement in the error term of the effective Chebotarev density theorem.

2.4 Analytic properties of Artin L-functions

2.4.1 Hadamard factorization

If we assume Artin's holomorphy conjecture, we can obtain analytic results for Artin L-functions which are analogous to similar properties of Dirchlet L-functions. We have the following product formula.

Theorem 2.4.1.1 *Assume $L(s, \chi, L/K)$ is holomorphic for $\chi \neq 1$. Let $\Lambda(s, \chi)$ be the extended Artin L-function. Then we have the factorization*

$$\Lambda(s, \chi) = (s(s-1))^{-\delta(\chi)} \exp(\alpha_\chi + \beta_\chi s) \prod_\rho \left(1 - \frac{s}{\rho}\right) e^{\frac{s}{\rho}}$$

where ρ ranges over all non-trivial zeros of $L(s, \chi)$. α_χ and β_χ are constants depending on χ.

Corollary 2.4.1.2 *The constant β_χ can be expressed in terms of the zeros by the expression*

$$\mathrm{Re}(\beta_\chi) = -\mathrm{Re} \sum_\rho \frac{1}{\rho}$$

Proof Observe that $\overline{\Lambda(s, \chi)} = \Lambda(\overline{s}, \overline{\chi})$. Taking logarithmic derivatives leads to

$$\overline{\frac{\Lambda'}{\Lambda}(s, \chi)} = \frac{\Lambda'}{\Lambda}(\overline{s}, \overline{\chi}).$$

In addition, taking the logarithmic derivative of the functional equation for Λ implies

$$\frac{\Lambda'}{\Lambda}(s, \chi) = -\frac{\Lambda'}{\Lambda}(1 - s, \overline{\chi}).$$

Substituting $s = \frac{1}{2}$ into the two equations and adding shows that

$$\mathrm{Re} \frac{\Lambda'}{\Lambda}\left(\frac{1}{2}, \chi\right) = 0.$$

Also, if ρ is a zero in the critical strip, then $1 - \overline{\rho}$ is a zero too. Consequently,

$$\mathrm{Re} \sum_\rho \left(\frac{1}{2} - \rho\right)^{-1} = 0$$

by grouping the terms ρ and $1-\overline{\rho}$. Taking the logarithmic derivative of the Hadamard factorization and evaluating at $s=\frac{1}{2}$ yields

$$\frac{\Lambda'}{\Lambda}\left(\frac{1}{2},\chi\right)=\beta_\chi+\sum_\rho\left(\frac{1}{2}-\rho\right)^{-1}+\sum_\rho\frac{1}{\rho}.$$

Taking the real part of this equation and applying the previous two equations implies the final result.

2.4.2 A bound for $\frac{L'}{L}(1,\rho)$

It was noted in [62] that for a Dirichlet L-function $L(s,\chi)$ where χ is a character of conductor q, $\frac{L'}{L}(1,\chi)\ll\log\log q$. In this section, we will prove the analogous statement for Artin L-functions. The proof is analogous to the one for Dirichlet L-functions and follows Littlewood's original idea. We will require some lemmas.

Lemma 2.4.2.1 *Let $\epsilon>0$, $\rho:\mathrm{Gal}(L/K)\to\mathbb{GL}_n(\mathbb{C})$ a non-trivial group representation. Assuming RH and AC for $L(s,\rho)$*

$$\left|\frac{L'}{L}(s,\rho)\right|\ll_\epsilon(n_K\chi(1)\log(|t|+2)+\log A(\rho))$$

for $\frac{1}{2}+\epsilon<Re(s)\le 2$, where the implied constant depends on ϵ and $A(\rho)=d_K^{\chi(1)}\mathbb{N}_{K/\mathbb{Q}}(f(\rho))$.

Proof By a lemma proven later in this thesis (Lemma 3.4.5)

$$\left|\frac{L'}{L}(s,\rho)-\sum_{\rho,|\gamma-t|\le 1}\frac{1}{s-\rho}\right|\ll(n_K\chi(1)\log(|t|+2)+\log A(\rho))$$

for $s=\sigma+it$ with $-\frac{1}{4}\le\sigma\le 3$, $|s|\ge\frac{1}{8}$. Now restrict s to the range $\frac{1}{2}+\epsilon\le\sigma\le 2$. In the sum over ρ we have

$$|s-\rho|=|(\sigma-\frac{1}{2})+i(t-\gamma)|\ge\epsilon$$

and hence

$$\left|\sum_{\rho,|\gamma-t|\le 1}\frac{1}{s-\rho}\right|\le\epsilon^{-1}\sum_{\rho,|\gamma-t|\le 1}1\ll\epsilon^{-1}\left(n_K\chi(1)\log(|t|+2)+\log A(\rho)\right).$$

where the last inequality is Lemma 3.4.3. Combining the two estimates proves the lemma. $\Box$.

Lemma 2.4.2.2 *Let $\epsilon > 0$, $0 < y \leq 1$, and $\rho : \mathrm{Gal}(L/K) \to \mathbb{GL}_n(\mathbb{C})$ a group representation. Assuming the Riemann Hypothesis for $L(s,\rho)$ and Artin's conjecture we have*

$$\left| \frac{L'}{L}(1,\rho) + \sum_{\mathfrak{p}^m} \frac{\chi(\sigma_{\mathfrak{p}}^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^m} e^{-\mathbb{N}\mathfrak{p}^m} \right| \ll_\epsilon y^{\frac{1}{2}-\epsilon} n_K \chi(1) \log A(\rho).$$

Proof For $y > 0$ and n a positive integer, consider the identity

$$e^{-ny} = \frac{1}{2\pi i}\int_{(2)} n^{-z} y^{-z}\Gamma(z)dz.$$

Replacing n by $\mathbb{N}\mathfrak{p}^m$ and summing over all prime powers, we obtain

$$\sum_{\mathfrak{p}^m} \frac{\chi(\sigma_{\mathfrak{p}}^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^m} e^{-\mathbb{N}\mathfrak{p}^m y} = \frac{1}{2\pi i}\int_{(2)} -\frac{L'}{L}(1+z,\rho) y^{-z}\Gamma(z)dz$$

valid for $y > 0$. Moving the contour to the left introduces an extra term from the pole at $z = 0$ of the gamma function. Hence,

$$\frac{L'}{L}(1,\rho) + \sum_{\mathfrak{p}^m} \frac{\chi(\sigma_{\mathfrak{p}}^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^m} e^{-\mathbb{N}\mathfrak{p}^m y} = \frac{1}{2\pi i}\int_{(-\frac{1}{2}+\epsilon)} -\frac{L'}{L}(1+z,\rho) y^{-z}\Gamma(z)dz$$

where $y > 0$ and ϵ a small positive number. The right hand side can be bounded by the preceding lemma.

$$\mathrm{RHS} \ll \int_{-\infty}^{\infty} \left|\frac{L'}{L}(1+z,\rho)\right| y^{\frac{1}{2}-\epsilon} \left|\Gamma(-\frac{1}{2}+\epsilon+it)\right| dt \ll_\epsilon \chi(1) y^{\frac{1}{2}-\epsilon}\log A(\rho).$$

This completes the lemma. □

Proposition 2.4.2.3 *Let $\rho : \mathrm{Gal}(L/K) \to \mathbb{GL}_n(\mathbb{C})$ a group representation and $L(s,\rho)$ the corresponding Artin L-function. Assuming RH and AC we have*

$$\left|\frac{L'}{L}(1,\rho)\right| \ll n_K \chi(1) \log\log(A(\rho)).$$

Proof Let x be a free parameter. By the preceding lemma, it suffices to estimate

$$\Sigma = \sum_{\mathfrak{p}^m} \frac{\chi(\sigma_{\mathfrak{p}}^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^m} e^{-\mathbb{N}\mathfrak{p}^m y} = \sum_{\mathbb{N}\mathfrak{p}\leq x} + \sum_{\mathbb{N}\mathfrak{p}>x} + \sum_{\mathbb{N}\mathfrak{p}^m,\, m\geq 2} \left(\frac{\chi(\sigma_{\mathfrak{p}}^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^m} e^{-\mathbb{N}\mathfrak{p}^m y}\right) \tag{2.4}$$

upon an appropriate choice of y. Denote the three sums as Σ_1, Σ_2, and Σ_3 respectively. We obtain

$$\begin{aligned}\Sigma_1 &\leq \chi(1) \sum_{\mathbb{N}\mathfrak{p}\leq x} \frac{\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}} e^{-\mathbb{N}\mathfrak{p}y} \leq n_K\chi(1) \sum_{p\leq x} \frac{\log p}{p} e^{-py} \\ &< n_K\chi(1) \sum_{p\leq x} \frac{\log p}{p} \leq n_K\chi(1)\log x\end{aligned} \tag{2.5}$$

since $\sum_{p\leq x} \frac{\log p}{p} = \log x + O(1)$ (see Davenport [11] p. 57). Set $P(t) = \sum_{p\leq t} \frac{\log p}{p}$ and $\phi(t) = e^{-ty}$. By partial summation, this implies

$$\begin{aligned}\Sigma_2 &\leq n_K\chi(1) \sum_{p>x} \frac{\log p}{p} e^{-py} \\ &\leq n_K\chi(1) \left(P(x)e^{-xy} + y\int_x^\infty P(t)e^{-ty}dt \right) \\ &\ll n_K\chi(1) \left(\log xe^{-xy} + y\int_x^\infty \log te^{-ty}dt \right) .\end{aligned} \tag{2.6}$$

Integrating by parts, the integral is bounded by $\frac{\log x}{ye^{xy}} + \frac{1}{xy^2e^{xy}}$. This leads to

$$\Sigma_2 \ll n_K\chi(1) \left(\frac{\log x}{e^{xy}} + \frac{\log x}{e^{xy}} + \frac{1}{xye^{xy}} \right) .$$

Assume that x and y are chosen such that $xy \gg 1$. Thus, $\Sigma_2 \ll n_K\chi(1)\log x$. Finally, we bound Σ_3:

$$\Sigma_3 \leq n_K\chi(1) \sum_{p^m,\, m\geq 2} \frac{\log p}{p^m} e^{-py} \leq n_K\chi(1) \sum_{p^m,\, m\geq 2} \frac{\log p}{p^m} = n_K\chi(1) \sum_p \frac{\log p}{p(p-1)} \ll n_K\chi(1) .$$

Combining the three estimates shows that

$$\Sigma \ll n_K\chi(1)\log x$$

subject to $xy \gg 1$. Let ϵ be a fixed positive number (One can take, for example, $\epsilon = \frac{1}{10}$). By the preceding proposition, we now have

$$\left| \frac{L'}{L}(1,\rho) \right| \ll y^{\frac{1}{2}-\epsilon} n_K\chi(1)\log A(\rho) + n_K\chi(1)\log x .$$

Now choose $x = (\log A(\rho))^{c_1}$ and $y = \left(\frac{\log\log A(\rho)}{\log A(\rho)}\right)^{c_1}$ where $c_1 = \frac{1}{\frac{1}{2}-\epsilon}$. This yields the stated bound

$$\left| \frac{L'}{L}(1,\rho) \right| \ll n_K\chi(1)\log\log A(\rho) .$$

2.5 Examples of $L(\frac{1}{2}, \rho) = 0$

The first appearance of an Artin L-function with a zero at $s = \frac{1}{2}$ occured in a paper written by Armitage [2]. His example was a degree twelve extension L/K where $[K : \mathbb{Q}] = 4$. The reason this function has a zero is because of a minus one root number. Later Serre discovered a simpler example. His example was an extension $K/\mathbb{Q}$ of degree eight. Serre never published this result, however there are references in the literature to his example. Works by Chowla [6], Friedlander [21], and Frölich [22] refer to this example. Friedlander wrote a short paper giving an application to certain fields whose Dedekind zeta functions have a zero at $s = \frac{1}{2}$. For these fields he gave an effective lower bound to the class number in terms of the field discriminant. Serre's example has Galois group H_8, the quarternion group. Here is Serre's example. Let $K = \mathbb{Q}(\sqrt{5}, \sqrt{41})$, $\theta = \frac{5+\sqrt{5}}{2}\frac{41+\sqrt{205}}{2}$ and consider $L = K(\sqrt{\theta})$. One of the reasons that the group H_8 is studied in this context is because it has a real symplectic character. In fact, a real character has Artin root number equal to ± 1. However, Serre proved that for the orthogonal real characters the root number equals plus one. Thus the initial search for Artin L-functions was amongst quarternion groups. Later Fröhlich proved that there are infinitely many normal extensions $N/\mathbb{Q}$ with Galois group H_8 and Artin root number $W_N = -1$. Serre and Armitage had noticed that these extensions had $W_N = -1$ precisely when $\mathcal{O}_N$ did not have a normal integral basis. (A normal integral basis contains an element $x \in \mathcal{O}_N$ such that the elements $\sigma(x)$ where $\sigma \in H_8$ form an integral basis of $\mathcal{O}_N$). Subsequently, Fröhlich [22] proved this result for tamely ramified extensions N.

Chapter 3
Chebotarev's Density Theorem

3.1 Frobenius' theorem

Frobenius' theorem is a precursor to Chebotarev's density theorem. Frobenius proved his theorem in 1880 and later published it in 1896. In that paper, Frobenius conjectured Chebotarev's density theorem. Before we state Frobenius' theorem, we present a numerical example that illustrates the result. Consider the polynomial $f(x) = x^4 + x - 1$. Let L_f be the splitting field of f over $\mathbb{Q}$. Its polynomial discriminant is $d_f = -283$. Moreover, the resolvent cubic is the irreducible polynomial $g(x) = x^3 + 4x + 1$. Since the resolvent cubic is irreducible and d_f is not a square, $\mathrm{Gal}(L_f/\mathbb{Q}) \cong S_4$ (See [19] pp. 527-529 for more details). We factor $f(x)$ modulo p for some small prime numbers. The possible factorizations are:

1. irreducible
2. a linear factor and an irreducible cubic factor
3. two irreducible quadratic factors
4. one irreducible quadratic factor and two linear factors
5. four linear factors .

Each of these cases are denoted as C_1, C_2, C_3, C_4, and C_5 respectively. Using Maple, we found the following factorizations.

p	$f(x)$ mod p	cycle type
2	$x^4 + x + 1$	C_1
3	$x^4 + x + 2$	C_1
5	$x^4 + x + 4$	C_1
7	$(x^3 + 4x^2 + 2x + 2)(x + 3)$	C_2
11	$(x^3 + 8x^2 + 9x + 7)(x + 3)$	C_2
13	$(x^3 + 11x^2 + 4x + 6)(x + 2)$	C_2
17	$(x^2 + 7x + 5)(x + 12)(x + 15)$	C_4
19	$x^4 + x + 18$	C_1
23	$(x^3 + 11x^2 + 6x + 21)(x + 12)$	C_2

29	$(x^3+7x^2+20x+25)(x+22)$	C_2
31	x^4+x+30	C_1
37	$(x^2+9x+24)(x+32)(x+33)$	C_4
41	$(x^3+19x^2+33x+13)(x+22)$	C_2
43	x^4+x+42	C_1
47	x^4+x+46	C_1
53	$(x^2+6x+40)(x+12)(x+35)$	C_4
59	$(x^3+32x^2+21x+24)(x+27)$	C_2
61	$(x^3+53x^2+3x+38)(x+8)$	C_2
67	$(x^2+13x+2)(x+8)(x+46)$	C_4
71	$(x^2+15x+32)(x^2+56x+51)$	C_3
73	$(x^2+38x+16)(x^2+35x+41)$	C_3
79	$(x^2+77x+22)(x+17)(x+64)$	C_4
83	$(x+24)(x+69)(x+76)(x+80)$	C_5

In this list, each of the five types of splitting occurs. The reason we compute the polynomial mod p is because we want to determine empirically the relative frequency for which each of the 5 different factorizations occurs. In fact, if a large enough sample of prime numbers are tested, we would observe the relative frequencies to be $\frac{6}{24}$, $\frac{8}{24}$, $\frac{3}{24}$, $\frac{6}{24}$, and $\frac{1}{24}$. These numbers are interesting as they indicate a connection to the five conjugacy classes of S_4. Representatives of these five classes are (1234), (123)(4), (12)(34), (12)(3)(4), and (1)(2)(3)(4). Note that the five classes have sizes 6, 8, 3, 6, and 1. The above example suggests that the various factorizations mod p of an irreducible polynomial somehow depends on the Galois group of polynomial's splitting field. This is not a coincidence and Frobenius' theorem explains the exact connection.

Suppose $f(x)$ is a monic polynomial with integer coefficients. Let the degree of f be n. Suppose its discriminant $d_f \neq 0$. Thus it has distinct zeros $\alpha_1, \alpha_2, \ldots, \alpha_n$. Let the splitting field of f be denoted L_f and $G = \mathrm{Gal}(L_f/\mathbb{Q})$ the corresponding Galois group. For each element $g \in G$, g permutes the roots of f. Moreover, g is completely determined by its action on the roots. Suppose that $g(\alpha_i) = \alpha_{\sigma(i)}$ for some permutation $\sigma \in S_n$. This gives a faithful map from $G \to S_n$ by sending $g \to \sigma$. Thus, G is viewed as a subgroup of S_n. Therefore, each element of G is interpreted as a product of disjoint cycles, say

$$g = (a_1 a_2 \ldots a_{n_1})(b_1 b_2 \ldots b_{n_2}) \ldots$$

Here g is product of t cycles of lengths $n_1, n_2, \ldots, n_t$ such that $n = n_1 + n_2 + \cdots + n_t$. We say that g has cycle type $n_1, n_2, \ldots, n_t$. On the other hand, suppose p is a prime and $p \nmid d_f$. This makes f seperable over $\mathbb{F}_p$. Consequently, f factors in $\mathbb{F}_p$ as $f(x) = p_1(x)p_2(x)\cdots p_k(x)$ for some irreducible polynomials p_i. Assume that p_i has degree n_i. We say that f has factorization type $n_1, n_2, \ldots, n_k$ mod p. We can now state Frobenius' theorem.

Theorem 3.1.1 *Given a polynomial f as above, let $P_{n_1,n_2,\ldots,n_t}$ be the set of primes P such that f has factorization type $n_1, n_2, \ldots, n_t$ mod p. The Dirichlet density of $P_{n_1,n_2,\ldots,n_t}$ exists. Moreover,*

$$\mathcal{D}(P_{n_1,n_2,\ldots,n_t}) = \frac{\#\{g \in G \mid g \text{ has cycle type } n_1, n_2, \ldots, n_t\}}{\#G}.$$

Comment Recall that Dirichlet density of a set of primes P is defined as follows:

$$\mathcal{D}(P) := \lim_{s\to 1^+} \frac{\sum_{p\in P} \frac{1}{p^s}}{\sum_{\text{all } p} \frac{1}{p^s}} \text{ if this limit exists .}$$

In fact, the existence of natural density for a set of primes implies the existence of Dirichlet density. However, the converse is not true. Serre [64] p. 76 mentions that the set P_1 consisting of primes whose first decimal digit is equal to 1 has a Dirichlet density, but not a natural density.

3.2 What is Chebotarev's density theorem?

In algebraic number theory, we are interested in a normal extension of number fields L/K and $G = \mathrm{Gal}(L/K)$. Let $\mathfrak{p}$ be a prime in $\mathcal{O}_K$. Extend $\mathfrak{p}$ to an ideal in $\mathcal{O}_L$ by considering $\mathfrak{p}\mathcal{O}_L$. We would like to know if this ideal is still prime or how it factors in the larger Dedekind domain $\mathcal{O}_L$. Since $\mathcal{O}_L$ is a Dedekind domain, we know that $\mathfrak{p}\mathcal{O}_L$ factors as $\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1\mathfrak{q}_2\ldots\mathfrak{q}_g$ for some primes $\mathfrak{q}_i \subset \mathcal{O}_L$ for $1 \le i \le g$. We would like to understand how a given prime splits. This question can be answered, by considering the Frobenius substitution and Chebotarev's density theorem. Recall that, for a prime $\mathfrak{p}$ in K, the Frobenius substitution is a canonical conjugacy class $\sigma_\mathfrak{p}$, of G. In fact, the Frobenius substitution describes the splitting of a prime $\mathfrak{p}$ in the larger ring. An example of how Frobenius affects the splitting of a prime is as follows:

Example 3.2.1

$$\sigma_\mathfrak{p} = 1 \;\Leftrightarrow\; \mathfrak{p} \text{ splits completely in } \mathcal{O}_L \Leftrightarrow \mathfrak{p} = \mathfrak{q}_1\mathfrak{q}_2\ldots\mathfrak{q}_n$$

where $n = [L : K]$.

The Frobenius symbol appears to be a strange and non-intuitive symbol. One may ask how $\sigma_\mathfrak{p}$is distributed as $\mathfrak{p}$ ranges over all prime ideals of $\mathcal{O}_K$. This is answered by Chebotarev's density theorem which was proven in 1926 by Chebotarev, 42 years after Frobenius conjectured the theorem.

Theorem 3.2.1 *Let $C \subset \mathrm{Gal}(L/K)$ be a fixed conjugacy class. Then we have*

$$\pi_C(x) = \sum_{\mathbb{N}\mathfrak{p} \leq x,\ \mathfrak{p} \nmid d_L} 1 \sim \frac{|C|}{|G|}\mathrm{Li}(x) \sim \frac{|C|}{|G|}\frac{x}{\log x} \ .$$

This theorem shows that for a given fixed prime $\mathfrak{p}$, the probability that $\sigma_{\mathfrak{p}}$ equals C is $\frac{|C|}{|G|}$. Consequently, the $\sigma_{\mathfrak{p}}$ are more likely to lie in the larger conjugacy classes. Here are a few examples that help explain this theorem.

Example 3.2.2 *If $L = \mathbb{Q}(\zeta_n)$ and $K = \mathbb{Q}$ then $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^*$. This isomorphism can be realized by noting that elements of the Galois group are of the form $\sigma_a : \zeta_n \to \zeta_n^a$ for $(a, n) = 1$. Note that σ_a is independent of a's residue class modulo n. Thus the map sending $a \bmod n \ \to \sigma_a$ induces the above isomorphism. In addition, it can be shown [39] p.56 that the ramified primes must divide n. Suppose $p \nmid n$ and σ_p is the corresponding Frobenius element. Using local methods , it is possible to show that $\sigma_p : \zeta_n \to \zeta_n^p$. See [39] pp. 132-133. Now choose $C \subset \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ to be an arbitrary conjugacy class where $C = \sigma_a$ for $(a, n) = 1$. By the above comments*

$$\sigma_p = C = \sigma_a \Longleftrightarrow p \equiv a \bmod n \ .$$

However, by Chebotarev's density theorem, the density of this set of primes is $\frac{|C|}{|G|} = \frac{1}{\phi(n)}$. This shows that Chebotarev's theorem includes Dirichlet's theorem on primes in arithmetic progressions. In fact, it is a vast generalization that includes non-abelian Galois groups.

Here is a non-abelian example of Chebotarev's density theorem. This just means that the Galois group of the field extensions is a non-abelian group.

Example 3.2.3 *Consider the polynomial $q(x) = x^3 - x - 1$. Its polynomial discriminant is $d_q = -23$. Let L be the splitting field of q over $\mathbb{Q}$. As the discriminant is non-square, $\mathrm{Gal}(L/\mathbb{Q}) \cong S_3$ where S_3 is the symmetric group on 3 letters. The conjugacy classes of this group are*

$$C_1 = \{1\} \ , \ C_2 = \{(12), (13), (23)\} \ , \ C_3 = \{(123), (132)\}.$$

Applying Kummer's Theorem (see [39] p. 37) and some properties of the Frobenius substitution one can show:

$$\begin{aligned}
&\sigma_p = C_1 \Longleftrightarrow q(x) \text{ splits completely mod } p \Longleftrightarrow p = \mathfrak{q}_1\mathfrak{q}_2\mathfrak{q}_3\mathfrak{q}_4\mathfrak{q}_5\mathfrak{q}_6 \\
&\sigma_p = C_2 \Longleftrightarrow q(x) \text{ has a unique root mod } p \Longleftrightarrow p = \mathfrak{q}_1\mathfrak{q}_2\mathfrak{q}_3 \\
&\sigma_p = C_3 \Longleftrightarrow q(x) \text{ is irreducible mod } p \Longleftrightarrow p = \mathfrak{q}_1\mathfrak{q}_2
\end{aligned}$$

where the $\mathfrak{q}_i$ *are prime ideals in* $\mathcal{O}_L$. *Chebotarev tells us that each of these cases occur with natural density* $\frac{1}{6}, \frac{1}{2}$, *and* $\frac{1}{3}$ *respectively. Now define Dedekind's eta function*

$$\eta(z) = e^{\frac{2\pi iz}{24}} \prod_{n=1}^{\infty}(1 - e^{2n\pi iz})$$

for $z \in \mathbb{H}$. *If we let* $f(z) = \sum_{n=1}^{\infty} a_n q^n = \eta(z)\eta(23z)$, *then it can be shown that* $L(s, f) = L(s, \rho)$ *where* ρ *is the irreducible two dimensional representation of* $\mathrm{Gal}(L/\mathbb{Q})$. *Consequently, it follows that for* $p \neq 23$

$$\begin{aligned}\sigma_p = C_1 &\Longleftrightarrow a_p = 2\\ \sigma_p = C_2 &\Longleftrightarrow a_p = 0\\ \sigma_p = C_3 &\Longleftrightarrow a_p = -1\end{aligned}$$

The above equivalences shows that the distribution of the Frobenius symbol can be detected by computing the prime Fourier coefficients of the modular form $f(z)$. *Checking the value of each* a_p *determines whether* σ_p *equals* C_1, C_2, *or* C_3. *This example gives an explicit connection between a modular form and certain prime numbers.*

3.3 Effective versions of Chebotarev's density theorem

Effective versions of Chebotarev's density theorem provide explicit error terms depending on field constants. These forms of Chebotarev's density theorem did not appear in the literature until the 1970's. Lagarias-Odlyzko and Serre were the first to prove these theorems. Effective versions are important because number theoretic applications depend on the size of the error terms involved. Serre [70] p. 133 proved the following version of Chebotarev.

Theorem 3.3.1 *Let* L/K *be a normal extension of number fields with* $G = \mathrm{Gal}(L/K)$. *Let* $C \subset G$ *be a conjugacy class. Assume the GRH for the Dedekind zeta function of* L. *Then there exists an absolute constant* $c_1 > 0$ *such that*

$$\left|\pi_C(x) - \frac{|C|}{|G|}\mathrm{Li}(x)\right| \leq c_1 \frac{|C|}{|G|} x^{\frac{1}{2}} (\log d_L + n_L \log x)$$

for all $x \geq 2$.

The above theorem by Serre is an improvement of a similar theorem by Lagarias-Odlyzko. Serre's improvement uses a proposition by Hensel that estimates the valuation of the different of the field extension at various primes. The Lagarias-Odlyzko version of this theorem is slightly weaker and under the same assumptions gives the estimate

$$\left|\pi_C(x) - \frac{|C|}{|G|}\mathrm{Li}(x)\right| \leq c_2 \left(\log d_L + \frac{|C|}{|G|} x^{\frac{1}{2}} (\log d_L + n_L \log x)\right).$$

In the Lagarias-Odlyzko paper [43] there is also an unconditional version of effective Chebotarev. For the unconditional version of the theorem they had to take into consideration the effect of a possible Stark zero. The Stark zero is an analogue of the Siegel zero for Dirichlet L-functions. For a real non-principal Dirichlet character χ, there is the possibility that the Dirichlet L-function $L(s, \chi)$ has a zero close to one. Specifically, if q is the conductor of χ then there exists an absolute constant c_3 such that there is at most one real simple zero in the region

$$\mathrm{Re}(s) \geq 1 - \frac{c_3}{\log q} \ , \ \mathrm{Im}(s) \leq 1.$$

Stark generalized this result to Dedekind zeta functions in [73] pp. 139-140.

Theorem 3.3.2 *Let $L \neq \mathbb{Q}$ be a number field with absolute disccriminant d_L. Let $\zeta_L(s)$ be the corresponding Dedekind zeta function. Then $\zeta_L(s)$ has at most one zero in the region*

$$\mathrm{Re}(s) \geq 1 - \frac{1}{4\log|d_L|} \ , \ \mathrm{Im}(s) \leq \frac{1}{4\log|d_L|}$$

If such a zero exists, it is real and simple. In addition, we call this a Stark zero.

Set β_0 to denote the possible Stark zero in the above region. Lagarias-Odlyzko prove

Theorem 3.3.3 *There exist absolute effectively computable constants c_4, c_5 such that*

$$\left|\pi_C(x) - \frac{|C|}{|G|}\mathrm{Li}(x)\right| \leq \frac{|C|}{|G|}\mathrm{Li}(x^{\beta_0}) + c_4 x \exp(-c_5 n_L^{-\frac{1}{2}} (\log x)^{\frac{1}{2}})$$

if $x \geq \exp(10 n_L (log d_L)^2)$. Also, the β_0 term is only present when β_0 exists.

Lagarias and Odlyzko note that their proof "is a direct descendant of de la Vallée Poussin's proof of the prime number theorem." One of the difficulties in giving an unconditional proof of effective Chebotarev is that Artin's holomorphy conjecture is not known to be true. When proving the Siegel-Walfisz theorem on primes in arithmetic progressions, one can use the fact that non-trivial Dirichlet L-functions are holomorphic. However, Lagarias-Odlyzko get around the holomorphy problem by using a trick invented by Deuring. Deuring's trick allows one to change field extensions from L/K to an abelian sub-extension L/E where $E \supset K$. Since L/E is abelian, all Artin L-functions of this extension are holomorphic. The problem with the Deuring trick is that the field constants that appear in the error term now depend on the field E. Hence, they can only be bounded uniformly by field constants in L. In reality, it is expected that field constants appearing in the error depend on both L and K. If Artin's conjecture and GRH are both assumed, much better results can be obtained. Murty, Murty, and Saradha [51] showed that, on average, a much better bound is obtained. Consider the field constants $P(L/K)$ and $M(L/K)$ where

$$P(L/K) = \{p \in \mathbb{Z} \mid \text{there exists } \mathfrak{p} \in K \text{ s.t. } \mathfrak{p}|p \text{ and } \mathfrak{p} \text{ ramifies in } L\}$$

and

$$M(L/K) = n d_K^{\frac{1}{n_K}} \prod_{p \in P(L/K)} p \ .$$

Specifically, they proved

Theorem 3.3.4 *Suppose that all irreducible Artin L-functions of the extension L/K are holomorphic for $s \neq 1$, and that GRH holds for $\zeta_L(s)$. Then*

$$\sum_C \frac{1}{|C|} \left(\pi_C(x) - \frac{|C|}{|G|} \mathrm{Li}(x) \right)^2 \ll x n_K^2 (\log M(L/K) x)^2 \ .$$

The above theorem, immediately yields the following corollary. It is significant to note that the error term contains the term $|C|^{\frac{1}{2}}$ rather than $|C|$.

Corollary 3.3.5 *Under the same assumptions there exists an absolute effectively computable constant c_6 such that*

$$\left| \pi_C(x) - \frac{|C|}{|G|} \mathrm{Li}(x) \right| \leq c_6 (|C|^{\frac{1}{2}} x^{\frac{1}{2}} n_K \log(M(L/K)x)) \ .$$

3.4 An explicit formula

In this section, we assume throughout that L/K is a normal extension with Galois group G. Furthermore, χ denotes some irreducible character of G. The Artin L-function $L(s, \chi, L/K)$ is abbreviated to $L(s, \chi)$. Using techniques from Lagarias-Odlyzko [43] we obtain explicit formulas for the function $\psi(x, \chi)$ which is defined as

$$\psi(x, \chi) = \sum_{\mathbb{N}\mathfrak{p}^m \leq x,\ \mathfrak{p} \text{ unramified}} \chi(\sigma_p^m) \log(\mathbb{N}\mathfrak{p}) \ .$$

Some preliminary lemmas are required. This section follows the Lagarias-Odlyzko paper very closely. The only difference is that we assume Artin's conjecture, so we can avoid Deuring's trick of changing to relative field extensions where Artin's conjecture is known to be true. This is all standard material and is presented for the sake of completeness.

Lemma 3.4.1 *Let* $\sigma = \mathrm{Re}(s) > 1$*, then*

$$\left|\frac{L'}{L}(s,\chi)\right| \ll \frac{\chi(1)n_K}{\sigma - 1} \ .$$

Proof By definition of the Artin L-function,

$$\left|\frac{L'}{L}(s,\chi)\right| \leq -\chi(1)\frac{\zeta_K'}{\zeta_K}(\sigma) \leq \frac{\chi(1)n_K}{\sigma - 1} \ .$$

Lemma 3.4.2 *If* $\sigma = \mathrm{Re}(s) > -\frac{1}{4}$ *and* $|s| \geq \frac{1}{8}$*, then*

$$\left|\frac{\gamma_\chi'}{\gamma_\chi}(s)\right| \ll \chi(1)n_K \log(|s|+2) \ .$$

Proof From the definition of $\gamma_\chi(s)$ we obtain

$$\frac{\gamma_\chi'}{\gamma_\chi}(s) = -r_2\chi(1)\log 2 - \frac{1}{2}n_K\chi(1)\log\pi + r_2\chi(1)\frac{\Gamma'}{\Gamma}(s) + \frac{a_1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right) + \frac{a_2}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s+1}{2}\right) \ .$$

However, observe that $\frac{\Gamma'}{\Gamma}(z) \ll \log(|z|+2)$ for z satisfying $|z| \geq \frac{1}{16}$, $\mathrm{Re}(z) > -\frac{1}{4}$. This leads to

$$\begin{aligned}\frac{\gamma_\chi'}{\gamma_\chi}(s) &\ll r_2\chi(1) + n_K\chi(1) + r_2\chi(1)\log(|s|+2) \\ &\quad + \frac{a_1}{2}\log\left(\frac{|s|}{2}+2\right) + \frac{a_2}{2}\log\left(\frac{|s+1|}{2}+2\right) \\ &\ll r_2\chi(1)\log(|s|+2) + \frac{r_1}{2}\log(|s|+2)) \\ &\ll \chi(1)n_K\log(|s|+2)\end{aligned} \tag{3.1}$$

for s in the stated range. □

Define $n_\chi(t)$ to be

$$n_\chi(t) = \#\{\rho = \beta + i\gamma \mid L(\rho,\chi) = 0 \ , \ |\gamma - t| \leq 1\} \ .$$

Lemma 3.4.3 *For all t*

$$n_\chi(t) \ll \log A(\chi) + \chi(1) n_K \log(|t|+2) \ .$$

Proof Consider the identity

$$\frac{L'}{L}(s,\chi) + \frac{L'}{L}(s,\overline{\chi}) = \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{s-\overline{\rho}} \right) - \log A(\chi) - 2\delta(\chi) \left(\frac{1}{s} + \frac{1}{s-1} \right) - 2\frac{\gamma_\chi'}{\gamma_\chi}(s) \ .$$

Substituting $s = 2+it$ and applying the previous two lemmas implies

$$\begin{aligned} \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{s-\overline{\rho}} \right) &\ll \log A(\chi) + \chi(1) n_K + \chi(1) n_K \log(|t|+2) \\ &\ll \log A(\chi) + \chi(1) n_K \log(|t|+2) \ . \end{aligned} \tag{3.2}$$

On the other hand, the sum on the left is positive. We truncate this sum to only include those $\rho = \beta + i\gamma$ with $|\gamma - t| \le 1$. That is,

$$\begin{aligned} \sum_\rho \mathrm{Re} \left(\frac{1}{s-\rho} + \frac{1}{s-\overline{\rho}} \right) &= \sum_\rho \frac{2-\beta}{(2-\beta)^2 + (t-\gamma)^2} \\ &\ge \sum_{\rho,\, |\gamma - t| \le 1} \frac{2-\beta}{(2-\beta)^2 + (t-\gamma)^2} \\ &\ge \sum_{\rho,\, |\gamma - t| \le 1} \frac{1}{5} = \frac{1}{5} n_\chi(t) \end{aligned} \tag{3.3}$$

since $1 < 2 - \beta < 2$ implies $\frac{2-\beta}{(2-\beta)^2+(t-\gamma)^2} \ge \frac{1}{5}$. Thus, we obtain the required bound for $n_\chi(t)$. $\square$

Lemma 3.4.4 *For any ϵ with $0 < \epsilon \le 1$*

$$B(\chi) + \sum_{\rho,\, |\rho| < \epsilon} \frac{1}{\rho} \ll \epsilon^{-1} (\log A(\chi) + \chi(1) n_K) \ .$$

Proof Consider the logarithmic derivative

$$\frac{L'}{L}(s,\chi) = B(\chi) + \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right) - \frac{1}{2} \log A(\rho) - \delta(\chi) \left(\frac{1}{s} + \frac{1}{s-1} \right) - \frac{\gamma_\chi'}{\gamma_\chi}(s) \ .$$

Setting $s = 2$ and applying the previous lemmas gives

$$B(\chi) + \sum_\rho \left(\frac{1}{2-\rho} + \frac{1}{\rho} \right) \ll \log A(\chi) + \chi(1) n_K \ .$$

Observe that

$$\left|\frac{1}{2-\rho}+\frac{1}{\rho}\right| = \frac{2}{|\rho(2-\rho)|} \le \frac{2}{|\rho|^2} \ ,$$

and so we obtain

$$\sum_{\rho,\,|\rho|\ge 1}\left|\frac{1}{2-\rho}+\frac{1}{\rho}\right| \ll \sum_{j=1}^{\infty}\frac{n_\chi(j)}{j^2} \ll \log A(\chi)+\chi(1)n_K \ .$$

Likewise, $|2-\rho|\ge 1$, so

$$\sum_{|\rho|<1}\left|\frac{1}{2-\rho}\right| \ll \log A(\chi)+\chi(1)n_K \ .$$

Putting this together, we obtain

$$B(\chi)+\sum_{\rho,\,|\rho|<\epsilon}\frac{1}{\rho} \ll \sum_{\rho,\,\epsilon\le|\rho|<1}\frac{1}{|\rho|}+\log A(\chi)+\chi(1)n_K \ .$$

Finallly, the remaining sum is bounded by the preceding lemma.

$$\sum_{\rho,\,\epsilon\le|\rho|<1}\frac{1}{|\rho|} \le \epsilon^{-1}\sum_{\rho,\epsilon\le|\rho|<1} 1 \le \epsilon^{-1}n_\chi(0) \ll \epsilon^{-1}(\log A(\chi)+\chi(1)n_K) \ .$$

This proves the lemma. □

Lemma 3.4.5 *If $s=\sigma+it$ with $-\frac{1}{4}\le\sigma\le 3$, $|s|\ge\frac{1}{8}$, then*

$$\left|\frac{L'}{L}(s,\chi)+\frac{\delta(\chi)}{s-1}-\sum_{\rho,\,|\gamma-t|\le 1}\frac{1}{s-\rho}\right| \ll \log A(\chi)+\chi(1)n_K\log(|t|+5) \ .$$

Proof As in the previous lemma, we evaluate the logarithmic derivative of $L(s,\chi)$ at $s=\sigma+it$ and $3+it$ and subtract. This is done to remove the $B(\chi)$ term. We obtain

$$\begin{aligned}\frac{L'}{L}(s,\chi)-\frac{L'}{L}(3+it,\chi) &= \sum_{\rho}\left(\frac{1}{s-\rho}-\frac{1}{3+it-\rho}\right)-\frac{\gamma_\chi'}{\gamma_\chi}(s)\\ &+\frac{\gamma_\chi'}{\gamma_\chi}(3+it)-\delta(\chi)\left(\frac{1}{s}+\frac{1}{s-1}-\frac{1}{2+it}-\frac{1}{3+it}\right) \ .\end{aligned} \tag{3.4}$$

Applying Lemmas 3.4.1 and 3.4.2 yields

$$\begin{aligned}\left|\frac{L'}{L}(s,\chi)+\frac{\delta(\chi)}{s-1}-\sum_{\rho,\,|\gamma-t|\le 1}\frac{1}{s-\rho}\right| &\ll \chi(1)n_K+\chi(1)n_K\log(|t|+5)\\ +\sum_{\rho,\,|\gamma-t|\ge 1}\left|\frac{1}{s-\rho}-\frac{1}{3+it-\rho}\right| &+ \sum_{\rho,\,|\gamma-t|\le 1}\left|\frac{1}{3+it-\rho}\right| \ .\end{aligned} \tag{3.5}$$

In the last sum observe that $|3+it-\rho| > 1$ and there are $n_\chi(t)$ terms in this sum. Hence, the last sum is bounded by $\log A(\chi) + \chi(1) n_K \log(|t|+5)$. For the other sum on the right we obtain

$$\begin{aligned}\sum_{\rho,\,|\gamma-t|\geq 1} \left| \frac{1}{s-\rho} - \frac{1}{3+it-\rho} \right| &= \sum_{\rho,\,|\gamma-t|\geq 1} \frac{3-\sigma}{|s-\rho||3+it-\rho|} \\ &\ll \sum_{j=1}^{\infty} \frac{n_\chi(t+j)+n_\chi(t-j)}{j^2} \qquad (3.6) \\ &\ll \log A(\chi) + \chi(1) n_K \log(|t|+5)\ .\end{aligned}$$

Combining the estimates proves the lemma. □

Lemma 3.4.6 *If $|z+k| \geq \frac{1}{8}$ for all non-negative integers k, then*

$$\frac{\Gamma'}{\Gamma}(z) \ll \log(|z|+2)\ .$$

Proof See Lagarias-Odlyzko [43] p. 441.

Lemma 3.4.7 *If $s = \sigma + it$ with $\sigma \leq -\frac{1}{4}$, and $|s+m| \geq \frac{1}{4}$ for all non-negative integers m, then*

$$\frac{L'}{L}(s,\chi) \ll \log A(\chi) + \chi(1) n_K \log(|s|+3)\ .$$

Proof Logarithmically differentiating the functional equation for $L(s,\chi)$ yields

$$\frac{L'}{L}(s,\chi) = -\frac{L'}{L}(1-s,\overline{\chi}) - \log A(\chi) - \frac{\gamma_\chi'}{\gamma_\chi}(1-s) - \frac{\gamma_\chi'}{\gamma_\chi}(s)\ .$$

Assuming that s lies in that stated range implies $\mathrm{Re}(s) \geq \frac{5}{4}$. Thus,

$$\left| \frac{L'}{L}(1-s,\overline{\chi}) \right| \ll \chi(1) n_K$$

and

$$\left| \frac{\gamma_\chi'}{\gamma_\chi}(1-s) \right| \ll \chi(1) n_K \log(|1-s|+2) \leq \chi(1) n_K \log(|s|+3)$$

by earlier lemmas. Applying the previous lemma to γ_χ implies that

$$\left| \frac{\gamma_\chi'}{\gamma_\chi}(s) \right| \ll \chi(1) n_K \log(|s|+2)$$

for s in the stated range. Combining estimates implies the theorem. □

Lemma 3.4.8 *Let $\rho = \beta + i\gamma$ have $0 < \beta < 1$, $\gamma \neq t$. If $|t| \geq 2$, $x \geq 2$, and $1 < \sigma_1 \leq 3$, then*

$$\int_{-\frac{1}{4}}^{\sigma_1} \frac{x^{\sigma+it}}{(\sigma+it)(\sigma+it-\rho)} d\sigma \ll |t|^{-1} x^{\sigma_1} (\sigma_1 - \beta)^{-1} \ .$$

Proof see Lagarias-Odlyzko [43] pp. 444-445.

In the main theorem of this section, it will be important to evaluate the integrals

$$I_\chi(x,T) = \frac{1}{2\pi i} \int_{\sigma - iT}^{\sigma+iT} -\frac{L'}{L}(s,\chi) \frac{x^s}{s} \ ds$$

and

$$I_\chi(x,T,U) = \frac{1}{2\pi i} \int_{B_{T,U}} -\frac{L'}{L}(s,\chi) \frac{x^s}{s} \ ds$$

where $B_{T,U}$ is the positively oriented rectangle with vertices at $\sigma_0 - iT, \sigma_0 + iT, -U + iT$, and $-U - iT$. The parameter U will be chosen as $U = j + \frac{1}{2}$ where j is an integer and $\sigma_0 = 1 + (\log x)^{-1}$. We would now like to show that the difference between these two integrals is small. Set

$$R_\chi(x,T,U) = I_\chi(x,T,U) - I_\chi(x,T) \ .$$

We can write $R_\chi(x,T,U) = V_\chi(x,T,U) + H_\chi(x,T,U) + H^*_\chi(x,T,U)$, where

$$\begin{aligned} V_\chi(x,T,U) &= \frac{1}{2\pi} \int_T^{-T} \frac{x^{-U+it}}{-U+it} \frac{L'}{L}(-U+it,\chi) \ dt \\ H_\chi(x,T,U) &= \frac{1}{2\pi i} \int_{-U}^{-\frac{1}{4}} \left(\frac{x^{\sigma - iT}}{\sigma - iT} \frac{L'}{L}(\sigma - iT, \chi) - \frac{x^{\sigma+iT}}{\sigma+iT} \frac{L'}{L}(\sigma+iT,\chi) \right) dt \\ H^*_\chi(x,T) &= \frac{1}{2\pi i} \int_{-\frac{1}{4}}^{\sigma_0} \left(\frac{x^{\sigma - iT}}{\sigma - iT} \frac{L'}{L}(\sigma - iT, \chi) - \frac{x^{\sigma+iT}}{\sigma+iT} \frac{L'}{L}(\sigma+iT,\chi) \right) dt \ . \end{aligned} \tag{3.7}$$

Our immediate goal is to bound each of these integrals. The first two can be estimated easily using previous lemmas. By the choice of $U = j + \frac{1}{2}$, it follows that $|-U+it+m| \geq \frac{1}{4}$ for all integers m. Hence,

$$\begin{aligned} V_\chi(x,T,U) &\ll \frac{x^{-U}}{U} \int_{-T}^{T} \left| \frac{L'}{L}(-U+it,\chi) \right| \ dt \\ &\ll \frac{x^{-U}}{U} T \left(\log A(\chi) + n_K \chi(1) \log(T+U) \right) \ , \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} H_\chi(x,T,U) &\ll \int_{-\infty}^{-\frac{1}{4}} \frac{x^\sigma}{T} \left(\log A(\chi) + n_K\chi(1)\log(|\sigma|+2) + n_K\chi(1)\log T\right)\, d\sigma \\ &\ll \frac{x^{-\frac{1}{4}}}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right) . \end{aligned} \tag{3.9}$$

Finally, we need to bound $H^*(x,T)$. By Lemma 3.4.5 we have the following inequality

$$\frac{L'}{L}(\sigma+iT,\chi) - \sum_{\rho,\,|\gamma-T|\le 1} \frac{1}{\sigma+iT-\rho} \ll \log A(\chi) + n_K\chi(1)\log T$$

valid for $-\frac{1}{4} \le \sigma \le \sigma_0 = 1 + (\log x)^{-1}, x \ge 2, T \ge 2$. Substituting this expression in the defining integral for $H^*_\chi(x,T)$ leads to

$$\begin{aligned} H^*_\chi(x,T) - \frac{1}{2\pi i}\int_{-\frac{1}{4}}^{\sigma_0} &\left(\frac{x^{\sigma-iT}}{\sigma-iT} \sum_{\rho,\,|\gamma+T|\le 1} \frac{1}{\sigma-iT-\rho} - \frac{x^{\sigma+iT}}{\sigma+iT} \sum_{\rho,|\gamma-T|\le 1} \frac{1}{\sigma+iT-\rho} \right)\, d\sigma \\ &\ll \int_{-\frac{1}{4}}^{\sigma_0} \frac{x^\sigma}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right)\, d\sigma \\ &\ll \frac{x}{T\log x} \left(\log A(\chi) + n_K\chi(1)\log T\right) . \end{aligned} \tag{3.10}$$

However, we can estimate the integral by applying Lemma 3.4.8.

$$\begin{aligned} \frac{1}{2\pi i}\int_{-\frac{1}{4}}^{\sigma_0} \frac{x^{\sigma-iT}}{\sigma-iT} \left(\sum_{\rho\ ,\ |\gamma+T|\le 1} \frac{1}{\sigma+iT-\rho} \right)\, d\sigma &\ll \frac{x^{\sigma_0}}{T}(\sigma_0-1)^{-1} n_\chi(-T) \\ &\ll \frac{x\log x}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right). \end{aligned} \tag{3.11}$$

The other part of the integral is treated entirely the same and gives the same error bound. We deduce

$$H^*_\chi(x,T) \ll \frac{x\log x}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right).$$

Therefore, combining the estimates of the three integrals shows that

$$\begin{aligned} R_\chi(x,T,U) &\ll \frac{x^{-U}}{U} T \left(\log A(\chi) + n_K\chi(1)\log(T+U)\right) \\ &+ \frac{x^{-\frac{1}{4}}}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right) + \frac{x\log x}{T} \left(\log A(\chi) + n_K\chi(1)\log T\right) . \end{aligned} \tag{3.12}$$

Theorem 3.4.9 *Assuming Artin's Conjecture for an irreducible character χ associated to the normal extension L/K, we have*

$$\psi(x,\chi) = \delta(\chi)x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} + \frac{a_1 + r_2\chi(1)}{2}\log\left(1-\frac{1}{x^2}\right) + \frac{a_2 + r_2\chi(1)}{2}\log\left(\frac{x-1}{x+1}\right)$$
$$+ O\left(\frac{x\log x}{T}(\log A(\chi) + n_K\chi(1)\log T)\right) + O\left(\chi(1)\log x(\log d_L + n_K\frac{x}{T}\log x)\right)$$
$$+ O\left(\log x(\log A(\chi) + n_K\chi(1))\right) + \sum_{|\rho|<\frac{1}{2}} \frac{1}{\rho}\ . \tag{3.13}$$

Proof Consider the integral

$$I_\chi(x,T) = \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} -\frac{L'}{L}(s,\chi)\frac{x^s}{s}\ ds$$

where $\sigma > 1, T \ge 1$. The integral $I_\chi(x,T)$ is essentially the prime number sum $\psi(x,\chi)$. To see this, observe that if we set

$$\delta(y) = \begin{cases} 1 \text{ if } y > 1 \\ \frac{1}{2} \text{ if } y = 1 \\ 0 \text{ if } 0 \le y \le 1 \end{cases}$$

then we have the following Lemma from p.105 of Davenport's book [11].

Lemma 3.4.10 *Let*

$$I(y,T) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\ ds.$$

Then, for $y > 0, c > 0, T > 0$,

$$|I(y,T) - \delta(y)| \le \begin{cases} y^c \min(1, T^{-1}|\log y|^{-1}) \text{ if } y \ne 1 \\ cT^{-1} \text{ if } y = 1 \end{cases}\ .$$

Therefore, if we substitute the Dirichlet series expansion $-\frac{L'}{L}(s,\chi) = \sum_{\mathfrak{p}^m} \frac{\chi(\sigma_p^m)\log(\mathbb{N}\mathfrak{p})}{\mathbb{N}\mathfrak{p}^{ms}}$ and integrate termwise in the integral $I_\chi(x,T)$ we obtain

$$\left| I_\chi(x,T) - \sum_{\mathbb{N}\mathfrak{p}^m \le x} \chi(\sigma_p^m)\log(\mathbb{N}\mathfrak{p}) \right| \le \chi(1)\sum_{\mathfrak{p}^m,\ \mathbb{N}\mathfrak{p}^m = x} \left(\log(\mathbb{N}\mathfrak{p}) + \sigma_0 T^{-1}\right) + R_0(x,T),$$

where

$$R_0(x,T) = \sum_{\mathfrak{p}^m,\, \mathbb{N}\mathfrak{p}^m \neq x} \chi(\sigma_{\mathfrak{p}}^m) \left(\frac{x}{\mathbb{N}\mathfrak{p}^m}\right)^{\sigma_0} \min(1, T^{-1}|\log \frac{x}{\mathbb{N}\mathfrak{p}^m}|^{-1}) \log(\mathbb{N}\mathfrak{p}) \ .$$

Observe that the first term is only present if there exists a prime ideal $\mathfrak{p}$ with $\mathbb{N}\mathfrak{p}^m = x$ for some integer m. Also, we see that $I_\chi(x,T)$ differs from $\psi(x,\chi)$ only by the ramified prime terms. Note that all ramified primes divide the discriminant of L over K. Thus,

$$\begin{aligned}\left|\sum_{\mathbb{N}\mathfrak{p}^m \leq x} \chi(\sigma_p^m)\log(\mathbb{N}\mathfrak{p}) - \psi(x,\chi)\right| &\leq \chi(1) \sum_{\mathbb{N}\mathfrak{p}^m \leq x,\ \mathfrak{p}\ \text{ramified}} \log(\mathbb{N}\mathfrak{p}) \\ &\leq \chi(1) \sum_{\mathfrak{p}\ \text{ramified}} \log(\mathbb{N}\mathfrak{p}) \sum_{m,\, \mathbb{N}\mathfrak{p}^m \leq x} 1 \\ &\leq \chi(1) 2\log x \sum_{\mathfrak{p}\ \text{ramified}} \log(\mathbb{N}\mathfrak{p}) \\ &\ll \chi(1)\log x \log d_L \ .\end{aligned} \tag{3.14}$$

In addition, there are at most n_K pairs $\mathfrak{p}, m$ satisfying $\mathbb{N}\mathfrak{p}^m = x$. Thus,

$$\sum_{\mathfrak{p},m} \left(\log(\mathbb{N}\mathfrak{p}) + \sigma_0 T^{-1}\right) \leq n_K \log x + n_K \sigma_0 T^{-1} \ .$$

Putting all this together shows that

$$\psi(x,\chi) = I_\chi(x,T) + R_1(x,T)$$

where

$$R_1(x,T) \leq \chi(1)\left(2\log x \log d_L + n_K \log x + n_K\sigma_0 T^{-1}\right) + R_0(x,T) \ .$$

We will now estimate $R_0(x,T)$ and then evaluate $I_\chi(x,T)$ with the residue theorem. For convenience, we will now choose $\sigma_0 = 1 + (\log x)^{-1}$. The reason we make this choice is because of the simple and useful identity $x^{\sigma_0} = ex$. Now divide up the sum $R_0(x,T)$ into $R_0(x,T) = S_1 + S_2 + S_3$. S_1 consists of those prime powers for which $\mathbb{N}\mathfrak{p}^m \leq \frac{3}{4}x$ or $\mathbb{N}\mathfrak{p}^m \geq \frac{5}{4}x$. S_2 consists of the terms for which $|x - \mathbb{N}\mathfrak{p}^m| \leq 1$ and S_3 consists of the remaining terms. For the first term, if $\mathbb{N}\mathfrak{p}^m \leq \frac{3}{4}x$ or $\mathbb{N}\mathfrak{p}^m \geq \frac{5}{4}x$, then

$$\left|\log \frac{x}{\mathbb{N}\mathfrak{p}^m}\right| \geq \min(\log\frac{5}{4}, \log\frac{4}{3}) = \log\frac{5}{4}$$

and

$$\min\left(1, T^{-1}\left|\log\frac{x}{\mathbb{N}\mathfrak{p}^m}\right|^{-1}\right) \ll T^{-1}.$$

Hence we have,

$$S_1 \ll \chi(1)xT^{-1}\sum_{\mathfrak{p},\, m}(\mathbb{N}\mathfrak{p})^{-m\sigma_0}\log(\mathbb{N}\mathfrak{p}) = \chi(1)xT^{-1}\left(-\frac{\zeta_K'}{\zeta_K}(\sigma_0)\right) .$$

However, one can show that for $\sigma > 1$, $-\frac{\zeta_K'}{\zeta_K}(\sigma) \le -n_K\frac{\zeta'}{\zeta}(\sigma)$ where ζ is the Riemann zeta function. Also, it's not difficult to show that $-\frac{\zeta'}{\zeta}(\sigma) \ll (\sigma-1)^{-1}$ for $\sigma > 1$. Combining these two facts yields

$$S_1 \ll \chi(1)n_KxT^{-1}(\sigma_0-1)^{-1} = \chi(1)n_KxT^{-1}\log x .$$

In the second sum, we are counting prime powers $\mathfrak{p}^m$ such that $0 < |\mathbb{N}\mathfrak{p}^m - x| < 1$. Observe that there are at most two integers in the interval $(x-1, x+1)$. Hence, there are at most $2n_K$ prime powers $\mathfrak{p}^m$ that lie in this interval. Also, we have the trivial bound $\min(1, T^{-1}|\log\frac{x}{\mathbb{N}\mathfrak{p}^m}|^{-1}) \le 1$. Using these facts we obtain,

$$S_2 \le 2\chi(1)n_K\log(x+1)\left(\frac{x}{x-1}\right)^{\sigma_0} \ll \chi(1)n_K\log x .$$

In the last sum we consider prime powers that satisfy $1 < |\mathbb{N}\mathfrak{p}^m - x| < \frac{1}{4}x$. We use the elementary estimate $|\log\frac{x}{n}| \le \frac{2n}{|x-n|}$ valid for $n \ge \frac{1}{2}x$.

$$\begin{aligned} S_3 &= \sum_{1<|\mathbb{N}\mathfrak{p}^m-x|<\frac{1}{4}x}\chi(\sigma_{\mathfrak{p}}^m)\left(\frac{x}{\mathbb{N}\mathfrak{p}^m}\right)^{\sigma_0}\min(1, T^{-1}\left|\log\frac{x}{\mathbb{N}\mathfrak{p}^m}\right|^{-1})\log(\mathbb{N}\mathfrak{p}) \\ &\ll \sum_{1<|N\mathfrak{p}^m-x|<\frac{1}{4}x}\chi(1)\left(\frac{x}{\mathbb{N}\mathfrak{p}^m}\right)T^{-1}\left|\log\frac{x}{\mathbb{N}\mathfrak{p}^m}\right|^{-1}\log x \\ &\ll \chi(1)T^{-1}\log x\sum_{1<|n-x|<\frac{1}{4}x}\left|\log\frac{x}{n}\right|^{-1}\sum_{\mathbb{N}\mathfrak{p}^m=n}1 \\ &\ll \chi(1)n_KxT^{-1}\log x\sum_{1\le k\le\frac{1}{4}x}\frac{1}{k} \\ &\ll \chi(1)n_KxT^{-1}(\log x)^2 . \end{aligned} \tag{3.15}$$

We finally have

$$\psi(x,\chi) = I_\chi(x,T) + R_1(x,T)$$

where

$$\begin{aligned} R_1(x,T) &\ll \chi(1)(\log x\log d_L + n_K\log x + n_K\sigma_0T^{-1}) \\ &\quad + \chi(1)(n_KxT^{-1}\log x + n_K\log x + n_KxT^{-1}(\log x)^2) \\ &\ll \chi(1)(\log x\log d_L + n_K\log x + n_KxT^{-1}(\log x)^2) . \end{aligned} \tag{3.16}$$

The above bound is valid for $x \geq 2, T \geq 1$. We now have

$$\psi(x,\chi) = I_\chi(x,T,U) + R_\chi(x,T,U) + R_1(x,T) \ .$$

We will now apply Cauchy's theorem. Note that in the box $B_{T,U}$, $\frac{L'}{L}(s,\rho)\frac{x^s}{s}$ has simple poles at the non-trivial zeros of the Artin L-function $L(s,\rho)$. In addition, there are simple poles at $s=-2m, m=1,2,\ldots$, where the residue is $(a_1+r_2\chi(1))\frac{x^{-2m}}{2m}$. At the negative odd integers, $s=-(2m-1), m=1,2,\ldots$, the residue is $(a_2+r_2\chi(1))\frac{x^{-(2m-1)}}{2m-1}$. Lastly, there is a double pole at $s=0$. In the following paragraph, we will write $h_i(s)$ for $1 \leq i \leq 5$ to denote an entire function in some neighbourhood of $s=0$. We have the expansion

$$\frac{x^s}{s} = \frac{1}{s} + \log x + sh_1(s) \ .$$

In addition, we can obtain a Laurent series for $\frac{L'}{L}(s,\chi)$ from the Hadamard factorization. We have the logarithmic derivative expression

$$\frac{L'}{L}(s,\chi) = B(\chi) + \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) - \frac{1}{2}\log A(\rho) - \delta(\chi)\left(\frac{1}{s} + \frac{1}{s-1}\right) - \frac{\gamma_\chi'}{\gamma_\chi}(s) \ .$$

The second term represents an analytic function in a small neighbourhood of zero. The value of this function at zero is zero. Thus, we can write

$$\frac{L'}{L}(s,\chi) = -\frac{\delta(\chi)}{s} + B(\chi) - \frac{1}{2}\log A(\rho) + \delta(\chi) + sh_2(s) - \frac{\gamma_\chi'}{\gamma_\chi}(s) \ .$$

where $h_2(s)$ is an analytic function in a neighbourhood of zero. It suffices to write down the Laurent series at $s=0$ for $\frac{\gamma_\chi'}{\gamma_\chi}(s)$. Recall that

$$\frac{\gamma_\chi'}{\gamma_\chi}(s) = -r_2\chi(1)\log 2 - \frac{1}{2}n_K\chi(1)\log\pi + r_2\chi(1)\frac{\Gamma'}{\Gamma}(s) + \frac{a_1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right) + \frac{a_2}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s+1}{2}\right) \ .$$

Also, note that $\frac{\Gamma'}{\Gamma}(s)$ has the Laurent series expansion

$$\frac{\Gamma'}{\Gamma}(s) = -\frac{1}{s} - \gamma + sh_3(s) \ .$$

Applying this formula shows

$$\begin{aligned}\frac{\gamma_\chi'}{\gamma_\chi}(s) = &-\frac{r_2\chi(1)+a_1}{s} - r_2\chi(1)\log 2 - \frac{1}{2}n_K\chi(1)\log\pi \\ &- (r_2\chi(1) + \frac{a_1}{2})\gamma + \frac{a_2}{2}\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}\right) + sh_4(s) \ .\end{aligned} \tag{3.17}$$

Therefore,

$$\frac{L'}{L}(s,\chi) = \frac{\kappa_1}{s} + \kappa_2 + sh_5(s)$$

where

$$\begin{aligned}
&\kappa_1 = -\delta(\chi) - r_2\chi(1) + a_1 \ , \\
&\kappa_2 = B(\chi) - \frac{1}{2}\log A(\rho) + \delta(\chi) - r_2\chi(1)\log 2 - \frac{1}{2}n_K\chi(1)\log\pi \\
&- (r_2\chi(1) + \frac{a_1}{2})\gamma + \frac{a_2}{2}\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}\right) \ .
\end{aligned} \tag{3.18}$$

Note that $\kappa_1 \ll n_K\chi(1)$ and that

$$\kappa_2 + \sum_{|\rho|<\frac{1}{2}} \frac{1}{\rho} \ll \log A(\rho) + n_K\chi(1)$$

by Lemma 3.4.4 Multiplying the two Laurent expansions shows that

$$\text{Res}\left(\frac{L'}{L}(s,\chi)\frac{x^s}{s}, s=0\right) = \kappa_1\log x + \kappa_2 \ .$$

By Cauchy's theorem this is

$$\begin{aligned}
\psi(x,\chi) &= \delta(\chi)x - \sum_{|\gamma|\le T}\frac{x^\rho}{\rho} - (a_1 + r_2\chi(1))\sum_{m=1}^{[\frac{U}{2}]}\frac{x^{-2m}}{2m} \\
&- (a_2 + r_2\chi(1))\sum_{m=1}^{[\frac{U+1}{2}]}\frac{x^{-(2m-1)}}{2m-1} - \kappa_1\log x - \kappa_2 + R_\chi(x,T,U) + R_1(x,T) \ .
\end{aligned} \tag{3.19}$$

Letting $U \to \infty$ yields

$$\begin{aligned}
\psi(x,\chi) &= \delta(\chi)x - \sum_{|\gamma|\le T}\frac{x^\rho}{\rho} + \frac{a_1 + r_2\chi(1)}{2}\log\left(1 - \frac{1}{x^2}\right) + \frac{a_2 + r_2\chi(1)}{2}\log\left(\frac{x-1}{x+1}\right) \\
&+ O\left(\frac{x\log x}{T}(\log A(\chi) + n_K\chi(1)\log T)\right) + O\left(\chi(1)\log x(\log d_L + n_K\frac{x}{T}\log x)\right) \\
&+ O\left(\log x(\log A(\chi) + n_K\chi(1))\right) + \sum_{|\rho|<\frac{1}{2}}\frac{1}{\rho} \ .
\end{aligned} \tag{3.20}$$

It should be observed that final sum is empty if GRH is true for $L(s,\chi)$.

3.5 Applications of Chebotarev

The point of this section is to give examples of applications of Chebotarev's density theorem. Chebotarev's density theorem is significant because it can give information on mathematical objects such as Artin L-functions and elliptic curves. Although Chebotarev's density theorem is a prime number theorem, it is a deeper theorem because of these applications. The applications listed in subsections 1-3 are independent of this thesis and are used to give the reader an idea of the significance of the Chebotarev theorem. Subsection 4 contains some calculations concerning the coefficients of Artin L-functions similar to those made by Serre in his long monograph. Many interesting applications can be found in Serre's monograph on effective Chebotarev [70].

3.5.1 Bounds for the least prime ideal

If we have a normal extension L/K with $G = \mathrm{Gal}(L/K)$ and C is some conjugacy class, we can ask how large is the least prime ideal $\mathfrak{p}$ with $\sigma_{\mathfrak{p}} = C$? This is analogous to the classical problem of finding the least prime in an arithmetic progression. Assuming GRH for $\zeta_L(s)$, Lagarias and Odlyzko [43] showed that the least prime ideal satisfying the above conditions satisfies $\mathbb{N}_{K/\mathbb{Q}}(\mathfrak{p}) \ll (\log d_L)^2(\log\log d_L)^2$. This was later improved by Lagarias, Montgomery, and Odlyzko [42] under the same assumptions to $\mathbb{N}_{K/\mathbb{Q}}(\mathfrak{p}) \ll (\log d_L)^2$. However, the best unconditional result is $\mathbb{N}_{K/\mathbb{Q}}(\mathfrak{p}) \ll d_L^c$ where c is some (not yet computed) absolute constant. It should be noted that the latter two results of Lagarias, Montgomery, and Odlyzko are not direct applications of Chebotarev, but can be proven using techniques similar to the previous proof of the Chebotarev density theorem. V.K. Murty proved the following result based on variations of the Lagarias, Montgomery, and Odlyzko results. Let E be a non-CM elliptic curve with conductor N. For $p \nmid N$ let $E(\mathbb{F}_p)$ be the group of $\mathbb{F}_p$ rational points on E. The action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on points of $E(\overline{\mathbb{Q}})$ which are in the kernel of multiplication by l gives a representation

$$\rho_l : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{GL}_2(\mathbb{F}_l) \ .$$

If p is a prime and $p \nmid lN$, then $\rho_l(\sigma_p)$ has trace $a(p)$ and determinant p modulo l. Set

$$T = \mathrm{lcm}_{E'} |E'(\mathbb{Q})_{tors}|$$

where the lcm ranges over elliptic curves E' which are $\mathbb{Q}$-isogeneous to E. Then there is the following result:

Theorem 3.5.1 *Suppose that E does not have complex multiplication and let $l \geq 5$ be a prime that does not divide T. Denote by N the conductor of E. Assume the GRH. Then there is a prime*

$$p \ll (l \log(lN))^2$$

such that $E(\mathbb{F}_p)$ does not have a point of order l.

3.5.2 Lang-Trotter estimates

The Lang-Trotter conjecture for modular forms describes how frequently the coeffecients of the L-function of a modular form take on certain values. Let $k \geq 2, N \geq 1$ be integers. Suppose f is a cusp form of weight k for $\Gamma_0(N)$, which is a normalized eigenform for all Hecke operators $T_p(p \nmid N)$ and the $U_p(p|N)$. f has a Fourier expansion at $i\infty$

$$f(z) = \sum_{n=1}^{\infty} a_n q^n$$

where $q = e^{2\pi i z}$. For simplicity, assume that all $a_n \in \mathbb{Z}$. For each $a \in \mathbb{Z}$ set

$$\pi_{f,a}(x) = \{p \leq x \mid a_p = a\} \ .$$

If $a = 0$ and f is a CM modular form, then we know that

$$\pi_{f,a} \sim \pi(x)/2 \ .$$

In all other cases, Lang and Trotter conjecture

$$\pi_{f,a}(x) \sim c_{f,a} \begin{cases} x^{\frac{1}{2}}/\log x \text{ if } k = 2 \\ \log\log x \text{ if } k = 3 \\ 1 \text{ if } k \geq 4 \end{cases} .$$

Assuming the Riemann Hypothesis for all Artin L-fuctions, Murty, Murty, and Saradha [51] obtained upper bounds for the functions $\pi_{f,a}(x)$. Their technique was an improvement of Serre's original method.

Theorem 3.5.2 *Suppose GRH holds. Then*

$$\pi_{f,a}(x) \ll \begin{cases} x^{\frac{4}{5}} \log x^{-\frac{1}{2}} \text{ if } a \neq 0 \\ x^{\frac{3}{4}} \text{ if } a = 0 \end{cases} .$$

However, the unconditional results are much weaker. For example if E is an elliptic curve V.K. Murty showed that

$$\#\{p \leq x \mid a_p(E) = 0\} \ll \frac{x(\log\log x)^2}{(\log x)^2}.$$

This is obviously a lot weaker than the expected truth and is due to the weak error term in the Chebotarev theorem.

3.5.3 Elliptic curves

Let E be a non-CM elliptic curve. Let S_E denote the set of primes for which E has bad reduction. This is a finite non-empty set after a theorem of Shafarevich. Set $N_E = \prod_{l \in S_E} l$. For each l not dividing N_E, there is the canonical representation

$$\rho_{l^m} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{GL}_2(\mathbb{F}_{l^m}) \ .$$

Taking inverse limits, gives the representation

$$\rho_l : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{GL}_2(\mathbb{Z}_l) \ .$$

Let $G_l = \rho_l(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$. A well-known result of Serre says that if l is sufficiently large, then $G_l = \mathbb{GL}_2(\mathbb{Z}_l)$. Using versions of effective Chebotarev, Serre [70] made this result effective.

Theorem 3.5.3 *Assuming GRH, one has $G_l = \mathbb{GL}_2(\mathbb{Z}_l)$ for all prime numbers l such that*

$$l \geq c(\log N_E)(\log\log 2N_E)^3 \ .$$

3.5.4 Coefficients of Artin L-functions

Serre discovered the amazing fact that degree two Artin L-functions have almost all coefficients zero. In computations of zeros of L-functions, it was necessary to compute coefficients of certain degree two Artin L- functions. This surprising fact greatly sped up the computations of the zeros. It should be noted that for some weight two modular forms, this phenomenon clearly does not hold. For instance, Lang and Trotter conjecture that for a non-CM elliptic curve that the number of Fourier coefficients of its L-series that equal zero for $p \leq x$ is $\ll \frac{\sqrt{x}}{\log x}$. This would imply that the number of non-zero Fourier coefficients of an elliptic curve L-function is of density one. In this section, we give an account of Serre's zero-density result for Artin L-functions and provide some interesting examples that demonstrate Serre's theorem.

Let f be a non-zero element of $S_1(N,\epsilon)$. Suppose that f is an eigenfunction of the Hecke-operators, T_p $(p \nmid N)$ and U_p $(p|N)$, then by the Deligne-Serre Theorem there exists a continuous irreducible representation

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \ \to \mathbb{GL}_2(\mathbb{C})$$

such that $L(s,f) = L(s,\rho)$. Consider the group $G = \mathrm{Im}(\rho)$. This is a finite subgroup of $\mathbb{GL}_2(\mathbb{C})$. Furthermore, consider the image of G in $\mathbb{PGL}_2(\mathbb{C}) = \mathbb{GL}_2(\mathbb{C})/\mathbb{C}^*$. As mentioned before, this must be isomorphic to one of the following groups:

$$\begin{cases} \text{Dihedral group } D_n \text{ of order } 2n \ (n \geq 2), \\ \text{Alternating group } A_4, \\ \text{Symmetric group } S_4, \\ \text{Alternating group } A_5. \end{cases}$$

A proof of this result is given in [46] p. 185 . We say that f is type D_n, A_4, S_4 or A_5 depending on which case $\mathrm{Im}(G)$ is. Let PG_2 denote the set of elements of order two in PG (elements that are the image of trace zero matrices). Set $\lambda = \frac{|PG_2|}{|PG|}$. A simple computation shows that

$$\lambda = \begin{cases} 1/2 + 1/2n \text{ if } PG \cong D_n \ (n \text{ even} \geq 2) \\ 1/2 \text{ if } PG \cong D_n \ (n \text{ odd} \geq 3) \\ 3/8 \text{ if } PG \cong S_4 \\ 1/4 \text{ if } PG \cong A_4 \text{ or } A_5 \end{cases} \tag{3.21}$$

Serre's zero density result may now be stated. If $L(s, \rho)$ is an Artin L-function and $L(s, \rho) = \sum_{n=1}^{\infty} \frac{a_\rho(n)}{n^s}$ then we set $M_\rho(x) = \#\{n \leq x \mid a_\rho(n) \neq 0\}$. Likewise, if f is a weight one modular form with Fourier expansion $f(z) = \sum_{n=1}^{\infty} a_f(n) q^n$ then we set $M_f(x) = \#\{n \leq x \mid a_f(n) \neq 0\}$.

Theorem 3.5.4
(i) Suppose f is a non-zero element of $S_1(N, \epsilon)$ which is an eigenvector of the Hecke operators T_p and U_p. There exists $\alpha > 0$ and $\lambda > 0$ such that

$$M_f(x) \sim \frac{\alpha x}{\log^\lambda x} \ .$$

(ii) Suppose f is a non-zero element of $S_1(N, \epsilon)$. There exists $\lambda > 0$ such that

$$M_f(x) \ll \frac{x}{\log^\lambda x} \ .$$

(iii) Suppose $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \mathbb{GL}_2(\mathbb{C})$ is irreducible and continuous. There exists $\alpha > 0$ and $\lambda > 0$ such that

$$M_\rho(x) \sim \frac{\alpha x}{\log^\lambda x} \ .$$

It should be noted that the λ in parts (i) and (iii) of the theorem agrees with the λ associated to each weight one modular form above. Also, it is possible to extend part (i) of the theorem to an asymptotic formula of the form

$$M_f(x) = \frac{\alpha x}{\log^\lambda x} \left(c_0 + \frac{c_1}{\log x} + \cdots + \frac{c_k}{\log^k x} + O\left(\frac{1}{\log^{k+1} x} \right) \right)$$

This can be achieved by using a technique invented by E. Landau. This is explained in [66] p. 233. In the above theorem, λ is determined by the calculation (3.21). However, α is the more interesting constant to compute. In Serre's long article [70], he gives the following example.

Example 3.5.5 *: f(z)* $= q \prod_{m=1}^{\infty} (1 - q^{12m})^2 = \eta^2(12z)$

Thus $f = q - 2q^{13} - q^{25} + 2q^{37} + q^{49} + 2q^{61} - 2q^{73} - 2q^{97} - 2q^{109} + q^{121} + 2q^{157} + 3q^{169} \cdots$
f is a cusp form of weight one, level $N = 144$ and character $\epsilon(p) = (-1)^{(p-1)/2} = (\frac{-1}{p})$.
Also, f is an eigenfunction of the Hecke-operators T_p $(p \nmid 2, 3)$, U_p $(p = 2, 3)$.

The corresponding Artin L-function is of type D_2. Hence, by the above theorem

$$\#\{n \leq x \mid a_f(n) \neq 0\} \sim \frac{\alpha x}{(\log x)^{\frac{3}{4}}} \ .$$

Serre [70], calculates

$$\alpha = \left(2^{-1}3^{-7}\pi^6 \log(2+\sqrt{3})\right)^{\frac{1}{4}} \prod_{p \equiv 1 \mod 12} (1-p^{-2})^{\frac{1}{2}}/\Gamma(\frac{1}{4}) \ .$$

Example 3.5.6 : $f(z) = q\prod_{m=1}^{\infty}(1-q^m)(1-q^{23m}) = \eta(z)\eta(23z)$

Thus $f = q - q^2 - q^3 + q^6 + q^8 - q^{13} - q^{16} + q^{23} - q^{24} + q^{25} + q^{26} + q^{27} - q^{29} - q^{31} + \cdots$
f is a cusp form of weight one, level $N = 23$ and character $\epsilon(p) = (\frac{p}{23})$. Also, f is an eigenfunction of the Hecke-operators T_p $(p \nmid 23), U_{23}$. Consider the polynomial $q(x) = x^3 - x - 1$ and let L be its splitting field over $\mathbb{Q}$. The Galois group is S_3 as the polynomial has discriminant -23. Let ρ be the irreducible 2-dimensional Artin L-function of this extension. It's possible to check that $L(s, f) = L(s, \rho)$. Using this fact, Serre's result shows that

$$\#\{n \leq x \mid a_\rho(n) \neq 0\} \sim \frac{\alpha x}{(\log x)^{\frac{1}{2}}}$$

where α is some positive constant. We now give a brief outline of how to compute α. Write $L(s, \rho) = \sum_{n=1}^{\infty} a(n)n^{-s}$. Define

$$\phi(s) = \sum_{n=1}^{\infty} a^0(n)n^{-s}$$

where

$$a^0(n) = \begin{cases} 1 \text{ if } a^0(n) \neq 0 \\ 0 \text{ if } a^0(n) = 0 \end{cases}$$

is a multiplicative function. We can write $\phi(s) = \prod_p \phi_p(s)$ where

$$\phi_p(s) = \begin{cases} 1/(1-23^{-s}) \text{ if } p = 23 \\ 1/(1-p^{-s}) \text{ if } \sigma_p = C_1 \\ 1/(1-p^{-2s}) \text{ if } \sigma_p = C_2 \\ (1+p^{-s})/(1-p^{-3s}) \text{ if } \sigma_p = C_3 \end{cases} \ .$$

(For more details on these local factors, one should refer to the discussion of this Artin L-function in Section 4.3.2). On the other hand, consider the Dedekind zeta function of the field $K = \mathbb{Q}(\sqrt{-23})$. Note that $\zeta_K(s) = \zeta(s)L(s, (\frac{-23}{\cdot}))$. Hence, $\zeta_K(s) = \prod_p \zeta_{K,p}(s)$ where

$$\zeta_{K,p}(s) = \begin{cases} 1/(1-23^{-s}) \text{ if } p = 23 \\ 1/(1-p^{-s})^2 \text{ if } \sigma_p = C_1 \text{ or } \sigma_p = C_3 \\ 1/(1-p^{-2s}) \text{ if } \sigma_p = C_2 \end{cases}.$$

Comparing the local Euler factors shows that $\phi(s)^2 = \zeta_K(s)\psi(s)$ where

$$\psi(s) = (1-23^{-s})^{-1} \prod_{\sigma_p = C_2} (1-p^{-2s})^{-1} \prod_{\sigma_p = C_3} \left(\frac{1-p^{-2s}}{1-p^{-3s}} \right)^2.$$

Notice that $\psi(s)$ is holomorphic for $\mathrm{Re}(s) > \frac{1}{2}$. Let κ be the residue at $s = 1$ of $\zeta_K(s)$. Precisely, $\kappa = L(1, (\frac{-23}{\cdot})) = \frac{3\pi}{\sqrt{23}}$. This shows that

$$\psi(s) \sim (\kappa\psi(1))^{\frac{1}{2}}/(s-1)^{\frac{1}{2}} \text{ for } s \to 1.$$

By the Wiener-Ikehara Tauberian theorem it follows that $\alpha = (\kappa\psi(1))^{\frac{1}{2}}/\Gamma(\frac{1}{2})$ and this simplifies to

$$\alpha = \left(\frac{3\sqrt{23}}{22} \prod_{\sigma_p = C_2} (1-p^{-2})^{-1} \prod_{\sigma_p = C_3} \left(\frac{1-p^{-2}}{1-p^{-3}} \right)^2 \right)^{\frac{1}{2}}.$$

Dihedral Examples 3.5.7

Lastly, we construct a family of examples. Each of these examples will correspond to a 2-dimensional irreducible representation of a dihedral group. Let l run through all odd prime numbers. We want to find normal extensions L_l of $\mathbb{Q}$ such that $\mathrm{Gal}(L_l/\mathbb{Q}) \cong D_l$, where D_l is the dihedral group of order $2l$. Nagell showed that for each prime number l there exist infinitely many imaginary quadratic fields K such that $l | h(K)$ where $h(K)$ is the class number of K. In fact, Ram Murty recently proved the following stronger quantatitive result.

Theorem 3.5.8 *Let $g \geq 3$. The number of imaginary quadratic fields whose absolute discriminant is $\leq x$ and whose class group has an element of order g is $\gg x^{\frac{1}{2}+\frac{1}{g}}$.*

For each prime number l, choose a discriminant $-d_l < 0$ such that $\mathbb{Q}(\sqrt{-d_l})$ has class number divisible by l. In addition, choose the d_l so that they are an increasing sequence

$$d_2 < d_3 < d_5 < \cdots < d_l < \cdots$$

Let H_l be the Hilbert class field of $\mathbb{Q}(\sqrt{-d_l})$. This is a relative extension of degree $h_l = h(\mathbb{Q}(\sqrt{-d_l}))$ where h_l is the class number of the quadratic field. By class field theory, we know that

$$\mathrm{Gal}(H_l/\mathbb{Q}(\sqrt{-d_l})) \cong \mathcal{H}_{\mathbb{Q}(\sqrt{-d_l})}.$$

However, $\mathcal{H}_{\mathbb{Q}(\sqrt{-d_l})}$ is a finite abelian group of order h_l. By assumption, $l|h_l$. Consequently, we find a subgroup of the class group with index l. Call this subgroup S_l. Therefore, the field $H_l^{S_l}$ is a normal extension of degree l over $\mathbb{Q}(\sqrt{-d_l})$. Thus we take $L_l = H_l^{S_l}$ as the extensions under consideration. Furthermore, it can be shown that L_l is normal over $\mathbb{Q}$ (see Cox [10] p. 129 exercise 6.4). In addition, it can be proven that

$$\mathrm{Gal}(L_l/\mathbb{Q}) \cong D_l$$

(see Cox [10] pp. 119-122 for a proof of this). Recall that

$$D_l = \{r, s \mid r^l = s^2 = 1,\ rs = sr^{-1}\}.$$

The generators r and s are known as a rotation and reflection respectively. The set of conjugacy classes of D_l can be written down as follows:

$$C_0 = \{1\}\ ,\ C_i = \{r^{\pm i}\} \text{ for } 1 \le i \le l,\ \mathrm{Re} = \{sr^i \mid 1 \le i \le l\}$$

Define the conjugacy set of rotations Ro as the union of C_i for $1 \le i \le l$. For each prime l, let $\rho_l : D_l \to \mathbb{GL}_2(\mathbb{C})$ be the irreducible two-dimensional representation defined as follows:

$$\rho_l(r) = \begin{pmatrix} w_l & 0 \\ 0 & w_l^{-1} \end{pmatrix}\ ,\ \rho_l(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

where w_l is the l-th root of unity $w_l = \exp(\frac{2\pi i}{l})$. Attached to the representation ρ_l is the Artin L-function

$$L(s, \rho_l) = \sum_{n=1}^{\infty} \frac{a_l(n)}{n^s}\ .$$

Define the multiplicative function $a_l^0(n)$ by

$$a_l^0(n) = \begin{cases} 1 \text{ if } a_l(n) \ne 0 \\ 0 \text{ if } a_l(n) = 0 \end{cases}\ .$$

Observe that a_l^0 is a multiplicative function. Define the Dirichlet series

$$\phi_l(s) = \sum_{n=1}^{\infty} \frac{a_l^0(n)}{n^s}\ .$$

By multiplicativity, we obtain an Euler product

$$\phi_l(s) = \prod_p \sum_{m=0}^{\infty} \frac{a_l^0(p^m)}{p^{ms}}\ .$$

Using this product, we claim that the following identity can be proven. This will not be proven as it is very similar to the preceding example.

Claim 3.5.9

$$\phi_l^2(s) = \zeta_{\mathbb{Q}(\sqrt{-d_l})}(s) f_{\mathrm{ram}}(s) \prod_{\sigma_p \subset \mathrm{Ro}} \left(\frac{1 - p^{-(l-1)s}}{1 - p^{-ls}} \right)^2 \prod_{\sigma_p \subset \mathrm{Re}} (1 - p^{-2s})^{-1}$$

where $f_{\mathrm{ram}}(s)$ is holomorphic function at $s = 1$ depending only on the primes that ramify in $\mathbb{Q}(\sqrt{-d_l})$. Furthermore, $f_{\mathrm{ram}}(1) \in \mathbb{Q}$. From the above identity it follows that as $s \to 1^+$

$$\phi_l(s) \sim \frac{1}{(s-1)^{\frac{1}{2}}} \left(L(1, (\frac{-d_l}{\cdot})) f_{\mathrm{ram}}(1) \prod_{\sigma_p \subset \mathrm{Ro}} \left(\frac{1 - p^{-(l-1)}}{1 - p^{-l}} \right)^2 \prod_{\sigma_p \subset \mathrm{Re}} (1 - p^{-2})^{-1} \right)^{\frac{1}{2}} .$$

From the class number formula, $L(1, (\frac{-d_l}{\cdot})) = \frac{\pi h_l}{\sqrt{d_l}}$ and the Wiener-Ikehara Tauberian theorem we obain

$$\#\{ n \leq x \mid a_l(n) \neq 0 \} \sim \frac{\alpha_l x}{(\log x)^{\frac{1}{2}}}$$

where

$$\alpha_l = \left(\frac{h_l f_{\mathrm{ram}}(1)}{\sqrt{d_l}} \prod_{\sigma_p \subset \mathrm{Ro}} \left(\frac{1 - p^{-(l-1)}}{1 - p^{-l}} \right)^2 \prod_{\sigma_p \subset \mathrm{Re}} (1 - p^{-2})^{-1} \right)^{\frac{1}{2}} .$$

The constant α_l is the same for the other irreducible two dimensional representations of D_l. This is because each of these Artin L-functions have an identical $\phi_l(s)$ function. It would be interesting to understand the behaviour of α_l as $l \to \infty$. It should be noted that the products in the above terms are inconsequential to the behaviour of α_l as l grows. The first product approaches 1 as l gets large. In any case, it is bounded unformly with respect to l. The second product is bounded by $\prod_p (1 - p^{-2})^{-\frac{1}{2}} = \frac{1}{\zeta(2)}$.

Chapter 4

Computing L-functions

In this section, a variety of techniques that are used in computing values of various L-functions are described. The L-functions computed in this thesis are either Dirichlet L-functions or degree two Artin L-functions. Historically, the Riemann zeta function was the first L-function to be computed numerically. The zeta function is the most famous L-function and the easiest one to compute. The original interest in computing $\zeta(s)$ was to test the Riemann Hypothesis.

4.1 The prototype: Riemann's zeta function

In Riemann's notes, Siegel found calculations of the first few zeros of the zeta function. In order to compute values of the zeta function, Riemann developed the Riemann-Siegel formula. This formula demonstrates Riemann's deep and insightful understanding of the zeta function. Furthermore, it is a remarkable computation of contour integrals applying the saddle point method. Siegel's name is attached to the formula because he was able to decipher Riemann's unpublished notes and fill in the gaps. When computing values of an L-function, it is often convenient to normalize it on the critical line. This is done by multiplying the L-function by a normalizing factor that makes it real on the critical line. For example, recall that the zeta function satisfies the functional equation

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s)\ .$$

Substituting the value $s = \frac{1}{2} + it$, we obtain

$$\pi^{-\frac{1}{4}-i\frac{t}{2}}\Gamma\left(\frac{1}{4}+i\frac{t}{2}\right)\zeta\left(\frac{1}{2}+it\right) = \pi^{-\frac{1}{4}+i\frac{t}{2}}\Gamma\left(\frac{1}{4}-i\frac{t}{2}\right)\zeta\left(\frac{1}{2}-it\right)\ .$$

Using $\overline{\Gamma(s)} = \Gamma(\overline{s})$ and defining $\theta(t) = -\frac{t}{2}\log(\pi) + \mathrm{Im}(\log(\Gamma(\frac{1}{4}+i\frac{t}{2}))$, it follows that

$$Z(t) = \exp(i\theta(t))\zeta\left(\frac{1}{2}+it\right)$$

is a real-valued function. Finding zeros of the zeta function on the critical line is now reduced to finding zeros of the real-valued function $Z(t)$.

4.1.1 Riemann-Siegel formula

The Riemann-Siegel formula can be derived from a variation of one of Riemann's proofs of the functional equation. The proof uses an appropriate choice of a contour integral. Let $\tau = \frac{t}{2\pi}$, $N = \lfloor \tau \rfloor$, $z = 2(\tau^{\frac{1}{2}} - N) - 1$. Then for any $k \geq 0$

$$Z(t) = 2\sum_{n=1}^{N} n^{-\frac{1}{2}} \cos(\theta(t) - t\log(n)) + (-1)^{N+1}\tau^{-\frac{1}{4}} \sum_{j=0}^{k} \Phi_j(z)(-1)^j \tau^{-\frac{j}{2}} + R_k(\tau)$$

where $\Phi_j(z)$ are certain entire functions and the remainder $R_k(\tau) \ll \tau^{-\frac{2k+3}{4}}$. The benefit of the Riemann-Siegel formula is that the main sum is of length $O(\sqrt{t})$. This makes a single evaluation of the zeta function considerably faster than other methods of computing the zeta-function. For example, Euler-Maclaurin summation applied to $\zeta(\frac{1}{2}+it)$ requires computing a main sum of length $O(t)$. The Riemann-Siegel formula has been used by many people to verify the Riemann Hypothesis. It is known that the Riemann Hypotheis is true for the first $1.5 \cdot 10^9$ zeros. However, A. Odlyzko and A. Schönage [55] have used the Riemann-Siegel formula in conjuction with a technique invented by them to compute many values of the zeta function to heights in the critical strip in the order of $t \approx 10^{21}$. For a very nice discussion of computing zeros of the zeta function and an easy to read derivation of the Riemann-Siegel formula see Glen Pugh's Master's thesis [58].

4.2 Dirichlet *L*-functions

The technique used to compute Dirichlet L-functions is Euler-Maclaurin summation. However, there are Riemann-Siegel formulae that exist for Dirichlet L-functions. Unfortunately, no explicit bounds have been computed for the remainder terms. The technique sketched in subsection 2 was used by Robert Rumely [63] to compute the zeros of many Dirichlet L-functions to small height ($t \leq 10000$). Some of the zeros used in this thesis were computed by Rumely's method. Others were kindly provided by Robert Rumely.

4.2.1 Special values

The discovery of special values of L-functions dates back to Dirichlet's proof of the class number formula. Dirichlet's goal was to show the non-vanishing of Dirichlet

L-functions at $s = 1$. This was significant because it would imply Dirichlet's famous theorem on the infinitude of primes in arithmetic progressions. For real Dirichlet characters χ, there are some beautiful formulae describing $L(1, \chi)$. Let d be a fundamental discriminant and suppose χ is the character $(d|n)$. The class number formula says

$$L(1,\chi) = \begin{cases} 2\pi h(d)/w|d|^{\frac{1}{2}} \text{ if } d < 0 \\ h(d)\log \epsilon/d^{\frac{1}{2}} \text{ if } d > 0 \end{cases}$$

where $h(d)$ is the class number, w is the number of roots of one in $\mathbb{Q}(\sqrt{d})$, and $\log \epsilon$ is the regulator. Another form of this equation is

$$L(1,\chi) = \begin{cases} -\frac{\pi}{|d|^{\frac{3}{2}}} \sum_{m=1}^{|d|} m(d|m) \text{ if } d < 0 \\ -\frac{1}{d^{\frac{1}{2}}} \sum_{m=1}^{d} (d|m) \log \sin \frac{m\pi}{d} \text{ if } d > 0 \end{cases}.$$

This second formula provides a relatively easy but inefficient method of computing real Dirichlet L-functions at $s = 1$. We also require some formulas for $L^{'}(1, \chi)$. Rubinstein-Sarnak use the formula

$$L^{'}(1,\chi) = \gamma L(1,\chi) + \int_0^\infty \frac{h(e^{-u})}{1 - e^{-uq}} \log ue^{-u} \, du$$

where χ is a real Dirichlet character of period q and $h(x) = \sum_{m=1}^{q-1} \chi(m)x^{m-1}$ to compute these special values. However, there are simpler formulas that exist. Consider the field $K = \mathbb{Q}(\sqrt{-d})$ of discriminant $-d$. Let w be the number of roots of unity in K and $\chi_{-d}(n) = (\frac{-d}{n})$ is the associated character, then there is the formula

$$L^{'}(1,\chi_{-d}) = \frac{2\pi}{w\sqrt{d}} \left(h(\gamma + \log 2\pi) - \frac{w}{2} \sum_{n=1}^{d-1} \left(\frac{-d}{n}\right) \log \Gamma\left(\frac{n}{d}\right) \right).$$

The above formula was discovered by many people and appears in a well-known article of Chowla and Selberg. For real quadratic characters, Deninger [13] derived an analagous formula. In this case, let $K = \mathbb{Q}(\sqrt{d})$ be a real quadratic field of discriminant d. Let $\chi = \chi_d$ be the associated character defined by $\chi_d(n) = (\frac{d}{n})$. Deninger discovered the formula

$$L^{'}(1,\chi_d) = L(1,\chi_d)(\gamma + \log(2\pi)) + \frac{1}{\sqrt{d}} \sum_{n=1}^{d-1} \left(\frac{d}{n}\right) R\left(\frac{n}{d}\right)$$

where $R(x) := -(\frac{\partial \zeta(s,x)}{\partial s^2})|_{s=0}$ and $\zeta(s, x)$ is the Hurwitz zeta function. Note that this is defined by

$$\zeta(s,x) = \sum_{n=0}^{\infty} \frac{1}{(n+x)^s}$$

for Re(s) > 1 and $x > 0$. It should be observed that this function has a pole at $s = 1$ and is analytic everywhere else. One way to analytically continue this functions is to apply Euler-Maclaurin summation. Using the aformentioned formulae, we computed the following table of data.

χ	h	$L(1,\chi)$	$L^{'}(1,\chi)$
$(-3\|\cdot)$	1	0.6045997880780726168645	0.2226629869686015094866
$(-4\|\cdot)$	1	0.7853981633974483096158	0.1929013167969124293634
$(-7\|\cdot)$	1	1.187410411723725948785	0.0185659810930280571722 6
$(-8\|\cdot)$	1	1.110720734539591561754	$-0.0230045878627360103171 2$
$(-11\|\cdot)$	1	0.9472258250994829364300	$-0.0797737527762439195438 1$
$(-23\|\cdot)$	3	1.965202054107859165903	$-0.8295529541671135684326$
$(-31\|\cdot)$	3	1.692740092179276090282	$-0.7636917992698947825824$
$(-47\|\cdot)$	5	2.291241928528615936699	-1.469050657121794989610
$(-59\|\cdot)$	3	1.227001578948635475258	$-0.6541524534839911010475$
$(-71\|\cdot)$	7	2.609869177157845864345	-2.042419052345417779933
$(-83\|\cdot)$	3	1.034503778430993813932	$-0.4748405532715693972689$
$(-107\|\cdot)$	3	0.9111276755829139131380	$-0.3227283614369443729385$
$(-139\|\cdot)$	3	0.7993992331016359259937	$-0.3215125570700064115168$
$(-211\|\cdot)$	3	0.6488284725382536365639	$-0.0989083378673081696173 1$
$(-283\|\cdot)$	3	0.5602448972645525149179	0.0017651457227207804593 98

χ	(a,b)	$L(1,\chi)$	$L^{'}(1,\chi)$
$(5\|\cdot)$	$(0,1)$	0.4304089409640040388895	0.3562406470307614988656
$(8\|\cdot)$	$(1,1)$	0.6232252401402305133940	0.3939500015064181287680
$(12\|\cdot)$	$(2,1)$	0.7603459963009463475309	0.3624949106605562121170
$(13\|\cdot)$	$(1,1)$	0.6627353910718455897135	0.3114667901362450908270
$(17\|\cdot)$	$(3,2)$	1.016084833842840752195	0.3364169850639022493590
$(29\|\cdot)$	$(2,1)$	0.6117662895623068698260	0.1897143093476905329860
$(37\|\cdot)$	$(5,2)$	0.8192921687254318779237	0.1277170050822472878060
$(41\|\cdot)$	$(27,10)$	1.299093061575098921650	0.0502999783012077864320
$(61\|\cdot)$	$(17,5)$	0.9383101982488353661439	-0.004151780673833061164
$(89\|\cdot)$	$(447,106)$	1.464441402264019404650	$-0.2538497654985935447790$

In the above table, each of the discriminants has class number one. The column (a, b) denotes the coordinates of the fundamental unit ϵ of the field. Precisely,

$$\epsilon = a + b \cdot \omega \text{ where } \omega = \begin{cases} \sqrt{d}/2 \text{ if } d \equiv 0 \bmod 4 \\ (1+\sqrt{d})/2 \text{ if } d \equiv 1 \bmod 4 \end{cases}$$

The reason we are interested in calculating the values of these Dirichlet L-functions $L(s,\chi)$ and their derivatives at $s=1$ is because there are explicit formulas relating the sum $\sum_{\gamma>0}\frac{1}{\frac{1}{4}+\gamma^2}$ to the quotient $\frac{L'}{L}(1,\chi)$. For instance, there is the formula

$$R_\chi := \sum_{\gamma>0}\frac{1}{\frac{1}{4}+\gamma^2} = \frac{1}{2}\sum_{\gamma}\frac{1}{\frac{1}{4}+\gamma^2} = \frac{1}{2}\log\frac{q}{\pi} - \frac{\gamma}{2} - (1+\chi(-1))\frac{\log 2}{2} + \frac{L'}{L}(1,\chi)$$

for a real primitive Dirichlet character χ mod q, where γ is Euler's constant (see Rubinstein-Sarnak [62] p. 191). The observant reader should note that in the sum γ refers to the imaginary ordinate of a zero of the L-function. The other instance of γ is Euler's constant. Using this formula we computed the following values for R_χ.

χ	R_χ
$(-3\vert\cdot)$	0.0566149849287361709559
$(-4\vert\cdot)$	0.0777839899617929644310
$(-7\vert\cdot)$	0.1276179891459105141348
$(-8\vert\cdot)$	0.1580365896651604878807
$(-11\vert\cdot)$	0.2537565567266778295614
$(-23\vert\cdot)$	0.2846533845506610025162
$(-31\vert\cdot)$	0.4048636747453660565112
$(-47\vert\cdot)$	0.4229419061258128645776
$(-59\vert\cdot)$	0.6446650486982674010033
$(-71\vert\cdot)$	0.4877919022006494081390
$(-83\vert\cdot)$	0.8894443202486700527320
$(-107\vert\cdot)$	1.121234026297686916567
$(-139\vert\cdot)$	1.204071464756366281607
$(-211\vert\cdot)$	1.662514864924585110179
$(-283\vert\cdot)$	1.964901341540300528290

χ	R_χ
$(5\vert\cdot)$	0.0782784769971432485011
$(8\vert\cdot)$	0.1177157809443544633677
$(12\vert\cdot)$	0.1650833123055380477515
$(13\vert\cdot)$	0.1983262896261366805427
$(17\vert\cdot)$	0.1935781504482023305875
$(29\vert\cdot)$	0.4396370850173859318395
$(37\vert\cdot)$	0.4072260086574323116915
$(41\vert\cdot)$	0.3413853807438039974910
$(61\vert\cdot)$	0.4968922340120434424005
$(89\vert\cdot)$	0.5168558438368311756474

4.2.2 Computing Dirichlet L-functions

Suppose we have a Dirichlet character χ defined on $(\mathbb{Z}/q\mathbb{Z})^*$. Consider the L-function $L(s,\chi)$. By the periodicity of χ, $L(s,\chi)$ can be written as a linear combination of Hurwitz zeta functions.

$$L(s,\chi)=\sum_{a=1}^{q-1}\chi(a)\sum_{n=0}^{\infty}\frac{1}{(a+nq)^s}=\sum_{a=1}^{q-1}\chi(a)\zeta(s,a,q).$$

Note that $\zeta(s,a,q)=\frac{1}{q^s}\zeta(s,\frac{a}{q})$ where $\zeta(s,\alpha)=\sum_{n=0}^{\infty}\frac{1}{(n+\alpha)^s}$ is the Hurwitz zeta function. Because of this expansion, it suffices to compute the Hurwitz zeta function. The Hurwitz zeta function can be computed by applying the Euler-Maclaurin summation formula. Assume f is a reasonably nice function (C^∞). We would like to estimate $\sum_{n=a}^{b} f(n)$. Euler-Maclaurin describes precisely the difference between this sum and the integal $\int_a^b f(t)\,dt$. By partial integration, it can be shown that

$$\sum_{n=a}^{b} f(n)=\int_a^b f(t)\,dt+\frac{f(a)+f(b)}{2}+\sum_{k=1}^{\nu}\frac{B_{2k}}{(2k)!}f^{2k-1}(x)|_a^b+R_{2\nu}$$

where B_{2k} is the sequence of Bernoulli numbers and $R_{2\nu}$ is a small remainder term, if ν is chosen appropriately. For a single evaluation of $L(s,\chi)$ it suffices to apply Euler-Maclaurin to each of the $\phi(q)$ Hurwitz zeta functions. For multiple evaluations, a slightly different technique is required. Rumely computes the Taylor series of $L(s,\chi)$ at equally spaced points $s=\frac{1}{2}+it$ on the critical line. He chooses $t=0,\frac{1}{2},1,\frac{3}{2},\ldots$. At a given point $s_0=\frac{1}{2}+it_0$ there is the Taylor series expansion

$$L(s,\chi)=\sum_{n=0}^{\infty}a_n(s_0)(s-s_0)^n$$

and it is the Taylor series coefficients that need to be computed. For real numbers $|t-t_0|\le 1/4$ the Taylor series is used to compute $L(\frac{1}{2}+it,\chi)$. This Taylor-series is approximated by the truncation $\sum_{n=0}^{N}a_n(s_0)(s-s_0)^n$ where N will be chosen appropriately. Rumely [63] p. 419 gives the numerical criterion

Proposition 4.2.1 *Suppose* $\mathrm{Re}(s_0)=\frac{1}{2}$. *In order for* $\sum_{n=0}^{N}a_n(s_0)(s-s_0)^n$ *to approximate* $L(s,\chi)$ *within* 10^{-20} *on* $D(s_0,\frac{1}{2})=\{s\in\mathbb{C}\mid |s-s_0|\le\frac{1}{2}\}$, *it suffices that* N *be large enough that* $P(\lambda)/(11\cdot 12^N)\le 10^{-20}$ *where* $P(x)=0.303x^3+0.221x^2+0.605x+0.687$ *and* $\lambda=\lfloor q/2\pi\rfloor\lfloor|t|+4\rfloor$.

From complex analysis, we know that $a_n(s_0)=L^{(n)}(s_0,\chi)/n!$. Differentiating the Dirichlet series of $L(s,\chi)$ for $\mathrm{Re}(s)>1$ leads to the formula

$$a_n(s_0)=\frac{1}{n!}\sum_a\chi(a)\zeta^{(n)}(s,a,q).$$

For $\mathrm{Re}(s) > 1$ term-wise differentiation yields

$$(-1)^n \zeta^{(n)}(s,a,q) = \sum_{m=0}^{\infty} \frac{\log(a+mq)^n}{(a+mq)^s}.$$

To evaluate the functions $\zeta^{(n)}(s,a,q)$ in the region $\mathrm{Re}(s) \leq 1$ it suffices to apply partial summation to the functions

$$f_n(s;x,a,q) = \frac{\log(a+xq)^n}{(a+xq)^s}.$$

In applying Euler-Maclaurin this function is considered a function of x with the variables s, a, and q regarded as fixed. In addition, define

$$f_n^{(k)}(s;x,a,q) = (d/dx)^k \left(\frac{\log(a+xq)^n}{(a+xq)^s} \right).$$

A direct application of Euler-Maclaurin shows that for $\mathrm{Re}(s) > 1 - 2L$

$$\begin{aligned}(-1)^n \zeta^{(n)}(s,a,q) \approx & \sum_{m=0}^{M-1} \frac{\log(a+mq)^n}{(a+mq)^s} + \frac{1}{2}\frac{\log(a+Mq)^n}{(a+Mq)^s} + \int_M^{\infty} f_n(s;x,a,q)dx \\ & - \sum_{\nu=1}^{L} \frac{B_{2\nu}}{(2\nu)!} f_n^{(2\nu-1)}(s;M,a,q)\end{aligned} \tag{4.1}$$

where M and L are integers yet to be specified. Set $I_n = I_n(s;M,a,q)$ to be the analytic continuation of $\int_M^{\infty} f_n(s;x,a,q)dx$. It would also be possible to write down an exact formula for the error. Rumely notes that I_n can be evaluated using the recurrence

$$I_0 = \frac{1}{q(s-1)(a+Mq)^{s-1}} \text{ and } I_n = \log(a+Mq)^n I_0 + n/(s-1) I_{n-1}.$$

Likewise, the derivatives f_n^k are computed using the recurrence

$$f_n^{(k)}(s;x,a,q) = nq f_{n-1}^{(k-1)}(s+1;x,a,q) - sq f_n^{(k-1)}(s;x,a,q).$$

We now need to choose M and L to ensure our approximation is sufficiently good. Again we will skip the details and refer to Rumely's paper. Rumely uses the value $M \approx 1.3(|s| + 40)$. He gives the following criterion for choosing L [63] p. 421.

Proposition 4.2.2 *Suppose* $\mathrm{Re}(s_0) = \frac{1}{2}$. *To compute* $\zeta^{(n)}(s_0;a,q)$ *accurately enough that for* $s \in D(s_0, \frac{1}{4})$ *each term* $a_n(s_0)(s-s_0)^n$ *contributes an error at most* 10^{-20} *to the sum* $\sum_{n=0}^{N} a_n(s_0)(s-s_0)^n$, *it is enough to choose* M *and* L *so that* $(2L - \frac{1}{2})\log(a+Mq) \geq 2N$, *and so that for* $0 \leq n \leq N$,

$$\frac{1}{2L-\frac{1}{2}} \cdot \frac{|B_{2L}|}{(2L)!} \cdot \frac{(|s|+n)\cdots(|s+2L-1|+n)}{(M+a/q)^{2L}} \leq \frac{10^{-20} \cdot q^{\frac{1}{2}} \cdot 4^n \cdot n!}{2 \cdot \phi(q) \cdot \log(a+Mq)^n \cdot (M+a/q)^{\frac{1}{2}}}.$$

For further details on this technique, see Rumely's paper [63]. Another interesting article is by Deuring [14]. His article contains the Riemann-Siegel formula for Dirichlet L-functions. Perhaps, the remainder terms can be computed and bounds could be found for the error terms. This would greatly speed up the computation of Dirichlet L-functions.

4.2.3 Zero searching strategy

The strategy to find zeros of Dirichlet L-functions is the Gram point method. Let $L(s,\chi)$ be a fixed real Dirichlet character of conductor q. Set $\delta = (1-\chi(-1))/2$. Note that if we set $\Lambda(s,\chi) = \left(\frac{q}{\pi}\right)^{s/2}\Gamma(\frac{s+\delta}{2})L(s,\chi)$ then the functional equation can be written as

$$\Lambda(s,\chi) = \Lambda(1-s,\chi)\ .$$

If we write

$$Z(t) = e^{i\theta(t)}L(\frac{1}{2}+it)$$

where

$$\theta(t) = \frac{t}{2}\log(q/\pi) + \mathrm{Im}(\log(\Gamma(\frac{1}{4}+\frac{\delta}{2}+i\frac{t}{2})))$$

then it follows from the functional equation that $Z(t)$ is a real-valued function for t real. Hence, real zeros of Z correspond to zeros of $L(s,\chi)$ on the critical line. Gram discovered a remarkable heuristic to speed up searches for zeros. Define the j-th Gram point g_j to satisfy

$$\theta(g_j) = j\pi$$

for $j \geq 0$. One should note that $\theta(t)$ is an increasing function. This can be observed by applying Stirling's formula to the second sum. Gram observed that quite frequently

$$(-1)^j Z(g_j) > 0.$$

Define g_j to be a good Gram point if the above condition is satisfied. Otherwise, g_j is a bad Gram point. If there are two consecutive good Gram points g_j and g_{j+1} , then Z has an opposite sign at these points. This guarantees at least one zero in the interval (g_j, g_{j+1}). The above condition is known as Gram's Law. For the zeta function this is true roughly seventy percent of the time. Rosser discovered another heuristic to take care of the case when Gram's Law fails (when a good Gram point is followed by a bad Gram point). We call a set of Gram points

$$\{g_j, g_{j+1}, \ldots, g_{j+k-1}, g_{j+k}\}$$

a Gram block of length k if g_j and g_{j+k} are good Gram points and all the interior Gram points are bad. Rosser observed that in a Gram block of length k there appears to be exactly k zeros of the L-function. One should note that there are at least $k-2$

zeros in a Gram block of length k. This is since all interior gram points are bad and each satisfies $(-1)^j Z(g_j) < 0$. Hence, the consecutive bad Gram points contain zeros in between them. Thus, one needs to find the 2 stray zeros. They usually can be found in one of the outer Gram intervals. Yet, this is not always the case. Applying these two observations, one can find zeros faster than randomly searching for zeros.

For the zeta function, the most common zero searching routine used is Newton's method. However, Newton's method requires the computation of the derivative of Z also. For the zero searches in this thesis the methods used only depended on knowing the values of the function Z. Some of these methods can be found in the programming book Numerical Recipes in C [57] .

4.3 Artin L-functions

Zeros of Artin L-functions were originally computed by Lagarias and Odlyzko [44]. In their article, they use a classical expansion of the Artin L-functions into sums of incomplete gamma functions. They computed the incomplete gamma functions using a formula obtained by partial integration. Unfortunately, their method allowed only the computation of a few small zeros of the Artin L-functions. In Rubinstein's recent thesis, he generalizes the work of Lagarias-Odlyzko. Using two other techniques to compute incomplete gamma functions, it is now possible to compute values of Artin L-functions (and many other L-functions) higher into the critical strip.

4.3.1 Computing coefficients

S_3 examples

Most Artin L-functions we are considering are attached to a weight one modular form. For the S_3 examples, the modular forms considered are the difference of two theta functions. For the primes $l = 23, 31, 59, 83, 107, 139, 211, 283,$ and 307 we considered the modular forms

$$f_l(z) = \frac{1}{2}\left(\sum_{m,n\in\mathbb{Z}} q^{Q_{1,l}(m,n)} - \sum_{m,n\in\mathbb{Z}} q^{Q_{2,l}(m,n)}\right).$$

where $Q_{i,l}$ are binary quadratic forms of discriminant $-l$. To compute the coefficients of f_l a simple program was written to compute the coefficients of an arbitrary theta series. Computing coeffcients of each of the above theta series yields the coefficients of f_l. Listed below is a table of the corresponding binary quadratic forms.

A second way of computing the coefficients of f_l is to realize that it is attached to a certain Artin L-function by the Serre-Deligne theorem. Note that each of the above primes l has class number three. Hence, the corresponding Hilbert class fields of $\mathbb{Q}(\sqrt{-l})$ are of degree six over the rationals with Galois group S_3. Consequently we guess that the corresponding Artin L-function is the two dimensional irreducible representation of $\mathrm{Gal}(H_l/\mathbb{Q})$. By the Langlands-Weil theorem, this Artin L-function corresponds to some weight one modular form with level equal to its Artin conductor. Hence, we can check that the first few coefficients of each modular form agree to show equality of modular forms. It suffices to check that the first B coefficients agree, where

$$B \geq \frac{N}{12} \prod_{p|N} (1 + p^{-1}).$$

and N is the level of the modular form. The polynomial generators for each H_l are listed below.

l	$Q_{1,l}$	$Q_{2,l}$	$q_l(x)$	$disc(q_l)$
23	[1,1,6]	[2,1,3]	$x^3 - x - 1$	-23
31	[1,1,8]	[2,1,4]	$x^3 + x - 1$	-31
59	[1,1,15]	[3,1,5]	$x^3 + 2x - 1$	-59
83	[1,1,21]	[3,1,7]	$x^3 - x^2 + x - 2$	-83
107	[1,1,27]	[3,1,9]	$x^3 - x^2 + 3x - 2$	-107
139	[1,1,35]	[5,1,7]	$x^3 + x^2 + x - 2$	-139
211	[1,1,53]	[5,3,11]	$x^3 - 2x - 3$	-211
283	[1,1,71]	[7,5,11]	$x^3 + 4x - 1$	-283

where $[a, b, c]$ refers to the binary quadratic form $ax^2 + bxy + cy^2$. Using the above polynomials, we can compute the coefficients of the corresponding Artin L-functions quite easily. Note that S_3 has the three conjugacy classes,

$$C_1 = \{1\}, C_2 = \{(12), (13), (23)\}, C_3 = \{(123), (132)\}.$$

One of the irreducible two dimensional representation of S_3 sends

$$(12) \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } (123) \to \begin{pmatrix} e^{2\pi i/3} & 0 \\ 0 & e^{-2\pi i/3} \end{pmatrix}.$$

It follows from the definition of an Artin L-function that for p unramified

$$\sigma_p = C_1 \Longleftrightarrow L_p(s, \rho) = (1 - p^{-s})^{-2} = \sum_{m=0}^{\infty} \frac{(m+1)}{p^{ms}} \Longleftrightarrow a_{p^m} = m + 1 \ ,$$

$$\sigma_p = C_2 \Longleftrightarrow L_p(s, \rho) = (1 - p^{-2s})^{-1} = \sum_{m=0}^{\infty} \frac{1}{p^{2ms}} \Longleftrightarrow a_{p^m} = \begin{cases} 1 \text{ if } m \equiv 0 \pmod 2 \\ 0 \text{ if } m \equiv 1 \pmod 2 \end{cases} \ ,$$

$$\sigma_p = C_3 \Longleftrightarrow L_p(s,\rho) = (1+p^{-s}+p^{-2s})^{-1} = \sum_{m=0}^{\infty} \frac{\left(\frac{m+1}{3}\right)}{p^{ms}} \Longleftrightarrow a_{p^m} = \left(\frac{m+1}{3}\right) .$$

Lastly, it is possible to show in each of these cases that

$$L_{ram}(s,\rho_l) = (1-l^{-s})^{-1} = \sum_{m=0}^{\infty} \frac{1}{l^{ms}}.$$

Hence we can compute the coefficients of the Artin L-functions at all prime powers. Applying multiplicativity, we wrote a short Maple program to compute the non-prime-power coefficients.

D_4 example

For $f_{144}(z)$, we used the following expansions mentioned in Serre [69].

$$f_{144}(z) = \sum (-1)^n q^{m^2+n^2}$$

where the final sum is over pairs $(m,n) \in \mathbb{Z} \times \mathbb{Z}$ such that:

$$m \equiv 1 \,(\mathrm{mod}\ 3),\ n \equiv 0 \ (\mathrm{mod}\ 3),\ m+n \equiv 1 \,(\mathrm{mod}\ 2).$$

H_8 example

The final two Artin L-functions that are computed are of the group H_8. The first example is Serre's example of an Artin L-function with a zero at $s = \frac{1}{2}$. The polynomials we consider are

$$f(x) = x^8 - 205x^6 + 13940x^4 - 378225x^2 + 3404025$$

and

$$f(x) = x^8 - 24x^6 + 144x^4 - 288x^2 + 144.$$

Each has a splitting field with Galois group H_8. Recall that that H_8 has the five conjugacy classes

$$C_1 = \{1\},\ C_2 = \{-1\},\ C_3 = \{\pm i\},\ C_4 = \{\pm j\},\ C_5 = \{\pm k\}.$$

One of the irreducible representations sends

$$i \to \begin{pmatrix} e^{2\pi i/4} & 0 \\ 0 & e^{-2\pi i/4} \end{pmatrix} \text{ and } j \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Define the conjugacy set $D = C_3 \cup C_4 \cup C_5$. From the definition of the Artin L-function we see that

$$\sigma_p = C_1 \Longleftrightarrow L_p(s,\rho) = (1-p^{-2s})^{-1} = \sum_{m=0}^{\infty} \frac{1}{p^{2ms}} ,$$

$$\sigma_p = C_2 \Longleftrightarrow L_p(s,\rho) = (1+p^{-s})^{-2} = \sum_{m=0}^{\infty} \frac{(-1)^m(m+1)}{p^{ms}} \ ,$$

$$\sigma_p = C_3 \Longleftrightarrow L_p(s,\rho) = (1-p^{-s})^{-2} = \sum_{m=0}^{\infty} \frac{(m+1)}{p^{ms}} \ .$$

Finally, it is possible to show that $L_{ram}(s,\rho) = 1$ in each of these cases. An easy way to observe this is to note that no subgroup of H_8 fixes a non-trivial subspace of $V = \mathbb{C}^2$. We can detect which conjugacy class σ_p lies in by reducing the polynomials f mod p. We have,

$$\sigma_p = C_1 \Longleftrightarrow f \equiv f_1 f_2 f_3 f_4 f_5 f_6 f_7 f_8 \bmod p$$

where each f_i is linear,

$$\sigma_p = C_2 \Longleftrightarrow f \equiv f_1 f_2 f_3 f_4 \bmod p$$

where each f_i is quadratic, and

$$\sigma_p = C_3 \Longleftrightarrow f \equiv f_1 f_2 \bmod p \ .$$

where each f_i is quartic irreducible. Maple has a function called `Factors` that produces an array of information about a polynomial reduced mod p. For example, it gives the degrees of the reduced polynomials. This is all that is needed in determining σ_p. We wrote a short Maple program that computes the coefficients at the prime powers and then applies multiplicativity to obtain all of the coefficients. Lastly, one can show that

$$\#\{n \le x \mid a(n) \ne 0\} \sim \frac{\alpha x}{(\log x)^{\frac{3}{4}}}$$

where $\alpha = c^{\frac{1}{4}}/\Gamma(\frac{1}{4})$ and

$$c = \left(\frac{2^{11}}{5 \cdot 41^3} \log\left(\frac{1+\sqrt{5}}{2}\right) \log(32+5\sqrt{41}) \log\left(\frac{43+3\sqrt{205}}{2}\right) \prod_{\sigma_p \in D} \left(\frac{1}{1-p^{-2}}\right)^2 \right).$$

This is similar to some examples in Serre's article [70]. The constant c was computed using Maple. The infinite product was approximated by a truncation. Note that we have the estimate

$$\left| \prod_{\sigma_p \in D} \left(\frac{1}{1-p^{-2}}\right)^2 - \prod_{\sigma_p \in D,\, p \le X} \left(\frac{1}{1-p^{-2}}\right)^2 \right| \ll \frac{\log X}{X}$$

which follows by multiplying out the products and then partially summing (one uses $\sum_{n \le x} d(n) \sim x \log x$). This shows that the convergence is rather slow and our calculation of the constant is only accurate to a few digits. Taking $X = 10^6$ we found that $\alpha = 0.159198....$ Note that $\frac{\log 10^6}{10^6} = 1.38.. \cdot 10^{-5}$.

4.3.2 Special values

In this thesis, we needed to compute $L(1, \rho)$ and $L'(1, \rho)$ for various degree two Artin L-functions. Because of Langlands' theorem, in most instances we only need to consider how to compute $L(1, f)$ and $L'(1, f)$ for certain weight one newforms f. Recall that if f is a weight one modular form of $S_1(N, \epsilon)$ then

$$f\left(\frac{az+b}{cz+d}\right) = \epsilon(d)(cz+d)f(z)$$

for all integers a, b, c, and d with $ad - bc = 1, N|d$. Assume that f has Fourier expansion at $i\infty$, $f(z) = \sum_{n=1}^{\infty} a_n q^n$ where $q = e^{2\pi i z}$. Set

$$\alpha_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}.$$

In addition, assume $f|[\alpha_N] = Ci^{-1}f$ where $C = \pm 1$.

From Hecke's proof of the analytic continuation of $L(s, f)$ one can derive

$$L(1, f) = 2\pi \int_{1/\sqrt{N}}^{\infty} f(iy)\left(1 + \frac{C}{\sqrt{N}y}\right) dy \ .$$

Truncating the Fourier expansion of f and integrating allows us to compute $L(1, f)$. Likewise, we can derive a similar formula for the derivative. Namely,

$$L'(1, f) = (\gamma + \log(2\pi))L(1, f) + 2\pi \int_{1/\sqrt{N}}^{\infty} f(iy)\left(\log(y) - \frac{C\log(Ny)}{\sqrt{N}y}\right) dy \ .$$

Let $i_1 = \int_{1/\sqrt{N}}^{\infty} f(iy)\left(1 + \frac{C}{\sqrt{N}y}\right) dy$ and $i_2 = \int_{1/\sqrt{N}}^{\infty} f(iy)\left(\log(y) - \frac{C\log(Ny)}{\sqrt{N}y}\right) dy$. Using Maple, these integrals were computed by direct integration and substituting a truncated sum for the modular form.

f	i_1	i_2
f_{23}	0.0586341644730512351395	−0.0819722040214841519829
f_{31}	0.06865324409055009927331	−0.1039653360910070385616
f_{59}	0.1029775695713633351984	−0.1820415574436797577441
f_{83}	0.1142308530543222459635	−0.2152571060757523997243
f_{107}	0.1214275864957923711948	−0.2386720934323237854665
f_{144}	0.1385738242548850896377	−0.2874114275894826594421
f_{139}	0.1411008710454903962602	−0.2921655468432146614807
f_{211}	0.1540449997868547133077	−0.3385555664397623716576
f_{283}	0.166602309670010402045	−0.381923355949097494643
f_{307}	0.1688304386204930431932	−0.390870618588258866212

f	$L(1,f)$	$L'(1,f)$
f_{23}	0.3684093207158268211117	0.3746961247030549165233
f_{31}	0.4313610545599581555372	0.3885434752553471475272
f_{59}	0.6470271520998537745988	0.4188297329709176310 47
f_{83}	0.7177336175375079251374	0.3808929565711321761 97
f_{107}	0.7629520273566409964743	0.3429789049201517899 96
f_{139}	0.8865629197832667947139	0.3054013922289413689 51
f_{144}	0.8706850165179802030771	0.2969257957081189562 33
f_{211}	0.9678932793052480651984	0.2103446628165771568 23
f_{283}	1.046793184260792898508	0.1284073919259211810 90
f_{307}	1.060792931344966901757	0.1060007701952448996 48

As in the last section, these special values are needed since they are connected to sums over the non-trivial zeros of the corresponding L-functions. For example, consider $f \in S_1(N, \epsilon)$ and ϵ an odd Dirichlet character such that $\epsilon(-1) = -1$. If we define

$$\Lambda(s,f) = \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma(s)L(s,f)$$

then the completed function satisfies $\Lambda(s,f) = i\Lambda(s,f')$. Moreover, a calculation similar to the Rubinstein-Sarnak article gives the formula

$$R_f = \frac{1}{2}\sum_\gamma \frac{1}{\frac{1}{4}+\gamma^2} = \log\left(\frac{\sqrt{N}}{2\pi}\right) - \gamma + \frac{L'}{L}(1,f)\ .$$

Using this formula we computed the following table of values.

f	R_f
f_{23}	0.1697191023099657484203
f_{31}	0.2026392921406046646806
f_{59}	0.2709899891529447741350
f_{83}	0.3250160442768871524576
f_{107}	0.3708635871434746226282
f_{144}	0.4108394214719667710376
f_{139}	0.3966222022594503086794
f_{211}	0.4781585002918207182 43
f_{283}	0.5302981114684091836101
f_{307}	0.5482571196108368083953

4.3.3 Computing Artin L-functions

Lagarias and Odlyzko restricted their studies to Artin L-functions that were linear combinations of of the L-functions attached to certain quadratic forms. For example they would consider an Artin L-function having the form

$$L(s, \rho, K/\mathbb{Q}) = \sum_{j=1}^{n} c(\rho, Q_j) L(s, Q_j)$$

where $Q_1, Q_2, \ldots Q_n$ are quadratic forms of discriminant df^2 and $c(\rho, Q_j) \in \mathbb{C}$. Examples of these types of L-functions are listed in section 4.3.1. For such Artin L-functions Lagarias-0dlyzko [44] applied the following incomplete gamma function expansion.

Proposition 4.3.1 *Let $Q(x,y) = Ax^2 + Bxy + Cy^2$, with determinant $D = AC - B^2/4 > 0$, and let $\delta \in \mathbb{C}$ have $Re(\delta) > 0$. Set*

$$\Lambda(s, Q) = (\pi\delta)^{-s}\Gamma(s)L(s, Q) = \pi^{-s}\Gamma(s)L(s, \delta Q)\ .$$

Then

$$\begin{aligned}\Lambda(s, Q) = &-\frac{1}{s} - \frac{(\delta^2 D)^{-\frac{1}{2}}}{1-s} \\ &+ \sum_{(x,y)\neq 0} \left(G(s, \pi\delta Q(x,y)) + (\delta^2 D)^{-\frac{1}{2}} G(1-s, \pi\delta^{-1}D^{-1}Q(x,y)) \right)\end{aligned} \tag{4.2}$$

where

$$G(s, \alpha) = \alpha^{-s}\Gamma(s, \alpha) = \int_1^\infty t^{s-1}e^{-\alpha t}dt,$$

valid for $s \in \mathbb{C}$ and $\mathrm{Re}(\alpha) > 0$.

Note that $\Gamma(s, \alpha)$ is the incomplete gamma function, traditionally defined by the contour integral

$$\Gamma(s, \alpha) = \int_\alpha^\infty t^{s-1}e^{-t}dt,$$

where the contour is required to stay within the region $|\arg(t)| < \pi/2$. This proposition shows we can compute the L-functions attached to a quadratic form if we have some technique to compute the incomplete gamma functions $G(s, \delta)$ or $\Gamma(s, \delta)$. Originally, Lagarias and Odlyzko only applied this formula with the parameter δ real and $|s|$ small. Rubinstein shows in his thesis [60] that in order to apply the above formula for $|s|$ large one needs to choose δ to be an appropriate complex number. This will be explained shortly.

I will now discuss a particular example of a formula for L-functions that Rubinstein [60] pp. 62-69 derives. Consider

$$L(s) = \sum_{n=1}^{\infty} \frac{b(n)}{n^s}.$$

Assume this Dirichlet series converges absolutely in $\mathrm{Re}(s) > \sigma_1$. Complete this L-function to

$$\Lambda(s) = Q^s\Gamma(\alpha s + \lambda)L(s), \; Q \in \mathbb{R}^+, \alpha \in \{1/2, 1\}, \mathrm{Re}(\lambda) \geq 0.$$

In additon, assume that:

1. $\Lambda(s)$ has a meromorphic continuation to all of $\mathbb{C}$ with simple poles at $s_1, \ldots s_l$ and corresponding residues $r_1, \ldots r_l$.
2. (functional equation) $\Lambda(s) = w\overline{\Lambda(1-\overline{s})}$ for some $w \neq 0 \in \mathbb{C}$.
3. For any $\alpha \leq \beta$, $L(\sigma + it) \ll_{\alpha,\beta} \exp(t^A)$ for some $A = A(\alpha, \beta) > 0$, as $|t| \to \infty$, $\alpha \leq \sigma \leq \beta$.

Then we have the following expansion:

Proposition 4.3.2

$$\begin{aligned}\Lambda(s)\delta^{-s} = &\sum_{k=1}^{l} \frac{r_k\delta^{-s_k}}{s-s_k} + (\delta/Q)^{\lambda/\gamma}\sum_{n=1}^{\infty} b(n)n^{\lambda/\gamma}G(\gamma s + \lambda, (n\delta/Q)^{1/\lambda}) \\ &+ \frac{w}{\delta}(Q\delta)^{-\overline{\lambda}/\gamma}\sum_{n=1}^{\infty}\overline{b(n)}n^{\overline{\lambda}/\gamma}G(\lambda(1-s)+\overline{\lambda}, (n/(\delta Q))^{1/\gamma}).\end{aligned} \tag{4.3}$$

Using this formula, with an appropriately chosen $\delta \in \mathbb{C}$ allows one to make a single evaluation of this L-function. In addition, we need some way to compute the functions $G(s, \alpha)$. In some papers, the $G(s, \alpha)$ are computed by direct numerical integration. However, this proves unwieldy when $|s|$ becomes large. Another recent approach was to use certain continued fraction expansions for the $G(s, \alpha)$. Rubinstein uses three different classical expansions to compute $G(s, \alpha)$ depending on various cases for the sizes of $|s|$ and $|\alpha|$. We also need to define the following related functions

$$\begin{aligned}\gamma(s,\alpha) &:= \Gamma(s) - \Gamma(s,\alpha) = \int_0^\alpha e^{-x}x^{\alpha-1}dx, \; \mathrm{Re}(s) > 0, \; |\mathrm{arg}\alpha| < \pi \\ g(s,\alpha) &= \alpha^{-s}\gamma(s,\alpha) = \int_0^1 e^{-\alpha t}t^{s-1}dt.\end{aligned} \tag{4.4}$$

Here are the three expressions Rubinstein uses to compute the various incomplete gamma functions.

$$\begin{aligned}g(s,\alpha) &= e^{-\alpha}\sum_{j=0}^{M-1}\frac{\alpha^j}{(s)_{j+1}} + R_M(s,\alpha), \; \mathrm{Re}(s) > -M \;, \\ \gamma(s,\alpha+d) &= \gamma(s,\alpha) + \alpha^{s-1}e^{-\alpha}\sum_{j=0}^{\infty}\frac{(1-s)_j}{(-\alpha)_j}(1-e^{-d}e_j(d)), \; |d| < |\alpha| \;, \\ G(s,\alpha) &= \frac{e^{-\alpha}}{\alpha}\sum_{j=0}^{M-1}\frac{(1-z)_j}{(-w)_j} + \epsilon_M(s,\alpha) \;,\end{aligned} \tag{4.5}$$

where

$$
\begin{aligned}
(z)_j &= \begin{cases} z(z+1)\cdots(z+j-1) \text{ if } j>0 \\ 1 \text{ if } j=0 \end{cases} , \\
R_M(s,\alpha) &= \frac{\alpha^M}{(s)_M} g(s+M,\alpha) , \\
e_j(d) &= \sum_{m=0}^{j} \frac{d^m}{m!} , \\
\epsilon_M(s,\alpha) &= \frac{(1-s)_M}{(-w)^M} G(s-M,\alpha) .
\end{aligned} \tag{4.6}
$$

Rubinstein gives conditions on s and α for when each of the above expansions is applied. For more details on the computation of the incomplete gamma functions see Chapter Three of Rubinstein's thesis [60].

4.3.4 Verifying the Riemann Hypothesis

There are at least two ways of verifying the Riemann Hypothesis to a given height. One technique requires applying the argument principle directly (see [76] p. 212). This requires evaluating the Artin L-function off the critical line. Instead, we will use an elegant method invented by Turing. This was originally used to check the Riemann Hypothesis for the zeta function. The details of this method can be found in [58] pp. 38-48 and [77] pp. 1313-1315. The Turing method can be developed for Artin L-functions. However, to save time we used a theorem by Tollis [77] that applies to Dedekind zeta functions. Consider the Artin L-function attached to the modular form f_{31}. Let $K = \mathbb{Q}(\theta)$ where θ is the real root of x^3+x-1. One can show that

$$\zeta_K(s) = \zeta(s)L(s, f_{31}).$$

Hence, if we have a way of computing the zeros of $\zeta(s)$, we can check RH. Glen Pugh generously supplied his programs to compute zeros of the zeta function. Using his programs we found the zeros of the zeta function. In fact, his programs also verify RH for the zeta function to a given height. We need some notation. Set $Q = \sqrt{31}/2\pi^{\frac{3}{2}}$. Complete the zeta function to

$$\Lambda_K(s) = Q^s\Gamma(s)\Gamma(\frac{s}{2})\zeta_K(s)$$

then we have the functional equation

$$\Lambda_K(1-s) = \Lambda_K(s).$$

Let $N_K(T)$ be the number of zeros satifying $0 \le \mathrm{Re}(s) \le 1$ and $0 \le \mathrm{Im}(s) \le T$. Define

$$\theta_K(t) = t \cdot \log Q + \mathrm{Im}\log\Gamma(\frac{1}{2}+it) + \mathrm{Im}\log\Gamma(\frac{1}{4}+i\frac{t}{2})$$

and $Z_K(t) = \exp(i\theta_K(t))\zeta_K(\frac{1}{2}+it)$. Note that $Z_K(t)$ is a real-valued function. A Gram point, g_j is a number that satisfies

$$\theta_K(g_j) = j\pi \ .$$

Let ϵ be a number satisfying $0 < \epsilon < \frac{1}{2}$. An ϵ-Gram block consists of numbers,

$$g_n, g_{n+1}, \ldots, g_{n+l-1}, g_l$$

such that

$$(-1)^j Z_K(g_j) > 0 \text{ for } j = n \text{ and } j = n+l,$$
$$(-1)^j Z_K(g_j) < 0 \text{ for } n < j < n+l,$$

and

$$|\theta_K(g_j) - j\pi| \le \epsilon.$$

Tollis [77] p. 1314 shows the following theorem.

Theorem 4.3.1 *Assume that $g_n > 40$ and assume also that the interval (g_n, g_{n+l}) is the union of p disjoint ϵ-Gram blocks, each containing at least as many zeros of $Z_K(t)$ as its length. If*

$$\begin{aligned} p + (\frac{1}{2} - \epsilon)l > (0.2928 \cdot N_K + 0.0419) \log\left(|D_K|(\frac{g_{n+l}}{2\pi})^3\right) \\ + 0.0195 \log^2\left(|D_K|(\frac{g_{n+l}}{2\pi})^3\right) \end{aligned} \tag{4.7}$$

then $N_K(g_n) \le n+1$ and $N_K(g_{n+l}) \ge n+l+1$.

We will now show how to apply the above theorem. Let $T = 1000$. For the function $L(s, f_{31})$ our program found 1842 zeros on the critical line to height T. For the zeta function $\zeta(s)$ there are precisely 649 zeros of $\zeta(s)$ to this height. We would like to show that $N_K(1000) = 649 + 1842 = 2491$. If this is true then RH is true for $L(s, f_{31})$ to height $T = 1000$. Let $\epsilon = 0.001$. Using our program, we found that

$$g_{2490} = 1000.123...$$

and is a good Gram point. We will take $n = 2490$ in the above theorem and by trial and error we chose $l = 30$. It turns out that

$$g_{2490+30} = g_{2520} = 1011.907...$$

There are $p = 21$ ϵ-Gram blocks in (g_n, g_{n+l}). Thus we obtain

$$\text{LHS} = 30.149...$$

and

$$\text{RHS} = 28.663...$$

From this we deduce that $N_K(g_{2490}) = 2491$ and there are 1842 zeros of $L(s, f_{31})$ to height $T = 1000$. This example shows how RH was verified for f_{23}, f_{31}, f_{59}, and f_{83}. Lastly, the function f_{144} can also be treated in this way. In this case, we let K be the splitting field of $x^4 - 12$. We have the identity

$$\zeta_K(s) = \zeta(s) L(s, (\frac{-4}{\cdot})) L(s, (\frac{12}{\cdot})) L(s, (\frac{-3}{\cdot})) L(s, f_{144})^2.$$

From this identity, we similarly apply Tollis' theorem and show RH for $L(s, f_{144})$ to a given height. This technique only works if we have programs to compute zeros of each of the L-functions on the right.

4.3.5 Zero data

Listed below is some data on the the Artin L-functions whose zeros were computed. T denotes the height to which they were computed. $N(T)$ is the number of zeros to height T. This has been verified by Tollis' theorem in the preceding section.

$L(s, \rho_p)$	T	$N(T)$	least zero
23	972.2	1736	5.1156833287
31	1071.83	1999	4.1662147526
59	793.8	1485	3.4318050042
83	737.33	1403	2.9160681285
144	667.25	1307	2.5166582692

Chapter 5
Chebyshev's Bias in Galois Groups

5.1 Existence of the limiting distribution

Let L/K be a normal extension of number fields. Let $G = \mathrm{Gal}(L/K)$ be the corresponding Galois group and C be a conjugacy class of G. We will now derive an explicit formula for $\frac{|G|}{|C|}\pi_C(x) - \pi_K(x)$. This formula explains when there should be a bias. As in the Rubinstein-Sarnak article, a bias can be caused by the behaviour of the squares of the Frobenius elements attached to each prime ideal. However, in this setting an additional bias term can arise in the case where the corresponding Artin L-functions vanish at the central point $s = \frac{1}{2}$. Consider the following functions

$$\psi_C(x) = \sum_{\mathbb{N}\mathfrak{p}^m \leq x,\, \sigma_{\mathfrak{p}}^m = C} \log(\mathbb{N}\mathfrak{p}) \text{ and } \theta_C(x) = \sum_{\mathbb{N}\mathfrak{p} \leq x,\, \sigma_{\mathfrak{p}} = C} \log(\mathbb{N}\mathfrak{p}) \ .$$

These are the natural analogues of $\psi(x; a, q)$ and $\theta(x; a, q)$ in the number field setting. By definition, it follows that

$$\psi_C(x) = \theta_C(x) + \sum_{\mathbb{N}\mathfrak{p}^2 \leq x,\, \sigma_{\mathfrak{p}}^2 = C} \log(\mathbb{N}\mathfrak{p}) + O(x^{\frac{1}{3}}) \ .$$

The error term is obtained by observing that

$$\theta_C(x) \leq \theta_K(x) = \sum_{\mathbb{N}\mathfrak{p} \leq x} \log(\mathbb{N}\mathfrak{p}) \ll x$$

by the prime ideal theorem. Let $C_1, C_2, \ldots, C_t$ be conjugacy classes such that $C_i^2 \subset C$ for $i = 1 \ldots t$. Let $\mathrm{sq}^{-1}(C)$ be the conjugacy set defined by

$$\mathrm{sq}^{-1}(C) = \bigcup_{i=1}^{t} C_i \ .$$

Rearranging the above formula we obtain

$$\theta_C(x) = \psi_C(x) - \sum_{\mathbb{N}\mathfrak{p} \leq x^{\frac{1}{2}},\, \sigma_{\mathfrak{p}} \subset \mathrm{sq}^{-1}(C)} \log(\mathbb{N}\mathfrak{p}) + O(x^{\frac{1}{3}}) \ .$$

However, in proving an effective version of Chebotarev's density theorem we obtain

$$\sum_{\mathbb{N}\mathfrak{p}\leq x^{\frac{1}{2}},\, \sigma_{\mathfrak{p}}\subset \mathrm{sq}^{-1}(C)} \log(\mathbb{N}\mathfrak{p}) = \frac{|\mathrm{sq}^{-1}(C)|}{|G|} x^{\frac{1}{2}} + O(x^{\frac{1}{4}+\epsilon}) \ .$$

under the assumption of GRH. Substituting this into the previous equation and multiplying by $\frac{|G|}{|C|}$ we obtain

$$\frac{|G|}{|C|}\theta_C(x) = \frac{|G|}{|C|}\psi_C(x) - \frac{|\mathrm{sq}^{-1}(C)|}{|C|} x^{\frac{1}{2}} + O(x^{\frac{1}{3}}) \ . \tag{5.1}$$

Now we make the following observation from representation theory. Let $\mathrm{Irr}(G)$ denote the set of irreducible characters of G. Let s, t be elements of G and $C = C(s)$ the conjugacy class of s. There is the following orthogonality formula for group characters.

$$\frac{|C|}{|G|}\sum_{\chi\in\mathrm{Irr}(G)} \overline{\chi(s)}\chi(t) = \begin{cases} 1 \text{ if } t \in \ C \\ 0 \text{ otherwise} \end{cases} .$$

Applying this formula we obtain

$$\psi_C(x) = \sum_{\mathbb{N}\mathfrak{p}^m\leq x, \sigma_{\mathfrak{p}}^m\subset C} \log(\mathbb{N}\mathfrak{p}) = \sum_{\mathbb{N}\mathfrak{p}^m\leq x} \log(\mathbb{N}\mathfrak{p}) \left(\frac{|C|}{|G|} \sum_{\chi\in\mathrm{Irr}(G)} \overline{\chi(C)}\chi(\sigma_{\mathfrak{p}}^m) \right) \ .$$

Switching order of summation yields

$$\psi_C(x) = \frac{|C|}{|G|}\sum_{\chi\in\mathrm{Irr}(G)} \overline{\chi(C)}\psi(x,\chi) \ . \tag{5.2}$$

This formula is very useful since the functions $\psi(x, \chi)$ are intimately related to the logarithmic derivatives of the Artin L-functions. Applying (5.1) and (5.2) we obtain

$$\frac{|G|}{|C|}\theta_C(x) = \sum_{\chi\in\mathrm{Irr}(G)} \overline{\chi(C)}\psi(x,\chi) - \frac{|\mathrm{sq}^{-1}(C)|}{|C|} x^{\frac{1}{2}} + O(x^{\frac{1}{3}}) \ . \tag{5.3}$$

Recall that by definition of the Riemann-Stieltjes integral we also have

$$\pi_C(x) = \int_{2^-}^{x} \frac{d\theta_C(t)}{\log t} \ . \tag{5.4}$$

Hence combining (5.3) and (5.4) leads to

$$\begin{aligned}
\frac{|G|}{|C|}\pi_C(x) &= \sum_{\chi}\overline{\chi(C)}\int_{2^-}^{x}\frac{d\psi(t,\chi)}{\log t} - \frac{|\mathrm{sq}^{-1}(C)|}{|C|}\frac{x^{\frac{1}{2}}}{\log x} + O\left(\frac{x^{\frac{1}{2}}}{\log^2 x}\right)\\
&= \int_{2^-}^{x}\frac{d\psi(t,1)}{\log t} + \sum_{\chi\neq 1}\overline{\chi(C)}\frac{\psi(x,\chi)}{\log x} - \frac{|\mathrm{sq}^{-1}(C)|}{|C|}\frac{x^{\frac{1}{2}}}{\log x} + O\left(\int_{2^-}^{x}\frac{\psi(t,\chi)dt}{t\log^2 t} + \frac{x^{\frac{1}{2}}}{\log^2 x}\right)\\
&= \pi_K(x) + \frac{\sqrt{x}}{\log x} + O\left(\frac{x^{\frac{1}{3}}}{\log x}\right) + \sum_{\chi\neq 1}\overline{\chi(C)}\frac{\psi(x,\chi)}{\log x} - \frac{|\mathrm{sq}^{-1}(C)|}{|C|}\frac{x^{\frac{1}{2}}}{\log x}\\
&\quad + O\left(\int_{2^-}^{x}\frac{\psi(t,\chi)dt}{t\log^2 t} + \frac{x^{\frac{1}{2}}}{\log^2 x}\right)\ .
\end{aligned} \tag{5.5}$$

We have used the notation

$$\pi_K(x) = \#\{\mathfrak{p}\subset\mathcal{O}_K \mid \mathbb{N}\mathfrak{p}\leq x\ \}\ .$$

Therefore,

$$\begin{aligned}
\frac{|G|}{|C|}\pi_C(x) - \pi_K(x) &= \sum_{\chi\neq 1}\overline{\chi(C)}\frac{\psi(x,\chi)}{\log x} + \left(1 - \frac{|\mathrm{sq}^{-1}(C)|}{|C|}\right)\frac{x^{\frac{1}{2}}}{\log x}\\
&\quad + O\left(\int_{2^-}^{x}\frac{\psi(t,\chi)dt}{t\log^2 t} + \frac{x^{\frac{1}{2}}}{\log^2 x}\right)\ .
\end{aligned} \tag{5.6}$$

It now suffices to show that

$$\int_{2^-}^{x}\frac{\psi(t,\chi)dt}{t\log^2 t} \ll O\left(\frac{x^{\frac{1}{2}}}{\log^2 x}\right)\ . \tag{5.7}$$

The proof of this statement is completely analogous to [62] p. 179. It is possible to show, assuming Artin's conjecture on the holomorphy of the Artin L-function $L(s,\chi)$ that

$$\psi(x,\chi) = \delta(\chi) - \sum_{|\gamma_\chi|\leq X}\frac{x^\rho}{\rho} + O_{K,L}\left(\frac{x\log^2(xX)}{X} + \log x\right) \tag{5.8}$$

where $\rho = \beta_\chi + i\gamma_\chi$ runs over the zeros of $L(s,\chi)$ in $0 < \mathrm{Re}(s) < 1$ and the implied constant in the O depends on the fields L and K. Assuming the RH for $L(s,\chi)$, we have $\beta_\chi = \frac{1}{2}$ and the preceding equation becomes

$$\psi(x,\chi) = -\sqrt{x}\sum_{|\gamma_\chi|\leq X}\frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + O_{K,L}\left(\frac{x\log^2(xX)}{X} + \log x\right)\ . \tag{5.9}$$

Let $G(x, \chi) = \int_{2^-}^{x} \psi(t, \chi)dt$. After subsituting (5.9), integrating, and letting $X \to \infty$, we obtain

$$G(x, \chi) = -\sum_{\gamma_\chi} \frac{x^{\frac{3}{2}+i\gamma_\chi}}{(\frac{1}{2}+i\gamma_\chi)(\frac{3}{2}+i\gamma_\chi)} + O(x \log x) \tag{5.10}$$

where the constant in the error term depends on field constants of L and K. The above sum is absolutely convergent in light of the following formula for the number of zeros of $L(s, \chi)$ in the critical strip. Applying the argument principle yields

$$\#\{|\gamma_\chi| \leq T\} \ll_{L,K} T \log T$$

where the implied constant depends on the fields L and K. This shows that $G(x, \chi) \ll_{L,K} x^{\frac{3}{2}}$. Thus, integrating the left hand side of (5.7) leads to the appropriate error term $O(\frac{x^{\frac{1}{2}}}{\log^2 x})$. We have established the following lemma.

Lemma 5.1.1 *Assume GRH. As $x \to \infty$ we have*

$$\begin{aligned} \frac{\log x}{\sqrt{x}} \left(\frac{|G|}{|C|}\pi_C(x) - \pi_K(x) \right) = \sum_{\chi \neq 1} \overline{\chi(C)} \frac{\psi(x,\chi)}{\sqrt{x}} + \left(1 - \frac{|\mathrm{sq}^{-1}(C)|}{|C|} \right) \\ + O\left(\frac{1}{\log x} \right) . \end{aligned} \tag{5.11}$$

Comment The above method can be used to obtain a formula to compare sets of primes whose Frobenius elements are in two different conjugacy classes. Using this formula, we can now compare two conjugacy classes C_1 and C_2. Define

$$E_{G;1,2}(x) = \frac{\log x}{x^{\frac{1}{2}}} \left(\frac{|G|}{|C_1|}\pi_{C_1}(x) - \frac{|G|}{|C_2|}\pi_{C_2}(x) \right).$$

We obtain

$$\begin{aligned} E_{G;1,2}(x) = \left(\frac{|\mathrm{sq}^{-1}(C_2)|}{|C_2|} - \frac{|\mathrm{sq}^{-1}(C_1)|}{|C_1|} \right) - \sum_{\chi \neq 1} (\overline{\chi(C_1)} - \overline{\chi(C_2)}) \sum_{|\gamma_\chi| \leq X} \frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} \\ + O\left(\frac{\sqrt{x}\log^2 X}{X} + \frac{1}{\log x} \right) \end{aligned} \tag{5.12}$$

where $2 \leq x \leq X$ and O depends on L and K. In the inner sum on the right, we are summing over zeros of a non-trivial Artin L-function. In the corresponding sum in the Rubinstein-Sarnak article, it is assumed that no terms arise from the vanishing of a Dirichlet L-function at $s = \frac{1}{2}$. In fact, this conjecture is widely accepted to be true.

There are results of Iwaniec-Sarnak [37] and Soundararajan [72] that show a large proportion of the Dirichlet L-functions do not vanish at the critical point. Contrary to the situation in [62], there are known examples of Artin L-functions vanishing at $s = \frac{1}{2}$. Armitage and Serre were among the first to find such examples. Because of the above comments, set

$$n_\chi = \mathrm{ord}_{s=\frac{1}{2}} L(s,\chi) \ .$$

Note that if Artin's holomorphy conjecture is true, then $n_\chi \geq 0$. In (5.12), removing the contributions coming from possible zeros at the critical point leads to the following bias factor

$$c_{\frac{1}{2}}(G,C) = 2\sum_{\chi\neq 1} \overline{\chi(C)}\ n_\chi \ . \tag{5.13}$$

Similarly, we will set $c_{\mathrm{sq}}(G,C) = \frac{|\mathrm{sq}^{-1}(C)|}{|C|}$. Therefore, we define the modified bias factor as

$$c(G,C) = c_{\mathrm{sq}}(G,C) + c_{\frac{1}{2}}(G,C) \ . \tag{5.14}$$

When the group is understood to be fixed, we will abbreviate $c(G,C)$ to $c(C)$, $c_{\mathrm{sq}}(G,C)$ to $c_{\mathrm{sq}}(C)$, and $c_{\frac{1}{2}}(G,C)$ to $c_{\frac{1}{2}}(C)$. The bias factor c consists of two terms. The first term is a result of the behaviour of the squares of the Frobenius substitutions. This is the classical Chebyshev bias. The second term reflects the possible vanishing of Artin L-functions. This bias factor is only valid for the two-way races. Otherwise, we consider the bias factor in the form

$$\alpha(G,C) = 1 - c(G,C).$$

We can now rewrite (5.12) as

$$\begin{aligned} E_{G;1,2}(x) =& c(G,C_2) - c(G,C_1) - \sum_{\chi\neq 1}(\overline{\chi(C_1)} - \overline{\chi(C_2)}) \sum_{0<|\gamma_\chi|\leq X} \frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} \\ &+ O\left(\frac{\sqrt{x}\log^2 X}{X} + \frac{1}{\log x}\right) \ . \end{aligned} \tag{5.15}$$

Notice that the inner sum on the right now excludes the possible zeros at $s = \frac{1}{2}$. Define the vector-valued function

$$E_{G;1,2,\ldots,r}(x) = \frac{\log x}{\sqrt{x}}\left(\frac{|G|}{|C_1|}\pi_{C_1}(x) - \pi_K(x), \ldots, \frac{|G|}{|C_r|}\pi_{C_r}(x) - \pi_K(x)\right) \ .$$

Also define

$$\begin{aligned} E(y) &= E_{G;1,2,\ldots,r}(e^y) = (E_{G;C_1}(e^y), \ldots, E_{G;C_r}(e^y)) \ , \\ E^{(T)}(y) &= (E_1^{(T)}(y), \ldots, E_r^{(T)}(y)) \text{ where } , \\ E_j^{(T)}(y) &= \alpha(G;C) - \sum_{\chi\neq 1}\overline{\chi}(C_j) \sum_{0<|\gamma_\chi|\leq T} \frac{e^{iy\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} \ , \ 1\leq j\leq r \ , \\ \epsilon^{(T)}(y) &= E(y) - E^{(T)}(y) = (\epsilon_1^{(T)}(y), \ldots, \epsilon_r^{(T)}(y)) \ . \end{aligned} \tag{5.16}$$

Theorem 5.1.2 *Assume GRH for the Artin L-functions and Artin's holomorphy conjecture. Then $E(y)$ has a limiting distribution $\mu_{1,2,\ldots,r}$ on $\mathbb{R}^r$, that is,*

$$\lim_{Y\to\infty} \frac{1}{Y}\int_2^Y f(E_{G;1,2,\ldots,r}(e^y))\ dy = \int_{\mathbb{R}^r} f(x) d\mu_{1,2,\ldots,r}(x)$$

for all bounded Lipschitz continuous functions on $\mathbb{R}^r$.

Proof This proof is the same as in the Rubinstein-Sarnak article and is only presented for completeness. We will specialize to $f : \mathbb{R}^r \to \mathbb{R}$, a Lipschitz function. This satisfies

$$|f(x) - f(y)| \leq c_f |x - y| \ .$$

We will need some lemmas.

Lemma 5.1.3 *For $T \geq 1$ and $Y \geq \log 2$,*

$$\int_{\log 2}^Y |\epsilon_i^{(T)}(y)|^2\ dy \ll Y\frac{\log^2 T}{T} + \frac{\log^3 T}{T} \ .$$

Proof See Rubinstein-Sarnak [62] p. 179 where the argument is identical.

The next step is to prove the existence of a probability measure for the approximation $E^{(T)}(y)$. We need the following result:

Lemma 5.1.4 *For each T there is a probability measure ν_T on $\mathbb{R}^r$ such that*

$$\nu_T(f) := \int_{\mathbb{R}^r} f(x)\ d\nu_T(x) = \lim_{Y\to\infty}\frac{1}{Y}\int_{\log 2}^Y f(E^{(T)}(y))\ dy$$

Proof For all $\chi \neq 1$ list the zeros $\frac{1}{2} + i\gamma_\chi$ of the $L(s,\chi)$ satisfying $0 < \gamma_\chi \leq T$ by size as $\gamma_1, \gamma_2, \ldots, \gamma_N$. We only need to consider the positive values since if χ is a real character, $L(\frac{1}{2} + i\gamma, \chi) = 0 \Leftrightarrow L(\frac{1}{2} - i\gamma, \chi) = 0$. Likewise, if χ is complex, then $L(\frac{1}{2} + i\gamma, \chi) = 0 \Leftrightarrow L(\frac{1}{2} - i\gamma, \overline{\chi}) = 0$. Using these symmetries we obtain

$$E^{(T)}(y) = 2\mathrm{Re}\left(\sum_{l=1}^N b_l e^{iy\gamma_l}\right) + b_0 \ ,$$

where $b_0, b_1, \ldots, b_N \in \mathbb{C}$ with

$$\begin{aligned} b_0 &= (\alpha(G;C_1), \ldots, \alpha(G;C_r)), \\ b_l &= -\left(\frac{\chi(C_1)}{\frac{1}{2}+\gamma_l}, \ldots, \frac{\chi(C_r)}{\frac{1}{2}+\gamma_l}\right) \ . \end{aligned} \tag{5.17}$$

Define the function $g(\theta_1, \ldots, \theta_N)$ on the N-torus $\mathbb{T}^N = \mathbb{R}^N/\mathbb{Z}^N$ by

$$g(\theta_1, \ldots, \theta_N) = f\left(2\mathrm{Re}\left(\sum_{l=1}^{N} b_l e^{2\pi i\theta_l}\right) + b_0\right) .$$

Observe that g is a continuous function on $\mathbb{T}^N$ and

$$f(E^{(T)}(y)) = g\left(\frac{\gamma_1 y}{2\pi}, \ldots, \frac{\gamma_N y}{2\pi}\right)$$

Let A be the topological closure in $\mathbb{T}^N$ of the one-parameter subgroup

$$\Gamma(y) := \{(\frac{\gamma_1 y}{2\pi}, \ldots, \frac{\gamma_N y}{2\pi}) \mid y \in \mathbb{R}\} .$$

Observe that A is a torus. Also, there is a normalized Haar measure on A, da. This is the canonical probability measure. The Kronecker-Weyl Theorem shows that $\Gamma(y)$ is equidistributed in A. Since $g|_A$ is continuous on A, we have the following equivalent formulation of $\Gamma(y)$ being equidistributed in A,

$$\lim_{Y\to\infty} \frac{1}{Y} \int_{\log 2}^{Y} f(E^{(T)}(y))\ dy = \int_A g(a)\ da .$$

This almost completes the proof. Observe that $g|_A = f \circ X_T|_A$ where $X_T : \mathbb{T}^N \to \mathbb{R}^r$ is defined by

$$X_T(\theta_1, \ldots, \theta_N) = 2\mathrm{Re}\left(\sum_{l=1}^{N} b_l e^{2\pi i\theta_l}\right) + b_0 .$$

Since $X_T|_A$ is a random variable defined on the probability space A, there is a canonical probability measure, ν_T, defined on $\mathbb{R}^r$. If B is a Borel subset of $\mathbb{R}^r$, then $\nu_T(B)$ is defined by

$$\nu_T(B) = a(X_T|_A^{-1}(B))$$

where a is the normalized Haar measure on A. It now follows from the standard change of variable formula from probability that

$$\int_A g(a)\ da = \int_{\mathbb{R}^r} f(x)\ d\nu_T(x) .$$

This completes the proof of the lemma. □
We can now complete the main theorem.

$$\begin{aligned}
\frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\ dy &= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y)+\epsilon^{(T)}(y))\ dy \\
&= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\ dy + O\left(\frac{c_f}{Y}\int_{\log 2}^{Y} |\epsilon^{(T)}(y)|\ dy\right) \\
&= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\ dy + O\left(\frac{c_f}{\sqrt{Y}}\left(\int_{\log 2}^{Y} |\epsilon^{(T)}(y)|^2\ dy\right)^{\frac{1}{2}}\right) \\
&= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\ dy + O\left(c_f\left(\frac{\log T}{\sqrt{T}} + \frac{\log^2 T}{Y\sqrt{T}}\right)\right)\ .
\end{aligned} \tag{5.18}$$

The above equalities are obtained by applying the Lipschitz condition, Cauchy-Schwarz, and Lemma 5.1.3. Letting $Y \to \infty$ and applying Lemma 5.1.4 yields

$$\begin{aligned}
\nu_T(f) - O\left(\frac{c_f \log T}{\sqrt{T}}\right) &\leq \liminf_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\ dy \\
&\leq \limsup_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\ dy \leq \nu_T(f) + O\left(\frac{c_f \log T}{\sqrt{T}}\right)\ .
\end{aligned} \tag{5.19}$$

Since T can be chosen arbitrarily large, we deduce that the lim sup and lim inf are equal. Therefore, the limit

$$\lim_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\ dy$$

exists. We will now refer to Helly's Theorems to construct a limiting Borel measure. The result states:

Lemma - Helly's Theorems 5.1.5
(i) Every sequence of $\{F_n(x)\}$ of uniformly bounded non-decreasing functions contains a subsequence $\{F_{n_k}(x)\}$ which converges weakly to some non-decreasing bounded function $F(x)$.
(ii) Let $f(x)$ be a continuous function and assume that $\{F_n(x)\}$ is a sequence of uniformly bounded, non-decreasing functions which converge weakly to some function $F(x)$ at all points of a continuity interval $[a, b]$ of $F(x)$, then

$$\lim_{n\to\infty} \int_a^b f(x)\ dF_n(x) = \int_a^b f(x)\ dF(x)\ .$$

Apply Lemma 5.1.5 (i) to $\nu_T(x)$, to obtain a countable sequence of distribution functions such that

$$\{\nu_{T_k} \mid k \geq 1\} \text{ such that } \nu_{T_k} \to \mu \text{ weakly}$$

where μ is a Borel measure on $\mathbb{R}^r$. By Lemma 5.1.5 (ii), this implies that

$$\lim_{k\to\infty} \int_B f(x)\ d\nu_{T_k}(x) = \int_B f(x)\ d\mu(x)$$

for f a continuous function and B a Borel set. Hence, by the above inequalities we observe that

$$\lim_{Y\to\infty} \frac{1}{Y} \int_{\log 2}^{Y} f(E(y))\ dy = \int_{\mathbb{R}^r} f(x)\ d\mu(x) = \mu(f)\ .$$

In addition, we have proven that the Borel measure μ satisfies

$$|\mu(f) - \nu_T(f)| \ll \frac{c_f \log T}{\sqrt{T}}$$

for all continuous Lipschitz functions f. Taking $f = 1$ shows that

$$|\mu(1) - 1| \ll \frac{\log T}{\sqrt{T}}$$

since ν_T are probability measures. Letting $T \to \infty$ implies that $\mu(1) = 1$ and we conclude that μ is also a probability measure. This completes the proof of the theorem in the case that f is bounded continuous and Lipschitz.

5.2 Applications of LI

In this section, we will derive an explicit formula for the Fourier transform of the probability measure μ.

Theorem 5.2.1 *Assuming LI, the Fourier transform of $\mu_{G;1,\ldots,r}$ can be explicitly computed to be*

$$\hat{\mu}(\xi) = \exp\left(-i\sum_{j=1}^{r} \alpha(G, C_j)\xi_j\right) \prod_{\chi\neq 1} \prod_{\gamma_\chi>0} J_0\left(\frac{2|\sum_{j=1}^{r} \chi(C_j)\xi_j|}{\sqrt{\frac{1}{4} + \gamma_\chi^2}}\right)\ .$$

Proof By definition the Fourier transform of μ is defined by

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^r} e^{-i<\xi,t>}\ d\mu(t)\ .$$

In the previous section, it was proven that $\nu_T \to \mu$. By Levy's Theorem, we have $\hat{\nu}_T \to \hat{\mu}$. We saw that ν_T is constructed from the canonical probability measure a on the torus A. However, the assumption of LI implies that $A = \mathbb{T}^N$. Let $dP_N = d\theta_1 \ldots d\theta_N$ denote Lebesgue measure on $\mathbb{T}^N$. Hence we see by the change of variable formula mentioned in the last section that

$$\hat{\nu}_T(\xi) = \int_{\mathbb{R}^r} e^{-i<\xi,t>}\, d\nu_T(t) = \int_{\mathbb{T}^N} e^{-i<\xi,X_T(\theta)>}\, dP_N(\theta)\ .$$

However, we can write $X_T(\theta) = (X_1(\theta), \ldots, X_r(\theta))$ where

$$X_i(\theta) = 2\mathrm{Re}\sum_{k=1}^{N} \frac{\overline{\chi}(C_i)}{\frac{1}{2}+i\gamma_j} e^{2\pi i\theta_k} + \alpha(G, C_i)\ .$$

This implies that

$$\begin{aligned}
&\int_{\mathbb{T}^N} e^{-i<\xi,X_T(\theta)>}\, dP_N(\theta) \\
&= \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\int_{\mathbb{T}^N} e^{-i\sum_{i=1}^{r}\xi_i X_i(\theta)}\, dP_N(\theta) \\
&= \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\int_{\mathbb{T}^N} e^{-i\sum_{i=1}^{r}\xi_i 2\mathrm{Re}\sum_{j=1}^{N}\frac{\overline{\chi}(C_i)}{\frac{1}{2}+i\gamma_j}e^{2\pi i\theta_j}}\, dP_N(\theta) \\
&= \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\int_{\mathbb{T}^N} e^{-i\sum_{j=1}^{N}\sum_{i=1}^{r}\xi_i 2\mathrm{Re}\left(\frac{\overline{\chi}(C_i)}{\frac{1}{2}+i\gamma_j}e^{2\pi i\theta_j}\right)}\, dP_N(\theta) \\
&= \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\prod_{j=1}^{N}\int_{\mathbb{T}} e^{-i<\xi,\Sigma_j>}\, d\theta_j \\
&= \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\prod_{j=1}^{N}\int_{\mathbb{R}^r} e^{-i<\xi,t>}\, d\mu_{\gamma_j}(t)
\end{aligned} \tag{5.20}$$

where Σ_j is the random vector defined on $\mathbb{T}$ by

$$\Sigma_j(\theta) = -\left(\frac{\overline{\chi}(C_1)e^{2\pi i\theta}}{\frac{1}{2}+i\gamma_j} + \frac{\chi(C_1)e^{-2\pi i\theta}}{\frac{1}{2}-i\gamma_j}, \ldots, \frac{\overline{\chi}(C_r)e^{2\pi i\theta}}{\frac{1}{2}+i\gamma_j} + \frac{\chi(C_r)e^{-2\pi i\theta}}{\frac{1}{2}-i\gamma_j}\right)$$

and

$$\mu_{\gamma_j}(t) = \mathrm{meas}(\theta \in \mathbb{T} \mid \Sigma_j(\theta) \le t)\ .$$

We have now shown that

$$\hat{\mu}(\xi) = \exp\left(-i\sum_{i=1}^{r}\alpha(G,C_i)\xi_i\right)\lim_{N\to\infty}\prod_{j=1}^{N}\hat{\mu}_{\gamma_j}(\xi)\ .$$

Notice that we can write

$$\frac{\overline{\chi}(C)e^{2\pi i\theta}}{\frac{1}{2}+i\gamma}+\frac{\chi(C)e^{-2\pi i\theta}}{\frac{1}{2}-i\gamma}=\frac{\overline{\chi}(C)e^{2\pi i\theta}}{|\frac{1}{2}+i\gamma|e^{i\beta}}+\frac{\chi(C)e^{-2\pi i\theta}}{|\frac{1}{2}-i\gamma|e^{-i\beta}}=\frac{2}{\sqrt{\frac{1}{4}+\gamma^2}}\mathrm{Re}\left(\overline{\chi}(C)e^{2\pi i\theta-i\beta}\right) .$$

Setting $\overline{\chi}(C_j)=u_j+iv_j$ and $R_{\gamma_j}=\frac{2}{\sqrt{\frac{1}{4}+\gamma_j^2}}$ shows that

$$\begin{aligned}\Sigma_j(\theta)=-R_{\gamma_j}(&u_1\sin(2\pi\theta-\beta)+v_1\cos(2\pi\theta-\beta),\dots\\&,u_r\sin(2\pi\theta-\beta)+v_r\cos(2\pi\theta-\beta)) .\end{aligned} \tag{5.21}$$

Consider the integral

$$\begin{aligned}\int_{\mathbb{T}}e^{-i<\xi,\Sigma_j>}\,d\theta&=\int_{\mathbb{T}}\exp(iR_{\gamma_j}\sum_{m=1}^{r}\xi_m(u_m\sin(2\pi\theta-\beta)+v_m\cos(2\pi\theta-\beta))\,d\theta\\&=\int_{I_1}\exp(iR_{\gamma_j}\sum_{m=1}^{r}\xi_m(u_m\sin(2\pi\theta-\beta)+v_m\cos(2\pi\theta-\beta))\,d\theta\\&+\int_{I_2}\exp(iR_{\gamma_j}\sum_{m=1}^{r}\xi_m(u_m\sin(2\pi\theta-\beta)-v_m\cos(2\pi\theta-\beta))\,d\theta\end{aligned} \tag{5.22}$$

where we have written $\mathbb{T}=I_1\bigcup I_2$ and $I_1=\{\theta\in\mathbb{T}\mid\cos(2\pi\theta-\beta)\geq 0\}$ and $I_2=\{\theta\in\mathbb{T}\mid\cos(2\pi\theta-\beta)<0\}$. In each integral, make the variable change $t=\sin(2\pi\theta-\beta)$. Notice that $d\theta=\pm\frac{dt}{2\pi\sqrt{1-t^2}}$ depending on whether $\theta\in I_1$ or I_2. Setting $U=\sum_{m=1}^{r}\xi_m u_m$ and $V=\sum_{m=1}^{r}\xi_m v_m$ we arrive at

$$\begin{aligned}&\int_{\mathbb{T}}e^{-i<\xi,\Sigma_j>}\,d\theta\\&=\frac{1}{2}\int_{-1}^{1}\exp(iR_{\gamma_j}\sum_{m=1}^{r}\xi_m(u_mt+v_m\sqrt{1-t^2}))\,\frac{dt}{\pi\sqrt{1-t^2}}\\&\ +\frac{1}{2}\int_{-1}^{1}\exp(iR_{\gamma_j}\sum_{m=1}^{r}\xi_m(u_mt-v_m\sqrt{1-t^2}))\,\frac{dt}{\pi\sqrt{1-t^2}}\\&=\frac{1}{\pi}\int_{-1}^{1}\frac{1}{2}\Big(\exp(iR_{\gamma_j}(Ut+V\sqrt{1-t^2}))+\exp(iR_{\gamma_j}(Ut-V\sqrt{1-t^2}))\Big)\,\frac{dt}{\sqrt{1-t^2}}\\&=\frac{1}{\pi}\int_{-1}^{1}\exp(iR_{\gamma_j}Ut)\cos(R_{\gamma_j}V\sqrt{1-t^2})\,\frac{dt}{\sqrt{1-t^2}}\\&=J_0(R_{\gamma_j}\sqrt{U^2+V^2}) .\end{aligned} \tag{5.23}$$

This completes the theorem as $\sqrt{U^2+V^2} = |\sum_{j=1}^{r}\chi(C_j)\xi_j|$. We have now shown that

$$\hat{\mu}(\xi) = \exp\left(-i\sum_{j=1}^{r}\alpha(G,C_j)\xi_j\right)\prod_{\chi\neq 1}\prod_{\gamma_\chi>0} J_0\left(\frac{2|\sum_{j=1}^{r}\chi(C_j)\xi_j|}{\sqrt{\frac{1}{4}+\gamma_\chi^2}}\right).$$

Comment The above formula contains the formula from the Rubinstein-Sarnak article (see [62] p. 184). We can see this by taking the field extension $L/\mathbb{Q}$ where $L=\mathbb{Q}(\zeta_q)$.

5.3 Examples

5.3.1 S_3

There are three conjugacy classes in this group. Namely,

$$C_1 = \{1\},\ C_2 = \{(12),(13),(23)\},\ \text{and } C_3 = \{(123),(132)\}.$$

Note that,

$$C_1^2 = C_1,\ C_2^2 = C_1,\ \text{and } C_3^2 = C_3.$$

This leads to the bias terms,

$$c_{\mathrm{sq}}(S_3,C_1) = 4,\ c_{\mathrm{sq}}(S_3,C_2) = 0,\ \text{and } c_{\mathrm{sq}}(S_3,C_3) = 1.$$

The S_3 examples considered correspond to the Hilbert class fields of $\mathbb{Q}(\sqrt{-l})$ for $l = 23, 31, 59, 83$. (This list may be extended). In Chapter 4, the polynomials whose roots generate these Hilbert class fields are listed. The three irreducible characters of this group will be denoted χ_1, χ_2, and χ_3. S_3 has the following character table:

	C_1	C_2	C_3
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

Set $\pi_1(x) = 6\pi_{C_1}(x)$, $\pi_2(x) = 2\pi_{C_2}(x)$, and $\pi_3(x) = 3\pi_{C_3}(x)$. Because of our normalization, $\pi_i(x) \sim \mathrm{Li}(x)$ in each case. We are interested in the sets of real numbers described by

$$P_{i,j} = \{x \geq 2 \mid \pi_i(x) > \pi_j(x)\ \}\ .$$

We will first consider the two-way races. The following identities for the non-trivial Artin L-functions can be proven.

$$L(s,\chi_2) = L(s,(\frac{\cdot}{l}))\ ,\ L(s,\chi_3) = L(s,f_l)$$

where the f_l are the modular forms mentioned in the last chapter. In all cases, it was checked numerically that these Artin L-functions are non-zero at $s = 1/2$. This implies that $c(S_3, C_i) = c_{\mathrm{sq}}(S_3, C_i)$ for $i = 1, 2, 3$. We will use the convention that non-trivial zeros of $L(s, \chi_2)$ are of the form $\frac{1}{2} + i\gamma_2$ and non-trivial zeros of $L(s, \chi_3)$ are of the form $\frac{1}{2} + i\gamma_3$. From the last section, we obtain the following formulae.

$$E_{1,2}(x) = \frac{\log x}{x^{\frac{1}{2}}}(\pi_1(x) - \pi_2(x)) = (0-4) - 2\sum_{|\gamma_2| \le X} \frac{x^{i\gamma_2}}{\frac{1}{2} + i\gamma_2} - 2\sum_{|\gamma_3| \le X} \frac{x^{i\gamma_3}}{\frac{1}{2} + i\gamma_3} + \text{small} ,$$

$$E_{1,3}(x) = \frac{\log x}{x^{\frac{1}{2}}}(\pi_1(x) - \pi_3(x)) = (1-4) - 3\sum_{|\gamma_3| \le X} \frac{x^{i\gamma_3}}{\frac{1}{2} + i\gamma_3} + \text{small} ,$$

$$E_{3,2}(x) = \frac{\log x}{x^{\frac{1}{2}}}(\pi_3(x) - \pi_2(x)) = (0-1) - 2\sum_{|\gamma_2| \le X} \frac{x^{i\gamma_2}}{\frac{1}{2} + i\gamma_2} + \sum_{|\gamma_3| \le X} \frac{x^{i\gamma_3}}{\frac{1}{2} + i\gamma_3} + \text{small} ,$$

where small $\ll \frac{\sqrt{x}\log^2 X}{X} + \frac{1}{\log x}$ for $2 \le x \le X$. This notation will be used throughout the following sections. As shown in the previous section, there exist limiting distributions of the functions $E_{i,j}(e^y)$ denoted as $\mu_{i,j}$. As in the Rubinstein-Sarnak article, their Fourier transforms may be computed. Doing a similar calculation we get

$$\hat{\mu}_{1,2}(\xi) = e^{4i\xi} \prod_{\gamma_2 > 0} J_0\left(\frac{4\xi}{\sqrt{\frac{1}{4} + \gamma_2^2}}\right) \prod_{\gamma_3 > 0} J_0\left(\frac{4\xi}{\sqrt{\frac{1}{4} + \gamma_3^2}}\right) ,$$

$$\hat{\mu}_{1,3}(\xi) = e^{3i\xi} \prod_{\gamma_3 > 0} J_0\left(\frac{6\xi}{\sqrt{\frac{1}{4} + \gamma_3^2}}\right) ,$$

$$\hat{\mu}_{3,2}(\xi) = e^{i\xi} \prod_{\gamma_2 > 0} J_0\left(\frac{4\xi}{\sqrt{\frac{1}{4} + \gamma_2^2}}\right) \prod_{\gamma_3 > 0} J_0\left(\frac{2\xi}{\sqrt{\frac{1}{4} + \gamma_3^2}}\right) .$$

Assuming RH and LI we have

$$\delta(P_{1,2}) = \int_0^\infty d\mu_{1,2}(t) \quad \text{and} \quad \delta(P_{2,1}) = \delta(P_{1,2}^c) = \int_{-\infty}^0 d\mu_{1,2}(t) .$$

If we let $f_{1,2}(t)$ be the density function of $\mu_{1,2}$ then $g(t) = f_{1,2}(t-4)$ is the density function of the distribution $w_{1,2}$. This has an even Fourier transform

$$\widehat{w}_{1,2}(\xi) = \prod_{\gamma_2 > 0} J_0\left(\frac{4\xi}{\sqrt{\frac{1}{4} + \gamma_2^2}}\right) \prod_{\gamma_3 > 0} J_0\left(\frac{4\xi}{\sqrt{\frac{1}{4} + \gamma_3^2}}\right) .$$

Therefore,

$$\delta(P_{2,1}) = \int_{-\infty}^{4} dw_{1,2}(t) = \frac{1}{2}\left(\int_{-\infty}^{4} + \int_{-4}^{\infty}\right) dw_{1,2}(t) = \frac{1}{2}\left(\int_{-\infty}^{\infty} + \int_{-4}^{4}\right) dw_{1,2}(t)$$
$$= \frac{1}{2} + \frac{1}{2}\int_{-4}^{4} dw_{1,2}(t) \ . \tag{5.24}$$

Applying the inversion formula for distribution functions, we obtain

$$\delta(P_{2,1}) = \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(4t)}{t}\hat{w}_{1,2}(t)\ dt \ .$$

Using the same techniques as in the Rubinstein-Sarnak article, we can compute the above integral very precisely thus obtaining values for the logarithmic densities. Details for all density calculations can be found in section 5.4. We computed the following densities for $l = 23, 31, 59, 83$:

(i,j)	$\delta_{23;i,j}$	(i,j)	$\delta_{23;i,j}$	(i,j)	$\delta_{23;i,j}$
21	0.9830	23	0.7246	31	0.9570
12	0.0170	32	0.2754	13	0.0430

(i,j)	$\delta_{31;i,j}$	(i,j)	$\delta_{31;i,j}$	(i,j)	$\delta_{31;i,j}$
21	0.9664	23	0.6863	31	0.9414
12	0.0336	32	0.3137	13	0.0586

(i,j)	$\delta_{59;i,j}$	(i,j)	$\delta_{59;i,j}$	(i,j)	$\delta_{59;i,j}$
21	0.9287	23	0.6511	31	0.9108
12	0.0713	32	0.3489	13	0.0892

(i,j)	$\delta_{83;i,j}$	(i,j)	$\delta_{83;i,j}$	(i,j)	$\delta_{83;i,j}$
21	0.8953	23	0.6238	31	0.8891
12	0.1047	32	0.3762	13	0.1109

Finally, we could also race conjugacy sets. Note that

$$\sigma_p \in C_1 \cup C_3 \iff (\frac{p}{l}) = 1 \text{ and } \sigma_p \in C_2 \iff (\frac{p}{l}) = -1$$

for p an unramified prime. This shows that if we define conjugacy sets $D_1 = C_1 \cup C_3$ and $D_2 = C_2$, then the race between primes that satisfy $\sigma_p \in D_1$ versus those that satisfy $\sigma_p \in D_2$ is just the race between primes that are quadratic residues mod l versus quadratic non-residues mod l. This is the case studied in the Rubinstein-Sarnak [62] paper. Listed below are the densities of the sets

$$P_{l;N,R} = \{x \geq 2 \mid \pi_N(x;l) > \pi_R(x;l)\ \}$$

where

$$\pi_N(x;l) = \sum_{p \leq x,\, (\frac{p}{l})=-1} 1 = \pi_{D_1}(x) \text{ and } \pi_R(x;l) = \sum_{p \leq x,\, (\frac{p}{l})=1} 1 = \pi_{D_2}(x).$$

l	$\delta_{l;N,R}$
23	0.90318
31	0.85507
59	0.79420
83	0.74696

5.3.2 D_4

Let $q(x) = x^4 - 12$. The splitting field of this polynomial is easily seen to be $L = \mathbb{Q}(i, \sqrt[4]{12})$ and it is possible to show that $\mathrm{Gal}(L/\mathbb{Q}) \cong D_4$. D_4 is the dihedral group of order eight and has the group presentation

$$D_4 = \{r, s \mid r^4 = s^2, rs = sr^{-1}\ \}\ .$$

The isomorphism can be specified by defining elements of the Galois group as

$$r : \begin{cases} \sqrt[4]{12} \to i\sqrt[4]{12} \\ i \to i \end{cases} \text{ and } s : \begin{cases} \sqrt[4]{12} \to \sqrt[4]{12} \\ i \to -i \end{cases} .$$

It has the five conjugacy classes,

$$C_1 = \{1\}, C_2 = \{r^2\}, C_3 = \{s, sr^2\}, C_4 = \{r, r^3\}, \text{and } C_5 = \{sr, sr^3\}$$

which satisfy

$$C_1^2 = 1,\ C_2^2 = 1,\ C_3^2 = 1,\ C_4^2 = C_2,\ C_5^2 = 1.$$

Therefore, we have the Chebyshev bias terms

$$c_{\mathrm{sq}}(D_4, C_1) = \frac{6}{1} = 6,\ c_{\mathrm{sq}}(D_4, C_2) = \frac{2}{1} = 2,\ c_{\mathrm{sq}}(D_4, C_i) = 0 \text{ for } i = 3, 4, 5\ .$$

In the example cited, none of the Artin L-functions vanish at $s = \frac{1}{2}$ and so we obtain $c(D_4, C_i) = c_{\mathrm{sq}}(D_4, C_i)$ for $1 \leq i \leq 5$. It is clear that C_1 loses to all conjugacy clases. C_2 beats C_1 but loses to the other three conjugacy classes. Lastly, there are no biases between C_3, C_4, and C_5. In fact,

$$\delta(P_{D_4;i,j}) = \frac{1}{2} \text{ for } i \neq j \in \{3, 4, 5\}\ .$$

Listed below is the character table for D_4.

	C_1	C_2	C_3	C_4	C_5
χ_1	1	1	1	1	1
χ_2	1	1	-1	1	-1
χ_3	1	1	1	-1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Serre notes in [69] pp. 242-243 that the Artin L-function $L(s,\chi_5)$ equals the modular form L-function $L(s,f)$ where $f(z)=\eta^2(12z)$. We can also show that

$$L(s,\chi_2)=L(s,(\frac{-4}{n})), L(s,\chi_3)=L(s,(\frac{12}{n})), \text{ and } L(s,\chi_4)=L(s,(\frac{-3}{n})),$$

As an example the first identity is proven. Consider $\chi_2 : \mathrm{Gal}(L/\mathbb{Q}) \to \mathbb{C}^*$. Let $H=\ker\chi_2 = C_1\bigcup C_2\bigcup C_4 = \{1,r,r^2,r^3\}$. Note that H is a normal subgroup and D_4/H is a group of order two. The fixed field of H is $\mathbb{Q}(i)$, thus

$$D_4/H \cong \mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \ .$$

We have the projection map

$$p : D_4 \to D_4/H \ .$$

Let λ be the non-trivial character of D_4/H. It's easy to see that $\chi_2 = p\circ\lambda$. Consequently, by properties of Artin L-functions,

$$L(s,\chi_2,L_q/\mathbb{Q}) = L(s,\lambda,\mathbb{Q}(i)/\mathbb{Q}) \ .$$

However, the Artin L-function on the right is precisely the Dirichlet L-function $L(s,(\frac{-4}{n}))$. The other two identities are similar.

In this example, there is a natural race between C_1 and C_2 since they have the same number of elements. Note that $\chi(C_1)-\chi(C_2)=0$ for $\chi\neq\chi_5$ and $\chi_5(C_1)-\chi_5(C_2)=2-(-2)=4$. From the formulas in the previous section,

$$8\frac{\log x}{\sqrt{x}}(\pi_{C_1}(x)-\pi_{C_2}(x)) = (2-6) - 4\sum_{0<|\gamma_5|<X}\frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + \text{small} \ .$$

Thus, we have

$$\frac{\log x}{\sqrt{x}}(\pi_{C_1}(x)-\pi_{C_2}(x)) = -\frac{1}{2} - \frac{1}{2}\sum_{0<|\gamma_5|<X}\frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + \text{small} \ .$$

This shows a bias towards C_2.

(ii) In a similar fashion one derives for $j = 3, 4, 5$

$$E_{j,2}(x) = \frac{\log x}{\sqrt{x}}(\pi_{C_j}(x) - 8\pi_{C_2}(x)) = 2 - \sum_{\chi \neq 1} (\chi(C_j) - \chi(C_2)) \sum_{0<|\gamma_\chi|<X} \frac{x^{i\gamma}}{\frac{1}{2} + i\gamma} + \text{small}$$

and

$$E_{j,1}(x) = \frac{\log x}{\sqrt{x}}(\pi_{C_j}(x) - 8\pi_{C_1}(x)) = 6 - \sum_{\chi \neq 1} (\chi(C_j) - \chi(C_1)) \sum_{0<|\gamma_\chi|<X} \frac{x^{i\gamma}}{\frac{1}{2} + i\gamma} + \text{small}\ .$$

Both formulas indicate a bias towards C_j. Each of the two way logarithmic densities were calculated using the zeros of the Artin L-functions. We obtained

(i,j)	$\delta_{144;i,j}$	(i,j)	$\delta_{144;i,j}$	(i,j)	$\delta_{144;i,j}$	(i,j)	$\delta_{144;i,j}$
21	0.8597	42	0.8098	31	0.9985	51	0.9963
12	0.1403	24	0.1902	13	0.0015	15	0.0037
32	0.8272	52	0.8060	41	0.9968	34,35,45	1/2
23	0.1728	25	0.1940	14	0.0031	43,53,54	1/2

5.3.3 H_8

We will now consider two examples of field extensions of $\mathbb{Q}$ having Galois group H_8. The first example we will consider is Serre's example of a Galois extension whose degree two Artin L-function has a zero at the centre of the crtical line. Set $L_1 = \mathbb{Q}(\theta_1)$ where $\theta_1 = \sqrt{\frac{5+\sqrt{5}}{2}\frac{41+\sqrt{5\cdot 41}}{2}}$. The minimal polynomial of θ is

$$f(x) = x^8 - 205x^6 + 13940x^4 - 378225x^2 + 3404025.$$

The second example is given by $L_2 = \mathbb{Q}(\theta_2)$ where $\theta_2 = \sqrt{(2+\sqrt{2})(3+\sqrt{3})}$. The minimal polynomial of θ_2 is

$$f(x) = x^8 - 24x^6 + 144x^4 - 288x^2 + 144.$$

It is not too difficult to check that these extensions are normal and have Galois group H_8 (see p. 498 of Dummit and Foote [19] for a sketch of a proof). Furthermore, H_8 has a group presentation

$$H_8 = \{i, j \mid i^4 = 1, j^2 = i^2, jij^{-1} = i^{-1}\}$$

and has the five conjugacy classes,

$$C_1 = \{1\}\ ,\ C_2 = \{-1\}\ ,\ C_3 = \{\pm i\}\ ,\ C_4 = \{\pm j\}\ ,\ C_5 = \{\pm k\}.$$

which satisfy

$$C_1^2 = C_1,\ C_2^2 = C_1, C_3^2 = C_2,\ C_4^2 = C_2,\ C_5^2 = C_2.$$

Hence, we have the constants

$$c_{\mathrm{sq}}(H_8, C_1) = \frac{2}{1} = 2,\ c_{\mathrm{sq}}(H_8, C_2) = \frac{6}{1} = 6,\ \text{and } c_{\mathrm{sq}}(H_8, C_i) = 0 \text{ for } i = 3, 4, 5.$$

The above computations show that if there are no contributions from central zeros of an Artin L-function then C_i beats C_1 and C_2 for $i = 3, 4, 5$ and C_1 beats C_2. Also, there are no biases between C_3, C_4, and C_5. It is interesting to note that in the D_4 case C_1 loses to C_2, whereas in this example C_1 beats C_2. The race between C_1 and C_2 is natural as they have the same number of elements. Given a prime p we have

$$\sigma_p = C_1 \Leftrightarrow\ f(x) \text{ splits completely mod } p.$$

and

$$\sigma_p = C_2 \Leftrightarrow\ f \equiv f_1 f_2 f_3 f_4 \bmod p.$$

where the $f_i(x)$ are quadratic irreducible polynomials mod p. Listed below is the character table for H_8

	C_1	C_2	C_3	C_4	C_5
χ_1	1	1	1	1	1
χ_2	1	1	-1	1	-1
χ_3	1	1	1	-1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

It is interesting to note that this character table is identical to the character table for D_4, yet these two groups are non-isomorphic.

(i) We will now analyze these examples using the explicit formulae. It should be observed that L_1 is a totally real field and contains the biquadratic field $\mathbb{Q}(\sqrt{5}, \sqrt{41})$. Consequently, it contains the three real quadratic fields $\mathbb{Q}(\sqrt{5}), \mathbb{Q}(\sqrt{41})$, and $\mathbb{Q}(\sqrt{205})$. As in the D_4 example, we have

$$L(s, \chi_2) = L(s, (\frac{5}{\cdot})),\ L(s, \chi_3) = L(s, (\frac{41}{\cdot})),\ \text{and } L(s, \chi_4) = L(s, (\frac{205}{\cdot}))\ ,$$

up to some reordering of the characters. Unfortunately, the 2-dimensional character cannot be matched up with a weight one modular form as in the D_4 example. This is because L_1 is totally real and this forces ρ_5 to be an even representation. We now consider the contribution to the bias terms from the zero at $s = \frac{1}{2}$ of $L(s, \chi_5)$. We should notice that $L(\frac{1}{2},\chi_i) \neq 0$ for $i = 2,3,4$. This can just be checked numerically. Set $\eta_{\chi_5} = \mathrm{ord}_{\mathrm{s}=\frac{1}{2}}\ L(\mathrm{s},\chi_5)$. This will be determined numerically. Therefore, we have

$$c_{\frac{1}{2}}(H_8, C_3) = 0,\ c_{\frac{1}{2}}(H_8, C_4) = 0,\ \text{and } c_{\frac{1}{2}}(H_8, C_5) = 0$$

since $\chi_5(C_i) = 0$ in each of these cases. Looking at the character table we also obtain

$$c_{\frac{1}{2}}(H_8, C_1) = 4\eta_{\chi_5} \text{ and } c_{\frac{1}{2}}(H_8, C_2) = -4\eta_{\chi_5}.$$

We conclude that $c(H_8, C_1) = 2 + 4\eta_{\chi_5}$, $c(H_8, C_2) = 6 - 4\eta_{\chi_5}$ and $c(H_8, C_i) = 0$ for $i = 3,4,5$. We should first observe that there are no biases between any two of the conjugacy classes of order two. A race between any two of these would have logarithmic density equal to $\frac{1}{2}$. Let's consider the more interesting case of racing C_1 and C_2. From the explicit formulas, we have

$$8\frac{\log x}{\sqrt{x}}(\pi_{C_1}(x) - \pi_{C_2}(x)) = (6 - 4\eta_{\chi_5}) - (2 + 4\eta_{\chi_5}) - 4\sum_{0<|\gamma_5|<X} \frac{x^{i\gamma}}{\frac{1}{2} + i\gamma} + \text{small}.$$

Consequently, we obtain

$$\frac{\log x}{\sqrt{x}}(\pi_{C_1}(x) - \pi_{C_2}(x)) = \frac{1}{2} - \eta_{\chi_5} - \frac{1}{2}\sum_{0<|\gamma_5|<X} \frac{x^{i\gamma}}{\frac{1}{2} + i\gamma} + \text{small}.$$

Since $\eta_{\chi_5} \geq 1$ we see that there is a bias towards C_2. What is interesting about this formula is that without knowledge of the central zero of $L(s,\chi_5)$, the classical formulas suggest a bias towards C_1. In fact, the next example will show that there are quaternion groups in which the bias is towards C_1. Originally, we had planned to compute the logarithmic density of the set

$$P_{205^2;2,1} = \{x \geq 2 \mid \pi_{C_2}(x) > \pi_{C_1}(x)\ \}.$$

Unfortunately, this project could not be completed in time. A future project is to compute $\delta(P_{205^2;2,1})$. One of the difficulties in computing this density is that the Artin L-function $L(s,\chi_5)$ has Artin conductor $\mathfrak{f}(\chi_5) = 205^2 = 42025$. This is rather large and may limit the number of zeros that can be computed. Following Rubinstein's thesis, to compute $L(s,\chi_5)$ at a point $\frac{1}{2} + it$ would require evaluating a sum of special functions of length $O(\frac{205}{\pi} \cdot t)$. If the constant in the O term is 10 then computing the L-function at height $t = 100$ would require computing $10 \cdot \frac{205}{\pi} \cdot 100 \approx 65253$ special functions. (This is not quite correct as we know that non-zero coefficients of Artin L-functions have zero density). However, this may still be too many terms to reasonably compute zeros at a larger height.

Instead, we wrote a short progam using PARI to detect sign changes of the function $\pi_{C_1}(x) - \pi_{C_2}(x)$. Running this program for a few days, we found all sign changes less than 100 million. We found 1375 sign changes. We then used the sign changes to estimate the density by the integral on the right. That is,

$$\delta(P_{205^2;2,1}) \approx \frac{1}{\log X}\int_2^X 1_{\{x\geq 2:\pi_{C_2}(x)>\pi_{C_1}(x)\}}\,\frac{dt}{t}.$$

where $X = 10^8$. With this choice we found the right hand side to be 0.8454.... This seems to indicate a bias towards C_2 as expected. In addition, it may be noticed by looking at the list of sign changes that were many more long intervals where $\pi_{C_2}(x)$ beats $\pi_{C_1}(x)$.

(ii) As in the previous example, this field is normal over $\mathbb{Q}$ with Galois group H_8. Clearly it contains the biquadratic field $\mathbb{Q}(\sqrt{2},\sqrt{3})$ and the quadratic fields $\mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3})$, and $\mathbb{Q}(\sqrt{6})$. As before, the one-dimensional Artin L-functions are Dirichlet L-functions. Namely,

$$L(s,\chi_2) = L(s,(\frac{8}{\cdot})),\ L(s,\chi_3) = L(s,(\frac{12}{\cdot})),\ \text{and}\ L(s,\chi_4) = L(s,(\frac{24}{\cdot})).$$

Once again, the explicit formula is

$$\frac{\log x}{\sqrt{x}}(\pi_{C_1}(x) - \pi_{C_2}(x)) = \frac{1}{2} - \eta_{\chi_5} - \frac{1}{2}\sum_{0<|\gamma_5|<X}\frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + \text{small}$$

where $\eta_{\chi_5} = \mathrm{ord}_{s=\frac{1}{2}} L(s,\chi_5)$. As in the previous example, we counted sign changes of $\pi_{C_1}(x) - \pi_{C_2}(x)$ up to $X = 10^8$. We found 1478 signs changes less than 10^8. In addition we found that

$$\delta(P_{48^2;1,2}) \approx 0.7391...$$

(48^2 is in fact, the Artin conductor of $L(s,\chi_5)$.) This seems to indicate a bias towards C_1 contrary to the previous example. In addition, this suggests $L(\frac{1}{2},\chi_5) \neq 0$ in this case. It would be interesting to determine how close this approximation is to the actual size of the logarithmic density of $P_{48^2;1,2}$.

5.3.4 S_4

No numerical examples were computed in this case. The symmetric group on four numbers, S_4, can be generated by the elements $a = (12)$ and $b = (1234)$. It has the presentation

$$S_4 = \{a, b \mid a^2 = b^4 = 1\ ,\ (ab)^3 = 1\ \}.$$

Certainly, there exist infinitely many number fields $L/\mathbb{Q}$ with Galois group S_4. For example, take

$$f_p(x) = x^4 + px + p$$

for $p \geq 7$ and prime. Let L_p be the splitting field of p. The resolvent cubic is $h_p(x) = x^3 - 4px + p^2$. This is irreducible for p in the stated range, since the possible rational roots are $\pm 1, \pm p$, and $\pm p^2$. It can be checked that $h_p(x) \neq 0$ for any of these numbers. In addition, the polynomial discriminant of f_p is

$$D_p = -p^3(27p - 256)$$

which is not a square. Therefore, $\mathrm{Gal}(L_p/\mathbb{Q}) \cong S_4$ by the condition stated on p. 529 of [19].

S_4 has five conjugacy classes represented by the following elements

$$C_1 = \{1\}\ ,\ C_2 = \{(12)\}\ ,\ C_3 = \{(123)\}\ ,\ C_4 = \{(1234)\}\ ,\ C_5 = \{(12)(34)\}.$$

Note that these conjugacy classes have sizes 1, 6, 8, 6, and 3. In addition,

$$C_1^2 = C_1,\ C_2^2 = C_1,\ C_3^2 = C_3,\ C_4^2 = C_5,\ \text{and } C_5^2 = C_1.$$

thus

$$\begin{aligned} &c_{\mathrm{sq}}(S_4, C_1) = \frac{10}{1},\ c_{\mathrm{sq}}(S_4, C_2) = 0,\ c_{\mathrm{sq}}(S_4, C_3) = \frac{3}{3} = 1, \\ &c_{\mathrm{sq}}(S_4, C_4) = 0,\ \text{and } c_{\mathrm{sq}}(S_4, C_5) = \frac{6}{3} = 2\ . \end{aligned} \tag{5.25}$$

This shows that C_2 and C_4 each beat C_1, C_3 and C_5. However, there is no bias between C_2 and C_2 as $\delta(P_{S_4;2,4}) = \frac{1}{2}$. C_3 beats C_1 and C_5. Lastly, C_5 beats C_1. C_1 loses to all conjugacy classes.

5.3.5 S_n

The symmetric group on n letters is generated by (12) and $(12 \cdots n)$. If n is prime it is generated by any transposition and any n-cycle. The conjugacy classes of S_n can be characterized by the partitions of n. For example, if we have the partition

$$n = n_1 + n_2 + n_3 + \cdots + n_k$$

then the corresponding conjugacy class of S_n would consist of elements $\sigma \in S_n$ of the form

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_k$$

where σ_i is an n_i-cycle and each of the cycles contains distinct elements from one another. This is explained in the graduate algebra textbook [19]. For each positive integer n, there exist infinitely many irreducible $f(x) \in \mathbb{Z}[x]$ with S_n as the Galois group of the corresponding splitting field (see [19] p. 555). For $n = l$ a prime, we can even write down a specific family of polynomials for which the Galois group is the whole symmetric group S_l. It is shown on pp. 150-151 of [23] that if $p_l(x) = x^l - ax + b$ where $((l-1)a, lb) = 1$ then the splitting field of the Galois group is S_l.

We compute some bias terms for conjugacy classes in the symmetric group S_n. Consider the conjugacy classes $C_1 = 1$ and

$$C_q = \langle (12\cdots q) \rangle$$

the conjugacy class of q-cycles where $q \leq n$. Let $n = 2q + r$. We claim that

$$\begin{aligned}
c_{\mathrm{sq}}(S_n, C_1) &= 1 + \sum_{1 \leq k \leq \frac{n}{2}} \binom{n}{2k} \frac{(2k)!}{2^k k!} \\
c_{\mathrm{sq}}(S_n, C_q) &= 1 + \sum_{1 \leq k \leq \frac{n-q-r}{2}} \binom{n-q}{2k} \frac{(2k)!}{2^k k!} \text{ if } q \text{ is odd} \\
c_{\mathrm{sq}}(S_n, C_q) &= 0 \text{ if } q \text{ is even.}
\end{aligned} \tag{5.26}$$

We will compute the first identity. Note that $\mathrm{sq}^{-1}(S_n, C_1)$ consists of the identity element and products of transpositions. Suppose we are considering the conjugacy class of elements of the form

$$(a_1 a_2)(a_3 a_4) \cdots (a_{2k-1} a_{2k})$$

where $2k \leq n$. Note that there are

$$n \cdot (n-1) \cdot (n-2) \cdots (n-2k-2) \cdot (n-2k-1)$$

choices for the elements $a_1, \ldots, a_{2k}$. However, each transposition can be written in 2 ways. In addition, an element of the above form is independent of the ordering of the 2-cycles. That is, there are $k!$ ways to permute the transpositions. Hence, the number of elements in this conjugacy class is

$$\frac{n \cdot (n-1) \cdot (n-2) \cdots (n-2k-2) \cdot (n-2k-1)}{2^k k!} = \binom{n}{2k} \frac{(2k)!}{2^k k!}.$$

Summing over $k \leq n/2$ and adding in the contribution from the identity element shows that

$$c_{\mathrm{sq}}(S_n, C_1) = 1 + \sum_{1 \leq k \leq \frac{n}{2}} \binom{n}{2k} \frac{(2k)!}{2^k k!}.$$

A similar counting argument gives the other formula. More generally, we could compute bias factors for an arbitrary conjugacy class in the symmetric group. However, some of the formulas can become quite complicated. Note that in the second example, we only considered the conjugacy classes C_q for q odd. In that case, the permutations that square to a q-cycle must be a q-cycle multiplied by transpositions. However, if $q = 2t$ is even, then $c(S_n, C_{2t}) = 0$. This is because

$$\mathrm{sq}((a_1 a_2 \dots a_{2t-1} a_{2t})) = (a_1 a_3 \dots a_{2t-1})(a_2 a_4 \dots a_{2t}).$$

Clearly, the square of an odd q-cycle is a q-cycle. On the other hand, the square of an even q-cycle is the product of two $q/2$ cycles. Thus, no permutation can square to an even q cycle. In general, this shows that if we write out a partition of a number n and we group the partition numbers into odds and evens

$$n = n_{odd} + n_{even} = (u_1 + u_2 + \cdots + u_a) + (v_1 + v_2 + \cdots + v_b)$$

where the u_i's are odd and the v_j's are even then if b is odd the Chebyshev bias term corresponding to the conjugacy class of n's cycle type is zero. It seems difficult to write a nice precise formula for an arbitrary conjugacy class.

Note that in the above identities, we have shown that

$$c_{\mathrm{sq}}(S_n, C_1) > c_{\mathrm{sq}}(S_n, C_q)$$

for q odd as $\binom{n}{2k} > \binom{n-q}{2k}$. Hence, C_1 loses to C_q for q odd. Likewise, we trivially have $c_{\mathrm{sq}}(S_n, C_1) > c_{\mathrm{sq}}(S_n, C_q) = 0$ and C_1 loses to C_q for q even. It would be interesting to know if C_1 loses to every conjugacy class. However, this cannot be true for every group as we know that C_1 can beat C_2 in H_8 if the degree two Artin L-function does not vanish at $1/2$.

Lastly, it is of interest to note that Chowla et al. [7] studied the asymptotic behaviour of the Chebyshev bias term $c_{\mathrm{sq}}(S_n, C_1)$. They denoted this as T_n. Originally, Chowla, Herstein, and Moore [7] vol.2 pp. 772-778 showed that

$$T_n \sim \frac{(n/e)^{\frac{n}{2}} e^{n^{\frac{1}{2}}}}{2^{\frac{1}{2}} e^{\frac{1}{4}}}.$$

In a later article, Chowla, Herstein, and Scott [7] vol.2 pp. 826-828 studied the numbers $A_{n,d}$ where

$$A_{n,d} = \#\{\sigma \in S_n \mid \sigma^d = 1\ \}.$$

They found a generating function for the numbers $A_{n,d}$. Precisely,

$$\sum_{n=0}^{\infty} \left(\frac{A_{n,d}}{n!} \right) x^n = \exp\left(\sum_{k|d} \frac{x^k}{k} \right).$$

The numbers $A_{n,d}$ are very similar to the Chebyshev bias terms. However, we are interested in counting the number of elements in S_n which square into a fixed conjugacy class. It would be interesting to investigate whether the techniques of Chowla et al. give any information on the Chebyshev bias constants for S_n. For example, can we find a generating function or asymptotic formula for some of these constants.

5.4 Computing the bias

This section will outline how to compute the logarithmic densities of the sets we are interested in. Suppose P is one of these sets. We are able to write

$$\delta(P) = \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(au)}{u}\hat{w}(u)\ du$$

where a is an integer and $\hat{w}$ is an infinite product of Bessel functions of the form

$$\hat{w}(u) = \prod_{\gamma>0} J_0\left(\frac{2\alpha_\gamma u}{\sqrt{\frac{1}{4}+\gamma^2}}\right)$$

where γ ranges over a countable set and $\alpha_\gamma \in \mathbb{C}$. The technique in evaluating these integrals involves applying Poisson summation to replace the the improper integral by an infinite sum, then replacing the infinite sum by a finite sum, and finally replacing the infinite product for $\hat{w}(u)$ by a finite product. Here are the details:

5.4.1 Step 1: Poisson summation

The Poisson summation formula states that if ϕ is a sufficiently nice function (C^∞) and $\epsilon > 0$ is some parameter then

$$\epsilon\sum_{n\in\mathbb{Z}} \phi(\epsilon n) = \sum_{n\in\mathbb{Z}} \hat{\phi}(\frac{n}{\epsilon}) = \hat{\phi}(0) + \sum_{n\in\mathbb{Z},\, n\neq 0} \hat{\phi}(\frac{n}{\epsilon}).$$

Typically, we will choose ϵ to be a small number like $\epsilon = \frac{1}{20}$. We will apply this formula to

$$\begin{aligned}\phi(u) &= \frac{1}{2\pi}\frac{\sin(au)}{u}\hat{w}(u)\\ \hat{\phi}(x) &= \frac{1}{2}(\chi_{[-a,a]} * g)(x) = \frac{1}{2}\int_{x-1}^{x+1} g(u)\ du = \frac{1}{2}\int_{x-1}^{x+1} dw(u).\end{aligned} \tag{5.27}$$

Substituting these into the Poisson summation formula shows

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(au)}{u}\hat{w}(u)\ du = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\hat{w}(\epsilon n) - \sum_{n\in\mathbb{Z},\, n\neq 0} \hat{\phi}(\frac{n}{\epsilon}).$$

The first approximation in computing the integral will be to drop the second sum on the RHS of the equation. We will later show this sum is small. This gives us

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(au)}{u}\hat{w}(u)\ du = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\hat{w}(\epsilon n) + \mathrm{error}_1(\epsilon)$$

where $\mathrm{error}_1(\epsilon) = -\sum_{n\in\mathbb{Z},n\neq 0}\hat{\phi}(\frac{n}{\epsilon})$. Note that

$$\hat{\phi}(\frac{n}{\epsilon}) = \frac{1}{2}\int_{\frac{n}{\epsilon}-1}^{\frac{n}{\epsilon}+1} g(u)\ du \leq \frac{1}{2}w[\frac{n}{\epsilon}-1,\infty) \leq \frac{1}{2}\exp\left(\frac{-3(\frac{n}{\epsilon}-1-2\gamma_1)^2}{16R}\right)$$

where $R = \sum_{\gamma>0}\frac{\alpha_\gamma^2}{\frac{1}{4}+\gamma^2}$. Thus, we see that

$$|\mathrm{error}_1| = 2\sum_{n=1}^{\infty}\hat{\phi}(\frac{n}{\epsilon}) \leq \sum_{n=1}^{\infty}\exp\left(\frac{-3(\frac{n}{\epsilon}-1-2\gamma_1)^2}{16R}\right).$$

5.4.2 Step 2 : Truncating the sum

The next approximation will be to replace the infinite sum by a finite sum. We choose an appropriate C such that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin(au)}{u}\hat{w}(u)\ du = \frac{1}{2\pi}\sum_{-C\leq n\epsilon\leq C}\epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\hat{w}(\epsilon n) + \mathrm{error}_1(\epsilon) + \mathrm{error}_2(\epsilon, C).$$

In the following approximation of error_2 we will need the estimate $|J_0(x)| \leq \min(1, \sqrt{\frac{2}{\pi|x|}})$. Therefore

$$\begin{aligned}\mathrm{error}_2(\epsilon, C) =& \frac{1}{2\pi}\left(\sum_{-\infty<n\epsilon<\infty} - \sum_{-C\leq n\epsilon\leq C}\right)\epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\hat{w}(\epsilon n)\\ =& \frac{1}{\pi}\sum_{n\epsilon>C}\epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\prod_{j=1}^{\infty}J_0\left(\frac{2\alpha_j\epsilon n}{\sqrt{\frac{1}{4}+\gamma_j^2}}\right) \qquad (5.28)\\ \leq& \frac{1}{\pi}\sum_{n\epsilon>C}\epsilon\frac{1}{\epsilon n}\left|\prod_{j=1}^{M}J_0\left(\frac{2\alpha_j\epsilon n}{\sqrt{\frac{1}{4}+\gamma_j^2}}\right)\right|\end{aligned}$$

for any $M \geq 1$. Using the bound for $J_0(x)$, we obtain

$$\begin{aligned}|\mathrm{error}_2| \leq& \frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{\frac{1}{4}}}{\pi^{\frac{M}{2}+1}\left(\prod_{j=1}^{M}|\alpha_j|\right)^{\frac{1}{2}}}\sum_{n\epsilon>C}\frac{\epsilon}{(\epsilon n)^{\frac{M}{2}+1}}\\ <& \frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{\frac{1}{4}}}{\pi^{\frac{M}{2}+1}\left(\prod_{j=1}^{M}|\alpha_j|\right)^{\frac{1}{2}}}\left(\int_C^{\infty}\frac{1}{x^{\frac{M}{2}+1}}\ dx + \frac{\epsilon}{C^{\frac{M}{2}+1}}\right) \qquad (5.29)\\ =& \frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{\frac{1}{4}}}{\pi^{\frac{M}{2}+1}\left(\prod_{j=1}^{M}|\alpha_j|\right)^{\frac{1}{2}}}\left(\frac{2}{MC^{\frac{M}{2}}} + \frac{\epsilon}{C^{\frac{M}{2}+1}}\right)\ .\end{aligned}$$

5.4.3 Step 3 : Replacing the infinite product

The last approximation will to be replace the infinite product expansion for $\hat{w}(u)$ by a finite product and a compensating polynomial. More precisely, we want

$$\hat{w}(u) \approx p(u) \prod_{0<\gamma\leq X} J_0\left(\frac{2\alpha_\gamma u}{\sqrt{\frac{1}{4}+\gamma^2}}\right)$$

where $p(u)$ is some compensating polynomial. In fact, we will take the polynomial to be of the form $p(u) = \sum_{m=0}^{A} b_m u^{2m}$ where we have the expansion

$$\prod_{\gamma>X} J_0\left(\frac{2\alpha_\gamma u}{\sqrt{\frac{1}{4}+\gamma^2}}\right) = \sum_{m=0}^{\infty} b_m u^{2m}$$

valid for $-C \leq u \leq C$. In all of our cases, we will choose X as large as possible. Clearly, this will depend on how many zeros of the appropriate L-functions we are able to compute. In addition, we will use the simplest approximation $p(u) = 1 + b_1 u^2$. Therefore, our final expression for the integral will be

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin(au)}{u}\hat{w}(u)\,du = \frac{1}{2\pi}\sum_{-C\leq n\epsilon\leq C} \epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\left(1+b_1(n\epsilon)^2\right)\prod_{0<\gamma\leq X} J_0\left(\frac{2\alpha_\gamma\epsilon n}{\sqrt{\frac{1}{4}+\gamma^2}}\right) + \mathrm{error}_1(\epsilon) + \mathrm{error}_2(\epsilon, C) + \mathrm{error}_3(\epsilon, C). \tag{5.30}$$

and

$$\mathrm{error}_3(\epsilon, C) = \frac{1}{2\pi}\sum_{-C\leq n\epsilon\leq C} \epsilon\frac{\sin(a\epsilon n)}{\epsilon n}\left(\sum_{m\geq 2} b_m(\epsilon n)^{2m}\right)\prod_{0<\gamma\leq X} J_0\left(\frac{2\alpha_\gamma\epsilon n}{\sqrt{\frac{1}{4}+\gamma^2}}\right).$$

As we are truncating after the second term, we need to bound the error arising from the higher order terms. If we consider the function defined by

$$\prod_{\gamma>X} \exp\left(\frac{1}{4}\left(\frac{2\alpha_\gamma u}{\sqrt{\frac{1}{4}+\gamma^2}}\right)\right) = \sum_{m=0}^{\infty} c_m u^{2m},$$

then we see by comparing Taylor series expansions that $|b_m| \leq |c_m|$. On the other hand, if we set $T_1 = T_1(X) = \sum_{\gamma>X} \frac{\alpha_\gamma^2}{\frac{1}{4}+\gamma^2}$, we have

$$\prod_{\gamma>X} \exp\left(\frac{1}{4}\left(\frac{2\alpha_\gamma u}{\sqrt{\frac{1}{4}+\gamma^2}}\right)\right) = \exp\left(u^2\sum_{\gamma>X}\frac{\alpha_\gamma}{\frac{1}{4}+\gamma^2}\right) = \exp(u^2 T_1).$$

This shows that $|b_m| \leq \frac{|T_1|^m}{m!}$. Therefore

$$\begin{aligned}\left|\sum_{m=A+1}^{\infty} b_m u^{2m}\right| < \sum_{m=A+1}^{\infty} \frac{T_1^m}{m!}|u|^{2m} &< \frac{(T_1u^2)^{A+1}}{(A+1)!}(1 + T_1u^2 + (T_1u^2)^2 + \cdots) \\ &= \frac{(T_1u^2)^{A+1}}{(A+1)!}\frac{1}{1-T_1u^2}\end{aligned} \tag{5.31}$$

if $T_1u^2 < 1$. In all cases we will take $A = 1$ and the above bound shows that

$$|\mathrm{error}_3| \leq \frac{1}{2\pi} \sum_{-C\leq n\epsilon\leq C} \epsilon \frac{|\sin(a\epsilon n)|}{\epsilon n} \prod_{0<\gamma\leq X} \left| J_0\left(\frac{2\alpha_\gamma \epsilon n}{\sqrt{\frac{1}{4}+\gamma^2}}\right)\right| \frac{(T_1(\epsilon n)^2)^2}{2(1-T_1(\epsilon n)^2)}$$

as long as $T_1C^2 < 1$.

5.4.4 Numerical examples

Here are some sample calculations of the various densities. All of these calculations were done using Maple.

S_3 examples

	X	ϵ	C	M	error$_1$	error$_2$	error$_3$	δ
$P_{23;2,1}$	972.2	$\frac{1}{30}$	20	76	$< 10^{-15}$	$< 10^{-28}$	$< 3.4 \times 10^{-6}$	0.98309
$P_{23;2,3}$	972.2	$\frac{1}{30}$	20	76	$< 10^{-21}$	$< 10^{-22}$	$< 2.2 \times 10^{-6}$	0.72469
$P_{23;3,1}$	972.2	$\frac{1}{30}$	20	42	$< 10^{-28}$	$< 10^{-42}$	$< 9.6 \times 10^{-6}$	0.95704
$P_{31;2,1}$	1071.83	$\frac{1}{30}$	20	80	$< 10^{-14}$	$< 10^{-30}$	$< 1.5 \times 10^{-6}$	0.96647
$P_{31;2,3}$	1071.83	$\frac{1}{30}$	20	80	$< 10^{-19}$	$< 10^{-23}$	$< 1.4 \times 10^{-6}$	0.68634
$P_{31;3,1}$	1071.83	$\frac{1}{30}$	20	44	$< 10^{-27}$	$< 10^{-44}$	$< 5.5 \times 10^{-6}$	0.94144
$P_{59;2,1}$	793.8	$\frac{1}{30}$	20	90	$< 10^{-10}$	$< 10^{-34}$	$< 1.4 \times 10^{-6}$	0.92876
$P_{59;2,3}$	793.8	$\frac{1}{30}$	20	90	$< 10^{-14}$	$< 10^{-27}$	$< 8.0 \times 10^{-7}$	0.65110
$P_{59;3,1}$	793.8	$\frac{1}{30}$	20	49	$< 10^{-22}$	$< 10^{-50}$	$< 6.1 \times 10^{-6}$	0.91087
$P_{83;2,1}$	737.33	$\frac{1}{30}$	20	96	$< 10^{-8}$	$< 10^{-37}$	$< 1.4 \times 10^{-6}$	0.89532
$P_{83;2,3}$	737.33	$\frac{1}{30}$	20	96	$< 10^{-11}$	$< 10^{-29}$	$< 8.7 \times 10^{-7}$	0.62386
$P_{83;3,1}$	737.33	$\frac{1}{30}$	20	52	$< 10^{-19}$	$< 10^{-53}$	$< 5.7 \times 10^{-6}$	0.88910

Class group examples

	X	ϵ	C	M	error$_1$	error$_2$	error$_3$	δ
$P_{\mathbb{Q}(\sqrt{-15});n,p}$	10000	$\frac{1}{20}$	20	68	$< 10^{-19}$	$< 7.0 \times 10^{-10}$	$< 2.4 \times 10^{-7}$	0.973286
$P_{\mathbb{Q}(\sqrt{-20});n,p}$	10000	$\frac{1}{20}$	20	41	$< 10^{-25}$	$< 3.3 \times 10^{-10}$	$< 1.8 \times 10^{-7}$	0.963473

Non-residues versus Residues mod p

	X	ϵ	C	M	error$_1$	error$_2$	error$_3$	δ
$P_{3;N,R}$	10000	$\frac{1}{20}$	20	26	$< 10^{-12}$	$< 1.4 \times 10^{-5}$	$< 3.0 \times 10^{-7}$	0.999063
$P_{4;N,R}$	10000	$\frac{1}{20}$	20	30	$< 10^{-50}$	$< 3.1 \times 10^{-6}$	$< 1.7 \times 10^{-7}$	0.995928
$P_{5;N,R}$	10000	$\frac{1}{20}$	20	30	$< 10^{-30}$	$< 1.5 \times 10^{-5}$	$< 1.7 \times 10^{-7}$	0.995422
$P_{23;N,R}$	2549.42	$\frac{1}{20}$	25	50	$< 10^{-50}$	$< 10^{-11}$	$< 2.3 \times 10^{-7}$	0.90318
$P_{29;N,R}$	2520.02	$\frac{1}{20}$	25	50	$< 10^{-43}$	$< 10^{-12}$	$< 2.1 \times 10^{-7}$	0.83894
$P_{31;N,R}$	2523.03	$\frac{1}{20}$	25	50	$< 10^{-44}$	$< 10^{-12}$	$< 1.8 \times 10^{-7}$	0.85507
$P_{37;N,R}$	2527.23	$\frac{1}{20}$	25	50	$< 10^{-44}$	$< 10^{-12}$	$< 1.7 \times 10^{-7}$	0.85460
$P_{59;N,R}$	2499.86	$\frac{1}{20}$	25	50	$< 10^{-31}$	$< 10^{-13}$	$< 7.4 \times 10^{-8}$	0.79420
$P_{61;N,R}$	2519.68	$\frac{1}{20}$	25	50	$< 10^{-37}$	$< 10^{-13}$	$< 1.1 \times 10^{-7}$	0.83013
$P_{83;N,R}$	2864.34	$\frac{1}{20}$	25	50	$< 10^{-25}$	$< 10^{-14}$	$< 5.6 \times 10^{-8}$	0.74696
$P_{89;N,R}$	2670.21	$\frac{1}{20}$	25	50	$< 10^{-36}$	$< 10^{-14}$	$< 9.3 \times 10^{-8}$	0.82555

Chapter 6

Chebyshev's Bias in Class Groups

6.1 Bias formulae

Let K be a number field. Let $\mathcal{H}_K$ be the corresponding class group with class number $h = h_K = |\mathcal{H}_K|$. Denote the ideal classes of K as $\mathfrak{a}_1, \mathfrak{a}_2, \ldots \mathfrak{a}_r$. As in the classical race between primes in arithmetic progressions and in the Chebotarev case, we can race primes in ideal classes. Specifically, we are considering sets of the form

$$\{x \geq 2 \mid \pi_{\mathfrak{a}_1}(x) > \pi_{\mathfrak{a}_2}(x) > \cdots > \pi_{\mathfrak{a}_r}(x)\}$$

where

$$\pi_{\mathfrak{a}_i}(x) = \sum_{\mathbb{N}\mathfrak{p} \leq x,\ \mathfrak{p} \in \mathfrak{a}_i} 1$$

is the prime counting function and $1 \leq i \leq r$. We would like to derive a formula for $\pi_{\mathfrak{a}_i}(x) - \pi_{\mathfrak{a}_j}(x)$ and be able to determine when there is a bias. The reasoning is analogous to the reasoning in Chapter 5. Fix an arbitrary ideal class $\mathfrak{a}$. Define the functions $\psi_{\mathfrak{a}}(x) = \sum_{\mathbb{N}\mathfrak{p}^m \leq x,\ \mathfrak{p}^m \in \mathfrak{a}} \log(\mathbb{N}\mathfrak{p})$ and $\theta_{\mathfrak{a}}(x) = \sum_{\mathbb{N}\mathfrak{p} \leq x,\ \mathfrak{p} \in \mathfrak{a}} \log(\mathbb{N}\mathfrak{p})$. As before,

$$\psi_{\mathfrak{a}}(x) = \theta_{\mathfrak{a}}(x) + \sum_{\mathbb{N}\mathfrak{p}^2 \leq x,\ \mathfrak{p}^2 \in \mathfrak{a}} \log(\mathbb{N}\mathfrak{p}) + O(x^{\frac{1}{3}}) \ . \tag{6.1}$$

Now consider the group homomorphism defined by the square map

$$\mathrm{sq} : \mathcal{H}_K \to \mathcal{H}_K \text{ where } \mathrm{sq}(\mathfrak{b}) = \mathfrak{b}^2$$

and $\mathfrak{b}$ is an ideal class. If $\mathfrak{a}$ is an ideal class, then $\mathrm{sq}^{-1}(\mathfrak{a})$ denotes the inverse image of $\mathfrak{a}$ under this map. Precisely,

$$\mathrm{sq}^{-1}(\mathfrak{a}) = \{\mathfrak{b} \subset \mathcal{H}_K \mid \mathfrak{b}^2 = \mathfrak{a}\ \} \ .$$

Therefore, we have

$$\sum_{\mathbb{N}\mathfrak{p}^2 \leq x,\ \mathfrak{p}^2 \in \mathfrak{a}} \log(\mathbb{N}\mathfrak{p}) = \sum_{\mathbb{N}\mathfrak{p} \leq x^{\frac{1}{2}},\ \mathfrak{p} \in \mathrm{sq}^{-1}(\mathfrak{a})} \log(\mathbb{N}\mathfrak{p}) = \frac{|\mathrm{sq}^{-1}(\mathfrak{a})|}{|\mathcal{H}_K|} x^{\frac{1}{2}} + O(x^{\frac{1}{4}+\epsilon}) \tag{6.2}$$

by the prime number theorem for ideal classes. In the preceding equation, it is possible for $\mathrm{sq}^{-1}(\mathfrak{a})$ to be empty. In that case, set $|\mathrm{sq}^{-1}(\mathfrak{a})| = 0$. Therefore, rearranging (6.1) gives

$$\theta_{\mathfrak{a}}(x) = \psi_{\mathfrak{a}}(x) - \frac{|\mathrm{sq}^{-1}(\mathfrak{a})|}{|\mathcal{H}_K|} x^{\frac{1}{2}} + O(x^{\frac{1}{3}}) \ . \tag{6.3}$$

Now let $\widehat{\mathcal{H}}_K = \{\phi \mid \phi : \mathcal{H}_K \to \mathbb{C}^*\}$ be the character group of $\mathcal{H}_K$. This is a group of group homomorphisms. In fact, $\widehat{\mathcal{H}}_K \cong \mathcal{H}_K$. Define $\psi(x,\chi) = \sum_{\mathbb{N}\mathfrak{p}^m \leq x} \chi(\mathfrak{p}^m) \log(\mathbb{N}\mathfrak{p})$. It should be understood that in the expression $\phi(\mathfrak{p}^m)$, the ideal class of $\mathfrak{p}^m$ is evaluated. From orthogonality properties of characters we have,

$$\psi_{\mathfrak{a}}(x) = \frac{1}{|\mathcal{H}_K|} \sum_{\chi \in \widehat{\mathcal{H}}_K} \overline{\chi(\mathfrak{a})} \psi(x,\chi) \ . \tag{6.4}$$

Combining the preceding formulae and using $\pi_{\mathfrak{a}}(x) = \int_{2-}^{x} \frac{dt}{\log t}$ we obtain upon integration

$$\pi_{\mathfrak{a}}(x) = \frac{1}{|\mathcal{H}_K|} \sum_{\chi \in \widehat{\mathcal{H}}_K} \overline{\chi(\mathfrak{a})} \frac{\psi(x,\chi)}{\log x} - \frac{|\mathrm{sq}^{-1}(\mathfrak{a})|}{|\mathcal{H}_K|} \frac{x^{\frac{1}{2}}}{\log x} + O\left(\frac{x^{\frac{1}{2}}}{\log^2 x}\right) \ . \tag{6.5}$$

Since the derivation of the above formula is identical to Chapter 5, the details have been left out. As before, $\psi(x,\chi)$ is related to the logarithmic derivative of $L(s,\chi)$, the class group L-function associated to χ. The class group L function is defined as

$$L_K(s,\chi) = \sum_{\mathfrak{a} \subset \mathcal{O}_K} \frac{\chi(\mathfrak{a})}{\mathbb{N}\mathfrak{a}^s}$$

where $\mathfrak{a}$ ranges over ideals in $\mathcal{O}_K$. Unlike Artin L-functions, there is no question about the holomorphy of the class group L-function. A class group L-function is a particular case of a Hecke L-function. It is classical work due to Hecke that these functions satisfy a functional equation and are holomorphic if χ is not the identity character. Hecke's proof uses a multi-dimensional version of the Poisson summation formula (see Chapter 8 of Lang [45]). Consequently, the proof is directly related to Riemann's original proof of the functional equation and analytic continuation of the Riemann zeta function. Interestingly, Tate, in his Ph.D. thesis, generalized Hecke's results working from an idelic point of view. As in the previous chapter, it is standard to prove that

$$\psi(x,\chi) = -\sum_{|\gamma_\chi| \leq X} \frac{x^\rho}{\rho} + O\left(\frac{x\log^2(xX)}{X} + \log x\right) \tag{6.6}$$

where $\rho_\chi = \beta_\chi + i\gamma_\chi$ runs over the zeros of $L_K(s,\chi)$ in $0 < Re(s) < 1$ and O depends on K. Assuming the Riemann Hypothesis for $L_K(s,\chi)$ we get

$$\psi(x,\chi) = -\sqrt{x} \sum_{|\gamma_\chi| \leq X} \frac{x^{i\gamma_\chi}}{\frac{1}{2} + i\gamma_\chi} + O\left(\frac{x\log^2(xX)}{X} + \log x\right) \ . \tag{6.7}$$

From equations (6.5) and (6.7) we obtain the following formula for the normalized function

$$E_{K;1,\dots,r}(x) = \frac{\log x}{x^{\frac{1}{2}}}(h_K\pi_{\mathfrak{a}_1}(x) - \pi(x), \dots, h_K\pi_{\mathfrak{a}_r}(x) - \pi(x))$$

where $r \le h_K$. As in the previous chapter, the existence of a limiting distribution $\mu_{K;1,\dots,r}$ defined on $\mathbb{R}^r$ can be proven. As before, it satisfies

$$\lim_{X\to\infty} \frac{1}{\log X}\int_{\log 2}^{X} f(E_{K;1,\dots,r}(x))\frac{dx}{x} = \int_{\mathbb{R}^r} f(x)\ d\mu_{K;1,\dots,r}(x)$$

for all bounded Lipschitz continous functions f. If we just wanted to compare primes that lie in two ideals classes $\mathfrak{a}_1$ and $\mathfrak{a}_2$ we can derive the formula

$$\begin{aligned} E_{K;1,2}(x) &= \frac{\log x}{\sqrt{x}}(\pi_{\mathfrak{a}_1}(x) - \pi_{\mathfrak{a}_2}(x)) \\ &= \frac{|\mathrm{sq}^{-1}(\mathfrak{a}_2)| - |\mathrm{sq}^{-1}(\mathfrak{a}_1)|}{|\mathcal{H}_K|} - \frac{1}{|\mathcal{H}_K|}\sum_{\chi\in\hat{\mathcal{H}}_K^*}(\overline{\chi(\mathfrak{a}_1)} - \overline{\chi(\mathfrak{a}_2)})\sum_{|\gamma_\chi|\le X}\frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} \\ &\quad + O\left(\frac{\sqrt{x}\log^2(X)}{X} + \frac{1}{\log x}\right)\ . \end{aligned} \tag{6.8}$$

valid for $2 \le x \le X$. This formula will be analyzed in the next section. In the prime number case, we know the complete answer for two way races. The answer depends on whether a residue class is a square or non-square. We can also discover the complete answer for two way races in the complex quadratic field setting. This will depend on which genus an ideal class lies in.

6.1.1 Interpreting the Chebyshev bias term

We will now use the above formula to analyze when there are biases. In [62], the authors mention that there is no bias for odd class number and if the class number is two, then there is a bias towards non-principal ideals. Let's look at these two examples so that we may understand the Chebyshev bias term. Note that $\mathrm{sq} : \mathcal{H}_K \to \mathcal{H}_K$ is a group homomorphism. Furthermore, $\mathcal{H}_K$ is a finite abelian group.

Example h_K odd

Observe that, if the class number h_K is odd, sq is an isomorphism. If not, there would be a non-trivial element in the kernel of the map. This element would have order two and also would divide h_K. This is a contradiction. In this case, $\mathrm{sq}^{-1}(\mathfrak{a})$ would consist of a single element. Hence,

$$\frac{|\mathrm{sq}^{-1}(\mathfrak{a}_2)| - |\mathrm{sq}^{-1}(\mathfrak{a}_1)|}{|\mathcal{H}_K|} = \frac{1-1}{|\mathcal{H}_K|} = 0\ .$$

Therefore, no bias appears from the Chebyshev bias term. However, despite the comment in the Rubinstein-Sarnak paper that there is no bias in this case, there is still the possibility of a class group L-function having a central zero. If this were the case, then there would be an algebraic bias term arising from the central zero.

Example $h_K = 2$

In this example, there are only two ideal classes. They are $\mathfrak{a}_1 = 1$, the class of principal ideals, and $\mathfrak{a}_2$, the class of non-principal ideals. As the class number is two, $\mathrm{sq}(\mathfrak{a}) = \mathfrak{a}^2 = 1$ for all ideal classes. Thus $\mathrm{sq}^{-1}(\mathfrak{a}_1) = \mathcal{H}_K$ and $\mathrm{sq}^{-1}(\mathfrak{a}_2)$ is empty. Consequently,

$$\frac{|\mathrm{sq}^{-1}(\mathfrak{a}_2)| - |\mathrm{sq}^{-1}(\mathfrak{a}_1)|}{|\mathcal{H}_K|} = \frac{0-2}{2} = -1 \ .$$

The -1 explains the bias towards non-principal ideals.

Example The general case

Suppose K has class number $h_K = h_2 h_{odd}$ where $h_2 = 2^{e(K)}$ and $2 \nmid h_{odd}$. Since $\mathcal{H}_K$ is a finite abelian group, there is the decomposition

$$\mathcal{H}_K \cong \mathcal{H}_{K,2} \times \mathcal{H}_{K,odd}$$

where $\mathcal{H}_{K,odd}$ has order h_{odd} and $\mathcal{H}_{K,2}$ has order h_2. Since $\mathcal{H}_K$ can be written in this way, it suffices to consider the restriction of sq to the 2-part of $\mathcal{H}_K$. This is because $\mathrm{sq}|_{\mathcal{H}_{K,odd}}$ is an isomorphism. Note that $\mathcal{H}_{K,2}$ is a product of cyclic groups each having order equal to a power of two. Assume

$$\mathcal{H}_{K,2} \cong C_{2^{\alpha_1}} \times C_{2^{\alpha_2}} \times \ldots \times C_{2^{\alpha_k}}$$

where $e(K) = \alpha_1 + \alpha_2 + \ldots + \alpha_k$ and C_n is a cyclic group of n elements. Also assume that $C_{2^{\alpha_i}} =< x_i >$ where x_i is a generator of order 2^{α_i} for $1 \le i \le k$. Consider $sq|_{\mathcal{H}_{K,2}} : \mathcal{H}_{K,2} \to \mathcal{H}_{K,2}$. The image of this map is

$$< x_1^2 > \times < x_2^2 > \times \ldots \times < x_k^2 > \ .$$

The image has order $2^{(\alpha_1 - 1)+\ldots+(\alpha_k - 1)} = \frac{h_2}{2^k}$ where k is the number of cyclic factors in the 2-part of the class group. It follows that the kernel of $\mathrm{sq}|_{\mathcal{H}_{K,2}}$ has order 2^k. The kernel can be explicitly written down. Observe that the elements $y_i = x_i^{2^{\alpha_i - 1}}$ for $1 \le i \le k$ are elements of order two. Thus, the square map sends them to 1. Hence,

$$\ker(\mathrm{sq}|_{\mathcal{H}_{K,2}}) =< y_1 > \times < y_2 > \times \ldots \times < y_k >$$

It follows that $\ker(\mathrm{sq}) = \ker(\mathrm{sq}|_{\mathcal{H}_{K,2}}) \times 1$, where 1 is the trivial subgroup of $\mathcal{H}_{K,odd}$. Likewise $\mathrm{im}(\mathrm{sq}) = \mathrm{im}(\mathrm{sq}|_{\mathcal{H}_{K,2}}) \times \mathcal{H}_{K,odd}$. This is a group of order $\frac{h_K}{2^k}$. From these observations, we see that the existence of a Chebyshev bias term will depend on the number of cyclic factors in the 2-part of the class group. Suppose $\mathfrak{a} \notin \mathrm{im}(\mathrm{sq})$. Then $\mathrm{sq}^{-1}(\mathfrak{a})$ is empty and $|\mathrm{sq}^{-1}(\mathfrak{a})| = 0$. On the other hand, if $\mathfrak{a} \in \mathrm{im}(\mathrm{sq})$, then there exists an element $\mathfrak{b}$ with $\mathfrak{b}^2 = \mathfrak{a}$. Therefore, every element of the coset $\mathfrak{b}\ker(\mathrm{sq})$ maps to $\mathfrak{a}$. This leads to $|\mathrm{sq}^{-1}(\mathfrak{a})| = |\mathfrak{b}\ker(\mathrm{sq})| = 2^k$. Let $\mathfrak{a}_1 \in \mathrm{im}(\mathrm{sq})$ and $\mathfrak{a}_2 \notin \mathrm{im}(\mathrm{sq})$, then the bias term is,

$$\frac{|\mathrm{sq}^{-1}(\mathfrak{a}_2)| - |\mathrm{sq}^{-1}(\mathfrak{a}_1)|}{|\mathcal{H}_K|} = \frac{0 - 2^k}{h_K} = -\frac{2^k}{h_K} \ .$$

This shows that the Chebyshev bias term is more pronounced if either there are many cyclic factors in the 2-part of the class group or if the class number is small.

6.1.2 Complex quadratic fields

For complex quadratic fields, there is a simple interpretation of when a bias occurs. Assume that $d < 0$ and $K = \mathbb{Q}(\sqrt{d})$. Consider the ideal class group,

$$\mathcal{H}_d := \mathcal{H}(\mathbb{Q}(\sqrt{d})) \ .$$

Instead of interpreting the class group as ideal classes, we will use the classical interpretation of binary quadratic forms. Hence, $\mathcal{H}_d$ consists of reduced primitive binary quadratic forms, $Q_1, Q_2, \ldots, Q_h$, where $Q_i(x,y) = a_i x^2 + b_i xy + c_i y^2$ for $1 \le i \le h$. Note that Q_i is reduced if $-|a_i| < b_i \le |a_i| < c_i$. For a given form Q, consider the set of values of $Q(x,y)$ mod d as x and y vary over all integers. We say that two forms Q_i and Q_j are in the same genus if they represent the same values in $(\mathbb{Z}/d\mathbb{Z})^*$. Let Q_1 denote the principal form. Set

$$H = \{Q_1(x,y) \bmod d \mid x, y \in \mathbb{Z}\ \} \ .$$

H is the set of values represented by Q_1. This is a subgroup of $(\mathbb{Z}/d\mathbb{Z})^*$. Furthermore, let $\chi : (\mathbb{Z}/d\mathbb{Z})^* \to \{\pm 1\}$ be the character defined by $\chi(p) = (\frac{d}{p})$ if $(p,d) = 1$. The kernel of this map, $\ker(\chi)$ is also a subgroup of $(\mathbb{Z}/d\mathbb{Z})^*$. In fact, there are the inclusions

$$H \subset \ker(\chi) \subset (\mathbb{Z}/D\mathbb{Z})^* \ .$$

It is proven in [10] p.34 that, the set of values of a given form $f(x,y)$ is represented by a coset of H in $\ker(\chi)$. This gives a map from $\Phi : \mathcal{H}_d \to \ker(\chi)/H$ sending a quadratic form $f(x,y)$ to the coset of H which represents its values mod d. Given a coset H' of $\ker(\chi)/H$, we can describe the various genera as the inverse image under Φ. Each genus is of the form $\Phi^{-1}(H')$. This description shows that all genera of quadratic forms consist of the same number of classes. This follows from Φ being a group homomorphism. Also, it is known that the number of genera of forms is a power of two. Interestingly, the number of genera is exactly equal to h_2, the order of the 2-part of the class group. The reason we are discussing genera is because it was proven by Gauss that the genus class of the principal form Q_1 exactly equals

$\mathcal{H}_d^2 = \mathrm{sq}(\mathcal{H}_d)$. It follows from our previous discussion, that an ideal class belonging to the principal genus will always lose to an ideal class in a different genus class. We will now present an example to demonstrate this phenomenon. The following example was taken from a paper by Chowla [6] pp. 29-30.

Chowla's example

Let $K = \mathbb{Q}(\sqrt{-2700})$. Chowla states that $\mathcal{H}_{-2700} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ and writes down the following reduced primitive binary quadratic forms

$$\begin{aligned} &C_1 = [1, 0, 675],\ C_2 = [25, 0, 27],\ C_3 = [13, 2, 52],\ C_4 = [4, 2, 169],\ C_5 = [7, 4, 97], \\ &C_6 = [9, 6, 76],\ C_7 = [19, 6, 36],\ C_8 = [25, 10, 28],\ C_9 = [25, 20, 31], C_{10} = [27, 18, 28] \end{aligned} \tag{6.9}$$

and, for $3 \le n \le 10$,

$$C_{n+8} = \overline{C_n}\ .$$

Note that, if $C = [a, b, c]$, then $\overline{C} = [a, -b, c]$. C_7 and C_3 are generators of this group of orders 3 and 6 respectively. There are the following relations

$C_1 = C_3^0$	$C_6 = C_7^2 C_3^2$
$C_2 = C_3^3$	$C_7 = C_7$
$C_3 = C_3$	$C_8 = C_7^2 C_3$
$C_4 = C_3^2$	$C_9 = C_7 C_3^2$
$C_5 = C_7^2 C_3^3$	$C_{10} = C_7 C_3$

Observe that $\mathcal{H}_{-2700}^2 = \mathrm{sq}(\mathcal{H}_{-2700})$ consists of elements of the form $C_7^a C_3^{2b}$ for $0 \le a \le 2$ and $0 \le b \le 2$. Using the above relations one determines that

$$\mathcal{H}_{-2700}^2 = \{C_1, C_4, C_6, C_7, C_9, C_{12}, C_{14}, C_{15}, C_{17}\}\ .$$

Note that there are two genus classes and this is the principal genus. For any $C \in \mathcal{H}_{-2700}^2$, $\mathrm{sq}^{-1}(C) = 2$, otherwise $\mathrm{sq}^{-1}(C) = 0$. Hence, each of the forms $C \in \mathcal{H}_{-2700}^2$ lose to the forms in the non-principal genus.

6.2 Central limit theorem

This section will focus on the limiting behaviour of primes in different ideal classes. We will specialize to imaginary quadratic fields. Consider a number field $K = \mathbb{Q}(d)$ with class number $h_K = h$ and discriminant $d_K = d$. Let $r \le h_K$ and $\mathfrak{a}_1, \ldots, \mathfrak{a}_r$ arbitrary ideal classes in $\mathcal{H}_K$. The function $E_{K;1,\ldots,r}(e^y)$ has a limiting distribution $\mu = \mu_{K;1,\ldots,r}(x)$ which has a Fourier transform

$$\hat{\mu}(\xi) = \exp\left(-i \sum_{m=1}^{r} \alpha(K, \mathfrak{a}_i)\right) \prod_{\chi \ne 1} \prod_{\gamma_\chi > 0} J_0 \left(\frac{2|\sum_{j=1}^{r} \chi(\mathfrak{a}_j)\xi_j|}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \right).$$

where

$$\alpha(K,\mathfrak{a}) = 1 - \mathrm{sq}^{-1}(\mathfrak{a}) + 2\sum_{\chi\neq 1}\chi(\mathfrak{a})\mathrm{ord}_{s=\frac{1}{2}}L(s,\chi)\ .$$

Let $\tilde{\mu}$ be the measure on $\mathbb{R}^r$ whose Fourier transform is

$$\hat{\tilde{\mu}}(\xi) = \hat{\mu}\left(\frac{\xi}{\sqrt{2h(d)\log d}}\right)\ .$$

Our goal is to prove a limit theorem analogous to Rubinstein-Sarnak [62] pp. 185-187. In the following theorem, we will need the modified LI. For each field discriminant $d < 0$, the non-trivial zeros of each class group L-function $L(s,\chi)$ can be written as $\frac{1}{2} + i\gamma_\chi$. The modified LI assumes that the non-zero γ_χ are linearly independent over $\mathbb{Q}$ for all characters χ of the class group. It should be noted that in the case of complex quadratic fields each of the functional equations of the class group L-functions have root number one. The real class group L-functions correspond to products of real Dirichlet L-functions. On the other hand, the complex class-group L-functions correspond to weight one modular forms which are linear combinations of theta functions. In fact, these are modular forms of weight one with odd character. Consequently, it seems unlikely that these L-functions can vanish at $\frac{1}{2}$. Below, we will also assume an upper bound on the sum $\sum_{\chi\neq 1}\mathrm{ord}_{s=\frac{1}{2}}L(s,\chi)$. One could dispense with the bound by assuming the ordinary LI for these class group L-functions as this would force the sum to be zero.

Theorem 6.2.1 *Assume GRH, the modified LI, and the upper bound,*

$$\sum_{\chi\neq 1}\mathrm{ord}_{s=\frac{1}{2}}L(s,\chi) = o(\sqrt{h\log d})\ .$$

Let $\tilde{\mu}$ be the probability measure defined above. Then $\tilde{\mu}$ converges in measure to the Gaussian $\frac{e^{-(x_1^2+\cdots+x_r^2)}}{(2\pi)^{\frac{r}{2}}}dx_1\ldots dx_r$ as $d\to\infty$, independently of the choice of ideal classes $\mathfrak{a}_1,\ldots,\mathfrak{a}_r$.

This theorem shows that as the field discriminant d gets large, biases between different ideal classes begin to disappear. Specifically, we can deduce that for r fixed,

$$\max_{\mathfrak{a}_1,\ldots,\mathfrak{a}_r\subset\mathbb{Q}(d)}\left|\delta(P_{\mathbb{Q}(d);1,\ldots,r}) - \frac{1}{r!}\right| \to 0 \text{ as } d\to\infty\ .$$

This is analogous to a result by Rubinstein-Sarnak in which they show biases disappear in the case of primes in residue classes mod q.

Proof Fix a large parameter A and consider $\xi\in\mathbb{R}^r$ with $|\xi|\le A$. Taking logarithms of the infinite product, we have

$$\log\hat{\tilde{\mu}}(\xi) = \frac{-i}{\sqrt{2h\log d}}\sum_{j=1}^{r}\alpha(K,\mathfrak{a}_j)\xi_j + \sum_{\chi\neq 1}\sum_{\gamma_\chi>0}\log J_0\left(\frac{2\left|\sum_{j=1}^{r}\chi(\mathfrak{a}_j)\xi_j\right|}{\sqrt{h\log d(\frac{1}{4}+\gamma_\chi^2)}}\right). \quad (6.10)$$

Recall that

$$\begin{aligned}\alpha(K,\mathfrak{a}_j) &= 1-\mathrm{sq}^{-1}(\mathfrak{a}_j)+2\sum_{\chi\neq 1}\chi(\mathfrak{a}_j)\mathrm{ord}_{s=\frac{1}{2}}\ L(s,\chi)\\ &\ll 2^{w(d)}+\sum_{\chi\neq 1}\mathrm{ord}_{s=\frac{1}{2}}\ L(s,\chi) \\ &\ll d^{\frac{(1+\epsilon)\log 2}{\log\log d}}+o(\sqrt{h\log d})\end{aligned} \tag{6.11}$$

where $w(d)$ is the number distinct prime divisors of d. The first bound is an elementary prime number estimate and the second bound was assumed to be true. To estimate the second term, note that $\log J_0(z)=\frac{-z^2}{4}+O(|z|^4)$ if $|z|$ is sufficiently small. In this case, we have

$$z=\frac{2\left|\sum_{j=1}^{r}\chi(\mathfrak{a}_j)\xi_j\right|}{\sqrt{2h\log d(\frac{1}{4}+\gamma_\chi^2)}}\leq\frac{2\sqrt{r}A}{\sqrt{h\log d}}\to 0\ .$$

We certainly can choose d large enough to guarantee that the right hand side is smaller than $\frac{1}{10}$ for each of the terms. Combining these observations we arrive at

$$\begin{aligned}\log\hat{\hat{\mu}}(\xi) &= -\sum_{\chi\neq 1}\sum_{\gamma_\chi>0}\frac{|\sum_{j=1}^{r}\chi(\mathfrak{a}_j)\xi_j|^2}{2h\log d(\frac{1}{4}+\gamma_\chi^2)}+\frac{o(\sqrt{h\log d})\sqrt{r}A}{\sqrt{2h\log d}}\\ &\quad +O\left(\frac{A^4}{(h\log d)^2}\sum_{\chi\neq 1}\sum_{\gamma_\chi>0}\frac{1}{(\frac{1}{4}+\gamma_\chi^2)^2}\right)\ .\end{aligned} \tag{6.12}$$

Applying the estimate of Chapter 4

$$\frac{1}{2}\sum_{\text{all }\gamma}\frac{1}{\frac{1}{4}+\gamma^2}=\log d+O\left(\log\log d\right)$$

shows that the third term is $\ll\frac{1}{h\log d}$. In addition, one shows that the main term equals

$$\begin{aligned}&-\frac{1}{2h\log d}\sum_{\chi\neq 1}\left|\sum_{j=1}^{r}\chi(\mathfrak{a}_j)\xi_j\right|^2\left(\log d+O\left(\log\log d\right)\right)\\ &=\left(-\frac{1}{2}+O\left(\frac{\log\log d}{\log d}\right)\right)\sum_{j=1}^{r}\sum_{k=1}^{r}\xi_j\xi_k\frac{1}{h}\sum_{\chi\neq 1}\chi(\mathfrak{a}_j)\overline{\chi(\mathfrak{a}_k)}\ .\end{aligned} \tag{6.13}$$

Orthogonality of characters yields

$$\frac{1}{h}\sum_{\chi\neq 1}\chi(\mathfrak{a}_j)\overline{\chi(\mathfrak{a}_k)}=\begin{cases}\frac{h-1}{h} & \text{if } j=k\\ -1 & \text{otherwise}\end{cases}\ .$$

Split the main terms into two terms I and II where the first term consists of pairs $j = k$ and the second term consists of all other pairs. Clearly,

$$\begin{aligned} \mathrm{I} &= \sum_{j=1}^{r} \xi_j^2 \left(1 - \frac{1}{h}\right) \\ \mathrm{II} &= -\frac{1}{h} \sum_{j \neq k} \xi_j \xi_k \ll \frac{1}{h} \left(\sum_{j=1}^{r} |\xi_j|\right)^2 \leq \frac{rA^2}{h} \end{aligned} \tag{6.14}$$

Combining these formulas shows that as $|d| \to \infty$

$$\log \hat{\tilde{\mu}}(\xi) \to -\sum_{j=1}^{r} \frac{1}{2} \xi_j^2 \text{ and } \hat{\tilde{\mu}}(\xi) \to \exp\left(-\sum_{j=1}^{r} \frac{1}{2} \xi_j^2\right)$$

for $|\xi| \leq A$. Hence, Levy's theorem implies that $\hat{\tilde{\mu}}$ converges in measure to the Gaussian.

Rubinstein-Sarnak assume LI for Dirichlet L-functions which implies $L(\frac{1}{2}, \chi) \neq 0$ for all Dirchlet characters. However, the proof of their theorem would remain valid under the assumption

$$\sum_{\chi} \mathrm{ord}_{s=\frac{1}{2}} L(s, \chi) = o(\sqrt{\phi(q) \log q})$$

and the linear independence of the non-zero imaginary ordinates. Note that this assumption is quite strong. It says that as $q \to \infty$, 0-density of the Dirichlet L-functions vanish at $\frac{1}{2}$. Currently, it is only known that more than one half of these functions do not vanish at $\frac{1}{2}$. This is work due to Iwaniec-Sarnak [37].

6.2.1 Weil's explicit formulae

This section will explore what kind of upper bound can be obtained for the sum

$$\sum_{\chi \neq 1} \mathrm{ord}_{s=\frac{1}{2}} L(s, \chi) \ .$$

We will need to use Weil's explicit formulae.

Let M and M' be two non-negative integers, A and B two positive real numbers, $(a_i)_{1 \leq i \leq M}$ and $(a_i')_{1 \leq i \leq M}$ two sequences of non-negative real numbers such that $\sum_{i=1}^{M} a_i = \sum_{i=1}^{M} a_i'$. Finally, let $(b_i)_{1 \leq i \leq M}$ and $(b_i')_{1 \leq i \leq M}$ be two sequences of complex numbers with non-negative real part.

Suppose there exist two meromorphic functions Λ_1 and Λ_2 verifying the following conditions:

1. There exists $w \in \mathbb{C}^*$ so that $\Lambda_1(1-s) = w\Lambda_1(s)$,
2. Λ_1 and Λ_2 have only a finite number of poles ,
3. For $i = 1, 2$, Λ_i minus its singular terms is bounded inside any vertical strip of the form $-\infty < \sigma_0 <\leq \mathrm{Re}(s) \leq \sigma_1 < \infty$,
4. There exists $c \geq 0$ such that, for $\mathrm{Re}(s) > 1 + c$ we have:
   $\Lambda_1(s) = A^s \prod_{j=1}^{M} \Gamma(a_j s + b_j) \prod_p \prod_{i=1}^{M'} (1 - \alpha_i(p)p^{-s})^{-1}$
   $\Lambda_2(s) = B^s \prod_{j=1}^{M} \Gamma(a_j' s + b_j') \prod_p \prod_{i=1}^{M'} (1 - \beta_i(p)p^{-s})^{-1}$
   where p runs over all prime numbers and $\alpha_i(p)$, $\beta_i(p)$ are complex numbers of modulus $\leq p^c$.

Let $F : \mathbb{R} \to \mathbb{R}$ be a function satisfying the following conditions:

1. There exists $\epsilon > 0$ such that $F(x)\exp((\frac{1}{2} + c + \epsilon)x)$ is integrable and has bounded variation, the value at each point being the average of the left-hand limit and the right-hand limit.
2. $(F(x) - F(0))/x$ has bounded variation.

We define

$$\begin{aligned} \Phi(s) &= \int_{-\infty}^{\infty} F(x) e^{(s-\frac{1}{2})X}\ dx \ , \\ I(a,b) &= -\frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \Phi(s) \frac{\Gamma'}{\Gamma}(as+b)\ ds \ , \\ J(a,b) &= -\frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \Phi(1-s) \frac{\Gamma'}{\Gamma}(as+b)\ ds \ . \end{aligned} \tag{6.15}$$

Weil generalized Riemann's original explicit formula to obtain the following useful result. In recent years, many analytic number theorists have applied this formula to obtain results about non-vanishing of L-functions and the zeros of primitive L-functions.

Riemann-Weil explicit formula 6.2.1.1

$$\begin{aligned} \sum_\rho \Phi(\rho) - \sum_\mu \Phi(\mu) + \sum_{i=1}^{M} I(a_i, b_i) + \sum_{i=1}^{M} J(a_i', b_i') = \\ F(0)\log(AB) - \sum_{p,i,k} (\alpha_i^k(p) F(k\log p) + \beta_i^k(p) F(-k\log p)) \frac{\log p}{p^{k/2}} \end{aligned} \tag{6.16}$$

where ρ runs over the zeros of Λ_1 and μ runs over the poles of Λ_1 in the critical strip $-c \leq \mathrm{Re}(s) \leq 1 + c$, each of them counted with multiplicity.

Comment The sum $\sum_\rho \Phi(\rho)$ should be interpreted as

$$\sum_\rho \Phi(\rho) = \lim_{T\to\infty} \sum_{|\mathrm{Im}(\rho)|<T} \Phi(\rho) \ .$$

Proof See pp. 16-19 of Sica's [71] thesis.

This following result follows from the Weil explicit formula and depends on a technique used by Ram Murty to bound the analytic rank of $J_0(q)$. One limitation of this method is that it depends on the Riemann Hypothesis. It should be observed that upper bound is at least $\frac{h^{\frac{1}{2}}}{\sqrt{\log d}}$ larger than the hypothesis made in the previous theorem.

Lemma 6.2.1.2 *Let $K = \mathbb{Q}(d)$ be an imaginary quadratic field of discriminant d and class number h. Let $\mathcal{H}_K$ denote the class group and $\widehat{\mathcal{H}}_K$ its set of class group characters. Assume RH for the class group L-functions. Then*

$$\sum_{\chi \neq 1} \operatorname{ord}_{s=\frac{1}{2}} L(s,\chi) \leq h + o(h)\ .$$

Proof We will now apply the Weil formula to some specific functions. Note that if $F(u) = \max(1-|u|, 0)$ then its Fourier transform is

$$\phi(t) = \hat{F}(t) = \int_{\mathbb{R}} F(u) e^{-iut}\, du = \left(\frac{\sin(\frac{t}{2})}{\frac{t}{2}}\right)^2 .$$

Instead of using this function in the explicit formula, we will work with the scaled function $F_x(u) = F(\frac{u}{\log x})$ where x is a free parameter to be determined later. It will be chosen as some function of D. This function has the Fourier transform

$$\phi_x(t) = \hat{F_x}(t) = \log x \cdot \phi(t \log x) = \log x \left(\frac{\sin(\frac{t\log x}{2})}{\frac{t \log x}{2}})\right)^2 = \frac{4}{\log x}\left(\frac{\sin(\frac{t\log x}{2})}{t}\right)^2 .$$

Now let χ be a fixed character of the class-group $\mathcal{H}_K$. Recall that if $\chi \neq 1$ then $L_K(s,\chi)$ is holomorphic and hence its completed L-function has no poles. However, if $\chi = 1$ the class group L-function is just the Dedekind zeta function of K. Therefore, its completed L-function has poles at $s = 0, 1$. Lastly, observe that the class group L-functions of K can all be completed to have the form

$$\Lambda_K(s,\chi) = \left(\frac{|d|}{2\pi}\right)^s \Gamma(s) L_K(s,\chi)\ .$$

Furthermore, it satisfies the functional equation

$$\Lambda_K(s,\chi) = \Lambda_K(1-s,\chi)\ .$$

Notice that the above functional equation indicates the root number is one. Further information on these functions can be found on pp. 211-222 of Iwaniec [36]. Observe that for $\rho = \frac{1}{2} + i\gamma$

$$\Phi_x(\rho) = \int_{\mathbb{R}} F_x(u) e^{i\gamma u}\, du = \phi_x(-\gamma) = \phi_x(\gamma)\ ,$$

and

$$\begin{aligned}\sum_{L(\frac{1}{2}+i\gamma,\chi)=0} \phi_x(\gamma) =& \delta_\chi(\Phi(0)+\Phi(1)) - I(1,0) - J(1,0) + F_x(0)\log\left(\frac{|d|}{4\pi^2}\right) \\ & - \sum_{N\mathfrak{a}\geq 1} (\chi(\mathfrak{a}) + \overline{\chi(\mathfrak{a})}) \frac{\Lambda(\mathfrak{a})}{N\mathfrak{a}^{\frac{1}{2}}} F_x(\log N\mathfrak{a}) \ .\end{aligned} \tag{6.17}$$

One can show that $I(1,0)+J(1,0) \ll 1$. Also, notice that

$$\begin{aligned}\Phi(1) = \Phi(0) = \Phi(\frac{1}{2}+i(i\frac{1}{2})) &= \frac{16}{\log x}\left(\frac{\sin(\frac{i\log x}{4})}{i}\right)^2 \\ &= \frac{16}{\log x}\sinh^2\left(\frac{\log x}{4}\right) = \frac{4\sqrt{x}}{\log x} + O\left(\frac{1}{\log x}\right) \ .\end{aligned} \tag{6.18}$$

Let $r_\chi = \mathrm{ord}_{s=\frac{1}{2}} L(s,\chi)$. Summing over characters χ we obtain

$$\begin{aligned}\log x \sum_\chi r_\chi \leq & \frac{4\sqrt{x}}{\log x} + O(h) + h\log\left(\frac{|d|}{4\pi^2}\right) \\ & - \sum_\chi \sum_{N\mathfrak{a}\geq 1} (\chi(\mathfrak{a}) + \overline{\chi(\mathfrak{a})}) \frac{\Lambda(\mathfrak{a})}{N\mathfrak{a}^{\frac{1}{2}}} F_x(\log N\mathfrak{a}) \ .\end{aligned} \tag{6.19}$$

Notice that the last sum is positive since it equals

$$\begin{aligned}& \sum_{1\leq N\mathfrak{a}\leq x} \left(\sum_\chi \chi(\mathfrak{a}) + \overline{\chi(\mathfrak{a})}\right) \frac{\Lambda(\mathfrak{a})}{N\mathfrak{a}^{\frac{1}{2}}} F_x(\log N\mathfrak{a}) \\ & = 2h \sum_{1\leq N\mathfrak{a}\leq x,\ \mathfrak{a}\text{ principal}} \frac{\Lambda(\mathfrak{a})}{N\mathfrak{a}^{\frac{1}{2}}} F_x(\log N\mathfrak{a}) \ \geq 0 \ .\end{aligned} \tag{6.20}$$

Thus, the final term can be discarded from the inequality to yield

$$\sum_\chi r_\chi \leq \frac{4\sqrt{x}}{(\log x)^2} + O\left(\frac{h}{\log x}\right) + \frac{h\log|d|}{\log x} \ .$$

Now choose $x = |d|$ and we have

$$\sum_\chi r_\chi \leq h + O\left(\frac{h}{\log|d|}\right) + O\left(\frac{|d|^{\frac{1}{2}}}{(\log|d|)^2}\right) \ .$$

Recall that under the assumption of GRH Littlewood [47] proved that the L-function at $s=1$ satisfies the inequalities

$$\frac{1}{\log\log|d|} \ll L(1, (\frac{d}{\cdot})) \ll \log\log|d|$$

which translates to

$$\frac{|d|^{\frac{1}{2}}}{\log\log|d|} \ll h(d) \ll |d|^{\frac{1}{2}} \log\log|d|$$

by the class number formula. These inequalities show that

$$\sum_{\chi} r_\chi \leq h + o(h) \ .$$

6.3 Numerical examples

6.3.1 Class number two

We will consider imaginary quadratic fields, $K = \mathbb{Q}(\sqrt{d})$ where d is the field discriminant and $h(d) = 2$. There are only a finite number of such fields. This was originally proven by Stark [74]. The complete list of field discriminants with $h(d) = 2$ is

$$\begin{aligned}&-15, -20, -24, -35, -40, -51, -52, -88, -91, -115, -123, -148, -187, -232,\\ &-235, -267, -403, -427.\end{aligned} \tag{6.21}$$

For such fields, there are only two ideal classes. The identity class is represented by principal ideals and the other class is represented by non-principal ideals. We will consider the race between primes that are principal and those which are non-principal. It was shown earlier that the Chebyshev bias is towards the non-principal primes. We will compute some examples of this. Let

$$\pi_n(x) = \sum_{N\mathfrak{p}\leq x,\, \mathfrak{p}\subset\mathcal{O}_K,\, \mathrm{non-principal}} 1 \quad \text{and} \quad \pi_p(x) = \sum_{N\mathfrak{p}\leq x,\, \mathfrak{p}\subset\mathcal{O}_K,\, \mathrm{principal}} 1.$$

Since the class number of these fields is two, there is only one non-trivial character of the class group. Denote this character by χ. Specifically,

$$\chi : \mathcal{H}_K \to \pm 1 \text{ where } \ \chi(\mathrm{principal}) = 1,\ \chi(\mathrm{non-principal}) = -1$$

From the formulas in the last section, we can show

$$\frac{\log x}{\sqrt{x}}\left(\pi_n(x) - \pi_p(x)\right) = 1 + \frac{\psi(x,\chi)}{\sqrt{x}} + \text{small} = 1 - \sum_{|\gamma_\chi|\leq X} \frac{x^{i\gamma_\chi}}{\frac{1}{2}+\gamma_\chi} + \text{small}. \tag{6.22}$$

Denote the function on the left as $E_K(x)$. The plus one in the above equation denotes the bias towards non-principal ideals. Also, the sum on the right hand side of the equation is over non-trivial zeros of the class group L-function, $L(s,\chi)$. The above equation is only valid under the assumption of RH for $L(s,\chi)$. For the fields we are considering, the class group L-function has a particularly simple interpretation. It is classical work due to Siegel that we can write $d = d_1 d_2$ where each d_i is a field discriminant. In addition, if we consider the Dirichlet characters $(\frac{d_i}{\cdot})$ for $i = 1, 2$ then

$$L(s,\chi) = L(s, (\frac{d_1}{\cdot}))L(s, (\frac{d_2}{\cdot})).$$

Because of this simple identity, it will be relatively easy to compute the logarithmic density of the sets $P_{K;np,p}$. As in previous sections, $E_K(e^y)$ has a limiting distribution, μ_K assuming RH. The Fourier transform of this distribution can be calculated to be

$$\hat{\mu}_K(\xi) = e^{i\xi} \prod_{\gamma_{d_1}>0} J_0\left(\frac{2\xi}{\sqrt{\frac{1}{4}+\gamma_{d_1}^2}}\right) \prod_{\gamma_{d_2}>0} J_0\left(\frac{2\xi}{\sqrt{\frac{1}{4}+\gamma_{d_2}^2}}\right).$$

Using the same techniques as the last chapter enables one to compute the following logarithmic densities. In the following table, d_K denotes the field discriminant of the field $\mathbb{Q}(d_K)$. We only considered fields with class number two. The second and third columns denote field discriminants d_1, d_2 such that

$$d_K = d_1 d_2 \text{ and } L(s,\chi) = L(s, (\frac{d_1}{\cdot}))L(s, (\frac{d_2}{\cdot})).$$

T denotes the height to which zeros of $L(s,\chi)$ have been computed and $N_{d_i}(T)$ denotes the number of zeros of $L(s(\frac{d_i}{\cdot}))$ to height T in the critical strip for $i = 1, 2$. The logarithmic densities in this section were computed using a program written in C. We wrote our own code to compute the J_0 Bessel function. The computation of J_0 is good to roughly 18 decimal digits. However, most of the computed zeros are only good to 10 decimal digits. The J_0 program is based on a similar version in Numerical Recipes in C [57] and some tables of rational approximations in Computer Approximations [28]. The C version of the program that computes the logarithmic densities is significantly faster than the corresponding Maple version.

d_K	d_1	d_2	T	$N_{d_1}(T)$	$N_{d_2}(T)$	$\delta_{K;n,p}$
-15	-3	5	10000	11891	12703	0.973286
-20	-4	5	10000	12349	12703	0.963473
-24	-3	8	10000	11891	13452	0.954865
-35	-7	5	10000	13239	12703	0.939922
-40	-8	5	10000	13452	12703	0.925852
-51	-3	17	2500	2421	3112	0.919848
-52	-4	13	10000	12349	14224	0.908316
-88	-11	8	10000	13959	13452	0.871237
-91	-7	13	10000	13239	14224	0.889205
-115	-23	5	2500	3232	2624	0.875912
-123	-3	41	2500	2421	3461	0.863227
-148	-4	37	2500	2536	3420	0.835824
-187	-11	17	2500	2938	3112	0.849390
-232	-8	29	2500	2811	3323	0.808134
-235	-47	5	2500	3516	2624	0.834246
-267	-3	89	2500	2421	3770	0.814913
-403	-31	13	2500	3351	3004	0.811309
-427	-7	61	2500	2759	3619	0.806438

Chapter 7

The Summatory Function of the Möbius Function

7.1 Introduction to $M(x)$

The function $M(x) = \sum_{n \le x} \mu(n)$ has close connections to the Riemann zeta function. In fact, the Riemann Hypothesis would follow from the famous Mertens conjecture which states that

$$|M(x)| \le x^{\frac{1}{2}} \text{ for } x \ge 1 \ .$$

Unfortunately, this conjecture was disproven by Odlyzko and te Riele [56]. Although the Mertens conjecture is false, there is still the possibility that the weak Mertens conjecture is true. The weak or averaged Mertens hypothesis states that

$$\int_2^X \left(\frac{M(t)}{t} \right)^2 dt \ll \log X \ .$$

It implies RH, all zeros of $\zeta(s)$ are simple, and that $\sum_{\gamma > 0} \frac{1}{|\rho \zeta'(\rho)|^2}$ converges. These are proven in Titchmarsh's book [76] (pp. 376 -380). It is very plausible that the weak Mertens conjecture is true. In fact, we will show that the weak Mertens conjecture is a consequence of RH and the Gonek-Hejhal conjecture. Moreover, it seems likely that the following asymptotic should be true.

$$\int_2^X \left(\frac{M(t)}{t} \right)^2 dt \sim \sum_{\gamma > 0} \frac{2}{|\rho \zeta'(\rho)|^2} \log X \ .$$

The reason for believing the above asymptotic formula is that Cramér proved assuming RH that

$$\int_2^X \left(\frac{\psi(t) - t}{t} \right)^2 dt \sim \sum_{\gamma > 0} \frac{2}{|\rho|^2} \log X \ .$$

The best known unconditional upper bound for $M(x)$ is

$$M(x) = O\left(x \exp(-c \log^{\frac{3}{5}} x (\log \log x)^{-\frac{1}{5}}) \right)$$

for some constant c (see Ivic [35] pp. 71-73). Whereas, RH implies

$$M(x) = O\left(x^{\frac{1}{2}} \exp\left(A\frac{\log x}{\log\log x}\right)\right)$$

for some constant A (see Titchmarsh [76] p. 371). However, it is known that the Riemann Hypothesis is equivalent to the estimate $|M(x)| \ll x^{\frac{1}{2}+\epsilon}$ for every $\epsilon > 0$ (see Titchmarsh [76] p.370). In fact, the true order of $M(x)$ appears to be something of a mystery. The best known unconditional omega result is $M(x) = \Omega(x^{\frac{1}{2}})$. However, if RH is false, then

$$M(x) = \Omega(x^{\theta})$$

for some $\theta > \frac{1}{2}$. In contrast, superior upper and lower bounds are known for the prime number sum $\psi(x)$. Recall that Littlewood proved

$$\psi(x) = x + \Omega_{\pm}(x^{\frac{1}{2}} \log\log\log x)\ .$$

One may ask what type of omega result is true for $M(x)$. In an attempt to understand this question we will give a conditional proof of the existence of a limiting distribution function for the function $\phi(y) = e^{-\frac{y}{2}} M(e^y)$. Heath-Brown writes in [30] "It appears to be an open question whether

$$x^{-\frac{1}{2}} M(x) = x^{-\frac{1}{2}} \sum_{n \le x} \mu(n)$$

has a distribution function. To prove this one would want to assume the Riemann Hypothesis and the simplicity of the zeros, and perhaps also a growth condtion on $M(x)$."

The constructed limiting distribution reveals significant information about $M(x)$. Studying tails of this distribution will lead to the conjecture

$$M(x) = \Omega_{\pm}\left(x^{\frac{1}{2}} (\log\log\log x)^{\frac{5}{4}}\right)\ .$$

This conjecture depends on deep conjectures of Gonek and Hejhal concerning the zeros of the Riemann zeta function. Additionally assuming the linear independence of the zeros of the zeta function almost leads to a proof of this result. Recall that in the Chebyshev's Bias article [62], the authors consider functions like $\phi(x) = \frac{\log x}{\sqrt{x}}(\pi(x) - \mathrm{Li}(x))$. They show limiting distributions exist for $\phi(e^y)$ rather than $\phi(x)$. In this case, a limiting distribution does not exist for $\phi(x)$. The technique that will be used in showing $\phi(y) = e^{-\frac{y}{2}} M(e^y)$ has a limiting distribution is essentially the same as in the Rubinstein-Sarnak paper. The existence of limiting distributions in their article depends on the Riemann Hypothesis. For the function, $\phi(y)$, we will need to assume the Riemann Hypothesis and the Gonek-Hejhal conjecture mentioned earlier.

7.2 Gonek-Hejhal conjectures

Originally, Hejhal [32] derived these conjectures heuristically from his results on the value distribution of the function $\log \zeta'(\frac{1}{2}+it)$.

Hejhal's Conjectures 7.2.1
(i)

$$\sum_{\gamma>0} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}}$$

converges for all $\lambda > 1$ *and*
(ii)

$$\sum_{T\leq\gamma\leq 2T} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} \asymp T^{1-\lambda}(\log T)^{\lambda^2/4+1-\lambda}.$$

Note that part *(ii)* implies that

$$\sum_{T\leq\gamma} \frac{1}{|\rho\zeta'(\rho)|^{2}} \asymp \frac{1}{T}+\frac{1}{2T}+\frac{1}{4T}+\frac{1}{8T}+\cdots \ll \frac{1}{T}\ .$$

Gonek made some related conjectures, assuming that all zeros of the zeta function are simple. Define the function

$$J_{-k}(T)=\sum_{\gamma\leq T} \frac{1}{|\zeta'(\rho)|^{2k}}\ .$$

Independently and using different techniques Gonek conjectured that

Gonek's Conjectures 7.2.2
(i)

$$J_{-1}(T) \sim \frac{3}{\pi^3}T$$

and
(ii)

$$J_{-k}(T) \asymp T \log T^{(k-1)^2}\ .$$

In Gonek's article [26], one of the main results is that

$$J_{-1}(T) \gg T$$

assuming the Riemann Hypothesis. However, no non-trivial upper bounds for this function are known. Even a weak upper bound of the form $J_{-1}(T) \ll T^c$ for some $1 \leq c \leq 2$ would be an interesting result. This type of problem can be considered as a discrete analogue of finding bounds for negative moments of the zeta function. Evidence for the first part of the conjecture was presented in a recent talk [27] at the MSRI. The techniques used are very similar to Montgomery's work on the Pair Correlation conjecture. It should be observed that part (i) of Gonek's conjecture implies part (i) of Hejhal's conjecture. Recently, there have been some more precise conjectures made about $J_{-k}(T)$. The applied mathematicians Hughes, Keating, and O'Connell [33], using random matrix models have made the following conjecture.

Random Matrix Model Conjecture 7.2.3
For $k > -\frac{3}{2}$ and bounded,

$$J_k(T) = \sum_{\gamma \leq T} |\zeta^{'}(\rho)|^{2k} \sim \frac{G^2(k+2)}{G(2k+3)} \cdot a(k) \cdot N(T) \cdot \left(\log \frac{T}{2\pi}\right)^{k(k+2)}$$

as $T \to \infty$,where G is Barnes' function defined by

$$G(z+1) = (2\pi)^{z/2} \exp\left(-\frac{1}{2}(z^2 + \gamma z^2 + z)\right) \prod_{n=1}^{\infty} (1+\frac{z}{n})^n e^{-z+z^2/2n}$$

and

$$a(k) = \prod_p \left(1 - \frac{1}{p}\right)^{k^2} \sum_{m=0}^{\infty} \left(\frac{\Gamma(m+k)}{m!\Gamma(k)}\right)^2 p^{-m}.$$

One should note that in the above definition of $a(k)$, the fraction $\frac{\Gamma(m+k)}{m!\Gamma(k)}$ may be indeterminate if $k = 0$ or -1. In these cases, set

$$\frac{\Gamma(m+k)}{m!\Gamma(k)} = \lim_{u \to k} \frac{\Gamma(m+u)}{m!\Gamma(u)} \ .$$

Furthermore, one can check that $G(1) = 1$ and $a(-1) = \frac{6}{\pi^2}$ and hence the above conjecture implies that $J_{-1}(T) \sim \frac{3}{\pi^3} T$. Consequently, the random matrix model conjecture implies part (i) of Gonek's conjecture. This is an amazing agreement, considering that Gonek's techniques are completely different from the random matrix model techniques.

The following lemma demonstrates the connection between Gonek and Hejhal's aforementioned conjectures.

Lemma 7.2.4
(i)

$$J_{-1}(T) \ll T \Longrightarrow \sum_{\gamma>0} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} \text{ converges for all } \lambda > 1 \ .$$

(ii)

$$\begin{aligned} &J_{-k}(T) \ll T \log T^{(k-1)^2} \text{ for } k \geq 0, k \neq \frac{1}{2} \\ &\Longleftrightarrow \\ &\sum_{T\leq\gamma\leq 2T} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} \ll T^{1-\lambda}(\log T)^{\lambda^2/4+1-\lambda} \text{ for } \lambda \geq 0, \lambda \neq 1 \ . \end{aligned} \tag{7.1}$$

Proof The proof will be broken into two cases. First consider $1 < \lambda \leq 2$. Let $\mu = 2 - \lambda$. Consider the tail, $\mathrm{Tail}(\lambda, T) = \sum_{\gamma\geq T} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}}$. Rewrite as

$$\sum_{\gamma\geq T} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} = \sum_{\gamma\geq T} \frac{|\zeta'(\rho)|^{\mu}}{|\rho|^{\lambda}|\zeta'(\rho)|^2} \ll \sum_{\gamma\geq T} \frac{|\rho|^{\epsilon}}{|\rho|^{\lambda}|\zeta'(\rho)|^2} \ .$$

The last step is a consequence of the Riemann Hypothesis. Applying partial summation we obtain

$$\mathrm{Tail}(\lambda, T) \ll \left[\frac{J_{-1}(t)}{|\frac{1}{2}+it|^{\lambda-\epsilon}}\right]_T^{\infty} + \lambda \int_T^{\infty} \frac{J_{-1}(t)}{(\frac{1}{4}+t^2)^{\lambda/2+1}} \cdot t \ dt \ .$$

The first term approaches zero as $T \to \infty$ since $\lambda > 1$. Also, the integrand is asymptotic to $\frac{3}{\pi^3 t^{\lambda}}$ as $t \to \infty$. Consequently, the integral is less than a constant times $\frac{1}{T^{\lambda-1}}$. Therefore, the tail goes to zero and we have the convergence of the series for $1 < \lambda \leq 2$. Now consider the case $\lambda \geq 2$. The convergence of the series for $\lambda = 2$ implies that

$$\frac{1}{|\rho\zeta'(\rho)|} \to 0 \ .$$

Choose T such that $\frac{1}{|\rho\zeta'(\rho)|} \leq 1$ for $\gamma \geq T$. Then there is the trivial bound

$$\sum_{\gamma\geq T} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} \leq \sum_{\gamma\geq T} \frac{1}{|\rho\zeta'(\rho)|^2} \ .$$

However, the second sum goes to zero by the previous discussion. In the second part of the argument, the convergence of $\sum_{\gamma} \frac{1}{|\rho\zeta'(\rho)|^2}$ is all that is required. The asymptotic is only required for the range $1 < \lambda \leq 2$.

(ii) For $\Rightarrow$ we partially sum to obtain

$$\sum_{T\le\gamma\le 2T}\frac{1}{|\rho\zeta'(\rho)|^\lambda} \ll \sum_{T\le\gamma\le 2T}\frac{1}{|\gamma\zeta'(\rho)|^\lambda} = \left[\frac{J_{\lambda/2}(t)}{t^\lambda}\right]_T^{2T} + \lambda\int_T^{2T}\frac{J_{\lambda/2}(t)}{t^{\lambda+1}}\,dt\ .$$

The first term is

$$\ll \frac{T(\log T)^{(\lambda/2-1)^2}}{T^\lambda} = T^{1-\lambda}(\log T)^{(\lambda/2-1)^2}$$

and the second term is

$$\ll \int_T^{2T}\frac{t(\log t)^{(\lambda/2-1)^2}}{t^{\lambda+1}}\,dt = \int_T^{2T}\frac{(\log t)^{(\lambda/2-1)^2}}{t^\lambda}\,dt = T^{1-\lambda}(\log T)^{(\lambda/2-1)^2}\ .$$

For the other direction, we substitute the values $\frac{T}{2}, \frac{T}{4}, \ldots \frac{T}{2^j}$ where j is the least integer such that $\frac{T}{2^j} \le 14$ and obtain

$$\sum_{\gamma<T}\frac{1}{|\rho\zeta'(\rho)|^\lambda} \ll \sum_{k=1}^{j}\left(\frac{T}{2^k}\right)^{1-\lambda}\left(\log\left(\frac{T}{2^k}\right)\right)^\mu$$

where $\mu = \lambda^2/4 + 1 - \lambda$. We make the variable change $U = \frac{T}{2^j}$. Notice that this is a number that satisfies $7 \le U \le 14$ and obtain

$$\sum_{\gamma<T}\frac{1}{|\rho\zeta'(\rho)|^\lambda} \ll \sum_{k=1}^{j}\left(2^kU\right)^{1-\lambda}\left(\log\left(2^kU\right)\right)^\mu \ll \sum_{k=1}^{j}\left(2^k\right)^{1-\lambda}k^\mu \ll \int_1^j x^\mu e^{x(1-\lambda)}\,dx\ .$$

However, integrating the last term by parts shows that

$$\sum_{\gamma<T}\frac{1}{|\rho\zeta'(\rho)|^\lambda} \ll j^\mu e^{j(1-\lambda)} \ll T^{1-\lambda}(\log T)^\mu$$

for $\lambda \ne 1$. We now apply partial summation to obtain

$$\begin{aligned}\sum_{\gamma<T}\frac{1}{|\zeta'(\rho)|^\lambda} &\ll \sum_{\gamma<T}|\gamma|^\lambda\frac{1}{|\rho\zeta'(\rho)|^\lambda} = O\left(T^\lambda T^{1-\lambda}(\log T)^\mu\right) + O\left(\int_1^T t^{\lambda-1}t^{1-\lambda}(\log T)^\mu\,dt\right)\\ &\ll T(\log T)^\mu\ .\end{aligned}\tag{7.2}$$

□

Lemma 7.2.5 *Let $\rho = \frac{1}{2} + i\gamma$ denote a zero of the zeta function with $\gamma > 0$.*
(a)

$$\sum_{\gamma<T}\frac{1}{|\rho|} = \sum_{\gamma<T}\frac{1}{\sqrt{\frac{1}{4}+\gamma^2}} = \frac{1}{4\pi}(\log T)^2 - \frac{\log 2\pi}{2\pi}\log T + O(1)\ .$$

(b) For $c > 1$

$$\sum_{\gamma > T} \frac{1}{|\rho|^c} = \sum_{\gamma > T} \frac{1}{(\frac{1}{4}+\gamma^2)^{\frac{c}{2}}} = \frac{1}{2\pi(c-1)} \frac{\log T}{T^{c-1}} + O\left(\frac{1}{T^{c-1}}\right) .$$

(c) $J_{-\frac{1}{2}}(T) = \sum_{0<\gamma<T} |\zeta'(\frac{1}{2}+i\gamma)|^{-1} \sim \alpha T (\log T)^{\frac{1}{4}}$ *implies*

$$\sum_{\gamma < T} \frac{1}{|\rho\zeta'(\rho)|} \sim \frac{4\alpha}{5} (\log T)^{\frac{5}{4}} .$$

(d) $J_{-1}(T) = \sum_{0<\gamma<T} |\zeta'(\frac{1}{2}+i\gamma)|^{-2} \sim \frac{3}{\pi^3} T$ *implies*

$$\sum_{\gamma > T} \frac{1}{|\rho\zeta'(\rho)|^2} \sim \frac{3}{\pi^3 T} .$$

(e) $J_{-\frac{1}{2}}(T) = \sum_{0<\gamma<T} |\zeta'(\frac{1}{2}+i\gamma)|^{-1} \ll T^u (\log T)^v$ *implies*

$$\sum_{\gamma > T} \frac{(\log \gamma)^a}{\gamma^b |\zeta'(\rho)|} \ll \frac{(\log T)^{a+v}}{T^{b-u}} .$$

Proof Note that $\frac{1}{\sqrt{\frac{1}{4}+x^2}} = \frac{1}{x} \cdot (1 + \frac{1}{4x^2})^{-\frac{1}{2}} = \frac{1}{x} \cdot (1 - \frac{1}{8x^2} + O(\frac{1}{x^4})) = \frac{1}{x} + O(\frac{1}{x^3})$. In this case, $x = \gamma$ the imaginary ordinate of a zero. Hence, $x \geq 14$ as the first zero occurs at $\gamma = 14.13....$ Evaluate the first term by partial summation and use Riemann's classical formula $N(t) = \frac{t}{2\pi} \log t - \frac{1+\log 2\pi}{2\pi} t + O(\log t)$ to obtain

$$\begin{aligned}\sum_{\gamma < T} \frac{1}{\gamma} &= \left[\frac{N(t)}{t}\right]_{\gamma_1^-}^{T} + \int_{\gamma_1^-}^{T} \frac{N(t)}{t^2}\, dt \\ &= \frac{1}{2\pi} \log T + \frac{1}{2\pi} \int_{\gamma_1^-}^{T} \frac{t \log(t)}{t^2}\, dt - \frac{1+\log 2\pi}{2\pi} \log T + O(1) \qquad (7.3) \\ &= \frac{1}{4\pi} (\log T)^2 - \frac{\log 2\pi}{2\pi} \log T + O(1) .\end{aligned}$$

In addition, note that $\sum_{\gamma>0} \frac{1}{\gamma^k}$ converges for $k > 1$. Hence, the other term only contributes $O(1)$ and we obtain part (a). Part (b) is proven likewise. Observe that $\frac{1}{|\rho|^c} = \frac{1}{\gamma^c}(1 + \frac{1}{4\gamma^2})^{-\frac{c}{2}} = \frac{1}{\gamma^c}\left(1 - \frac{c}{8\gamma^2} + O(\frac{1}{\gamma^3})\right) = \frac{1}{\gamma^c} + O(\frac{1}{\gamma^{c+2}})$ and hence

$$\begin{aligned}\sum_{T<\gamma<\infty} \frac{1}{|\gamma|^c} &= \left[\frac{N(t)}{t^c}\right]_T^{\infty} + c \int_T^{\infty} \frac{N(t)}{t^{c+1}} dt \\ &= -\frac{1}{2\pi} \frac{\log T}{T^{c-1}} + \frac{c}{2\pi(c-1)} \frac{\log T}{T^{c-1}} + O\left(\frac{1}{T^{c-1}}\right) \qquad (7.4) \\ &= \frac{1}{2\pi(c-1)} \frac{\log T}{T^{c-1}} + O\left(\frac{1}{T^{c-1}}\right) .\end{aligned}$$

For part (c) apply partial summation again. Set $\epsilon(T) = J_{-\frac{1}{2}}(T) - \alpha T(\log T)^{\frac{1}{4}} = o(T(\log T)^{\frac{1}{4}})$. Note that

$$\sum_{\gamma<T} \frac{1}{|\rho\zeta'(\rho)|} = \sum_{\gamma<T} \frac{1}{|\zeta'(\rho)\sqrt{\frac{1}{4}+\gamma^2}|} = \sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|\gamma\sqrt{1+\frac{1}{4\gamma^2}}} \ .$$

Using $(1+\frac{1}{4\gamma^2})^{-\frac{1}{2}} = 1 + O(\frac{1}{\gamma^2})$ we see that

$$\sum_{\gamma<T} \frac{1}{|\rho\zeta'(\rho)|} = \sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|\gamma} + O\left(\sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|\gamma^3}\right) \ .$$

Let Σ_1 denote the first sum. We have

$$\begin{aligned}
\Sigma_1 &= \left[\frac{J_{-\frac{1}{2}}(t)}{t}\right]_{\gamma_1^-}^{T} + \int_{\gamma_1^-}^{T} \frac{J_{-\frac{1}{2}}(t)}{t^2}\ dt \\
&= \frac{aT(\log T)^{\frac{1}{4}} + \epsilon(T)}{T} + \int_{\gamma_1^-}^{T} \frac{\alpha t(\log t)^{\frac{1}{4}} + \epsilon(t)}{t^2}\ dt \\
&= \alpha(\log T)^{\frac{1}{4}} + \frac{\epsilon(T)}{T} + \alpha\int_{\gamma_1^-}^{T} \frac{(\log t)^{\frac{1}{4}}}{t}\ dt + \int_{\gamma_1^-}^{T} \frac{\epsilon(t)}{t^2}\ dt \ .
\end{aligned} \tag{7.5}$$

Since $\epsilon(t) = o(t(\log t)^{\frac{1}{4}})$ it is clear that

$$\sum_{\gamma<T} \frac{1}{|\rho\zeta'(\rho)|} \sim \Sigma_1 \sim \alpha\int_{\gamma_1^-}^{T} \frac{(\log t)^{\frac{1}{4}}}{t}\ dt = \alpha\left[\frac{4}{5}(\log t)^{\frac{5}{4}}\right]_{\gamma_1^-}^{T} \sim \frac{4\alpha}{5}(\log T)^{\frac{5}{4}} \ .$$

Similarly, in part (d) we can show that

$$\sum_{\gamma<T} \frac{1}{|\rho\zeta'(\rho)|^2} \sim \sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|^2\gamma^2} \ .$$

By partial summation,

$$\begin{aligned}
\sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|^2\gamma^2} &= \left[\frac{J_{-1}(t)}{t^2}\right]_{T}^{\infty} + 2\int_{T}^{\infty} \frac{J_{-1}(t)}{t^3}\ dt \\
&= 0 - \frac{3}{\pi^3 T} - \frac{\epsilon_2(T)}{T^2} + 2\int_{T}^{\infty} \frac{\frac{3t}{\pi^3} + \epsilon_2(t)}{t^3}\ dt \ .
\end{aligned} \tag{7.6}$$

Since $\epsilon_2(t) = o(t)$ we obtain

$$\sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|^2\gamma^2} \sim \frac{6}{\pi^3}\int_{T}^{\infty} t^{-2}\ dt - \frac{3}{\pi^3 T} = \frac{3}{\pi^3 T} \ .$$

(e) Let $\phi(t) = (\log t)^a t^{-b}$. Its derivative is $\phi'(t) = (a(\log t)^{a-1} - b(\log t)^a)/t^{b+1}$. Partial summation implies

$$\sum_{\gamma>T} \frac{(\log\gamma)^a}{\gamma^b|\zeta'(\rho)|} = \Big[\phi(t)J_{-\frac{1}{2}}(t)\Big]_T^\infty - \int_T^\infty J_{-\frac{1}{2}}(t)\phi'(t)\ dt\ .$$

The first term is $\ll \phi(T)J_{-\frac{1}{2}}(T) = (\log T)^{a+v}/T^{b-u}$. Assuming the bound on $J_{-\frac{1}{2}}(T)$, the second term is bounded by

$$\ll \int_T^\infty \frac{(t^u(\log t)^v)(\log t)^a}{t^{b+1}}\ dt = \int_T^\infty \frac{(\log t)^{a+v}}{t^{b-u+1}}\ dt\ .$$

However, if $B > 1$ the integral $I_{A,B} = \int_T^\infty \frac{(\log t)^A}{t^B}\ dt$ can be computed by partial integration. Choosing $u = t^{-B}$ and $dv = (\log t)^A$ yields

$$I_{A,B} \sim \frac{(\log T)^A}{(B-1)T^{B-1}}\ .$$

If desired, lower order terms can be computed. This shows that the second term is also bounded by $(\log T)^{a+v}/T^{b-u}$ and this completes the proof. □

Comment 1 Curiously, Keating et al. [40] have suggested the constant α in part (c) of the Lemma is

$$\alpha = \frac{1}{\sqrt{\pi}}e^{3\zeta'(-1)-\frac{11}{12}\log 2}\prod_{p\ \text{prime}}\left(\left(1-\frac{1}{p}\right)^{\frac{1}{4}}\sum_{m=0}^\infty\left(\frac{\Gamma(m-\frac{1}{2})}{m!\Gamma(-\frac{1}{2})}\right)^2 p^{-m}\right).$$

This is a most amazing constant!

Comment 2 Set $f(t) = \sum_{\gamma<t}\frac{1}{|\rho\zeta'(\rho)|}$. By the above lemma, $f(t) \sim \frac{4\alpha}{5}(\log t)^{\frac{5}{4}}$. If $f(t)$ has an asymptotic formula with smaller order terms, it seems plausible that $f'(t) \sim \alpha\frac{(\log t)^{\frac{1}{4}}}{t}$. Observe that

$$\sum_{t\le\gamma<t+1}\frac{1}{|\rho\zeta'(\rho)|} = f(t+1)-f(t) = f'(\theta) \approx \frac{(\log\theta)^{\frac{1}{4}}}{\theta} \asymp \frac{(\log t)^{\frac{1}{4}}}{t}$$

for some $\theta \in (t, t+1)$. If we assumed an upper bound of this form, many of the estimates in the next section could be dramatically improved.

7.3 Conditional results concerning $M(x)$

Let $\mathcal{T} = \{T_\nu \mid \nu \ge 0\ ,\ \nu \le T_\nu \le \nu+1\}$ denote a positive increasing sequence of real numbers such that if $T \in \mathcal{T}$ then

$$\frac{1}{\zeta(\sigma+iT)} = O(T^\epsilon)$$

for all $\frac{1}{2} \leq \sigma \leq 2$. The fact that such a sequence of numbers can be chosen is explained on pp. 357-358 of Titchmarsh [76].

Lemma 7.3.1 *Assume the Riemann Hypothesis. For $x \geq 2$ and $T \in \mathcal{T}$*

$$M(x) = \sum_{\gamma \leq T} \frac{x^\rho}{\rho \zeta'(\rho)} + E(x,T)$$

where

$$E(x,T) \ll \frac{x \log x}{T} + \frac{x}{T^{1-\epsilon} \log x} + 1.$$

Proof The starting point is to recall that $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$. Hence, applying a variant of a theorem from Titchmarsh's book [76] pp. 60-63 (See also Ivić [35] pp. 300-303)

$$M(x) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{x^s}{\zeta(s)s} ds + O\left(\frac{x^c}{T(c-1)}\right) + O\left(\frac{x \log x}{T}\right) + O(1).$$

Setting $c = 1 + (\log x)^{-1}$, this becomes

$$M(x) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{x^s}{\zeta(s)s} ds + O\left(\frac{x \log x}{T}\right) + O(1).$$

Introduce a large parameter U. Consider the positively oriented rectangle $B_{T,U}$. Its vertices are at $c - iT, c + iT, -U + iT$, and $-U - iT$. Thus we write

$$\begin{aligned} M(x) = \frac{1}{2\pi i} \int_{B_{T,U}} \frac{x^s}{\zeta(s)s} ds - \frac{1}{2\pi i} \left(\int_{c+iT}^{-U+iT} + \int_{-U+iT}^{-U-iT} + \int_{-U-iT}^{c-iT} \right) \frac{x^s}{\zeta(s)s} ds \\ + O\left(\frac{x \log x}{T}\right) + O(1). \end{aligned} \tag{7.7}$$

Denote the three integrals as $I_1(U), I_2(U), I_3(U)$. It's possible to show that $I_2 \to 0$ as $U \to \infty$. We have by the functional equation

$$\begin{aligned} I_2 = \int_{-U-iT}^{-U+iT} \frac{x^s}{s\zeta(s)} \, ds &= \int_{U+1-iT}^{U+1+iT} \frac{x^{1-s}}{(1-s)\zeta(1-s)} \, ds \\ &= \int_{U+1-iT}^{U+1+iT} \frac{x^{1-s}}{(1-s)} \frac{2^{s-1}\pi^s}{\cos(\pi s/2)\Gamma(s)\zeta(s)} \, ds. \end{aligned} \tag{7.8}$$

Set $V = U + 1$. Note that $\frac{1}{|\zeta(V+it)\cos(\frac{\pi}{2}(V+it))|} \ll 1$ and $\frac{1}{|1-(V+it)|} \ll \frac{1}{V}$ for $-T \leq t \leq T$. By Stirling's formula we can estimate the reciprocal of the gamma function as

$$\frac{1}{\Gamma(s)} = O(|e^{s-(s-\frac{1}{2})\log s}|) = O(e^{\sigma-(\sigma-\frac{1}{2})\log|s|+\frac{1}{2}\pi t}) = O(e^{\sigma-(\sigma-\frac{1}{2})\log \sigma+\frac{1}{2}\pi t}).$$

Combining estimates we obtain

$$I_2 \ll \int_{-T}^{T} \frac{1}{V}\left(\frac{2\pi}{x}\right)^{V-1} e^{V-(V-\frac{1}{2})\log V}\, dt = \frac{2T}{V}\left(\frac{2\pi}{x}\right)^{V-1} e^{V-(V-\frac{1}{2})\log V}.$$

This shows that $I_2 \to 0$ as $U \to \infty$. In the box $B_{T,U}$, $\frac{x^s}{s\zeta(s)}$ has poles at the zeros of the zeta function and $s = 0$. Thus we now have by Cauchy's Residue Theorem

$$M(x) = \frac{1}{2\pi i}\sum_{|\gamma|<T} \frac{x^\rho}{\rho\zeta'(\rho)} - 2 + \sum_{k\geq 1} \frac{x^{-2k}}{-2k\zeta'(-2k)} - \frac{1}{2\pi i}\left(\int_{c+iT}^{-\infty+iT} + \int_{-\infty-iT}^{c-iT}\right)\frac{x^s}{\zeta(s)s}ds + O\left(\frac{x\log x}{T}\right) + O(1). \tag{7.9}$$

In addition, the assumption that $T \in \mathcal{T}$ allows us to show that $I_1+I_3 = O\left(\frac{1}{T} + \frac{x}{T^{1-\epsilon}\log x}\right)$. Let's consider I_1. Write $I_1 = \left(\int_{c+iT}^{-1+iT} + \int_{-1+iT}^{\infty+iT}\right)\frac{x^s}{s\zeta(s)}\, ds$. Since $T \in \mathcal{T}$ we have

$$\left|\int_{-1+iT}^{c+iT} \frac{x^s}{s\zeta(s)}\, ds\right| \leq \int_{-1}^{c} \frac{x^\sigma T^\epsilon}{\sqrt{\sigma^2+T^2}}\, d\sigma \leq T^{\epsilon-1}\int_{-1}^{c} e^{\sigma\log x}\, d\sigma = \frac{T^{\epsilon-1}}{\log x}(x^c - x^{-1}) \ll \frac{xT^{\epsilon-1}}{\log x}.$$

For the other integral we use the functional equation

$$\int_{-\infty+iT}^{-1+iT} \frac{x^s}{s\zeta(s)}\, ds = \int_{2+iT}^{\infty+iT} \frac{x^{1-s}}{(1-s)\zeta(1-s)}\, ds = \int_{2+iT}^{\infty+iT} \frac{x^{1-s}}{(1-s)}\frac{2^{s-1}\pi^s}{\cos(\frac{\pi s}{2})\Gamma(s)\zeta(s)}$$
$$= O\left(\int_2^\infty \left(\frac{2\pi}{x}\right)^\sigma \frac{1}{\sqrt{(1-\sigma)^2+T^2}|\cos(\frac{\pi(\sigma+iT)}{2})\Gamma(\sigma+iT)\zeta(\sigma+iT)|}\, d\sigma\right). \tag{7.10}$$

By definition of cosine we see that $\frac{1}{|\cos(\frac{\pi(\sigma+iT)}{2})|} \ll e^{-\frac{\pi}{2}T}$ and Stirling's formula shows that $\frac{1}{|\Gamma(\sigma+iT)|} \ll e^{\sigma-(\sigma-\frac{1}{2})\log\sigma+\frac{1}{2}\pi T}$. Furthermore, $\frac{1}{\zeta(\sigma+iT)} \ll 1$ for $\sigma \geq 2$. Hence the integral is

$$O\left(\int_2^\infty \frac{1}{T}\left(\frac{2\pi}{x}\right)^\sigma e^{\sigma-(\sigma-\frac{1}{2})\log\sigma}\, d\sigma\right) = O\left(\frac{1}{T}\right).$$

Therefore, I_1 satisfies the required bound. The same argument works for I_3. Since we have the bound

$$\sum_{k\geq 1}\frac{x^{-2k}}{-2k\zeta'(-2k)} \ll \frac{1}{x^2}$$

for $x \geq 2$ we have now shown that

$$M(x) = \sum_{\gamma\leq T}\frac{x^\rho}{\rho\zeta'(\rho)} + O(1) + O\left(\frac{1}{T} + \frac{x}{T^{1-\epsilon}\log x}\right) + O\left(\frac{x\log x}{T}\right).$$

□

Corollary 7.3.2 *As before, let $\mathcal{T} = \{T_\nu\}_{\nu=0}^{\infty}$, then RH implies*

$$M(x) = \lim_{\nu\to\infty} \sum_{|\gamma|<T_\nu} \frac{x^\rho}{\rho\zeta'(\rho)} - 2 + \sum_{k\geq 1} \frac{x^{-2k}}{-2k\zeta'(-2k)}$$

Proof This is the same argument as above.

Corollary 7.3.3 *RH implies*

$$\sum_{\gamma>0} \frac{1}{|\rho\zeta'(\rho)|} \text{ diverges.}$$

Proof By the previous corollary, we see that if the above sum is absolutely convergent, then the sum over zeros is uniformly convergent and hence continous. This implies $M(t)$ is continous. However, $M(t)$ has discontinuities at the squarefree numbers. □

Corollary 7.3.4 *Assume RH and the Gonek-Hejhal type bound*

$$\sum_{T\leq\gamma\leq 2T} \frac{1}{|\rho\zeta'(\rho)|^2} \ll \frac{1}{T}.$$

(which is equivalent to $J_{-1}(T) \ll T$). For $x \geq 2$,$T \geq 2$

$$M(x) = \sum_{\gamma\leq T} \frac{x^\rho}{\rho\zeta'(\rho)} + E(x,T)$$

where

$$E(x,T) \ll \frac{x\log x}{T} + \frac{x}{T^{1-\epsilon}\log x} + \left(\frac{x\log T}{T}\right)^{\frac{1}{2}} + 1.$$

Proof Let $T \geq 2$ lie between arbitrary non-negative integers $\nu, \nu+1$. There exists T_ν in this interval such that

$$M(x) = \sum_{\gamma\leq T_\nu} \frac{x^\rho}{\rho\zeta'(\rho)} + O\left(\frac{x\log x}{T_\nu} + \frac{x}{T_\nu^{1-\epsilon}\log x} + 1\right).$$

Now suppose without loss of generality that $\nu \leq T_\nu \leq T \leq \nu+1$. Then we have,

$$M(x) = \sum_{\gamma\leq T} \frac{x^\rho}{\rho\zeta'(\rho)} - \sum_{T_\nu\leq\gamma\leq T} \frac{x^\rho}{\rho\zeta'(\rho)} + O\left(\frac{x\log x}{T} + \frac{x}{T^{1-\epsilon}\log x} + \frac{1}{x^2}\right).$$

However, by Cauchy-Schwarz we can bound the second sum

$$\begin{aligned}\left|\sum_{T_\nu \leq \gamma \leq T} \frac{x^\rho}{\rho\zeta'(\rho)}\right| &\leq x^{\frac{1}{2}} \sum_{T_\nu \leq \gamma \leq T} \frac{1}{|\rho\zeta'(\rho)|} \\ &\leq x^{\frac{1}{2}} \left(\sum_{T_\nu \leq \gamma \leq T} \frac{1}{|\rho\zeta'(\rho)|^2}\right)^{\frac{1}{2}} \left(\sum_{T_\nu \leq \gamma \leq T} 1\right)^{\frac{1}{2}} = \left(\frac{x \log T}{T}\right)^{\frac{1}{2}}.\end{aligned} \tag{7.11}$$

□

We will now apply the explicit formula to obtain results concerning $M(x)$. In fact, it can be shown that RH and the Gonek-Hejhal conjecture implies the weak Mertens conjecture. The proof is similar to Cramér's analysis of the integral $\int_2^X \left(\frac{\psi(x)-x}{x}\right)^2 dx$.

Lemma 7.3.5 *Let*

$$S(t) = \sum_\nu c(\nu) e^{2\pi i \nu t}$$

be an absolutely convergent exponential sum. Here the frequencies ν run over an arbitrary sequence of real numbers and the coefficients are complex. Let $\delta = \frac{\theta}{T}$, with $0 < \theta < 1$. Then

$$\int_{-T}^{T} |S(t)|^2 dt \ll_\theta \int_{-\infty}^{\infty} \left|\delta^{-1} \sum_{x \leq \nu \leq x+\delta} c(\nu)\right|^2 dx.$$

Proof See Gallagher [25] pp. 330-331.

Lemma 7.3.6 *(i) Assume the following localized version of the Gonek-Hejhal conjecture*

$$\sum_{T \leq \gamma \leq T+1} \frac{1}{|\rho\zeta'(\rho)|} \ll \frac{(\log T)^{\frac{1}{4}}}{T}.$$

Then we have the bound

$$\int_X^{eX} \left|\sum_{T \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)}\right|^2 \frac{dx}{x^2} \ll \frac{(\log T)^{\frac{1}{2}}}{T}$$

for $T \leq X$.
(ii) Assume the Gonek-Hejhal cojecture, $J_{-1}(T) \ll T$. Then

$$\int_X^{eX} \left|\sum_{T \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)}\right|^2 \frac{dx}{x^2} \ll \frac{(\log T)}{T^{\frac{1}{4}}}$$

for $T \leq X$.
If we assumed $J_{-\frac{1}{2}}(T) \ll T(\log T)^{\frac{1}{4}}$, the above bound could be replaced by $\frac{(\log T)^{\frac{3}{4}}}{T^{\frac{1}{4}}}$.

Proof Making the substitution $x = e^y$ in the above integral and squaring we obtain

$$\begin{aligned} \int_{\log X}^{\log X+1} \left| \sum_{T \leq |\gamma| \leq X} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \right|^2 dy &\leq 4 \int_{\log X}^{\log X+1} \left| \sum_{T \leq \gamma \leq X} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \right|^2 dy \\ &= 4 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \sum_{T \leq \gamma \leq X} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \right|^2 dy. \end{aligned} \tag{7.12}$$

The final equality comes from replacing y by $y - \log X - \frac{1}{2}$. Applying Gallagher's lemma with $T = \frac{1}{2}$ and $\theta = \frac{1}{2}$ shows that the preceding integral is bounded by

$$\int_{T-1}^{X} \left(\sum_{t \leq \gamma \leq t+1} \frac{1}{|\rho\zeta'(\rho)|} \right)^2 dt \ll \int_{T-1}^{X} \frac{(\log t)^{\frac{1}{2}}}{t^2} dt \ll \frac{(\log T)^{\frac{1}{2}}}{T} \tag{7.13}$$

which establishes the claim. We will now prove the weaker bound

$$\int_{X}^{eX} \left| \sum_{T \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} \right|^2 \frac{dx}{x^2} \ll \frac{(\log T)}{T^{\frac{1}{4}}}$$

for $T \leq X$. Making the substitution $x = e^y$ in the above integral and squaring we obtain

$$\begin{aligned} \int_{\log X}^{\log X+1} \left| \sum_{T \leq |\gamma| \leq X} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \right|^2 dy &\leq 4 \int_{\log X}^{\log X+1} \left| \sum_{T \leq \gamma \leq X} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \right|^2 dy \\ &= 4 \sum_{T \leq \gamma \leq X} \sum_{T \leq \lambda \leq X} \frac{1}{\rho\zeta'(\rho)\overline{\rho'\zeta'(\rho')}} \int_{\log X}^{\log X+1} e^{i(\gamma-\lambda)y}\, dy. \end{aligned} \tag{7.14}$$

Note that ρ and ρ' denote zeros of the form $\rho = \frac{1}{2} + i\gamma$ and $\rho' = \frac{1}{2} + i\lambda$. In addition, observe that

$$\int_{\log X}^{\log X+1} e^{i(\gamma-\lambda)y}\, dy \ll \min\left(1, \frac{1}{|\gamma - \lambda|}\right).$$

Now let $\eta = 1$. Divide the above sum into two pieces Σ_1 and Σ_2. Let Σ_1 consist of those terms for which $|\gamma - \lambda| \leq \eta$ and Σ_2 consist of those terms for which $|\gamma - \lambda| > \eta$.

Observe that the summands in Σ_1 satisfy $\min(1, \frac{1}{|\gamma-\lambda|}) = 1$ and the summands in Σ_2 satisfy $\min(1, \frac{1}{|\gamma-\lambda|}) = \frac{1}{|\gamma-\lambda|}$. We will now bound each of these sums.

$$\Sigma_1 \ll \sum_{T \leq \gamma \leq X} \frac{1}{|\rho\zeta'(\rho)|} \sum_{\gamma-\eta \leq \lambda \leq \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|}.$$

Consider the inner sum

$$\sum_{\gamma-\eta \leq \lambda \leq \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|} \leq \left(\sum_{\gamma-\eta \leq \lambda \leq \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|^2} \right)^{\frac{1}{2}} (N(\gamma+\eta) - N(\gamma-\eta))^{\frac{1}{2}} \ll \left(\frac{\log \gamma}{\gamma} \right)^{\frac{1}{2}}.$$

Thus, we find that

$$\Sigma_1 \ll \sum_{T \leq \gamma} \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{3}{2}} |\zeta'(\rho)|} \ll \frac{\log T}{T^{\frac{1}{2}}} \tag{7.15}$$

by Lemma 7.2.5 (e). Now consider the second sum. We have

$$\begin{aligned} \Sigma_2 &= \sum_{T \leq \gamma \leq X} \sum_{T \leq \lambda \leq X,\, |\gamma-\lambda| \geq \eta} \frac{1}{|\rho\zeta'(\rho)||\rho'\zeta'(\rho')||\gamma-\lambda|} \\ &= \sum_{T \leq \gamma \leq X} \frac{1}{|\rho\zeta'(\rho)|} \sum_{T \leq \lambda \leq X,\, |\gamma-\lambda| \geq \eta} \frac{1}{|\rho'\zeta'(\rho')||\gamma-\lambda|}. \end{aligned} \tag{7.16}$$

We will analyze the inner sum and apply the same technique originally used by Cramér [8]. Denote this inner sum as $S(\gamma)$ where $T \leq \gamma \leq X$. Consider the set of numbers, $\gamma^c, \gamma - \gamma^c$, and $\gamma - \eta$. Either $T \leq \gamma^c$, $\gamma^c \leq T \leq \gamma - \gamma^c$, or $\gamma - \gamma^c \leq T \leq \gamma - \eta$. Suppose the first case is true. i.e. $T \leq \gamma^c$. Then we can write the inner sum $S(\gamma)$ as six seperate sums

$$\begin{aligned} S(\gamma) = &\sum_{T \leq \lambda < \gamma^c} + \sum_{\gamma^c \leq \lambda < \gamma-\gamma^c} + \sum_{\gamma-\gamma^c \leq \lambda \leq \gamma-\eta} \\ &+ \sum_{\gamma+\eta \leq \lambda < \gamma+\gamma^c} + \sum_{\gamma+\gamma^c \leq \lambda < 2\gamma} + \sum_{2\gamma \leq \lambda < X} \frac{1}{|\rho'\zeta'(\rho')||\gamma-\lambda|} \end{aligned} \tag{7.17}$$

where $0 < c < 1$ is some constant to be determined later. Denote each of these sums as σ_i for $i = 1 \ldots 6$. Let's now bound each of these appropriately. We have

$$\begin{aligned} \sigma_1 \leq \frac{1}{\gamma - \gamma^c} \sum_{T \leq \lambda < \gamma^c} \frac{1}{|\rho'\zeta'(\rho')|} &\ll \frac{1}{\gamma} \left(\sum_{T \leq \lambda < \gamma^c} \frac{1}{|\rho'\zeta'(\rho)|^2} \right)^{\frac{1}{2}} \left(\sum_{T \leq \lambda < \gamma^c} 1 \right)^{\frac{1}{2}} \\ &\ll \frac{1}{\gamma} \cdot 1 \cdot (\gamma^c \log \gamma)^{\frac{1}{2}} = \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{1-\frac{c}{2}}}, \end{aligned} \tag{7.18}$$

$$\begin{aligned}\sigma_2 &\leq \frac{1}{\gamma^c}\sum_{\gamma^c\leq\lambda<\gamma-\gamma^c}\frac{1}{|\rho'\zeta'(\rho')|} \leq \frac{1}{\gamma^c}\left(\sum_{\gamma^c\leq\lambda<\gamma-\gamma^c}\frac{1}{|\rho'\zeta'(\rho')|^2}\right)^{\frac{1}{2}}\left(\sum_{\gamma^c\leq\lambda<\gamma-\gamma^c}1\right)^{\frac{1}{2}} \\ &\ll \frac{1}{\gamma^c}\left(\frac{1}{\gamma^c}\right)^{\frac{1}{2}}(\gamma\log\gamma)^{\frac{1}{2}} = \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^{\frac{3c}{2}-\frac{1}{2}}},\end{aligned} \tag{7.19}$$

and

$$\sigma_3 \leq \left(\sum_{\gamma-\gamma^c\leq\lambda\leq\gamma-\eta}\frac{1}{|\rho'\zeta'(\rho')|^2}\right)^{\frac{1}{2}}\left(\sum_{\gamma-\gamma^c\leq\lambda\leq\gamma-\eta}1\right)^{\frac{1}{2}} \ll \frac{1}{\gamma^{\frac{1}{2}}}(\gamma^c\log\gamma)^{\frac{1}{2}} = \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{2}-\frac{c}{2}}}.$$

The fourth sum, σ_4, gives the same error term as the third sum. Lastly,

$$\begin{aligned}\sigma_5 &\ll \frac{1}{\gamma^c}\sum_{\gamma+\gamma^c\leq\lambda<2\gamma}\frac{1}{|\rho'\zeta'(\rho')|} \ll \frac{1}{\gamma^c}\left(\sum_{\gamma+\gamma^c\leq\lambda}\frac{1}{|\rho'\zeta'(\rho)|^2}\right)^{\frac{1}{2}}\left(\sum_{\gamma+\gamma^c\leq\lambda\leq 2\gamma}1\right)^{\frac{1}{2}} \\ &\ll \frac{1}{\gamma^c}\left(\frac{\gamma\log\gamma}{\gamma}\right)^{\frac{1}{2}} = \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^c}\end{aligned} \tag{7.20}$$

and

$$\begin{aligned}\sigma_6 &\leq \sum_{k=1}^{\infty}\sum_{2^k\gamma\leq\lambda\leq 2^{k+1}\gamma}\frac{1}{|\rho'\zeta'(\rho')||\gamma-\lambda|} \\ &\leq \sum_{k=1}^{\infty}\frac{1}{2^{k-1}\gamma}\left(\sum_{2^k\gamma\leq\lambda\leq 2^{k+1}\gamma}\frac{1}{|\rho'\zeta'(\rho')|^2}\right)^{\frac{1}{2}}\left(\sum_{2^k\gamma\leq\lambda\leq 2^{k+1}\gamma}1\right)^{\frac{1}{2}} \\ &\leq \sum_{k=1}^{\infty}\frac{1}{2^{k-1}\gamma}\left(\frac{2^{k+1}\gamma\log(2^{k+1}\gamma)}{2^k\gamma}\right)^{\frac{1}{2}} \\ &\ll \sum_{k=1}^{\infty}\frac{\sqrt{k}+(\log\gamma)^{\frac{1}{2}}}{2^{k-1}\gamma} \\ &\ll \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma}.\end{aligned} \tag{7.21}$$

Putting together these bounds leads to

$$S(\gamma) \ll \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^{1-\frac{c}{2}}} + \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^{\frac{3c-1}{2}}} + \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^{\frac{1-c}{2}}} + \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma^c} + \frac{(\log\gamma)^{\frac{1}{2}}}{\gamma}.$$

The best choice is $c = \frac{1}{2}$ and this gives us

$$S(\gamma) \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{4}}}.$$

The assumption $J_{-1}(T) \ll T$ implies by Cauchy-Schwarz that $J_{-1/2}(T) \ll T(\log T)^{\frac{1}{2}}$. Applying Lemma 7.2.5 (e) yields the bound

$$\Sigma_2 \ll \sum_{\gamma > T} \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{5}{4}} |\zeta'(\rho)|} \ll \frac{\log T}{T^{\frac{1}{4}}}$$

and the lemma is proved.

Comment The only reason part (i) of the lemma is included is to deduce the true size of the sum in question. We observe that we are off by a factor of $T^{\frac{3}{4}}$ in the denominator. It is not apparent whether part (ii) gives the optimal bound under the assumption $J_{-1}(T) \ll T$. This needs to be investigated further.

Theorem 7.3.7
RH and $\sum_{\gamma \leq T} \frac{1}{|\rho\zeta'(\rho)|} \ll (\log T)^{\frac{5}{4}}$ *implies*
(i)

$$M(x) \ll x^{\frac{1}{2}} (\log \log x)^{\frac{3}{2}}$$

except on a set of finite logarithmic measure.
(ii) RH and $J_{-1}(T) \ll T$ *implies*

$$\int_2^X \frac{M(x)^2}{x}\, dx \ll X.$$

(iii) RH and $J_{-1}(T) \ll T$ *implies*

$$\int_2^X \left(\frac{M(x)}{x} \right)^2 dx \ll \log X.$$

Proof (i) The starting point is to consider the explicit formula. We have

$$M(x) = - \sum_{|\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} + O\left(X^\epsilon\right)$$

valid for $X \leq x \ll X$. From the preceding lemma we have for $T^4 \leq X$

$$\int_X^{eX} \left| \sum_{T^4 \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} \right|^2 \frac{dx}{x^2} \ll \frac{(\log T)}{T}. \tag{7.22}$$

Now consider the set

$$S = \{x \geq 2 \mid |\sum_{T^4 \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)}| \geq x^{\frac{1}{2}} (\log\log x)^{\frac{3}{2}}\}.$$

Then it follows that

$$(\log\log X)^3 \int_{S \cap [X, eX]} \frac{dx}{x} \leq \int_X^{eX} \left| \sum_{T^4 \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} \right|^2 \frac{dx}{x^2} \ll \frac{(\log T)}{T}$$

and thus

$$\int_{S \cap [X, eX]} \frac{dx}{x} \ll \frac{(\log T)}{T(\log\log X)^3} = \frac{1}{T(\log T)^2}$$

for $T = \log X$. Choosing $X = e^T$ with $T = 2, 3, \ldots$ it follows that S has finite logarithmic measure. Assuming the Gonek-Hejhal type bound we notice that

$$\sum_{0 \leq |\gamma| \leq T^4} \frac{x^\rho}{\rho\zeta'(\rho)} \ll X^{\frac{1}{2}} \sum_{0 \leq |\gamma| \leq T^4} \frac{1}{|\rho\zeta'(\rho)|} \ll X^{\frac{1}{2}} (\log T)^{\frac{5}{4}} \ll X^{\frac{1}{2}} (\log\log X)^{\frac{5}{4}}$$

for $X \leq x \leq eX$. Hence,

$$M(x) = -\sum_{T^4 \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} + O\left(X^{\frac{1}{2}} (\log\log X)^{\frac{5}{4}}\right)$$

for $X \leq x \leq eX$ and $T = \log X$. Define the set

$$S' = \{x \geq 2 \mid |M(x)| \geq cx^{\frac{1}{2}} (\log\log x)^{\frac{3}{2}}\}.$$

where c is any constant greater than one. Suppose $x \in S' \bigcap [X, eX]$. Then

$$\begin{aligned} \left| \sum_{T^4 \leq |\gamma| \leq X} \frac{x^\rho}{\rho\zeta'(\rho)} \right| &\geq |M(x)| - O\left(X^{\frac{1}{2}} (\log\log X)^{\frac{5}{4}}\right) \\ &\geq cx^{\frac{1}{2}} (\log\log x)^{\frac{3}{2}} - O\left(X^{\frac{1}{2}} (\log\log X)^{\frac{5}{4}}\right) \\ &\geq x^{\frac{1}{2}} (\log\log x)^{\frac{3}{2}} \end{aligned} \tag{7.23}$$

for $x \in [X, eX]$ as long as X is sufficiently large. Thus $S' \bigcap [X, eX] \subset S \bigcap [X, eX]$ for X sufficiently large and it follows that S' has finite logarithmic measure.

(ii) Setting $T = \frac{3}{2}$ in the above integral (7.22) and applying the formula for $M(x)$ yields

$$\int_X^{eX} \left(\frac{M(x)}{x}\right)^2 dx \ll 1$$

which in turn implies
$$\int_X^{eX} \frac{M(x)^2}{x}\, dx \ll X.$$
Substituting the values $\frac{X}{e}, \frac{X}{e^2}, \ldots$ and adding yields the required bound.

(iii) As noted above,
$$\int_X^{eX} \left(\frac{M(x)}{x}\right)^2 dx \ll 1.$$
Substituting the values $\frac{X}{e}, \frac{X}{e^2}, \ldots, \frac{X}{e^j}$ and summing shows
$$\int_{\frac{X}{e^j}}^{X} \left(\frac{M(x)}{x}\right)^2 dx \ll j.$$
Choose j to be the least integer such that $\frac{X}{e^j} < 2$. This condition implies
$$\log(X/2) < j \leq \log(X/2) + 1$$
and thus we obtain
$$\int_2^{X} \left(\frac{M(x)}{x}\right)^2 dx \ll \log X.$$
□

7.4 Existence of a limiting distribution

The goal of this section is to prove a limiting distribution for the function $\phi(y) = e^{-\frac{y}{2}}M(e^y)$. If we assume RH and write non-trivial zeros as $\rho = \frac{1}{2}+i\gamma$, then we obtain
$$x^{-\frac{1}{2}}M(x) = \sum_{\gamma\leq T} \frac{x^{i\gamma}}{\rho\zeta(\rho)} + x^{-\frac{1}{2}}E(x,T).$$
Making the variable change $x = e^y$, we have
$$\phi(y) = \sum_{|\gamma|\leq T} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} + \epsilon^{(T)}(y)$$
where
$$\phi^{(T)}(y) = \sum_{|\gamma|\leq T} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} \text{ and } \epsilon^{(T)}(y) = \sum_{T\leq|\gamma|\leq e^Y} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)} + e^{-\frac{y}{2}}E(e^y, e^Y)\ .$$
The next lemma shows that the error term $\epsilon^{(T)}(y)$ has small L_2 norm.

Lemma 7.4.1 *Assume RH and $J_{-1}(T) \ll T$. For $T \geq 1$ and $Y \geq \log 2$,*

$$\int_{\log 2}^{Y} |\epsilon^{(T)}(y)|^2 dy \ll Y \frac{(\log T)}{T^{\frac{1}{4}}}.$$

Proof First we will consider the error term. Note that

$$\begin{aligned} |e^{-\frac{y}{2}} E(e^y, e^Y)|^2 &\ll e^{-y} \left(\frac{y^2 e^{2y}}{e^{2Y}} + \frac{\frac{1}{y^2} e^{2y}}{(e^{2Y})^{1-\epsilon}} + \frac{e^y Y}{e^Y} + 1 \right) \\ &= \frac{y^2 e^y}{e^{2Y}} + \frac{\frac{1}{y^2} e^y}{(e^{2Y})^{1-\epsilon}} + \frac{Y}{e^Y} + \frac{1}{e^y}. \end{aligned} \tag{7.24}$$

Therefore,

$$\int_{\log 2}^{Y} |e^{-\frac{y}{2}} E(e^y, e^Y)|^2 dy = O(1)$$

due to the last term. We have

$$\begin{aligned} \int_{\log 2}^{Y} |\epsilon^{(T)}(y)|^2 dy &\ll \int_{\log 2}^{Y} \left| \sum_{T \leq |\gamma| \leq e^Y} \frac{e^{i\gamma y}}{\rho \zeta'(\rho)} \right|^2 dy + O(1) \\ &\leq 4 \int_{\log 2}^{Y} \left| \sum_{T \leq \gamma \leq e^Y} \frac{e^{i\gamma y}}{\rho \zeta'(\rho)} \right|^2 dy + O(1) \\ &= 4 \sum_{T \leq \gamma \leq e^Y} \sum_{T \leq \lambda \leq e^Y} \int_{\log 2}^{Y} \frac{e^{i(\gamma - \lambda) y}}{\rho \overline{\rho'} \zeta'(\rho) \overline{\zeta'(\rho')}} dy + O(1). \end{aligned} \tag{7.25}$$

Note that ρ and ρ' denote zeros of the form $\rho = \frac{1}{2} + i\gamma$ and $\rho' = \frac{1}{2} + i\lambda$. In addition, observe that

$$\int_{\log 2}^{Y} e^{i(\gamma - \lambda) y} dy = \begin{cases} Y - \log 2 \text{ if } \gamma = \lambda \\ \frac{(e^{i(\gamma-\lambda)Y} - 2^{i(\gamma-\lambda)})}{i(\gamma - \lambda)} \text{ otherwise} \end{cases}.$$

It now follows that

$$\int_{\log 2}^{Y} |\epsilon^{(T)}(y)|^2 dy \ll \sum_{T \leq \gamma \leq e^Y} \sum_{T \leq \lambda \leq e^Y} \frac{\min(Y, \frac{1}{|\gamma - \lambda|})}{|\rho \zeta'(\rho)||\rho' \zeta'(\rho')|} + O(1).$$

Now let $\eta = \frac{1}{Y}$. Divide the sum into two pieces Σ_1 and Σ_2. Let Σ_1 consist of those terms for which $|\gamma - \lambda| \leq \eta$ and Σ_2 consist of those terms for which $|\gamma - \lambda| > \eta$.

Observe that the summands in Σ_1 satisfy $\min(Y, \frac{1}{|\gamma-\lambda|}) = Y$ and the summands in Σ_2 satisfy $\min(Y, \frac{1}{|\gamma-\lambda|}) = \frac{1}{|\gamma-\lambda|}$. In particular, we obtain

$$\Sigma_1 = Y \sum_{T \le \gamma \le e^Y} \frac{1}{|\rho\zeta'(\rho)|} \sum_{\gamma-\eta \le \lambda \le \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|} \, .$$

Consider the inner sum

$$\sum_{\gamma-\eta \le \lambda \le \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|} \le \left(\sum_{\gamma-\eta \le \lambda \le \gamma+\eta} \frac{1}{|\rho'\zeta'(\rho')|^2} \right)^{\frac{1}{2}} (N(\gamma+\eta)-N(\gamma-\eta))^{\frac{1}{2}} \ll \left(\frac{\log \gamma}{\gamma} \right)^{\frac{1}{2}} \, .$$

Thus, we find that

$$\Sigma_1 \ll Y \sum_{T \le \gamma} \frac{\log(|\rho|)^{\frac{1}{2}}}{|\rho|^{\frac{3}{2}} |\zeta'(\rho)|} \ll Y \frac{(\log T)}{T^{\frac{1}{2}}}$$

by Lemma 7.2.5(e). Now consider the second sum. We have

$$\begin{aligned} \Sigma_2 &= \sum_{T \le \gamma \le e^Y} \sum_{T \le \lambda \le e^Y, \, |\gamma-\lambda| \ge \eta} \frac{1}{|\rho\zeta'(\rho)||\rho'\zeta'(\rho')||\gamma-\lambda|} \\ &= \sum_{T \le \gamma \le e^Y} \frac{1}{|\rho\zeta'(\rho)|} \sum_{T \le \lambda \le e^Y, \, |\gamma-\lambda| \ge \eta} \frac{1}{|\rho'\zeta'(\rho')||\gamma-\lambda|} . \end{aligned} \tag{7.26}$$

We will analyze the inner sum and apply the same technique originally used by Cramér [8]. Denote this inner sum as $S(\gamma)$ where $T \le \gamma \le e^Y$. Consider the set of numbers, $\gamma^{\frac{1}{2}}, \gamma - \gamma^{\frac{1}{2}}$, and $\gamma - \eta$. Either $T \le \gamma^{\frac{1}{2}}$, $\gamma^{\frac{1}{2}} \le T \le \gamma - \gamma^{\frac{1}{2}}$, or $\gamma - \gamma^{\frac{1}{2}} \le T \le \gamma - \eta$. Suppose the first case is true. i.e. $T \le \gamma^{\frac{1}{2}}$. Then we can write the inner sum $S(\gamma)$ as six seperate sums

$$\begin{aligned} S(\gamma) = &\sum_{T \le \lambda < \gamma^{\frac{1}{2}}} + \sum_{\gamma^{\frac{1}{2}} \le \lambda < \gamma - \gamma^{\frac{1}{2}}} + \sum_{\gamma - \gamma^{\frac{1}{2}} \le \lambda \le \gamma - \eta} \\ &+ \sum_{\gamma+\eta \le \lambda < \gamma + \gamma^{\frac{1}{2}}} + \sum_{\gamma + \gamma^{\frac{1}{2}} \le \lambda < 2\gamma} + \sum_{2\gamma \le \lambda < e^Y} \frac{1}{|\rho'\zeta'(\rho')||\gamma-\lambda|} . \end{aligned} \tag{7.27}$$

Denote each of these sums as σ_i for $i = 1, \ldots, 6$. In a calculation virtually identical to Lemma 7.4.6 one obtains

$$\begin{aligned} &\sigma_1 \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{3}{4}}}, \ \sigma_2 \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{4}}}, \ \sigma_3, \sigma_4 \ll Y \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{4}}} \\ &\sigma_5 \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{2}}}, \ \sigma_6 \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma} \, . \end{aligned} \tag{7.28}$$

Putting together these bounds leads to

$$S(\gamma) \ll \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{4}}} + Y \frac{(\log \gamma)^{\frac{1}{2}}}{\gamma^{\frac{1}{4}}}.$$

Recall that we assumed that $T \leq \gamma^{\frac{1}{2}}$. In the other two cases we get the same bound. The only difference is that there would be fewer sums. We immediately obtain

$$\begin{aligned} \Sigma_2 &\ll \sum_{T \leq \gamma} \frac{(\log |\rho|)^{\frac{1}{2}}}{|\rho|^{1+\frac{1}{4}}|\zeta'(\rho)|} + Y \sum_{T \leq \gamma} \frac{(\log |\rho|)^{\frac{1}{2}}}{|\rho|^{1+\frac{1}{4}}|\zeta'(\rho)|} \\ &\ll \frac{(\log T)}{T^{\frac{1}{4}}} + Y \frac{(\log T)}{T^{\frac{1}{4}}} \end{aligned} \tag{7.29}$$

by applying Lemma 7.2.5 (e) again. The result of the lemma now follows. □

Lemma 7.4.2 *For each T there is a probabililty measure ν_T on $\mathbb{R}$ such that*

$$\nu_T(f) := \int_{-\infty}^{\infty} f(x)\ d\nu_T(x) = \lim_{Y \to \infty} \frac{1}{Y} \int_{\log 2}^{Y} f(\phi^{(T)}(y))\ dy.$$

for all bounded Lipschitz continuous functions f on $\mathbb{R}$.

Proof This is identical to Rubinstein-Sarnak [62] p. 180 and Lemma 5.1.3. This will only be sketched. Let $N = N(T)$ denote the number of zeros of $\zeta(s)$ to height T. Label the imaginary ordinates of the zeros as $\gamma_1, \ldots \gamma_N$. Set $b_l = -\frac{1}{(\frac{1}{2}+i\gamma_l)\zeta'(\frac{1}{2}+i\gamma_l)}$. We have

$$\phi^{(T)}(y) = 2\mathrm{Re}\left(\sum_{l=1}^{N} b_l e^{iy\gamma_l}\right).$$

Define the function $g(\theta_1, \ldots \theta_N)$ on the N-torus $\mathbb{T}^N$ by

$$g(\theta_1, \ldots \theta_N) = f\left(2\mathrm{Re}\left(\sum_{l=1}^{N} b_l e^{2\pi i \theta_l}\right)\right).$$

Let A be the topological closure in $\mathbb{T}^N$ of the one-parameter subgroup

$$\Gamma(y) := \{(\frac{\gamma_1 y}{2\pi}, \ldots, \frac{\gamma_N y}{2\pi}) \mid y \in \mathbb{R}\ \}.$$

As noted earlier,

$$\lim_{Y \to \infty} \int_{\log 2}^{Y} f(\phi^{(T)}(y))dy = \int_A g(a)da$$

where a is Haar measure on A and the proof is complete. □

Theorem 7.4.3 *Assume RH and $J_{-1}(T) \ll T$. Then $\phi(y) = e^{-\frac{y}{2}} M(e^y) = e^{-\frac{y}{2}} \sum_{n \le e^y} \mu(n)$ has a limiting distribution ν on $\mathbb{R}$, that is,*

$$\lim_{Y \to \infty} \frac{1}{Y} \int_0^Y f(e^{-\frac{y}{2}} M(e^y))\ dy = \int_{-\infty}^{\infty} f(x)\ d\nu(x)$$

for all Lipschitz bounded continuous functions f on $\mathbb{R}$.

Proof Once again the proof is identical to Theorem 5.1.2. Let f be Lipschitz bounded continuous that satisfies $|f(x) - f(y)| \le c_f |x - y|$. Note that

$$\begin{aligned}
\frac{1}{Y} \int_{\log 2}^{Y} f(\phi(y)) dy &= \frac{1}{Y} \int_{\log 2}^{Y} f(\phi^{(T)}(y)) dy + O\left(\frac{c_f}{\sqrt{Y}} \left(\int_{\log 2}^{Y} |\epsilon^{(T)}(y)|^2\ dy \right)^{\frac{1}{2}} \right) \\
&= \frac{1}{Y} \int_{\log 2}^{Y} f(\phi^{(T)}(y)) dy + O\left(\frac{c_f}{\sqrt{Y}} \cdot \sqrt{Y} \frac{(\log T)^{\frac{1}{2}}}{T^{\frac{1}{8}}} \right) \qquad (7.30) \\
&= \frac{1}{Y} \int_{\log 2}^{Y} f(\phi^{(T)}(y)) dy + O\left(\frac{c_f (\log T)^{\frac{1}{2}}}{T^{\frac{1}{8}}} \right) .
\end{aligned}$$

Applying the preceding lemma, we deduce that the limit

$$\lim_{Y \to \infty} \frac{1}{Y} \int_{\log 2}^{Y} f(\phi(y)) dy$$

exists. For more details see Theorem 5.1.2 where the argument is the same.

Comment Suppose the above theorem remains valid for indicator functions. Let V be a fixed real number. We will take $f = f_V$ where

$$f_V(x) = \begin{cases} 1 \text{ if } x \ge V \\ 0 \text{ if } x < V \end{cases} .$$

The above formula becomes

$$\lim_{Y \to \infty} \frac{1}{Y} \text{meas}\{ y \in [\log 2, Y] \mid M(e^y) \ge e^{\frac{y}{2}} V \ \} = \nu([V, \infty)).$$

As noted in Rubinstein-Sarnak, the above identity would be true if $\nu(x)$ is absolutely continuous. Under the assumption of LI this is the case.

Theorem 7.4.4 *Assume RH, $J_{-1}(T) \ll T$, and LI. Then the Fourier transform $\widehat{\nu}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t}\ d\nu(t)$ exists and equals*

$$\widehat{\nu}(\xi) = \prod_{\gamma > 0} J_0 \left(\frac{2\xi}{\sqrt{\frac{1}{4} + \gamma^2} \left| \zeta'(\frac{1}{2} + i\gamma) \right|} \right).$$

Proof This is same as the proof of Theorem 5.2.1 and Rubinstein-Sarnak [62] p. 184.

Comment In proving the existence of the limiting distribution associated to $e^{-\frac{y}{2}}M(e^y)$ and the weak Mertens conjecture, we assumed RH and $J_{-1}(T) \ll T$. As explained earlier, there is considerable evidence to believe the latter conjecture. Consequently, in our initial attempts to prove the existence of the limiting distribution we just assumed the known conjectures about $J_{-1}(T)$. However, by examining the proofs more closely, it now appears that these two results would follow from two weaker conjectures. If we assumed

$$\begin{aligned} &J_{-1}(T) \ll T^{\theta} \ , \\ &\sum_{\gamma>0} \frac{1}{|\rho\zeta'(\rho)|^{\lambda}} \text{ converges for } \lambda \geq \lambda(\theta) \end{aligned} \tag{7.31}$$

where $1 < \theta < 2$ and $\lambda(\theta) > 1$, we could prove the results of this chapter. It seems that there are values of θ larger than one for which the proofs would remain valid. How large θ can be taken is not presently obvious. This needs to be investigated further. Note that the simplicity of the zeros of the zeta function implies that $J_{-1}(T) \ll T^{2+\epsilon}$. Hence, the closer we can take θ to two, the more likely we can reduce the hypotheses to RH and the simplicity of the zeros of the zeta function.

7.5 A heuristic lower bound for $M(x)$

The goal of this section is to study the true order of $M(x)$. We will attempt to find the size of the tail of the probability measure ν associated to $\phi(y) = e^{-\frac{y}{2}}M(e^y)$. The only tool we have in studying tails of ν are some probability type results due to Montgomery [48]. We will need to assume LI. Consider a random variable X, defined on the infinite torus $\mathbb{T}^{\infty}$ by

$$X(\theta) = \sum_{k=1}^{\infty} r_k \sin(2\pi\theta_k)$$

where $r_k \in \mathbb{R}$ for $k \geq 1$. We will also assume $\sum_{k\geq 1} r_k^2 < \infty$. By Komolgorov's theorem, we have X converges almost everywhere. (If we had $\sum_{k\geq 1} r_k^2 = \infty$, X would diverge almost everywhere). This is a map $X : \mathbb{T}^{\infty} \to \mathbb{R}$. In addition, $\mathbb{T}^{\infty}$ has a canonical probability measure P. Attached to the random variable X is the distribution function μ_X defined by

$$\mu_X(x) = P(X^{-1}(-\infty, x)).$$

For these random variables, Montgomery [48] pp. 14-16 proved the following result.

Theorem 7.5.1 *Let $X(\underline{\theta}) = \sum_{k=1}^{\infty} r_k \sin 2\pi\theta_k$ where $\sum_{k=1}^{\infty} r_k^2 < \infty$. For any integer $K \geq 1$*
(a)

$$P\left(X(\underline{\theta}) \geq 2\sum_{k=1}^{K} r_k\right) \leq \exp\left(-\frac{3}{4}\left(\sum_{k=1}^{K} r_k\right)^2 \left(\sum_{k>K} r_k^2\right)^{-1}\right)$$

(b) If $r_k \searrow 0$ then

$$P\left(X(\underline{\theta}) \geq \frac{1}{2}\sum_{k=1}^{K} r_k\right) \geq 2^{-40} \exp\left(-100\left(\sum_{k=1}^{K} r_k\right)^2 \left(\sum_{k>K} r_k^2\right)^{-1}\right)$$

(c) If δ is so small that $\sum_{r_k>\delta}(r_k - \delta) \geq V$, then

$$P(X(\underline{\theta}) \geq V) \geq \frac{1}{2}\exp\left(-\frac{1}{2}\sum_{r_k>\delta} \log\left(\frac{\pi^2 r_k}{2\delta}\right)\right)$$

Montgomery proved the previous theorem in order to study the tails of the limiting distribution connected to

$$\frac{\log e^y}{e^{y/2}}(\pi(e^y) - \mathrm{Li}(e^y))\ .$$

Letting μ denote this distribution, Montgomery [48] showed that

$$\exp(-c_2\sqrt{V}\exp\sqrt{2\pi V}) \leq \mu(B_V) \leq \exp(-c_1\sqrt{V}\exp\sqrt{2\pi V})$$

where $B_V = [V, \infty)$ or $(-\infty, -V]$ for V large. From these bounds, he conjectured

$$\underline{\overline{\lim}}\frac{\log x(\pi(x) - \mathrm{Li}(x))}{\sqrt{x}(\log\log\log x)^2} = \pm\frac{1}{2\pi}.$$

We would like to apply the same type of reasoning to the distribution ν attached to $e^{-y/2}M(e^y)$. Observe that LI conjecture implies that the limiting distribution $\nu = \nu_X$. Here we have denoted ν_X to be the distribution function attached to the random variable

$$X(\theta) = \sum_{\gamma>0}^{\infty} r_\gamma \sin(2\pi\theta_\gamma)$$

and $r_\gamma = \frac{2}{|\rho\zeta'(\rho)|}$. Therefore, by assuming LI, we can study ν via the random variable X. Using the Lemma 7.2.5 and Montgomery's Theorem 7.5.1 on sums of independent random variables we can estimate the tails of the limiting distribution ν. Set $a(T) = \sum_{\gamma<T} r_\gamma$ and $b(T) = \sum_{\gamma\geq T} r_\gamma^2$. By lemma 7.2.5, the conjectured formulas for $a(T)$ and $b(T)$ are

$$a(T) \sim A(\log T)^{\frac{5}{4}} \text{ and } b(T) \sim \frac{12}{\pi^3 T}.$$

where $A = \frac{4\alpha}{5}$. We will now bound the tail of the limiting distribution ν. Let V be a large parameter. Choose T such that $a(T^-) < V \le a(T)$. Note that T is the ordinate of a zero. Let ϵ be a small parameter. We have the chain of inequalities

$$(1-\epsilon)A(\log T)^{\frac{5}{4}} \le (1-\frac{\epsilon}{2})a(T) < V \le a(T) \le (1+\epsilon)A(\log T)^{\frac{5}{4}}.$$

Solving for T this shows that this implies

$$e^{\left(\frac{1}{A(1+\epsilon)}V\right)^{\frac{4}{5}}} \le T < e^{\left(\frac{1}{A(1-\epsilon)}V\right)^{\frac{4}{5}}} \ , \ \left(\frac{1}{A(1+\epsilon)}V\right)^{\frac{4}{5}} \le \log T < \left(\frac{1}{A(1-\epsilon)}V\right)^{\frac{4}{5}}.$$

Likewise, let δ be a small parameter. We also have the inequalities

$$(1-\delta)\frac{12}{\pi^3 T} \le b(T) \le (1+\delta)\frac{12}{\pi^3 T}.$$

An upper bound for the tail

$$\begin{aligned} &P\left(X(\underline{\theta}) \ge \frac{2}{1-\frac{\epsilon}{2}}V\right) \le P\left(X(\underline{\theta}) \ge 2a(T)\right) \le \exp\left(-\frac{3}{4}a(T)^2 b(T)^{-1}\right) \\ &\le \exp\left(-\frac{3}{4}V^2\frac{\pi^3}{12(1+\delta)}T\right) \le \exp\left(-\frac{3\pi^3}{48(1+\delta)}V^2 e^{\left(\frac{1}{A(1+\epsilon)}V\right)^{\frac{4}{5}}}\right). \end{aligned} \tag{7.32}$$

This implies that

$$P(X(\underline{\theta}) \ge V) \le \exp\left(-c_1 V^2 e^{c_2 V^{\frac{4}{5}}}\right)$$

for some effective constants c_1 and c_2. Note that the above does not require asymptotics for $a(T)$ and $b(T)$. We only require the estimates $a(T) \asymp (\log T)^{\frac{5}{4}}$ and $b(T) \asymp \frac{1}{T}$.

A lower bound for the tail

This is a more delicate analysis than the upper bound of the tail. In the following analysis, we will denote a set of effectively computable constants as $A_1, A_2, \ldots, A_i$. At certain points, some of these constants may be replaced by a constant multiple of themselves. We will apply Theorem 7.5.1 (c). V is considered fixed and large. We would like to choose δ small enough such that

$$\sum_{r_\gamma > \delta} (r_\gamma - \delta) \ge V \ . \tag{7.33}$$

Introduce the notation S_δ and N_δ such that

$$S_\delta = \{\gamma \mid r_\gamma > \delta \ \} \text{ and } N_\delta = \# S_\delta \ .$$

Note that RH implies $|\zeta'(\rho)| \le A_1 |\rho|^\epsilon$ where $A_1 = A_1(\epsilon)$. Thus,

$$\delta < \frac{2}{A_1|\rho|^{1+\epsilon}} \Longrightarrow \delta < \frac{2}{|\rho\zeta'(\rho)|} \ .$$

However,

$$\delta < \frac{2}{A_1|\rho|^{1+\epsilon}} \Longleftrightarrow |\rho| \leq \left(\frac{2}{A_1\delta}\right)^{\frac{1}{1+\epsilon}} .$$

Notice that $|\rho| \ll \gamma$. Hence,

$$\text{if } \gamma \leq A_2\left(\frac{1}{\delta}\right)^{\frac{1}{1+\epsilon}} \Longrightarrow \delta < \frac{2}{|\rho\zeta'(\rho)|}.$$

We deduce from Riemann's zero counting formula that there are at least

$$N_1 = A_2\left(\frac{1}{\delta}\right)^{\frac{1}{1+\epsilon}} \log\left(\frac{1}{\delta}\right) + O(\left(\frac{1}{\delta}\right)^{\frac{1}{1+\epsilon}})$$

zeros in the set S_δ. We will now find an upper bound for N. Gonek [27] has defined the number

$$\Theta = \text{l.u.b.}\{\theta \mid \left|\frac{1}{\zeta'(\rho)}\right| \ll |\gamma|^{\theta},\ \forall\rho\ \}.$$

We will assume $\Theta < 1$, however $J_{-1}(T) \ll T$ implies $\Theta \leq \frac{1}{2}$. Gonek has suggested that $\Theta = \frac{1}{3}$. Choose $\epsilon < 1 - \Theta$ and we will set $\kappa = 1 - \Theta - \epsilon > 0$. This implies that if $\gamma \in S_\delta$ then

$$\delta < \frac{2}{|\rho\zeta'(\rho)|} \ll \frac{|\rho|^{\Theta+\epsilon}}{|\rho|} = \frac{1}{|\rho|^{\kappa}} \leq \frac{1}{|\gamma|^{\kappa}} .$$

We deduce that if $\gamma \in S_\delta$ then $\gamma \leq \frac{A_3}{\delta^{\frac{1}{\kappa}}}$. We conclude that

$$A_2\left(\frac{1}{\delta}\right)^{\frac{1}{1+\epsilon}} \log\left(\frac{1}{\delta}\right) + O(\left(\frac{1}{\delta}\right)^{\frac{1}{1+\epsilon}}) \leq N(\delta) \leq A_3\left(\frac{1}{\delta}\right)^{\frac{1}{\kappa}} \log\left(\frac{1}{\delta}\right) + O(\left(\frac{1}{\delta}\right)^{\frac{1}{\kappa}})$$

where the upper and lower bounds have been defined as N_1 and N_2. We are trying to determine a condition on δ so that we can satisfy the stated inequality (7.33). Note that

$$\sum_{r_\gamma > \delta}(r_\gamma - \delta) \geq \sum_{\gamma \leq N_1}(r_\gamma - \delta) .$$

Before evaluating the second sum, note that

$$\delta N_1 = A_2\delta^{1-\frac{1}{1+\epsilon}} \log\left(\frac{1}{\delta}\right) + O(\delta^{1-\frac{1}{1+\epsilon}}) \to 0 \text{ as } \delta \to 0.$$

We will choose δ as a function of V and as $V \to \infty$ we have $\delta \to 0$. However,

$$\begin{aligned}
\sum_{\gamma \leq N_1}(r_\gamma - \delta) &= 2\sum_{\gamma \leq N_1}\frac{1}{|\rho\zeta'(\rho)|} - \delta\sum_{\gamma \leq N_1} 1 \\
&= A(\log N_1)^{\frac{5}{4}} + o\left((\log N_1)^{\frac{5}{4}}\right) - \frac{\delta N_1}{2\pi}\log N_1 + O(\delta N_1) \\
&\geq A_4(\log N_1)^{\frac{5}{4}}
\end{aligned} \tag{7.34}$$

where $0 < A_4 < A$. The last inequality holds for N_1 sufficiently large. Thus we see that

$$A_4(\log N_1)^{\frac{5}{4}} \geq V \Longleftrightarrow N_1 \geq \exp(\left(\frac{V}{A_4}\right)^{\frac{4}{5}}).$$

Therefore, choosing $N_1 = \exp(\left(\frac{V}{A_4}\right)^{\frac{4}{5}})$ implies that $\sum_{r_\gamma > \delta}(r_\gamma - \delta) \geq V$. We can apply the theorem to find a lower bound. Recall that

$$P(X(\underline{\theta}) \geq V) \geq \frac{1}{2}\exp\left(-\frac{1}{2}\sum_{r_\gamma > \delta}\log\left(\frac{\pi^2 r_\gamma}{2\delta}\right)\right).$$

An upper bound of the sum will provide a lower bound for the tail. Note that $\frac{1}{|\rho\zeta'(\rho)|} \to 0$ under the assumption that all zeros are simple. Assume $\frac{1}{|\rho\zeta'(\rho)|} \leq A_5$.

$$\begin{aligned}\sum_{r_\gamma > \delta}\log\left(\frac{\pi^2 r_\gamma}{2\delta}\right) &\leq \sum_{\gamma \leq N_2}\log\left(\frac{\pi^2 r_\gamma}{2\delta}\right) \leq \sum_{\gamma \leq N_2}\log\left(\frac{\pi^2 A_5}{\delta}\right) \\ &= \log\left(\frac{\pi^2 A_5}{\delta}\right)\sum_{\gamma \leq N_2} 1 \ll \log\left(\frac{\pi^2 A_5}{\delta}\right) N_2 \log N_2.\end{aligned} \tag{7.35}$$

By definition of N_1 and N_2 there exists a real number $t > 1$ such that $N_2 \ll N_1^t$ for V sufficiently large. Thus,

$$N_2 \log N_2 \ll N_1^t \log N_1 \ll \exp(A_6 V^{\frac{4}{5}}) V^{\frac{4}{5}}$$

and we obtain

$$\sum_{r_\gamma > \delta}\log\left(\frac{\pi^2 r_\gamma}{2\delta}\right) \ll V^{\frac{4}{5}}\exp(A_6 V^{\frac{4}{5}}) V^{\frac{4}{5}} = V^{\frac{8}{5}}\exp(A_6 V^{\frac{4}{5}}) .$$

We finally arrive at the lower bound

$$P(X(\underline{\theta}) \geq V) \geq \frac{1}{2}\exp\left(-c_3 V^{\frac{8}{5}}\exp(c_4 V^{\frac{4}{5}})\right) .$$

for certain effective constants c_3 and c_4.

7.5.1 A lower bound for $M(x)$

In the following section, we will use the iterated logarithm notation. Namely, $\log_1 x = \log x$ and for $k \geq 2$

$$\log_k x = \log(\log_{k-1} x).$$

(For example, $\log_2 x = \log\log x$ and $\log_3 x = \log\log\log x$.)

Let's now consider the consequences of a lower bound of the form

$$\exp\left(-c_3 V^N \exp(c_4 V^{\frac{4}{5}})\right) \leq P\left(X(\underline{\theta}) \geq V\right).$$

where $c_3, c_4, N > 0$ are fixed constants. Recall that assuming LI, RH, and $J_{-1}(T) \ll T$ we have,

$$\lim_{Y\to\infty} \frac{1}{Y}\text{meas}\{y \in [0, Y] \mid M(e^y) \geq e^{y/2}V\ \} = P\left(X(\underline{\theta}) \geq V\right).$$

Assume that the lower bound is sufficiently uniform in Y. That is, for Y and V large,

$$\exp\left(-c_3 V^N \exp(c_4 V^{\frac{4}{5}})\right) \leq \frac{1}{Y}\text{meas}\{y \in [0, Y] \mid M(e^y) \geq e^{\frac{y}{2}}V\ \}.$$

Choose

$$c_4 V^{\frac{4}{5}} = \theta(\log_2 Y) \iff V = \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y)^{\frac{5}{4}}$$

where $0 < \theta < 1$. Substituting these values shows that

$$\begin{aligned} &\exp\left(-c_3 \left(\frac{\theta}{c_4}\right)^{\frac{5N}{4}} (\log_2 Y)^{\frac{5N}{4}} (\log Y)^{\theta}\right) \leq \\ &\frac{1}{Y}\text{meas}\{y \in [0, Y] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y)^{\frac{5}{4}}\ \}. \end{aligned} \tag{7.36}$$

However, notice that

$$c_3 \left(\frac{\theta}{c_4}\right)^{\frac{5N}{4}} (\log_2 Y)^{\frac{5N}{4}} (\log Y)^{\theta} \leq \log Y + \log_3 Y$$

is true for Y sufficiently large. Thus, we have

$$\frac{1}{Y \log_2 Y} \leq \frac{1}{Y}\text{meas}\{y \in [0, Y] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y)^{\frac{5}{4}}\ \}$$

is true for $Y \geq Y_0$. Assume Y_0 is an integer and we get

$$\infty = \sum_{Y=Y_0}^{\infty} \frac{1}{Y \log_2 Y} \leq \sum_{Y=Y_0}^{\infty} \frac{1}{Y}\text{meas}\{y \in [0, Y] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y)^{\frac{5}{4}}\ \}.$$

The divergence of the above sum implies there exists an infinite sequence of real numbers Y_m, $m = 1, 2, \ldots$ such that $Y_m \to \infty$ and

$$\text{meas}\{y \in [0, Y_m] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}}\ \} > 0.$$

This, in turn, implies that there exists an increasing sequence of real numbers y_m such that $y_m \to \infty$ and

$$\frac{M(e^{y_m})}{e^{\frac{y_m}{2}}} \geq \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 y_m)^{\frac{5}{4}}.$$

Suppose by way of contradiction, that the above inequality is false. That is, there exists a real number u_0 such that

$$\frac{M(e^{y})}{e^{\frac{y}{2}}} < \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 y)^{\frac{5}{4}}$$

for all $y \geq u_0$. Assume that the Y_m are chosen such that $Y_m > u_0$. Then we have that

$$\begin{aligned} &\text{meas}\{y \in [0, Y_m] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}} \ \} \\ &= \text{meas}\{y \in [0, u_0] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}} \ \} \end{aligned} \tag{7.37}$$

since if $u_0 \leq y \leq Y_m$ then

$$\frac{M(e^{y})}{e^{\frac{y}{2}}} \leq \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 y)^{\frac{5}{4}} \leq \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}}.$$

Let

$$C_0 = \max_{0 \leq y \leq u_0} \frac{M(e^{y})}{e^{\frac{y}{2}}}.$$

Since $Y_m \to \infty$ we can choose m_0 such that

$$\left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}} > C_0$$

$m \geq m_0$. However, if there exists a real number y satisfying $0 \leq y \leq u_0$ with

$$\frac{M(e^{y})}{e^{\frac{y}{2}}} \geq \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}}$$

then we obtain

$$C_0 \geq \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}}.$$

This cannot happen for $m \geq m_0$. We now have shown that

$$\text{meas}\{y \in [0, Y_m] \mid M(e^y) \geq e^{\frac{y}{2}} \left(\frac{\theta}{c_4}\right)^{\frac{5}{4}} (\log_2 Y_m)^{\frac{5}{4}} \ \} = 0$$

for all $m \geq m_0$ and we have a contradiction. Hence, our original assumption is false and we obtain

$$M(e^y) = \Omega_+ \left(e^{\frac{y}{2}} (\log_2 y)^{\frac{5}{4}} \right).$$

which is

$$M(x) = \Omega_+ \left(x^{\frac{1}{2}} (\log\log\log x)^{\frac{5}{4}} \right).$$

A similar argument would show the other inequality. This explains the conjecture

$$M(x) = \Omega_\pm \left(x^{\frac{1}{2}} (\log\log\log x)^{\frac{5}{4}} \right).$$

We now have seen that the true size of $M(x)$ depends on the sizes of $J_{-\frac{1}{2}}(T)$ and $J_{-1}(T)$.

Chapter 8
Conclusion

8.1 Chebyshev's bias

It was shown in the Galois group setting that biases can occur either due to the behaviour of squares of primes (Chebyshev's bias) or because of a zero at the center of the critical strip of any of the Artin L-functions. This is interesting since it gives meaning to the zero at the center of the strip. Of course, this does not hold as deep a meaning as the vanishing or non-vanishing of an elliptic curve L-function as described by the Birch and Swinnerton-Dyer conjecture. This should be viewed more as a curiosity or an amusement. What is particularly interesting is that classically one could have studied the reduction mod p of the eighth degree polynomials with Galois group H_8. For example, Fröbenius could have looked at such a polynomial and not been able to give an adequate explanation of why σ_p has a bias towards C_1 in certain examples. The reason is that the naive and natural guess would be that there is always a bias towards C_2 in those cases. The knowledge of the Artin L-functions is essential in describing the aformentioned bias.

8.2 The summatory function of the Möbius function

The key point of the results from the final chapter was to investigate the true order of the summatory function of the Möbius function. To some extent this was achieved, but at the cost of assuming certain deep conjectures regarding the zeros of the Riemann zeta function. It should be noted that in studying the order of $\pi(x) - \mathrm{Li}(x)$ it is also necessary to at least assume RH in understanding the true order of this function. For the lower bound, there is Littlewood's amazing result which only assumes RH. However, when RH is not assumed an even better bound is obtained. As for $M(x)$, no new information can be deduced without any knowledge of the sizes of $|\zeta'(\rho)|^{-1}$. In fact, it has been conjectured by Gonek that $|\zeta'(\rho)|^{-1} \ll |\rho|^{\frac{1}{3}+\epsilon}$. It should be noted that even the assumption of this bound is rather crude and would not be sufficient in trying to prove the limiting distribution associated to $x^{-\frac{1}{2}}M(x)$. As was shown, we needed to employ conjectured upper bounds of average values of the form $J_{-k}(T) = \sum_{\gamma<T} \frac{1}{|\zeta'(\rho)|^{2k}}$ in order to show the existence of the limiting distribution. The other key result which can be proven very similarly is that Gonek's conjecture and the Riemann Hypothesis imply the weak Merten's conjecture. This appears not to have been known previously. The technique in proving the weak Merten's conjecture leads to the result

$$M(x) \ll x^{\frac{1}{2}}(\log\log x)^{\frac{3}{2}}$$

except on a set of finite logarithmic measure. On the other hand, the study of the tails of the limiting distribution leads to the conjecture

$$M(x) = \Omega_{\pm}(x^{\frac{1}{2}}(\log\log\log x)^{\frac{5}{4}}).$$

8.3 Future investigations and open problems

In the course of writing this thesis the following problems and questions arose.

1. Develop a function field analogue of Chebyshev's bias. Here is a typical example. Let $K = \mathbb{F}_q(T)$ be a function field. Let $f(T)$ be an irreducible polynomial over $\mathbb{F}_q$. Consider the quadratic extension $L = \mathbb{F}_q(T, \sqrt{f})$. Note that $\mathcal{O}_K = \mathbb{F}_q[T]$ and $\mathcal{O}_L = \mathbb{F}_q[T, \sqrt{f}]$. Clearly, $\mathrm{Gal}(L/K)$ is the cyclic group of order two. We could form the prime counting functions

$$\pi_1(x) = \sum_{\deg(p)\leq T,\, \sigma_p=1,\, (p,f)=1} 1 \text{ and } \pi_2(x) = \sum_{\deg(p)\leq T,\, \sigma_p\neq 1,\, (p,f)=1} 1.$$

The obvious question is whether there are any biases in this situation. Furthermore, is there a limiting distribution associated to the function $\pi_1(x) - \pi_2(x)$? There are explicit formulas known in this case and have been derived by Murty and Scherk [54]. What is intriguing about this situation is that the associated L-function satisfies the Riemann Hypothesis with roots lying on a circle. Perhaps in this case unconditional proofs may be available. Furthermore, the fact that the zeta function has finitely many zeros suggest that perhaps the Fourier transform of the conjectured limiting distribution may be a finite product. This is rather appealing considering the great difficulties encountered in approximating the limiting distribution in the classical case. It should be pointed out that all of these comments are rather premature and perhaps there are no interesting phenomenon in this case.

2. Although the Rubinstein-Sarnak article gave new insight into prime number races, they discovered the surprising fact that prime number races for $r > 3$ tend not to be symmetric. More precisely, $\delta(P_{q;a_1,a_2,\ldots,a_r}) \neq \frac{1}{r!}$. Recently, Feuerverger and Martin [20] developed a formula and computed some examples of r-way races for $r \geq 3$. They

also give an explanation of why some of these densities are either equal or unequal. However, this phenomenon is still not completely understood and is something of a mystery. An explanation of the behaviour of the r-way races remains an open problem.

3. Generalize the Chowla/folklore conjecture of the non-vanishing of Dirichlet L-functions to Artin L-functions. For example, is it true that an Artin L-function with root number not equal to minus one does not vanish at the center of the critical strip? If this is not true then find an example of an Artin L-function with root number equal to one and has a zero at $s = \frac{1}{2}$. In Murty and Murty [50] p. 37 they question what is the order of vanishing of an Artin L-function at $s = \frac{1}{2}$. If we set L/K a normal extension and χ a character of the Galois group is it true that

$$\mathrm{ord}_{s=\frac{1}{2}} L(s, \chi) \ll \chi(1) \ ?$$

In fact, it is quite likely an averaged version of this conjecture can be proven assuming the Riemann Hypothesis and Artin's Conjecture. The technique would require Weil's explicit formula. Another related problem would be to prove non-vanishing results for certain families of Galois groups and their corresponding Artin L-functions. This would be analogous to Iwaniec and Sarnak's result on the non-vanishing of at least one-third of the Dirichlet L-functions mod q at $s = \frac{1}{2}$. For example, assume L_q is a family of normal field extensions of $\mathbb{Q}$ parametrized by the prime numbers. Try to show some fraction of the corresponding Artin L-functions at $s = \frac{1}{2}$ do not have a zero there.

4. Prove non-vanishing results for families of L-functions attached to weight one modular forms. In the literature, there seems to be an absence of results in this case. The Iwaniec-Sarnak [37] paper only deals with weight greater or equal to two. Is it true that an L-function attached to weight one modular form does not vanish at $s = \frac{1}{2}$. One difficulty in studying weight one modular forms is that it is not known what the dimension of $S_1(q, \epsilon)$ is. Duke [17] has some upper bounds for the dimension of this space. However, his bounds are still far from the conjectured true size of the size. One may ask whether the L-functions attached to the newforms of dihedral, tetrahedral, octahedral, or icosahedral type vanish at $\frac{1}{2}$. Currently, some information about the dihedral case is known. Duke, Friedlander, and Iwaniec [18] considered imaginary quadratic fields $K = \mathbb{Q}(\sqrt{D})$ with class group $\mathcal{H}_K$. They proved the average value formulas

$$\begin{aligned} \frac{w}{h} \sum_{\chi \in \widehat{\mathcal{H}}_K} L_K(\frac{1}{2}, \chi) &= \frac{1}{2} \log D + c + O\left(e^{-\pi\sqrt{D}}\right) \\ \frac{1}{h} \sum_{\chi \in \widehat{\mathcal{H}}_K} \left| L_K(\frac{1}{2}, \chi) \right|^2 &= l_D\left(\frac{1}{2}\right) + O\left(L(1, \chi_D)\right) \end{aligned} \tag{8.1}$$

where $l_D(\frac{1}{2}) \sim cL(1, \chi_D) \log D$ under the Riemann hypothesis. However, Littlewood has shown that under the Riemann hypothesis $L(1, \chi_D) \ll \log\log D$. Applying the

Cauchy-Schwarz inequality one obtains

$$\#\{\chi \in \hat{\mathcal{H}}_K \mid L_K(\frac{1}{2}, \chi) \neq 0 \} \gg \frac{h}{\log D \log\log D}$$

assuming RH. Without assuming RH, the lower bound would be replaced by $\frac{h}{(\log D)^2}$. It would be interesting to find out if the introduction of mollifiers can make the above result into a positive density result.

5. Prove Gonek's conjecture. This seems to be a daunting task. A proof of Gonek's conjecure would establish the equivalence of the Riemann Hypothesis and the weak Merten's conjecture. Gonek proved the lower bound $J_{-1}(T) \gg T$ under the assumption of the Riemann Hypothesis. In addition, if all zeros of the zeta function are assumed to be simple, then we have the trivial bound $J_{-1}(T) \ll T^{2+\epsilon}$. Unfortunately, it is not yet known whether all zeros are simple. However, it is widely expected to be true. Conrey [9] has shown that more than $\frac{2}{5}$ of the zeros are simple and lie on the line $\sigma = \frac{1}{2}$. Assuming RH and the Generalized Lindelöf Hypothesis at least $\frac{19}{27}$ of the zeros of $\zeta(s)$ are simple. Moreover, if Montgomery's pair correlation conjecture is true then density one of the zeros would be simple. Nevertheless, one may try to obtain better upper bounds than the trivial one still assuming RH and the simplicity of the zeros of the Riemann zeta function.

6. Using the same techniques as the final chapter one can show that the function

$$\phi(y) = e^{-\frac{y}{2}} \sum_{n \le e^y} \lambda(n)$$

has a limiting distribution under suitable hypotheses. Note that this is the function that appears in Pólya's conjecture. The way to study this function is to make use of the Dirichlet series identity

$$\frac{\zeta(2s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s}.$$

In order to study the true order of this function we would require some knowledge of the corresponding discrete moments

$$K_{-1}(T) = \sum_{\gamma \le T} \left| \frac{\zeta(2\rho)}{\zeta'(\rho)} \right|^2 \text{ and } K_{-\frac{1}{2}}(T) = \sum_{\gamma \le T} \left| \frac{\zeta(2\rho)}{\zeta'(\rho)} \right|.$$

Note that the only difference between these moments and the $J_{-k}(T)$ that Gonek studied are the presence of the numbers $|\zeta(2\rho)|$. Assuming RH, the numbers 2ρ lie on the line $\mathrm{Re}(s) = 1$ and thus cannot be too large. Under the RH, there is Littlewood's estimate (see [76] pp. 344-347)

$$\frac{1}{\log\log t} \ll |\zeta(1+it)| \ll \log\log t.$$

Consequently, the contribution of the numbers $|\zeta(2\rho)|$ cannot contribute significantly to the sum. The technique's from Gonek's MSRI talk [27] seem to indicate that

$$K_{-1}(T) = \sum_{\gamma \leq T} \left| \frac{\zeta(2\rho)}{\zeta'(\rho)} \right|^2 \sim \frac{T}{2\pi}.$$

On the other hand, not much is known about the sum $K_{-\frac{1}{2}}(T)$. It would be interesting to see if any of Gonek's ideas [26], [27] or the random matrix models of Hughes, Keating, and O'Connell [33] lead to any reasonable conjectures regarding $K_{-\frac{1}{2}}(T)$. Is is true that $K_{-\frac{1}{2}}(T) \approx J_{-\frac{1}{2}}(T)$? The true order of $K_{-\frac{1}{2}}(T)$ would reveal the true order of $\sum_{n \leq x} \lambda(n)$ using Montgomery's [48] results on random variables.

Bibliography

[1] J.A. Antoniadis, *Diedergruppe und Reziprozitätsgesetz*, J. Reine Angew. Math. **377** (1987), 197-209.

[2] J.V. Armitage, *Zeta Functions with a zero at* $s = \frac{1}{2}$, Inventiones Math. **15** (1972), 199-205.

[3] P.T. Bateman et al., *Linear relations connecting the imaginary parts of the zeros of the zeta function*, in *Computers in Number Theory*, eds. A.O.L. Atkin and B.J. Birch, Academic Press, New York, 1971, 11-19.

[4] Carter Bays, Kevin B. Ford, Richard H. Hudson, and Michael Rubinstein, *Zeros of Dirichlet L-Functions near the real axis and Chebyshev's Bias*, preprint, 1999.

[5] J. Buhler, *Icosahedral Galois Representations*, Lecture Notes in Mathematics **654** (1978), Springer Verlag.

[6] S. Chowla, *L-series and elliptic curves*, in *Lecture Notes in Mathematics 626, Number Theory Day*, ed. M.B. Nathanson, Springer-Verlag, Berlin, 1977, 1-42.

[7] S. Chowla, *The collected papers of Sarvadaman Chowla, Volumes 1-3*, eds. J.G.Huard and K.S. Williams, Les Publications CRM, 1999.

[8] Harald Cramér, *Ein Mittelwertsatz in der Primzahltheorie*, Math. Z. **12** (1922), 147-153.

[9] 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.

[10] David A. Cox, *Primes of the form* $x^2 + ny^2$, John Wiley and Sons, New York, 1989.

[11] Harold Davenport, *Multiplicative number theory, Second Edition*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1980.

[12] M. Deléglise and J. Rivat, *Computing* $\pi(x)$*: The Meissel, Lehmer, Lagarias, Miller, Odlyzko method*, Math. Comp. **65** (1996), no. 213, 235-245.

[13] C. Deninger, *On the analogue of the formula of Chowla and Selberg for real quadratic fields*, J. Reine. Angew. Math **351** (1984), 171-191.

[14] Max Deuring, *Asymptotische Entwicklungen der Dirichletschen L-Reihen*, Math. Annalen **168** (1967), 1-30.

[15] M. Deléglise and J. Rivat, *Computing the summation of the Möbius function*, Experimental Mathematics **5** (1996), no. 4, 291-294.

[16] Pierre Deligne and Jean Pierre Serre, *Formes modulaires de poids 1*, Ann. scient. E.N.S. 4^e **7** (1974), 507-530.

[17] W. Duke, *The dimension of the space of cusp forms of weight one*, Internat. Math. Res. Notices **2** (1995), 99-109.

[18] W. Duke, J. Friedlander, and H. Iwaniec, *Class group L-functions*, Duke Math Journal **79** (1995), 1-56.

[19] David S. Dummit and Richard M. Foote, *Abstract algebra*, Prentice-Hall, Englewood Cliffs, New Jersey, 1991.

[20] Andrey Feuerverger and Greg Martin, *Biases in the Shanks-Rényi prime number race*, Experimental Mathematics (to appear).

[21] J.B. Friedlander, *On the class number of certain quadratic extensions*, Acta. Arith. **28** (1976), 391-393.

[22] A. Fröhlich, *Galois module structure of algebraic integers*, Springer-Verlag, Berlin, 1980.

[23] A. Fröhlich and M.J. Taylor, *Algebraic number theory*, Cambridge University Press, Cambridge, Great Britain, 1994.

[24] Gerhard Frey, *Modular forms of low weight*, in *Elliptic Curves and Related Topics*, eds. Hershy Kisilevsky and M. Ram Murty, American Mathematical Society, Rhode Island, 1994, 1-21.

[25] P.X. Gallagher, *Some consequences of the Riemann hypothesis*, Acta Arithmetica **37** (1980), 339-343.

[26] S.M. Gonek, *On negative moments of the Riemann zeta-function*, Mathematika **36** (1989), 71-88.

[27] S.M. Gonek, *The second moment of the reciprocal of the Riemann zeta-function and its derivative*, Notes from a talk at MSRI, summer 1999.

[28] John F. Hart et al., *Computer Approximations*, John Wiley and Sons, New York, 1968.

[29] C.B. Haselgrove, *A disproof of a conjecture of Pólya*, Mathematika, **5**, 141-145.

[30] D.R. Heath Brown, *The distribution and moments of the error term in the Dirichlet divisor problem*, Acta Arithmetica, **60** (4) (1992), 389-415.

[31] H. Heilbronn, *Zeta-functions and L-functions*, in *Algebraic Number Theory*, eds. J.W.S. Cassels and A. Frohlich, Academic Press, London, 1967, 204-230.

[32] D. Hejhal, *On the distribution of log*$|\zeta'(\frac{1}{2}+it)|$, in *Number Theory, Trace Formula and Discrete Groups*, eds. Aubert, Bombieri, and Goldfeld, Academic Press, San Diego, 1989, 343-370.

[33] C.P. Hughes, J.P. Keating, and Neil O'Connell, *Random matrix theory and the derivative of the Riemann zeta function*, preprint, 2000.

[34] A.E. Ingham, *On two conjectures in the theory of numbers*, American J. of Math., **64** (1942), 313-319.

[35] Aleksandar Ivić, *The theory of the Riemann zeta function with applications*, John Wiley and Sons, New York, 1985.

[36] Henryk Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, Volume 17, American Mathematical Society, USA, 1997.

[37] H. Iwaniec and P. Sarnak, *Dirichlet L-functions at the central point*, in *Number theory in progress, Vol.2*, eds. Zakopane and Kościelisko, de Gruyter, Berlin, 1999, 941-952.

[38] H. Iwaniec and P. Sarnak, *The non-vanishing of central values of automorphic L-functions and Siegel's zeros*, Israel J. Math (to appear).

[39] Gerald J. Janusz, *Algebraic number fields*, Graduate Studies in Mathematics, Volume 7, American Mathematical Society, USA, 1996.

[40] J.P. Keating, *Random matrix theory and* $\zeta(\frac{1}{2}+it)$, Notes from a talk at MSRI, summer 1999.

[41] E. Kowalski and P. Michel, *Explicit upper bound for the rank of* $J_0(q)$, Israel J. Math (to appear).

[42] J.C. Lagarias, H.L. Montgomery, and A.M. Odlyzko, *A bound for the least prime ideal in the Chebotarev density theorem*, Inventiones Math **54** (1979), 271-296.

[43] J. Lagarias and A.M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in *Algebraic number fields*, ed. A. Frohlich, Academic Press, New York, 1977, 409-464.

[44] J. Lagarias and A.M. Odlyzko, *On Computing Artin L-functions in the critical strip*, Math. Comp. **33** (1979), no. 147, 1081-1095.

[45] Serge Lang, *Algebraic number theory, Second Edition* , Graduate Texts in Mathematics, Springer Verlag, New York, 1994.

[46] Serge Lang, *Introduction to modular forms*, Grundlehren der mathematischen Wissenschaften 222, Springer Verlag, Berlin, 1976.

[47] J.E. Littlewood, *On the class number of the corpus* $P(\sqrt{-k})$, Proc. London Math. Soc. (2), **27** (1928), 358-372.

[48] H.L. Montgomery, *The zeta function and prime numbers*, Proceedings of the Queen's Number Theory Conference, 1979, Queen's Univ., Kingston, Ont., 1980, 1-31.

[49] M. Ram Murty, *Some remarks on the Riemann hypothesis*, preprint, 1999.

[50] M. Ram Murty and V. Kumar Murty, *Non-vanishing of L-functions and applications*, Progress in Mathematics, Birkhauser Verlag, Basel, 1997.

[51] M. Ram Murty, V. Kumar Murty, and N. Saradha, *Modular forms and the Chebotarev density theorem*, American Journal of Mathematics **110** (1988), 253-281.

[52] V. Kumar Murty, *Explicit formulae and the Lang-Trotter conjecture*, Rocky Mountain Journal of Mathematics **15** (1985), no. 2, 535-551.

[53] V. Kumar Murty, *Chebotarev's density theorem*, Notes from mini-course at CRM, Fall 1998.

[54] V.K. Murty and J. Scherk, *Effective versions of the Chebotarev density theorem for function fields*, C.R. Acad. Sci. Paris t. **319** (1994), Série I, 523-528.

[55] A.M. Odlyzko, *The* 10^{20}*-th zero of the Riemann zeta function and 175 Million of its neighbors*, Manuscript in Preparation (see `http://www.research.att.com/~amo/unpublished/index.html`.)

[56] A.M. Odlyzko and H.J.J. te Riele, *Disproof of the Mertens conjecture*, J. Reine Angew. Math, **357** (1985), 138-160.

[57] William H. Press et al., *Numerical recipes in C: the art of scientific computing, second edition*, Cambridge University Press, New York, 1992.

[58] Glen Pugh, *The Riemann-Siegel formula and large scale computations of the Riemann zeta function*, Masters Thesis, University of British Columbia, 1998.

[59] H.J.J. te Riele, *On the sign change of the difference* $\pi(x) - \mathrm{Li}(x)$, Math. Comp. **48** (1986), 667-681.

[60] Michael Rubinstein, *Evidence for a spectral interpretation of zeros of L-functions*, Ph.D. Thesis, Princeton University, 1998.

[61] Michael Rubinstein, *Zeros of L-functions: computations and the behaviour of low zeros*, Notes from a talk given at MSRI conference, June 1999.

[62] Michael Rubinstein and Peter Sarnak, *Chebyshev's Bias*, Journal of Experimental Mathematics **3** (1994), 173-197.

[63] Robert Rumely, *Numerical computations concerning the ERH*, Math. Comp. **61** (1993), no. 203, 415-440, S17-S23.

[64] Jean-Pierre Serre, *A course in arithmetic*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1973.

[65] Jean-Pierre Serre, *Conducteurs d'Artin des caractères réels*, Inventiones Math. **14** (1973), 173-183.

[66] Jean-Pierre Serre, *Divisibilité de certaines fonctions arithmétiques*, Enseignement Math. (2) **22** (1976), no. 3-4, 227-260.

[67] Jean-Pierre Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1996.

[68] Jean-Pierre Serre, *Local fields*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1979

[69] Jean-Pierre Serre, *Modular forms of weight one and Galois representations* , in *Algebraic Number Fields*, ed. A. Frohlich. Academic Press, New York, 1977, 193-268.

[70] Jean-Pierre Serre, *Quelques applications du théorème de densité de Chebotarev*, Publ. Math. IHES **54** (1982), 123-201.

[71] Francesco Sica, *The order of vanishing of L-functions at the center of the critical strip*, Ph.D. Thesis, McGill University, 1998.

[72] K. Soundararajan, *Non-vanishing of quadratic Dirichlet L-functions at* $s = \frac{1}{2}$, Annals of Math. (to appear).

[73] H.M. Stark, *Some effective cases of the Brauer-Siegel theorem*, Inventiones math. **23** (1974), 135-152.

[74] H.M. Stark, *On complex quadratic fields with class-number two*, Math. Comp. 29 (1975), no. 129, 289-302.

[75] P. Stevenhagen and H.W. Lenstra, Jr., *Chebotarev and his density theorem*, The Mathematical Intelligencer **18** (1996), no. 2, 26-37.

[76] E.C. Titchmarsh, *The Theory of the Riemann zeta function, Second Edition*, Oxford University Press, New York, 1986.

[77] Emmanuel Tollis, *Zeros of Dedekind zeta functions in the critical strip*, Math. Comp. **33** (1997), no. 219, 1295-1321.