# ON THE NUMBER OF INTEGERS EXPRESSIBLE AS THE SUM OF TWO SQUARES

by

Robert Richardson

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Mathematics

Program

YOUNGSTOWN STATE UNIVERSITY

December, 2009

On the Number of Integers Expressible as the Sum of Two Squares

Robert Richardson

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

*Robert Richardson*, Student Date

Approvals:

*Dr. Jacek Fabrykowski*, Thesis Advisor Date

*Dr. Eric Wingler*, Committee Member Date

*Dr. Thomas Smotzer*, Committee Member Date

Peter J. Kasvinsky, Dean of School of Graduate Studies and Research Date

## Abstract

We attempt to provide a reasonably complete work concerning the estimation of the number of integers expressible as the sum of two squares. We begin with some basic concepts from number theory, and progress rapidly through the theory necessary for Landau's theorem before presenting two proofs of his theorem.

# Contents

# Chapter 1

# Introduction

The goal of this paper is to collect in one location the results concerning the estimation of the number of integers less than a given number that are expressible as the sum of two squares. This has been a favorite pursuit of numerous mathematicians throughout the ages, as we will soon see, and it also makes use of some of the most powerful results available in number theory.

This paper is intended to be a complete reference, i.e. one can read this thesis and not need to refer to any other works, but further examination of the source material may be interesting to some, and so numerous references are given throughout the text. When more than one source possesses the same material (as has often been the case), the author has attempted to pick the “best” material to recommend, with comments about completeness, ease of reading, and other factors that led to the selection.

For readers very comfortable with their number theory, Chapter 2 may be omitted. Although it contains many advanced results that are not covered in the first few courses of number theory. Chapter 2 is intended to be used for quick and easy reference, and as such, proofs are omitted, until the section on the Prime Number Theorem. The analytic proof of the Prime Number Theorem has many parallels and analogs to the main proof of the paper and so is presented in its entirety. The interested reader may find every non-trivial excluded proof in Appendix A. Chapter 3 contains information on the geometrical representation of the problem, but unfortunately leads us down a few dead-ends and does not set us on a path towards a solution. For readers interested in brevity, Chapter 4 and Chapter 5 contain the meat and potatoes of this thesis and of this problem. Chapter 4 is concerned entirely with providing us with the necessary information about the types of numbers expressible as the sum of two squares so that in Chapter 5 we can progress to proving the result utilizing two separate approaches–one relying on the Generalized Wiener-Ikehara Theorem, and the other relying on the same methods of complex analysis that have proven both the Prime Number

Theorem and Dirichlet's theorem.

It has been said by G. H. Hardy in [HarRam] that "Almost every arithmetician of note since Fermat has contributed to the solution of the problem, and it has its puzzles for us still."

In fact, the pedigree of this problem has its roots in Euclid's proof of the infinitude of prime numbers and has progressed through the centuries capturing the imagination of mathematicians such as Fermat, Euler, Gauss, Landau, Hardy and Ramanujan (amongst others).

## 1.1 Notation and Conventions

Throughout this paper, $p$ and $q$ are reserved for primes. Following Riemann's notation, $s = \sigma + it$ is a complex variable with $\sigma$ and $t$ real. Also, $\log x$ refers to the natural logarithm of $x$, not the common logarithm. Finally, we use the notation $(a, b)$ to indicate the greatest common divisor of $a$ and $b$.

# Chapter 2

# Preliminaries

In this chapter, we attempt to cover the material necessary that readers may not have had a chance to see yet (or at least may not have seen for quite a while). We begin with a simple combinatorial result, and the beginnings of number theory before progressing through some of the ground-breaking results involved in proving the Prime Number Theorem utilizing the $\zeta$ and $L$ functions.

Most of the results and theorems in this chapter are presented without proof, the exception being those proofs whose inclusion is as important as the result of the theorem itself. For further reading, and information that the author did not feel was integral to the main aim of this thesis, any number of books on number theory may be consulted, including [LanHan], [HarWr], [Apos], [BatDia], [Burt], [Sier], [LeV1] and [LeV2] and [Kara].

For completeness, Edmund Landau's *Handbuch* [LanHan] is second to none, although, for readers that prefer a text written in English, G. H. Hardy and E. M. Wright's *Analytic Number Theory* [HarWr] is invaluable. For the purposes of this chapter, we have tried to use Tom Apostol's *Introduction to Analytic Number Theory* exclusively as it is both readily available, and an easy read. Once we begin leading up to the proof of the Prime Number Theorem, we switch to [LeV2], a text that has been integral in the main proof of the paper.

Finally, the omitted proofs may be found in Appendix A.

## 2.1 Cross-Classification Principle

We begin the preliminaries with a general combinatorial theorem. The *principle of cross-classification* will be used later in the heuristic argument of chapter 5, and it can be found in [LeV1]

**Theorem 1.** *Let $S$ be a set of $N$ distinct elements, and let $S_1, \dots, S_r$ be arbitrary subsets of $S$ containing $N_1, \dots, N_r$ elements, respectively. For $1 \leq i < j < \cdots < l \leq r$, let $S_{ij\cdots l}$ be the intersection of $S_i, S_j, \dots, S_l$; and let*

*$N_{ij\cdots l}$ be the number of elements of $S_{ij\cdots l}$. Then the number of elements of $S$ not in any of $S_1, \dots, S_r$ is*

$$K = N - \sum_{1\le i\le r} N_i + \sum_{1\le i<j\le r} N_{ij} - \sum_{1\le i<j<k\le r} N_{ijk} + \cdots \\ + (-1)^r N_{12\cdots r}.$$

**Proof.** *Let a certain element $s$ of $S$ belong to exactly $m$ of the sets $S_1, \dots, S_r$. If $m = 0$, $s$ is counted only once, in $N$ itself. If $0 < m \le r$, then $s$ is counted once, or $\binom{m}{0}$ times, in $N$, $\binom{m}{1}$ times in the terms $N_i$, $\binom{m}{2}$ times in the terms $N_{ij}$, etc. Thus, the total contribution to $K$ arising from the element $s$ is*

$$\binom{m}{0} - \binom{m}{1} + \binom{m}{2} - \cdots + (-1)^m \binom{m}{m} = (1-1)^m = 0.$$

## 2.2 Quadratic Residues

Quadratic residues commonly occur in number theory, and we will be using them to prove Dirichlet's Christmas Theorem. This theory can be found in nearly any book on number theory, but we have chosen [Burt] and [Sier] specifically as sources in this section.

**Definition 1.** *Let $p$ be an odd prime and $(a, p) = 1$. If the quadratic congruence $x^2 \equiv a \pmod p$ has a solution, then $a$ is said to be a* quadratic residue of $p$. *Otherwise, $a$ is called a* quadratic nonresidue of $p$.

**Theorem 2 (Euler's Criterion).** *Let $p$ be an odd prime and $(a, p) = 1$. Then $a$ is a quadratic residue of $p$ if and only if $a^{(p-1)/2} \equiv 1 \pmod p$.*

**Definition 2.** *Let $p$ be an odd prime and let $(a, p) = 1$. The* Legendre symbol $\left(\frac{a}{p}\right)$ *is defined by*

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue of } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue of } p. \end{cases}$$

Although this symbol bears no small resemblance to merely enclosing a fraction in parentheses (and all operations that parentheses entail), it should be clear by context what we intend.

**Theorem 3.** *If $p$ is a prime of the form $4k+1$ (where $k$ is a natural number), then*

$$p \,\Bigg|\, \left[\left(\frac{p-1}{2}\right)!\right]^2 + 1$$

Although this next theorem is not related to quadratic residues, it will be used along with the previous theorem to prove Dirichlet's Christmas Theorem. Both theorems can be found in [Sier].

**Theorem 4 (Thue Theorem).** *If $m$ is a natural number and $a$ an integer relatively prime to $m$, then there exist natural numbers $x$ and $y$ both less than $\sqrt{m}$ and such that the number $ax \pm y$ is divisible by $m$ for a suitable choice of the sign $\pm$.*

In other works, such as [BatDia], this theorem is known as Aubry-Thue, and phrased a bit differently. L. Aubry and A. Thue proved this theorem independently of each other, hence the slight confusion over the name. It is clear that both theorems as we shall use them are equivalent however.

**Theorem 5 (Aubry-Thue).** *Suppose $p$ is a prime and $a$ is an integer not divisible by $p$. Then there exist integers $x, y$ such that*

$$x \equiv ay \pmod{p},\ 0 < max\,(|x|, |y|) < p^{1/2}.$$

## 2.3 Dirichlet's Theorem for Primes of the Form $4k-1$ and $4k+1$

It is the case, as first proved by Dirichlet, that any arithmetic progression with first term $h$ and common difference $k$ where $(h,k) = 1$ will contain infinitely many primes. That is, if $(h,k) = 1$, then

$$kn + h,\ n = 0, 1, 2, \ldots$$

will contain infinitely many primes. This result is known as Dirichlet's theorem, but it is not necessary for the purpose of this paper to prove Dirichlet's theorem in full.

In fact, in this paper, we are only concerned with primes congruent to 1 and 3 modulo 4, and care not for general sequences. Thus, we can modify Euclid's proof that there are infinitely many primes to prove that the sequences $\{4k-1 : k \in \mathbb{N}\}$ and $\{4k+1 : k \in \mathbb{N}\}$ contain an infinitude of primes.

As the proofs are both simple and short, we include them in this chapter. Both can be found in [Apos].

**Theorem 6.** *There are infinitely many primes of the form $4k-1$.*

**Proof.** *Assume not. Let $p$ be the largest prime of the form $4k-1$, and consider*

$$N = 2^2 \cdot 3 \cdot 5 \cdots p - 1.$$

*Now, the product $3 \cdot 5 \cdots p$ contains all of the odd primes $\leq p$ as factors, and $N$ is of the form $4k-1$, so $N$ cannot be prime since $N > p$. However, there is no prime less than or equal to $p$ that divides $N$, so all prime factors of $N$ must exceed $p$. But all of the prime factors of $N$ cannot be of the form $4k+1$ since the product of two such numbers is of the same form, which is absurd. Therefore, some prime factor of $N$ must be of the form $4k-1$, contradicting our assumption that there is a maximum such prime.*

**Theorem 7.** *There are infinitely many primes of the form $4k+1$.*

**Proof.** *Let $N$ be any integer greater than $1$. We will show that there is a prime $p > N$ such that $p \equiv 1 \pmod 4$. Let*

$$m = (N!)^2 + 1.$$

*Note here that $m$ is odd and $m > 1$. Let $p$ be the smallest prime factor of $m$. None of the numbers $2, 3, \ldots, N$ divides $m$, so $p > N$. Furthermore, we have*

$$(N!)^2 \equiv -1 \pmod p.$$

*By raising both sides to the $(p-1)/2$ power, we obtain*

$$(N!)^{p-1} \equiv (-1)^{(p-1)/2} \pmod p,$$

*but then $(N!)^{p-1} \equiv 1 \pmod p$ by the Euler-Fermat theorem, thus*

$$(-1)^{(p-1)/2} \equiv 1 \pmod p.$$

*Now we have that the difference $(-1)^{(p-1)/2} - 1$ is either $0$ or $-2$, and it clearly cannot be $-2$, since it is divisible by $p$, so it must be $0$. That is,*

$$(-1)^{(p-1)/2} = 1.$$

*But then $(p-1)/2$ is even, and so $p \equiv 1 \pmod 4$. That is, for each integer $N > 1$, there exists a prime $p > N$ such that $p \equiv 1 \pmod 4$, and hence there are infinitely many primes of the form $4k+1$.*

Thankfully, for brevity's sake anyways, we did not need to prove Dirichlet's theorem, although with the machinery outlined in the next few chapters, this proves to be quite within our grasp. For the interested reader, a proof of Dirichlet's theorem using this machinery is presented at the end of Appendix A.

To summarize and bring our discussion back to the result we are truly concerned with, let $\pi(x; k, l)$ be the number of primes $p \equiv l \pmod k$ which do not exceed $x$. Then we have

$$\lim_{x \to \infty} \frac{\pi(x; 4, 1)}{\pi(x; 4, 3)} = 1.$$

In other words, there are asymptotically equally many primes in either sequence.

## 2.4 Functions

Here we attempt to summarize the needed mechanics of arithmetical functions, beginning with just what is an arithmetical function.

**Definition 3.** *A complex valued function defined on the domain of the integers is an* arithmetical function*; denoted* $f(s) \in \mathcal{A}$

We also have numerous properties that make our functions "nice".

**Definition 4.** *An arithmetical function* $f$ *is called* multiplicative *if* $f$ *is not identically zero and if*

$$f(mn) = f(m)f(n) \text{ whenever } (m, n) = 1.$$

Being multiplicative is all well and good, but some functions are even better: they are *completely multiplicative.*

**Definition 5.** *A multiplicative function* $f$ *is called* completely multiplicative *if*

$$f(mn) = f(m)f(n) \text{ for all } m, n.$$

For instance, the Euler totient function and the Möbius function are multiplicative.

**Definition 6.** *If* $n \geq 1$*, then the Euler totient* $\varphi(n)$ *is defined to be the number of positive integers not exceeding* $n$ *which are relatively prime to* $n$*; thus*

$$\varphi(n) = \sum_{\substack{k=1 \\ (k,n)=1}}^{n} 1.$$

The Möbius function, denoted $\mu(n)$ (and sometimes called Möbius mu) occurs commonly throughout number theory. It is defined as follows:

**Definition 7.** *Ifn* $> 1$*, write* $n = p_1^{a_1} \cdots p_k^{a_k}$*, then*
$\mu(1) = 1$

*(2)* $\mu(n) = (-1)^k$ *if* $a_1 = a_2 = \cdots = a_k = 1$

*(3)* $\mu(n) = 0$ *otherwise.*

Neither the Euler totient or Möbius mu functions are completely multiplicative, but the identity function $I(n)$ is.

**Definition 8.** *The arithmetical function $I$ defined by*

$$I(n) := \left\lfloor \frac{1}{n} \right\rfloor = \begin{cases} 1 \text{ if } n = 1, \\ 0 \text{ if } n > 1, \end{cases}$$

*is called the* identity function*.*

For $\varphi(n)$, we also have a product formula. This is common in number theory and this is merely the first of many equivalent formulas for the functions we will examine.

**Theorem 8.** *For $n \geq 1$ we have*

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Now we describe the operation of Dirichlet convolution, an operation defined on arithmetical functions and sometimes called the Dirichlet product of arithmetical functions

**Definition 9.** *If $f$ and $g$ are two arithmetical functions, we define their Dirichlet convolution (or Dirichlet product) as the arithmetical function $h$ defined by the equation*

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right),$$

*denoted $h = f * g$ or $h(n) = (f * g)(n)$.*

We can use Dirichlet convolution to do many things, one of the least being that:

**Theorem 9.** *For all $f$ we have $I * f = f * I = f$.*

Finally in this section, we will frequently have need to compare functions. One of the handiest notations we will use is $O$- and $o$- notation, read "big oh" and "little oh" respectively.

**Definition 10.** *Let $f, g$ be in $\mathcal{A}$, $g(x) > 0$ for all $x \geq a$ then $f = O(g)$ if and only if $f/g$ is bounded for all $x \geq a$ for some $a$. This is also denoted $f \ll g$ or $g \gg f$, and can also be expressed by saying there exists a constant $M > 0$ such that*

$$|f(x)| \leq Mg(x) \text{ for all } x \geq a.$$

*If we write*

$$f(x) = h(x) + O(g(x))$$

*, then we mean that $f(x) - h(x) = O(g(x))$. And note that $f(t) = O(g(t))$ for $t \geq a$ implies that*

$$\int_a^x f(t)dt = O\left(\int_a^x g(t)dt\right) \text{ for } x \geq a.$$

*We write, $f = o(g)$ as $x \to \infty$ if and only if $\lim_{x\to\infty} f(x)/g(x) = 0$. Again, an equation of the form $f(x) = h(x) + o(g(x))$ as $x \to \infty$ implies that $f(x) - h(x) = o(g(x))$ as $x \to \infty$.*

*Also note that $f(x) = O(1)$ implies that $f(x) = o(x)$ as $x \to \infty$.*

## 2.5 Dirichlet Series

Dirichlet series are among the most important tools in the mathematical toolbox of a number theorist. They are applicable in a broad range of areas, and both the Riemann zeta function and Dirichlet $L$-functions are examples of Dirichlet series.

**Definition 11.** *Let $f(n)$ be an arithmetical function, then*

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

*is called a* Dirichlet series, *denoted D.s., with coefficients $f(n)$.*

Note: $|n^s| = n^\sigma$ since $\left|e^{i\theta}\right| = 1$ for real $\theta$.

When $f(n)$ is an arithmetical function as described above, we will commonly refer to its associated Dirichlet series as $F(s)$ or if $F(n)$ is already taken, $\widehat{F}(s)$.

We have a number of results about Dirichlet series that will be applicable to Riemann zeta function and Dirichlet $L$-functions.

**Lemma 1.** *If $\sigma > a$, we have $|n^s| = n^\sigma \geq n^a$, therefore*

$$\left|\frac{f(n)}{n^s}\right| \leq \frac{|f(n)|}{n^a}.$$

**Proof.** *Clear.*

Therefore, if a Dirichlet series converges absolutely for $s = a + ib$, then by the comparison test it also converges absolutely for all $s$ with $\sigma \geq a$.

This leads us to the next theorem.

**Theorem 10.** *Suppose the series $\sum |f(n)n^{-s}|$ does not converge for all $s$ or diverge for all $s$. Then there exists a real number, $\sigma_a$, called the* abscissa of absolute convergence, *such that the series $\sum |f(n)n^{-s}|$ converges absolutely if $\sigma > \sigma_a$, but does not converge absolutely if $\sigma < \sigma_a$.*

Note 1: If a D.s. converges everywhere, $\sigma_a := -\infty$, and if the series converges nowhere, $\sigma_a := +\infty$.

Note 2: Analogous definitions and theorems pertaining to the *abscissa of convergence* and half-planes of convergence (instead of absolute convergence) exist. We shall assume the results of these theorems without mentioning them specifically. When we refer to absolute convergence we shall use the notation $\sigma_a$, and when referring to conditional convergence, we will use $\sigma_c$. Finally, it may interest the reader to know that for non-infinite $\sigma_a$ and $\sigma_c$, it is known that $0 \leq \sigma_a - \sigma_c \leq 1$.

### 2.5.1 The Function of a Dirichlet series

For this section, assume that $\sum f(n)n^{-s}$ converges absolutely for $\sigma > \sigma_a$ and let $F(s)$ denote the summation

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \text{ for } \sigma > \sigma_a \tag{2.1}$$

We have the following theorem

**Theorem 11 (Uniqueness theorem).** *Given two Dirichlet series*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \textit{ and } G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s},$$

*both absolutely convergent for $\sigma > \sigma_a$. If $F(s) = G(s)$ for each $s$ in an infinite sequence $\{s_k\}$ such that $\mathrm{Re}(s_k) = \sigma_k \to +\infty$ as $k \to \infty$, then $f(n) = g(n)$ for every $n$.*

Multiplication of Dirichlet series, maybe predictably, utilizes Dirichlet convolution.

**Theorem 12.** *Given two functions $F(s)$ and $G(s)$ represented by Dirichlet series,*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \textit{ for } \sigma > a,$$

*and*

$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s} \textit{ for } \sigma > b.$$

*Then in the half-plane where both series converge absolutely we have*

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s}, \tag{2.2}$$

*where $h = f * g$, the Dirichlet convolution of $f$ and $g$:*

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

*Conversely, if $F(s)G(s) = \sum \alpha(n) n^{-s}$ for all $s$ in a sequence $\{s_k\}$ with $\sigma_k \to +\infty$ as $k \to \infty$, then $\alpha = f * g$.*

### 2.5.2 Euler products

Sometimes called the analytic version of the fundamental theorem of arithmetic, this next theorem is from Euler in 1737.

**Theorem 13 (Analytic Fundamental Theorem of Arithmetic).** *Let $f$ be a multiplicative arithmetical function such that the series $\sum f(n)$ is absolutely convergent. Then the sum of the series can be expressed as an absolutely convergent infinite product,*

$$\sum_{n=1}^{\infty} f(n) = \prod_p \{1 + f(p) + f(p^2) + \cdots\} \tag{2.3}$$

*extended over all primes. If $f$ is completely multiplicative, the product simplifies and we have*

$$\sum_{n=1}^{\infty} f(n) = \prod_p \frac{1}{1 - f(p)}. \tag{2.4}$$

*Either version is referred to as the* Euler product *of the series.*

The next theorem is simply an application of the previous one to absolutely convergent Dirichlet series.

**Theorem 14.** *Assume $\sum f(n) n^{-s}$ converges absolutely for $\sigma > \sigma_a$. If $f$ is multiplicative, then we have*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left\{1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right\} \text{ if } \sigma > \sigma_a$$

*and if $f$ is completely multiplicative, then we have*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \frac{1}{1 - f(p) p^{-s}} \text{ if } \sigma > \sigma_a.$$

### 2.5.3 Analytic Properties of Dirichlet Series

We deduce the analytic properties of Dirichlet series from the following general theorem of complex function theory.

**Lemma 2.** *Let $\{f_n\}$ be a sequence of functions analytic on an open subset $S$ of the complex plane, and assume that $\{f_n\}$ converges uniformly on every compact subset of $S$ to a limit function $f$. Then $f$ is analytic on $S$ and the sequence of derivatives $\{f_n'\}$ converges uniformly on every compact subset of $S$ to the derivative $f'$.*

We wish to apply this lemma to Dirichlet series, but first we must show that we have uniform convergence on compact subsets of the half-plane of convergence. We require the following:

**Lemma 3.** *A Dirichlet series $\sum f(n)n^{-s}$ converges uniformly on every compact subset lying interior to the half-plane of convergence $\sigma > \sigma_c$.*

Now we have that every Dirichlet series is analytic in its half-plane of convergence, and we have its derivative.

**Theorem 15.** *The summatory function of a Dirichlet series, $F(s) = \sum f(n)n^{-s}$, is analytic in its half-plane of convergence $\sigma > \sigma_c$, and its derivative $F'(s)$ is represented in this half-plane of convergence by the Dirichlet series*

$$F'(s) = -\sum_{n=1}^{\infty} \frac{f(n)\log n}{n^s},$$

*which is obtained by differentiating term by term.*

Note: We have that $F'(s)$ has the same abscissa of convergence and the same abscissa of absolute convergence as the series for $F(s)$.

Our discussion on Dirichlet series is almost complete, but before we leave this subject, we will have need to talk about the singularity of a Dirichlet series so we can talk about the analytic continuation of the Riemann zeta function and of L-functions.

In fact, we find that the singularity of a Dirichlet series will occur at its abscissa of convergence on the real line thanks to this next theorem due to Landau.

**Theorem 16.** *Let $F(s)$ be represented in the half-plane $\sigma > c$ by the Dirichlet series*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

*where $c$ is finite, and assume that $f(n) \geq 0$ for all $n \geq n_0$. If $F(s)$ is analytic in some disk about the point $s = c$, then the Dirichlet series converges in the half-plane $\sigma > c - \varepsilon$ for some $\varepsilon > 0$. Therefore, if the Dirichlet series has a finite abscissa of convergence $\sigma_c$, then $F(s)$ has a singularity on the real axis at the point $s = \sigma_c$.*

The next two sections cover some more background before we can apply what we have learned about Dirichlet series to the Riemann zeta function and the Dirichlet $L$-functions.

## 2.6 Dirichlet Characters

Dirichlet characters will be absolutely necessary before continuing on to the Dirichlet $L$-functions, as they are part of the definition of any $L$-function.

But, before we get to Dirichlet characters, let us speak of general characters of a group, a definition from which Dirichlet characters will be defined, and with which we can prove some more general results that will serve us well when we are dealing with them.

**Definition 12.** *Let $G$ be an arbitrary group. A complex-valued function $f$ defined on $G$ is called a* character *of $G$, if $f$ has the multiplicative property, i.e.*

$$f(ab) = f(a)f(b)$$

*for all $a, b$ in $G$, and if $f(c) \neq 0$ for some $c$ in $G$.*

Now we will see that every character has a very special relationship with the roots of unity.

**Theorem 17.** *If $f$ is a character of a finite group $G$ with identity element $e$, then $f(e) = 1$, and each function value $f(a)$ is a root of unity. In fact, if $a^n = e$, then $[f(a)]^n = 1$.*

We always have the existence of at least one character, this is trivial, so we present this result as a definition.

**Definition 13.** *Every group $G$ has at least one character, i.e. the function which is identically $1$ on $G$. This is the* principal character *of $G$. That is*

$$f(g) \equiv 1, \textit{ for all } g \in G$$

*is the principal character of $G$.*

This is very nice, but if $G$ is "nice" as well, then we have the next result.

**Theorem 18.** *A finite abelian group $G$ of order $n$ has exactly $n$ distinct characters.*

In fact, for Dirichlet characters, the group that we wish to work with will be very "nice".

Recall that a reduced residue system modulo $k$ is a set of $\varphi(k)$ integers, $\{a_1, a_2, \ldots, a_{\varphi(k)}\}$ incongruent modulo $k$, each of which is relatively prime to $k$. For each integer $a$, the corresponding residue class $\hat{a}$ is the set of all integers congruent to $a$ modulo $k$:

$$\hat{a} = \{x : x \equiv a \pmod{k}\}.$$

We define multiplication of residue classes by the relation

$$\hat{a} \cdot \hat{b} = \widehat{ab}. \tag{2.5}$$

That is, the product of two residue classes $\hat{a}$ and $\hat{b}$ is the residue class of the product $ab$.

In fact, residue classes modulo a fixed positive integer $k$ form a group under this definition of multiplication.

**Theorem 19.** *With multiplication defined by* (2.5)*, the set of reduced residue classes modulo $k$ is a finite abelian group of order $\varphi(k)$. The identity is the residue class $\hat{1}$. The inverse of $\hat{a}$ is the residue class $\hat{b}$ where $ab \equiv 1 \pmod{k}$.*

Finally, we are ready to define Dirichlet characters.

**Definition 14.** *Let $G$ be the group of reduced residue classes modulo $k$. Corresponding to each character $f$ of $G$, we define an arithmetical function $\chi = \chi_f$ as follows:*

$$\begin{aligned} \chi(n) &= f(\hat{n}) \quad \text{if } (n,k) = 1, \\ \chi(n) &= 0 \quad\quad\ \text{if } (n,k) > 1. \end{aligned}$$

*The function $\chi$ is the* Dirichlet character modulo $k$*. The principal character $\chi_1$ has the properties*

$$\chi_1(n) = \begin{cases} 1 \text{ if } (n,k) = 1, \\ 0 \text{ if } (n,k) > 1. \end{cases}$$

In this paper, the only reduced residue system we will work with in-depth is the reduced residue system modulo 4. The Dirichlet characters for this system are summarized in the next table:

| $n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\chi_1(n)$ | 1 | 0 | 1 | 0 |
| $\chi_2(n)$ | 1 | 0 | $-1$ | 0 |

$\chi$ for $k = 4$

## 2.7 Properties of the Gamma function

Before continuing, we will require some of the properties of the gamma function $\Gamma(s)$. Although we will use only a handful of these, we provide this list for easy reference.

### Properties of $\Gamma(s)$

**(1)** For $\sigma > 0$, we have the integral representation

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x} dx. \tag{2.6}$$

**(2)** The function defined for $\sigma > 0$ can be continued beyond $\sigma = 0$ and so $\Gamma(s)$ exists as a function that is analytic everywhere except for simple poles at the points

$$s = 0, -1, -2, -3, \ldots$$

with residue $(-1)^n/(n!)$ at $s = -n$.

**(3)** We also have the following representation for gamma

$$\Gamma(s) = \lim_{n\to\infty} \frac{n^s n!}{s(s+1)\cdots(s+n)} \text{ for } s \neq 0, -1, -2, \ldots,$$

**(4)** The product formula for gamma

$$\frac{1}{\Gamma(s)} = se^{\gamma s} \prod_{n=1}^{\infty} \left(1 + \frac{s}{n}\right) e^{-s/n} \text{ for all } s,$$

where $\gamma$ is Euler's constant. Since the product converges for all $s$, $\Gamma(s)$ is never zero.

**(5)** The gamma function satisfies two functional equations,

$$\Gamma(s+1) = s\Gamma(s) \tag{2.7}$$

and

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}, \tag{2.8}$$

for all $s$.

**(6)** Perhaps the most well-known property is the last we mention

$$\Gamma(n+1) = n!, \text{ for } n \text{ a nonnegative integer.}$$

## 2.8 Generalized Wiener-Ikehara Theorem

The generalized Wiener-Ikehara theorem will be the workhorse that we use for the first proof of the main theorem of this paper. It is a Tauberian kind of theorem. This is the method that [BatDia] uses to prove their estimate of $B(x)$, and much of the material in this section can be found there. In order to talk about the generalized version of the Wiener-Ikehara theorem, we first speak about the original, but in order to do that, we must define Fejér kernels.

**Definition 15.** *For each positive real number $\lambda$, a* Fejér kernel *on $\mathbb{R}$, denoted $K_\lambda$, is defined by*

$$K_\lambda(x) := \frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{ixt}dt.$$

We normalize this function and use it as an approximate identity, while $1-|t|/(2\lambda)$ will help us improve the convergence of a specific integral. We are now in the realm of Fourier analysis, and require two lemmas from that field. The first will describe some properties of the Fejér kernel.

**Lemma 4.** *Let $x$ be a real number, $\lambda > 0$, and $\delta > 0$. Then*

$$K_\lambda(x) = \lambda\left(\frac{\sin\lambda x}{\lambda x}\right)^2, \tag{2.9}$$

$$0 \le K_\lambda(x) \le \min(\lambda, \frac{1}{\lambda x^2}), \tag{2.10}$$

$$\int_{-\infty}^{\infty} K_\lambda(u)du \text{ exists and is independent of } \lambda, \tag{2.11}$$

$$\int_{|u|>\delta} K_\lambda(u)du \le \frac{2}{\lambda\delta}. \tag{2.12}$$

**Proof.** *$K_\lambda(0) = \lambda$ by inspection. For $x \neq 0$, we have that*

$$K_\lambda(x) = \int_0^{2\lambda}\left(1-\frac{t}{2\lambda}\right)\cos(xt)dt.$$

*We integrate by parts to get*

$$K_\lambda(x) = \frac{1}{2\lambda x}\int_0^{2\lambda}\sin(xt)dt = \frac{1-\cos 2\lambda x}{2\lambda x^2} = \frac{(\sin\lambda x)^2}{\lambda x^2},$$

*which proves* (2.9).

*Now for* (2.10)*: we have the inequalities*

$$0 \le u^{-2}\sin^2 u \le \min(1, u^{-2}),$$

*from which* (2.10) *follows.*

*The integral of* (2.11) *converges by* (2.10). *By letting* $\lambda x = v$, *we see that*

$$\int_{-\infty}^{\infty} K_\lambda(x)dx = \int_{-\infty}^{\infty} \frac{\sin^2 v}{v^2} dv$$

*for any* $\lambda$. *In fact this integral is equal to* $\pi$, *but we are not concerned with this value here.*

*At last, we have* (2.12) *from the estimate* $K_\lambda(x) \leq \lambda^{-1}x^{-1}$.

The second result we require from Fourier analysis is a special case of the Riemann-Lebesgue lemma, which tells us that Fourier transforms vanish at infinity.

**Lemma 5.** *Let* $f$ *be a continuous complex valued function on* $\mathbb{R}$ *which is zero except on a bounded set, and let* $y$ *be a real number. Then*

$$\lim_{y \to \pm\infty} \int_{-\infty}^{\infty} f(t)e^{ity}dt = 0$$

**Proof.** *Let* $I(y) = \int f(t)e^{ity}dt$. *Now by changing the variable, we have*

$$I(y) = \int_{-\infty}^{\infty} f\left(t + \frac{\pi}{y}\right) e^{i(t+(\pi/y))y}dt = -\int_{-\infty}^{\infty} f\left(t + \frac{\pi}{y}\right) e^{ity}dt.$$

*Therefore,*

$$2I(y) = \int_{-\infty}^{\infty} \left\{ f(t) - f\left(t + \frac{\pi}{y}\right) \right\} e^{ity}dt.$$

*Now since* $f$ *is uniformly continuous, the expression in braces tends to zero uniformly as* $y \to \pm\infty$. *Furthermore, the last integrand vanishes outside a fixed bounded set, say if* $|y| \geq 1$. *Hence,* $I(y) \to 0$ *as* $|y| \to \infty$.

As we work towards an estimate of $B(x)$, we will be using functions of a particular class; one of the properties of functions in this class is that they are locally of bounded variation; that is,

**Definition 16.** *We say that* $f$ *is* locally of bounded variation *on* $(a, \infty)$ *if and only if the variation of* $f$ *on each compact subinterval* $[b, c] \subset (a, \infty)$ *is finite.*

**Definition 17.** *Let* $\mathcal{V}$ *denote the class of complex valued functions on* $\mathbb{R}$ *that possess the following properties:*

*They are zero in* $(-\infty, 1)$,

*continuous from the right,*

*and locally of bounded variation.*

Now we can prove the Wiener-Ikehara theorem. In fact, the theorem holds true for Mellin transforms, a more general type of Dirichlet series, but we shall not require this fact.

**Theorem 20.** *Let $F$ be a real valued monotone nondecreasing function in $\mathcal{V}$. Let $\sigma_c(\widehat{F}) = \alpha > 0$ and suppose that there exist a real number $L$ and a function $\phi$, continuous on the closed half plane $\sigma \geq \alpha$, such that*

$$\widehat{F}(s) = \int x^{-s} dF(x) = L(s-\alpha)^{-1} + \phi(s),$$

*where $\widehat{F}$ is the associated Dirichlet series of $F$, holds on the corresponding open half plane. Then*

$$F(x) = L\frac{x^\alpha}{\alpha} + o(x^\alpha).$$

**Proof.** *Note that $F(x) = O(x^{\alpha+\varepsilon})$ for any $\varepsilon > 0$, since the Dirichlet series converges for $\sigma > \alpha$. Now let $u = \alpha \log x$ and define $f$ by $f(u) := F(e^u/\alpha) = F(x)$. Then for $\sigma > 1$, we have*

$$\int_0^\infty e^{-su} df(u) = \widehat{F}(\alpha s) = \frac{L}{\alpha(s-1)} + \phi(\alpha s).$$

*Now, the above estimate of $F$ implies that $f(u) = O(e^{u+\varepsilon u})$ for any $\varepsilon > 0$. Therefore, for $\sigma > 1$,*

$$\begin{aligned} \int_0^\infty f(u)e^{-su}du &= s^{-1}\int_0^\infty e^{-su}df(u) \\ &= \frac{L}{\alpha s(s-1)} + \frac{\phi(\alpha s)}{s} = \frac{L}{\alpha(s-1)} + \phi_1(s), \end{aligned} \tag{2.13}$$

*where $\phi_1(s) = s^{-1}\phi(\alpha s) - L/(\alpha s)$, so that $\phi_1$ is continuous on $\sigma \geq 1$. If we express $L/(\alpha(s-1))$ as a Laplace integral, we obtain*

$$\phi_1(s) = \int_0^\infty e^{-su}\left\{f(u) - \frac{Le^u}{\alpha}\right\}du, \, \sigma > 1.(2.14)$$

*There remains two things to do. First, we establish an integral relation valid for each positive number $\lambda$, and second we give a tauberian argument based on this relation. To begin:*

$$\lim_{y\to\infty}\int_0^\infty e^{-x}f(x)K^*_\lambda(y-x)dx = \frac{L}{\alpha}, \tag{2.15}$$

*where*

$$K^*_\lambda(t) = K_\lambda(t)\Big/\int_{-\infty}^\infty K_1(u)du.$$

*We know that $K_\lambda^*$ has total integral* $1$*, independent of* $\lambda$*. When* $\lambda$ *is large,* $K_\lambda^*$ *is sharply peaked near the origin.*

*Let* $\varepsilon$ *and* $\lambda$ *be positive,* $s = 1+\varepsilon+it$*, and multiply* (2.14) *by* $(1/2)\left\{1-|t|/(2\lambda)\right\}e^{ity}$*, and integrate. We get that*

$$\begin{aligned}\frac{1}{2}\int_{-2\lambda}^{2\lambda}&(1-\frac{|t|}{2\lambda}e^{ity}\phi_1(1+\varepsilon+it)dt\\ &=\frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{ity}\left\{\int_0^\infty\left(f(u)-\frac{Le^u}{\alpha}\right)e^{-u(1+\varepsilon+it)}du\right\}dt\\ &=\int_0^\infty e^{-u-\varepsilon u}\left(f(u)-\frac{Le^u}{\alpha}\right)\left\{\frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{it(y-u)}dt\right\}du.\end{aligned}$$

*We have that the interchange of integration is justified since* $\int_0^\infty e^{-u-\varepsilon u}|f(u)-Le^u/\alpha|du$ *converges. Now, in terms of* $K_\lambda$*, we have*

$$\begin{aligned}\frac{1}{2}\int_{-2\lambda}^{2\lambda}&(1-\frac{|t|}{2\lambda}e^{ity}\phi_1(1+\varepsilon+it)dt\\ &=\int_0^\infty e^{-u-\varepsilon u}f(u)K_\lambda(y-u)du-\frac{L}{\alpha}\int_0^\infty e^{-\varepsilon u}K_\lambda(y-u)du.\end{aligned}\tag{2.16}$$

*Now we let* $\varepsilon\to 0^+$ *in equation* (2.16)*. Since* $\phi_1$ *is continuous and the range of integration is bounded, the limit may be taken inside the left integral. Also, we have that* $K_\lambda$ *is continuous, nonnegative, and has a finite integral, so*

$$\lim_{\varepsilon\to 0^+}\int_0^\infty e^{\varepsilon u}K_\lambda(y-u)du=\int_0^\infty K_\lambda(y-u)du.$$

*Therefore, for any real* $y$*,*

$$\begin{aligned}\lim_{\varepsilon\to 0^+}&\int_0^\infty e^{-u-\varepsilon u}f(u)K_\lambda(y-u)du\\ &=\frac{L}{\alpha}\int_0^\infty K_\lambda(y-u)du+\frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{ity}\phi_1(1+it)dt.\end{aligned}\tag{2.17}$$

*For any* $y$*, the limit in the last equation exists as a real number. We wish to take the limit inside the integral, but since we are not using Lebesgue theory, we give some details.*

*Let* $R(y)$ *denote the right hand side of* (2.17)*, and let* $\eta > 0$ *be given. Then* $f\geq 0$ *since* $f\in\mathcal{V}$ *and is nondecreasing. Furthermore,* $K_\lambda\geq 0$*, and therefore, for* $0<\varepsilon\leq\varepsilon_0$*,*

$$0\leq R(y)-\int_0^\infty e^{-u-\varepsilon u}f(u)K_\lambda(y-u)du<\eta.$$

*In view of* $e^{-u-\varepsilon_0 u}f(u)$ *being bounded, we have that*

$$\int_U^\infty e^{-u-\varepsilon_0 u}f(u)K_\lambda(y-u)du < \eta$$

*holds for sufficiently large* $U$*, and therefore,*

$$0 \leq R(y) - \int_0^U e^{-u-\varepsilon_0 u}f(u)K_\lambda(y-u)du < 2\eta.$$

*Since both* $f$ *and* $K_\lambda$ *are nonnegative, we have that this last inequality holds for* $\varepsilon_0 = 0$ *(for sufficiently large* $U$*); therefore,*

$$\begin{aligned}\int_0^\infty e^{-u}f(u)K_\lambda(y-u)du =& \frac{L}{\alpha}\int_0^\infty K_\lambda(y-u)du \\ &+ \frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{ity}\phi_1(1+it)dt.\end{aligned}$$

*Next, let* $y \to \infty$ *and apply Lemma 5 to obtain*

$$\lim_{y\to\infty}\frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1-\frac{|t|}{2\lambda}\right)e^{iyt}\phi_1(1+it)dt = 0.$$

*And it now follows that*

$$\lim_{y\to\infty}\left\{\int_0^\infty e^{-u}f(u)K_\lambda(y-u)du - \frac{L}{\alpha}\int_0^\infty K_\lambda(y-u)du\right\} = 0.$$

*By monotonicity, we have*

$$\lim_{y\to}\frac{L}{\alpha}\int_{-\infty}^{y} K_\lambda(x)dx = \frac{L}{\alpha}\int_{-\infty}^{\infty} K_1(x)dx,$$

*and* (2.15) *is obtained by dividing through by* $\int K_1(x)dx$*.*

*Now for the second stage of the argument. We begin by showing that* $f(y) = O(e^y)$*. Let* $\lambda$ *and* $\delta$ *be positive numbers and let* $J(y,\lambda)$ *denote the integral of* (2.15)*. Since the integrand of* $J(y,\lambda)$ *is nonnegative, and that* $f$ *and the exponential function are monotone, we have that for any* $y > \delta$*,*

$$J(y,\lambda) \geq \int_{y-\delta}^{y+\delta} e^{-x}f(x)K_\lambda^*(y-x)dx \geq f(y-\delta)e^{-y-\delta}\int_{-\delta}^{\delta} K_\lambda^*(u)du.$$

*We combine this with* (2.15) *to get*

$$f(y-\delta)e^{-(y-\delta)} \leq e^{2\delta}\frac{L}{\alpha}\Big/\left(\int_{-\delta}^{\delta} K_\lambda^*(u)du\right) + o(1), \tag{2.18}$$

*which tells us that* $f(y)/e^y$ *is bounded.*

*Now pick* $\delta = \left\{(\lambda/2)\int_{-\infty}^{\infty} K_1(u)du\right\}^{-1/2}$. *By* (2.12),

$$\int_{|u|>\delta} K_\lambda^*(u)du \le \frac{2}{\lambda\delta\int K_1(u)du} = \delta,$$

*and therefore,*

$$\int_{-\delta}^{\delta} K_\lambda^*(u)du \ge 1-\delta. \tag{2.19}$$

*Now let* $\varepsilon > 0$ *be given. Choose* $\lambda$ *large (and therefore* $\delta$ *small) to ensure that* $e^{2\delta}/(1-\delta) < 1+\varepsilon$. *Now from* (2.18) *and* (2.19), *we have, as* $y \to \infty$,

$$\frac{f(y)}{e^y} \le \frac{L(1+\varepsilon)}{\alpha} + o(1).$$

*Furthermore, this relation holds for arbitrary* $\varepsilon > 0$, *and thus,*

$$\limsup_{y\to\infty} \frac{f(y)}{e^y} \le \frac{L}{\alpha}.$$

*Now our goal is to obtain an inequality in the opposite direction. Since* $f(y)/e^y$ *is bounded, there exists a* $\delta > 0$ *such that* $f(y)/e^y \le b$ *for all* $y \ge 0$. *We have, for any positive* $\lambda$ *and* $y$,

$$\begin{aligned} J(y,\lambda) &\le b\int_{|u|>\delta} K_\lambda^*(u)du + f(y+\delta)e^{-(y+\delta)}\int_{-\delta}^{\delta} K_\lambda^*(u)du \\ &\le b\delta + f(y+\delta)e^{-(y+\delta)}e^{w\delta}. \end{aligned}$$

*Now, by the last inequality of* (2.15),

$$f(y)e^{-y} \ge \frac{Le^{-2\delta}}{\alpha} - b\delta e^{-2\delta} + o(1)$$

*as* $y \to \infty$ *for each fixed pair* $\lambda$, $\delta$ *that satisfies*

$$\lambda\delta^2 = 2/\int_{-\infty}^{\infty} K_1(u)du.$$

*Therefore, for each* $\delta > 0$, *we have*

$$\liminf_{y\to\infty} \frac{f(y)}{e^y} \ge \frac{Le^{-2\delta}}{\alpha} - b\delta e^{-2\delta},$$

*and hence*

$$\liminf_{y\to\infty} \frac{f(y)}{e^y} \ge \frac{L}{\alpha}.$$

*Finally, the two inequalities imply the Wiener-Ikehara theorem,* $\lim_{y\to\infty} f(y)/e^y = L/\alpha$ *or*

$$F(x) = \frac{Lx^\alpha}{\alpha} + o(x^\alpha).$$

Before we can generalize this result, we need another lemma in order to estimate integrals of the form $\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du$.

**Lemma 6.** *Let $\gamma$ denote a fixed number in the interval, $(1,2)$. Then there exist the functions $g = g_\lambda$ defined on $[2,\infty)$, satisfying that $g(\lambda) \to 0$ as $\lambda \to \infty$, and $\theta = \theta_\gamma(y,\lambda)$ with $|\theta| \le 1$, and such that for all $y \ge 2$ and $\lambda \ge 2$, we have*

$$\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du = y^{\gamma-1}\left\{1+\theta g(\lambda)\right\}\int_{-\infty}^\infty K_1(u)du. \tag{2.20}$$

*And for $\gamma$ fixed in the interval $(0,1)$ and any fixed $\lambda \ge 2$,*

$$\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du = o(1) \text{ as } y \to \infty.$$

**Proof.** *We noted before without proof that*

$$\int K_1 := \int_{-\infty}^\infty K_1(u)du = \pi.$$

*While we do not need this exact value, we do need a positive lower bound. Note that $|\sin u| \ge 2|u|/\pi$ for $|u| \le \pi/2$, which tells us that $\int K_1 > 4/\pi > 1$.*

*As in the proof of the Wiener-Ikehara theorem (Theorem 20), we define*

$$\delta := \left\{\frac{\lambda}{2}\int K_1\right\}^{-1/2} \tag{2.21}$$

*and observe that $0 < \delta < 1$ for $\lambda \ge 2$, and that $\delta \to 0$ as $\lambda \to \infty$.*

*For $1 < \gamma < 2$, we give a lower estimate of the integral by using the fact that $y - \delta > 0$ and applying (2.19). We get*

$$\begin{aligned}\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du &\ge (y-\delta)^{\gamma-1}\int_{y-\delta}^{y+\delta} K_\lambda(y-u)du\\ &\ge y^{\gamma-1}\left(1-\frac{\delta}{y}\right)^{\gamma-1}(1-\delta)\int K_1\\ &\ge y^{\gamma-1}\left(1-\frac{\delta}{2}\right)(1-\delta)\int K_1.\end{aligned}$$

*As for an upper estimate, we bound the integral over four intervals. By Lemma 4 and equation 2.21 (and the fact that $y \ge 2$), we obtain*

$$\begin{aligned}\int_0^{y-\delta} u^{\gamma-1}K_\lambda(y-u)du &\le (y-\delta)^{\gamma-1}\int_0^{y-\delta} K_\lambda(y-u)du\\ &\le y^{\gamma-1}\int_\delta^y K_\lambda(v)dv \le y^{\gamma-1}\frac{\delta}{\lambda\delta^2} = \frac{1}{2}y^{\gamma-1}\delta\int K_1,\end{aligned}$$

$$\begin{aligned}\int_{y-\delta}^{y+\delta} u^{\gamma-1}K_\lambda(y-u)du &= \int_0^\delta K_\lambda(v)\left\{(y-v)^{\gamma-1}+(y+v)^{\gamma-1}\right\}dv\\ &\le 2y^{\gamma-1}\int_0^\delta K_\lambda(v)dv < y^{\gamma-1}\int K_1,\end{aligned}$$

$$\int_{y+\delta}^{2y} u^{\gamma-1}K_\lambda(y-u)du \le (2y)^{\gamma-1}\int_{y+\delta}^{2y} K_\lambda(y-u)du \le y^{\gamma-1}\delta\int K_1,$$

$$\begin{aligned}\int_{2y}^{\infty} u^{\gamma-1}K_\lambda(y-u)du &\le \int_{2y}^{\infty} u^{\gamma-1}\lambda^{-1}(u-y)^{-2}du\\ &\le \frac{4}{\lambda}\int_{2y}^{\infty} u^{\gamma-3}du \le \frac{4y^{\gamma-2}}{\lambda(2-\gamma)} \le \frac{2y^{\gamma-1}}{\lambda(2-\gamma)}\int K_1,\end{aligned}$$

*where the second integral was estimated using the symmetry of $K_\lambda$ about the origin and the concavity of the function $u^{\gamma-1}$.*

*Now taking them altogether, we see that*

$$\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du \le y^{\gamma-1}\left\{1+\frac{3\delta}{2}+\frac{2}{\lambda(2-\gamma)}\right\}\int K_1.$$

*By combining the lower and upper estimates and recalling the definition of $\delta$ (equation 2.21), we obtain* (2.20). *For a fixed $\gamma\in(1,2)$, the error estimate satisfies $g_\gamma(\lambda)=O_\gamma(1/\sqrt{\lambda})$.*

*Now we turn to the case when $0<\gamma<1$ and separate $[0,\infty)$ into three segments to estimate each part separately. We have the following:*

$$\int_0^{y/2} u^{\gamma-1}K_\lambda(y-u)du \le \frac{4}{\lambda y^2}\int_0^{y/2} u^{\gamma-1}du < \frac{4}{\gamma\lambda}y^{\gamma-2},$$

$$\int_{y/2}^{2y} u^{\gamma-1}K_\lambda(y-u)du \le 2y^{\gamma-1}\int_{y/2}^{2y} K_\lambda(y-u)du < 2y^{\gamma-1}\int K_1,$$

$$\int_{2y}^{\infty} u^{\gamma-1}K_\lambda(y-u)du \le 4y^{\gamma-2}\lambda^{-1}(2-\gamma)^{-1},$$

*where the last integral was estimated similarly in the other case. Therefore, we have for each fixed $\gamma\in(0,1)$*

$$\int_0^\infty u^{\gamma-1}K_\lambda(y-u)du = O(y^{\gamma-1}) = o(1) \text{ as } y\to\infty.$$

We can now prove the generalized Wiener-Ikehara theorem.

**Theorem 21.** *Let $F$ be a real valued nondecreasing function in $\mathcal{V}$ with $\sigma_c(\widehat{F})=\alpha>0$. Let $\phi$ and $\vartheta$ be functions which are analytic on the closed half plane $\sigma\ge\alpha$ and assume that $\phi(\alpha)\ne 0$. For $\gamma$ a real number distinct* *from* $0$, $-1$, $-2$, $\ldots$, *let* $(s-\alpha)^{-\gamma}$ *be positive valued on the real ray,* $s > \alpha$. *Further suppose that*

$$\widehat{F}(s) = \int x^{-s} dF(x) = (s-\alpha)^{-\gamma}\phi(s) + \vartheta(s)$$

*holds on the open half plane* $\sigma > \alpha$. *Then*

$$F(x) \sim \phi(\alpha)\frac{x^{\alpha}(\log x)^{\gamma-1}}{\alpha\Gamma(\gamma)},$$

*where* $\Gamma$ *denotes the Euler gamma function.*

**Proof.** *The Wiener-Ikehara theorem (Theorem 20) covers the case of* $\gamma = 1$. *We consider the case* $1 < \gamma < 2$. *This case follows Theorem 20 closely. First, note that* $\phi(\alpha) > 0$ *in this case, since* $\widehat{F}(\sigma) > 0$ *on the ray* $\sigma > \alpha$ *and*

$$\phi(\alpha) = \lim_{\sigma\to\alpha^{+}} (\sigma-\alpha)^{\gamma}\widehat{F}(\sigma) \neq 0.$$

*Now set* $u = \alpha \log x$ *and define* $f(u) = F(e^{u/\alpha}) = F(x)$. *Then, for* $\sigma > 1$, *we have*

$$\int_0^{\infty} e^{-su} df(u) = \widehat{F}(\alpha s) = (s-1)^{-\gamma}\alpha^{-\gamma}\phi(\alpha s) + \vartheta(\alpha s).$$

*We integrate by parts and obtain*

$$\int_0^{\infty} e^{-su} f(u) du = \frac{\widehat{F}(\alpha s)}{s} \text{ for } \sigma > 1.$$

*Now, by expanding* $\phi(\alpha s)/s$ *in a Taylor series about* $s = 1$, *we find that*

$$\frac{\alpha^{-\gamma}}{(s-1)^{\gamma}}\frac{\phi(\alpha s)}{s} + \frac{\vartheta(\alpha s)}{s} = \frac{a}{(s-1)^{\gamma}} + \frac{b}{(s-1)^{\gamma-1}} + \phi_1(s), \tag{2.22}$$

*where* $a := \alpha^{-\gamma}\phi(\alpha) > 0$, $b$ *is some constant and* $\phi_1$ *is a continuous function on the closed half plane* $\sigma \geq 1$. *Now*

$$(s-1)^{-\beta} = \int_0^{\infty} e^{-su}\frac{e^{u}u^{\beta-1}du}{\Gamma(\beta)}$$

*for* $\sigma > 1$ *and* $\beta > 0$. *This identity can be established for* $s$ *real and* $s > 1$ *by changing the variable in the integral representation of the gamma function. The result follows for any complex* $s$ *with* $\sigma > 1$ *by analytic continuation. Therefore, for* $\sigma > 1$,

$$\phi_1(s) = \int_0^{\infty} e^{-su}\left\{f(u) - \frac{a}{\Gamma(\gamma)}e^{u}u^{\gamma-1} - \frac{b}{\Gamma(\gamma-1)}e^{u}u^{\gamma-2}\right\} du.$$

*We will show that $f(u) \sim ae^u u^{\gamma-1}/\Gamma(\gamma)$, which is equivalent to the theorem. As in the proof of Theorem 20, we take $s = 1 + \varepsilon + it$ with $\varepsilon > 0$, form the integral*

$$\frac{1}{2}\int_{-2\lambda}^{2\lambda} (1 - \frac{|t|}{2\lambda} e^{ity} \phi_1(1+\varepsilon+it)dt = \frac{1}{2}\int_{-2\lambda}^{2\lambda} \left(1 - \frac{|t|}{2\lambda}\right) e^{ity} \times$$

$$\int_{\infty \left\{ f(u) - \frac{a}{\Gamma(\gamma)} e^u u^{\gamma-1} - \frac{be^u}{\Gamma(\gamma-1)} u^{\gamma-2} \right\} \cdot e^{-u(1+\varepsilon+it)} dudt}$$

$$= \int_0^\infty e^{-u-\varepsilon u} \left\{ f(u) - \frac{a}{\Gamma(\gamma)} e^u u^{\gamma-1} - \frac{be^u}{\Gamma(\gamma-1)} u^{\gamma-2} \right\} K_\lambda(y-u)du,$$

*and let $\varepsilon \longrightarrow 0^+$. We use an argument similar as before to obtain*

$$\frac{1}{2}\int_{-2\lambda}^{2\lambda} (1 - \frac{|t|}{2\lambda} e^{ity} \phi_1(1+it)dt$$

$$= \int_0^\infty e^{-u} f(u) K_\lambda(y-u)du - \frac{a}{\Gamma(\gamma)} \int_0^\infty u^{\gamma-1} K_\lambda(y-u)du$$

$$- \frac{b}{\Gamma(\gamma-1)} \int_0^\infty u^{\gamma-2} K_\lambda(y-u)du.$$

*Now, Lemma 6 tells us that the last integral is $o(1)$ as $y \longrightarrow \infty$. The Riemann-Lebesgue lemma implies that the last integral containing $\phi_1$ also tends to zero as $y \longrightarrow \infty$. Therefore, for each $\lambda \geq 2$, as $y \longrightarrow \infty$ we have by (2.20) that*

$$\int_0^\infty e^{-u} f(u) K_\lambda(y-u)du = \frac{ay^{\gamma-1}}{\Gamma(\gamma)} \{1 + \theta g(\lambda)\} \int K_1 + o(1), \qquad (2.23)$$

*where $|\theta| \leq 1$ and $g(\lambda) = o(1)$ as $\lambda \longrightarrow \infty$.*

*Again, we follow in the same manner as in the proof of the Wiener-Ikehara theorem. We estimate the left hand side of (2.23) by using the range $y - \delta \leq u \leq y + \delta$, where $\delta = \delta(x)$ is chosen as it was in (2.21). Letting $y - \delta = w$, we obtain the estimate*

$$e^{-w} f(w) \leq \frac{ae^{2\delta}(w+\delta)^{\gamma-1}}{\Gamma(\gamma)(1-\delta)} \{1 + \theta g(\lambda)\} + o(1) \qquad (2.24)$$

*for fixed $\lambda \geq 2$. Therefore, (2.24) gives the bound $f(w)/e^w \leq Bw^{\gamma-1}$ for some $B > 0$ and for all $w >\geq 2$. Now we give upper estimates of the integrals $\int_2^{y-\delta}$ and $\int_{y+\delta}^\infty$ in the left hand side of (2.23) by using this bound and some inequalities from the proof of Lemma 2.16. Because the lemma uses only $y \geq 2$, we observe that*

$$\int_0^2 e^{-u} f(u) K_\lambda(y-u)du < f(2) \int_{y-2}^y K_\lambda(v)dv = o(1).$$

*We replace $y+\delta$ with $w$ to obtain*

$$\frac{f(w)}{e^w} \geq (w-\delta)^{\gamma-1}e^{-2\delta}\left\{\frac{a}{\Gamma(\gamma)} - \frac{ag(\lambda)}{\Gamma(\gamma)} - \frac{3B\delta}{2} - \frac{2B}{\lambda(2-\gamma)}\right\} + o(1).$$

*Now, $\lambda$ can be chosen sufficiently large, and $\delta \to 0^+$ as $\lambda \to \infty$. If we combine (2.24) with this last inequality, and take $\lambda$ large, we are able to conclude that*

$$\frac{f(w)}{e^w} = (1+o(1))\frac{a}{\Gamma(\gamma)}w^{\gamma-1} + o(1), \tag{2.25}$$

*which is equivalent to the assertion of the theorem for the case $1<\gamma<2$.*

*Now we turn our attention to the case $\gamma<1$ and $\gamma \neq 0,-1,-2,\ldots$. Let $N$ be the positive integer for which $1<\gamma+N<2$. Define $F_1(x) := \int_1^x L^N dF$. We form the Dirichlet series*

$$\begin{aligned}
\widehat{F_1}(s) &= \int x^{-s}\log^N x dF(x) = (-1)^N \widehat{F}^{(N)}(s) \\
&= (-1)^N \sum_{j=0}^{N} \binom{N}{j} \left\{(s-\alpha)^{-\gamma}\right\}^{(N-j)} \phi^{(j)}(s) + (-1)^N \vartheta^{(N)}(s) \\
&= (s-\alpha)^{-\gamma-N}\frac{\Gamma(\gamma+N)}{\Gamma(\gamma)}\phi(s) + \cdots + (-1)^N(s-\alpha)^{-\gamma}\phi^{(N)}(s) \\
&\quad (-1)^N\vartheta^{(N)}(s) \\
&= (s-\alpha)^{-\gamma-N}\Phi(s) + (-1)^N\vartheta^{(N)}(s),
\end{aligned}$$

*where $\Phi$ and $\vartheta^{(N)}$ are analytic functions on $\sigma \geq \alpha$, and*

$$\Phi(\alpha) = \frac{\Gamma(\gamma+N)\phi(\alpha)}{\Gamma(\gamma)}.$$

*As with $1<\gamma<2$, we have $\widehat{F_1}(\sigma) > 0$ on $\sigma > \alpha$, and therefore,*

$$\Phi(\alpha) = \lim_{\sigma\to\alpha^+}(\sigma-\alpha)^{\gamma+N}\widehat{F_1}(\sigma) > 0.$$

*And because $\Gamma(\gamma+N)>0$, we have that $\phi(\alpha)/\Gamma(\gamma)>0$ here.*

*Since $1<\gamma+N<2$, we apply to $F_1(x)$ the form of the theorem that we have already proved to obtain*

$$F_1(x) \sim \frac{\phi(\alpha)}{\alpha\Gamma(\gamma)}x^\alpha(\log x)^{\gamma+N-1}.$$

*Now for $x \geq e$,*

$$\begin{aligned}
F(x) &= F(e) + \int_e^x L^{-N}dF_1 \\
&= F_1(x)\log^{-N}x + N\int_e^x t^{-1}F_1(t)(t)^{-N-1}dt + O(1) \\
&= \{1+o(1)\}\frac{\phi(\alpha)}{\alpha\Gamma(\gamma)}x^\alpha(\log x)^{\gamma-1} + O\left\{x^\alpha(\log x)^{\gamma-2}\right\},
\end{aligned}$$

*which proves the theorem for* $\gamma < 1$, *and* $\gamma \neq 0, -1, -2, \ldots$.

*Finally, we can turn to the case* $\gamma \geq 2$. *We sketch the argument for* $2 \leq \gamma < 4$, *from which the general case follows.*

*Let* $f^+ := \max(f, 0)$, *then we have*

$$\begin{aligned}
\lambda^{-1} K_\lambda^2(x) &= \lambda \left( \frac{\sin \lambda x}{\lambda x} \right)^4 \\
&= \frac{1}{4\lambda} \int \left( 1 - \frac{|t|}{2\lambda} \right)^+ e^{ixt} dt \cdot \int \left( 1 - \frac{|u|}{2\lambda} \right)^+ e^{ixu} du \\
&= \int_{-4\lambda}^{4\lambda} h(v) e^{ixv} dv,
\end{aligned}$$

*where* $h$ *is the continuous function supported on* $[-4\lambda, 4\lambda]$ *and defined by*

$$h(v) = \frac{1}{4\lambda} \int \left( 1 - \frac{|t|}{2\lambda} \right)^+ \left( 1 - \frac{|v - t|}{2\lambda} \right)^+ dt.$$

*We do not need an explicit representation of* $h$, *and proceed using Lemma 4 concerning the properties of* $K_\lambda$. *The function* $K_\lambda^2/\lambda$ *satisfies analogous relations, and in particular*

$$\int_{-\infty}^{\infty} \frac{K_\lambda^2(u)}{\lambda} du = \int_{-\infty}^{\infty} K_1^2(u) du$$

*for all real* $\lambda$. *Now the analogue of Lemma 6 holds for* $2 \leq \gamma < 4$ *if we use* $K_\lambda^2/\lambda$ *instead of* $K_\lambda$.

*Now to prove the theorem for* $2 \leq \gamma < 4$, *we change* (2.22) *to exhibit all powers of* $s - 1$ *occurring with a negative exponent. In* (2.23), *we replace* $K_\lambda$ *with* $K_\lambda^2/\lambda$ *and the right hand side of* (2.23) *is changed by the inclusion of terms containing the factor* $y^{\gamma-2}$ *and* $y^{\gamma-3}$. *Now, these terms are of smaller order than the term containing the factor* $y^{\gamma-1}$, *and so the conclusion of the proof is as before.*

*For the general case* $\gamma \geq 2$, *we choose a positive integer* $N$ *for which* $2N > \gamma$, *and use the function* $\lambda^{1-N} K_\lambda^N$ *instead of* $K_\lambda$. *We have*

$$\lambda^{1-N} K_\lambda^N = \int e^{ixv} h_N(v) dv,$$

*where*

$$\begin{aligned}
h_N(v) =& \lambda^{1-N} 2^{-N} \int \cdots \int \left( 1 - \frac{|t_1|}{2\lambda} \right)^+ \times \\
& \cdots \times \left( 1 - \frac{|t_{N-1}|}{2\lambda} \right)^+ \left( 1 - \frac{|v - t_1 - \cdots - t_{N-1}|}{2\lambda} \right)^+ dt_1 \cdots dt_{N-1}.
\end{aligned}$$

## 2.9 Riemann Zeta Function and Dirichlet $L$-Functions

Both the Riemann zeta function and the Dirichlet L-functions are examples of Dirichlet series, and this information will be invaluable to for the first method of estimation. The next two subsections summarize the results when we apply what we know about Dirichlet series to the Riemann zeta function and Dirichlet $L$-functions.

### 2.9.1 Properties of $\zeta(s)$

We will begin with one of many definitions for the Riemann zeta function, denoted $\zeta(s)$.

**Definition 18.** *For $\sigma > 1$,*

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

We have the following properties of the zeta function due to the theorems on Dirichlet series applied to the zeta function.

**Properties of $\zeta(s)$**

**(1)** $\zeta(s)$ converges for all $s > 1$ and diverges at $s = 1$, so the abscissa of convergence is $\sigma_a = 1$.

**(2)** Both $\sum n^{-s}$ and $\sum \mu(n)n^{-s}$ converge absolutely for $\sigma > 1$. Utilizing our theorem concerning multiplication of Dirichlet series above, and taking $f(n) = 1$ and $g(n) = \mu(n)$, we get $h(n) = [1/n]$, thus for $\sigma > 1$

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

Particularly, this shows that $\zeta(s) \neq 0$ for $\sigma > 1$ and also that

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

**(3)** Applying the Euler product to $\zeta(s)$, i.e. when $f(n) = 1$, we obtain the product formula for $\zeta(s)$. For $\sigma > 1$

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

When $f(n) = \mu(n)$, $\sigma > 1$, we get the relation

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

**(4)** By differentiating the zeta function term by term, we get the following

$$\zeta'(s) = -\sum_{n=1}^{\infty} \frac{\log n}{n^s} \tag{2.26}$$

As mentioned in our properties, we have an inverse of $\zeta(s)$, but this deserves special note:

**Theorem 22 (Inverse of the Riemann zeta function).** *For $\sigma > 1$*

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

*where $\mu$ is the Möbius function.*

### 2.9.2 Properties of $L(s, \chi)$

**Definition 19.** *The* Dirichlet $L$-function *is defined as follows, for $\sigma > 1$,*

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

*where $\chi$ is a Dirichlet character.*

**Properties of $L(s, \chi)$**

**(1)** Since $\chi(n)$ is bounded, i.e. $|\chi(n)| \leq 1$, we have that $L(s, \chi)$ converges absolutely for $\sigma > 1$, and $\sigma_a \leq 1$.

In general, if $|f(n)| \leq M$ for some $M$ and for all $n \geq 1$, then $\sum f(n)n^{-s}$ converges absolutely for $\sigma > 1$.

**(2)** We utilize the theorem concerning multiplication of Dirichlet series, and assume $F(s) = \sum f(n)n^{-s}$ converges absolutely for $\sigma > \sigma_a$. If $f$ is completely multiplicative, then we have $f^{-1}(n) = \mu(n)f(n)$. Since $\left|f^{-1}(n)\right| \leq |f(n)|$, the series $\sum \mu(n)f(n)n^{-s}$ also converges absolutely for $\sigma > \sigma_a$ and we get

$$\sum_{n=1}^{\infty} \frac{\mu(n)f(n)}{n^s} = \frac{1}{F(s)} \text{ if } \sigma > \sigma_a.$$

In particular, for every Dirichlet character $\chi$, we have

$$\sum_{n=1}^{\infty} \frac{\mu(n)\chi(n)}{n^s} = \frac{1}{L(s, \chi)} \text{ if } \sigma > 1.$$

**(3)** Applying the Euler product formula to $\chi(n)$, we get an equivalent product formula of $L(s, \chi)$,

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1-\chi(p)p^{-s}} \text{ if } \sigma > 1.$$

We also have the particular case when $\chi = \chi_1$, the principal character modulo $k$. In this example, $\chi_1(p) = 0$ if $p|k$ and $\chi_1(p) = 1$ if $p \nmid k$, so the Euler product for $L(s, \chi_1)$ becomes

$$L(s,\chi_1) = \prod_{p \nmid k} \frac{1}{1-p^{-s}} = \prod_p \frac{1}{1-p^{-s}} \cdot \prod_{p|k}(1-p^{-s}) = \zeta(s) \prod_{p|k}(1-p^{-s}).$$

Therefore, $L(s, \chi_1)$ is equal to $\zeta(s)$ multiplied by a finite number of factors.

### 2.9.3 Hurwitz Zeta Function

In fact, not only are $\zeta$ and $L$ Dirichlet series, but they are also examples of what are called *Hurwitz zeta functions*. These functions are the workhorses we will use to prove the Prime Number Theorem.

**Definition 20.** *The Hurwitz zeta function, denoted $\zeta(s, w)$, defined for $\sigma > 1$ by*

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

*where $w$ is a fixed real number, $0 < w \leq 1$.*

When $w = 1$, this becomes the Riemann zeta function, i.e. $\zeta(s) = \zeta(s, 1)$. As mentioned before, we can also express $L(s, \chi)$ in terms of Hurwitz zeta functions.

**Theorem 23.** *For $\sigma > 1$, and where $\chi$ is a Dirichlet character modulo $k$,*

$$L(s,\chi) = \frac{1}{k^s} \sum_{w=1}^{k} \chi(w) \zeta\left(s, \frac{w}{k}\right).$$

It is important to note that Hurwitz zeta functions are not necessarily Dirichlet series, so that many of the things we have proven for Dirichlet series are not immediately applicable to this type of functions. However, it will be seen that we have enough of the same properties.

The remainder of this chapter will be devoted to exploring Hurwitz zeta functions while building up the machinery necessary to prove the prime number theorem, which happens to be the same tools necessary to prove our primary result. This material can be found particularly in [LeV2].

As expected, $\zeta(s, w)$ converges uniformly in its half-plane of convergence.

**Theorem 24.** *For any $\sigma_0 > 1$, the series*

$$\sum_{n=0}^{\infty} \frac{1}{(n+w)^{-s}}$$

*converges uniformly for $\sigma \geq \sigma_0$; thus $\zeta(s,w)$ is analytic for $\sigma > 1$.*

In order to give an analytic continuation of $\zeta(s,w)$, we have need of the following lemma.

**Lemma 7.** *If $a$ and $b$ are integers with $0 \leq a < b$, and if $f$ has a continuous derivative over $a \leq x \leq b$, then*

$$\sum_{n=a+1}^{b} f(n) = \int_a^b f(u)du + \int_a^b (u - \lfloor u \rfloor) f'(u)du.$$

Now we can illustrate the analytic continuation of a Hurwitz zeta function over the half-plane $\sigma > 0$.

**Theorem 25.** *If $m$ is a non-negative integer, and $\sigma > 1$, then*

$$\zeta(s,w) - \frac{1}{(s-1)(m+w)^{s-1}} = \sum_{n=0}^{m} \frac{1}{(n+w)^s} - s\int_m^{\infty} \frac{u - \lfloor u \rfloor}{(u+w)^{s-1}} du. \quad (2.27)$$

*It follows that $\zeta(s,w) - 1/(s-1)$ is analytic for $\sigma > 0$, and that equation 2.27 holds for $\sigma > 0$.*

In fact, the function is analytic over the entire plane except for the pole at $s = 1$, but we have no need of this fact.

For the remainder of this chapter, and when relevant in chapter 5, $c$ will play the part usually reserved for $\varepsilon$, denoting a positive constant which depends only on the arguments listed. It need not have the same value when it occurs in different results, unless indicated so by a particular subscript. When possible, the author has included figures to help describe the various $c$'s.

We have the next result, giving us useful bounds on $\zeta(s,w)$.

**Theorem 26.** *For $\frac{1}{2} \leq \sigma \leq 2$ and $t > c(w)$, where $c$ is a function of $w$,*

$$|\zeta(s,w)| < t^{3/4}.$$

*For $t \geq 8$ and $1 - (\log t)^{-1} \leq \sigma \leq 2$, we have*

$$|\zeta(s,w)| < c(w)\log t.$$

The next few theorems and lemmas attempt to narrow down the bounds and integrals for functions possessing properties like the Hurwitz zeta functions. We will use all of these bounds to rid ourselves of some very ugly integrals.

**Theorem 27.** *For $|x| \leq 1$, if*

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

*is analytic, and* $\operatorname{Re} f(x) \leq \frac{1}{2}$, *then* $|a_n| \leq 1$ *for* $n \geq 1$.

**Lemma 8.** *Let $R > 0$, and suppose that*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

*is analytic and* $\operatorname{Re} f(x) \leq M$ *for* $|x - x_0| \leq R$. *For* $n \geq 1$, *we have*

$$|a_n| \leq \frac{2}{R^n}(M - \operatorname{Re} a_0).$$

**Theorem 28.** *If $f$ satisfies the hypotheses of Lemma 8, and $0 < r < R$, then for $|x - x_0| \leq r$, we have*

$$|f(x)| \leq |a_0| + \frac{2r}{R-r}(|M| + |a_0|)$$

*and*

$$|f'(x)| \leq \frac{2R}{(R-r)^2}(|M| + |a_0|).$$

We will use this next theorem to eventually show that $\zeta$ does not vanish near $\sigma = 1$ and sufficiently far from $t = 0$.

**Theorem 29.** *Let $r > 0$ and $M \in \mathbb{R}$, and suppose that $f(s_0) \neq 0$ and that, for $|s - s_0| \leq r$, $f(s)$ is analytic and*

$$\left| \frac{f(s)}{f(s_0)} \right| < e^M.$$

*Further suppose that $f(s) \neq 0$ in the semicircular region $|s - s_0| \leq r$,* $\operatorname{Re} s > \operatorname{Re} s_0$. *Then*

$$-\operatorname{Re} \frac{f'}{f}(s_0) \leq \frac{4M}{r},$$

*and if there is a zero, say $\rho$, of $f$ on the open line segment between $s_0 - r/2$ and $s_0$, then*

$$-\operatorname{Re} \frac{f'}{f}(s_0) \leq \frac{4M}{r} - \frac{1}{s_0 - \rho}.$$

A note on notation before we continue: If $f$ is analytic on the vertical line $\sigma_0 + ti$, and if

$$\lim_{\substack{a\to\infty \\ b\to\infty}} \int_{\sigma_0 - ai}^{\sigma_0 + bi} f(s)ds = \lim_{\substack{a\to\infty \\ b\to\infty}} \int_{-a}^{b} f(\sigma_0 + ti)dt$$

exists, then we abbreviate this limit to

$$\int_{(\sigma_0)} f(s)ds,$$

as is the convention in [LeV2].

**Theorem 30.**

$$\frac{1}{2\pi i}\int_{(2)} \frac{y^s}{s^2}ds = \begin{cases} 0 \textit{ for } 0 < y < 1, \\ 1 \textit{ for } y \geq 1. \end{cases}$$

### 2.9.4 The Prime Number Theorem

The results of this section are easily modified to prove both Dirichlet's theorem on primes in arithmetical progressions and the primary result of this paper, i.e. the estimation of the number of integers expressible as a sum of two squares. Accordingly, the proofs of these results are included.

To begin with, we will need to learn something about the location of the zeroes of the Riemann zeta function. From the product formula for $\zeta(s)$

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

we see that $\zeta(s) \neq 0$ for $\sigma > 1$.

The following proof that $\zeta(1+ti) \neq 0$ is due to de la Vallée Poussin, but before we reach it, we should consider the following.

For $\sigma > 1$ we have

$$\log\zeta(s) = \sum_{m,p} \frac{1}{mp^{ms}} = \sum_p \frac{1}{p^s} + f(s),$$

where $f$ is clearly analytic for $\sigma > \frac{1}{2}$. Since $\zeta$ has a pole at $s = 1$, with residue 1, it follows that as $\sigma \to 1^+$,

$$\sum_p \frac{1}{p^\sigma} \sim \log\frac{1}{\sigma - 1}. \tag{2.28}$$

We continue heuristically. If $1 + t_0 i$ is a zero of $\zeta$, and we let $s = \sigma + t_0 i$, then, as $\sigma \to 1^+$,

$$\log|\zeta(s)| \sim \log(\sigma - 1)$$

and

$$\begin{aligned} \mathrm{Re}(\log\zeta(s)) - \mathrm{Re}\,(f(s)) &= \log|\zeta(s)| - \mathrm{Re}\,(f(s)) \\ &= \sum_p \frac{\cos(t_0 \log p)}{p^\sigma} \sim \log(\sigma - 1). \end{aligned}$$

We compare this with (2.28), and note that for most $p$, $\cos(t_0 \log p)$ is close to $-1$, but then $\cos(2t_0 \log p)$ will usually be near 1, and we have

$$\sum_p \frac{\cos(2t_0 \log p)}{p^\sigma} \sim \log \frac{1}{\sigma - 1}.$$

But then, this implies that $\zeta$ has a pole at $1 + 2t_0 i$, which is absurd.

For a rigorous argument, we note that for all real $\theta$,

$$3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geq 0.$$

Thus, for $\sigma > 1$,

$$\begin{aligned}
&\log \left|\zeta^3(\sigma)\zeta^4(\sigma + t_0 i)\zeta(\sigma + 2t_0 i)\right| \\
&\qquad = 3\log|\zeta(\sigma)| + 4\log|\zeta(\sigma + t_0 i)| + \log|\zeta(\sigma + 2t_0 i)| \\
&\qquad = 3\sum_{n,p} \frac{1}{np^{n\sigma}} + 4\sum_{n,p} \frac{\cos(t_0 n \log p)}{np^{n\sigma}} + \sum_{n,p} \frac{\cos(2t_0 n \log p)}{np^{n\sigma}} \\
&\qquad = \sum_{n,p} \frac{3 + 4\cos(t_0 n \log p) + \cos(2t_0 n \log p)}{np^{n\sigma}} \\
&\quad \geq 0.
\end{aligned}$$

Therefore,

$$((\sigma - 1)\zeta(\sigma))^3 \left|\frac{\zeta(\sigma + t_0 i)}{\sigma - 1}\right|^4 |\zeta(\sigma + 2t_0 i)| \geq \frac{1}{\sigma - 1},$$

and if $1 + t_0 i$ were a zero of $\zeta$, the left hand side would remaind bounded as $\sigma \to 1^+$, while the right hand side would increase without bound.

We are now in a position to show that $\zeta(s)$ does not vanish at any point close to the line $\sigma = 1$ and sufficiently far from $t = 0$ (the real axis). We utilize the argument above and Theorem 29

**Theorem 31.** *For $\sigma > 1$*

$$\mathrm{Re}\left(-3\frac{\zeta'}{\zeta}(\sigma) - 4\frac{\zeta'}{\zeta}(\sigma + ti) - \frac{\zeta'}{\zeta}(\sigma + 2ti)\right) \geq 0.$$

**Proof.** *We differentiate both sides of*

$$\log\zeta(s) = \sum_{m,p} \frac{1}{mp^{ms}}.$$

*We get that*

$$\frac{\zeta'}{\zeta}(s) = -\sum_{m,p} \frac{\log p}{p^{ms}} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}, \tag{2.29}$$

*where* $\Lambda(n)$ *is the Mangoldt function, that is,*

$$\Lambda(n) = \begin{cases} \log p \ \textit{if } n = p^m \textit{ for some prime } p \textit{ and some } m \geq 1, \\ \quad 0 \ \textit{else.} \end{cases}$$

*We have the following table to illustrate this function:*

| $n:$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Lambda(n)$ | 0 | $\log 2$ | $\log 3$ | $\log 2$ | $\log 5$ | 0 | $\log 7$ | $\log 2$ | $\log 3$ | 0 |

$\Lambda(n)$ *for* $1 \leq n \leq 10$

*Recall that the series for* $\log \zeta(s)$ *converges uniformly in any region to the right of* $\sigma = 1$*, thus we are allowed to differentiate termwise. We have the following:*

$$\begin{aligned}
\mathrm{Re}\Bigg( -3\frac{\zeta'}{\zeta}(\sigma) - 4\frac{\zeta'}{\zeta}(\sigma+ti) - \frac{\zeta'}{\zeta}(\sigma+2ti)\Bigg) \\
&= \mathrm{Re}\sum_{n=1}^{\infty} \frac{(3 + 4n^{-ti} + n^{-2ti})\Lambda(n)}{n^{\sigma}} \\
&= \sum_{n=1}^{\infty} \frac{(3 + 4\cos(t\log n) + \cos(2t\log n))\Lambda(n)}{n^{\sigma}} \\
&\geq 0,
\end{aligned}$$

*which was to be shown.*

From Theorem 26, we have the following:

**Theorem 32.** *(I) For* $\sigma \geq \frac{1}{2}$ *and* $t > c$*, we have* $|\zeta(s)| < 1$*, and*
*(II) For* $t \geq 8$ *and* $\sigma \geq 1 - (\log t)^{-1}$*, we have* $|\zeta(s)| < c\log t$*.*

**Proof.** *For* $\sigma \leq 2$*, both (I) and (II) follow directly from Theorem 26. For* $\sigma > 2$ *and* $t \geq 8$*,*

$$|\zeta(s)| < \sum_{n=1}^{\infty} \frac{1}{n^2} < 2 < \begin{cases} t, \\ \log t. \end{cases}$$

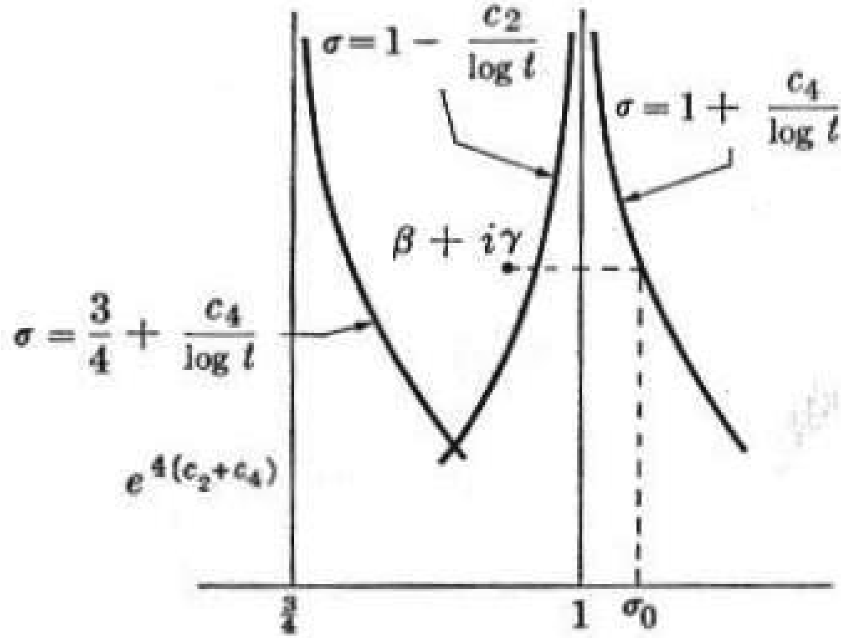
**Theorem 33.** *There exist constants* $c_1 > 8$ *and* $c_2 > 0$ *such that* $\zeta(s) \neq 0$ *for*

$$t > c_1 \textit{ and } \sigma > 1 - \frac{c_2}{\log t}.$$

**Proof.** *Following the premise of Theorem 32, choose* $c_3 > 8$ *such that*

$$|\zeta(s)| < t, \textit{ for } \sigma \geq \frac{1}{2},\ t > c_3.$$

*In view of the fact that*



$\sigma = 1 - \frac{c_2}{\log t}$
$\sigma = 1 + \frac{c_4}{\log t}$
$\beta + i\gamma$
$\sigma = \frac{3}{4} + \frac{c_4}{\log t}$
$e^{4(c_2+c_4)}$
$\frac{3}{4}$
1
$\sigma_0$

Figure 2.1: [LeV2], page 243

$$\frac{3}{4} + \frac{c_4}{\log x} < 1 - \frac{c_2}{\log x}, \textit{ for } x > e^{4(c_2+c_4)},$$

*it is enough to show that any zero $\beta + \gamma i$ of $\zeta$ with $\gamma$ sufficiently large (specifically, $\gamma > 8$) and that satisfies*

$$\beta > \frac{3}{4} + \frac{c_4}{\log \gamma},$$

*also satisfies*

$$\beta < 1 - \frac{c_2}{\log \gamma}.$$

*Let*

$$\sigma_0 = \sigma(\gamma) = 1 + \frac{c_4}{\log \gamma},$$

*and suppose that $\beta+\gamma i$ is a zero of $\zeta$ for which $\gamma > e^{1+c_2+c_4}$ and $\beta > \sigma_0 - 1/4$. We apply Theorem 29 twice, once with $s_0 = \sigma_0 + \gamma i$, and another time with*

*$s_0 = \sigma_0 + 2\gamma i$. In both cases, since $\sigma_0 > 1$, we have the circle $|s - s_0| \leq 1/2$ is contained in the quadrant $\sigma \geq 1/2$, $t \geq c_3$ for $\gamma \geq c_3 + 1/2$. Since $\gamma > e^{c_4}$, we get $\sigma_0 < 2$, and by the inverse of zeta identity, we have*

$$\left|\frac{1}{\zeta(s_0)}\right| \leq \sum_{n=1}^{\infty} \frac{1}{n^{\sigma_0}} < 1 + \int_1^{\infty} \frac{du}{u^{\sigma_0}} = 1 + \frac{1}{\sigma_0^{-1}} < \frac{2}{\sigma_0^{-1}} = \frac{2}{c_4} \log \gamma.$$

*Therefore, for each $\varepsilon_1 > 0$, there exists a $c_5$ such that for $\gamma > c_5 > c_3 + 1/2$; now the inequality*

$$\left|\frac{\zeta(s)}{\zeta(s_0)}\right| \leq \frac{2}{c_4}\left(2\gamma + \frac{1}{2}\right) \log \gamma < \gamma^{1+\varepsilon_1}$$

*holds at every point $s$ of the disk $|s - s_0| \leq 1/2$, since we have that at every point, $c_3 < t \leq 2\gamma + 1/2$. If $\gamma \geq c_5$, we can apply Theorem 29 with $r = 1/2$, $f(s) = \zeta(s)$, and $M = (1 + \varepsilon_1) \log \gamma$. We use the first inequality of that theorem with $s_0 = \sigma_0 + 2\gamma i$ to obtain*

$$-\operatorname{Re} \frac{\zeta'}{\zeta}(\sigma_0 + 2\gamma i) < 8(1 + \varepsilon_1) \log \gamma. \tag{2.30}$$

*With the second inequality of Theorem 29, we let $s_0 = \sigma_0 + \gamma i$ to get*

$$-\operatorname{Re} \frac{\zeta'}{\zeta}(\sigma_0 + \gamma i) < 8(1 + \varepsilon_1) \log \gamma - \frac{1}{\sigma_0 - \beta}, \tag{2.31}$$

*since*

$$\sigma - \frac{r}{2} = \sigma_0 - \frac{1}{4} < \beta \leq 1 < \sigma_0.$$

*Finally, because $\sigma_0 \to 1^+$ as $t \to \infty$, we get from (2.28) that for $\varepsilon_2 > 0$,*

$$-\frac{\zeta'}{\zeta}(\sigma_0) < \frac{1 + \varepsilon_2}{\sigma_0 - 1} = \frac{1 + \varepsilon_2}{c_4} \log \gamma \tag{2.32}$$

*for $\gamma > c_6$. Using (2.30), (2.31), and (2.32) in Theorem 31 yields*

$$\frac{3(1 + \varepsilon_2)}{c_4} \log \gamma + 4 \cdot 8(1 + \varepsilon_1) \log \gamma - \frac{4}{\sigma_0 - \beta} + 8(1 + \varepsilon_1) \log \gamma \geq 0,$$

*which is quickly simplified to*

$$\sigma_0 - \beta > \frac{c_7}{\log \gamma},$$

*where*

$$c_7 = \frac{4c_4}{3(1 + \varepsilon_2) + 40(1 + \varepsilon_1)c_4},$$

*and this gives us*

$$\beta < 1 - \frac{c_7 - c_4}{\log \gamma}.$$

*Clearly, $c_7 > c_4$ if $\varepsilon_1 < 1/3$, and $c_4$ small, and we can take $c_2 = c_7 - c_4$ and $c_1 = \max(c_5, c_6)$.*

**Theorem 34.** *If* $0 < c_8 < c_2$, *then*

$$|\log \zeta(s)| < \log^2 t \textit{ for } t > c_9 \textit{ and } \sigma \geq 1 - \frac{c_8}{\log t}.$$

**Proof.** *We utilize Theorem 28, concerning bounds on a more general function, with* $s_0 = 2 + t_0 i$, *for some* $t_0 > 8$ *to be determined. For* $t$ *sufficiently large, the region*

$$|s - s_0| \leq 1 + \frac{\frac{1}{2}(c_2 + c_8)}{\log t_0} \tag{2.33}$$

*is contained entirely in the region described in the previous theorem, in which* $\zeta$ *has no zeros. Thus,* $\log \zeta(s)$ *is analytic in this disk, and by Theorem 32 (II),*

$$\begin{aligned} \operatorname{Re} \log \zeta(s) = \log |\zeta(s)| &< \log(c \log t) \\ &< \log\left(c \log(t_0 + 2)\right) < c_{10} \log \log t_0. \end{aligned}$$

*Thus, by Theorem 28, we have for* $s$ *in the region described in* (2.33),

$$\begin{aligned} |\log \zeta(s)| &\leq |\zeta(s_0)| + \frac{2 \cdot 2\left(c_{10} \log \log t_0 + |\zeta(s_0)|\right)}{\frac{c_8 - c_2}{2} \cdot \frac{1}{\log t_0}} \\ &\leq c + c \log t_0 \log \log t_0 < \log^2 t_0, \end{aligned}$$

*if* $t_0$ *is sufficiently large. This inequality holds on the radius extending toward the left from* $s_0$, *for every large* $t_0$, *and thus throughout the region* $t \geq c_9$, $1 - c_8(\log t)^{-1} \leq \sigma \leq 2$. *Finally, both* $|\zeta(s)|$ *and* $|1/\zeta(s)|$ *are bounded in the half-plane* $\sigma > 2$, *and* $|\log \zeta(s)|$ *is therefore smaller than* $\log^2 t$ *for* $t$ *large and* $\sigma > 2$.

**Theorem 35.** *There exists a constant* $\alpha > 0$ *such that as* $x \to \infty$,

$$\sum_{p \leq x} \log \frac{x}{p} = \int_c^1 \frac{x^s}{s^2} ds + O\left(x e^{-\alpha\sqrt{\log x}}\right)$$

*for some* $c$ *such that* $0 < c < 1$.

**Proof.** *We use Theorem 30, which gives us*

$$\begin{aligned}
\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^2}\log\zeta(s)ds &= \frac{1}{2\pi i}\int_{(2)}\frac{1}{s^2}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n}\left(\frac{x}{n}\right)^s ds\\
&= \frac{1}{2\pi i}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n}\int_{(2)}\frac{(x/n)^s}{s^2}ds\\
&= \sum_{n\le x}\frac{\Lambda(n)}{\log n}\log\frac{x}{n}\\
&= \sum_{\substack{m,p\\ p^m\le x}}\frac{1}{m}\log\frac{x}{p^m}\\
&= \sum_{p\le x}\log\frac{x}{p}+\sum_{\substack{m,p\\ m\ge 2\\ p^m\le x}}\frac{1}{m}\log\frac{x}{p^m}.
\end{aligned}$$

*Clearly, the number of terms in the last sum is*

$$\pi(x^{1/2})+\pi(x^{1/3})+\cdots < x^{1/2}+x^{1/3}+\cdots+x^{1/u} < ux^{1/2},$$

*where $u$ is the smallest integer such that $x^{1/u}<2$. A word of caution here: the number of terms in the last sum is not the Riemann prime counting function, frequently denoted $\Pi(x)$ or $J(x)$, although both the subject and Riemann's influence on the subject may lead one to believe (or hope) that this is the case. In any case, we have*

$$\pi(x^{1/2})+\pi(x^{1/3})+\cdots = O\left(x^{1/2}\log x\right),$$

*so we obtain*

$$2\pi i\sum_{p\le x}\log\frac{x}{p}=\int_{(2)}\frac{x^s}{s^2}\log\zeta(s)ds+O\left(xe^{-\sqrt{\log x}}\right),$$

*since*

$$\sum_{\substack{m\ge 2\\ p^m\le x}}\frac{1}{m}\log\frac{x}{p^m}\le\sum_{\substack{m\ge 2\\ p^m\le x}}\log x = O\left(\sqrt{x}\log^2 x\right)=O\left(xe^{-\sqrt{\log x}}\right).$$

*Now we cut the complex plane along the real axis, continuing the cut from $s=1$ to the left, and consider the function $\log\zeta(s)$ in the cut plane. Now, if $\bar{z}$ denotes the complex conjugate of $z$, then $\zeta(\bar{z})=\overline{\zeta(s)}$ and $\log\bar{z}=\overline{\log z}$, so clearly $\log\zeta(\bar{s})=\overline{\log\zeta(s)}$.*

*Then, by Theorem 33, $\zeta(s)\ne 0$ for $|t|>c_9>c_1$ and $\sigma\ge 1-c_8(\log|t|)^{-1}$. Furthermore, since $\zeta(s)$ does not vanish on $\sigma=1$, and since its zeros have*

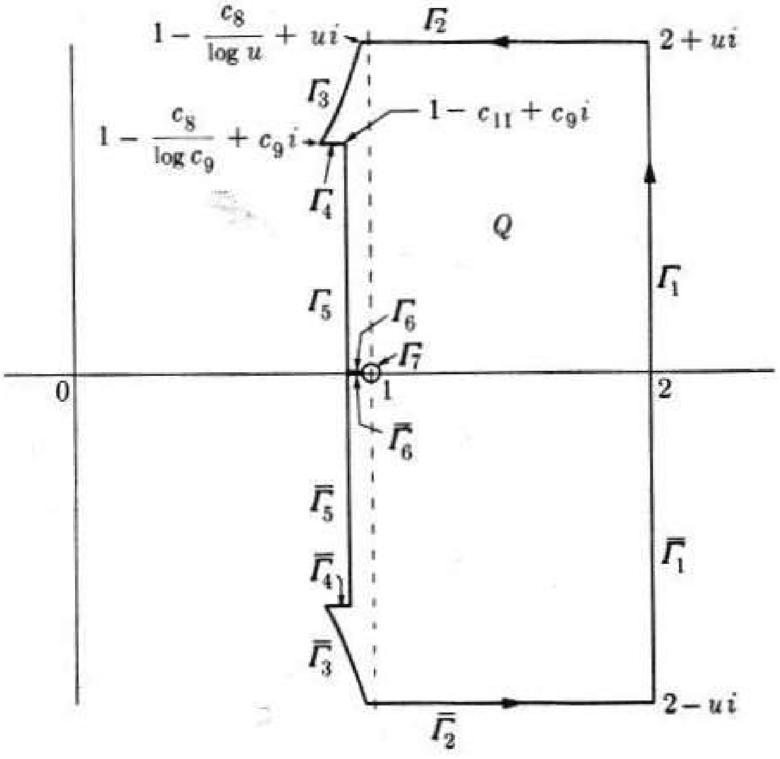


Figure 2.2: [LeV2], page 243

*no finite limit point in any half-plane $\sigma \geq \sigma_0 > 0$ (due to $\zeta(s)$ being analytic there), there exists a constant $c_{11} > 0$ such that $\zeta(s) \neq 0$ in the rectangle*

$$1 - c_{11} \leq \sigma \leq 1,\ |t| \leq c_9.$$

*Finally, $\zeta(s) \neq 0$ for $1 \leq \sigma \leq 2$, and the only singularity in the half-plane $\sigma > 0$ occurs at $s = 1$. Therefore, for an arbitrary $u > c_9$, $\log \zeta(s)$ is a single-valued analytic function in the region $Q$ outlined by Figure 2.2 bounded by $\Gamma_1, \Gamma_2, \ldots, \Gamma_6, \Gamma_7, \bar{\Gamma}_6, \bar{\Gamma}_5, \ldots, \bar{\Gamma}_1$. Now let $\Gamma$ denote the total boundary of $Q$, i.e. $\Gamma = \Gamma_1 + \cdots + \bar{\Gamma}_1$, then by Cauchy's theorem we have the following:*

$$\int_\Gamma \frac{x^s}{s^2} \log \zeta(s) ds = 0.$$

*It follows, by taking the integrals in the positive direction, that*

$$\begin{aligned}&\int_{(2)} \frac{x^s}{s^2}\log\zeta(s)ds\\ &\quad= \left(\int_{2-\infty i}^{2-ui} + \int_{\bar{\Gamma}_1} + \int_{\Gamma_1} + \int_{2+ui}^{2+\infty i}\right)\frac{x^s}{s^2}\log\zeta(s)ds\\ &\quad= \left(\int_{2-\infty i}^{2-ui} - \int_{\Gamma_2+\cdots\Gamma_6+\Gamma_7+\bar{\Gamma}_6+\cdots+\bar{\Gamma}_2} + \int_{2+ui}^{2+\infty i}\right)\frac{x^s}{s^2}\log\zeta(s)ds.\end{aligned}$$

*Now it remains for us to show that all of these integrals, with the exception of $\Gamma_6$ and $\bar{\Gamma}_6$ are small for $u$ large.*

*So we utilize Theorem 34. For $u > u_0(\varepsilon)$, we have*

$$\begin{aligned}\left|\int_{2+ui}^{2+\infty i}\frac{x^s}{s^2}\log\zeta(s)ds\right| &\le \int_{2+ui}^{2+\infty i}\frac{x^2}{|s|^2}|\log\zeta(s)||ds|\\ &\le x^2\int_u^\infty \frac{\log^2 t}{t^2}dt\\ &\le x^2\int_u^\infty \frac{dt}{t^{2-\varepsilon}}\\ &< \frac{cx^2}{u^{1-\varepsilon}}\end{aligned}$$

*so that*

$$\lim_{u\to\infty}\int_{2+ui}^{2+\infty i}\frac{x^s}{s^2}\log\zeta(s)ds = 0.$$

*Furthermore, this same estimate applies if we replace the "+" in the limits of integration with a "−" and interchange top and bottom, i.e. interchange $2+ui$ with $2-\infty i$ and $2+\infty i$ with $2-ui$.*

*For $\Gamma_2$, we have that the length of this arc is less than 2, and the integrand is still smaller than $x^2\log^2 u/u^2$ for $u$ large, so we have*

$$\lim_{u\to\infty}\int_{\Gamma_2}\frac{x^s}{s^2}\log\zeta(s)ds = 0,$$

*and the same applies to $\bar{\Gamma}_2$:*

$$\lim_{u\to\infty}\int_{\bar{\Gamma}_2}\frac{x^s}{s^2}\log\zeta(s)ds = 0.$$

*As for $\Gamma_3$ and its related arc, we have $s = 1 - c_8(\log t)^{-1} + ti$, so that*

$$\left|\int_{\Gamma_3}\frac{x^s}{s^2}\log\zeta(s)ds\right| \le \int_{c_9}^{u}\frac{x^{1-c_8(\log t)^{-1}}}{t^2}\log^2 t\left|\frac{c_8}{t\log^2 t}+i\right|dt.$$

*Now suppose that $x$ and then $u$ are chosen large such that*

$$c_9 < e^{\sqrt{2c_8\log x}} < u.$$

*Then*

$$\begin{aligned}
&\int_{\Gamma_3} \frac{x^s}{s^2} \log \zeta(s) ds \\
&\quad = O\left( \int_{c_9}^{e^{\sqrt{2c_8 \log x}}} x \cdot x^{-c_8(2c_8 \log x)^{-1/2}} \frac{\log^2 t}{t^2} dt + x \int_{e^{\sqrt{2c_8 \log x}}}^{u} \frac{\log^2 t}{t^{1/2} \cdot t^{3/2}} dt \right) \\
&\quad = O\left( x e^{-\sqrt{\frac{1}{2} c_8 \log x}} \int_{c_9}^{\infty} \frac{\log^2 t}{t^2} dt + \frac{x}{e^{\sqrt{\frac{1}{2} c_8 \log x}}} \cdot \int_{e^{\sqrt{2c_8 \log x}}}^{u} \frac{\log^2 t dt}{t^{3/2}} \right) \\
&\quad = O\left( x e^{-\alpha \sqrt{\log x}} \right),
\end{aligned}$$

*where* $\alpha = \sqrt{c_8/2}$.

*By symmetry, we have the same result for* $\bar{\Gamma}_3$,

$$\int_{\bar{\Gamma}_3} \frac{x^s}{s^2} \log \zeta(s) ds = O\left( x e^{-\alpha \sqrt{\log x}} \right).$$

*Now* $\Gamma_4$, $\Gamma_5$, $\bar{\Gamma}_4$, *and* $\bar{\Gamma}_5$ *are all of fixed lengths, and on these paths we have*

$$\frac{x^s}{s^2} \log \zeta(s) = O\left( x^{1-c_{11}} \right) = o\left( x e^{-\alpha \sqrt{\log x}} \right).$$

*Since these paths are of fixed lengths, we have that the same estimate holds for the integrals as well.*

$\Gamma_7$ *is described by the relations* $s = 1 + \delta e^{i\theta}$, $|\theta| \leq \pi$, *where* $\delta > 0$. *Since* $(s-1)\zeta(s) \to 1$ *as* $s \to 1$, *we have*

$$\begin{aligned}
&\operatorname{Re} \log \zeta(s) = \log |\zeta(s)| \sim -\log |s-1| = -\log \delta, \\
&\operatorname{Im} \log \zeta(s) = \arg \zeta(s) = O(1)
\end{aligned}$$

*as* $\delta \to 0^+$. *Therefore,*

$$\int_{\Gamma_7} \frac{x^s}{s^2} \log \zeta(s) ds = O\left( 2\pi\delta \frac{x^{1+\delta}}{(1-\delta)^2} \log \delta \right) = o(1).$$

*Combining all stated results in this proof, we take the limit as* $u \to \infty$ *and* $\delta \to 0^+$ *and obtain*

$$2\pi i \sum_{p \leq x} \log \frac{x}{p} = \int_{1-c_{11}}^{1} \frac{x^s}{s^2} \log \zeta(s) ds + \int_{1}^{1-c_{11}} \frac{x^s}{s^2} \log \zeta(s) ds + o\left( x e^{-\alpha \sqrt{\log x}} \right),$$

*where we have the first integral to be along the top edge of the cut and the second integral to be along the bottom edge, i.e.* $\Gamma_6$ *and* $\bar{\Gamma}_6$ *respectively. We know that* $(1-s)\zeta(s)$ *is analytic in the half-plane* $\sigma > 0$, *and that it has no zeros in the region* $\sigma > 1 - c_{11}$, $|t| < c_9$. *Therefore*

$$\log\left( (s-1)\zeta(s) \right) = \log(s-1) + \log \zeta(s)$$

*is single-valued in this region. Furthermore, since* $\log(s-1)$ *has values which differ by* $2\pi i$ *on the upper and lower edges of the cut, the same must be true of* $\log\zeta(s)$*, so long as the difference is taken in the reverse order. Now, let* $s^+$ *indicate the upper edge of the cut, and* $s^-$ *indicate the lower edge, then we have*

$$\begin{aligned}\int_{1-c_{11}}^{1} \frac{x^{s^+}}{(s^+)^2}\log\zeta(s^+)ds^+ + \int_{1}^{1-c_{11}} \frac{x^{s^-}}{(s^-)^2}\log\zeta(s^-)ds^- \\ = \int_{1-c_{11}}^{1} \frac{x^{s^+}}{(s^+)^2}\log\zeta(s^+)ds^+ - \int_{1-c_{11}}^{1} \frac{x^{s^+}}{(s^+)^2}\left(\log\zeta(s^+) - 2\pi i\right)ds^+ \\ = 2\pi i\int_{1-c_{11}}^{1} \frac{x^s}{s^2}ds,\end{aligned}$$

*and*

$$\sum_{p\le x}\log\frac{x}{p} = \int_{1-c_{11}}^{1}\frac{x^s}{s^2}ds + O\left(xe^{-\alpha\sqrt{\log x}}\right), \tag{2.34}$$

,

*which was to be shown.*

At this point, we have covered the necessary preliminary material, but we would be remiss if we came this far to stop short of the Prime Number Theorem. In fact, we prove a stronger result:

**Theorem 36.** *As* $x \to \infty$

$$\pi(x) = \int_2^x \frac{du}{\log u} + O\left(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}\right).$$

**Proof.** *We replace* $1-c_{11}$ *with* $C$ *in equation* (2.34)*, and let*

$$\delta = \delta(x) = e^{-\frac{1}{2}\alpha\sqrt{\log x}}.$$

*Now, since* $\log(1+\delta) \sim \delta$ *as* $x \to \infty$*, we get*

$$\begin{aligned}\sum_{p\le x(1+\delta)}\log\frac{x(1+\delta)}{}p - \sum_{p\le x}\log\frac{x}{p} \\ = \sum_{p\le x}\log(1+\delta) + \sum_{x<p\le x(1+\delta)}\log\frac{x(1+\delta)}{p} \\ = \log(1+\delta)\pi(x) + O\left(\log(1+\delta)\cdot\delta x\right) \\ = \int_C^1 \frac{x^s}{s^2}\left((1+\delta)^s - 1\right)ds + O\left(xe^{-\alpha\sqrt{\log x}}\right),\end{aligned}$$

*thus, by solving for* $\pi(x)$ *in the last two parts, we obtain*

$$\pi(x) = \frac{1}{\log(1+\delta)}\int_C^1 \frac{(1+\delta)^s - 1}{s^2}x^s ds + O(\delta x) + O\left(\frac{xe^{-\alpha\sqrt{\log x}}}{\delta}\right).$$

*Note*

$$(1+\delta)^s - 1 = s\delta + \frac{s(s-1)}{2!}(1+\vartheta\delta)^{s-2}\delta^2,$$

*where* $0 < \vartheta < 1$. *So now for* $0 < s < 1$,

$$|(1+\delta)^s - 1 - s\delta| \leq \frac{s|s-1|}{2}\delta^2 < \delta^2.$$

*Now we simply make a change of variables, let* $x^s = u$, *and we get the following:*

$$\begin{aligned}
\int_C^1 \frac{(1+\delta)^2 - 1}{s^2} x^s ds &= \delta \int_C^1 \frac{x^s}{s} ds + O\left(\delta^2 \int_C^1 \frac{x^s}{s^2} ds\right) \\
&= \delta \int_{x^C}^x \frac{du}{\log u} + O\left(\delta^2 \int_C^1 x^s ds\right) \\
&= \delta \int_2^x \frac{du}{\log u} + O\left(\delta^2 x\right).
\end{aligned}$$

*Finally, we have enough to make the last computations,*

$$\begin{aligned}
\pi(x) &= \frac{\delta}{\log(1+\delta)} \int_2^x \frac{du}{\log u} + O(\delta x) + O\left(\frac{xe^{-\alpha\sqrt{\log x}}}{\delta}\right) \\
&= (1 + O(\delta)) \int_2^x \frac{du}{\log u} + O(\delta x) + O\left(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}\right) \\
&= \int_2^x \frac{du}{\log u} + O(\delta x) \\
&= \int_2^x \frac{du}{\log u} + O\left(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}\right),
\end{aligned}$$

*which was to be shown.*

Now we have the Prime Number Theorem as a weak consequence of Theorem 36, since

$$\int_2^x \frac{du}{\log u} = \frac{u}{\log u}\bigg]_2^x + \int_2^x \frac{du}{\log^2 u} \sim \frac{x}{\log x}$$

and

$$xe^{-c\sqrt{\log x}} = o\left(\frac{x}{\log x}\right)$$

for every $c > 0$. Actually, if we apply repeated integration by parts, we get that

$$\pi(x) = \frac{x}{\log x} + \frac{2!x}{\log^2 x} + \cdots + \frac{m!x}{\log^m x} + o\left(\frac{x}{\log^m x}\right)$$

is true for every positive integer $m$.

# Chapter 3

# Geometrical Representation

This chapter focuses on illustrating the problem geometrically. Once we do so, the problem of figuring out the number of integers expressible as the sum of two squares can be understood by those possessing only a rudimentary knowledge of geometry.

We shall see that our problem easily relates to Pythagorean triangles and the number of lattice points within circles.

## 3.1 The Circle Problem of Gauss

The circle problem of Gauss is concerned with finding the number of lattice points within the circle

$$u^2 + v^2 = n.$$

Typically, most people do this by defining $r(n)$ as the number of lattice points on the circle, seeing that

$$\sum_{n \le x} r(n) = r(1) + r(2) + ... + r(\lfloor x \rfloor),$$

and realizing that this problem is concerned with the average order of the function $r(n)$.

It has been shown in many texts that

$$\sum_{n \le x} r(n) = \pi x + O(x^{\frac{1}{2}}).$$

That is, that the number of lattice points inside the circle is approximately its area with an error term on the order of its circumference.

This is almost intuitive and will not be proven here (for the interested reader, see [HarWr], pages 270-271), but what does this interesting relation have to do with our problem?

Consider $d$, the distance from any lattice point, $(u, v)$ to the origin. Then $d = u^2 + v^2$ is a number representable as the sum of two squares! Now the question as to how many $d$'s are there less than a given number $x$ becomes known as exactly the number of circles less than the square root of $x$ that contain lattice points!

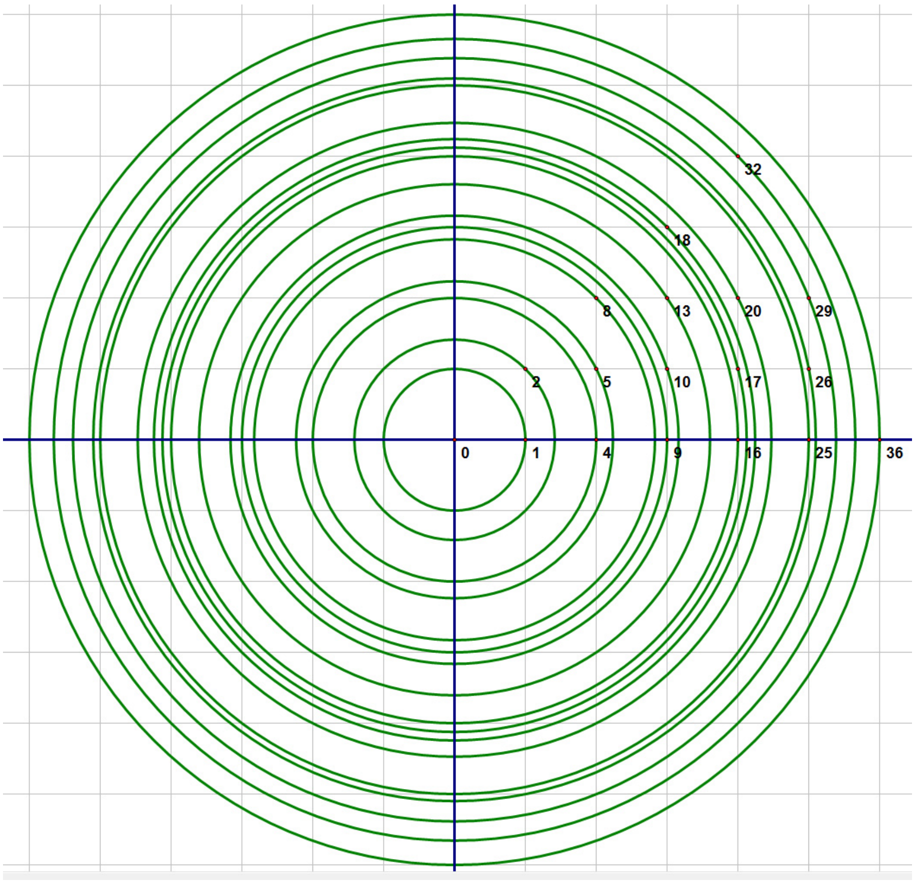


Figure 3.1: Number of integers expressible as sum of two squares less than or equal to 36

Each circle represents a new $d$ and the problem begins to resemble something relating to Pythagorean's theorem.

We have gone from one classical problem (the circle problem of Gauss) to another (numbers representable as the sum of two squares), but the transition is not without its drawbacks.

## 3.2 Difficulties with Geometrical Representation

One of the biggest problems with our transition is that we have lost the main tool mathematicians have used to attack the circle problem of Gauss, i.e. lattice points.

Although it is clear that to continue to work with lattice points would be cumbersome (take for instance, $x = 1$, then there are two circles (corresponding to 0 and 1, both of which have obvious decompositions into the sum of two squares), but there are five lattice points; at $x = 4$ the number of lattice points jumps to 13, and at $x = 36$ we have over 80 lattice points!), the very distribution of the circles is the problem we were trying to tackle in the first place.

The next section illustrates some of the work done continuing on with lattice points, recognizing that the number of lattice points on a particular circle is exactly the number of representations of that number as the sum of two squares. For $x > 0$, each circle will contribute at least 4 lattice points.

Finally, although it is true that the formula given for the average order of $r(n)$ is very clear-cut and has the wonderful relation to area and circumference of the circle, this is only the tip of the iceberg, so to speak, and the real meat of the circle problem of Gauss is concerned with finding $\theta$, the smallest value of $\xi$ such that

$$\sum_{n\leq x} r(x) = \pi x + O(x^{\xi+\varepsilon}).$$

It has been shown that $\frac{1}{4} \leq \theta \leq \frac{1}{3}$, numbers that have no nice relation to the circumference of our circles.

As a last note about this problem, many mathematicians attempt to delve into higher-dimensional figures, obtaining some rather nice results, but this has no bearing on our current subject.

## 3.3 Other Work Relating to the Lattice Points in a Circle

As mentioned in the previous section, each circle (after the trivial one) contributes at least 4 lattice points; e.g. If $x = 2$, then $2 = 1^2 + 1^2 = (-1)^2 + 1^2 = (-1)^2 + (-1)^2 = 1^2 + (-1)^2$. The question becomes then, once we know a number is representable as the sum of two squares, is there a way to unlock the number of representations?

Most work in this area uses the notation $r(n)$ to be the arithmetical counting function for the number of representations of $n$. I will also make use of the notation $r^*(n)$ to indicate the number of unique representations up to order and sign. E.g.

$$r(0) = 1,\ r^*(0) = 1$$
$$r(2) = 4,\ r^*(2) = 1$$
$$r(25) = 12,\ r^*(25) = 2.$$

With this notation, we can describe the number of lattice points as the summatory function associated with $r$, i.e.

$$N(x) = r(0) + r(1) + \cdots + r(\lfloor x \rfloor) = \sum_{0 \le n \le x} r(x),$$

and as before, we know that this function is on the order of the area of the circle with an error term the size of the circumference, i.e.

$$\sum_{n \le x} = \pi x + O(\sqrt{x}).$$

Now we turn to [Sier] to try to get a handle on the number of representations of a number $n$ such that $n = x^2 + y^2$. If $n$ is such a number, then $n \ge x^2$ and $n \ge y^2$, and so $|x| \le \sqrt{n}$ and $|y| \le \sqrt{n}$. So similar to when we learned to check primality in grade school, we need to go no further than $\sqrt{n}$. For $x$, we substitute integers whose values are not greater than $\sqrt{n}$, and solve for $y$. That is,

$$y^2 = n - x^2.$$

Hence, if we obtain a square, we obtain two representations (since order is important when computing the decomposition in this manner). We give the following example:

Let $n = 20$. We form the sequence, $20, 19, 16, 11, 4$. We know the second and fourth term of this sequence are squares, so $x = \pm 2$, $y = \pm 4$, or $x = \pm 4$ and $y = \pm 2$. Thus, 20 has eight decompositions. They are

$$\begin{aligned} 20 &= 2^2 + 4^2 = 2^2 + (-4)^2 = (-2)^2 + 4^2 = (-2)^2 + (-4)^2 \\ &= 4^2 + 2^2 = 4^2 + (-2)^2 = (-4)^2 + 2^2 = (-4)^2 + (-2)^2. \end{aligned}$$

Yet, $n = 20$ has only one unique decomposition up to order and sign.

We will explore lattice points more in Appendix B, after we have explored the type of numbers that are representable as the sum of two squares. Figuring out just what numbers generate these circles is the subject of the next chapter.

# Chapter 4

# Numerical Candidates

This chapter is concerned with discerning what numbers are representable as the sum of two squares. In the vein of Chapter 3, this is equivalent to discovering the squares of the radii of circles that contain lattice points. In the beginning, we are concerned only with what primes are representable, but we will see that these primes are wholly responsible for figuring out which composite numbers have representations as the sum of two squares.

We split the primes into three distinct subgroups that contain all primes, 2, primes of the form $4k + 1$, and primes of the form $4k + 3$ or $4k - 1$. For primes of the first two types we have a positive answer concerning the existence of a representation as the sum of two squares, but primes of the form $4k + 3$ tend to be the ugly stepbrother of the others (at least when it comes to being representable as the sum of two squares).

### 2

**Theorem 37.** *2 is representable as the sum of two squares.*

**Proof.** $2 = 1^2 + 1^2$.

## 4.1 Primes of the form $4k + 1$

Fermat was the first to claim a proof that every prime $p = 4k + 1$ is representable as the sum of two squares, and this theorem is sometimes called, Fermat's Christmas theorem, because of this.

**Theorem 38 (Fermat's Christmas Theorem).** *Every prime $p = 4k+1$, $k$ an integer, is representable as the sum of two squares.*

**Proof.** *This proof can be found in [Sier]. Let $p$ be a prime number of the form $4k + 1$ and $a = \left(\frac{p-1}{2}\right)!$. We have $p|a^2 + 1$, with $(a, p) = 1$ due to Theorem 3. Now according to Thue's theorem with $m = p$ in our case, we* *have that there exist two natural numbers $x, y$, with each $\leq \sqrt{p}$, such that for a suitable choice of sign, $ax \pm y$ is divisible by $p$. Hence, it follows that $a^2x^2 - y^2 = (ax - y)(ax + y)$ is divisible by $p$*

*$a^2x^2 + x^2 = (a^2 + 1)x^2$ is divisible by $p$ since $p|a^2 + 1$. Therefore, the number $x^2 + y^2 = a^2x^2 + x^2 - (a^2x^2 - y^2)$ is divisible by $p$. We have that $x, y$ are natural numbers $\leq \sqrt{p}$, so $x, y < \sqrt{p}$. Consequently, $x^2 + y^2$ is a natural number, $1 < x^2 + y^2 < 2p$, and is not divisible by $p$. So $p = x^2 + y^2$, proving that $p$ is the sum of two squares of natural numbers.*

This theorem is the first key in unlocking the puzzle of numbers representable as the sum of two squares, and it also opens up many other avenues of interest. For a few of these, please refer to Appendix 2.

## 4.2 Primes of the form $4k + 3$

**Theorem 39.** *No prime $p$ of the form $4k + 3$, $k$ an integer, is expressible as the sum of two squares.*

**Proof.** *This proof is relatively simple to demonstrate completely. Let $n$ be an integer. Then there are four possibilities.*

**(Case 1)** *If $n \equiv 0 \pmod 4$, then $n^2 \equiv 0 \pmod 4$.*

**(Case 2)** *If $n \equiv 1 \pmod 4$, then $n^2 \equiv 1 \pmod 4$.*

**(Case 3)** *If $n \equiv 2 \pmod 4$, then $n^2 \equiv 0 \pmod 4$.*

**(Case 4)** *If $n \equiv 3 \pmod 4$, then $n^2 \equiv 1 \pmod 4$.*

*In any case, the square of an integer is congruent to $0$ or $1$ modulo $4$. Thus, the sum of two squares of integers can only be congruent to $0, 1$, or $2$ modulo $4$. Therefore, no prime $p$ (or any integer for that matter) congruent to $3$ modulo $4$, can be the sum of two squares.*

## 4.3 The Numerical Winners

**Theorem 40.** *A positive integer $n$ has a representation as the sum of two squares if and only if each prime factor congruent to $3 \pmod 4$ occurs with even multiplicity.*

**Proof.** *Let $q \equiv 3 \pmod 4$ be prime, and let $q^\alpha || n$, where $n = x^2 + y^2$. We will show that $\alpha$ is even.*

*Let $d = (x, y)$ and set $x = dx_0, y = dy_0$. Then $(x_0, y_0) = 1$, and $n = d^2(x_0^2 + y_0^2)$. Let $\alpha$ be odd, then $x^2 + y^2 \equiv 0 \pmod q$. It follows that $q \nmid y_0$, since otherwise $q|x_0$, which is absurd since $(x_0, y_0) = 1$. Therefore,* *there must exist* $z_0$ *such that* $y_0 z_0 \equiv 1 \pmod q$*; for instance, we could take* $z_0 = y_0^{p-2}$ *by Fermat's Little theorem. Thus,*

$$(x_0 z_0)^2 + 1 \equiv (x_0^2 + y_0^2) z_0^2 \equiv 0 \pmod q,$$

*which is absurd, since* $-1$ *is not a quadratic residue for* $q \equiv 3 \pmod 4$*. So if* $q^\alpha || n$*, then* $\alpha$ *is even.*

*Next, we show that if each prime* $q \equiv 3 \pmod 4$ *in the factorization of* $n$ *occurs to an even power, then* $n$ *is expressible as the sum of two squares.*

*We have the identity*

$$\left(x^2 + y^2\right)\left(u^2 + v^2\right) = (xu - yv)^2 + (xv + yu)^2$$

*which shows that the product of two integers expressible as the sum of two squares is itself expressible as a sum of two squares.*

*Clearly,* $2$ *is the sum of two squares, and any square of an integer is as well, i.e.* $m^2 = m^2 + 0^2$*, finally we have already shown that for all primes* $p \equiv 1 \pmod 4$*,* $p$ *is the sum of two squares.*

*Thus, all natural numbers* $n$ *that can be expressed as the sum of two squares are of the form*

$$n = 2^\alpha P Q^2$$

*where* $0 \leq \alpha$*,* $P$ *is the product of primes congruent to* $1 \pmod 4$ *and* $Q$ *is the product of primes congruent to* $3 \pmod 4$*.*

# Chapter 5

# The Number of Integers Expressible as the Sum of Two Squares

Let $B(x)$ be the counting function for the number of integers expressible as the sum of two squares. That is, for $u, v \in \mathbb{Z}$

$$B(x) := \#\{n \leq x : n = u^2 + v^2\}.$$

In other words, if we have $b_n$ as the characteristic function for numbers representable as the sum of two squares, i.e.

$$b_n = \begin{cases} 1 \text{ if } n = x^2 + y^2 \text{ for some integers } x, y, \\ 0 \text{ else,} \end{cases}$$

then

$$B(x) = \sum_{n \leq x} b_n.$$

## 5.1 A Heuristic Argument

The following heuristic argument, given in [LeV2], will lead us to believe (correctly) that

$$B(x) \sim \frac{\beta x}{\sqrt{\log x}},$$

for some positive real, $\beta$. Although not a formal proof, it contains some interesting mathematics in and of itself.

To begin with, take $x$ very large. Then one out of every $p$ integers is divisible by $p$, so that the number of integers *not* divisible by $p$ is about $x - x/p$ or $x\left(1 - \frac{1}{p}\right)$.

Therefore, the number of integers not divisible by any $p \leq \sqrt{x}$ is about

$$x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right).$$

Now, by the Prime Number Theorem, the number of integers not divisible by any $p \leq \sqrt{x}$ is

$$x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) \approx \frac{x}{\log x}.$$

This gives us the number of integers not divisible by any $p$ in a sequence.

But to count up $B(x)$, we do not want to eliminate every composite number, as explained in Theorem 4.2, we wish only to eliminate integers divisible by an odd power of any prime $q \equiv 3 \pmod 4$.

This can be accomplished using the cross-classification principle: basically, we eliminate all integers divisible by $q$, then bring back those divisible by $q^2$ that we previously eliminated, and proceeding on in like fashion (eliminating odd powers and bringing back even powers). This yields the following:

$$x\left(1 - \frac{1}{q}\right)\left(1 + \frac{1}{q^2}\right)\left(1 - \frac{1}{q^3}\right)\cdots$$

as the number left after accounting for the prime $q$. Note that the product is finite since we do not proceed beyond $\sqrt{x}$.

Now we eliminate all primes $q \equiv 3 \pmod 4$ for an estimation of $B(x)$:

$$B(x) \approx x \prod_{q \leq \sqrt{x}} \left(1 - \frac{1}{q}\right) \prod_{q^2 \leq \sqrt{x}} \left(1 + \frac{1}{q^2}\right) \prod_{q^3 \leq \sqrt{x}} \left(1 - \frac{1}{q^3}\right) \cdots$$

But each product after the first converges as $x \to \infty$, so we get,

$$B(x) \approx x \prod_{q \leq \sqrt{x}} \left(1 - \frac{1}{q}\right).$$

It remains to get a handle on the product

$$x \prod_{q \leq \sqrt{x}} \left(1 - \frac{1}{q}\right).$$

For all primes $p$, it is true that

$$\log \prod_{p\leq\sqrt{x}} \left(1 - \frac{1}{p}\right) = \sum_{p\leq\sqrt{x}} \log \left(1 - \frac{1}{p}\right) \approx -\log(\log x).$$

And since, by Dirichlet's theorem, about half of all primes are 3 (mod 4), we have that

$$\log \prod_{q\leq\sqrt{x}} \left(1 - \frac{1}{q}\right) \approx -\frac{1}{2}\log(\log x) = \log\left(\frac{1}{\sqrt{\log x}}\right),$$

Which clearly gives us,

$$B(x) \approx x \prod_{q\leq\sqrt{x}} \left(1 - \frac{1}{q}\right) \approx \frac{x}{\sqrt{\log x}}.$$

This provides us very strong evidence to believe that, for some $\beta \in \mathbb{R}$

$$B(x) \sim \frac{\beta x}{\sqrt{\log x}}.$$

## 5.2 Estimate of $B(x)$

In the following subsections we examine two separate approaches to estimating the number of integers expressible as the sum of two squares. One approach makes use of the generalized Wiener-Ikehara theorem, while the other relies on contour integrals and the proof of the Prime Number Theorem. In this section we go through the first proof, but we only set up the second proof with comments on the asymptotic expansion of $B(x)$, leaving the proof in its entirety for the final section of the chapter.

### 5.2.1 Method 1 - Via the Generalized Wiener-Ikehara Theorem

This section makes use of the *Generalized Wiener-Ikehara Theorem* as described in the preliminaries and this is the method that [BatDia] makes use of to estimate $B(x)$.

Recall that this theorem says:

**Generalized Wiener-Ikehara Theorem** *Let $F$ be a real valued nondecreasing function in $\mathcal{V}$ with $\sigma_c\left(\widehat{F}\right) = \alpha > 0$. Let*

*phi and $\vartheta$ be functions which are analytic on the closed half-plane $\{s : \sigma \geq \alpha\}$ and assume that $\phi(\alpha) \neq 0$. For $\gamma$ a real number distinct from $0, -1, -2, ...,$ let $(s-\alpha)^{-\gamma}$ be positive valued on the real ray $\{s : s > \alpha\}$. Suppose that*

$$\widehat{F}(s) = \int x^{-s} dF(x) = (s-\alpha)^{-\gamma}\phi(s) + \vartheta(s)$$

*holds on the open half-plane* $\{s : \sigma > \alpha\}$*. Then*

$$F(x) \sim \frac{\phi(\alpha)x^{\alpha}(\log x)^{\gamma-1}}{\alpha\Gamma(\gamma)}.$$

*Where* $\Gamma$ *denotes the Euler gamma function.*

Now, let $k$ be a given integer, $k \geq 2$, and let $h = \varphi(k)$ (where $\varphi$ refers to the Euler totient function). Let $a_1, a_2, ..., a_h$ be a reduced residue system modulo $k$. Let $b_1, b_2, ..., b_h$ be real numbers in $[0, 1]$, not all 0, and let $f$ be a multiplicative function with the following properties:

(1) $f(p) := b_j$ if $p \equiv a_j \pmod{k}$, and

(2) $0 \leq f \leq 1$.

Let $\mathcal{B} := b_1 + b_2 + \cdot + b_h$. In the next lemma, given in [BatDia], we get an estimate for the summatory function of $f$, our first step towards finding an estimate for $B(x)$. Many of the intermediary results have already been shown in the preliminaries, but where possible, the author has aimed to provide separate proofs of certain results when they have differed from the proofs found in Appendix A. Also, comments relating these results back to those found in Chapter 2 are provided whenever necessary.

**Lemma 9.** *Let* $f$ *be as above. There exists a positive number* $c = c(f)$ *such that*

$$\sum_{n \leq x} f(n) \sim cx(\log x)^{(\mathcal{B}/h)-1}.$$

**Proof.** *Our goal is to apply the Generalized Wiener-Ikehara theorem to the Dirichlet series* $\widehat{F}(s) := \sum f(n)n^{-s}$*, but before we can do that, we must show that*

$$\widehat{F}(s) = (s-1)^{\mathcal{B}/h}F^*(s),$$

*where* $F^*(s)$ *is analytic on* $\{s : \sigma \geq 1\}$ *and* $F^*(1) \neq 0$*. Since* $f \geq 0$*, it follows that* $F^*(1) > 0$*.*

*Since* $f$ *is a multiplicative function by our premises,*

$$\widehat{F}(s) = \prod_{l=1}^{h} \prod_{p \equiv a_l(k)} \left(1 + \frac{b_l}{p^s} + \frac{f\left(p^2\right)}{p^{2s}} + \cdots\right) K \tag{5.1}$$

*with*

$$K := K(s) := \prod_{p|k} \left\{ 1 + \frac{f(p)}{p^s} + \frac{f\left(p^2\right)}{p^{2s}} + \cdots \right\}.$$

*Each factor of $K$ converges for $\operatorname{Re} s > 0$ and is positive at $s = 1$. Since $k$ has only a finite number of prime divisors, $K$ converges and defines an analytic function for $\operatorname{Re} s > 0$.*

*We express a general product:*

$$G(s)H(s) := \prod_{p \equiv a(k)} \left\{ 1 + bp^{-s} + f\left(p^2\right)p^{-2s} + \cdots \right\},$$

*where*

$$H(s) := H_{a,b}(s) := \prod_{p \equiv a(k)} \left\{ \left(1 - p^{-s}\right)^b \left( 1 + \frac{b}{p^s} + \frac{f\left(p^2\right)}{p^{2s}} + \cdots \right) \right\}$$

*and*

$$G(s) := G_{a,b}(s) := \prod_{p \equiv a(k)} \left(1 - p^{-s}\right)^{-b}.$$

*Given $p \equiv a \pmod{k}$ with $(a, k) = 1$, consider a single factor of $H$:*

$$\left(1 - p^{-s}\right)^b \left(1 + bp^{-s} + f\left(p^2\right)p^{-2s} + \cdots\right).$$

*The binomial expansion of $\left(1 - p^{-s}\right)^b$ has coefficients all of size at most 1, as are the coefficients of $b, f\left(p^2\right), ...,$ in the second factor. Formally multiplying, we see that each factor of $H$ is of the form*

$$1 + 0p^{-s} + 3\theta_2 p^{-2s} + 4\theta_3 p^{-3s} + \cdots,$$

*where each $|\theta_v| \leq 1$. Thus, $H(s) = \prod_{p \equiv a} \left\{1 + O\left(p^{-2\sigma}\right)\right\}$ converges uniformly for say, $\operatorname{Re} s > 2/3$, and is therefore analytic on this half-plane. Also, $H(1) > 0$, since each factor is positive.*

*Now let's take a closer look at $G$. We have*

$$\log G(s) = b \sum_{p \equiv a(k)} \log\left(1 - p^{-s}\right)^{-1} = b \sum_{p \equiv a(k)} \sum_{\alpha=1}^{\infty} \frac{p^{-\alpha s}}{\alpha}.$$

*As a comparison, consider $G^*$, where*

$$\begin{aligned}\log G^*(s) &:= \frac{b}{h}\sum_{\chi}\overline{\chi(a)}\log L(s,\chi)\\ &= \frac{b}{h}\sum_{\chi}\sum_{n=1}^{\infty}\kappa(n)n^{-s}\chi(n)\overline{\chi(a)}\\ &= b\sum_{n\equiv a(k)}\kappa(n)n^{-s}\\ &= b\sum_{p^\alpha\equiv a(k)}\frac{p^{-as}}{\alpha}.\end{aligned}$$

*This last sum extends over all primes p and positive integers $\alpha$ satisfying $p^\alpha \equiv a \pmod{k}$. Let*

$$b\phi_a(s) := \log G(s) - \log G^*(s),$$

*then this is a Dirichlet series with bounded coefficients extending over higher prime powers and thus is analytic for* $\operatorname{Re} s > 1/2$. *It follows that*

$$G(s) = \exp\{b\phi_a(s)\}\prod_{\chi} L(s,\chi)^{b\overline{\chi}(a)/h}.$$

*Now, for $\chi \neq \chi_1$, $L(s,\chi)$ is analytic and nonzero on the closed half-plane $\sigma \geq 1$, and thus all factors of $G_{a_i,b_i}$ except $L(s,\chi_1)^{b_i/h}$ have the same property.*

*For $\chi_1$, we simply multiply over all the reduced residue classes and obtain*

$$\begin{aligned}\prod_{i=1}^{h} L(s,\chi_1)^{b_i/h} &= L(s,\chi_1)^{\mathcal{B}/h}\\ &= (s-1)^{\mathcal{B}/h}\left\{(s-1)\zeta(s)\prod_{p|k}\left(1-p^{-s}\right)\right\}^{\mathcal{B}/h}.\end{aligned}$$

*Now, figuring out c from the preceding calculation is quite involved, but knowing the existence of c, we can determine it by an abelian estimate.*

*Finally, we have that*

$$\begin{aligned}c =& \frac{1}{\Gamma(\mathcal{B}/h)}\prod_{p|k}\left\{1+\frac{f(p)}{p}+\frac{f\left(p^2\right)}{p^2}+\cdots\right\}\times\\ &\lim_{s\to 1+}(s-1)^{\mathcal{B}/h}\prod_{i=1}^{h}\prod_{p\equiv a_i \pmod{k}}\left\{1+\frac{b_i}{p^s}+\frac{f\left(p^2\right)}{p^{2s}}+\cdots\right\}.\end{aligned}$$

*Now let*

$$J_{a,b}(s) = H_{a,b}(s) \exp\{b\phi_a(s)\} \prod_{\chi \neq \chi_1} L(s,\chi)^{b\overline{\chi}(a)/h}.$$

*Then we have* $\widehat{F}(s) = F^*(s)(s-1)^{-\mathcal{B}}$*, with*

$$F^*(s) = K(s) \left\{ (s-1)\zeta(s) \prod_{p|k} \left(1 - p^{-s}\right) \right\}^{\mathcal{B}/h} \prod_{i=1}^{h} J_{a_i,b_i}(s),$$

*an analytic function with no zeroes in* $\{s : \sigma \geq 1\}$*. Also,* $0 < \mathcal{B}/h \leq 1$*. Therefore, the generalized Wiener-Ikehara theorem applies, and we finally obtain*

$$\sum_{n \leq x} f(n) \sim cx \left(\log x\right)^{(\mathcal{B}-h)/h}$$

*with* $c := c_f := F^*(1)/\Gamma(\mathcal{B}/h)$*.*

Now we are ready to prove our theorem.

**Theorem 41.** *There is a constant* $\beta$ *such that*

$$B(x) \sim \frac{\beta x}{\sqrt{\log x}}.$$

**Proof.** *Let* $b \in \mathcal{A}$ *be the indicator function of* $\left\{n \in \mathbb{Z} : n = u^2 + v^2\right\}$*. By our lemmas, we see that* $b$ *is multiplicative and*

$$b\left(p^{\alpha}\right) = \begin{cases} 1, \text{ if } p = 2 \text{ or } p \equiv 1 \pmod 4, \\ 1, \text{ if } p \equiv 3 \pmod 4 \text{ and } \alpha \text{ is even,} \\ 0, \text{ if } p \equiv 3 \pmod 4 \text{ and } \alpha \text{ is odd.} \end{cases}$$

*We will typically denote this by* $b_n$ *for integral* $n$*. Let* $\widehat{B}$ *denote the associated Dirichlet series, i.e. if*

$$\widehat{B}(s) = \sum_{n=1}^{\infty} \frac{b_n}{n^s}$$

*then we have*

$$\widehat{B}(s) = \left(1 - 2^{-s}\right)^{-1} \prod_{p} \left(1 - p^{-s}\right)^{-1} \prod_{q} \left(1 - q^{-2s}\right)^{-1} \ (\sigma > 1),$$

*where* $p$ *runs through primes congruent to* $1 \pmod 4$ *and* $q$ *runs through primes congruent to* $3 \pmod 4$*. Now* $b$ *satisfies the previous lemma.*

*In the notation of that lemma, we have $k$, our modulus, equal to $4$, $h = \varphi(4) = 2$, and $\mathcal{B}$, $f = b$ and the sum of $b(p)$, for $p$ in the reduced residue system modulo $4$, equal to $1$.*

*It follows that $\sqrt{s-1}\widehat{B}(s)$ is analytic on the closed half-plane $\{s : \sigma \geq 1\}$ and that $B$ has an asymptotic formula of the stated form.*

*It will be easier to evaluate $\beta$ by expressing $\widehat{B}$ in terms of some other functions which we have some tools for. Observe that*

$$\begin{aligned}\widehat{B}(s)^2 = &\left\{\left(1-2^{-s}\right)^{-1}\prod_p\left(1-p^{-s}\right)^{-1}\prod_q\left(1-q^{-s}\right)^{-1}\right\}\times\\ &\left(1-2^{-s}\right)^{-1}\left\{\prod_p\left(1-p^{-s}\right)^{-1}\prod_q\left(1+q^{-s}\right)^{-1}\right\}\prod_q\left(1-q^{-2s}\right)^{-1}\\ &=\zeta(s)\left(1-2^{-s}\right)^{-1}L(s,\chi_2)\prod_q\left(1-q^{-2s}\right)^{-1},\end{aligned}\tag{5.2}$$

*where $\chi_2$ is the nonprincipal character modulo $4$. Furthermore, the last product converges for $\sigma > 1/2$.*

*Since the factors of (5.2) other than $\zeta$ are analytic on the open half-plane $\{s : \sigma > 1/2\}$, we conclude from the Generalized Wiener-Ikehara Theorem that*

$$\begin{aligned}B(x) &\sim x(\log x)^{-1/2}\frac{1}{\Gamma(1/2)}\left\{2L(1,\chi_2)\prod_{q\equiv(4)}\left(1-q^{-2}\right)^{-1}\right\}^{1/2}\\ &\sim\prod_{q\equiv 3(4)}\left(1-q^{-2}\right)^{-1/2}x/\sqrt{2\log x},\end{aligned}$$

*and in fact, this proves the theorem since we can let*

$$\beta = \frac{1}{\Gamma(1/2)}\left\{2L(1,\chi_2)\prod_{q\equiv 3(4)}\left(1-q^{-2}\right)^{-1}\right\}^{1/2}.$$

*It is a simple matter to estimate $\beta$ however, so we might as well.*

*We use the reflection formula for the gamma function to evaluate $\Gamma(1/2)$:*

$$(\Gamma(1/2))^2 = \frac{\pi}{\sin\pi/2}$$

$$\Rightarrow \Gamma(1/2) = \sqrt{\pi}\tag{5.3}$$

*The series for $L(1,\chi_2)$ tells us that it is equal to* $\arctan 1 = \pi/4$. *By multiplying by* 2 *and taking the square root, we get that*

$$\frac{1}{\Gamma(1/2)} \cdot \sqrt{2 \cdot L(1,\chi_2)} = \frac{1}{\sqrt{2}}.$$

*Now the product can be estimated as needed for more accuracy. Although it converges slowly (reaching only three decimal places of accuracy upon evaluating it over the first* 25 *primes congruent to* 3 *modulo* 4*), there exist methods to evaluate this product quicker.*

*Our estimation will serve fine for our purposes however; we get the result*

$$\beta = \frac{1}{\sqrt{2}} \left\{ \prod_{q \equiv 3(4)} \left(1 - q^{-2}\right)^{-1} \right\}^{1/2} = 0.764\ldots.$$

### 5.2.2 Asymptotic Expansion of $B(x)$

The original proof of our theorem was first done by Landau which can be found in [LanCol] or [LanZah]. The following is a sketch of Landau's proof given by G. H. Hardy in [HarRam]. In the next section we prove the theorem completely using methods developed in the Prime Number Theorem, but this argument serves us well by giving us another viewpoint and by reinforcing ideas developed up to this point.

As was shown earlier in the paper, it is necessary and sufficient that for $n$ to be expressible as a sum of two squares

$$n = 2^\alpha P Q^2,$$

where $P$ is a product of primes $p \equiv 1 \pmod 4$ and $Q$ a product of primes $q \equiv 3 \pmod 4$.

As before, we get the associated Dirichlet series $\widehat{B}$ such that

$$\widehat{B}(s) = \frac{1}{1-2^{-s}} \prod_{p \equiv 1 \pmod 4} \frac{1}{1-p^{-s}} \prod_{q \equiv 3 \pmod 4} \frac{1}{1-q^{-2s}}.$$

Also we have the nice representation of its square

$$\widehat{B}^2(s) = \psi(s)\zeta(s)\mathrm{L}(s). \tag{5.4}$$

where $\psi(s)$ is the ugly part, that is,

$$\psi(s) := \frac{1}{1-2^{-s}} \prod_{q \equiv 3(4)} \frac{1}{1-q^{-2s}}. \tag{5.5}$$

It is clear that $\psi(s)$ is analytic and has no zeroes for $\sigma > 1/2$. $L(s)$ is an integral function that has value $\pi/4$ for $s = 1$ and $\zeta(s)$ is analytic everywhere except at its pole at $s = 1$. It is also a well known fact that neither $\zeta(s)$ nor $L(s)$ vanishes in a region $D$ to the left of $\sigma = 1$ of the type

$$\sigma > 1 - \frac{A}{\{\log(|t| + 2)\}^{\Lambda}}.$$

And finally, for large $t$ in $D$, $\zeta(s)$ and $L(s)$ are

$$O\left((\log|t|)^{\Lambda}\right).$$

It follows that

$$\widehat{B}(s) = (s-1)^{-1/2} g(s)$$

,

where $g(s)$ is analytic in $D$ and

$$g(1) = \left\{\frac{\pi}{4}\psi(1)\right\}^{1/2} = \left\{\frac{\pi}{2}\prod_{q\equiv 3(4)}\left(\frac{1}{1-r^{-2}}\right)\right\}^{1/2} = \beta\sqrt{\pi},$$

Now, recalling that $B(x)$ is the number of integers representable as the sum of two squares up to $x$, we have

$$B^*(x) = \sum_{n\le x}{}^{*} b_n = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(s)\frac{x^s}{s}ds$$

for $c > 1$, where the $^*$ in $B^*(x)$ tells us that when $x$ is an integer, we take only half of the last term of the sum. This formula comes from Perron's formula.

Now the integrand has an algebraic singularity, so we will approximate $B^*(x)$ by a loop integral around $s = 1$. We will do so via a path of integration of the type $C$ in Figure 5.1. Our approximating function will be

$$\frac{1}{2\pi i}\int_L f(s)\frac{x^s}{s}ds = \frac{1}{2\pi i}\int_L \frac{x^s}{(s-1)^{1/2}s}g(s)ds = \frac{\beta\sqrt{\pi}}{2\pi i}\int_L \frac{x^s}{(s-1)^{1/2}}h(s)ds, \tag{5.6}$$

where

$$h(s) = 1 + a_1(s-1) + a_2(s-1)^2 + \cdots \tag{5.7}$$

near $s = 1$, and $L$ is the path from $1-\eta$ around 1 as shown in the figure. If we work with $h(s)$ as though it were 1, we get

$$\frac{\beta}{\sqrt{\pi}}\int_{1-\eta}^{1}\frac{x^s}{(1-s)^{1/2}}ds = \frac{\beta x}{\sqrt{\pi}}\int_0^{\eta}\frac{\exp(-u\log x)}{\sqrt{u}}du,$$

which is nearly

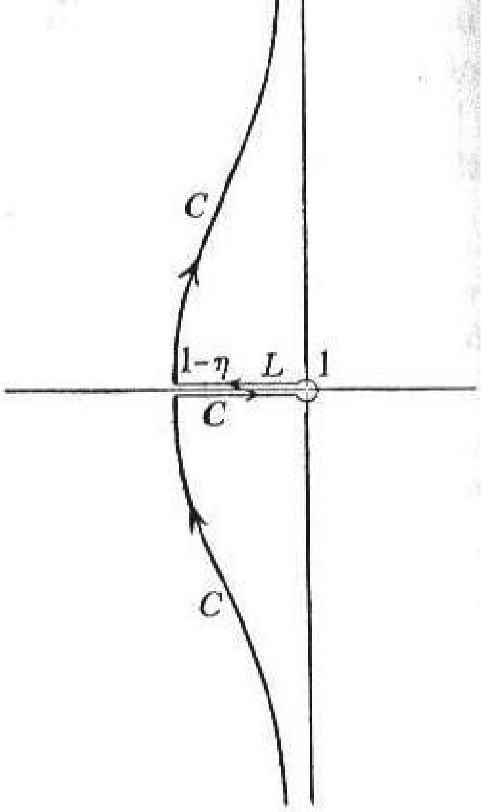


Figure 5.1: [HarRam], page 62

$$\frac{\beta x}{\sqrt{\log x}}$$

and when this is properly developed, leads us to the asymptotic expansion

$$B(x) = \frac{\beta x}{\sqrt{\log x}} \left\{ 1 + \frac{\alpha_1}{\log x} + \frac{\alpha_2}{(\log x)^2} + \cdots \right\}. \tag{5.8}$$

This is effectively Landau's result.

### 5.2.3 Method 2 - Via Contour Integration

Up to now, we have only outlined a sketch of the argument, that when developed leads us to the asymptotic expansion given above. As we shall see, the reason for doing this is to avoid some tedious calculations as well as delving too deeply into the complex analysis.

But we shall advance further utilizing the methods laid out in [LeV2] to prove the Prime Number Theorem applied to $B(x)$.

First, we require two lemmas, which, because $B(x)$ is composed of factors of $\zeta$ and $L$, we obtain immediately.

**Lemma 10.** *1. $\widehat{B}^2(s)$ is analytic and different from zero in the region $Q$ of Figure 5.2, for suitable $c_{11}$ and $c_9$, and it has a simple pole at $s = 1$ with residue*

$$\frac{\pi}{2} \prod_{q \equiv 3(4)} \left(1 - q^{-2}\right)^{-1} = r^2.$$

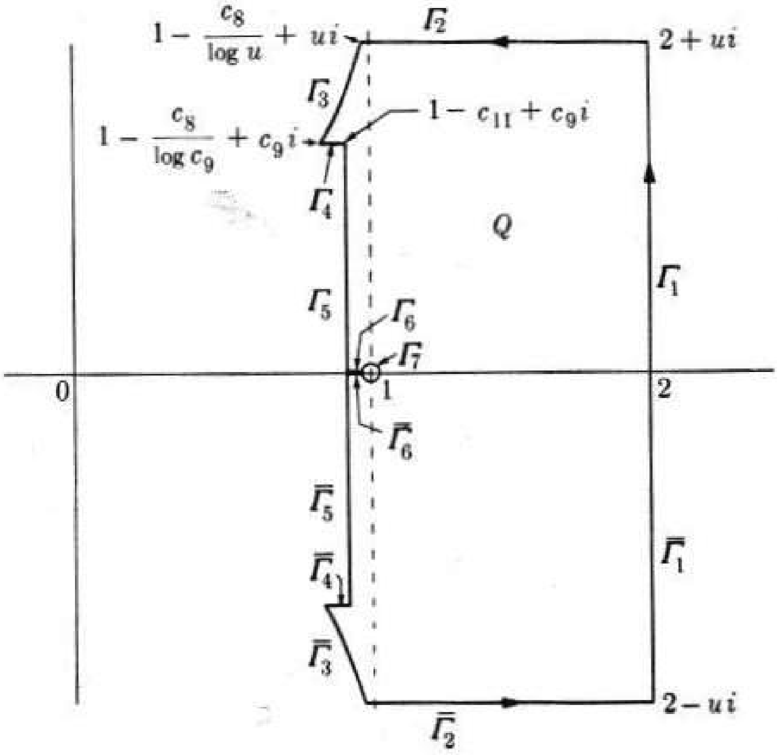


Figure 5.2: [LeV2], page 243

*Hence, $\widehat{B}$ is also analytic in $Q$ and $\widehat{B}^2(s) \cdot (s-1)$ is analytic in the uncut region of $Q'$ created from $Q$ by omitting $\Gamma_6$, $\Gamma_7$, and $\overline{\Gamma}_6$, and joining $\Gamma_5$ and $\overline{\Gamma}_5$.*

2. *For $|t| \geq 8$ and $s$ in $Q$, the inequality $\left|\widehat{B}(s)\right| < c_{14} \log |t|$ holds.*

From this, we obtain a similar result as we had before.

**Lemma 11.** *For appropriate $c < 1$,*

$$\sum_{n \leq x} b_n \log \frac{x}{n} = \frac{1}{\pi i} \int_c^1 \frac{x^s}{s^2} \widehat{B}(s) ds + O\left(x e^{-\alpha\sqrt{\log x}}\right).$$

**Proof.** *This proof begins the same as the proof of the similar result in chapter 2. That is, by changing the path of integration in the relation*

$$\sum_{n\le x} b_n \log\frac{x}{n} = \frac{1}{2\pi i}\int_{(2)} \frac{x^s}{s^2}\widehat{B}(s)ds$$

*and estimating the new integrals along paths which are bounded away from $s = 1$; the only change is that we use the estimate from the last lemma, i.e.* $|\widehat{B}(s)| \le c_{14}\log|t|$ *instead of* $|\log(\zeta(s))| < \log^2|t|$. *Excluding this tediousness, we obtain the relation*

$$\sum_{n=1}^{x} b_n \log\frac{x}{n} = \frac{1}{2\pi i}\left(-\int_{\overline{\Gamma}_6} - \int_{\Gamma_7} - \int_{\Gamma_6}\right)\frac{x^s}{s^2}\widehat{B}(s)ds + O\left(xe^{-\alpha\sqrt{\log x}}\right).$$

*Near* $s = 1$, $\widehat{B}(s)$ *has the expansion*

$$\widehat{B}(s) = \frac{r}{\sqrt{s-1}} + \cdots,$$

*with*

$$r = \sqrt{\frac{\pi}{2}\prod_q (1 - r^-2)^{-1}}$$

*as before. Here* $\sqrt{s-1} > 0$ *for* $s > 1$. *Putting* $s = 1 + \delta e^{i\theta}$, *we obtain*

$$\int_{\Gamma_7}\frac{x^s}{s^2}\widehat{B}(s) = O\left(\frac{x^{1+\delta}}{(1-\delta)^2}\cdot\frac{1}{\sqrt{\delta}}\cdot 2\pi\delta\right) = o(1)$$

*as* $\delta \to 0$.

*Since* $\widehat{B}(s)(s-1)$ *is single-valued in* $Q'$, *the quantity* $2\arg\widehat{B}(s)+\arg(s-1)$ *is unchanged by traveling a path in* $Q$ *from* $\overline{\Gamma}_6$ *to* $\Gamma_6$. *Since* $\arg(s-1)$ *increases by* $2\pi$, $\arg\widehat{B}(s)$ *decreases by* $\pi$. *Thus* $\widehat{B}(s)$ *has opposite signs on the two edges of the cut. Therefore*

$$\int_{\Gamma_6}\frac{x^s}{s^2}\widehat{B}(s)ds + \int_{\overline{\Gamma}_6}\frac{x^s}{s^2}\widehat{B}(s)ds = 2\int_{\overline{\Gamma}_6}\frac{x^s}{s^2}\widehat{B}(s)ds,$$

*and therefore*

$$\sum_{n=1}^{x} b_n \log\frac{x}{n} = \frac{1}{\pi i}\int_{1-c_{11}}^{1}\frac{x^s}{s^2}\widehat{B}(s)ds + O\left(xe^{-\alpha\sqrt{\log x}}\right),$$

*which was to be shown.*

**Theorem 42.** *As* $x \to \infty$,

$$B(x) = \frac{\beta x}{\sqrt{\log x}} + O\left(\frac{x}{(\log x)^3/4}\right),$$

*where*

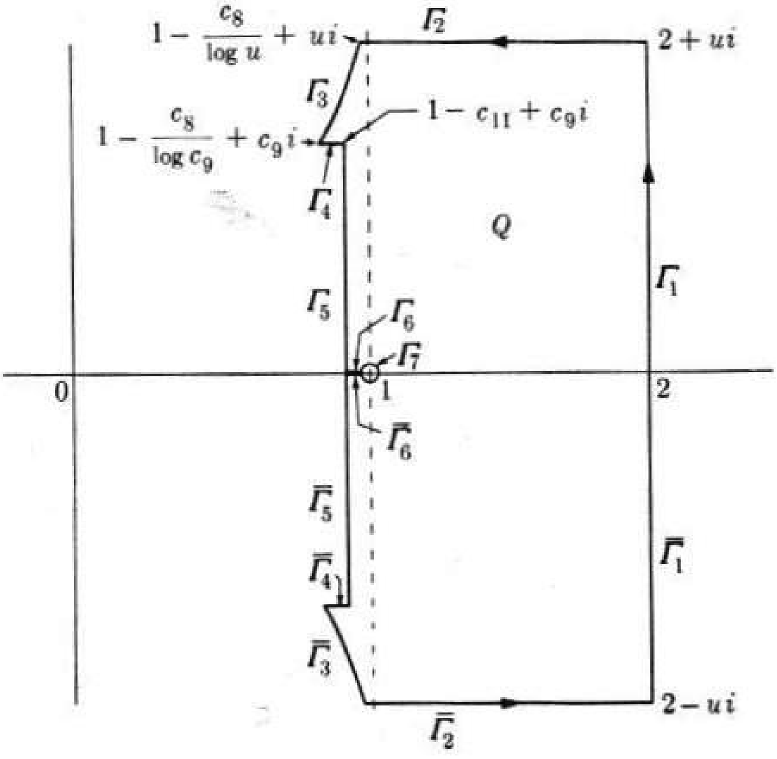


Figure 5.3: [LeV2], page 243

$$\beta = \frac{1}{\sqrt{2}} \prod_{q \equiv 3(4)} \left(1 - \frac{1}{q^2}\right)^{-1/2}.$$

**Proof.** *Referring to Figure 5.3, on* $\overline{\Gamma}_6$ *we have*

$$\begin{aligned}\frac{\widehat{B}(s)}{s^2} &= \frac{ri}{\sqrt{1-s}(1-(1-s))^2} + O(\sqrt{1-s}) \\ &= \frac{ri}{\sqrt{1-s}} + O(\sqrt{1-s})\end{aligned}$$

*as $s \to 1^-$. Then according to the previous lemma*

$$
\begin{aligned}
&\sum_{n=1}^{x} b_n \log\frac{x}{n} \\
&= \frac{1}{\pi i}\int_{1-c_{11}}^{1}\frac{x^s ri}{\sqrt{1-s}}ds + O\left(\int_{1-c_{11}}^{1} x^s\sqrt{1-s}ds\right) + O\left(\frac{x}{\log^2 x}\right) \\
&= \frac{r}{\pi}\int_{0}^{c_{11}} x^{1-u}u^{-1/2}du + O\left(\int_{0}^{c_{11}} x^{1-u}u^{1/2}du\right) + O\left(\frac{x}{\log^2 x}\right) \\
&= \frac{rx}{\pi}\int_{0}^{c_{11}} e^{-u\log x}u^{-1/2}du + O\left(x\int_{0}^{c_{11}} e^{-u\log x}u^{1/2}du\right) + O\left(\frac{x}{\log^2 x}\right) \\
&= \frac{rx}{\pi}\int_{0}^{c_{11}\log x} e^{-v}\left(\frac{v}{\log x}\right)^{-1/2}\frac{dv}{\log x} + O\left(x\int_{0}^{c_{11}\log x} e^{-v}\left(\frac{v}{\log x}\right)^{1/2}\frac{dv}{\log x}\right) \\
&\qquad + O\left(\frac{x}{\log^2 x}\right) \\
&= \frac{rx}{\pi\sqrt{\log x}}\int_{0}^{c_{11}\log x} e^{-v}v^{-1/2}dv + O\left(\frac{x}{\log^{3/2} x}\int_{0}^{\infty} e^{-v}v^{1/2}dv\right) + O\left(\frac{x}{\log^2 x}\right) \\
&= \frac{rx}{\pi\sqrt{\log x}}\left\{\Gamma\left(\frac{1}{2}\right) - \int_{c_{11}\log x}^{\infty} e^{-v}v^{-1/2}dv\right\} + O\left(\frac{x}{\log^{3/2} x}\right) \\
&= \frac{rx}{\pi\sqrt{\log x}}\left\{\sqrt{\pi} + O\left(\frac{x}{\log^{3/2} x}\right)\right\}.
\end{aligned}
$$

*Therefore,*

$$\sum_{n=1}^{x} b_n \log\frac{x}{n} = \frac{\beta x}{\sqrt{\log x}} + O\left(\frac{x}{\log^{3/2} x}\right),$$

*where*

$$\beta = \frac{r}{\sqrt{\pi}} = \frac{1}{\sqrt{2}}\prod_{q}(1-q^2)^{-1/2},$$

*as before. Note that this is not $B(x)$, but it is very closely related to $B(x)$, and we will be able to estimate our function more accurately than before by using it.*

*Let $\delta = \delta(x)$ be positive. Then we have*

$$
\begin{aligned}
&\sum_{n=1}^{x+\delta x} b_n \log\frac{x+\delta x}{n} - \sum_{n=1}^{x} b_n \log\frac{x}{n} \\
&\qquad = \log(1+\delta)\sum_{n=1}^{x} b_n + \sum_{n=x}^{x+\delta x} b_n \log\frac{x+\delta x}{n} \\
&\qquad = \log(1+\delta)B(x) + O\left(\log(1+\delta)\cdot\delta x\right),
\end{aligned}
\tag{5.9}
$$

*but we also have that this is equal to the following*

$$\begin{aligned}\frac{\beta x(1+\delta)}{\sqrt{\log(x+\delta x)}}-\frac{\beta x}{\sqrt{\log x}}&=\frac{\beta x}{\sqrt{\log x}}\left(\frac{1+\delta}{\sqrt{\frac{\log(x+\delta x)}{\log x}}}-1\right)\\&=\frac{\beta x}{\sqrt{\log x}}\left(\frac{1+\delta}{\sqrt{1+\frac{\log(1+\delta)}{\log x}}}-1\right)\\&=\frac{\beta x}{\sqrt{\log x}}\left(\frac{1+\delta}{1+O(\delta/\log x)}-1\right)\\&=\frac{\beta x}{\sqrt{\log x}}\left(\frac{\delta+O(\delta/\log x)}{1+O(\delta/\log x)}\right)\\&=\frac{\beta x}{\sqrt{\log x}}\left(\delta+O(\delta/\log x)\right).\end{aligned}\tag{5.10}$$

*Now since* $\log(1+\delta)=\delta+O(\delta^2)$ *as* $\delta\to 0$*, we can solve for* $B(x)$ *by setting the last relations of equations 5.9 and 5.10 equal, and we obtain,*

$$\begin{aligned}B(x)&=\frac{\beta x}{\sqrt{\log x}}\left\{\frac{\delta}{log(1+\delta)}+O\left(\frac{1}{\log x}\right)\right\}+O(\delta x)+O\left(\frac{x}{\delta\log^{3/2}x}\right)\\&=\frac{\beta x}{\sqrt{\log x}}\left(1+O(\delta)\right)+O\left(\frac{x}{\log^{3/2}x}\right)+O(\delta x)+O\left(\frac{x}{\delta\log^{3/2}x}\right).\end{aligned}$$

*Let* $\delta(x)=1/\log^{3/4}x$*, we obtain the following:*

$$\begin{aligned}B(x)&=\frac{\beta x}{\sqrt{\log x}}+O\left(\frac{x}{\log^{5/4}x}\right)+O\left(\frac{x}{\log^{3/2}x}\right)\\&\quad+O\left(\frac{x}{\log^{3/4}x}\right)+O\left(\frac{x}{\log^{3/4}x}\right)\\&=\frac{\beta x}{\sqrt{\log x}}+O\left(\frac{x}{\log^{3/4}x}\right),\end{aligned}$$

*which was to be proved.*

# Chapter 6

# Conclusion

Using the same methods, an analytic proof of Dirichlet's theorem can be obtained (Appendix A), and clearly other functions that can be obtained in terms of $\zeta(s)$ or $L(s, \chi)$ and fit our criteria can be explored as well.

As we have seen, the mathematics going into proving the estimate of $B(x)$ are significant and able to be utilized in other situations as well. The machinery of $\zeta$ and of $L$ are used throughout number theory and are on the forefront of research being done today, not just in the proof of the Prime Number Theorem. To anyone attempting to delve the mysteries of $\zeta$ or $L$, this material provides ample proving grounds, and prompts even more questions.

## Acknowledgements

I would like to thank Dr. Fabrykowski for his invaluable guidance and assistance with this paper and in all things number theory. I also thank the professors I am defending this thesis to: Dr. Wingler and Dr. Smotzer.

The mathematics department at Youngstown State University has monopolized the best experiences of my academic career for the last six years and I am indebted to the wonderful characters populating it. In particular, I would like to acknowledge my colleagues in the graduate program of YSU.

I would like to thank my family members in supporting me in pursuing this degree, and without whom I would surely have entered into some artistic endeavor. Especially, my fiancé Samantha has been my heart and inspiration in everything I do.

Finally, I give my regards to all that supported me in any manner during this thesis.

# Bibliography

[Apos] Tom M. Apostol, *Introduction to Analytic Number Theory*, Springer Science+Business Media, LLC, New York, 1976.

[BatDia] Paul Bateman and Harold Diamond, *Analytic Number Theory: An Introductory Course*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2004.

[BamChow] R. P. Bambah and S. Chowla, *On Numbers Which Can Be Expressed As A Sum Of Two Squares*, `http://www.new.dli.ernet.in/rawdataupload/upload/insa/INSA_1/20005b85_101.pdf`

[Burt] David Burton, *Elementary Number Theory*, McGraw-Hill, New York, 2007.

[HarRam] G. H. Hardy, *Ramanujan:Twelve lectures on subjects suggested by his life and work*, Cambridge University Press, Cambridge, 1940. Reprinted by American Mathematical Society, Providence, 1999.

[HarWr] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, University Press, Oxford, 1975.

[Hool] Christopher Hooley, *On the Intervals Between Numbers that are Sums of Two Squares*,

[Kara] Anatoly A. Karatsuba, *Complex Analysis in Number Theory*, CRC Press, Inc., Boca Raton, 1995.

[LanCol] Edmund Landau, *Collected Works*, Vol. 4, Thales Verlag, Essen, 1987.

[LanHan] Edmund Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Chelsea Publishing Company, New York, 1974.

[LanZah] Edmund Landau, *Vorlesungen über Zahlentheorie*, Chelsea Pub. Co., New York, 1947.

[LeV1] William Judson LeVeque, *Topics in Number Theory, Volume I*, Addison-Wesley Publishing Company, Inc, Reading, 1958.

[LeV2] William Judson LeVeque, *Topics in Number Theory, Volume II*, Addison-Wesley Publishing Company, Inc, Reading, 1961.

[Sier] W. Sierpinski, *Elementary Theory of Numbers*, PWN - Polish Scientific Publishers, Warszawa, 1964.

[Shiu] P. Shiu, *Counting Sums of Two Squares: The Meissel-Lehmer Method*, Mathematics of Computation, Vol. 47, No. 175, Jul., 1986.

# Appendix A

# Proofs

This appendix contains all proofs omitted from chapter 2. Like that chapter, this material can be found in any number of texts, including, [LanHan], [HarWr], [Apos], [BatDia], [Burt], [Sier] and [Kara].

When possible, we have relied on [Apos] exclusively.

## A.1 Cross-Classification Principle

**Theorem.** *Let $S$ be a set of $N$ distinct elements, and let $S_1, \ldots, S_r$ be arbitrary subsets of $S$ containing $N_1, \ldots, N_r$ elements, respectively. For $1 \leq i < j < \cdots < l \leq r$, let $S_{ij\cdots l}$ be the intersection of $S_i, S_j, \ldots, S_l$; and let $N_{ij\cdots l}$ be the number of elements of $S_{ij\cdots l}$. Then the number of elements of $S$ not in any of $S_1, \ldots, S_r$ is*

$$K = N - \sum_{1\leq i\leq r} N_i + \sum_{1\leq i<j\leq r} N_{ij} - \sum_{1\leq i<j<k\leq r} N_{ijk} + \cdots \\ + (-1)^r N_{12\cdots r}.$$

**Proof.** *Let a certain element $s$ of $S$ belong to exactly $m$ of the sets $S_1, \ldots, S_r$. If $m = 0$, $s$ is counted only once, in $N$ itself. If $0 < m \leq r$, then $s$ is counted once, or $\binom{m}{0}$ times, in $N$, $\binom{m}{1}$ times in the terms $N_i$, $\binom{m}{2}$ times in the terms $N_{ij}$, etc. Thus, the total contribution to $K$ arising from the element $s$ is*

$$\binom{m}{0} - \binom{m}{1} + \binom{m}{2} - \cdots + (-1)^m \binom{m}{m} = (1-1)^m = 0.$$

## A.2 Quadratic Residues

**Theorem (Euler's Criterion).** *Let $p$ be an odd prime and $(a, p) = 1$. Then $a$ is a quadratic residue of $p$ if and only if $a^{(p-1)/2} \equiv 1 \pmod{p}$.*

**Proof.** *Suppose that $a$ is a quadratic residue of $p$, thus $x^2 \equiv a \pmod{p}$ has a solution, say $x_0$. Because $(a,p)=1$, we have that $(x_0,p)=1$. We may therefore apply Fermat's little theorem to obtain*

$$a^{(p-1)/2} \equiv \left(x_0^2\right)^{(p-1)/2} \equiv x_0^{p-1} \equiv 1 \pmod{p}.$$

*Now for the other direction. Assume the congruence $a^{(p-1)/2} \equiv 1 \pmod{p}$ holds, and let $r$ be a primitive root of $p$. Then $a \equiv r^k \pmod{p}$ for some integer $k$ with $1 \leq k \leq p-1$. It follows that*

$$r^{k(p-1)/2} \equiv a^{(p-1)/2} \equiv 1 \pmod{p}.$$

*Now, the order of $r$ (i.e. $p-1$), must divide the exponent $k(p-1)/2$. Thus, $k$ is an even integer, say $k = 2m$. Therefore,*

$$(r^m)^2 = r^{2m} = r^k \equiv a \pmod{p}.$$

*Thus, $r^m$ is a solution of the congruence $x^2 \equiv a \pmod{p}$. Thus proving that $a$ is a quadratic residue of the prime $p$.*

**Theorem.** *If $p$ is a prime of the form $4k+1$ (where $k$ is a natural number), then*

$$p \left| \left[\left(\frac{p-1}{2}\right)!\right]^2 + 1\right.$$

**Proof.** *We have $\frac{p-1}{2} = 2k$ since $p = 4k+1$, which implies that*

$$\begin{aligned} 1 \cdot 2 \cdot 3 \cdots \frac{p-1}{2} &= (-1)(-2) \cdots - \frac{p-1}{2} \\ &\equiv (p-1)(p-2) \cdots \frac{p+1}{2} ( \pmod{p}). \end{aligned}$$

*Therefore we obtain*

$$\left[\left(\frac{p-1}{2}\right)!\right]^2 \equiv 1 \cdot 2 \cdots \frac{p-1}{2} \cdot \frac{p+1}{2} \cdots (p-1) \equiv (p-1)! \equiv -1 \pmod{p},$$

*which yields our theorem after adding 1 to both sides.*

**Theorem (Thue Theorem).** *If $m$ is a natural number and $a$ an integer relatively prime to $m$, then there exist natural numbers $x$ and $y$ both less than $\sqrt{m}$ and such that the number $ax \pm y$is divisible by $m$ for a suitable choice of the sign $\pm$.*

**Proof.** *If $m = 1$, then the theorem holds since we may set $x = y = 1$. Now suppose that $m$ is a natural number greater than 1. Let $q = \lfloor\sqrt{m}\rfloor$. Then $q + 1 > \sqrt{m}$ and $(q + 1)^2 > m$.*

*Consider the expressions $ax - y$ for $x, y$ taking the values $0, 1, 2, \ldots, q$. There are $(q + 1)^2$ of these expressions, and only $m$ different remainders obtained by division of $m$. Thus, for two different pairs $x_1, y_1$ and $x_2, y_2$, and without loss of generality $x_1 > x_2$, we obtain the same remainders by divison of $ax - y$ by $m$.*

*Consequently, we have*

$$m|ax_1 - y_1 - (ax_2 - y_2) = a(x_1 - x_2) - (y_1 - y_2)$$

*Now, we cannot have $x_1 = x_2$, since if so, then $y_1 - y_2$ would be divisible by $m$ but $0 \leq y_1 \leq q \leq \sqrt{m} < m$ and $0 \leq y_2 < m$.*

*Furthermore, the equality $y_1 = y_2$ is impossible, since then $a(x_1 - x_2)$ would be divisible by $m$ and $(a, m) = 1$ implies that $m\,|x_1 - x_2$, but we have that $0 \leq x_1 < m$ and $0 \leq x_2 < m$. Thus we have both $x_1 \neq x_2$ and $y_1 \neq y_2$. Now since, $x_1 \geq x_2$, $x = x_1 - x_2$ is a natural number. The number $y_1 - y_2$ can be negative, but it is different from zero, so $y = |y_1 - y_2|$ is a natural number. Thus*

$$x = x - x_1 - x_2 \leq x_1 \leq q \leq \sqrt{m},\, y \leq q \leq \sqrt{m}$$

.

*Finally, we have that for the appropriate sign, + or −, the number $a\,(x_1 - x_2) - (y_1 - y_2) = ax \pm y$ is divisible by $m$.*

## A.3 Functions

Before we can prove the next theorem, involving the product formula of $\varphi(n)$, we need a lemma relating $\varphi(n)$ to $\mu(n)$.

**Lemma 12.** *If $n \geq 1$ we have*

$$\varphi(n) = \sum_{d|n} \mu(d)\frac{n}{d}.$$

Now we are prepared to prove our theorem.

**Theorem.** *For $n \geq 1$ we have*

$$\varphi(n) = n\prod_{p|n}\left(1 - \frac{1}{p}\right).$$

**Proof.** *For $n = 1$, the product, $\prod_{p|n}(1 - 1/p)$ is empty since there are no primes that divide $1$. Therefore the product is assigned the value $1$ by definition.*

*Now suppose $n > 1$ and let $p_1, \dots, p_r$ be the distinct prime divisors of $n$. Now the product can be written as*

$$\begin{aligned}\prod_{p|n}\left(1 - \frac{1}{p}\right) &= \prod_{i=1}^{r}\left(1 - \frac{1}{p_i}\right)\\ &= 1 - \sum \frac{1}{p_i} + \sum \frac{1}{p_i p_j} - \sum \frac{1}{p_i p_j p_k} + \cdots + \frac{(-1)^r}{p_1 p_2 \cdots p_r}\end{aligned}$$

*Now, on the right hand side, a term like $\sum 1/(p_i p_j p_k)$ is understood to consider all possible products $p_i p_j p_k$ of distinct prime factors of $n$ taken three at a time. Now each term on the right is of the form $\pm 1/d$ where $d$ is a divisor of $n$ which is either $1$ or a product of distinct primes. The numerator $\pm 1$ is $\mu(d)$. Now we have that $\mu(d) = 0$ if $d$ is divisible by the square of any $p_i$ and so the sum is exactly*

$$\sum_{d|n} \frac{\mu(d)}{d} = \varphi(n),$$

*proving the theorem.*

**Proof.** ***(via Cross-Classification Principle)****.*
*We have a second proof of the product formula, utilizing the principle of cross-classification from the first section.*

*In the notation of that theorem, take $S$ to be the set of integers $1, \dots, n$, and for $1 \leq k \leq r$, take $S_k$ to be the set of elements of $S$ which are divisible by $p_k$, where $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$. Then if $d|n$, the number of integers $s \leq n$ such that $d|s$ is $n/d$; therefore*

$$\varphi(n) = n - \sum_{1 \leq i \leq r} \frac{n}{p_i} + \sum_{1 \leq i < j \leq r} \frac{n}{p_i p_j} - \cdots = n \prod_{p|n}\left(1 - \frac{1}{p}\right).$$

## A.4 Dirichlet Series

**Theorem.** *Suppose the series $\sum |f(n)n^{-s}|$ does not converge for all $s$ or diverge for all $s$. Then there exists a real number, $\sigma_c$, called the* abscissa of absolute convergence*, such that the series $\sum |f(n)n^{-s}|$ converges absolutely if $\sigma > \sigma_c$, but does not converge absolutely if $\sigma < \sigma_c$.*

**Proof.** *Let $D$ be the set of all real $\sigma$ such that $\sum |f(n)n^{-s}|$ diverges. $D$ is not empty because the series does not converge for all $s$, and $D$ is bounded above because the series does not diverge for all $s$. Therefore, $D$ has a least upper bound, say $\sigma_c$. If $\sigma < \sigma_c$, then $\sigma \in D$, otherwise $\sigma$ would be an upper bound for $D$ smaller than the least upper bound, which is absurd. If $\sigma > \sigma_c$, then $\sigma \notin D$ since $\sigma_c$ is an upper bound for $D$. Hence the theorem is proved.*

Again, we are in need of another lemma before proving the next theorem.

**Lemma 13.** *If $N \geq 1$ and $\sigma \geq b > \sigma_c$, we have*

$$\left|\sum_{n=N}^{\infty} f(n)n^{-s}\right| \leq N^{-(\sigma-b)} \sum_{n=N}^{\infty} |f(n)|\, n^{-b}.$$

**Theorem (Uniqueness theorem).** *Given two Dirichlet series*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \text{ and } G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s},$$

*both absolutely convergent for $\sigma > \sigma_c$. If $F(s) = G(s)$ for each $s$ in an infinite sequence $\{s_k\}$ such that $\mathrm{Re}(s_k) = \sigma_k \to +\infty$ as $k \to \infty$, then $f(n) = g(n)$ for every $n$.*

**Proof.** *Let $h(n) = f(n) - g(n)$ and let $H(s) = F(s) - G(s)$. Then $H(s_k) = 0$ for each $k$. To prove $h(n) = 0$ for all $n$ we assume that $h(n) \neq 0$ for some $n$ and obtain an absurdity.*

*Let $n_0$ be the smallest integer for which $h(n) \neq 0$. Then we have the following*

$$H(s) = \sum_{n=n_0}^{\infty} \frac{h(n)}{n^s} = \frac{h(n_0)}{n_0^s} + \sum_{n=n_0+1}^{\infty} \frac{h(n)}{n^s}.$$

*Therefore,*

$$h(n_0) = n_0^s H(s) - n_0^s \sum_{n=n_0+1}^{\infty} \frac{h(n)}{n^s}.$$

*Now assign $s = s_k$, we get $H(s_k) = 0$. It follows that*

$$h(n_0) = -n_0^{s_k} \sum_{n=n_0+1}^{\infty} h(n)n^{-s_k}.$$

*Now choose $k$ such that $\sigma_k > b$ where $b > \sigma_c$. Our lemma implies that*

$$|h(n_0)| \leq n_0^{\sigma_k}(n_0+1)^{-(\sigma_k-b)} \sum_{n=n_0+1}^{\infty} |h(n)|\, n^{-c} = \left(\frac{n_0}{n_0+1}\right)^{\sigma_k} A$$

*with $A$ independent of $k$. Allowing $k \to \infty$, we find that $(n_0/(n_0+1))^{\sigma_k} \to 0$, but this implies that $h(n_0) = 0$, which is absurd. Therefore, $n_0$ does not exist and our theorem is proved.*

**Theorem.** *Given two functions $F(s)$ and $G(s)$ represented by Dirichlet series,*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \text{ for } \sigma > a,$$

*and*

$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s} \text{ for } \sigma > b.$$

*Then in the half-plane where both series converge absolutely we have*

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s}, \tag{A.1}$$

*where $h = f * g$, the Dirichlet convolution of $f$ and $g$:*

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

*Conversely, if $F(s)G(s) = \sum \alpha(n) n^{-s}$ for all $s$ in a sequence $\{s_k\}$ with $\sigma_k \to +\infty$ as $k \to \infty$, then $\alpha = f * g$.*

**Proof.** *For any $s$ for which both series converge absolutely we have*

$$F(s)G(s) = \sum_{n=1}^{\infty} f(n) n^{-s} \sum_{m=1}^{\infty} g(m) m^{-s} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(n) g(m) (mn)^{-s}.$$

*Due to absolute convergence, we can multiply these series together and rearrange terms as we see fit without altering the value of the sum. We collect together those terms for which $mn$ is constant, say $mn = k$. The possible values of $k$ are $1, 2, \ldots$, therefore,*

$$F(s)G(s) = \sum_{k=1}^{\infty} \left( \sum_{mn=k} f(n) g(m) \right) k^{-s} = \sum_{k=1}^{\infty} h(k) k^{-s}$$

*where $h(k) = \sum_{mn=k} f(n) g(m) = (f * g)(k)$, proving the first claim, while the second follows from uniqueness.*

**Theorem.** *For all $f$ we have $I * f = f * I = f$.*

**Proof.** *We have the following*

$$(f * I)(n) = \sum_{d|n} f(d) I\left(\frac{n}{d}\right) = \sum_{d|n} f(d) \left[\frac{d}{n}\right] = f(n)$$

*since* $[d/n] = 0$ *if* $d < n$.

**Theorem (Analytic Fundamental Theorem of Arithmetic).** *Let* $f$ *be a multiplicative arithmetical function such that the series* $\sum f(n)$ *is absolutely convergent. Then the sum of the series can be expressed as an absolutely convergent infinite product,*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \left\{1 + f(p) + f(p^2) + \cdots\right\} \tag{A.2}$$

*extended over all primes. If* $f$ *is completely multiplicative, the product simplifies and we have*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \frac{1}{1 - f(p)}. \tag{A.3}$$

*Either version is referred to as the* Euler product *of the series.*

**Proof.** *Consider the finite product*

$$P(x) = \prod_{p \leq x} \left\{1 + f(p) + f(p^2) + \cdots\right\}$$

*extended over all primes* $p \leq x$. *Since this is the product of a finite number of absolutely convergent series, we can multiply the series and rearrange the terms in any fashion without altering the sum. A typical term is of the form*

$$f\left(p_1^{a_1}\right) f\left(p_2^{a_2}\right) \cdots f\left(p_r^{a_r}\right) = f\left(p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}\right)$$

*since* $f$ *is multiplicative. By the fundamental theorem of arithmetic we can write*

$$P(x) = \sum_{n \in A} f(n)$$

*where* $A$ *consists of those* $n$ *having all their prime factors* $\leq x$. *Therefore,*

$$\sum_{n=1}^{\infty} f(n) - P(x) = \sum_{n \in B} f(n),$$

*where* $B$ *is the set of* $n$ *having at least one prime factor* $> x$. *Now we have*

$$\left|\sum_{n=1}^{\infty} f(n) - P(x)\right| \leq \sum_{n \in B} |f(n)| \leq \sum_{n > x} |f(n)| .$$

*As $x \to \infty$, the last sum on the right goes to $0$ since $\sum |f(n)|$ is convergent. Hence, $P(x) \to \sum f(n)$ as $x \to \infty$.*

*An infinite product of the form $\prod(1 + a_n)$ converges absolutely whenever the corresponding series $\sum a_n$ converges absolutely. In this case, we have*

$$\sum_{p \leq x} \left|f(p) + f(p^2) + \cdots\right| \leq \sum_{p \leq x} \left(|f(p)| + \left|f(p^2)\right| + \cdots\right) \leq \sum_{n=2}^{\infty} |f(n)| .$$

*Since all of the partial sums are bounded, the series of positive terms*

$$\sum_{p} \left|f(p) + f(p^2) + \cdots\right|$$

*converges, and this implies absolute convergence of the first product in the theorem.*

*Finally, when $f$ is completely multiplicative, we have $f(p^n) = [f(p)]^n$ and each series on the right in the theorem is a convergent geometric series with sum $(1 - f(p))^{-1}$*

**Lemma.** *Let $\{f_n\}$ be a sequence of functions analytic on an open subset $S$ of the complex plane, and assume that $\{f_n\}$ converges uniformly on every compact subset of $S$ to a limit function $f$. Then $f$ is analytic on $S$ and the sequence of derivatives $\{f_n'\}$ converges uniformly on every compact subset of $S$ to the derivative $f'$.*

**Proof.** *$f_n$ analytic on $S$ allows us to use Cauchy's integral formula*

$$f_n(a) = \frac{1}{2\pi i} \int_{\partial D} \frac{f_n(z)}{z - a} dz$$

*where $D$ is any compact disk in $S$, $\partial D$ is its positively oriented boundary, and $a$ is any interior point of $D$. Uniform convergence allows us to pass to the limit under the integral sign and obtain*

$$f(a) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(z)}{z - a} dz$$

*which implies that $f$ is analytic inside $D$. For the derivatives, we have*

$$f_n'(a) = \frac{1}{2\pi i} \int_{\partial D} \frac{f_n(z)}{(z - a)^2} dz \text{ and } f'(a) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(z)}{(z - a)^2} dz$$

*from which it follows that $f_n'(a) \to f'(a)$ uniformly on every compact subset of $S$ as $n \to \infty$.*

Before we prove the next lemma, we in fact have need of a separate one and also of Abel's identity, frequently called Abel summation.

**Theorem (Abel's identity).** *For any arithmetical function $a(n)$, let*

$$A(x) = \sum_{n\leq x} a(n)$$

*,*

*where $A(x) = 0$ if $x < 1$. Assume $f$ has a continuous derivative on the interval $[y, x]$, where $0 < y < x$. Then we have*

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt$$

*.*

**Proof.** *Let $k = [x]$ and $m = [y]$, so now $A(x) = A(k)$ and $A(y) = A(m)$. Then*

$$\begin{aligned}
\sum_{y<n\leq x} a(n)f(n) &= \sum_{n=m+1}^{k} a(n)f(n) = \sum_{n=m+1}^{k} \{A(n) - A(n-1)\}f(n) \\
&= \sum_{n=m+1}^{k} A(n)f(n) - \sum_{n=m}^{k-1} A(n)f(n+1) \\
&= \sum_{n=m+1}^{k-1} A(n)\{f(n) - f(n+1)\} + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} A(n)\int_n^{n+1} f'(t)dt + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} \int_n^{n+1} A(t)f'(t)dt + A(k)f(k) - A(m)f(m+1) \\
&= -\int_{m+1}^{k} A(t)f'(t)dt + A(x)f(x) - \int_k^x A(t)f'(t)dt \\
&\quad - A(y)f(y) - \int_y^{m+1} A(t)f'(t)dt \\
&= A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt.
\end{aligned}$$

**Lemma.** *Let $s_0 = \sigma_0 + it_0$ and assume that the D.s. $\sum f(n)n^{-s_0}$ has bounded partial sums, say*

$$\left|\sum_{n\leq x} f(n)n^{-s_0}\right| \leq M$$

*for all $x \geq 1$. Then for each $s$ with $\sigma > \sigma_0$, we get*

$$\left| \sum_{a<n\leq b} f(n)n^{-s} \right| \leq 2Ma^{\sigma_0-\sigma} \left( 1 + \frac{|s-s_0|}{\sigma-\sigma_0} \right).$$

**Proof.** *Let $a(n) = f(n)n^{-s_0}$ and let $A(x) = \sum_{n\leq x} a(n)$. Then $f(n)n^{-s} = a(n)n^{s_0-s}$, then through Abel summation (with $f(x) = x^{s_0-s}$), we obtain*

$$\sum_{a<n\leq b} f(n)n^{-s} = A(b)b^{s_0-s} - A(a)a^{s_0-s} + (s-s_0) \int_a^b A(t)t^{s_0-s-1}dt.$$

*We have $|A(x)| \leq M$, so*

$$\begin{aligned} \left| \sum_{a<n\leq b} f(n)n^{-s} \right| &\leq Mb^{\sigma_0-\sigma} + Ma^{\sigma_0-\sigma} + |s-s_0|M \int_a^b t^{\sigma_0-\sigma-1}dt \\ &\leq 2Ma^{\sigma_0-\sigma} + |s-s_0|M \left| \frac{b^{\sigma_0-\sigma} - a^{\sigma_0-\sigma}}{\sigma_0-\sigma} \right| \\ &\leq 2Ma^{\sigma_0-\sigma} \left( 1 + \frac{|s-s_0|}{\sigma-\sigma_0} \right). \end{aligned}$$

**Lemma.** *A Dirichlet series $\sum f(n)n^{-s}$ converges uniformly on every compact subset lying interior to the half-plane of convergence $\sigma > \sigma_c$.*

**Proof.** *It is enough to show that $\sum f(n)n^{-s}$ converges uniformly on every compact rectangle $R = [\alpha, \beta] \times [c, d]$ with $\alpha > \sigma_c$. In order to do this we use the estimate obtained in the previous lemma.*

$$\left| \sum_{a<n\leq b} f(n)n^{-s} \right| \leq 2Ma^{\sigma_0-\sigma} \left( 1 + \frac{|s-s_0|}{\sigma-\sigma_0} \right) \tag{A.4}$$

*where $s_0 = \sigma_0 + it_0$ is any point in the half-plane $\sigma > \sigma_c$ and $s$ is any point with $\sigma > \sigma_0$. We choose $s_0 = \sigma_0$ where $\sigma_c < \sigma_0 < \alpha$. i.e. Our rectangle $R$ will occur above and to the right of $\sigma_0 > \sigma_c$ as in the figure due to [Apos] (pg. 235).*

*Now if $s \in R$, then we have $\sigma - \sigma_0 \geq \alpha - \sigma_0$ and $|s_0 - s| < C$, where $C$ is a constant depending on $s_0$ and $R$, but not on $s$. Then our lemma implies*

$$\left| \sum_{a<n\leq b} f(n)n^{-s} \right| \leq 2Ma^{\sigma_0-\alpha} \left( 1 + \frac{C}{\alpha-\sigma_0} \right) = Ba^{\sigma_0-\alpha}$$

*where $B$ is independent of $s$. Now since $a^{\sigma_0-\alpha} \to 0$ as $a \to +\infty$, the Cauchy criterion for uniform convergence is satisfied.*

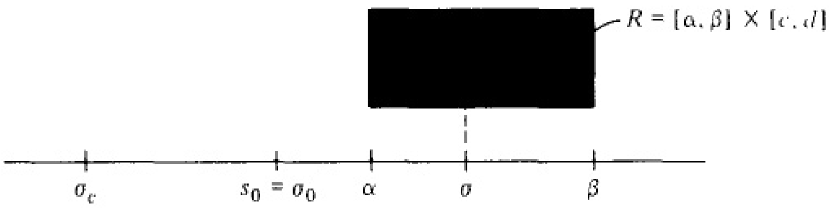


Figure A.1:

**Theorem.** *The summatory function of a Dirichlet series, $F(s) = \sum f(n)n^{-s}$, is analytic in its half-plane of convergence $\sigma > \sigma_c$, and its derivative $F'(s)$ is represented in this half-plane of convergence by the Dirichlet series*

$$F'(s) = -\sum_{n=1}^{\infty} \frac{f(n)\log n}{n^s},$$

*which is obtained by differentiating term by term.*

**Proof.** *Simply apply that last few lemmas to the sequence of partial sums.*

**Theorem.** *Let $F(s)$ be represented in the half-plane $\sigma > c$ by the Dirichlet series*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

*where $c$ is finite, and assume that $f(n) \geq 0$ for all $n \geq n_0$. If $F(s)$ is analytic in some disk about the point $s = c$, then the Dirichlet series converges in the half-plane $\sigma > c - \varepsilon$ for some $\varepsilon > 0$. Therefore, if the Dirichlet series has a finite abscissa of convergence $\sigma_c$, then $F(s)$ has a singularity on the real axis at the point $s = \sigma_c$.*

**Proof.** *Let $a = 1 + c$. Since $F$ is analytic at $a$, it can be represented by an absolutely convergent power series expansion about $a$,*

$$F(s) = \sum_{k=0}^{\infty} \frac{F^{(k)}(a)}{k!}(s-a)^k, \tag{A.5}$$

*and the radius of convergence for the power series is greater than 1 since $F$ is analytic at $c$ (see A.2). By the last theorem, we know that the derivatives $F^{(k)}(a)$ can be determined by repeated differentiation. This gives us*

$$F^{(k)}(a) = (-1)^k \sum_{n=1}^{\infty} \sum f(n)(\log n)^k n^{-a}.$$

*Thus we can rewrite the power series as*

$$F(s) = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{(a-s)^k}{k!} f(n)(\log n)^k n^{-a}.$$

*Now since the radius of convergence exceeds 1, this formula is valid for some real $s = c - \varepsilon$ where $\epsilon > 0$ as in Figure A.2. Then $a - s = 1 + \varepsilon$ for this $s$ and the double series has nonnegative terms for $n \geq n_0$. Therefore, we may interchange the order of summation in order to obtain*

$$F(c-\varepsilon) = \sum_{n=1}^{\infty} \frac{f(n)}{n^a} \sum_{k=0}^{\infty} \frac{\{(1+\varepsilon)\log n\}^k}{k!} = \sum_{n=1}^{\infty} \frac{f(n)}{n^a} e^{(1+\varepsilon)\log n} = \sum_{n=1}^{\infty} \frac{f(n)}{n^{c-\varepsilon}}.$$

*That is, the Dirichlet series $\sum f(n)n^{-s}$ converges for $s = c - \varepsilon$, hence it also converges in the half-plane $\sigma > c - \varepsilon$.*

## A.5 Dirichlet Characters

**Theorem.** *If $f$ is a character of a finite group $G$ with identity element $e$, then $f(e) = 1$, and each function value $f(a)$ is a root of unity. In fact, if $a^n = e$, then $[f(a)]^n = 1$.*

**Proof.** *Choose $c$ in $G$ such that $f(c) \neq 0$. Since $ce = c$, we have*

$$f(c)f(e) = f(c)$$

*so $f(e) = 1$. If $a^n = e$, then $[f(a)]^n = f(a^n) = f(e) = 1$.*

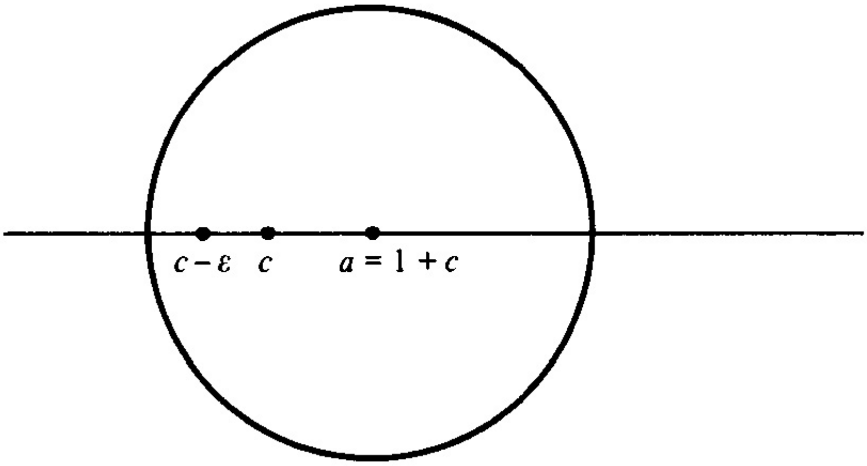


Figure A.2:

**Theorem.** *A finite abelian group $G$ of order $n$ has exactly $n$ distinct characters.*

**Proof.** *We prove the theorem by induction on the order of $G$, but first we must construct proper subgroups of $G$. We will use the symbol $< G'; a >$ to indicate a proper subgroup $G''$ of $G$ constructed from the proper subgroup $G'$ with the element $a$ not in $G'$ as follows:*

$$G'' :=< G'; a >:= \{xa^k : x \in G' \text{ and } 0 \leq k < h\}$$

*where $h$ is the indicator of $a$ in $G'$.*

*We apply this construction from the bottom up, so to speak, beginning with $G_1 = \{e\}$ and progressing until we reach $G$. That is, if $G_1 \neq G$, let $a_1$ be an element of $G$ other than $e$ and define $G_2 =< G_1; a_1 >$. If $G_2 \neq G$, let $a_2$ be an element of $G$ not in $G_2$ and define $G_3 =< G_2; a_2 >$. Continue to obtain a finite set of elements $a_1, a_2, \ldots, a_t$ and a corresponding set of* subgroups $G_1, G_2, \ldots, G_{t+1}$ such that

$$G_{r+1} =< G_r; a_r >$$

with

$$G_1 \subset G_2 \subset \cdots \subset G_{t+1} = G.$$

We are guaranteed $t < \infty$ since $G$ is finite. Now we are set up to begin our proof by induction.

It is obvious that there is only one character for $G_1$, i.e. the function which is identically $1$. Assume the inductive hypothesis of the subgroup $G_r$ has order $m$ and that there are exactly $m$ distinct characters for $G_r$. Let us consider $G_{r+1} =< G_r; a_r >$, and let $h$ be the indicator of $a_r$ in $G_r$, that is, let $h$ be the smallest positive integer such that $a_r^h \in G_r$.

We will show that there are exactly $h$ different ways to extend each character of $G_r$ to obtain a character of $G_{r+1}$ and that each character of $G_{r+1}$ is the extension of some character of $G_r$, proving that there are exactly $mh$ characters of $G_{r+1}$, which happens to be its order as well.

A typical element of $G_{r+1}$ has the form

$$xa_r^k, \text{ where } x \in G_r \text{ and } 0 \leq k < h.$$

Suppose that it is possible to extend a character $f$ of $G$ to $G_{r+1}$. Call this extension $\tilde{f}(x)$ and let us examine $\tilde{f}(xa_r^k)$. The multiplicative property requires

$$\tilde{f}(xa_r^k) = \tilde{f}(x)\tilde{f}(a_r)^k,$$

but $x \in G_r$, so $\tilde{f}(x) = f(x)$ and thus

$$\tilde{f}(xa_r^k) = f(x)\tilde{f}(a_r)^k.$$

Which tells us that $\tilde{f}(xa_r^k)$ is known as soon as $\tilde{f}(a_r)$ is.

Now we are concerned with the possible values of $\tilde{f}(a_r)$. Let $c = a_r^h$. Since $c \in G_r$, we have $\tilde{f}(c) = f(c)$, and since $\tilde{f}$ is multiplicative, we also have $\tilde{f}(c) = \tilde{f}(a_r)^h$. Thus

$$\tilde{f}(a_r)^h = f(c),$$

so $\tilde{f}(a_r)$ is an $h$th root of $f(c)$. Therefore, there are at most $h$ choices for $\tilde{f}(a_r)$.

Now we can define $\tilde{f}$. If $f$ is a given character of $G$, then choose one of the $h$th roots of $f(c)$, where $c = a_r^k$, and define $\tilde{f}(a_r)$ to be this root. Then define $\tilde{f}$ on the rest of $G_{r+1}$ by

$$\tilde{f}(xa_r^k) := f(x)\tilde{f}(a_r)^k. \tag{A.6}$$

*The $h$ choices for $\tilde{f}(a_r)$ are all different, so this gives us $h$ different ways to define $\tilde{f}(xa_r^k)$. Now we verify that $\tilde{f}$ has the multiplicative property. From* (A.6) *we obtain*

$$\begin{aligned}\tilde{f}(xa_r^k \cdot ya_r^j) &= \tilde{f}(xy \cdot a_r^{k+j}) = f(xy)\tilde{f}(a_r)^{k+j}\\ &= f(x)f(y)\tilde{f}(a_r)^k\tilde{f}(a_r)^j\\ &= \tilde{f}(xa_r^k)\tilde{f}(ya_r^j),\end{aligned}$$

*so $\tilde{f}$ is a character on $G_{r+1}$. Furthermore, no two of the extensions $\tilde{f}$ and $\tilde{g}$ can be identical on $G_{r+1}$ because the functions $f$ and $g$ which they extend would then be identical on $G_r$. Therefore, each of the $m$ characters of $G_r$ can be extended in $h$ different ways to produce a character on $G_{r+1}$. Finally, if $\psi$ is any character of $G_{r+1}$, then restricting it to $G_r$ results in a character of $G_r$, so the extension process gives us all of the characters of $G_{r+1}$, which was to be shown.*

**Theorem.** *With multiplication defined by* (2.5)*, the set of reduced residue classes modulo $k$ is a finite abelian group of order $\varphi(k)$. The identity is the residue class $\hat{1}$. The inverse of $\hat{a}$ is the residue class $\hat{b}$ where $ab \equiv 1 \pmod{k}$.*

**Proof.** *The closure property is automatically satisfied by the definition of multiplication of residue classes. The class $\hat{1}$ is obviously the identity element also. If $(a, k) = 1$, then there is a unique $b$ such that $ab \equiv 1 \pmod{k}$. Hence, the inverse of $\hat{a}$ is $\hat{b}$. Finally, it is clear that the group is abelian and of order $\varphi(k)$.*

## A.6 Properties of $\zeta(s)$

**Lemma.** *If $f$ is a completely multiplicative function, and the series*

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

*converges absolutely for $s > s_0$, then*

$$\left(\sum_{n=1}^{\infty} \frac{f(n)}{n^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{f(n)\mu(n)}{n^s}$$

*for $s > s_0$.*

**Proof.** *We have the following,*

$$\sum_{m=1}^{\infty}\frac{f(m)}{m^s}\sum_{n=1}^{\infty}\frac{f(n)\mu(n)}{n^s} = \sum_{m,n=1}^{\infty}\frac{f(mn)\mu(n)}{(mn)^s}$$
$$= \sum_{j=1}^{\infty}\frac{\sum_{d|j}\mu(d)}{j^s}f(j) = 1.$$

*which proves the lemma.*

**Theorem.** *For $\sigma > 1$*

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

*where $\mu$ is the Möbius function.*

**Proof.** *This theorem follows immediately by the previous lemma for all s real and greater than 1. Furthermore, by the analytic continuation of $\zeta$, we also have that this theorem holds for all $\sigma > 1$.*

## A.7 Riemann Zeta Function and Dirichlet $L$-Functions

### A.7.1 Hurwitz Zeta Function

**Theorem.** *For $\sigma > 1$*

$$L(s,\chi) = \frac{1}{k^s}\sum_{w=1}^{k}\chi(a)\zeta\left(s,\frac{w}{k}\right). \tag{A.7}$$

**Proof.** *Since $\chi$ is periodic with period k,*

$$L(s,\chi) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$$
$$= \sum_{w=1}^{k}\chi(w)\sum_{m=0}^{\infty}\frac{1}{(km+w)^s}$$
$$= \frac{1}{k^s}\sum_{w=1}^{k}\chi(w)\zeta\left(s,\frac{w}{k}\right).$$

**Theorem.** *For any $\sigma_0 > 1$, the series*

$$\sum_{n=0}^{\infty}(n+w)^{-s}$$

*converges uniformly for $\sigma \geq \sigma_0$; thus $\zeta(s,w)$ is analytic for $\sigma > 1$.*

**Proof.** *We have*

$$\left|(n+w)^{-s}\right| = \left|e^{-(\sigma+it)\log(n+w)}\right| = e^{-\sigma\log(n+w)} = (n+w)^{-\sigma},$$

*hence, for* $\sigma \geq \sigma_0$,

$$\sum_{n=0}^{\infty}(n+w)^{-s} \ll \sum_{n=0}^{\infty}(n+w)^{-\sigma_0}.$$

*Therefore, we have a series of analytic functions which is dominated throughout the region* $\sigma \geq \sigma_0$ *by a convergent series of positive constants, which is therefore uniformly convergent. Thus, by Weierstrass' M-test, the theorem follows.*

**Lemma.** *If* $a$ *and* $b$ *are integers with* $0 \leq a < b$, *and if* $f$ *has a continuous derivative over* $a \leq x \leq b$, *then*

$$\sum_{n=a+1}^{b} f(n) = \int_a^b f(u)du + \int_a^b (u - \lfloor u \rfloor) f'(u)du.$$

**Proof.**

$$\begin{aligned}\int_{n-1}^{n} uf'(u)du &= nf(n) - (n-1)f(n-1) - \int_{n-1}^{n} f(u)du\\ &= f(n) + (n-1)\int_{n-1}^{n} f'(u)du - \int_{n-1}^{n} f(u)du\\ &= f(n) + \int_{n-1}^{n} \lfloor u \rfloor f'(u)du - \int_{n-1}^{n} f(u)du,\end{aligned}$$

*and the result follows by summing from* $a+1$ *to* $b$ *on* $n$.

**Theorem.** *If* $m$ *is a non-negative integer, and* $\sigma > 1$, *then*

$$\zeta(s,w) - \frac{1}{(s-1)(m+w)^{s-1}} = \sum_{n=0}^{m} \frac{1}{(n+w)^s} - s\int_m^{\infty} \frac{u - \lfloor u \rfloor}{(u+w)^{s-1}} du. \quad \text{(A.8)}$$

*It follows that* $\zeta(s,w) - 1/(s-1)$ *is analytic for* $\sigma > 0$, *and that* (A.8) *holds for* $\sigma > 0$.

**Proof.** *If* $\sigma > 1$ *and*

$$f(u) = \frac{1}{(u+w)^s},$$

*then the equation of the previous theorem continues to hold for* $b \to \infty$, *and, if we replace* $a$ *by* $m$, *we obtain*

$$\sum_{n=m+1}^{\infty} \frac{1}{(n+w)^s} = \frac{1}{(s-1)(m+w)^{s-1}} - s\int_m^{\infty} \frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}du,$$

*from which* (A.8) *follows. Now, since*

$$\left|\frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}\right| < \frac{1}{(u+w)^{\sigma+1}} < \frac{1}{u^{\sigma+1}},$$

*the integral on the right hand side of* (A.8) *converges absolutely for $\sigma > 0$ and uniformly for $\sigma \geq \sigma_0 > 0$. Now, for arbitrary $n \geq 0$, we have that the quantity*

$$\int_n^{n+1} \frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}du = \int_n^{n+1} \frac{u-n}{(u+w)^{s+1}}du$$

*is an analytic function of $s$ for $\sigma > 0$, and the same is true for*

$$\sum_{n=m}^{\infty} \int_n^{n+1} \frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}du = \int_m^{\infty} \frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}du,$$

*with $m \geq 0$. At last, taking $m = 0$ in* (25), *we have that*

$$\zeta(s,w) - \frac{1}{s-1} = \frac{1}{w^s} + \frac{w^{1-s}-1}{s-1} - s\int_0^{\infty} \frac{u-\lfloor u\rfloor}{(u+w)^{s+1}}du$$

*where the right hand side is analytic for $\sigma > 0$.*

**Theorem.** *For $\frac{1}{2} \leq \sigma \leq 2$ and $t > c(w)$, where $c$ is a function of $w$,*

$$|\zeta(s,w)| < t^{3/4}.$$

*For $t \geq 8$ and $1-(\log t)^{-1} \leq \sigma \leq 2$, we have*

$$|\zeta(s,w)| < c(w)\log t.$$

**Proof.** *For $1/2 \leq \sigma \leq 2$, and $t \geq 3$, we have $|s| < 2+t < 2t$ and $|s-1| >\geq t > 1$. Therefore, if we let $m = \lfloor t\rfloor + 1$ in the previous theorem, we get*

$$\begin{aligned}|\zeta(s,w)| &< \frac{1}{(\lfloor t\rfloor+1+w)^{\sigma-1|}} + \sum_{n=1}^{\lfloor t\rfloor+1} \frac{1}{n^\sigma} + 2t\int_t^{\infty} \frac{du}{u^{\sigma+1}} \\ &\leq \frac{1}{(\lfloor t\rfloor+1+w)^{\sigma+1}} + c(w) + \sum_{n=1}^{\lfloor t\rfloor} \frac{1}{n^\sigma} + \frac{2t}{\sigma t^\sigma},\end{aligned}$$

*or rather,*

$$|\zeta(s,w)| < \frac{1}{(\lfloor t\rfloor + 1 + w)^{\sigma+1}} + c(w) + \sum_{n=1}^{\lfloor t\rfloor} \frac{1}{n^{\sigma}} + 4t^{1-\sigma}. \tag{A.9}$$

*Hence, for the same range of $\sigma$ and $t$, we have*

$$\begin{aligned}|\zeta(s,w)| &< \frac{1}{(\lfloor t\rfloor + 1 + w)^{-1/2}} + c(w) + \sum_{n=1}^{\lfloor t\rfloor} \frac{1}{\sqrt{n}} + 4\sqrt{t} \\ &< 2\sqrt{t} + c(w) + \int_0^t \frac{du}{\sqrt{u}} + 4\sqrt{t} \leq 8\sqrt{t} + c(w),\end{aligned}$$

*and that this is smaller than $t^{3/4}$ for $t > c(w)$.*

*Now let $t \geq 8 > e^2$. Then $1-(\log t)^{-1} \geq 1/2$; therefore, if $1-(\log t)^{-1} \leq \sigma \leq 2$, (A.9) gives us*

$$\begin{aligned}|\zeta(s,w)| &< (2t)^{1/\log t} + c(w) + \sum_{n=1}^{\lfloor t\rfloor} \frac{n^{1/\log t}}{n} + 4t^{1/\log t} \\ &< 2^{1/2}e + c(w) + e\sum_{n=1}^{\lfloor t\rfloor} \frac{1}{n} + 4e \\ &< c(w)\log t.\end{aligned}$$

*Proving the theorem.*

**Theorem.** *For $|x| \leq 1$, if*

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

*is analytic, and $\operatorname{Re} f(x) \leq \frac{1}{2}$, then $|a_n| \leq 1$ for $n \geq 1$.*

**Proof.** *We have $|f(x)| \leq |1 - f(x)|$ for $|x| \leq 1$, therefore, the function*

$$\frac{f(x)}{1-f(x)} = \frac{a_1 x + \cdots}{1 - a_1 x - \cdots} = a_1 x + b_2 x^2 + \cdots$$

*is analytic and has modulus at most $1$ for $|x| \leq 1$. But if we define a function, $f_1$,*

$$f_1(x) := \frac{f(x)}{x\,(1-f(x))},$$

*then this is also analytic for $|x| \leq 1$, and its value at $x = 0$ is $a_1$. By the maximum-modulus principle, its absolute value is at least as large at some point on $|x| = 1$. Because for $|x| = 1$,*

$$|f_1(x)| = \left| \frac{f(x)}{1 - f(x)} \right|,$$

*it follows that*

$$|a_1| \leq 1. \tag{A.10}$$

*So it remains to show that each of the functions*

$$F_n(x) = a_n x + a_{2n} x^2 + \cdots$$

*fulfills the same hypotheses as* $f(x)$*. This depends on the fact that if* $\eta = e^{2\pi i/n}$*, then*

$$\sum_{l=0}^{n-1} \eta^{lk} = \begin{cases} n \text{ if } n \big| k, \\ \Big(\eta^{kn} - 1\Big) \Big/ \Big(\eta^k - 1\Big) = 0 \text{ if } n \nmid k. \end{cases}$$

*We have that*

$$\begin{aligned} \sum_{l=0}^{n-1} f(\eta^l x) &= \sum_{l=0}^{n-1} \sum k = 1^\infty a_k \eta^{kl} x^k \\ &= \sum_{k=1}^{\infty} a_k x^k \sum_{l=0}^{n-1} \eta^{kl} \\ &= n \sum_{n|k} a_k x^k \\ &= n F_n(x^n), \end{aligned}$$

*therefore,* $F_n(x)$ *is analytic for* $|x| \leq 1$*, and for such* $x$*, we have*

$$\operatorname{Re} F_n(x^n) = \frac{1}{n} \sum_{l=0}^{n-1} \operatorname{Re} f(n^l x) \leq \frac{1}{n} \sum_{l=0}^{n-1} \frac{1}{2} = \frac{1}{2}.$$

**Lemma.** *Let* $R > 0$*, and suppose that*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

*is analytic and* $\operatorname{Re} f(x) \leq M$ *for* $|x - x_0| \leq R$*. For* $n \geq 1$*, we have*

$$|a_n| \leq \frac{2}{R^n}(M - \operatorname{Re} a_0).$$

**Proof.** *If* $\operatorname{Re} a_0 = M$*, then* $a_n = 0$ *for* $n \geq 1$*, by the maximum-modulus principle.*

*If* $\operatorname{Re} a_0 < M$*, set*

$$g(x) = \frac{f(x_0 + Rx) - a_0}{2(M - \operatorname{Re} a_0)}.$$

*Then we have* $g$ *analytic for* $|x| \leq 1$*,* $g(0) = 0$*, and that*

$$\operatorname{Re} g(x) = \frac{\operatorname{Re} f(x_0 + Rx) - \operatorname{Re} a_0}{2(M - \operatorname{Re} a_0)} \leq \frac{M - \operatorname{Re} a_0}{2(M - \operatorname{Re} a_0)} = \frac{1}{2}.$$

*Therefore,* $g$ *satisfies the hypotheses of Theorem 27 and we have*

$$\left| \frac{a_n R^n}{2(M - \operatorname{Re} a_0)} \right| \leq 1,$$

*from which the theorem follows.*

**Theorem.** *If* $f$ *satisfies the hypotheses of Lemma 8, and* $0 < r < R$*, then for* $|x - x_0| \leq r$*, we have*

$$|f(x)| \leq |a_0| + \frac{2r}{R - r}(|M| + |a_0|)$$

*and*

$$|f'(x)| \leq \frac{2R}{(R - r)^2}(|M| + |a_0|).$$

**Proof.** *The theorem follows quickly: we have*

$$\begin{aligned} |f(x)| &\leq |a_0| + \sum_{n=1}^{\infty} |a_n| r^n \\ &\leq |a_0| + 2\left(|M| + |a_0|\right) \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^n \\ &= |a_0| + \frac{2r}{R - r}\left(|M| + |a_0|\right), \end{aligned}$$

*and*

$$\begin{aligned} |f'(x)| &\leq \sum_{n} = 1^{\infty} |a_n| n r^{n-1} \\ &\leq \frac{2\left(|M| + |a_0|\right)}{R} \sum_{n=1}^{\infty} n \left(\frac{r}{R}\right)^{n-1} \\ &= \frac{2R}{(R - r)^2}\left(|M| + |a_0|\right). \end{aligned}$$

**Theorem.** *Let $r > 0$ and $M \in \mathbb{R}$, and suppose that $f(s_0) \neq 0$ and that, for $|s - s_0| \leq r$, $f(s)$ is analytic and*

$$\left| \frac{f(s)}{f(s_0)} \right| < e^M.$$

*Further suppose that $f(s) \neq 0$ in the semicircular region $|s - s_0| \leq r$, $\operatorname{Re} s > \operatorname{Re} s_0$. Then*

$$-\operatorname{Re} \frac{f'}{f}(s_0) \leq \frac{4M}{r},$$

*and if there is a zero, say $\rho$, of $f$ on the open line segment between $s_0 - r/2$ and $s_0$, then*

$$-\operatorname{Re} \frac{f'}{f}(s_0) \leq \frac{4M}{r} - \frac{1}{s_0 - \rho.}$$

**Theorem.**

$$\frac{1}{2\pi i} \int_{(2)} \frac{y^s}{s^2} = \begin{cases} 0 \text{ for } 0 < y < 1, \\ 1 \text{ for } y \geq 1. \end{cases}$$

### A.7.2 Analytic Proof of Dirichlet's Theorem

As mentioned in the preliminaries, with the machinery of the prime number theorem and its proof, we are easily able to prove Dirichlet's theorem. This work comes from [LeV2] in particular, but like many things concerning the Prime Number Theorem, can be found in a variety of places.

Let $k$ and $l$ be relatively prime integers, and $\pi(x; k, l)$ be the number of primes $p \equiv l \pmod k$ which do not exceed $x$. For a given $k$, there are $\varphi(k) = h$ choices of $l$ which are distinct modulo $k$, so that if the primes are more or less evenly dispersed amongst the various arithmetic progressions, it is expected that

$$\pi(x; k, l) \sim \frac{1}{h} \frac{x}{\log x}.$$

In fact this is the case, and this result is known as Dirichlet's theorem. To obtain an estimate for $\pi(x; k, l)$, we go through similar arguments for our proof of the Prime Number Theorem, which was concerned with $\pi(x) = \pi(x; 1, 0)$.

We have already that for $\sigma > 1$,

$$L(s, \chi) = \frac{1}{k^s} \sum_{a=1}^{k} \chi(a) \zeta\left(s, \frac{a}{k}\right).$$

But in fact, the domain for this can be extended. If we let

$$E(\chi) = \begin{cases} 1, \text{ if } \chi = \chi_0 \\ 0, \text{ if } \chi \neq \chi_0, \end{cases}$$

then

$$\sum_{a=1}^{k} \chi(a) = E(\chi) \cdot h.$$

Therefore, for $\sigma > 1$, we have

$$\begin{aligned} L(s,\chi) &- \frac{E(\chi) \cdot h}{k} \cdot \frac{1}{s-1} \\ &= \frac{E(\chi) \cdot h}{s-1}\left(\frac{1}{k^s} - \frac{1}{k}\right) + \frac{1}{k^s}\sum_{a=1}^{k} \chi(a)\left\{\zeta\left(s, \frac{a}{k}\right) - \frac{1}{s-1}\right\}. \end{aligned}$$

Now, by Theorem 25, which concerns the analyticity of functions like the Hurwitz zeta function, we have that each term on the right is analytic for $\sigma > 0$, and also

$$\frac{1}{s-1}\left(\frac{1}{k^s} - \frac{1}{k}\right) = \frac{k^{1-s} - 1}{k(s-1)}$$

is an integral function. Now by analytic continuation, we have the next theorem.

**Theorem 43.**

$$L(s,\chi) = \frac{1}{k^s}\sum_{a=1}^{k} \chi(a)\zeta\left(s, \frac{a}{k}\right).$$

*holds for $\sigma > 0$ except at $s = 1$. Furthermore,*

$$\lim_{s\to 1}(s-1)L(s,\chi) = \frac{h \cdot E(\chi)}{k}. \tag{A.11}$$

*Therefore, $L(s,\chi)$ is analytic for $\sigma > 0$, except that $L(s,\chi_0)$ has a simple pole at $s = 1$.*

Now we utilize Theorem 26. As is to be expected, we are concerned with $\sigma > 2$ and $t \geq 8$. We have

$$|L(s,\chi)| < \sum_{n=1}^{\infty} \frac{1}{n^2} < 2 < \begin{cases} t, \\ \log t, \end{cases}$$

while for $\sigma > 0$ and $t > 0$, we have

$$|L(s,\chi)| \leq \sum_{a=1}^{k} \left| \zeta\left(s, \frac{a}{k}\right) \right|.$$

Now Theorem 26 gives us the following:

**Theorem 44.** *(I) For $\sigma > 1/2$ and $t > c_{12}(k)$, we have $|L(s,\chi)| < t$. (II) For $t \geq 8$ and $\sigma > 1 - (\log t)^{-1}$, we have $|L(s,\chi)| < c_{13}(k) \log t$.*

Now we can generalize the proof that $\zeta(s)$ does not vanish on $\sigma = 1$ in a simple way.

**Theorem 45.** *$L(s,\chi)$ does not vanish on the line $\sigma = 1$.*

**Proof.** *For $\sigma > 1$,*

$$L(s,\chi) = \prod_p \left(1 - \chi(p)p^{-s}\right)^{-1},$$

*so we can choose*

$$\log L(s,\chi) = \sum_{m,p} \frac{\chi(p^m)}{mp^{ms}}.$$

*Therefore,*

$$\begin{aligned}
&\log \left|L^3(\sigma,\chi_0)L^4(\sigma+ti,\chi)L(\sigma+2ti,\chi^2)\right| \\
&\quad = 3\log|L(\sigma,\chi_0)| + 4\log|L(\sigma+ti,\chi)| + \log|L(\sigma+2ti,\chi^2)| \\
&\quad = 3\log L(\sigma,\chi_0) + 4\operatorname{Re}\log L(\sigma+ti,\chi) + \operatorname{Re}\log L(\sigma+2ti,\chi^2)| \\
&\quad = \sum_{m,p} \left( \frac{3\chi_0(p^m)}{mp^{m\sigma}} + \operatorname{Re}\frac{4\chi(p^m)}{mp^{m(\sigma+ti)}} + \operatorname{Re}\frac{\chi^2(p^m)}{mp^{m(\sigma+2ti)}} \right) \\
&\quad = \sum_{\substack{m,p \\ p\nmid k}} \frac{3 + 4\cos(\eta(p^m) - t\log p^m) + \cos 2(\eta(p^m) - t\log p^m)}{mp^{m\sigma}} \\
&\quad \geq 0,
\end{aligned}$$

*where $\chi(p^m) = e^{i\eta(p^m)}$. Hence,*

$$((\sigma-1)L(\sigma,\chi_0))^3 \left|\frac{L(\sigma+ti,\chi)}{\sigma-1}\right|^4 \left|L(\sigma+2ti,\chi^2)\right| \geq \frac{1}{\sigma-1},$$

*but then if the theorem were false, then Theorem 43 would be also, which is absurd.*

Now we clearly have that Theorem 31 has the following analog:

**Theorem 46.** *For $\sigma > 1$,*

$$-3\frac{L'}{L}(\sigma, \chi_0) - 4\operatorname{Re}\frac{L'}{L}(\sigma + ti, \chi) - \operatorname{Re}\frac{L'}{L}(\sigma + 2ti, \chi^2) \geq 0.$$

Similarly, Theorem 33 has the analog:

**Theorem 47.** *There is a $c_1(k) \geq 8$ such that $L(s, \chi) \neq 0$ for $t > c_1(k)$ and $\sigma \geq 1 - c_2/\log t$.*

The only differences between the proofs of these theorems and their analogs is that we use $f(s) = L(s, \chi^2)$ and $s_0 = \sigma + 2ti$, and also $f(s) = L(s, \chi)$ with $s_0 = \sigma + ti$, and the constants now depend on $k$.

In the same manner, we replace $\zeta(s)$ with $L(s, \chi)$ in Theorem 34 to get the next theorem.

**Theorem 48.** *For $t \geq c_9(k) > 8$ and $\sigma \geq 1 - c_8(\log t)^{-1}$, $|\log L(s, \chi)| < \log^2 t$.*

Here the analogs are not quite so clear, so for Theorem 35, we break the argument into two parts.

**Theorem 49.** *For $(k, l) = 1$, we have*

$$\sum_{\substack{p \leq x \\ p \equiv l \pmod k}} \log\frac{x}{p} = \frac{1}{2\pi i h}\sum_{\chi}\frac{1}{\chi(l)}\int_{(2)}\frac{x^s}{s^2}\log L(s, \chi)ds + O(\sqrt{x}\log^2 x).$$

**Proof.** *We utilize the series expansion for $L(s, \chi)$ to obtain,*

$$\begin{aligned}
\frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^2}\log L(s, \chi)ds &= \frac{1}{2\pi i}\int_{(2)}\frac{x^s}{s^2}\sum_{m,p}\frac{\chi(p^m)}{mp^{ms}}ds \\
&= \frac{1}{2\pi i}\sum_{m,p}\frac{\chi(p^m)}{m}\int_{(2)}\frac{(x/p^m)^s}{s^2}ds \\
&= \sum_{\substack{m,p \\ p^m \leq x}}\frac{\chi(p^m)\log(x/p^m)}{m} \\
&= \sum_{\substack{p \leq x \\ p \nmid k}}\chi(p)\log\frac{x}{p} + \sum_{\substack{p^m \leq x \\ m \geq 2}}\frac{\chi(p^m)\log(x/p^m)}{m} \\
&= \sum_{\substack{p \leq x \\ p \nmid k}}\chi(p)\log\frac{x}{p} + O(\sqrt{x}\log^2 x).
\end{aligned}$$

*Now we multiply by $1/\chi(l)$ and sum over all characters modulo $k$:*

$$\sum_{\chi} \frac{1}{\chi(l)} \sum_{\substack{p \le x \\ p \nmid k}} \chi(p) \log \frac{x}{p} = h \sum_{\substack{p \le x \\ p \equiv l \pmod{k}}} \log \frac{x}{p}$$

$$= \frac{1}{2\pi i} \sum_{\chi} \frac{1}{\chi(l)} \int_{(2)} \frac{x^s}{s^2} \log L(s,\chi) ds + O(\sqrt{x} \log^2 x),$$

*which is the theorem. Note that the implied constant in the $O$ may depend on $k$.*

To estimate these integrals, we need to consider two cases. First, the case when $\chi = \chi_0$. Fortunately, every property we used to estimate

$$\int_{(2)} \frac{x^s}{s^2} \log \zeta(s) ds$$

carries over for

$$\int_{(2)} \frac{x^s}{s^2} \log L(s,\chi_0) ds.$$

Quickly, we have that for suitable $c$ with $0 < c < 1$,

$$\int_{(2)} \frac{x^s}{s^2} \log L(s,\chi_0) ds = 2\pi i \int_c^1 \frac{x^s}{s^2} ds + O\left(x e^{-\alpha\sqrt{\log x}}\right).$$

However, if $\chi \neq \chi_0$, then $L(s,\chi)$ has no pole at $s = 1$, but the other properties still apply thankfully. Therefore, if we do not cut the plane, but instead consider the line segments $\Gamma_5$ and $\bar{\Gamma}_5$ as a single segment, say $\Gamma_8$ and omit $\Gamma_6, \Gamma_7$, and $\bar{\Gamma}_6$, then

$$\frac{x^s}{s^2} \log L(s,\chi)$$

is analytic in the region bounded by $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_8, \bar{\Gamma}_4, \bar{\Gamma}_3, \bar{\Gamma}_2, \bar{\Gamma}_1$. Thus,

$$\int_{(2)} \frac{x^s}{s^2} \log L(s,\chi) ds = \left( \int_{2-\infty i}^{2+ui} - \int_{\bar{\Gamma}_2+\bar{\Gamma}_3+\bar{\Gamma}_4+\Gamma_8+\Gamma_4+\Gamma_3+\Gamma_2} + \int_{2+ui}^{2+\infty i} \right) \frac{x^s}{s^2} \log L(s,\chi) ds.$$

Furthermore, the integral along each of these new arcs either tends to zero or is

$$O\left(x e^{-\alpha\sqrt{\log x}}\right).$$

Now

$$\sum_{\substack{p \le x \\ p \equiv l \pmod{k}}} \log \frac{x}{p} = \frac{1}{h} \int_c^1 \frac{x^s ds}{s^2} + O\left(x e^{-\alpha\sqrt{\log x}}\right),$$

follows, and leads us to the analog of the Prime Number Theorem in the same way its predecessor led to the Prime Number Theorem itself.

**Theorem 50.** *If $k$ is a fixed integer and $(k,l)=1$, then, as $x \to \infty$,*

$$\pi(x;k,l) = \frac{1}{\varphi(k)} \int_2^x \frac{du}{\log u} + O\left(xe^{-\alpha\sqrt{\log x}}\right).$$

And in particular, this leads us to

$$\pi(x;k,l) \sim \frac{1}{\phi(k)} \frac{x}{\log x},$$

and that if $(k,l_1) = (k,l_2) = 1$, then

$$\lim_{x\to\infty} \frac{\pi(x;k,l_1)}{\pi(x;k,l_2)} = 1,$$

so that asymptotically there are as many primes in the arithmetic progression $kt + l_1$ as there are in $kt + l_2$.

And as we have seen several times throughout this paper,

$$\lim_{x\to\infty} \frac{\pi(x;4,1)}{\pi(x;4,3)} = 1.$$

# Appendix B

# Concerning Integers $n$ such that $n = x^2 + y^2$

This appendix is primarily concerned with manipulating the integers that are expressible as a sum of two squares. While interesting, none of this information has a direct influence on the primary result of this paper, hence its inclusion here.

Most of our work here comes from [Sier]. First, we have a corollary to Theorem 40, which told us the type of integers capable of being expressed as the sum of two squares.

**Corollary 1.** *If a natural number is not representable as the sum of two squares of integers, then neither is it the sum of two squares of rational numbers.*

**Proof.** *Let $n$ be a natural number such that it is not a sum of two squares of integers. By Theorem 40, there must be a prime $p \equiv 3 \pmod 4$ that divides $n$ to an odd power.*

*We prove by contradiction. Assume $n = \left(\frac{l_1}{m_1}\right)^2 + \left(\frac{l_2}{m_2}\right)^2$ where $m_1, m_2$ are natural numbers and $l_1, l_2$ are integers. Then we clear the denominators and get*

$$(m_1 m_2)^2 n = (l_1 m_2)^2 + (l_2 m_1)^2.$$

*But $p$ must appear with an odd exponent in the factorization of the left-hand side, but this is absurd for the right-hand side. Hence the corollary is proved.*

## B.1 Decompositions

At least four decompositions of a prime into the sum of two squares are known, due to Legendre (1808), Gauss (1825), Serret (1848) and Jacobsthal

(1906). We examine only the construction of Gauss, which is the most elementary to formulate.

**Theorem 51.** *If $p = 4k + 1$ is a prime number, then let $x$, $y$ be integers, such that*

$$x \equiv \frac{(2k)!}{(k!)^2} \pmod{p} \text{ and } y \equiv (2k)!x \pmod{p},$$

*with $|x| < p/2$ and $|y| < p/2$. Then $p = x^2 + y^2$.*

A proof of this theorem has been given by Cauchy and another proof has been given by Jacobsthal, but neither is simple and we shall not attempt a proof here. We do however give an example of the difficulty in calculating $x$ and $y$.

Let $p = 37$. Then $k = 9$ and so

$$x \equiv \frac{18!}{2 \cdot (9!)^2} = \frac{6402373705728000}{263363788800} = 24310 \equiv 1 \pmod{37},$$

while

$$y \equiv 6402373705728000 \cdot 1 \equiv 6 \pmod{37}.$$

**Theorem 52.** *If $a$ and $b$ are natural numbers, then the representation of a prime $p$ of the form $p = ax^2 + by^2$, where $x, y$ are natural numbers, if it exists, is unique, apart from the possibility of interchanging $x$ and $y$ in the case of $a = b = 1$.*

**Proof.** *Suppose for a prime $p$,*

$$p = ax_1^2 + by_1^2 = ax_2^2 + by_2^2,$$

*where $x_1, x_2, y_1, y_2$ are natural numbers. Clearly, $(x_1, y_1) = (x_2, y_2) = 1$. We have*

$$\begin{aligned} p^2 &= (ax_1x_2 + by_1y_2)^2 + ab(x_1y_2 - x_2y_1)^2 \\ &= (ax_1x_2 - by_1y_2)^2 + ab(x_1y_2 + x_2y_1)^2, \end{aligned}$$

*but*

$$\begin{aligned} (ax_1x_2 + by_1y_2)(x_1y_2 + x_2y_1) &= (ax_1^2 + by_1^2)x_2y_2 + (ax_2^2 + by_2^2)x_1y_1 \\ &= p(x_1y_1 + x_2y_2). \end{aligned}$$

*Therefore at least one of the factors on the left hand side of the equality must be divisible by $p$.*

*If $p \big| ax_1x_2 + by_1y_2$, then the first of the equations for $p^2$ gives us $x_1y_2 - x_2y_1 = 0$, and therefore $x_1/y_1 = x_2/y_2$, which since they both have greatest common divisor of $1$, this proves that $x_1 = x_2$ and $y_1 = y_2$.*

*Now if $p \big| x_1y_2 + x_2y_1$, then the second of the equations for $p^2$ gives us $p^2 \geq abp^2$, which implies $a = b = 1$, but then $x_1x_2 - y_1y_2 = 0$, and now $x_1/y_1 = y_2/x_2$. Therefore $x_1 = y_2$ and $x_2 = y_1$, where the decompositions differ only in the order of summands. Hence the theorem is proved.*

**Corollary 2.** *Decomposition of any prime into the sum of two squares is unique up to order and sign.*

**Proof.** *In the previous theorem take $a = 1$ and $b = 1$, then we are done.*

**Corollary 3.** *If a natural number $n$ admits two or more different representations in the form $ax^2 + by^2$, where $x, y$ are natural numbers, then $n$ must be composite.*

**Proof.** *By our theorem, every prime number admits a unique representation, hence if a number does not admit a unique representation, then it is composite.*

Unfortunately, the converse of this corollary is not true. For instance, 14 has a unique representation with $a = 2$, $b = 3$; namely, $14 = 2 \cdot 1^2 + 3 \cdot 2^2$.

Thankfully, we do have a theorem concerning numbers of the form $4k+1$, but we shall save this for the primality test at the end of the chapter.

## B.2 Unsolved Questions

According to [Sier], the following items are unknown to be true or false:

**I.** There exist infinitely many primes $p$ such that $p = x^2 + (x+1)^2$, where $x$ is a natural number.

e.g. $5 = 1^2 + 2^2$, $13 = 2^2 + 3^2$, $41 = 4^2 + 5^2$, $61 = 5^2 + 6^2$, etc.

**II.** (I.) is equivalent to the conjecture that there exist infinitely many primes $p$ for which $2p = a^2 + 1$, where $a$ is a natural number.

**III.** There exist infinitely many primes $p$ such that $p = a^2 + b^2$, where $a$ and $b$ are prime.

e.g. $13 = 2^2 + 3^2$, $29 = 2^2 + 5^2$, etc.

**IV.** There exist infinitely many primes that are the sum of three squares of consecutive natural numbers.

e.g. $29 = 2^2 + 3^2 + 4^2$, $149 = 6^2 + 7^2 + 8^2$, etc.

**V.** There exist infinitely many primes that are the sum of three squares of primes.

e.g. $83 = 3^2 + 5^2 + 7^2$, $179 = 3^2 + 7^2 + 11^2$, etc.

Furthermore, if this conjecture holds, then one of the squares must always be $3^2$.

**VI.** For every natural number $n$, there exist infinitely many natural numbers $x$ such that $x^2 + n^2$ are primes.

i.e. for every $n \in \mathbb{N}$, the set $\{n^2 + x^2 : x \in \mathbb{N}\}$ contains infinitely many primes.

## B.3 A Primality Test

As promised, we have a theorem concerning numbers of the form $4k+1$, but first we need a lemma:

**Lemma 14.** *If each of two given natural numbers of the form $4k + 1$ with $k > 0$ is the sum of two squares of integers, then their product does not admit a unique representation as the sum of two squares of integers greater than or equal to zero and the squares are relatively prime.*

**Proof.** *Let $m = a^2 + b^2$, $n = c^2 + d^2$, where $a, b, c,$ and $d$ are integers. Then we have*

$$mn = (ac + bd)^2 + (ad - bc)^2 = (ac - bd)^2 + (ad + bc)^2.$$

*So we have two decompositions. Let us suppose they differ only by the order of factors. Then either $ac+bd = ad+bc$ or $ac+bd = |ac-bd|$. In our first case, we have $a(c - d) = b(c - d)$, but $c \neq d$, since otherwise $n = 2c^2$, which contradicts $n$ odd. But then we have $a = b$, and this is also absurd since $m$ is an odd number.*

*In the case where $ac+bd = |ac-bd|$, we have either $ac+bd = ac-bd$ or $ac+bd = bd-ac$. If the former holds, then $bd = 0$ and so $b = 0$ or $d = 0$. If $b = 0$, then $m = a^2$, where $a > 1$, and $mn = (ac)^2 + (ad)^2$, where $ac$ and $bd$ have a common divisor greater than $1$, and so does not satisfy our premise. Now if the latter holds, then $ac = 0$ and so $a = 0$ or $c = 0$, and a similar result follows.*

*Finally, if the decompositions of $mn$ differ in more than merely the order of the terms, then we have that $mn$ clearly does not admit a unique representation as the sum of two squares of integers greater than or equal to $0$ whose greatest common divisor is $1$.*

Now we can prove the following:

**Theorem 53.** *A natural number of the form $4k + 1 > 1$ is a prime if and only if it admits a unique representation (apart from the order of the terms)* *as the sum of two squares of integers $\geq 0$ and in this representation, the squares are relatively prime.*

**Proof.** *Suppose $p = 4k + 1$ is prime, then $p$ admits a unique representation of the form $p = x^2 + y^2$ where $x$ and $y$ are natural numbers. Thus the conditions of the theorem are necessary, and for sufficiency we have our lemma.*

As an application of Theorem 53, we have a primality test. In order to check if a number $n$ of the form $4k + 1$ is a prime, one must form the sequence of numbers

$$n - 0^2,\, n - 1^2,\ \ldots,\, n - (\lfloor\sqrt{n}\rfloor)^2$$

and check to see how many squares there are. [Sier] points out that this is the method utilized by T. Kulikowski to find that the number $2^{39} - 7$ is a prime number. That is,

$$2^{39} - 7 = 64045^2 + 738684^2$$

where 64045 and 738684 are relatively prime. In fact, the problem whether numbers of the form $2^n - 7$ was formulated by P. Erdös in 1956. For $n = 4, 5, \ldots, 38$, $2^n - 7$ is composite.

# Index