On the Hilbert-Pólya and Pair Correlation Conjectures

Author: Mathias Hudoba de Badyn
Supervisor: Dr. Greg Martin

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE
COMBINED HONOURS IN PHYSICS AND MATHEMATICS

_________________________________________  _______________________________________
Author                                      Date

_________________________________________  _______________________________________
Supervisor                                  Date
The complexities of the Riemann Zeta Function. [19]
1 Introduction
  1.1 Motivation for Studying $\zeta(s)$ ........................................... 5
  1.2 The Euler Product Formula ..................................................... 6
  1.3 $\zeta(s)$ and $\xi(s)$ .......................................................... 8
  1.4 The Conjecture of Hilbert and Pólya, and of Montgomery ............... 11

2 Riemann’s Wonderful Function ............................................... 14
  2.1 Introduction ................................................................. 14
  2.2 Prime Counting Functions ................................................. 14
    2.2.1 Perron’s Formula and Mellin Transform Theory .................... 16
    2.2.2 First Proof of $\psi(x)$ Formula .................................. 17
    2.2.3 Second Proof of $\psi(x)$ Formula ................................ 20
  2.3 Density of the Non-Trivial Roots ....................................... 22

3 A Primer on Physics ............................................................. 23
  3.1 Introduction ................................................................. 23
  3.2 Classical Mechanics ......................................................... 23
    3.2.1 Lagrangian Formalism ................................................ 24
    3.2.2 Hamiltonian Formalism .............................................. 26
  3.3 Quantum Mechanics .......................................................... 27
    3.3.1 Hilbert Spaces ......................................................... 27
    3.3.2 Operators in Quantum Mechanics .................................. 29
    3.3.3 Matrix Representation of an Operator ............................ 30
CONTENTS

3.3.4 Fundamental Equations ........................................... 31

4 The Hilbert-Pólya Conjecture .................................. 33
  4.1 Introductory Notes ................................................. 33
  4.2 Random Matrix Theory ............................................ 33
    4.2.1 Gaussian Unitary Matrices .................................. 34
    4.2.2 Pair Correlation Function for Gaussian Unitary Matrices .... 34
  4.3 Zeroes of $\zeta(s)$ and Random Matrices ....................... 36
  4.4 The Berry-Keating Operator ..................................... 40

5 Bibliography ......................................................... 43

Appendices ............................................................. 45

A Appendix ............................................................. 46

ABSTRACT

The Hilbert-Pólya Conjecture supposes that there exists an operator in a Hilbert space whose eigenvalues are the zeroes of the Riemann Zeta function $\zeta(s)$. This conjecture, if true, would very likely expedite the proof of the Riemann Hypothesis, namely that the non-trivial zeroes of $\zeta(s)$ have real part $\frac{1}{2}$. In this thesis we summarize work by Berry, Keating and others in constructing such an operator. Although the work so far has not yet yielded such an operator, some have been found that have properties very close to what is desired. We also summarize a (partially proven) conjecture by Montgomery that motivates the search for this operator. He conjectures that the pair correlation function for the spacing between the imaginary parts of the Riemann zeroes is the same as the correlation function for the spacing between eigenvalues of random Gaussian unitary matrices.

ACKNOWLEDGEMENTS

I would like to first thank my parents for the many things they have done that have helped contribute to my successes so far, one of which was their decision to give me a name which contains ‘math’ as a subsequence. I would also like to thank the laws of physics and mathematics which allow the Universe to exist the way it does and for allowing the atoms which comprise my body to have
evolved over the last 13.something billion years to the point where they can begin to comprehend
said laws in the form of the biological computation unit that is me.

I must also give thanks to my supervisor Dr. Greg Martin for his patience, guidance and
uniformly continuous support, and also his excellent clarity and wit which made the study of the
material in this thesis all the more interesting. There have also been several other mathematicians
and physicists at the University of British Columbia which have helped me a great deal in my
undergraduate years. I must thank Dr. Joanna Karczmarek for supervising my concurrently-
written physics Honours thesis in string theory, as well as Dr. Fok-Shuen Leung for mentoring me
in my first year in the Science One program, and Dr. Brian Marcus for advising my junior seminar
course in entropy, information theory and graph theory. I also thank all the professors in both the
physics and mathematics departments who have taught the various courses I have taken (and have
yet to take). In particular, I must thank Dr. David Morrissey, Dr. Mark Van Raamsdonk and Dr.
Ariel Zhitnitsky for teaching me elementary quantum mechanics, and quantum field theory which
is relevant in Chapters 3 and 4 of this thesis.

I should also acknowledge the ‘support’ of my peers, in particular my fellow students in the
Combined Honours in Physics and Mathematics program, past and present, and the Honours
Mathematics program. In alphabetical order (I think), these are Daniel Baker, Kahan Dare, Deshin
Finlay, Andrew Kuba Karpierz, Paul Liu, Kevin Martin, Emily Neufeld and Kai Ogasawara. In
particular, I should thank Justin Scarfy for lending me his copy of Edwards. I also thank the many
other students in Honours Physics, and Combined Honours in Physics with [not math], who are
too numerable to list here.
Chapter 1

INTRODUCTION

With great power, comes great
privilege. I mean responsibility.
Not Voltaire

1.1. Motivation for Studying $\zeta(s)$

In Bernhard Riemann’s classic paper *On the Number of Primes Less Than a Given Magnitude* [15], he attempted to find a formula for $\pi(x)$, the function whose value at $x$ is the number of prime numbers in the interval $[1, x]$. In the paper, he defined the seemingly innocent function:

**Definition 1 (The Riemann Zeta Function).** Let $n \in \mathbb{Z}$ and $s > 1$. The Riemann Zeta Function is defined as:

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

(1.1)

This was in part motivated due to a conjecture by Gauss, namely that this prime counting function behaves as

$$
\pi(x) \sim \frac{x}{\log x}.
$$

(1.2)

This was later formally proven by Hadamard and de la Vallée Poussin as the Prime Number Theorem:
Theorem 2 (The Prime Number Theorem).

Let the prime counting function be given by \( \pi(x) \). Then,

\[
\lim_{n \to \infty} \frac{\pi(x) \log x}{x} = 1.
\] (1.3)

Equivalently, as \( x \) goes to infinity,

\[
\pi(x) \sim \frac{x}{\log x}
\]

The proof may be found in Chapter 4 of [8]. The connection between the Riemann Zeta function and prime numbers lies primarily in the Euler Product.

1.2. The Euler Product Formula

Theorem 3 (The Euler Product Theorem). Let the Riemann Zeta Function be

\[
\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.
\]

Then,

\[
\zeta(s) = \prod_p \frac{1}{1 - p^{-s}},
\] (1.4)

where \( p \) denotes prime numbers, and \( \Re(s) > 1 \).

Proof. Note that

\[
\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots
\]

and

\[
\frac{1}{2^s} \zeta(s) = \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \frac{1}{8^s} + \ldots.
\]

The right-hand side of the second line contains all of the terms of the Riemann Zeta function corresponding to the integers that have 2 as a factor. Subtracting these two equations yields

\[
\left(1 - \frac{1}{2^s}\right) \zeta(s) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \ldots
\]

The right-hand side of this equation now has all the terms from \( \zeta(s) \) with factor \( \frac{1}{2^s} \) removed. Note that these right-hand sides are all absolutely convergent Dirichlet series for \( \Re(s) > 1 \),
since the numerators are all 1, and thus bounded. Now, repeat the multiplication and
subtraction steps for $\frac{1}{3^s}$:

$$\frac{1}{3^s} \left( 1 - \frac{1}{2^s} \right) \zeta(s) = \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{15^s} + \ldots .$$

$$\left( 1 - \frac{1}{3^s} \right) \left( 1 - \frac{1}{2^s} \right) \zeta(s) = 1 + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{13^s} + \ldots .$$

We can see that all numbers in this infinite series that have a factor of 2 or 3 are ‘sieved’
in a manner like the Sieve of Erastosthenes. By the Fundamental Theorem of Arithmetic
every number that is not prime can be decomposed into prime factors, and hence repeating
this algorithm for all primes $p$ ad infinitum yields:

$$\zeta(s) \left( 1 - \frac{1}{2^s} \right) \left( 1 - \frac{1}{3^s} \right) \left( 1 - \frac{1}{5^s} \right) \cdots = 1 .$$

The convergence of the right-hand side follows from the absolute convergence of the Dirichlet
series for $\Re(s) > 1$. It is now clear that

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}, \quad (1.5)$$

however we must now show that the right-hand side of 1.5 converges. Consider

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

$$= \log \left( \frac{1}{1 - 2^{-s}} \right) + \log \left( \frac{1}{1 - 3^{-s}} \right) + \log \left( \frac{1}{1 - 5^{-s}} \right) + \ldots$$

$$= - \left[ \log(1 - 2^{-s}) + \log(1 - 3^{-s}) + \log(1 - 5^{-s}) + \ldots \right]$$
We claim that this is absolutely convergent.

\[ |\log(1 - 2^{-s}) + \log(1 - 3^{-s}) + \log(1 - 5^{-s}) + \ldots| \]
\[ \leq |\log(1 - 2^{-s})| + |\log(1 - 3^{-s})| + |\log(1 - 5^{-s})| + \ldots \]
\[ \leq \left| \frac{1}{2^2} \right| + \left| \frac{1}{3^s} \right| + \left| \frac{1}{5^s} \right| + \ldots \]
\[ \leq \left| \frac{1}{1^s} \right| + \left| \frac{1}{2^s} \right| + \left| \frac{1}{3^s} \right| + \ldots, \]

where the bound \(|\log(1 + x)| \leq |x|\) is given from Equation 2 in [18]. Hence, by the absolute convergence of a Dirichlet series with \(\Re(s) > 1\) yields the convergence of the Euler product for \(\Re(s) > 1\).

This theorem should make clear that there is indeed a very intimate relationship between prime numbers and \(\zeta(s)\). In chapter 2, we utilize this relationship to outline the construction of \(\pi(x)\) from \(\zeta(s)\). For now, we discuss some properties of the Riemann Zeta function, including its convergence, analytic continuation and its relationship with its ‘sister’ function \(\xi(s)\). In the last section of this chapter, we give a brief overview of the Hilbert Pólya and Pair Correlation Conjectures, which are the main topics of this thesis.

1.3. \(\zeta(s)\) and \(\xi(s)\)

We can write the Riemann Zeta function in implicit functional form, and relate it to a function \(\xi(s)\). This will be useful in further discussions in Chapter 2. First, recall the definition of the Gamma Function.

**Definition 4 (The Gamma Function).** Let \(s\) be a complex number with real part greater than 0. Define

\[ \Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx. \quad (1.6) \]

Then, via integration by parts, we have

\[ \Gamma(s+1) = s\Gamma(s). \quad (1.7) \]
CHAPTER 1. INTRODUCTION

Noting that $\Gamma(1) = 1$, we can see that for non-negative integer values of $\Gamma$, we have that $\Gamma(n) = (n - 1)!$. Via coordinate substitution, we can transform $x \rightarrow \pi^n x$ for integer $n$ into Equation 1.6 to get

$$\pi^{s} \Gamma \left( \frac{s}{2} \right) n^{-s} = \int_{0}^{\infty} x^{s-1} e^{-n^{2}\pi x} dx. \quad (1.8)$$

Next, we can sum both sides over all positive integers $n$.

$$\sum_{n=1}^{\infty} \pi^{s} \Gamma \left( \frac{s}{2} \right) n^{-s} = \sum_{n=1}^{\infty} \int_{0}^{\infty} x^{s-1} e^{-n^{2}\pi x} dx. \quad (1.9)$$

giving a factor of $\zeta(s)$ on the left-hand-side. Note that by doing this, we lose convergence for $\Re(s) \in (0, 1)$. Thus, $\Re(s) > 1$. On the right-hand side, we have absolute convergence of the sum, so we can interchange the order of summation and integration. We now state a lemma that will lead us towards defining a function $\xi(s)$ which we will show to have some interesting and useful properties related to $\zeta(s)$.

**Lemma 5 (A Lemma Pertaining to $\zeta(s)$).** Define the useful function

$$\phi(x) = \sum_{n=1}^{\infty} e^{-n^{2}\pi x}. \quad (1.10)$$

and let $\Gamma(s)$ be defined as before. Then,

$$\pi^{s} \Gamma \left( \frac{s}{2} \right) \zeta(s) = -\frac{1}{s(1-s)} + \int_{1}^{\infty} \left( x^{s-1} + x^{1+s-1} \right) \phi(x) dx. \quad (1.11)$$

**Proof.** Following the proof in [16], the $\phi(x)$ function

$$\phi(x) = \sum_{n=1}^{\infty} e^{-n^{2}\pi x} \quad (1.12)$$

will give

$$\pi^{s} \Gamma \left( \frac{s}{2} \right) \zeta(s) = \int_{0}^{\infty} x^{s-1} \phi(x) dx. \quad (1.13)$$

We can now split the integral on the right-hand-side into an integral from $x \in (0, 1]$ and $x \in [1, \infty)$:

$$\pi^{s} \Gamma \left( \frac{s}{2} \right) \zeta(s) = \int_{0}^{1} x^{s-1} \phi(x) dx + \int_{1}^{\infty} x^{s-1} \phi(x) dx. \quad (1.14)$$
Substituting \( x \rightarrow \frac{1}{x} \) in the integral from \( x \in (0, 1] \) yields an integral from \( x \in [1, \infty) \). This is justified because the integral is still convergent under this coordinate transformation. We get:
\[
\pi^{-\frac{1}{2}} \Gamma \left( \frac{s}{2} \right) \zeta(s) = \int_1^\infty x^{-\frac{1}{2}-1} \phi(1/x) \, dx + \int_1^\infty x^{-\frac{1}{2}-1} \phi(x) \, dx.
\]
(1.15)

With some work, one may derive from Equation 1.10 that
\[
\phi \left( \frac{1}{x} \right) = \frac{\sqrt{x}}{2} - \frac{1}{2} + \sqrt{x} \phi(x).
\]
(1.16)

Substituting this into the integral in Equation 1.15 and integrating out the constant terms will give:
\[
\pi^{-\frac{1}{2}} \Gamma \left( \frac{s}{2} \right) \zeta(s) = -\frac{1}{s(1-s)} + \int_1^\infty \left( x^{-\frac{1}{2}-1} + x^{1-s} \right) \phi(x) \, dx.
\]
(1.17)

The right-hand side of this equation is in fact meromorphic on \( \mathbb{C} \), since the integral converges absolutely for all complex numbers due to the behaviour of \( \phi(x) \) as \( x \) approaches infinity. We can bound \( \phi(x) \) as
\[
\phi(x) = \sum_{n=1}^\infty \frac{e^{-\pi n^2 x}}{n!} \leq \sum_{n=1}^\infty e^{-\pi n^2} = \frac{1}{e^{\pi x} - 1}
\]
(1.18)

Therefore, \( \phi(x) \) behaves as \( O(e^{-\pi x}) \) as \( x \) approaches \( \infty \) and imposes absolute convergence on the integral in Equation 1.17. Furthermore, the right-hand-side is invariant under the transformation \( s \rightarrow 1 - s \). From this, we can define the function
\[
\xi(s) = \frac{1}{2} \frac{\pi^{-\frac{1}{2}} \Gamma \left( \frac{1}{2} \right)}{s(s-1)} \zeta(2),
\]
(1.19)

which from Equation 1.17 satisfies:
\[
\xi(s) = \xi(1 - s).
\]
(1.20)

The \( \xi(s) \) function is entire for \( \Re(s) > 0 \), since the factor of \( (s-1) \) cancels the pole of \( \zeta(s) \) at \( s = 1 \). We can also see from this definition that there are zeros of \( \zeta(s) \) at the poles of \( \Gamma(s/2) \) at the negative even integers (the factor of \( s \) counters the pole at \( s = 0 \)). These are referred
to as the 'trivial zeros', in that they are trivial due to rather well-understood behaviour of \( \Gamma(s) \). A colour plot of the Riemann zeta function (colours corresponding to the argument of \( \zeta(s) \)) can be seen in Figure 1.1.

We are now in a position to state the famous hypothesis of Riemann about the other (non-trivial) zeros of \( \zeta(s) \).

**Conjecture 6 (The Riemann Hypothesis).** Let \( \zeta(s) \) be defined by some analytic continuation on \( \mathbb{C} \) (perhaps Equation 1.17. Then, all non-trivial zeros of \( \zeta(s) \), namely those that do not come from the poles of \( \Gamma(s/2) \), have real part \( \frac{1}{2} \).

Hopefully one day we (I) will also be in a position to prove Conjecture 6.

1.4. The Conjecture of Hilbert and Pólya, and of Montgomery

The Hilbert-Pólya conjecture is one method of tackling the problem of proving or disproving the Riemann Hypothesis. The conjecture is as follows.

**Conjecture 7 (The Hilbert-Pólya Conjecture).** There exists a self-adjoint operator in an infinite-dimensional Hilbert space whose eigenvalues are the non-trivial Riemann zeros.

If this conjecture is true, then it would be a (presumably) relatively simple transition to equate the Riemann Hypothesis to a statement about the eigenvalues of this operator (perhaps other than the obvious statement ‘the real part of the eigenvalues is \( \frac{1}{2} \)’). Since linear algebra has been extremely well studied in the past century, in part due to the advent of quantum mechanics and the need for engineers to study and solve large systems of linear equations, it is a promising idea that reducing the Riemann Hypothesis to a statement about the eigenvalues of this operator would simplify the proof or disproof significantly. Although not formally discussed in this thesis, Connes has created a trace formula equivalent to the Riemann Hypothesis.

Montgomery was studying the pair correlation between spacings of the zeros of the Riemann Zeta function. Upon presenting his work at the Institute for Advanced Study, Freeman Dyson, who was one of the early figures in the study of random matrices\(^1\), announced that the pair correlation function Montgomery announced was identical to the one for the spacings between eigenvalues of random Gaussian unitary matrices. Thus we have the final conjecture discussed in the thesis:

\(^1\)Along with Eugene Wigner and others
**Conjecture 8** (Montgomery’s Pair Correlation Conjecture [13]). Let ζ(s) be the Riemann Zeta function. The zeros of the Riemann Zeta function, normalized to have unit spacing on average, have pair correlation function:

\[ 1 - \left( \frac{\sin(\pi u)}{\pi u} \right)^2. \] (1.21)

As discussed in Chapter 4, this conjecture was already partially proven by Montgomery in his paper introducing it. This fact strengthens the idea that the Hilbert-Pólya Conjecture offers a suitable approach to a proof or disproof of the Riemann Hypothesis.
Figure 1.1: Colour plot of $\zeta(s)$. Colour corresponds to Arg($\zeta(s)$). Note the trivial zeros on the bottom, and the first 3 non-trivial zeros on the line $\Re(s) = \frac{1}{2}$. Also note that while circling around a small contour around each zero, every colour is encountered once. Figure code from [7].
2.1. Introduction

In this chapter, we will expand further on the properties of $\zeta(s)$ and develop further intuition about the Riemann zeros. We begin with an overview of prime-counting functions and will see where $\zeta(s)$ arises naturally out of this theory. This will turn into a discussion of the $J(x)$ and $\psi(x)$ functions which will then end with a walkthrough of the theory behind the density of the Riemann zeros. This will give enough background to sufficiently study the ideas laid out in Chapter 4.

2.2. Prime Counting Functions

Let us begin by making a few informal statements and then later making them more precise. First, take the logarithm of both sides in the result of Theorem 3. This yields:

$$\log \zeta(s) = \sum_{p} \left( \sum_{n} \frac{p^{-ns}}{n} \right)$$

(2.1)
after noting the series expansion \( \log(1 - x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \ldots \), we can define the function

\[
J(x) = \frac{1}{2} \left[ \sum_{p^n < x} \frac{1}{n} + \sum_{p^n \leq x} \frac{1}{n} \right]
\] (2.2)

and apply it as a Stieltjes measure for Equation 2.1. Recall the definition of Riemann-Stieltjes integration.

**Definition 9** (Riemann-Stieltjes Integration (Definition 6.2 in [17])). Let \([a, b]\) be an integration domain with partition \(P\), such that \(P = \{a = x_0 < x_1 < \cdots < x_n = b\}\), without loss of generality. The Riemann-Stieltjes sum is defined as

\[
S(P, f, g) = \sum_{i=0}^{n-1} f(c_i)[g(x_{i+1}) - g(x_i)]
\] (2.3)

for some \(c_i\) in the interval \([x_i, x_{i+1}]\), and for some real functions \(f, g\). The Riemann-Stieltjes integral \(\int_a^b f(x)dg(x)\) equals \(A\) if for every \(\epsilon > 0\), there exists a \(\delta\) such that \(\max(|x_{i+1} - x_i|) < \delta\) implies that \(|S(P, f, g) - A| < \epsilon\) for all choices of \(c_i\).

Using \(J(x)\) as our \(g\) function, and taking \(b \to \infty\) gives the integral

\[
\log \zeta(s) = s \int_0^\infty x^{-1-s}J(x)dx.
\] (2.4)

Von Mangoldt realized that this could be written as

\[
\frac{\zeta'(s)}{\zeta(s)} = -\int_0^\infty x^{-s} \log(x)dJ(x).
\] (2.5)

Here, \(\log(x)dJ(x)\) is a measure that weights \(\frac{\log(p^n)}{n}\) to prime powers \(p^n\). We can write it as a Stieltjes measure \(\psi(x)\):

\[
\psi(x) = \sum_{p^n < x} \log(p).
\] (2.6)

this means that via integrating Equation 2.5 by parts we get

\[
-\frac{\zeta'(s)}{\zeta(s)} = s \int_0^\infty x^{-s-1}d\psi(x),
\] (2.7)
since the boundary terms are zero. In the next few subsections, we would like to give a more precise proof of these statements. First, we will take a tangent and talk about Perron’s Formula and Mellin Transforms, which will prove to be useful.

2.2.1. Perron’s Formula and Mellin Transform Theory

Perron’s Formula is an expression for the infinite sum of arithmetic functions. It is given in terms of the inverse of what will be subsequently defined as a Mellin transform.

**Theorem 10** (Perron’s Formula (Theorem 11.18 in [2])). Let \( \{a(n)\} \) be an arithmetic sequence, and let \( h(s) \) be an absolutely convergent Dirichlet series for \( \Re(s) > c \) for some positive \( c \) given by

\[
h(s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}. \tag{2.8}
\]

Let \( x > 0 \) be a real number. Then, Perron’s formula is

\[
A(x) = \sum_{n \leq x} a(n) = \frac{1}{2\pi i} \int_{d-i\infty}^{d+i\infty} h(z) \frac{x^z}{z} dz, \tag{2.9}
\]

where \( d > c \), and if \( x \) is an integer, \( a(x) \) is multiplied by a factor of \( \frac{1}{2} \). Further, \( x > 0 \) for real \( x \).

**Proof.** We defer the reader to page 245 of [2]. \( \square \)

**Definition 11** (Mellin Transform). Let \( f \) be a real-valued function. The Mellin transform of \( f \) is given by

\[
\{Mf\}(s) = \phi(s) = \int_{0}^{\infty} x^{s-1} f(x) dx. \tag{2.10}
\]

The inverse of the Mellin transform is given by

\[
\{M^{-1}\phi\}(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} \phi(s) ds. \tag{2.11}
\]

The conditions under which one may take the inverse Mellin transform of a function \( \phi(s) \) are given by the Mellin Inversion Theorem.

**Theorem 12** (Mellin Inversion). Let \( \phi(s) \) be a complex-valued function. Suppose that \( \phi(s) \) is analytic in the strip \( a < \Re(s) < b \), and that \( \phi(s) \to 0 \) uniformly as \( \Im(s) \to \pm\infty \) for
any $c \in (a,b)$. Further, suppose that integral of $\phi(s)$ on the line $(c - i\infty, c + i\infty)$ converges absolutely. Then, if
\[
\{M^{-1}\phi\}(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s}\phi(s)ds,
\] we have that
\[
\{Mf\}(s) = \phi(s) = \int_0^\infty x^{s-1}f(x)dx.
\] (2.13)

Proof. We refer the reader to [14] for the proof, as it does not offer any useful insights or methodology.

We can thus see that Equation 2.9 in Theorem 10 is given by an inverse Mellin transform.

2.2.2. First Proof of $\psi(x)$ Formula

In this subsection, we will give a proof of the formula for $\psi(x)$. This will provide useful background when discussing the density of the Riemann zeros. First, let us define the logarithmic integral.

Definition 13 (Logarithmic Integral). Let $x$ be a real variable greater than or equal to 1. The logarithmic integral function is defined as
\[
\text{Li}(x) = \lim_{\epsilon \to 0} \left[ \int_0^{1-\epsilon} \frac{dt}{\log t} + \int_{1+\epsilon}^x \frac{dt}{\log t} \right].
\] (2.14)

Riemann proved the following theorem, which will require a proof of an important integration formula in its own proof.

Theorem 14 (Logarithmic Integral Theorem). Let $a$ be real and greater than 1. Then,
\[
\text{Li}(x) = \frac{1}{2\pi i} \log x \int_{a-i\infty}^{a+i\infty} \frac{d}{ds} \left[ \frac{\log(s-1)}{s} \right] x^s ds.
\] (2.15)

Before we prove Theorem 14, let us first prove the following.

Theorem 15 (An Integration Formula). Let $a$ be real and greater than 1. Further, let $a > \Re(\beta)$. Then,
\[
\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{1}{s-\beta} y^s ds = \begin{cases} 
 y^\beta & \text{if } y > 1 \\
 \frac{1}{2} y^\beta & \text{if } y = 1 \\
 0 & \text{if } 0 < y < 1
\end{cases}
\] (2.16)
Proof. Following the proof from Section 1.14 of [8], consider the following integral for \( \Re(s - \beta) > 0 \). Applying the Fourier Inversion Theorem for \( a > \Re(\beta) \)

\[
\frac{1}{s - \beta} = \int_1^\infty x^{-s} x^{\beta-1} dx. \tag{2.17}
\]

\[
\frac{1}{a + i\mu - \beta} = \int_0^\infty e^{-i\lambda\mu} e^{\lambda(\beta-a)} d\lambda \tag{2.18}
\]

we get

\[
\int_{-\infty}^\infty \frac{1}{a + i\mu - \beta} e^{i\beta x} d\mu = \begin{cases} 
2\pi e^{x(\beta-a)} & \text{if } x > 0 \\
0 & \text{if } x < 0 \end{cases}. \tag{2.19}
\]

Assuming that \( a > \Re(\beta) \), taking the integral from \( \pm i\infty \) offset by the positive real number \( a \), we get the result:

\[
\frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \frac{1}{s - \beta} \frac{1}{y^s} ds = \begin{cases} 
y^\beta & \text{if } y > 1 \\
\frac{1}{2} y^\beta & \text{if } y = 1 \\
0 & \text{if } 0 < y < 1 \end{cases}. \tag{2.20}
\]

\[\square\]

Next, let us consider, without proof, some results for the exact value of \( J(x) \). Von Mangoldt proved Riemann’s original formula for \( J(x) \):

\[
J(x) = \text{Li}(x) - \sum_{\rho} \text{Li}(x^\rho) - \log 2 + \int_x^\infty \frac{dt}{t(t^2 - 1) \log t}, \tag{2.21}
\]

where \( \rho \) corresponds to the Riemann zeros. We would like to use this result to relate \( J(x) \) to \( \psi(x) \) in a more rigorous way than considered in the introduction of this chapter. This result gives the Stieltjes measure for \( x > 1 \):

\[
dJ = \left( \frac{1}{\log x} - \sum_{\rho} \frac{x^{\rho-1}}{\log x} - \frac{1}{x(x^2 - 1) \log x} \right) dx \tag{2.22}
\]

Then, informally we can say \( dJ \log x = d\psi \):

\[
d\psi = \left( 1 - \sum_{\rho} x^{\rho-1} - \sum_{\rho} x^{-2n-1} \right) dx. \tag{2.23}
\]
This led Von Mangoldt to guess
\[
\psi = x - \sum_{\rho} \frac{x^\rho}{\rho} + \sum_n \frac{x^{-2n}}{2n} + C
\]  
(2.24)
as the formula for \( \psi(x) \). We now begin the derivation of this claim. First, define the \( \Pi \) function.

**Proof.** Following Chapter 3 in [8]:

**Definition 16** *(The (Capital) Pi Function).* Let \( \Gamma(z) \) be defined as before. Then,

\[
\Pi(z) = \Gamma(z + 1) = z\Gamma(z) = \int_0^\infty e^{-t} t^z dt.
\]  
(2.25)

Recall that we have as in Equation 2.7

\[
-\frac{\zeta'(s)}{\zeta(s)} = s \int_0^\infty \psi(x)x^{-s-1} dx.
\]  
(2.26)

Using a now-well-defined Mellin transform, we get

\[
\psi(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left[ -\frac{\zeta'(s)}{\zeta(s)} \right] x^s \frac{ds}{s}.
\]  
(2.27)

Now let's define the \( \Lambda \) function as the weight assigned to \( n \) by the \( d\psi \) measure:

\[
\Lambda(n) = \begin{cases} 
\log p & \text{if } n = p^\alpha \\
0 & \text{else}
\end{cases}
\]  
(2.28)

Then, we can rewrite Equation 2.26 as

\[
-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=2}^\infty \frac{\Lambda(n)}{n^s}.
\]  
(2.29)

Inserting Equation 2.29 into Equation 2.27 yields:

\[
\psi(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left[ \sum_{n=2}^\infty \frac{\Lambda(n)}{n^s} \right] x^s \frac{ds}{s}.
\]  
(2.30)
\[
\psi(x) = \sum_{n=2}^{\infty} \Lambda(n) \int_{a-i\infty}^{a+i\infty} \left(\frac{x}{n}\right)^s \frac{ds}{s}.
\]

**Theorem 17** (Term-by-term integration (Section 9.9 of [1]).

Suppose that \( u_k \) is integrable on \([a, b]\) for all \( k \geq 1 \) and that \( \sum_{k=1}^{\infty} \) converges uniformly on \([a, b]\). Then, the limit function \( f(x) \) is integrable on \([a, b]\) and

\[
\int_a^b f(x)dx = \sum_{k=1}^{\infty} \int_a^b u_k(x)\]

Noting that Equation 2.30 converges uniformly, by Theorem 17 it follows that

\[
\psi(x) = \sum_{n<x} \Lambda(n),
\]

using the integration formula proven in Theorem 15.

\[\Box\]

2.2.3. Second Proof of \( \psi(x) \) Formula

Partially following Section 3.2 in [8], we can use the symmetry of \( \zeta(s) \) and \( \xi(s) \) to write

\[
\Pi \left( \frac{s}{2} \right) \pi \bar{\zeta}(s - 1) \zeta(s) = \xi(0) \prod_p \left( 1 - \frac{s}{p} \right).
\]

Taking the logarithmic derivative will give

\[
\frac{d}{ds} \log \Pi \left( \frac{s}{2} \right) - \frac{1}{2} \log \pi + \frac{1}{s - 1} + \frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{1}{1 - \frac{s}{p}} \left( -\frac{1}{p} \right)
\]

Next, solve for \(-\frac{\zeta'(s)}{\zeta(s)}\) and substitute

\[
\Pi(s) = \prod_{n=1}^{\infty} \left( 1 + \frac{s}{n} \right)^{-1} \left( 1 + \frac{1}{n} \right)^s
\]
to get

\[-\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{s - 1} - \sum_p \frac{1}{s - p} + \sum_{n=1}^{\infty} \left[ -\frac{1}{s + 2n} + \frac{1}{2} \log \left(1 + \frac{1}{n}\right) \right] - \frac{\log \pi}{2}, \tag{2.35}\]

which gives us

\[-\frac{\zeta'(0)}{\zeta(0)} = -1 - \sum_p \frac{1}{p} - \sum_{n=1}^{\infty} \left[ -\frac{1}{2n} + \frac{1}{2} \log \left(1 + \frac{1}{n}\right) \right] - \frac{\log \pi}{2}. \tag{2.36}\]

Using equation 2.36 and equation 2.35, we get

\[-\frac{\zeta'(s)}{\zeta(s)} + \frac{\zeta'(0)}{\zeta(0)} = \left[ 1 + \frac{1}{s - 1} \right] - \sum_p \left[ \frac{1}{p} + \frac{1}{s - p} \right] - \sum_{n=1}^{\infty} \left[ \frac{1}{s + 2n} - \frac{1}{2n} \right]. \tag{2.37}\]

Simplifying this will give us

\[-\frac{\zeta'(s)}{\zeta(s)} = \frac{s}{s - 1} - \sum_p \frac{s}{p(s - p)} + \sum_{n=1}^{\infty} \frac{s}{2n(s + 2n)} - \frac{\zeta'(0)}{\zeta(0)} \tag{2.38}\]

Plug equation 2.38 into equation 2.27 to get

\[
\psi(x) = \frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \left[ -\frac{\zeta'(s)}{\zeta(s)} \right] x^s \frac{ds}{s} \tag{2.39} \]

\[
= \frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \frac{x^s}{s - 1} ds - \sum_p \int_{a - i\infty}^{a + i\infty} \frac{x^s}{p(s - p)} ds + \frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \left( \sum_n \frac{x^s}{2n(s + 2n)} + \left[ -\frac{\zeta'(0)}{\zeta(0)} \right] \frac{1}{s} \right) ds.
\]

Let \( t = s - \beta \) and use the integration formula proven in Theorem 15 which then gives

\[
\frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \frac{1}{s - \beta} y^s ds = \left\{ \begin{array}{ll} y^\beta & \text{if } y > 1 \\ 0 & \text{if } y < 1 \end{array} \right. \tag{2.40} \]

and

\[
\frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} \frac{x^s}{s - \beta} = \frac{1}{2\pi i} \int_{a - i\infty - \beta}^{a + i\infty - \beta} x^t \frac{dt}{t} = \frac{x^\beta}{2\pi i} \int_{\Re(a - \beta) - i\infty}^{\Re(a - \beta) + i\infty} \frac{x^t}{t} dt = x^\beta.
\]
2.3. Density of the Non-Trivial Roots

Let $N$ denote the number of roots a function $f$ has in a complex contour $\gamma$, and let $P$ be the number of poles in this contour. Then by Cauchy’s Argument Principle (Theorem 2 in Section 3.1 of [9]),

$$N - P = \frac{1}{2\pi i} \oint_{\gamma} \frac{f'(z)}{f(z)} \, dz. \tag{2.41}$$

Consider the complex rectangle $RT$ with $0 < \Im(s) < T$ and $0 < \Re(s) < 1$, with $T > 0$. Then, integrating around $RT$ gives:

$$N(T) = \frac{1}{2\pi i} \int_{RT} \frac{\zeta'(s)}{\zeta(s)} \, ds \tag{2.42}$$

**Theorem 18.** Let $RT$ be as before with $T > 0$. Then,

$$N(T) = \frac{T}{2\pi} \log \left( \frac{T}{2\pi} \right) - \frac{T}{2\pi} + O(\log T) \tag{2.43}$$

In particular, we are interested in Backlund’s estimate (page 31 of [6]):

$$\left| N(T) - \left( \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} \right) \right| < 0.137 \log T + 0.443 \log \log T + 4.350 \tag{2.44}$$

for $T > 0$. We leave these results without proof, however the proof is readily available in Section 6.7 of [8].
3.1. Introduction

In this section we outline some of the main principles of physics, and in particular quantum mechanics so that the reader may have sufficient background to understand the terminology in the discussion of Berry’s Hilbert-Pólya operator. Due to the non-rigorous nature of physics as a discipline, some of the following is non-rigorous. The author, caught in the crossfire between the two disciplines, tries his best to introduce rigour whenever possible.

3.2. Classical Mechanics

Classical mechanics is the study of particle motion with the assumption that the motion is represented by a continuous function. Here, a ‘particle’ is any object whose volume can be neglected. Furthermore, the various observables (mathematical quantities that an experimentalist can measure) are also given by continuous functions, such as energy, position, and momentum. In contrast, quantum mechanics ‘discretizes’ the values of energy a particle can have. All material in this section can be found in [10].
3.2.1. LAGRANGIAN FORMALISM

The Lagrangian formalism is the standard variational calculus method for determining the ‘equations of motion’ of a system. These equations of motion are usually in the form of a second-order differential equation in position. The most familiar of these should be Newton’s Second Law:

**Definition 19** (Newton’s Second Law). The equations of motion of a particle with mass $m$ are of the form

$$m\frac{d^2 x}{dt^2} = f\left(x, \frac{dx}{dt}\right)$$  \hspace{1cm} (3.1)

for some continuous and differentiable function $f$.

The Lagrangian formalism begins with the definition of the Lagrangian:

**Definition 20** (The Lagrangian). The Lagrangian is a function of variables $q$ and $\dot{q}$ given by

$$\mathcal{L}(q, \dot{q}) = T - V,$$  \hspace{1cm} (3.2)

where $T$ is the kinetic energy function, $V$ the potential energy, and the dot referring to a time derivative.

In this definition, $q$ and $\dot{q} = \frac{d}{dt}q$ are some arbitrary variables that characterize the system. Typically we take $q$ to be position (of say, a particle), and $\dot{q}$ to be velocity. Before we define the functional action, we have to summarize the Fundamental Lemma of Calculus of Variations.

**Definition 21** (Differentiability Classes). A function $f$ is said to be of differentiability class $C^k$ if $f$ is $k$ times continuously differentiable.

**Lemma 22** (Fundamental Lemma of Calculus of Variations). Let $f$ be a $k$-times continuously differentiable function (of class $C^k$) on $[a,b]$. If for every function $g$ on $[a,b]$ of class $C^k$ with $g(a) = g(b) = 0$ we have that

$$\int_a^b f(x)g(x)dx = 0,$$  \hspace{1cm} (3.3)

then $f(x) = 0$ on $[a,b]$. 

Now, we can define the action and examine some of its properties.

**Definition 23 (The Action).** The action is given by the functional

\[ S[\mathcal{L}] = \int_{t_0}^{t_1} \mathcal{L}(q, \dot{q}) dt. \]  

(3.4)

The action is an integral over all ‘paths’ parameterized by \( \mathcal{L} \), and particularly \( q(t) \). \( q(t) \) is normally associated with position, so a ‘path’ here just means some trajectory of a particle. By assumption, \( q(t) \) has continuous first derivatives. We will identify some form of \( q(t) \) as the \( g(x) \) function in Lemma 22. We wish to find an extremal value of this action in order to derive the equations of motion. In order to do so, we assume what physicists call the ‘Principle of Least Action’. This essentially means that at the boundary of the integral, namely at \( t_0, t_1 \), we assume that the variation of the action \( \delta S \), and its variables \( \delta q, \delta \dot{q} \) are zero. We can then take the variation of the action:

\[ \delta S = 0 = \int_{t_0}^{t_1} \left( \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} \right) dt. \]  

(3.5)

Integrating the \( \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} \) term by parts yields:

\[ \delta S = 0 = \int_{t_0}^{t_1} \left( \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q \right) dt + \left. \frac{\partial \mathcal{L}}{\partial q} \delta q \right|_{t_0}^{t_1}. \]  

(3.6)

The boundary term is zero by the Principle of Least action. Since the variation of the action is zero, the integrand is also zero, and within the range \( (t_0, t_1) \), we have that \( \delta q > 0 \) which implies that by Lemma 22

\[ \frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0. \]  

(3.7)

This is the coveted Euler-Lagrange equation. Furthermore, this definition is easily generalizable to multiple \( q \)-variables. The general Euler-Lagrange equations are:

**Definition 24 (The Euler-Lagrange Equations).** Let the Lagrangian \( \mathcal{L} \) be defined as before. The Euler-Lagrange equations are given by:

\[ \sum_{i=1}^{n} \left[ \frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] = 0. \]  

(3.8)
3.2.2. Hamiltonian Formalism

The Hamiltonian formalism is equivalent to the Lagrangian formalism under a transformation of coordinates. Roughly speaking, one transforms the velocity coordinate into momentum. This simplifies the transition from classical physics to quantum physics immensely, because momentum is a ‘conserved’ quantity whereas velocity is not. A conserved quantity is one that is constant in time. We begin by defining the Hamiltonian:

**Definition 25** (The Hamiltonian). The Hamiltonian \( \mathcal{H} \) is a function of variables \( q \) and \( p \), where \( p \) and its time derivative are given by

\[
p = \frac{\partial L}{\partial \dot{q}}
\]
\[
\dot{p} = \frac{\partial L}{\partial q}.
\]  

(3.9)

The Hamiltonian is then defined as

\[
\mathcal{H} = \frac{\partial L}{\partial \dot{q}} \dot{q} - L.
\]  

(3.10)

Taking a variation of the Hamiltonian gives the following:

\[
\delta \mathcal{H} = \frac{\partial \mathcal{H}}{\partial q} \delta q + \frac{\partial \mathcal{H}}{\partial \dot{q}} \delta \dot{q}
\]

(3.11)

\[
\delta \mathcal{H} = \dot{q} \delta p + p \delta \dot{q} - \frac{\partial L}{\partial q} \delta q - \frac{\partial L}{\partial p} \delta \dot{p}
\]  

(3.12)

Matching coefficients from Equations 3.11 and 3.12 gives Hamilton’s equations.

**Definition 26** (Hamilton’s Equations). Let \( \mathcal{H} \) denote the Hamiltonian. Hamilton’s equations are given by:

\[
\frac{\partial \mathcal{H}}{\partial q} = -\dot{p}
\]

\[
\frac{\partial \mathcal{H}}{\partial p} = \dot{q}.
\]  

(3.13)

We can write these equations in a more concise form using Poisson brackets.
Definition 27 (Poisson Brackets). Let $\mathcal{H}$ be the Hamiltonian, and let $f$ be some continuous and differentiable function of $q, p$. Then the Poisson brackets of $f, \mathcal{H}$ with respect to $q, p$ are given by

$$\{f, \mathcal{H}\}_{q,p} = \frac{\partial f}{\partial q} \frac{\partial \mathcal{H}}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial \mathcal{H}}{\partial q}$$

(3.14)

We can then write Hamilton’s equations with some function $f$ as:

$$\frac{df}{dt} = \{f, \mathcal{H}\}_{q,p} + \frac{\partial f}{\partial t}.$$  

(3.15)

Admittedly, doing this is rather unmotivated. Once we get to the quantum formalism of physics, we shall see that we can make a formal analogy to this equation by identifying the Poisson brackets to a commutation relation, and the function $f$ to some infinite-dimensional, self-adjoint and Hermitian operator.

3.3. Quantum Mechanics

In this section, we outline the relevant components of quantum mechanics. We begin with a rigorous treatment of Hilbert spaces, and determine a matrix representation of infinite-dimensional, self-adjoint and Hermitian operators. We then outline how this mathematical formalism is used in the study of quantum mechanics. Again, the purpose of this section is to familiarize the reader with the language of quantum mechanics that will be used in later sections. For reference, one may read [11].

3.3.1. Hilbert Spaces

A Hilbert space is a generalization of a vector space. Given two elements $x, y$ in some Hilbert space $H$ with inner product $\langle x|y \rangle$, we have the following axioms. Note that the following notation convention is set by physicists, and is backwards in axioms 2 and 4 from the convention set by mathematicians. If you are unfamiliar with either convention, this is probably a good thing. Bars denote complex conjugates.

Definition 28 (Axioms of Hilbert Spaces).

1. $\langle x|y \rangle = \overline{\langle y|x \rangle}$

2. $\langle ax_1 + bx_2|y \rangle = \overline{a}\langle x_1|y \rangle + \overline{b}\langle x_2|y \rangle$, where $a, b \in \mathbb{C}$.  


CHAPTER 3. A PRIMER ON PHYSICS

3. \langle x|x \rangle \geq 0.

4. \langle x|ay_1 + by_2 \rangle = a\langle x|y_1 \rangle + b\langle x|y_2 \rangle, \text{ where } a, b \in \mathbb{C}.

5. \( H \) is separable.

6. \{\ket{\psi_n}\} is complete, so we can write

\[ ||\psi||^2 = \sum_{n=1}^{\infty} |\langle \phi_n | \psi \rangle|^2. \]  

(3.16)

From Axiom 5, we have that there exists a countable set \( A \) of vectors dense in \( H \). For example, let

\[ A = \{\ket{\psi_1} + \ket{\psi_2} + \cdots + \ket{\psi_k} + \ldots\}. \]

Take out every vector \( \ket{\psi_k} \), which can be a linear combination of \( \ket{\psi_1}, \ldots, \ket{\psi_{k-1}} \). Then, we obtain an infinite set of linearly independent vectors

\[ B = \{\ket{\phi_1} + \ket{\phi_2} + \cdots + \ket{\phi_n} + \ldots\} \subset A. \]  

(3.17)

By construction, \( B \) is dense and can be made orthogonal by the Gram-Schmidt orthogonalization process such that

\[ \langle \phi_n | \phi_m \rangle = \delta^m_n = \begin{cases} 1 & m = n \\ 0 & \text{else} \end{cases} \]

Thus, for all \( \ket{\psi_i} \) in \( H \), we have:

\[ \ket{\psi_i} = \sum_{n=1}^{\infty} \ket{\phi_n} a_n. \]  

(3.18)

Thus, we have

\[ \langle \phi_m | \psi \rangle = \sum_{n=1}^{\infty} \langle \phi_m | \phi_n \rangle a_n = \sum_{n=1}^{\infty} \delta^m_n a_n = a_m, \]

giving the result

\[ \ket{\psi} = \sum_{n=1}^{\infty} \ket{\phi_n} \langle \phi_n | \psi \rangle. \]  

(3.19)
From Axiom 6, it follows that

$$\langle \psi | \psi \rangle = \langle \psi | 1 | \psi \rangle = \sum_{n=1}^{\infty} \langle \psi | \phi_n \rangle \langle \phi_n | \psi \rangle.$$  \hspace{1cm} (3.20)

We have thus found a diagonal representation of the identity in this Hilbert space:

$$1 = \sum_{n=1}^{\infty} |\phi_n \rangle \langle \phi_n|$$  \hspace{1cm} (3.21)

Now, we can define the norm and metric.

**Definition 29 (Norm).** The norm of $x$ is given by

$$||x|| = \sqrt{\langle x | x \rangle}.$$  \hspace{1cm} (3.22)

**Definition 30 (Metric).** The metric is given by

$$d(x, y) = ||x - y|| = \sqrt{\langle x - y | x - y \rangle}.$$  \hspace{1cm} (3.23)

### 3.3.2. Operators in Quantum Mechanics

An operator is a mapping between vector spaces. In this thesis, we will be considering operators as mappings between infinite-dimensional Hilbert spaces. First, some elementary definitions.

**Definition 31 (Transpose Operation).** Let $A$ be an $n \times m$ matrix. $A^t$, called $A$-transpose is defined such that

$$A^t_{ij} = A_{ji}.$$  \hspace{1cm} (3.24)

This definition can be readily extended to infinite dimensions, however not all infinite-dimensional matrices have well-defined transposes. For the purposes of quantum mechanics, one always works with matrices that do have well-defined transposes.

**Definition 32 (Hermitivity).** An operator $A$ is Hermitian if $A^\dagger = A$, where the ‘$\dagger$’ symbol represents the conjugate transpose operation. In other words,

$$\langle A \psi | \phi \rangle = \langle \psi | A | \phi \rangle$$
Definition 33 (Self-Adjointivity). An operator $A$ is self-adjoint if:

$$\langle A^\dagger | \psi \rangle = \langle \psi | A | \phi \rangle. \quad (3.25)$$

For an infinite-dimensional Hilbert space, we have that if $A$ is Hermitian and self-adjoint, $(A^\dagger)^\dagger = A$. In quantum mechanics, all physical observables (energy, position, momentum, etc.) are Hermitian and self-adjoint operators in some infinite-dimensional Hilbert space. We now define some useful quantities.

Definition 34 (Expectation Value). For some vector $|\psi\rangle$ with norm $||\psi|| = 1$, the expectation value of an operator $A$ with respect to $|\psi\rangle$ is given by

$$\langle A \rangle_{|\psi\rangle} = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle} = \langle \psi | A | \psi \rangle. \quad (3.26)$$

As this is a measurable quantity by an experimenter, we require this quantity to be real, hence the restriction that $A$ be Hermitian and self-adjoint.

Definition 35 (Eigenvalue of an Operator). If

$$A |\psi\rangle = a |\psi\rangle, \quad (3.27)$$

where $a \in \mathbb{C}$, then $a$ is called the eigenvalue of $A$. For a given vector $|\psi\rangle$, if $a$ exists, $a$ is unique.

3.3.3. Matrix Representation of an Operator

Suppose we have an operator acting on a vector to produce another vector: $A |\psi\rangle = |\psi'\rangle$. Then,

$$A (\mathbb{1}) |\psi\rangle = |\psi'\rangle = A \left( \sum_{n=1}^{\infty} |\phi_n\rangle \langle \phi_n | \right) |\psi\rangle.$$

Multiplying both sides on the left by $\langle \phi_m |$ gives:

$$\langle \phi_m | A \left( \sum_{n=1}^{\infty} |\phi_n\rangle \langle \phi_n | \right) |\psi\rangle = \sum_{n=1}^{\infty} \langle \phi_m | A |\phi_n\rangle \langle \phi_n | \psi \rangle = \langle \phi_m | \psi' \rangle.$$
Recall that the expansion of a vector $|\psi\rangle$ is given by:

$$|\psi\rangle = \sum_{n=1}^{\infty} |\phi_n\rangle \langle \phi_n| \psi\rangle.$$  \hspace{1cm} (3.28)

Hence, the expansion coefficients $a_n$ in Equation 3.18 are given by:

$$\langle \phi_m|\psi'\rangle$$ \hspace{1cm} (3.29)

with respect to the orthonormal basis $\{\phi_m\}$. Therefore, the matrix elements of an operator representation of $A$ are:

$$A_{mn} = \langle \phi_m|A|\phi_n\rangle.$$ \hspace{1cm} (3.30)

Further, we can write a vector as:

$$|\psi\rangle \rightarrow \begin{pmatrix} \langle \phi_1|\psi\rangle \\ \langle \phi_2|\psi\rangle \\ \vdots \\ \langle \phi_n|\psi\rangle \\ \vdots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$ \hspace{1cm} (3.31)

3.3.4. Fundamental Equations

We have stated that in quantum mechanics, observables such as position, momentum and energy are given by Hermitian, self-adjoint operators. Further, the quantum state of a particle is given by some vector $|\psi\rangle$. The time-dependence on both of these quantities are given by differential equations.

**Definition 36** (The Schrödinger Equation). Let $H$ be the Hermitian and self-adjoint operator in an infinite-dimensional Hilbert space corresponding to the classical Hamiltonian, and let $|\psi\rangle$ be some vector in this space. The time dependence of $|\psi\rangle$ is given by the Schrodinger equation:

$$H|\psi(t)\rangle = i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle$$ \hspace{1cm} (3.32)

where $i^2 = -1$ and $\hbar$ is a numerical constant with units of [energy][time].

**Definition 37** (Operator Time Dependence). Let $A$ be a Hermitian and self-adjoint operator
in an infinite-dimensional Hilbert space, and let $H$ be the Hamiltonian operator as before. The time dependence of $A$ is given by

$$\frac{d}{dt} A = \frac{i}{\hbar} [H, A] + \frac{\partial A}{\partial t}$$

(3.33)

where $[H, A]$ is the commutation relation $[H, A] = HA - AH$.

Note that this equation is directly analogous to the classical equation for the time dependence of a continuous function $f$ as in Equation 3.15, if we identify the Poisson bracket to the commutator, and $f$ to the operator $A$. 
Chapter 4

THE HILBERT-PÓLYA CONJECTURE

4.1. Introductory Notes

The Hilbert-Pólya Conjecture offers a promising methodology in regards to proving the Riemann Hypothesis. The fields of linear operators, matrices and linear algebra are arguably more well-understood than the interface between complex analysis and analytic number theory. For one, the calculations themselves are somewhat simpler. Apart from this, the conjecture also points to connections between prime number theory, and the study of quantum mechanics.

4.2. Random Matrix Theory

Random matrices were first studied by Wigner and later by Dyson as a model for the energy levels of atomic nuclei. There are three main types of random matrices, namely Gaussian orthogonal, unitary and symplectic. We will be interested mainly in the Gaussian unitary matrices. All of the following can be found in Chapter 6 of [12].
4.2.1. Gaussian Unitary Matrices

A random Gaussian unitary matrix can be described by the joint probability distribution of its eigenvalues. This is given for an $N \times N$ matrix as

$$P_N(x_1, \ldots, x_N) = C_N \exp \left( -\sum_{j=1}^{N} x_j^2 \right) \prod_{j<k} |x_j - x_k|^2,$$  \hspace{1cm} (4.1)

where $C_N$ is a normalization constant given by

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P_N(x_1, \ldots, x_N) dx_1 \ldots dx_N = 1$$  \hspace{1cm} (4.2)

which yields

$$(C_N)^{-1} = (2\pi)^{\frac{N}{2}} \beta^{\frac{N}{2}} \frac{\beta N(N-1)}{4} (\Gamma(1 + \frac{\beta}{2}))^{-N} \prod_{j=1}^{N} \Gamma(1 + \frac{\beta j}{2})$$  \hspace{1cm} (4.3)

4.2.2. Pair Correlation Function for Gaussian Unitary Matrices

A pair correlation function of order $n$ (otherwise known as an $n$-point correlation function) gives the probability of finding an eigenvalue of a random matrix around points $x_1, x_2, \ldots, x_n$ given no information about the location of any other eigenvalues. The function is defined as follows.

**Definition 38 (The $n$-point Correlation Function).** Let $A$ be an $N \times N$ matrix. The $n$-point correlation function is given by

$$R_n(x_1, x_2, \ldots, x_n) = \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P_N(x_1, x_2, \ldots, x_N) dx_{n+1} \ldots dx_N.$$  \hspace{1cm} (4.4)

Note that the integrals are taken over all $n_i$ in between $n$ and $N$.

It is often more apt to work with the $n$-point cluster function $T_n$, defined in terms of the $n$-point correlation function.

**Definition 39 ($n$-point Cluster Function).** Let $A$ be an $N \times N$ matrix. The $n$-point cluster
function is defined as

\[ T_n(x_1, \ldots, x_n) = \sum_{G} (-1)^{n-m} (m-1)! \prod_{j=1}^{M} R_{G_j}(x_k, k \in G_j) \]  

(4.5)

where \( G \) is any division of the integer indices \((1, 2, 3, \ldots, n)\) into \( m \) partitions \((G_1, G_2, \ldots, G_m)\). For example,

\[
\begin{align*}
T_1(x_1) &= R_1(x) \\
T_2(x_1, x_2) &= -R_2(x_1, x_2) + R_1(x_1)R_2(x_2) \\
T_3(x_1, x_2, x_3) &= R_3(x_1, x_2, x_3) - R_1(x_1)R_2(x_2, x_3) - \cdots + 2R_1(x_1)R_1(x_2)R_1(x_3)
\end{align*}
\]  

(4.6)

and so on.

We can invert Equation 4.5 to express the correlation function in terms of the cluster function:

\[ R_n(x_1, \ldots, x_n) = \sum_{G} (-1)^{n-m} \prod_{j=1}^{m} T_{G_j}(x_k, k \in G). \]  

(4.7)

By extending \( N \to \infty \) and considering a ‘mean spacing’ \( \alpha \) between eigenvalues of a random matrix, we can define the mean-spaced cluster function.

**Definition 40** (Mean-Spaced \( n \)-point Cluster Function). Let \( N \) be defined as before, and let \( \alpha \) be the mean spacing between eigenvalues of an \( N \times N \) matrix. Further, let \( T_n \) be the \( n \)-point cluster function. The mean-spaced \( n \)-point cluster function is given as

\[ Y_n(y_1, y_2, \ldots, y_n) = \lim_{N \to \infty} \alpha^n T_n(x_1, \ldots, x_n), \]  

(4.8)

where \( y_i = \frac{x_i}{\alpha} \).

We now wish to find some of the values of these functions, in particular for the Gaussian unitary matrices. We will find that the values of these functions are eerily similar to that of the statistics corresponding to the distribution of the imaginary parts of the non-trivial Riemann Zeta zeros. First, let us define the oscillator functions.

**Definition 41** (Oscillator Function). For integer \( j \), the oscillator function is given by

\[ \phi_j(x) = (2^j j! \sqrt{\pi i})^{-1/2} \exp \left( \frac{-x^2}{2} \right) \left( -\frac{d}{dx} \right)^j \exp (-x^2). \]  

(4.9)
With some work, it is possible to show that for Gaussian unitary matrices, we can write the mean-spaced 2-point cluster function in term of these oscillator functions as

\[
Y_2(x_1, x_2) = \lim_{N \to \infty} \left( \frac{\pi}{(2N)^{1/2}} \sum_{j=0}^{N-1} \phi_j(x_1)\phi_j(x_2) \right)^2 = \left( \frac{\sin(\pi r)}{\pi r} \right)^2, \tag{4.10}
\]

and the 2-point correlation function as

\[
R_2(x_1, x_2) = 1 - \left( \frac{\sin(\pi r)}{\pi r} \right)^2. \tag{4.11}
\]

The pair correlation function is plotted in Figure 4.1.

4.3. Zeroes of \( \zeta(s) \) and Random Matrices

In this section, we wish to describe the pair correlation of the imaginary parts of the non-trivial Riemann zeros, in particular show some evidence that it is the same function as Equation 4.11. This is the conjecture put forth by Montgomery in his seminal paper. We then consider a formal analogy between the spectrum (eigenvalues of the Hamiltonian) of
a chaotic quantum system considered by Berry in [3] and the Riemann zeta zero counting function given by Von Mangoldt and Backlund earlier in the thesis. First, we begin by considering the distribution of Riemann zeros. The number of zeros in the critical strip with imaginary part less than or equal to $E$ is given by the ‘Riemann staircase’.

**Definition 42** (Riemann Staircase).

$$N_R(E) = \sum_{j=1}^{\infty} \Theta(E - E_j),$$  \hspace{1cm} (4.12)

where $\Theta$ is the Heaviside step function, and not the $\Theta$ function related to $\Gamma(s)$.

As in Chapter 2, we see that there are estimates and known asymptotic behaviour for this function. We resummarize a familiar theorem:

**Theorem 43** (Backlund’s Estimate for the Riemann Staircase.). Let $\zeta(s)$ denote the Riemann Zeta function, and let $E$ denote the imaginary part of some complex number $z$ in the critical strip. Then, the number of zeros of $\zeta(s)$ with imaginary part less than or equal to $E$ is given by Equation 4.12. Furthermore, an approximation to this expression is given by

$$\langle N_R(E) \rangle = \frac{E}{2\pi} \left( \log \left[ \frac{E}{2\pi} \right] - 1 \right) + \frac{7}{8},$$  \hspace{1cm} (4.13)

following the notation of [3].

This is a slight modification to Theorem 18 to chapter 2. The spacing between eigenvalues of random Gaussian unitary matrices have a similar form as the probability distribution $P(S)$ of the normalized spacing between adjacent zeros $S_j$:

$$S_j = \frac{(E_{j+1} - E_j)}{\langle d_r \left( \frac{E_j + E_{j+1}}{2} \right) \rangle}$$  \hspace{1cm} (4.14)

where $\langle D_r(E) \rangle$ is the average density of zeros:

$$\langle D_r(E) \rangle = \frac{d}{dE} \langle N_R(E) \rangle = \frac{1}{2\pi} \log \left( \frac{E}{2\pi} \right).$$  \hspace{1cm} (4.15)

The Fourier Transform of the pair correlation function of the Riemann zeros is given by
CHAPTER 4. THE HILBERT-PÓLYA CONJECTURE

Equation 9 in [3]:

\[
K(\tau) = \lim_{M \to \infty} \left\{ \frac{1}{M} \sum_{j=1}^{M} \sum_{k=1}^{M} \exp \left[ 2\pi i \tau (x_j - x_k) \right] - \frac{\sin M\pi \tau}{\pi \tau} \right\} \tag{4.16}
\]

where \( \{x_j\} \) are the Riemann zeros scaled to have unit mean spacing. Montgomery’s conjecture in [13] can be restated as follows:

**Conjecture 44** (Montgomery’s Conjecture on the Fourier Transform of the Pair Correlation Function). Let \( K(\tau) \) be the Fourier transform of the pair correlation function as defined in Equation 4.16. Further, note that the Fourier transform of Equation 4.11 is given by

\[
K_{\text{GUE}}(\tau) = \begin{cases} 
|\tau| & \text{for } |\tau| < 1 \\
1 & \text{for } |\tau| > 1.
\end{cases} \tag{4.17}
\]

In [13], it is shown that \( K(\tau) = K_{\text{GUE}}(\tau) \) for \( |\tau| < 1 \). The Conjecture is that Equation 4.17 holds for \( |\tau| > 1 \).

We now look at a chaotic quantum system described by Berry in [3]. This system follows three properties:

1. The Hamiltonian \( H \) has a classical limit. This means that the quantum-mechanical operator equations have a direct classical representation under Hamilton’s equations.
2. The classical orbits given by Hamilton’s equations are chaotic (unstable)
3. The classical orbits are not time-reversible (they are not symmetric under a coordinate change \( t \to -t \))

Much of the following exposition will be schematic. Readers are forwarded to read [3] whenever there is ambiguity. First, we consider the analogue of Backlund’s zero counting function \( N \) for the real part \( E \) of the eigenvalues of the Hamiltonian of this quantum system. We work with a general Hamiltonian. We can separate \( N(E) \) into an ‘average’ part and a ‘fluctuating’ (or oscillating) part:

\[
N(E) = \langle N(E) \rangle + N_{\text{osc}}(E). \tag{4.18}
\]
Then we get Equation 12 in [3]:

\[ N_{\text{osc}}(E) = N(E) - \langle N(E) \rangle = 3 \left[ \sum_p \sum_{m=L}^{\infty} B_{pm} \exp \left( \frac{iS_{pm}(E)}{\hbar} \right) \right] . \hspace{1cm} (4.19) \]

The double sum is over orbits labelled with prime numbers \( p \), and integers \( m \) traversals of that orbit. \( S_{pm} \) is the action given by:

\[ S_{pm}(E) = \oint p \mu dq_{\mu}. \hspace{1cm} (4.20) \]

This action describes some chaotic trajectories obtained by Hamilton’s equations. Equation 14 in [3] gives the amplitude coefficient

\[ B_{pm} = 2\pi m \sinh \left( \frac{m\lambda_p(E)}{2} \right). \hspace{1cm} (4.21) \]

Then,

\[ N_{\text{osc}} = \frac{1}{2\pi} \sum_p \sum_{m=1}^{\infty} \frac{\sinh \{ mS_{p1} \}}{m \sinh \{ m\lambda_p(E) \}}. \hspace{1cm} (4.22) \]

After some work, we can show this is equal to

\[ N_{\text{osc}}(E) \approx \frac{1}{\pi} \sum_p \sum_{m=1}^{\infty} \frac{1}{m} \exp \left( - \frac{m\lambda_p(E)}{2} \right) \sin \left( \frac{mS_{p1}(E)}{\hbar} \right). \hspace{1cm} (4.23) \]

Unfortunately, the corresponding expression for \( \zeta(s) \) is given by Equation 18 in [3]:

\[ N_{R,\text{osc}}(E) \approx -\frac{1}{\pi} \sum_p \sum_{m=1}^{\infty} \frac{\sin(mE \log p)}{mp^2}, \hspace{1cm} (4.24) \]

if we identify the actions of the closed orbits

\[ S_{pm} = mR \log p \]

and the instability exponent

\[ \lambda_p = \log p. \]
Implicitly, what we are suggesting is that the closed orbits are identified to prime numbers $p$. Unfortunately, now this equation equals Equation 4.22, but differs by a sign. We can strengthen this analogy by considering the product

$$
P(E) = \prod_{k=0}^{\infty} \zeta\left(\frac{1}{2} + k + iE\right). \tag{4.25}
$$

**Theorem 45 (Convergence of $P(E)$).**

Let $P(E)$ be defined as in Equation 4.25. Then,

1. $P(E)$ converges for real $E$.
2. $P(E)$ has the same zeros as $\zeta\left(\frac{1}{2} + iE\right)$.

Then, we get that the corresponding expression for $N_{P,osc}$ is given by a more immediately analogous equation:

$$
N_{P,osc} = -\frac{1}{\pi} \lim_{\eta \to 0} \Im \log P(E + i\eta) \approx -\frac{1}{2\pi} \sum_{p} \sum_{m=1}^{\infty} \frac{\sinh\{mE \log p\}}{m \sinh\{m \log p / 2\}}. \tag{4.26}
$$

We can directly compare factors in the hyperbolic sines between this equation and with Equation 4.22. This concludes the schematic introduction to the formal analogy between the statistics of the eigenvalues of the chaotic quantum system studied by Berry, and the statistics of the distribution of the imaginary parts of the Riemann zeros. The oscillating parts of the respective functions differ by a minus sign, which apart from lack of rigour, is a weakness in this approach. A slightly more rigorous discussion of this analogy is given by Berry and Keating in [4] and [5].

4.4. **The Berry-Keating Operator**

Montgomery’s conjecture tantalizingly points a finger and says ‘hey, look over there’ to the idea that the Riemann zeros are somehow connected to the eigenvalues of a self-adjoint unitary operator. We recall the Hilbert-Pólya conjecture:

**Conjecture 46 (The Hilbert-Pólya Conjecture).** There exists a self-adjoint operator in an infinite-dimensional Hilbert space whose eigenvalues are the non-trivial Riemann zeros.
Berry and Keating have speculated that a possible operator that may satisfy the Hilbert-Pólya conjecture is given by the Hamiltonian

$$H = \frac{1}{2}(XP + PX) = -i \left( X \frac{d}{dx} + \frac{1}{2} \right). \quad (4.27)$$

The eigenvectors of this Hamiltonian with eigenvalues $E$ corresponding to functions $\psi(X)$ are given from

$$H\psi(X) = E\psi(X) \quad (4.28)$$

as

$$\psi(X) = \frac{A}{X^{\frac{1}{2} - iE}}. \quad (4.29)$$

Using a Fourier transform, we can compute the eigenfunction in momentum space:

$$\phi(P) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dX \psi(X) e^{-iPX}, \quad (4.30)$$

giving

$$\phi(P) = \frac{A}{|P|^{\frac{1}{2} + iE}} 2^{iE} \frac{\Gamma\left(\frac{1}{4} + \frac{1}{2}iE\right)}{\frac{1}{4} - \frac{1}{2}iE} = \frac{A}{\sqrt{2\pi}|P/(2\pi)|^{\frac{1}{2} + iE}} e^{2\theta(E)} \quad (4.31)$$

where $\theta(E)$ is defined as in Equation 4.32:

$$\Theta(E) = \arg \Gamma\left(\frac{1}{4} + \frac{E}{2}\right) - \frac{E}{2} \log \pi \approx \frac{E}{2} \log \frac{E}{2\pi} - \frac{E}{2} - \frac{\pi}{8}. \quad (4.32)$$

Now, consider the following coordinate tranformation:

$$X \rightarrow X_1 = \frac{2\pi}{P}, \quad P \rightarrow P_1 = \frac{XP^2}{2\pi}. \quad (4.33)$$

We can then rewrite the transformed wavefunction $\psi_1$ in terms of the untransformed momentum wavefunction $\phi$:

$$\psi_1(X_1) = \frac{(2\pi)^{1/4}}{|X_1|} \phi \left( \frac{\sqrt{2\pi}}{|X_1|} \right). \quad (4.34)$$

Berry and Keating in [5] use this, Equation 4.30 and the functional equation for $\zeta(s)$ to write
the identity (Equation 6.12 in their paper)

\[ X^{1/2} \zeta \left( \frac{1}{2} - iE \right) \psi_E(X) - P^{1/2} \zeta \left( \frac{1}{2} + iE \right) \phi_E(P) = 0. \] (4.35)

Unfortunately, this approach fails at this last stage: the term on the right has a plus inside the zeta function. If this were a minus sign, then this would be precisely the condition relating the Riemann zeros to the eigenvalues of the Hamiltonian in Equation 4.27.
Bibliography


Appendices
Appendix A

Appendix

Code for Figure 1 (Mathematica). Reproduced from [7]

Warning: This requires a decent (at time of writing, ignore this if Moore’s law has taken over) computer with lots of RAM and processing power (takes about 20 seconds with quad-core Intel® i7® and 8GB of RAM). If reading from future, do not have lulz over this.

Parametric

PlotPoints -> 200, MaxRecursion -> 0,
Mesh -> 65, ColorFunctionScaling -> False,
ColorFunction -> (Hue[Rescale[Arg[Zeta[# + I #2]],
{-Pi, Pi}, {0, 1} + 0.5], 1,
Rescale[Log[Abs[Zeta[# + I #2]]],
{-Infinity, Infinity}, {0.1, 2}]] &),
MeshFunctions -> {Log[
Abs[Zeta[#1 +
I #2]]] &}, Axes -> False]