University
LECTURE
Series

Volume 62

Lectures on the Riemann Zeta Function

H. Iwaniec

American Mathematical Society

http://dx.doi.org/10.1090/ulect/062

Lectures on the Riemann Zeta Function

University
LECTURE
Series

Volume 62

Lectures on the Riemann Zeta Function

H. Iwaniec

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification*. Primary 11N05; Secondary 11N37.

For additional information and updates on this book, visit
www.ams.org/bookpages/ulect-62

Library of Congress Cataloging-in-Publication Data

Iwaniec, Henryk.
Lectures on the Riemann zeta function / H. Iwaniec.
pages cm. — (University lecture series ; volume 62)
Includes bibliographical references and index.
ISBN 978-1-4704-1851-9 (alk. paper)
1. Riemann, Bernhard, 1826–1866. 2. Functions, Zeta. 3. Riemann hypothesis. 4. Numbers, Prime. I. Title. II. Title: Riemann zeta function.

QA351.I93 2014
515′.56—dc23

2014021164

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at http://www.ams.org/

10 9 8 7 6 5 4 3 2 1 19 18 17 16 15 14

Contents

Preface

The Riemann zeta function $\zeta(s)$ in the real variable s was introduced by L. Euler (1737) in connection with questions about the distribution of prime numbers. Later B. Riemann (1859) derived deeper results about the prime numbers by considering the zeta function in the complex variable. He revealed a dual correspondence between the primes and the complex zeros of $\zeta(s)$, which started a theory to be developed by the greatest minds in mathematics. Riemann was able to provide proofs of his most fundamental observations, except for one, which asserts that all the non-trivial zeros of $\zeta(s)$ are on the line $\operatorname{Re} s = \frac{1}{2}$. This is the famous Riemann Hypothesis – one of the most important unsolved problems in modern mathematics.

These lecture notes cover closely the material which I presented to graduate students at Rutgers in the fall of 2012. The theory of the Riemann zeta function has expanded in different directions over the past 150 years; however my goal was limited to showing only a few classical results on the distribution of the zeros. These results include the Riemann memoir (1859), the density theorem of F. Carlson (1920) about the zeros off the critical line, and the estimates of G. H. Hardy - J. E. Littlewood (1921) for the number of zeros on the critical line.

Then, in Part 2 of these lectures, I present in full detail the result of N. Levinson (1974), which asserts that more than one third of the zeros are critical (lie on the line $\operatorname{Re} s = \frac{1}{2}$). My approach had frequent detours so that students could learn different techniques with interesting features. For instance, I followed the stronger construction invented by J. B. Conrey (1983), because it reveals clearly the esssence of Levinson's ideas.

After establishing the principal inequality of the Levinson-Conrey method, it remains to evaluate asymptotically the second power-moment of a relevant Dirichlet polynomial, which is built out of derivatives of the zeta function and its mollifier. This task was carried out differently than by the traditional arguments and in greater generality than it was needed. The main term coming from the contribution of the diagonal terms fits with results in sieve theory and can be useful elsewhere.

I am pleased to express my deep appreciation to Pedro Henrique Pontes, who actively participated in the course and he gave valuable mathematical comments, which improved my presentation. He also helped significantly in editing these notes in addition to typing them. My thanks also go to the Editors of the AMS University Lecture Series for publishing these notes in their volumes, and in particular to Sergei Gelfand for continuous encouragements.

Part 1

Classical Topics

http://dx.doi.org/10.1090/ulect/062/01

CHAPTER 1

Panorama of Arithmetic Functions

Throughout these notes we denote by $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ the sets of integers, rationals, real and complex numbers, respectively. The positive integers are called *natural numbers*. The set

$$\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$$

of natural numbers contains the subset of *prime numbers*

$$\mathcal{P} = \{2, 3, 5, 7, 11, \dots\}.$$

We will often denote a prime number by the letter p.

A function $f : \mathbb{N} \to \mathbb{C}$ is called an *arithmetic function*. Sometimes an arithmetic function is extended to all $\mathbb{Z}$. If f has the property

$$f(mn) = f(m) + f(n) \tag{1.1}$$

for all m, n relatively prime, then f is called an *additive function*. Moreover, if (1.1) holds for all m, n, then f is called *completely additive*; for example, $f(n) = \log n$ is completely additive. If f has the property

$$f(mn) = f(m)f(n) \tag{1.2}$$

for all m, n relatively prime, then f is called a *multiplicative function*. Moreover, if (1.2) holds for all m, n, then f is called *completely multiplicative*; for example, $f(n) = n^{-s}$ for a fixed $s \in \mathbb{C}$, is completely multiplicative.

If $f : \mathbb{N} \to \mathbb{C}$ has at most a polynomial growth, then we can associate with f the *Dirichlet series*

$$D_f(s) = \sum_1^\infty f(n)n^{-s}$$

which converges absolutely for $s = \sigma + it$ with σ sufficiently large. The product of Dirichlet series is a Dirichlet series

$$D_f(s)D_g(s) = D_h(s),$$

with $h = f * g$ defined by

$$h(l) = \sum_{mn=l} f(m)g(n) = \sum_{d|l} f(d)g(l/d), \tag{1.3}$$

which is called the *Dirichlet convolution*.

The constant function $f(n) = 1$ for all $n \in \mathbb{N}$ has the Dirichlet series

$$\zeta(s) = \sum_1^\infty n^{-s} \tag{1.4}$$

which is called the *Riemann zeta-function.* Actually $\zeta(s)$ was first introduced by L. Euler who studied the distribution of prime numbers using the infinite product formula

$$(1.5)\qquad \begin{aligned}\zeta(s) &= \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right)\\ &= \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}.\end{aligned}$$

The series (1.4) and the product (1.5) converge absolutely in the half-plane $s = \sigma + it$, $\sigma > 1$. Since $\zeta(s)$ for $s > 1$ is well approximated by the integral

$$\int_1^\infty y^{-s}\,\mathrm{d}y = \frac{1}{s-1}$$

as $s \to 1$, it follows from (1.5) that

$$(1.6)\qquad \sum_p \frac{1}{p^s} \sim \log \frac{1}{s-1}, \qquad \text{as } s > 1,\ s \to 1.$$

By the Euler product for $\zeta(s)$ it follows that $1/\zeta(s)$ also has a Dirichlet series expansion

$$(1.7)\qquad \frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \sum_1^\infty \frac{\mu(m)}{m^s}$$

where $\mu(m)$ is the multiplicative function defined at prime powers by

$$(1.8)\qquad \mu(p) = -1, \qquad \mu(p^\alpha) = 0, \quad \text{if } \alpha \geqslant 2.$$

This is a fascinating function (introduced by A. F. Möbius in 1832) which plays a fundamental role in the theory of prime numbers.

Translating the obvious formula $\zeta(s) \cdot \zeta(s)^{-1} = 1$ into the language of Dirichlet convolution we obtain the δ-function

$$(1.9)\qquad \delta(n) = \sum_{m|n} \mu(m) = \begin{cases} 1 & \text{if } n = 1\\ 0 & \text{if } n \neq 1.\end{cases}$$

Clearly the two relations

$$(1.10)\qquad g = 1 * f, \qquad f = \mu * g$$

are equivalent. This equivalence is called the *Möbius inversion*; more explicitly,

$$(1.11)\qquad g(n) = \sum_{d|n} f(d) \quad \Longleftrightarrow \quad f(n) = \sum_{d|n} \mu(d) g(n/d).$$

If f, g are multiplicative, then so are $f \cdot g$, $f * g$. If g is multiplicative, then

$$(1.12)\qquad \sum_{d|n} \mu(d) g(d) = \prod_{p|n} \big(1 - g(p)\big).$$

In various contexts one can view the left side of (1.12) as an "exclusion-inclusion" procedure of events occurring at divisors d of n with densities $g(d)$. Then the right side of (1.12) represents the probability that none of the events associated with prime divisors of n takes place.

Now we are going to give some basic examples of arithmetic functions. We begin by exploiting Dirichlet convolutions. The first one is the *divisor function*

$$\tau = 1 * 1, \qquad \tau(n) = \sum_{d|n} 1, \qquad \zeta^2(s) = \sum_1^\infty \tau(n) n^{-s}.$$

This is multiplicative with $\tau(p^\alpha) = \alpha + 1$. We have

$$\frac{\zeta^4(s)}{\zeta(2s)} = \sum_1^\infty \tau(n)^2 n^{-s}$$
$$\frac{\zeta^3(s)}{\zeta(2s)} = \sum_1^\infty \tau(n^2) n^{-s}$$
$$\zeta^3(s) = \sum_1^\infty \Big(\sum_{d^2 m = n} \tau(m^2) \Big) n^{-s}.$$

Note that

$$\sum_{d^2 m = n} \tau(m^2) = \sum_{klm = n} 1 = \tau_3(n), \tag{1.13}$$

say. Next we get

$$\frac{\zeta^2(s)}{\zeta(2s)} = \sum_1^\infty 2^{\omega(n)} n^{-s}$$

where $\omega(n)$ denotes the number of distinct prime divisors of n, so $2^{\omega(n)}$ is the number of squarefree divisors of n. The characteristic function of squarefree numbers is

$$|\mu(n)| = \mu^2(n) = \sum_{d^2|n} \mu(d),$$

its Dirichlet series is

$$\frac{\zeta(s)}{\zeta(2s)} = \sum_1^\infty |\mu(n)| n^{-s} = \prod_p \Big(1 + \frac{1}{p^s}\Big).$$

Inverting this we get the generating series for the *Liouville function* $\lambda(n)$

$$\frac{\zeta(2s)}{\zeta(s)} = \sum_1^\infty \lambda(n) n^{-s}.$$

Note that

$$\lambda(n) = (-1)^{\Omega(n)}$$

where $\Omega(n)$ is the total number of prime divisors of n (counted with the multiplicity).

The *Euler φ-function* is defined by $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^*|$; it is the number of reduced residue classes modulo n. This function satisfies

$$\varphi(n) = n \prod_p \Big(1 - \frac{1}{p}\Big) = n \sum_{d|n} \frac{\mu(d)}{d}.$$

Hence

$$\frac{\zeta(s-1)}{\zeta(s)} = \sum_1^\infty \varphi(n) n^{-s}.$$

A different class of arithmetic functions (very important for the study of prime numbers) is obtained by differentiating relevant Dirichlet series

$$-D_f'(s) = \sum_1^\infty f(n)(\log n)n^{-s}.$$

In particular

$$-\zeta'(s) = \sum_1^\infty (\log n)n^{-s}.$$

Since $L(n) = \log n$ is additive, we have the formula

$$L \cdot (f * g) = (L \cdot f) * g + f * (L \cdot g)$$

which says that the multiplication by L is a derivation in the Dirichlet ring of arithmetic functions.

By the Euler product we have

$$\log \zeta(s) = \sum_{l=1}^\infty \sum_p l^{-1} p^{-ls}. \tag{1.14}$$

Hence differentiating we get

$$-\frac{\zeta'}{\zeta}(s) = \sum_1^\infty \Lambda(n)n^{-s} \tag{1.15}$$

with $\Lambda(n)$ (popularly named *von Mangoldt function*) given by

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^\alpha,\ \alpha \geqslant 1 \\ 0, & \text{otherwise.} \end{cases} \tag{1.16}$$

From the left side of (1.15) we get

$$\Lambda = \mu * L, \qquad \Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d} = -\sum_{d|n} \mu(d) \log d. \tag{1.17}$$

Hence, by Möbius inversion we get

$$L = 1 * \Lambda, \qquad \log n = \sum_{d|n} \Lambda(d). \tag{1.18}$$

Similarly we define the von Mangoldt functions Λ_k of any degree $k \geqslant 0$ by

$$\Lambda_k = \mu * L^k, \qquad \Lambda_k(n) = \sum_{d|n} \mu(d) \left(\log \frac{n}{d}\right)^k. \tag{1.19}$$

We have

$$L^k = 1 * \Lambda_k, \qquad (\log n)^k = \sum_{d|n} \Lambda_k(d). \tag{1.20}$$

Note that $\Lambda_0 = \delta$, $\Lambda_1 = \Lambda$, and we have the recurrence formula

$$\Lambda_{k+1} = L \cdot \Lambda_k + \Lambda * \Lambda_k. \tag{1.21}$$

This follows by writing

$$\Lambda_{k+1}(n) = \sum_{d|n} \mu(d) \left(\log \frac{n}{d}\right)^k (\log n - \log d)$$

and using (1.18).

From (1.21) we derive by induction in k that $\Lambda_k(n) \geqslant 0$ and $\Lambda_k(n)$ is supported on positive integers having at most k distinct prime divisors. Moreover we get by (1.20) that

$$0 \leqslant \Lambda_k(n) \leqslant (\log n)^k. \tag{1.22}$$

EXERCISE. Prove the formula

$$\Lambda_k(mn) = \sum_{0 \leqslant j \leqslant k} \binom{k}{j} \Lambda_j(m)\Lambda_{k-j}(n). \tag{1.23}$$

EXERCISE. Prove the formula

$$\sum_p \frac{1}{p^s} = \sum_1^\infty \frac{\mu(n)}{n} \log \zeta(ns), \qquad \text{if } s > 1. \tag{1.24}$$

http://dx.doi.org/10.1090/ulect/062/02

CHAPTER 2

The Euler–Maclaurin Formula

Some arithmetic functions f are naturally defined on segments of real numbers, say $f : [a,b] \to \mathbb{C}$. If f is continuous and has relatively slow variation, then one should expect that the sum

$$\sum_{a<n\leqslant b} f(n)$$

is well approximated by the corresponding integral. Indeed we have the following exact formula

$$\sum_{a<n\leqslant b} f(n) = \int_a^b \big(f(x) + \psi(x)f'(x)\big)\,\mathrm{d}x + \frac{1}{2}\big(f(b) - f(a)\big) \tag{2.1}$$

provided $a < b$, $a, b \in \mathbb{Z}$, and f is of class $\mathcal{C}^1$ on $[a,b]$. Here $\psi(x)$ is the saw function

$$\psi(x) = x - [x] - \frac{1}{2}. \tag{2.2}$$

This classical formula of Euler-Maclaurin follows easily by partial integration. Estimating $\psi(x)$ trivially we infer the following approximation

$$\left|\sum_{a<n\leqslant b} f(n) - \int_a^b f(x)\,\mathrm{d}x\right| \leqslant \frac{1}{2}\int_a^b |f'(x)|\,\mathrm{d}x + \frac{1}{2}\big|f(b) - f(a)\big|. \tag{2.3}$$

This approximation is particularly useful for functions $f(x)$ with $f'(x)$ relatively small, in which case the error term is of order of magnitude no larger than the terms of the summation.

Exercise. Derive the formulas (for $x \geqslant 2$)

$$\sum_{n\leqslant x} \log n = x\log x - x + O(\log x) \tag{2.4}$$

$$\sum_{n\leqslant x} \log\frac{x}{n} = x + O(\log x) \tag{2.5}$$

$$\sum_{n\leqslant x} \frac{1}{n} = \log x + \gamma + O(x^{-1}). \tag{2.6}$$

Lemma 2.1. *For* $\operatorname{Re} s > 1$ *we have*

$$\zeta^{(l)}(s) = \frac{(-1)^l l!}{(s-1)^{l+1}} + O\left((\log 2|s|)^{l+1}\right) \tag{2.7}$$

where the implied constant depends only on l.

PROOF. By the Euler-Maclaurin formula we derive (with any $X \geqslant 2$)

$$\begin{aligned}(-1)^l \zeta^{(l)}(s) &= \sum_1^\infty \frac{(\log n)^l}{n^s} \\ &= \sum_{n \leqslant X} \frac{(\log n)^l}{n^s} + \int_X^\infty \frac{(\log y)^l}{y^s}\,\mathrm{d}y + O\left(\frac{|s|}{X}(\log X)^{l+1}\right) \\ &= \int_1^\infty \frac{(\log y)^l}{y^s}\,\mathrm{d}y + O\left(\left(1 + \frac{|s|}{X}\right)(\log X)^{l+1}\right).\end{aligned}$$

Choosing $X = 2|s|$ we get (2.7). □

COROLLARY 2.2. *For* $\operatorname{Re} s > 1$ *we have*

$$\big((s-1)\zeta(s)\big)^{(l)} \ll |s|\big(\log 2|s|\big)^{l+1}. \tag{2.8}$$

PROOF. This follows from (2.7) by the formula

$$\big((s-1)\zeta(s)\big)^{(l)} = (s-1)\zeta^{(l)}(s) + l\zeta^{(l-1)}(s).$$ □

Some arithmetic functions have no natural extension to real numbers, for example the divisor function $\tau(n)$. In this case there is no single steady function which well approximates $\tau(n)$. However, since $\tau = 1 * 1$ is the convolution of smooth functions (constant functions), one can still evaluate the sum

$$D(x) = \sum_{n \leqslant x} \tau(n) \tag{2.9}$$

by opening the convolution. We obtain

$$\begin{aligned}D(x) &= \sum_{lm \leqslant x} 1 = \sum_{l \leqslant x} \left[\frac{x}{l}\right] = x \sum_{l \leqslant x} \frac{1}{l} + O(x) \\ &= x \log x + O(x).\end{aligned}$$

Dirichlet had the great idea of improving the above approximation by switching divisors (Dirichlet hyperbola method). We have

$$\begin{aligned}D(x) &= 2 \sum_{lm \leqslant x,\, l \leqslant \sqrt{x}} 1 - \sum_{l, m \leqslant \sqrt{x}} 1 = 2 \sum_{l \leqslant \sqrt{x}} \left[\frac{x}{l}\right] - [\sqrt{x}]^2 \\ &= 2 \sum_{l \leqslant \sqrt{x}} \frac{x}{l} - x + O\left(\sqrt{x}\right) = 2x \log \sqrt{x} + 2\gamma - x + O\left(\sqrt{x}\right).\end{aligned}$$

Hence

$$D(x) = x(\log x + 2\gamma - 1) + \Delta(x) \tag{2.10}$$

with

$$\Delta(x) \ll \sqrt{x}. \tag{2.11}$$

LEMMA 2.3. *Suppose $g(x)$ is a real-valued function on $[a, b]$ with continuous derivatives $g'(x), g''(x)$ such that $g'(x)g''(x) \neq 0$. Then*

$$\left|\int_a^b \mathrm{e}\left(g(x)\right)\mathrm{d}x\right| \leqslant \frac{1}{\pi|g'(a)|} + \frac{1}{\pi|g'(b)|}, \tag{2.12}$$

where $\mathrm{e}(x) = \exp(2\pi i x)$.

PROOF. Since $g''(x) \neq 0$ we can assume that $g''(x) > 0$. This shows that $(1/g'(x))' = -g''(x)(g'(x))^{-2} < 0$. Then by partial integration

$$2\pi i \int_a^b \mathrm{e}\big(g(x)\big)\,\mathrm{d}x = \frac{\mathrm{e}\big(g(b)\big)}{g'(b)} - \frac{\mathrm{e}\big(g(a)\big)}{g'(a)} - \int_a^b \mathrm{e}\big(g(x)\big)\,\mathrm{d}\left(\frac{1}{g'(x)}\right).$$

This yields (2.12) by trivial estimation. □

Using partial summation we derive from (2.12)

COROLLARY 2.4. *Let $h(x)$ be a smooth function on $[a,b]$, and $g(x)$ as in Lemma 2.3. Then*

$$\left|\int_a^b h(x)\,\mathrm{e}\big(g(x)\big)\,\mathrm{d}x\right| \leqslant \frac{H}{\pi}\left(\frac{1}{|g'(a)|} + \frac{1}{|g'(b)|}\right) \tag{2.13}$$

with

$$H = |h(b)| + \int_a^b |h'(y)|\,\mathrm{d}y. \tag{2.14}$$

COROLLARY 2.5. *Suppose $g(x)$ is a real-valued function on $[a,b]$ with continuous derivatives $g'(x), g''(x)$ satisfying $g''(x) \neq 0$ and*

$$|g'(x)| \leqslant \theta, \qquad 0 < \theta < 1. \tag{2.15}$$

Let $h(x)$ be smooth on $[a,b]$. Then for any $l \geqslant 1$ we have

$$\left|\int_a^b h(x)\,\mathrm{e}\big(g(x) - lx\big)\,\mathrm{d}x\right| \leqslant \frac{2H}{\pi(l-\theta)}. \tag{2.16}$$

PROOF. This follows from (2.13) when $g(x)$ is replaced by $g(x) - lx$. □

COROLLARY 2.6. *Suppose $g(x)$ satisfies the conditions of Corollary 2.5. Then*

$$\left|\int_a^b h(x)\psi(x)\,\mathrm{e}\big(g(x)\big)\,\mathrm{d}x\right| \leqslant \frac{H}{1-\theta}. \tag{2.17}$$

PROOF. We obtain this by applying (2.16) to every term in the Fourier expansion

$$\psi(x) = -\sum_{0<|l|\leqslant L} \frac{\mathrm{e}(lx)}{2\pi i l} + O\left(\frac{1}{1+\|x\|L}\right)$$

where $\|x\|$ is the distance of x to the nearest integer. Hence the integral in (2.17) is bounded by

$$2\pi^{-2} H \sum_{l=1}^{\infty} l^{-1}(l-\theta)^{-1} \leqslant H(1-\theta)^{-1}.$$ □

Finally, applying the Euler-Maclaurin formula (2.1) we get

THEOREM 2.7. *Let $g(x), h(x)$ be real-valued smooth functions on $[a,b]$ with $g''(x) \neq 0$ and $g'(x)$ satisfying (2.15). Then*

$$\sum_{a<n\leqslant b} h(n)\,\mathrm{e}\big(g(n)\big) = \int_a^b h(x)\,\mathrm{e}\big(g(x)\big)\,\mathrm{d}x + O\big(H(1-\theta)^{-1}\big) \tag{2.18}$$

where the implied constant is absloute.

http://dx.doi.org/10.1090/ulect/062/03

CHAPTER 3

Tchebyshev's Prime Seeds

Euler, Legendre, and Gauss tried to get hold on the distribution of prime numbers with very little success. The first remarkable results were obtained in the 1850's by Pafnucy Lvovich Tchebyshev. His ideas are elementary and elegant. We begin by writing the sum (2.4) in the following way

$$S(x) = \sum_{n \leqslant x} \log n = \sum_{lm \leqslant x} \Lambda(l) = \sum_{l \leqslant x} \Lambda(l) \left[\frac{x}{l} \right].$$

Replacing $[x/l]$ by $x/l + O(1)$ we get

$$S(x) = x \sum_{l \leqslant x} \frac{\Lambda(l)}{l} + O\big(\psi(x)\big) \tag{3.1}$$

where $\psi(x)$ is defined by (not the saw function (2.2))

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n). \tag{3.2}$$

Changing the order of summation we get another expression

$$S(x) = \sum_{m \leqslant x} \psi\left(\frac{x}{m}\right) = x \log x - x + O(\log x).$$

Hence we compute $S(x) - 2S(\frac{x}{2})$, getting

$$\psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) - \psi\left(\frac{x}{4}\right) + \cdots = x \log 2 + O(\log x).$$

By the monotonicity we infer two estimates:

$$x \log 2 + O(\log x) < \psi(x) < x \log 4 + O\big((\log x)^2\big). \tag{3.3}$$

This shows that

$$\psi(x) \asymp x. \tag{3.4}$$

Inserting (3.4) to (3.1) we derive

$$\sum_{n \leqslant x} \frac{\Lambda(n)}{n} = \log x + O(1). \tag{3.5}$$

Exercise. Derive from (3.5) the following formulas of Mertens

$$\sum_{p \leqslant x} \frac{\log p}{p} = \log x + O(1), \tag{3.6}$$

$$\sum_{p \leqslant x} \frac{1}{p} = \log \log x + c + O\left(\frac{1}{\log x}\right), \tag{3.7}$$

$$\prod_{p \leqslant x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right) \tag{3.8}$$

where c and γ are constants.

Similar (actually simpler) arguments work for the sum of the Möbius function

$$M(x) = \sum_{m \leqslant x} \mu(m). \tag{3.9}$$

For $x \geqslant 1$ we have

$$1 = \sum_{lm \leqslant x} \mu(m) = \sum_{l \leqslant x} M\left(\frac{x}{l}\right) = \sum_{m \leqslant x} \mu(m) \left[\frac{x}{m}\right]. \tag{3.10}$$

Hence

$$\left|\sum_{m \leqslant x} \frac{\mu(m)}{m}\right| \leqslant 1. \tag{3.11}$$

http://dx.doi.org/10.1090/ulect/062/04

CHAPTER 4

Elementary Prime Number Theorem

It was conjectured by Legendre and Gauss that the number of primes $p \leqslant x$ satisfies the asymptotic formula

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

This assertion is called the *Prime Number Theorem*, it has been proved much later independently by Hadamard and de la Vallée Poussin (1896). We shall give an elementary proof of a stronger statement which is inspired by the ideas of Hadamard and de la Vallée Poussin. Our arguments are elementary in the sense that the complex function theory (specially the concept of analytic function and contour integration) is not used. But we do not hesitate to use continuity over the complex numbers, and we employ infinite series to the extent of absolute convergence.

THEOREM 4.1. *For $x \geqslant 2$ and $A \geqslant 0$ we have*

$$\psi(x) = x + O\big(x(\log x)^{-A}\big) \tag{4.1}$$

where the implied constant depends only on A.

We shall derive (4.1) from a similar estimate for the sum of the Möbius function

THEOREM 4.2. *For $x \geqslant 2$ and $A \geqslant 0$ we have*

$$M(x) \ll x(\log x)^{-A} \tag{4.2}$$

where the implied constant depends only on A.

One can show by elementary means that the following two statements are equivalent:

$$\psi(x) \sim x, \quad \text{as } x \to \infty, \tag{4.3}$$

$$M(x) = o(x), \quad \text{as } x \to \infty. \tag{4.4}$$

To this end we develop two approximate formulas. The first one starts by the exact relation

$$\sum_{n \leqslant x} \mu(n) \log \frac{x}{n} = M(x) \log x + \sum_{mn \leqslant x} \mu(m)\Lambda(n),$$

which follows from $(1/\zeta)' = -(\zeta'/\zeta) \cdot (1/\zeta)$. Applying (2.5) we get

$$M(x) \log x = -\sum_{m \leqslant x} \mu(m) \psi\left(\frac{x}{m}\right) + O(x).$$

Next, by (3.4) and (3.11)

$$M(x) \log x = \sum_{m \leqslant x} \mu(m) \left(\frac{x}{m} - \psi\left(\frac{x}{m}\right)\right) + O(x \log 2K)$$

for any $1 \leqslant K \leqslant x$. Hence it is easy to see that (4.3) implies (4.4).

To establish the converse we start from $\Lambda - 1 = \mu * (L - \tau)$. This yields

$$\begin{aligned}\psi(x) - [x] + 2\gamma &= \sum_{dk\leqslant x} \mu(d)\big(\log k - \tau(k) + 2\gamma\big) \\ &= \sum_{k\leqslant K} \big(\log k - \tau(k) + 2\gamma\big) M\left(\frac{x}{k}\right) \\ &\quad + \sum_{d\leqslant x/K} \mu(d)\left(\Delta\left(\frac{x}{d}\right) - \Delta(K)\right) + O\left(\frac{x}{K}\right).\end{aligned}$$

Recall that $\Delta(y)$ is the error term in the divisor problem, see (2.10). Hence for any $1 \leqslant K \leqslant x$ we have

$$\psi(x) - x = \sum_{k\leqslant K} \big(\log k - \tau(k) + 2\gamma\big) M\left(\frac{x}{k}\right) + O\left(\frac{x}{\sqrt{K}}\right). \tag{4.5}$$

This shows that (4.4) implies (4.3), by letting $K \to \infty$ very slowly. Moreover, choosing $K = \sqrt{x}$ in (4.5) we see that Theorem 4.2 implies Theorem 4.1.

Now we proceed to the proof of Theorem 4.2. First we are going to estimate the series

$$G(s) = \sum_{1}^{\infty} \frac{\mu(m)}{m^s}(\log m)^k = (-1)^k \left(\frac{1}{\zeta(s)}\right)^{(k)} \tag{4.6}$$

for $k \geqslant 0$ and $s = \sigma + it$, $\sigma > 1$. Put $\zeta^*(s) = (s-1)\zeta(s)$ and recall that its derivatives were estimated in (2.8). We also need a lower bound for $\zeta^*(s)$. To this end we use the Euler product for $\zeta(s)$ giving

$$\begin{aligned}1 &\leqslant \prod_p \left(1 + \big(1 + p^{it} + p^{-it}\big)^2 p^{-\sigma}\right) \\ &= \prod_p \left(1 + \big(3 + 2p^{it} + 2p^{-it} + p^{2it} + p^{-2it}\big)p^{-\sigma}\right) \\ &\asymp \zeta(\sigma)^3 \big|\zeta(\sigma+it)\big|^4 \big|\zeta(\sigma+2it)\big|^2. \end{aligned} \tag{4.7}$$

If $|s-1|$ is not very small, then $|\zeta(\sigma+2it)| \ll \log 2|s|$ by (2.7) with $l = 0$, so (4.7) yields $|\zeta(s)| \gg (\sigma-1)^{3/4}(\log 2|s|)^{-1/2}$. Hence

$$|\zeta^*(s)| \gg (\sigma-1)^{3/4}|s|\big(\log 2|s|\big)^{-1/2}. \tag{4.8}$$

Clearly (4.8) also holds if $|s-1|$ is small by (2.7) with $l = 0$. We have

$$(-1)^k G(s) = \left(\frac{s-1}{\zeta^*(s)}\right)^{(k)} = (s-1)\left(\frac{1}{\zeta^*(s)}\right)^{(k)} + k\left(\frac{1}{\zeta^*(s)}\right)^{(k-1)}$$

and by the formula from the differential calculus

$$\left(\frac{1}{f}\right)^{(k)} = \frac{k!}{f} \sum_{a_1+2a_2+\cdots=k} \frac{(a_1+a_2+\cdots)}{a_1!a_2!\cdots}\left(\frac{-f'}{1!f}\right)^{a_1}\left(\frac{-f''}{2!f}\right)^{a_2}\cdots$$

with $f = \zeta^*$ we get

$$\left(\frac{1}{\zeta^*(s)}\right)^{(k)} \ll (\sigma-1)^{-\frac{3}{4}(k+1)} \frac{\big(\log 2|s|\big)^{\kappa}}{|s|},$$

with some κ depending on k, specifically $\kappa = \frac{1}{2}(5k+1)$. Hence

$$G(s) \ll (\sigma-1)^{-\frac{3}{4}(k+1)}\big(\log 2|s|\big)^{\kappa}. \tag{4.9}$$

Next from the estimate (4.9) for the infinite series (4.6) we derive an estimate for the finite sum

$$F(x) = \sum_{m\leqslant x} \frac{\mu(m)}{m^\sigma}(\log m)^k, \qquad x>1,\ k\geqslant 1. \tag{4.10}$$

But first we smooth out at the endpoint of summation by means of the function $\Delta(y)$ whose graph is given by Figure 4.1 with $0<\delta\leqslant 1$ to be chosen later.

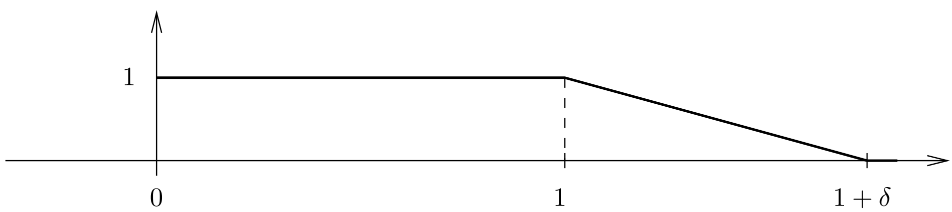

FIGURE 4.1

We have

$$F(x) = \sum_{m} \frac{\mu(m)}{m^\sigma}(\log m)^k \Delta\left(\frac{\log m}{\log x}\right) + O\big(\delta(\log x)^{k+1}\big). \tag{4.11}$$

Let $\hat{\Delta}(u)$ be the Fourier transform of $\Delta(|y|)$, that is;

$$\begin{aligned} \hat{\Delta}(u) &= \int_{-\infty}^{\infty} \Delta(|y|)\,\mathrm{e}(-uy)\,\mathrm{d}y = \frac{\sin \pi u(2+\delta)}{\pi u}\cdot\frac{\sin \pi u\delta}{\pi u\delta} \\ &\ll \min\left(1, \frac{1}{|u|}, \frac{1}{\delta u^2}\right) \ll \frac{1}{1+|u|+\delta u^2}, \end{aligned} \tag{4.12}$$

where $\mathrm{e}(u) = \exp(2\pi i u)$. By Fourier inversion we have

$$\Delta(y) = \int_{-\infty}^{\infty} \hat{\Delta}(u)\,\mathrm{e}(uy)\,\mathrm{d}u. \tag{4.13}$$

Introducing (4.13) to (4.11) we get

$$\begin{aligned} F(x) &= \int_{-\infty}^{\infty} \hat{\Delta}(u) G\left(\sigma + \frac{2\pi i u}{\log x}\right)\mathrm{d}u + O\big(\delta(\log x)^{k+1}\big) \\ &\ll (\sigma-1)^{-\frac{3}{4}(k+1)} \int_0^\infty \frac{\log^\kappa(2+u)}{1+u+\delta u^2}\,\mathrm{d}u + \delta(\log x)^{k+1} \\ &\ll (\sigma-1)^{-\frac{3}{4}(k+1)}\left(\log\frac{1}{\delta}\right)^{\kappa+1} + \delta(\log x)^{k+1}. \end{aligned}$$

Choosing $\delta = (\log x)^{-k-1}$ we get (for all $x>1$)

$$F(x) \ll (\sigma-1)^{-\frac{3}{4}(k+1)}(\log\log 3x)^{\kappa+1}. \tag{4.14}$$

Finally, by partial summation we obtain

$$\begin{aligned} M(x) = \sum_{m\leqslant x}\mu(m) &= \int_1^x \frac{y^\sigma}{(\log y)^k}\,\mathrm{d}F(y) \\ &\ll x^\sigma(\log x)^{-k}(\sigma-1)^{-\frac{3}{4}(k+1)}(\log\log 3x)^{\kappa+1} \\ &\ll x(\log x)^{\frac{3-k}{4}}(\log\log 3x)^{\kappa+1} \end{aligned}$$

by choosing $\sigma = 1 + 1/\log x$. This completes the proof of (4.2).

Exercise. Derive the estimates

$$\sum_{m\leqslant x}\frac{\mu(m)}{m} \ll (\log x)^{-A}, \qquad \sum_{m\leqslant x}\frac{\mu(m)}{m}\log m = -1 + O\big((\log x)^{-A}\big)$$

for all $x \geqslant 2$ and any fixed $A \geqslant 0$, the implied constant depending only on A.

For the lower bound of $\zeta(s)$ we have used the trigonometric inequality

$$(1 + 2\cos x)^2 = 3 + 4\cos x + 2\cos 2x \geqslant 0. \tag{4.15}$$

The important feature of (4.15) is that all the coefficients are non-negative and at least one of them is strictly larger than the coefficient at $\cos 0 = 1$. Specifically in the above case we have $3 < 4$. If we could reduce the ratio $3/4$, then we would only need to use derivatives of $\zeta(s)$ of smaller order.

Now we are going to show weaker results by applying the following inequality

$$\begin{aligned} &(1+\cos 3x)(1+2\cos x)^2 \\ &\qquad = 3 + 5\cos x + 4\cos 2x + 3\cos 3x + 2\cos 4x + \cos 5x \geqslant 0, \end{aligned} \tag{4.16}$$

which has a better ratio $3/5$. This will be sufficient to employ only $\zeta(s)$, $\zeta'(s)$, and $\zeta''(s)$. By the Euler product we get

$$\begin{aligned} 1 &\leqslant \prod_p \left(1 + \frac{1}{2}(p^{3it}+p^{-3it})(1+p^{it}+p^{-it})^2 p^{-\sigma}\right) \\ &\asymp \zeta(\sigma)^3\big|\zeta(\sigma+it)\big|^5\big|\zeta(\sigma+2it)\big|^4\big|\zeta(\sigma+3it)\big|^3\big|\zeta(\sigma+4it)\big|^2\big|\zeta(\sigma+5it)\big|. \end{aligned}$$

If $|s-1|$ is not very small, then $\zeta(\sigma + nit) \ll \log 2|s|$ for $n = 2, 3, 4, 5$, giving

$$|\zeta(s)| \gg (\sigma-1)^{3/5}\big(\log 2|s|\big)^{-2}. \tag{4.17}$$

Hence by (2.7) we get

$$G(s) = \left(\frac{1}{\zeta(s)}\right)'' = 2\frac{\zeta'(s)^2}{\zeta(s)^3} - \frac{\zeta''(s)}{\zeta(s)^2} \ll (\sigma-1)^{-9/5}\big(\log 2|s|\big)^{10}. \tag{4.18}$$

If $|s-1|$ is very small, then by (2.7) we have $\zeta^{(l)}(s) \asymp |s-1|^{-l-1}$ and the formula (4.18) yields $G(s) \ll |s-1|^{-1} \ll (\sigma-1)^{-1}$. Actually one can see that $G(s) \ll 1$ if $|s-1|$ is small. Therefore

$$G(s) = \sum_1^\infty \frac{\mu(m)}{m^s}(\log m)^2 \ll (\sigma-1)^{-9/5}(\log 2|s|)^{10}$$

for all $s = \sigma + it$ with $\sigma > 1$. Hence we derive an estimate for the finite sum

$$F(x) = \sum_{m\leqslant x}\frac{\mu(m)}{m^\sigma}(\log m)^2$$

as follows:

$$\begin{aligned} F(x) &= \int_{-\infty}^{\infty} \hat{\Delta}(u) G\left(\sigma + \frac{2\pi i u}{\log x}\right) \mathrm{d}u + O\big(\delta(\log x)^3\big) \\ &\ll (\sigma-1)^{-9/5} \int_0^{\infty} \frac{\log^{10}(2+u)}{1+u+\delta u^2} \,\mathrm{d}u + \delta(\log x)^3 \\ &\ll (\sigma-1)^{-9/5} \left(\log \frac{1}{\delta}\right)^{11} + \delta(\log x)^3. \end{aligned} \tag{4.19}$$

Choosing $\delta = (\log x)^{-3}$ we find that

$$F(x) \ll (\sigma-1)^{-9/5} (\log\log 3x)^{11}, \qquad x \geqslant 2. \tag{4.20}$$

Note that $F(x) = 0$ if $1 \leqslant x < 2$. By partial summation we get

$$\begin{aligned} M(x) = \sum_{m \leqslant x} \mu(m) &= \int_2^x \frac{y^\sigma}{(\log y)^2} \,\mathrm{d}F(y) \\ &\ll x^\sigma (\log x)^{-2} (\sigma-1)^{-9/5} (\log\log 3x)^{11}. \end{aligned}$$

Choosing $\sigma = 1 + 1/\log x$ we obtain

$$M(x) \ll x(\log x)^{-1/5} (\log\log 3x)^{11}, \qquad x \geqslant 2. \tag{4.21}$$

Finally choosing $K = \log x$ in (4.5) we infer by (4.21) the PNT in the following form

$$\psi(x) = x + O\big(x(\log x)^{-1/5} (\log\log 3x)^{13}\big). \tag{4.22}$$

There are trigonometric polynomials of the above type which have still smaller ratio of the first two coefficients. For example

$$(1 + \cos 4x)(1 + 2\cos x)^4 = 20 + 36\cos x + \ldots \geqslant 0$$

has the ratio $20/36 = 5/9 < 3/5$. For $0 < a < 1$ we have

$$1 + 2\sum_{m=1}^{\infty} a^m \cos(mx) = \frac{1-a^2}{1 - 2a\cos x + a^2} > 0.$$

This has the ratio $1/2a$, which is close to $\frac{1}{2}$ as $a \to 1$.

QUESTION. Does there exist a trigonometric polynomial

$$T(x) = 1 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx \geqslant 0$$

with $a_m \geqslant 0$ for all m and $a_m > 2$ for some m? This does not happen for squares of polynomials. Indeed, if

$$P(x) = \sum_l a_l \cos lx, \qquad a_l = a_{-l},$$

then

$$2P^2(x) = \sum_k \sum_l a_k a_l \big(\cos(k+l)x + \cos(k-l)x\big) = \sum_m A_m \cos mx,$$

with

$$A_m = \sum_{k+l=m} a_k a_l + \sum_{k-l=m} a_k a_l = \sum_k a_k (a_{m-k} + a_{m+k}).$$

Therefore we have

$$|A_m| \leqslant \frac{1}{2} \sum_k (a_k^2 + a_{m-k}^2 + a_k^2 + a_{m+k}^2) = 2 \sum_k a_k^2 = A_0$$

for every m. Hence $|A_m + A_{-m}| \leqslant 2A_0$.

http://dx.doi.org/10.1090/ulect/062/05

CHAPTER 5

The Riemann Memoir

In 1859, B. Riemann [**Rie59**] wrote a short paper (8 pages) called "On the number of primes less than a given magnitude" (in German) in which he expressed fundamental properties of $\zeta(s)$ in the complex variable $s = \sigma + it$. We state these in the modern forms.

A. The function $\zeta(s)$ defined in $s > 1$ by

$$\zeta(s) = \sum n^{-s}$$

has analytic continuation to the whole complex plane, and it is holomorphic except for a simple pole at $s = 1$ with residue 1.

B. The functional equation holds

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

where $\Gamma(s)$ denotes the gamma function of Euler, see Appendix B

C. The zeta function has simple real zeros at $s = -2, -4, -6, \ldots$, which are called *trivial zeros*, and infinitely many non-trivial zeros of the form

$$\rho = \beta + i\gamma, \qquad 0 \leqslant \beta \leqslant 1,\ \gamma \in \mathbb{R}.$$

The number $N(T)$ of non-trivial zeros of height $0 < \gamma < T$ satisfies

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi e} + O(\log T), \qquad T \geqslant 2.$$

D. The product formula holds

$$s(s-1)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = e^{-Bs}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{s/\rho}$$

with

$$B = 1 + \frac{\gamma}{2} - \log 2\sqrt{\pi}.$$

E. The prime numbers formula holds for $x > 1$,

$$\psi^\flat(x) = {\sum_{n\leqslant x}}^{\flat}\Lambda(n) = x - \sum_{\rho}\frac{x^\rho}{\rho} - \log 2\pi - \frac{1}{2}\log\left(1-\frac{1}{x^2}\right).$$

We shall explain the convergence issues in Corollary 10.3.

F. The Riemann Hypothesis. Every non-trivial zero of $\zeta(s)$ is on the line $\operatorname{Re} s = 1/2$, i.e.,

$$\rho = \frac{1}{2} + i\gamma.$$

http://dx.doi.org/10.1090/ulect/062/06

CHAPTER 6

The Analytic Continuation

By collecting the odd and the even terms separately we arrange the following alternating series

$$(1-2^{1-s})\zeta(s) = -\sum_1^\infty (-1)^n n^{-s}.$$

This yields the analytic continuation of $\zeta(s)$ to the half-plane $\sigma > 0$. It also shows that $s = 1$ is a simple pole of $\zeta(s)$ with residue 1, and that $\zeta(s)$ is negative on the segment $0 < \sigma < 1$, $t = 0$.

We have

$$s\int_n^{n+1} x^{-s-1}\,\mathrm{d}x = \frac{1}{n^s} - \frac{1}{(n+1)^s}.$$

Hence

$$s\int_n^{n+1} \frac{[x]}{x^{s+1}}\,\mathrm{d}x = n\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right).$$

Summing over $n = 1, 2, \ldots$ we obtain

$$\begin{aligned}\zeta(s) = s\int_1^\infty \frac{[x]}{x^{s+1}}\,\mathrm{d}x &= \frac{s}{s-1} - \frac{1}{2} - s\int_1^\infty \frac{\psi(x)}{x^{s+1}}\,\mathrm{d}x \\ &= \frac{1}{s-1} + \frac{1}{2} - s(s+1)\int_1^\infty \left(\int_1^x \psi(y)\,\mathrm{d}y\right)\frac{\mathrm{d}x}{x^{s+2}}\end{aligned}$$

where $\psi(x) = x - [x] - \frac{1}{2}$. Since the integral of $\psi(y)$ is bounded, the last integral in x converges absolutely if $\sigma > -1$ giving the analytic continuation of $\zeta(s)$ to the half-plane $\sigma > -1$. Note that

$$\zeta(0) = -\frac{1}{2}. \tag{6.1}$$

Moreover,

$$\begin{aligned}\lim_{s\to 1}\left(\zeta(s) - \frac{1}{s-1}\right) &= \frac{1}{2} - \int_1^\infty \frac{\psi(x)}{x^2}\,\mathrm{d}x = \frac{1}{2} - \lim_{N\to\infty}\int_1^N \frac{x-[x]-\frac{1}{2}}{x^2}\,\mathrm{d}x \\ &= \frac{1}{2} - \lim_{N\to\infty}\left[\log N - \sum_{n=1}^N \frac{1}{n} + \frac{N-1}{N} - \frac{1}{2} + \frac{1}{2N}\right] \\ &= \lim_{N\to\infty}\left(\sum_{n=1}^N \frac{1}{n} - \log N\right) = \gamma,\end{aligned}$$

where $\gamma = 0.577\ldots$ is the *Euler constant*. Hence

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(|s-1|), \quad \text{as } s \to 1. \tag{6.2}$$

If we start summing from $n = N, N+1, \ldots$, we get the formula

$$\zeta(s) = \sum_1^N n^{-s} + \frac{N^{1-s}}{s-1} - \frac{1}{2}N^{-s} - s(s+1)\int_N^\infty \left(\int_N^x \psi(y)\,\mathrm{d}y\right)\frac{\mathrm{d}x}{x^{s+2}}. \tag{6.3}$$

Hence we get the approximation

$$\zeta(s) = \sum_{n \leqslant N} n^{-s} + \frac{N^{1-s}}{s-1} - \frac{1}{2}N^{-s} + O\left(\frac{|s(s+1)|}{\sigma+1}N^{-\sigma-1}\right), \tag{6.4}$$

which is valid for $\sigma > -1$ and $N \geqslant 1$, the implied constant being absolute.

Suppose $s = \sigma + it$ with $\sigma \geqslant 0$ and $|t| \leqslant 2T$. If $N \geqslant T \geqslant 1$, then we can evaluate the partial sum in (6.4) by applying (2.18) with $h(n) = n^{-\sigma}$ and $g(n) = (t/2\pi)\log n$. We get

$$\int_T^N x^{-s}\,\mathrm{d}x + O(T^{-\sigma}) = \frac{N^{1-s} - T^{1-s}}{1-s} + O(T^{-\sigma}).$$

Hence (6.4) yields

PROPOSITION 6.1. *For $s = \sigma + it$ with $\sigma \geqslant 0$, $|t| \leqslant 2T$ and $T \geqslant 1$ we have*

$$\zeta(s) = \sum_{n \leqslant T} n^{-s} + \frac{T^{1-s}}{s-1} + O(T^{-\sigma}), \tag{6.5}$$

where the implied constant is absolute.

http://dx.doi.org/10.1090/ulect/062/07

CHAPTER 7

The Functional Equation

We follow the original ideas of Riemann which make use of the modularity of the *theta series*

$$\theta(x) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 x}, \qquad x > 0. \tag{7.1}$$

To this end we apply the Poisson formula

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$$

with $f(u) = e^{-\pi u^2 x}$. The Fourier transform of this function is

$$\begin{aligned}\hat{f}(v) &= \int_{-\infty}^{\infty} e^{-\pi u^2 x}\, \mathrm{e}(-uv)\, \mathrm{d}u = \int_{-\infty}^{\infty} \exp(-\pi u^2 x - 2\pi i u v)\, \mathrm{d}u \\ &= e^{-\pi v^2/x} \int_{-\infty}^{\infty} e^{-\pi x(u+iv/x)^2}\, \mathrm{d}u \\ &= e^{-\pi v^2/x} \int_{-\infty}^{\infty} e^{-\pi x u^2}\, \mathrm{d}u = x^{-\frac{1}{2}} e^{-\pi v^2/x}.\end{aligned}$$

Hence the theta series satisfies the modular equation

$$\theta(x) = x^{-\frac{1}{2}}\theta(x^{-1}). \tag{7.2}$$

Putting

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

we get

$$\omega(x^{-1}) = x^{\frac{1}{2}}\omega(x) + \frac{1}{2}(x^{\frac{1}{2}} - 1). \tag{7.3}$$

Now we are ready to implement the ideas of Riemann. We start with the equation

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right) n^{-s} = \int_0^{\infty} e^{-\pi n^2 x} x^{\frac{s}{2}} \frac{\mathrm{d}x}{x}$$

for $n = 1, 2, 3, \ldots$. Hence

$$\begin{aligned}\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) &= \int_0^{\infty} \omega(x) x^{\frac{s}{2}} \frac{\mathrm{d}x}{x} = \int_1^{\infty} \omega(x) x^{\frac{s}{2}} \frac{\mathrm{d}x}{x} + \int_0^{1} \omega(x) x^{\frac{s}{2}} \frac{\mathrm{d}x}{x} \\ &= \int_1^{\infty} \omega(x) x^{\frac{s}{2}} \frac{\mathrm{d}x}{x} + \int_1^{\infty} \left(x^{\frac{1}{2}}\omega(x) + \frac{1}{2}(x^{\frac{1}{2}} - 1)\right) x^{-\frac{s}{2}} \frac{\mathrm{d}x}{x}.\end{aligned}$$

In the last integral we changed x to x^{-1} and applied (7.3). This yields

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)=\int_1^\infty \omega(x)(x^{\frac{s}{2}}+x^{\frac{1-s}{2}})\frac{\mathrm{d}x}{x}+\frac{1}{s(s-1)}. \tag{7.4}$$

Since $\omega(x) \ll e^{-\pi x}$ for $x \geqslant 1$, the above integral converges absolutely in the whole complex s-plane. This gives the analytic continuation of $\zeta(s)$ to the whole complex s-plane. In equation (7.4), the poles of $\Gamma(\frac{s}{2})$ at the negative even numbers are cancelled by the zeros of $\zeta(s)$ at $s=-2,-4,-6,\ldots$, the trivial zeros of $\zeta(s)$.

The right-hand side of (7.4) is invariant under the change s to $1-s$, so we get the *functional equation*

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)=\pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s). \tag{7.5}$$

http://dx.doi.org/10.1090/ulect/062/08

CHAPTER 8

The Product Formula over the Zeros

We consider the entire function

$$(8.1)\quad \begin{aligned}\xi(s) &= s(s-1)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)\\ &= 1+s(s-1)\int_1^\infty \omega(x)(x^{\frac{s}{2}}+x^{\frac{1-s}{2}})x^{-1}\,\mathrm{d}x.\end{aligned}$$

Our goal is to prove the formula

$$(8.2)\quad \xi(s) = e^{-Bs}\prod_\rho\left(1-\frac{s}{\rho}\right)e^{s/\rho}$$

where $\rho=\beta+i\gamma$ runs over the zeros of $\zeta(s)$ that are in the strip $0\leqslant\beta\leqslant 1$.

We begin by showing some basic results for entire functions of *finite order*, that is for holomorphic functions $f:\mathbb{C}\to\mathbb{C}$ which satisfy the upper bound

$$(8.3)\quad |f(z)|\leqslant \alpha e^{\beta|z|^\gamma}\qquad \text{for all } z\in\mathbb{C},$$

with some positive constants α,β,γ. The infimum of the exponents γ in (8.3) is called the *order* of f.

For example our function $\xi(s)$ given by the integral (8.1) satisfies

$$\begin{aligned}|\xi(s)| &\ll 1+\big|s(s-1)\big|\int_1^\infty \omega(x)(x^{\frac{\sigma}{2}}+x^{\frac{1-\sigma}{2}})x^{-1}\,\mathrm{d}x\\ &\ll 1+\big|s(s-1)\big|\int_1^\infty e^{-\pi x}(x^{\frac{\sigma}{2}}+x^{\frac{1-\sigma}{2}})x^{-1}\,\mathrm{d}x\\ &\ll 1+\big|s(s-1)\big|\big(1+|\sigma|\big)^{\pi|\sigma|}\\ &\ll \big(1+|s|\big)^{\pi|s|} = \exp\Big(\pi|s|\log\big(1+|s|\big)\Big).\end{aligned}$$

Therefore $\xi(s)$ has order at most 1. However, for $s=\sigma\to\infty$ we have

$$\xi(\sigma)\sim\sigma^2\pi^{-\frac{\sigma}{2}}\Gamma\left(\frac{\sigma}{2}\right)\sim 2\pi^{\frac12}\sigma^{\frac32}\left(\frac{\sigma}{2\pi e}\right)^{\frac{\sigma}{2}}$$

so $\xi(s)$ has order just 1.

Lemma 8.1. *An entire function $f(z)$ of order κ which does not have zeros is of type*

$$(8.4)\quad f(z)=e^{g(z)}$$

where $g(z)$ is a polynomial of degree κ.

Proof. The function $\log f(z)$ exists and is entire, so

$$g(z)=\log f(z)=\sum_{n=0}^\infty a_n z^n$$

where the power series converges absolutely. By (8.3) we get

$$\operatorname{Re} g(z) = \log |f(z)| \leqslant \beta |z|^\gamma + \log \alpha.$$

We use this inequality on the circle $z = r\,\mathrm{e}(\theta)$, $0 \leqslant \theta < 1$. Letting $a_n = b_n + ic_n$ we get

$$\operatorname{Re} g(z) = b_0 + \sum_1^\infty (b_n \cos 2\pi n\theta - c_n \sin 2\pi n\theta) r^n.$$

Next, by the orthogonality of the trigonometric functions we obtain

$$b_n r^n = 2 \int_0^1 \operatorname{Re} g\big(r\,\mathrm{e}(\theta)\big) \cos(2\pi n\theta)\, \mathrm{d}\theta.$$

Hence

$$\begin{aligned}
|b_n| r^n &\leqslant 2 \int_0^1 \Big|\operatorname{Re} g\big(r\,\mathrm{e}(\theta)\big)\Big|\, \mathrm{d}\theta \\
&= 2 \int_0^1 \Big\{\Big|\operatorname{Re} g\big(r\,\mathrm{e}(\theta)\big)\Big| + \operatorname{Re} g\big(r\,\mathrm{e}(\theta)\big)\Big\}\, \mathrm{d}\theta - 2b_0 \\
&= 4 \int_0^1 \max\big\{0, \operatorname{Re} g\big(r\,\mathrm{e}(\theta)\big)\big\}\, \mathrm{d}\theta - 2b_0 \\
&\leqslant 4(\beta r^\gamma + \log \alpha) - 2b_0.
\end{aligned}$$

Since r can be arbitrarily large we get $b_n = 0$ for $n > \gamma$. Similarly, we show that $c_n = 0$ for $n > \gamma$. Therefore $g(z)$ is a polynomial of degree less than or equal to γ. This shows (8.4) and that $\kappa = \deg g$ is the order of f. □

Our next result (Jensen's formula) gives a connection between the growth of a holomorphic function and its zeros in a disc.

LEMMA 8.2. *Let $f(z)$ be holomorphic in $|z| \leqslant R$. If $f(0) \neq 0$ and $f(z) \neq 0$ for $|z| = R$, then*

$$\int_0^1 \log \big|f\big(R\,\mathrm{e}(\theta)\big)\big|\, \mathrm{d}\theta = \log \frac{|f(0)| R^n}{|z_1 \cdots z_n|} = \int_0^R \frac{n(r)}{r}\, \mathrm{d}r + \log |f(0)|$$

where $z_1, \dots, z_n$ are all the zeros of $f(z)$ (counted with multiplicities) and $n(r)$ is the number of zeros in $|z| < r$.

PROOF. Let $F(z)$ be such that

$$f(z) = (z - z_1) \cdots (z - z_n) F(z),$$

so $F(z)$ is holomorphic and it does not vanish in $|z| \leqslant R$. It suffices to prove the formula for every factor separately. For $F(z)$ we use $\log F(z)$ and get by Cauchy's theorem

$$\log F(0) = \frac{1}{2\pi i} \int_{|z|=R} \log F(z) \frac{\mathrm{d}z}{z} = \int_0^1 \log F\big(R\,\mathrm{e}(\theta)\big)\, \mathrm{d}\theta.$$

Taking the real parts we conclude the formula

$$\int_0^1 \log \big|F\big(R\,\mathrm{e}(\theta)\big)\big|\, \mathrm{d}\theta = \log |F(0)| = \log \left|\frac{f(0)}{z_1 \cdots z_n}\right|.$$

For $z - z_j$ we have

$$\int_0^1 \log\big|R\,\mathrm{e}(\theta) - z_j\big|\,\mathrm{d}\theta = \frac{1}{2\pi i}\int_{|z|=R} \log|z - z_j|\frac{\mathrm{d}z}{z}$$
$$= \frac{1}{2\pi i}\int_{|z|=R} \Big\{\log|z| + \log\Big|1 - \frac{z_j}{z}\Big|\Big\}\,\frac{\mathrm{d}z}{z} = \log R$$

because the integral of the second part vanishes by Cauchy's theorem since it is the real part of the contour integral of $z^{-1}\log(1 - z_j/z)$. □

Corollary 8.3. *If $f(z)$ is entire and satisfies* (8.3), *then the number of zeros in the circle $|z| \leqslant R$ satisfies*

$$n(R) \ll R^\gamma, \qquad R \geqslant 1. \tag{8.5}$$

Proof. We apply Jensen's formula in the circle $|z| \leqslant eR$ for the function $z^{-m}f(z)$, where m is the order of the zero of $f(z)$ at $z = 0$. We get

$$n(R) \leqslant m + \int_0^{eR} \frac{n(r)}{r}\,\mathrm{d}r \ll R^\gamma$$ □

Corollary 8.4. *If $f(z)$ is entire and satisfies* (8.3), *then*

$$\sum_{\rho\neq 0} |\rho|^{-\gamma-\varepsilon} < \infty, \tag{8.6}$$

where ρ runs over the zeros of $f(z)$.

From now on we restrict our analysis to entire functions $f(z)$ with order at most one. Since the series

$$\sum_{\rho\neq 0} |\rho|^{-1-\varepsilon} \tag{8.7}$$

converges, we can define the product

$$P(z) = \prod_{\rho\neq 0}\Big(1 - \frac{z}{\rho}\Big)\,e^{z/\rho}. \tag{8.8}$$

Moreover, there are arbitrarily large numbers R such that

$$\big|R - |\rho|\big| > |\rho|^{-2} \qquad \text{for every } \rho \neq 0. \tag{8.9}$$

Hence, by the convergence of the series (8.7), one can show the following lower bound for the product (8.8):

$$|P(z)| \gg \exp(-R^{1+\varepsilon}), \qquad \text{on } |z| = R.$$

Dividing $f(z)$ by the product in (8.8) we get an entire function $F(z) = f(z)/P(z)$ which has no zeros and it satisfies

$$|F(z)| \ll \exp(R^{1+\varepsilon}), \qquad \text{on } |z| = R.$$

Since R can be arbitrarily large, by the maximum principle this estimate implies that $F(z)$ is of order at most 1. So by Lemma 8.1 we conclude

Theorem 8.5. *An entire function $f(z)$ of order at most 1 with $f(0) \neq 0$ has the following product representation:*

$$f(z) = e^{A+Bz}\prod_{\rho}\Big(1 - \frac{z}{\rho}\Big)\,e^{z/\rho}. \tag{8.10}$$

Using the inequality $|(1-\eta)e^{\eta}| \leqslant e^{2|\eta|}$ for $\eta \in \mathbb{C}$ we derive from (8.10)

COROLLARY 8.6. *If the series* $\sum |\rho|^{-1}$ *converges, then*

$$|f(z)| \ll e^{c|z|} \tag{8.11}$$

for all $z \in \mathbb{C}$, *where* c *is a positive constant.*

This result has a more surprising formulation; if an entire function $f(z)$ has order 1, but fails to satisfy (8.11), then

$$\sum_{\rho} |\rho|^{-1} = \infty. \tag{8.12}$$

In particular $f(z)$ must have infinitely many zeros.

As an example we have shown that $\xi(s)$ is entire of order 1. Therefore Theorem 8.5 yields the product (8.2) (we have $\xi(0) = 1$ by (8.1), so $A = 0$ in (8.10)). By the above remarks we also learn that $\zeta(s)$ has infinitely many zeros $\rho = \beta + i\gamma$ in the *critical strip* $0 \leqslant \operatorname{Re} s \leqslant 1$.

Actually we can already see that $\zeta(s)$ does not vanish on the line $\operatorname{Re} s = 1$. Indeed, if $\rho = 1 + i\gamma$ was a zero of $\zeta(s)$, then for $s = \sigma + i\gamma$ we would have the upper bound $\zeta(s) \ll \sigma - 1$ as $\sigma \to 1+$, which contradicts (4.17). By the functional equation (7.5) it follows that $\zeta(s)$ does not vanish on the line $\operatorname{Re} s = 0$.

Now, by taking the logarithmic derivative of (8.2) we get

$$\frac{\xi'}{\xi}(s) = B + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right). \tag{8.13}$$

Hence

$$\frac{\zeta'}{\zeta}(s) = -B - \frac{1}{s-1} + \frac{1}{2}\log \pi - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}+1\right) + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right). \tag{8.14}$$

For $s = \sigma + it$ with $-1 \leqslant \sigma \leqslant 2$, $t \geqslant 2$, this together with Stirling's bound, yields

$$\frac{\zeta'}{\zeta}(s) = \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right) + O(\log t). \tag{8.15}$$

In particular for $s = 2 + iT$, $T \geqslant 2$, we get

$$\sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right) \ll \log T. \tag{8.16}$$

The real parts of the above fractions are

$$\operatorname{Re} \frac{1}{s-\rho} = \frac{2-\beta}{(2-\beta)^2 + (T-\gamma)^2} \geqslant \frac{1}{4 + (T-\gamma)^2},$$
$$\operatorname{Re} \frac{1}{\rho} = \frac{\beta}{\beta^2 + \gamma^2} \geqslant 0.$$

Hence, discarding some positive terms in (8.16) we get

$$\sum_{\rho} \frac{1}{4 + (T-\gamma)^2} \ll \log T. \tag{8.17}$$

Ignoring more terms in (8.17) we conclude

COROLLARY 8.7. *The number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ in the box $0 < \beta < 1$, $T < \gamma \leqslant T+1$, is*

$$N(T+1) - N(T) \ll \log T. \tag{8.18}$$

Now, evaluating (8.15) at $s = 2+iT$, subtracting, and applying (8.18) we derive

THEOREM 8.8. *For s in the strip $-1 \leqslant \operatorname{Re} s \leqslant 2$ we have*

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{|s-\rho|\leqslant 1} \frac{1}{s-\rho} + O\big(\log(2+|s|)\big). \tag{8.19}$$

Similar arguments yield more general results of type (8.19). In Chapter 19 we shall need the following

THEOREM 8.9. *Let $f(z)$ be an entire function such that*

$$|f(0)| \geqslant 1 \tag{8.20}$$

$$|f(z)| \leqslant e^b(|z|+3)^{c|z|} \tag{8.21}$$

for all $z \in \mathbb{C}$, where b, c are positive constants.
Then for $|z| \leqslant R$ we have

$$\left|\frac{f'}{f}(z) - \frac{f'}{f}(0) - \sum_{|\rho|<2R}\left(\frac{1}{z-\rho} + \frac{1}{\rho}\right)\right| \leqslant bR^{-1} + 64c\,\log\,(R+3). \tag{8.22}$$

PROOF. By Jensen's formula (see Lemma 8.2) the number $n(R)$ of zeros ρ of $f(z)$ in the circle $|z| \leqslant R$ is bounded by

$$n(R) \leqslant \int_0^{3R} \frac{n(r)}{r}\,\mathrm{d}r \leqslant b + 3cR\,\log\,(3R+3).$$

Then, by (8.10) the left-hand side of (8.22) is

$$\left|\sum_{|\rho|\geqslant 2R}\left(\frac{1}{z-\rho} + \frac{1}{\rho}\right)\right| \leqslant \sum_{|\rho|\geqslant 2R} \frac{R}{(|\rho|-R)|\rho|} \leqslant bR^{-1} + 64c\,\log\,(R+3). \qquad \square$$

We can apply Theorem 8.9 for Dirichlet polynomials.

COROLLARY 8.10. *Let $M(s)$ be given by*

$$M(s) = \sum_{1\leqslant m\leqslant M} b_m m^{-s} \tag{8.23}$$

with $b_1 = 1$ and $|b_m| \leqslant 1$ if $2 \leqslant m \leqslant M$. Then for $\operatorname{Re} s \geqslant 0$ we have

$$\frac{M'}{M}(s) = \sum_{|s-\rho|<1} \frac{1}{s-\rho} + O(\log M), \tag{8.24}$$

where ρ runs over the zeros of $M(s)$ and the implied constant is absolute. Moreover, the number of zeros ρ with $|s-\rho| < 1$ is $O(\log M)$.

PROOF. First note that $M(s)$ does not vanish in $\operatorname{Re} s \geqslant 2$; in fact $|M(s)| \geqslant 2 - \zeta(2)$. We take $f(z) = 2M(z+s+3)$, so for any $z \in \mathbb{C}$ we have

$$|f(z)| \leqslant 2\sum_{m\leqslant M} m^{-\operatorname{Re}(z+s+3)} \leqslant 2\sum_{m\leqslant M} m^{-\operatorname{Re}(z)} \leqslant 2\sum_{m\leqslant M} m^{|z|} \leqslant 2M^{1+|z|}.$$

Moreover, we have $|f(0)| = 2|M(s+3)| > 2(2-\zeta(3)) > 1$ and $|f'(0)| < 2|\zeta'(3)|$. By (8.22) we get

$$\frac{f'}{f}(z) = \sum_{|\rho-s-3|<6} \left(\frac{1}{z-(\rho-s-3)} + \frac{1}{\rho-s-3} \right) + O(\log M)$$

for any $|z| \leqslant 3$. Taking $z = -3$ this gives

$$\frac{M'}{M}(s) = \sum_{|\rho-s-3|<6} \left(\frac{1}{s-\rho} + \frac{1}{\rho-s-3} \right) + O(\log M).$$

Now, the number of zeros ρ of $M(s)$ with $|\rho - s - 3| < 6$ is equal to the number of zeros of $f(z)$ in the circle $|z| < 6$. Since this number is $O(\log M)$ by Jensen's fomula, the result follows. $\square$

Remark. Modifying the above lines slightly, one can extend the result for Dirichlet polynomials (8.23) with $b_1 = 1$ and $|b_m| \leqslant m$.

http://dx.doi.org/10.1090/ulect/062/09

CHAPTER 9

The Asymptotic Formula for $N(T)$

Recall that $N(T)$ denotes the number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ in the rectangle

$$0 < \beta < 1, \quad 0 < \gamma \leqslant T. \tag{9.1}$$

Because of symmetry, the number of zeros with $|\gamma| < T$ is just $2N(T)$. Our goal is to prove the following approximation

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi e} + O(\log T), \qquad T \geqslant 2. \tag{9.2}$$

Since $\zeta(s)$ and $\xi(s)$ (see (8.1)) have the same zeros (counted with multiplicity), we shall work with $\xi(s)$ rather than with $\zeta(s)$. By the principle of argument variation we obtain

$$2\pi N(T) = \Delta_{\mathcal{L}} \arg \xi(s)$$

where the curve $\mathcal{L}$ consists of four segments with endpoints 2, $2 + iT$, $-1 + iT$, -1, as in Figure 9.1. We assume that $\mathcal{L}$ is positively oriented, which means that as one starts from the point 2, one goes upwards to the point $2 + iT$ and continues along the curve, the rectangle (9.1) is on the left-hand side. Note that $\mathcal{L}$ surrounds the boundary of (9.1), however there are no additional zeros of $\xi(s)$ in the excess area. We also assumed that $\zeta(s)$ has no zeros on the horizontal segments.

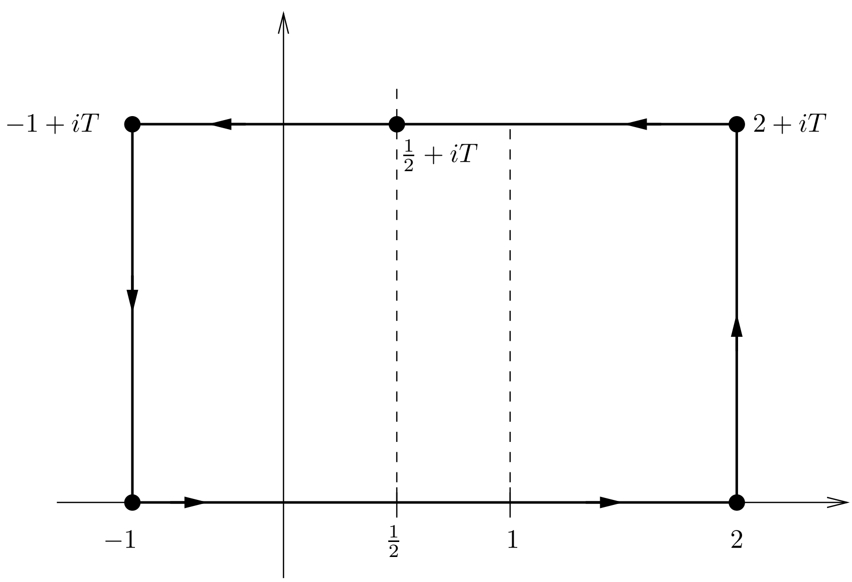

FIGURE 9.1

By the functional equation $\xi(s) = \xi(1-s)$ it follows that $\Delta_{\mathcal{L}} \arg \xi(s)$ is equal to $2\Delta_{\mathcal{M}} \arg \xi(s)$, where $\mathcal{M}$ is the part of $\mathcal{L}$ to the right of the critical line $\operatorname{Re} s = 1/2$. Now we compute the variation of the argument of

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

for every factor separately along every segment. There is no variation along the segment $1/2 \leqslant s \leqslant 2$, because every factor is real. We have

$$\Delta_{\mathcal{M}} \arg s(s-1) = \arg\left(\frac{1}{2}+iT\right)\left(-\frac{1}{2}+iT\right) = \arg -\left(\frac{1}{4}+T^2\right) = \pi,$$

$$\Delta_{\mathcal{M}} \arg \pi^{-\frac{s}{2}} = \arg \pi^{-\frac{1}{4}-\frac{iT}{2}} = -\frac{T}{2}\log \pi,$$

and by Stirling's formula

$$\begin{aligned}\Delta_{\mathcal{M}}\Gamma\left(\frac{s}{2}\right) &= \operatorname{Im} \log \Gamma\left(\frac{1}{2}\left(\frac{1}{2}+iT\right)\right)\\ &= \operatorname{Im}\left\{\left(-\frac{1}{4}+\frac{iT}{2}\right)\log\left(\frac{1}{4}+\frac{iT}{2}\right) - \left(\frac{1}{4}+\frac{iT}{2}\right)\right\} + O\left(\frac{1}{T}\right)\\ &= -\frac{\pi}{8}+\frac{T}{2}\log\frac{T}{2}-\frac{T}{2}+O\left(\frac{1}{T}\right).\end{aligned}$$

Adding up the above results we arrive at the formula

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O\left(\frac{1}{T}\right) \tag{9.3}$$

where

$$S(T) = \frac{1}{\pi}\Delta_{\mathcal{M}} \arg \zeta(s).$$

Since $\zeta(s)$ does not vanish in the half-plane $\sigma > 1$ we have

$$S(T) = \frac{1}{\pi}\arg\zeta\left(\frac{1}{2}+iT\right) \tag{9.4}$$

where the argument is defined as the continuous variation along the horizontal segment from $\infty + iT$ to $\frac{1}{2}+iT$ starting at $\infty + iT$ with the value 0.

It remains to estimate $S(T)$ which is one of the central problems in the theory of the zeta function. The formula (9.2) will follow from (9.3) if we show that

$$S(T) \ll \log T. \tag{9.5}$$

To this end we use the integral

$$\pi S(T) = \int_{\infty+iT}^{\frac{1}{2}+iT} \operatorname{Im}\frac{\zeta'}{\zeta}(s)\,\mathrm{d}s. \tag{9.6}$$

For $s = \sigma + iT$ with $2 \leqslant \sigma < \infty$ we have

$$\left|\frac{\zeta'}{\zeta}(s)\right| \leqslant \sum_n \Lambda(n)n^{-\sigma}$$

so the integral (9.6) along the segment from $\infty + iT$ to $2 + iT$ is bounded by the constant

$$\int_2^\infty\left(\sum_n \Lambda(n)n^{-\sigma}\right)\mathrm{d}\sigma = \sum_n \frac{\Lambda(n)}{n^2\log n}.$$

On the segment $\mathcal{N}$ from $2 + iT$ to $\frac{1}{2} + iT$ we use the expansion (8.19). Since

$$\left| \int_{2+iT}^{\frac{1}{2}+iT} \mathrm{Im}\left(\frac{1}{s-\rho}\right) \mathrm{d}s \right| = \left|\Delta_{\mathcal{N}} \arg(s-\rho)\right| \leqslant \pi$$

and the number of zeros ρ with $|s-\rho| \leqslant 1$ is $O(\log T)$ we conclude the bound (9.5).

http://dx.doi.org/10.1090/ulect/062/10

CHAPTER 10

The Asymptotic Formula for $\psi(x)$

For $x > 0$ we consider the sum

$$\psi^\flat(x) = \sum_{n \leqslant x}{}^{\flat}\, \Lambda(n) \tag{10.1}$$

where the superscript $\flat$ indicates that the last term for $n = x$ is taken with half of the value $\Lambda(x)$. Note that this restriction makes a difference to $\psi(x)$ only when x is a power of a prime number. Our goal is to expand $\psi^\flat(x)$ into a series over the zeros of $\zeta(s)$. Specifically we shall prove the following

Theorem 10.1. *For $x \geqslant 2$ and $T \geqslant 2$ we have*

$$\psi^\flat(x) = x - \sum_{|\gamma|<T} \frac{x^\rho}{\rho} - \log 2\pi - \frac{1}{2}\log\left(1 - \frac{1}{x^2}\right) + R(x,T) \tag{10.2}$$

with

$$R(x,T) \ll \frac{x}{T}(\log xT)^2 + \min\left(\log x,\ \frac{x \log x}{\{x\}T}\right) \tag{10.3}$$

where $\{x\}$ denotes the distance of x to the nearest prime power other than x, and the implied constant is absolute.

We begin by proving a technical lemma which is known as *Perron's formula.* For $y > 0$ put

$$\delta(y) = \begin{cases} 0 & \text{if } 0 < y < 1 \\ \frac{1}{2} & \text{if } y = \frac{1}{2} \\ 1 & \text{if } y > 1. \end{cases} \tag{10.4}$$

and for $T \geqslant \alpha > 0$ put

$$\delta_\alpha(y,T) = \frac{1}{2\pi i}\int_{\alpha - iT}^{\alpha + iT} y^s \frac{\mathrm{d}s}{s}. \tag{10.5}$$

Lemma 10.2. *If $y \neq 1$ then*

$$\big|\delta(y) - \delta_\alpha(y,T)\big| \leqslant y^\alpha \min\left(1,\ \frac{1}{T|\log y|}\right) \tag{10.6}$$

and for $y = 1$

$$\big|\delta(1) - \delta_\alpha(1,T)\big| \leqslant \frac{\alpha}{T}. \tag{10.7}$$

Proof. If $0 < y < 1$ we move the integration in (10.5) to the horizontal segments $s = \sigma \pm iT$ with $\alpha < \sigma < \infty$ getting

$$\big|\delta_\alpha(y,T)\big| \leqslant \frac{1}{\pi T}\int_\alpha^\infty y^\sigma\, \mathrm{d}\sigma = \frac{y^\alpha}{\pi T|\log y|}.$$

If we move the integration to the right arc of the circle $|s| = |\alpha + iT|$ then we get

$$\left|\delta_\alpha(y, T)\right| \leqslant \frac{1}{2} y^\alpha.$$

Combining both estimates we get (10.6).

The case $y > 1$ is similar but we move the integration to the left side of the segment $s = \alpha + it$, $-T < t < T$. The pole at $s = 0$ contributes the residue 1, showing (10.6).

Finally for $y = 1$ we compute as follows

$$\begin{aligned}\delta_\alpha(1, T) &= \frac{1}{2\pi i}\big(\log(\alpha + iT) - \log(\alpha - iT)\big)\\ &= \frac{1}{2\pi}\big(\arg(\alpha + iT) - \arg(\alpha - iT)\big)\\ &= \frac{1}{\pi}\arg(\alpha + iT) = \frac{1}{2} + \frac{1}{\pi}\arg(T - i\alpha).\end{aligned}$$

This gives (10.7). □

Now we proceed to estimate $\psi^\flat(x)$. We start by

$$\frac{-\zeta'}{\zeta}(s) = \sum_n \Lambda(n) n^{-s}, \qquad \text{if } \operatorname{Re}(s) > 1.$$

Let $1 < \alpha \leqslant 2$, $x \geqslant 2$, $T \geqslant 2$. Putting

$$\psi(x, T) = \frac{1}{2\pi i}\int_{\alpha - iT}^{\alpha + iT} \frac{-\zeta'}{\zeta}(s)\frac{x^s}{s}\,\mathrm{d}s \tag{10.8}$$

we derive by Lemma 10.2

$$\left|\psi^\flat(x) - \psi(x, T)\right| \leqslant \sum_{n \leqslant x} \Lambda(n)\left(\frac{x}{n}\right)^\alpha \min\left(1, \frac{1}{T|\log(x/n)|}\right) + \frac{\alpha}{T}\Lambda(x)$$

where $\Lambda(x) = 0$ unless x is a prime power. Let $q \neq x$ be the prime power which is the nearest to x. This single term contributes

$$\min\left(\log x, \frac{1}{T}\frac{\log x}{\log x/q}\right) \ll \min\left(\log x, \frac{x\log x}{\{x\}T}\right).$$

For $n \neq x$, $n \neq q$, we have $|n - x| \geqslant 1$, so the sum is estimated by the integral

$$\int_{\substack{u \geqslant 1\\ |u - x| \geqslant 1}} (\log u)\left(\frac{x}{u}\right)^\alpha \min\left(1, \frac{1}{T|\log(x/u)|}\right)\mathrm{d}u \ll \frac{x}{T}(\log x)^2$$

by choosing $\alpha = 1 + 1/\log x$. Adding these estimates we get

$$\psi^\flat(x) - \psi(x, T) \ll \frac{x}{T}(\log x)^2 + \min\left(\log x, \frac{x\log x}{\{x\}T}\right). \tag{10.9}$$

Next we need to evaluate $\psi(x, T)$. To this end we move the integration in (10.8) to the left arbitrarily far away, getting two integrals over the horizontal segments $s = \sigma + it$ with $-\infty < \sigma < \alpha$, $t = \pm T$, and the contribution of the simple poles of $-\frac{\zeta'}{\zeta}(s)x^s/s$ at $s = 1$, $s = \rho = \beta + i\gamma$ with $|\gamma| < T$, $s = 0$, and $s = -2, -4, -6, \ldots$ The polar contribution is equal to

$$x - \sum_{|\gamma| < T}\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \sum_{m=1}^{\infty}\frac{x^{-2m}}{2m} = x - \sum_{|\gamma| < T}\frac{x^\rho}{\rho} - \log 2\pi - \frac{1}{2}\log\left(1 - \frac{1}{x^2}\right).$$

For the estimation of the horizontal integral we first assume that T is chosen so that no zero ρ of $\zeta(s)$ has height near T. Specifically, changing T by adding an absolutely bounded positive number, we may require

$$|T-\gamma| \gg (\log T)^{-1}$$

because $N(T+1)-N(T) \ll \log T$. For s on such segment we get

$$\frac{\zeta'}{\zeta}(s) \ll \big(\log|s|\big)^2 \tag{10.10}$$

by (8.19). Well, (8.19) yields the above only for $s=\sigma \pm iT$ with $-1 \leqslant \sigma \leqslant 2$, however, an extension of (10.10) to all σ can be easily deduced by the functional equation. Using (10.10) we find that the horizontal integrals are bounded by

$$T^{-1}\int_{-\infty}^{\alpha} x^{\sigma}\log^2\big(|\sigma|+T\big)\,\mathrm{d}\sigma \ll \frac{x}{T}\frac{(\log T)^2}{\log x}.$$

Adding the above estimates we arrive at (10.2), but only for T specially chosen. However, if T is changed by adding an absolutely bounded number, then the approximation (10.2) remains valid. This completes the proof of (10.2) for all $T \geqslant 2$.

Corollary 10.3. *For $x \geqslant 2$ we have*

$$\psi^{\flat}(x) = x - \sum_{\rho}\frac{x^{\rho}}{\rho} - \log 2\pi - \frac{1}{2}\log\left(1-\frac{1}{x^2}\right) \tag{10.11}$$

where the series over the nontrivial zeros is evaluated as the limit

$$\lim_{T\to\infty}\sum_{|\gamma|<T}\frac{x^{\rho}}{\rho}. \tag{10.12}$$

Remark. Our estimation (10.2) shows that the limit (10.12) does exist for every $x \geqslant 2$, but of course it is not a continuous function of x; a jump by $\frac{1}{2}\log p$ occurs at $x=p^m$. The formula (10.11) also holds true for $1<x<2$, but the above arguments require more attention.

By the Riemann hypothesis we get from (10.2) the estimate

$$\psi(x) = x + (x^{\frac{1}{2}}\log^2 x). \tag{10.13}$$

Conversely, if

$$\psi(x) = x + O(x^{\frac{1}{2}+\varepsilon}), \tag{10.14}$$

then the Riemann hypothesis is true. In fact, for any $\alpha \geqslant \frac{1}{2}$ the estimate

$$\psi(x) = x + O(x^{\alpha+\varepsilon}) \tag{10.15}$$

is equivalent to the non-vanishing of $\zeta(s)$ in $\operatorname{Re} s > \alpha$. This can be easily seen by using (10.2) in one direction, and in the other direction from the integral expressions

$$\begin{aligned}\frac{-\zeta'}{\zeta}(s) &= s\int_1^{\infty}\psi(x)x^{-1-s}\,\mathrm{d}x\\ &= \frac{s}{s-1} + s\int_1^{\infty}\big(\psi(x)-x\big)x^{-1-s}\,\mathrm{d}x.\end{aligned}$$

If (10.15) holds, then the last integral converges absolutely in $\operatorname{Re} s > \alpha$, so $\zeta(s)$ has no zeros in this half-plane.

http://dx.doi.org/10.1090/ulect/062/11

CHAPTER 11

The Zero-free Region and the PNT

It is clear from our previous considerations that the Prime Number Theorem

$$\psi(x) \sim x \qquad \text{as } x \to \infty$$

is equivalent to

$$\zeta(1+it) \neq 0 \qquad \text{for } t \in \mathbb{R}.$$

The latter was established in 1896 by Hadamard and de la Vallée Poussin. In fact it was shown (by de la Vallée Poussin) that

$$\zeta(s) \neq 0 \qquad \text{for } \sigma > 1 - c/\log(2+|t|), \tag{11.1}$$

where c is an absolute positive constant. Here we present an elegant proof (given by Mertens) of this fact, which uses our familiar inequality

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

For $\sigma > 1$ we have

$$\frac{-\zeta'}{\zeta}(s) = \sum \Lambda(n) n^{-s}$$

so by the above inequality we get (because $\Lambda(n) \geqslant 0$)

$$3\frac{-\zeta'}{\zeta}(\sigma) + 4\frac{-\zeta'}{\zeta}(\sigma+it) + 2\frac{-\zeta'}{\zeta}(\sigma+2it) \geqslant 0. \tag{11.2}$$

From the pole of $\zeta(s)$ at $s=1$ we get

$$\frac{-\zeta'}{\zeta}(\sigma) = \frac{1}{\sigma-1} + O(1).$$

By (8.19) and the positivity of $\operatorname{Re}(1/(s-\rho))$ we get

$$\operatorname{Re}\frac{-\zeta'}{\zeta}(\sigma+it) < \operatorname{Re}\frac{1}{\sigma+it-\rho} + O\big(\log(2+|t|)\big)$$
$$\operatorname{Re}\frac{-\zeta'}{\zeta}(\sigma+2it) \ll \log\big(2+|t|\big).$$

In the first inequality we kept only one term with a selected zero $\rho = \beta + i\gamma$, whereas in the second inequality we dropped every term. Note that

$$\operatorname{Re}\frac{1}{\sigma+it-\rho} = \frac{\sigma-\beta}{(\sigma-\beta)^2+(t-\gamma)^2} = \frac{1}{\sigma-\beta}$$

for $s = \sigma + i\gamma$. Inserting these results to (11.2) we get

$$\frac{3}{\sigma-1} > \frac{4}{\sigma-\beta} + O\big(\log(2+|\gamma|)\big).$$

Since $4 > 3$, this shows that β cannot be close to 1 if $\sigma > 1$ is close to one. Specifically, choosing $\sigma = 1 - c_1/\log(2+|\gamma|)$ with c_1 sufficiently small to absorb the term $O(\log(2+|\gamma|))$, we get

$$\beta < 1 - c/\log(2+|\gamma|)$$

where c is a small, positive constant. This proves (11.1).

Now we can derive in many ways the following

THEOREM 11.1. *For $x \geqslant 2$ we have*

$$\psi(x) = x + O\big(x\exp(-c\sqrt{\log x})\big) \tag{11.3}$$

where c is an absolute, positive constant.

PROOF. One can get (11.3) quickly by the formula (10.2) with $2 \leqslant T \leqslant x$ giving

$$\psi(x) = x - \sum_{|\gamma|<T} \frac{x^\rho}{\rho} + O\left(\frac{x}{T}(\log x)^2\right).$$

By (11.1) we have $|x^\rho| = |x^\beta| \leqslant x^{1-c/\log T}$ and by (8.18)

$$\sum_{|\gamma|<T} \frac{1}{|\rho|} \ll (\log T)^2.$$

Hence

$$|\psi(x) - x| \ll x(x^{-c/\log T} + T^{-1})(\log x)^2.$$

Choosing $T = \exp\sqrt{c\log x}$ we get (11.3) with a different constant c. □

REMARKS. The best known zero-free region and corresponding error term for the PNT are due to Korobov and Vinogradov (1957):

$$\zeta(s) \neq 0 \qquad \text{if } \sigma > 1 - c(\log t)^{-2/3},\ t \geqslant 2; \tag{11.4}$$

$$\psi(x) - x \ll \exp\big(-c(\log x)^{3/5}(\log\log x)^{-1/5}\big), \qquad x \geqslant 3. \tag{11.5}$$

EXERCISE. Using (11.1) or (11.3) prove the following estimates (for $x \geqslant 1$)

$$\sum_{m\leqslant x} \mu(m) \ll x\exp\big(-c\sqrt{\log x}\big) \tag{11.6}$$

$$\sum_{m\leqslant x} \frac{\mu(m)}{m} \ll \exp\big(-c\sqrt{\log x}\big), \tag{11.7}$$

where c is an absolute, positive constant.

http://dx.doi.org/10.1090/ulect/062/12

CHAPTER 12

Approximate Functional Equations

Given two Dirichlet series

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad \text{and} \qquad g(s) = \sum_{n=1}^{\infty} b_n n^{-s},$$

assume that they are connected by a functional equation of the type

$$f(s) = \varepsilon(s) g(1-s),$$

with $\varepsilon(s)$ being the product of a constant, a power, and some gamma functions. Then one can derive useful approximations to $f(s)$ and $g(s)$ by their partial sums of considerably short length. For example, in case of the Riemann zeta function Hardy and Littlewood proved in 1921 the following

THEOREM 12.1. *Let $s = \sigma + it$ with $0 \leqslant \sigma \leqslant 1$ and let $2\pi xy = t$ with $x \geqslant 1$, $y \geqslant 1$. Then*

$$\zeta(s) = \sum_{n \leqslant x} n^{-s} + \varepsilon(s) \sum_{n \leqslant y} n^{s-1} + O(x^{-\sigma} + y^{\sigma-1} t^{\frac{1}{2}-\sigma}) \tag{12.1}$$

where

$$\varepsilon(s) = \frac{\gamma(1-s)}{\gamma(s)}, \qquad \gamma(s) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right), \tag{12.2}$$

and the implied constant is absolute.

By (6.5) we have the approximate expansion

$$\zeta(s) = \sum_{n \leqslant T} n^{-s} + O\big(T^{-\frac{1}{2}}\big), \tag{12.3}$$

for $s = \frac{1}{2} + it$ with $T < |t| < 2T$. Hence estimating trivially we derive

$$\zeta\left(\frac{1}{2} + it\right) \ll \big(|t| + 1\big)^{\frac{1}{2}}. \tag{12.4}$$

On the other hand, the formula of Hardy-Littlewood offers approximations by two Dirichlet polynomials which are much shorter. For example, choosing $x = y = \sqrt{t/2\pi}$ for $t \geqslant 2\pi$ we obtain

$$\zeta\left(\frac{1}{2} + it\right) = \sum_{n \leqslant \sqrt{t/2\pi}} n^{-\frac{1}{2}-it} + \varepsilon\left(\frac{1}{2} + it\right) \sum_{n \leqslant \sqrt{t/2\pi}} n^{it-\frac{1}{2}} + O\big(t^{-\frac{1}{4}}\big) \tag{12.5}$$

with $|\varepsilon(\frac{1}{2} + it)| = 1$. Estimating trivially we derive

$$\zeta\left(\frac{1}{2} + it\right) \ll \big(|t| + 1\big)^{\frac{1}{4}}. \tag{12.6}$$

This is called the "convexity bound," because it can be obtained directly by the convexity principle of Phragmén-Lindelöf (with an extra factor $\log(2+|t|)$).

On the critical line $\sigma = 1/2$, the approximation of Hardy-Littlewood contains an error term of order of magnitude $x^{-1/2}+y^{-1/2}$, which is as small as the last terms of the partial sums, therefore this error term cannot be improved. However, stronger approximations are possible by smoothly weighted partial sums. For example, we have

THEOREM 12.2. *Let $s = \frac{1}{2}+it$ and $\lambda = s(1-s) = \frac{1}{4}+t^2$. Choose $x > 0$, $y > 0$ with $2\pi xy = \sqrt{\lambda}$. Then we have*

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}e^{-n/x} + \varepsilon(s)\sum_{n=1}^{\infty} n^{s-1}e^{-n/y} - \Gamma(1-s)x^{1-s} + O\big(|s|^{-3/4}\big), \tag{12.7}$$

where the implied constant is absolute.

PROOF. In fact we shall prove (12.7) with the error term $O(|s|^{\alpha-1})$, where α is the exponent in the bound

$$\zeta(s) \ll |s|^{\alpha}, \qquad \text{for } \operatorname{Re} s = \frac{1}{2}. \tag{12.8}$$

We begin by the first sum

$$\begin{aligned}\sum_n n^{-s}e^{-n/x} &= \sum_n n^{-s}\frac{1}{2\pi i}\int_{(1)} \Gamma(w)\left(\frac{x}{n}\right)^w \mathrm{d}w \\ &= \frac{1}{2\pi i}\int_{(1)} \Gamma(w)x^w\zeta(s+w)\,\mathrm{d}w.\end{aligned}$$

Now we move the integration to the line $\operatorname{Re} w = -1/4$. The simple poles at $w = 1-s$ and $w = 0$ contribute

$$\Gamma(1-s)x^{1-s} + \zeta(s). \tag{12.9}$$

Next, on the line $\operatorname{Re} w = -1/4$ we use the functional equation

$$\zeta(s+w) = \varepsilon(s+w)\zeta(1-s-w).$$

We write

$$\varepsilon(s+w) = \varepsilon(s)\left(\frac{2\pi}{\sqrt{s(1-s)}}\right)^w \big(1+w\rho(s,w)\big),$$

say, where $\rho(s,w)$ is holomorphic in the strip $|\operatorname{Re} w| \leqslant \frac{1}{4}$. Accordingly we split the resulting integral

$$\frac{1}{2\pi i}\int_{(-1/4)} \Gamma(w)x^w\varepsilon(s+w)\zeta(1-s-w)\,\mathrm{d}w = \varepsilon(s)I_1(s) + \varepsilon(s)I_\rho(s),$$

with

$$\begin{aligned}I_1(s) &= \frac{1}{2\pi i}\int_{(-1/4)} \Gamma(w)y^{-w}\zeta(1-s-w)\,\mathrm{d}w \\ I_\rho(s) &= \frac{1}{2\pi i}\int_{(-1/4)} \Gamma(1+w)y^{-w}\rho(s,w)\zeta(1-s-w)\,\mathrm{d}w.\end{aligned}$$

To compute $I_1(s)$ we move the integration further to the line $\operatorname{Re} w = -3/4$, then expand $\zeta(1-s-w)$ into Dirichlet series, and interchange the summation and integration to get

$$I_1(s) = -\sum_n n^{s-1} e^{-n/y}. \tag{12.10}$$

To estimate $I_\rho(s)$ we move to the line $\operatorname{Re} w = 0$ and obtain

$$|I_\rho(s)| \leqslant \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \Gamma(1+iu)\rho(s,iu)\zeta(1-s-iu) \right| \mathrm{d}u.$$

Here we have

$$\zeta(1-s-iu) \ll |1-s-iu|^\alpha \ll |s|^\alpha (1+|u|)^\alpha$$
$$\Gamma(1+iu) \ll (1+|u|)^{\frac{1}{2}} \exp\left(-\frac{\pi}{2}|u|\right).$$

To estimate $\rho(s,iu)$ we use Stirling's formula with a good error term. First we get

$$\frac{\varepsilon(s+w)}{\varepsilon(s)} = \pi^w \frac{\Gamma\big(\frac{1}{2}(1-s-w)\big)}{\Gamma\big(\frac{1}{2}(1-s)\big)} \cdot \frac{\Gamma(\frac{s}{2})}{\Gamma(\frac{s+w}{2})}$$
$$= (2\pi)^w \big(s(1-s)\big)^{-w/2} \left\{ 1 + O\left((1+|w|) \frac{|w|}{|s|} \right) \right\}.$$

For $w = iu$ this yields $\rho(s,iu) \ll |1+iu|/|s|$. Collecting the above results we conclude that

$$I_\rho(s) \ll |s|^{\alpha-1}. \tag{12.11}$$

Adding the results we obtain the formula (12.7) with the error term (12.11).

The resulting formula (12.7) is self-improving. Indeed, applying (12.7) for $x = y$ and estimating trivially we get $\zeta(s) \ll |s|^{1/4} + |s|^{\alpha-1}$. This yields (12.8) with $\alpha = 1/4$, and completes the proof of (12.7). □

Our proof of (12.7) can be generalized to obtain similar formulas in which the smooth weight functions $e^{-n/x}$, $e^{-n/y}$ are replaced by other pairs of functions satisfying some functional properties. Actually one can keep all the resulting integrals quite explicit and obtain an exact formula for $\zeta(s)$. Here is one of many formulas of such form:

PROPOSITION 12.3. *Let $G(u)$ be a holomorphic function in the strip $|\operatorname{Re} u| < 4$, $G(u)$ even, and $G(0) = 1$ for normalization. Let $X > 0$. Then for $0 \leqslant \operatorname{Re} s \leqslant 1$ we have*

$$\zeta(s) = \sum_n n^{-s} V_s\left(\frac{n}{X}\right) + \varepsilon(s) \sum_n n^{s-1} V_{1-s}(nX) + R(s,X). \tag{12.12}$$

Here $V_s(y)$ is the smooth weight function given by

$$V_s(y) = \frac{1}{2\pi i} \int_{(3)} \frac{\gamma(s+u)}{\gamma(s)} y^{-u} \frac{G(u)}{u} \mathrm{d}u \tag{12.13}$$

and $R(s,X)$ is the polar term given by

$$\gamma(s) R(s,X) = \frac{G(1-s)}{1-s} X^{1-s} + \frac{G(-s)}{-s} X^{-s}. \tag{12.14}$$

REMARKS. The exact formula (12.12) discloses the functional equation $\zeta(s) = \varepsilon(s)\zeta(1-s)$, nevertheless it should not be called "an approximate functional equation." One can differentiate or integrate (12.12) in the parameter X to produce new desirable features.

There are many interesting choices of the test function $G(u)$, for example

$$G(u) = \left(\cos \frac{\pi u}{A}\right)^{-A}, \qquad A \geqslant 4. \tag{12.15}$$

http://dx.doi.org/10.1090/ulect/062/13

CHAPTER 13

The Dirichlet Polynomials

1. General Properties

Let $\mathcal{A} = (a_n)$ be a sequence of complex numbers. The finite sum

$$A(s) = \sum_{1 \leqslant n \leqslant N} a_n n^{-s}$$

is called a *Dirichlet polynomial* of length N and coefficients a_n. The Dirichlet polynomials have distinct properties which yield pretty good estimates on vertical lines.

First observe that the product of Dirichlet polynomials of length $N_1, \ldots, N_k$ is a Dirichlet polynomial of length $N = N_1 \cdots N_k$;

$$\Big(\sum_{n_1 \leqslant N_1} a_{n_1}^{(1)} n_1^{-s}\Big) \cdots \Big(\sum_{n_k \leqslant N_k} a_{n_k}^{(k)} n_k^{-s}\Big) = \sum_{n \leqslant N} a_n n^{-s}$$

with

$$a_n = \sum_{\substack{n_1 \cdots n_k = n \\ n_1 \leqslant N_1, \ldots, n_k \leqslant N_k}} a_{n_1}^{(1)} \cdots a_{n_k}^{(k)}. \tag{13.1}$$

Hence the *absolute mean value* of the coefficients is bounded by the product of the absolute mean values;

$$\sum_{n \leqslant N} |a_n| \leqslant \Big(\sum_{n_1 \leqslant N_1} |a_{n_1}^{(1)}|\Big) \cdots \Big(\sum_{n_k \leqslant N_k} |a_{n_k}^{(k)}|\Big). \tag{13.2}$$

For the *square mean value* of the coefficients

$$G(\mathcal{A}) = \sum_{n \leqslant N} |a_n|^2 \tag{13.3}$$

we get

$$\begin{aligned} G(\mathcal{A}) &\leqslant \sum \cdots \sum_{n_1 \cdots n_k = n_{k+1} \cdots n_{2k}} |a_{n_1}^{(1)} \cdots a_{n_k}^{(k)}| \cdot |a_{n_{k+1}}^{(1)} \cdots a_{n_{2k}}^{(k)}| \\ &\leqslant \sum_{n_1} \cdots \sum_{n_k} |a_{n_1}^{(1)} \cdots a_{n_k}^{(k)}|^2 \tau_k(n_1 \cdots n_k). \end{aligned}$$

Since $\tau_k(m) \ll m^{\varepsilon}$, this gives

$$G(\mathcal{A}) \ll N^{\varepsilon} G(\mathcal{A}_1) \cdots G(\mathcal{A}_k) \tag{13.4}$$

where the implied constant depends only on k and ε.

The derivatives of a Dirichlet polynomial are also Dirichlet polynomials given by

$$A^{(k)}(s) = (-1)^k \sum_{n\leqslant N} a_n(\log n)^k n^{-s}.$$

Note that the coefficients do not change much; the $\log n$ is a slowly increasing function.

If $A(s)$ is a Dirichlet polynomial not identically zero, then

$$\frac{1}{A(s)} = \sum_m b_m m^{-s}$$

is a Dirichlet series which converges absolutely in $\operatorname{Re} s \geqslant \alpha$ for sufficiently large α depending on the coefficients $\mathcal{A} = (a_n)$. To be explicit let us consider the series

$$A(s) = 1 + \sum_{n=2}^{\infty} a_n n^{-s} \tag{13.5}$$

which converges absolutely for sufficiently large $\sigma = \operatorname{Re} s$. Let α be such that

$$\sum_{n=2}^{\infty} |a_n| n^{-\alpha} \leqslant \frac{1}{2}. \tag{13.6}$$

Then for $\operatorname{Re} s \geqslant \alpha$ we have

$$\frac{1}{A(s)} = 1 + \sum_{k=1}^{\infty}\left(-\sum_{n=2}^{\infty} a_n n^{-s}\right)^k = 1 + \sum_{m=2}^{\infty} b_m m^{-s}$$

with

$$b_m = \sum_{\substack{n_1\cdots n_r = m\\ n_1,\ldots,n_r \geqslant 2}} (-1)^r a_{n_1}\cdots a_{n_r}.$$

We estimate b_m as follows:

$$|b_m| \leqslant m^{\alpha} \sum_{r=1}^{\infty}\left(\sum_{n=2}^{\infty} |a_n| n^{-\alpha}\right)^r \leqslant m^{\alpha}.$$

Therefore the series

$$\frac{1}{A(s)} = 1 + \sum_{m=2}^{\infty} b_m m^{s} \tag{13.7}$$

converges absolutely for $\operatorname{Re} s > \alpha + 1$.

Similarly one can prove that the series

$$\frac{-A'(s)}{A(s)} = \sum_{q=2}^{\infty} c_q q^{-s} \tag{13.8}$$

converges absolutely for $\operatorname{Re} s > \alpha + 1$. Indeed we have

$$\frac{-A'(s)}{A(s)} = -\big(\log A(s)\big)' = \sum_{k=1}^{\infty} \frac{1}{k}\left(-\sum_{n=2}^{\infty} a_n n^{-s}\right)^k.$$

Hence (13.8) holds with

$$|c_q| \leqslant q^{\alpha} \sum_{r=1}^{\infty} \frac{1}{r}\left(\sum_{n=2}^{\infty} |a_n| n^{-\alpha}\right)^r \leqslant q^{\alpha} \log 2.$$

2. The Mean Value of Dirichlet Polynomials

Using Cauchy's inequality we can estimate the Dirichlet polynomial

$$A(s) = \sum_{n \leqslant N} a_n n^{-s} \tag{13.9}$$

on the line $\mathrm{Re}\, s = 0$ trivially getting

$$|A(it)|^2 \leqslant GN \tag{13.10}$$

for any $t \in \mathbb{R}$, where

$$G = G(\mathcal{A}) = \sum_{n \leqslant N} |a_n|^2. \tag{13.11}$$

This trivial bound can be improved on average:

THEOREM 13.1. *For $T > 0$ we have*

$$\int_0^T |A(it)|^2 \, \mathrm{d}t = TG + O(G^{\frac{1}{2}} H^{\frac{1}{2}}) \tag{13.12}$$

where

$$H = H(\mathcal{A}) = \sum_{n \leqslant N} n^2 |a_n|^2, \tag{13.13}$$

and the implied constant is absolute.

PROOF. Let $f(t)$ be the piecewise linear function whose graph is given by Figure 13.1. Then

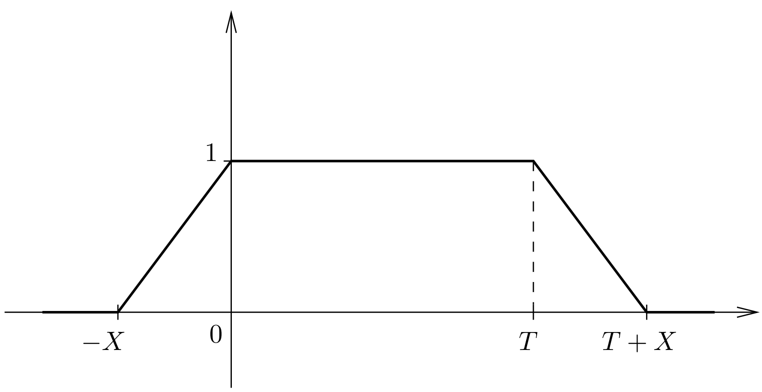

FIGURE 13.1

$$\int_0^T |A(it)|^2 \, \mathrm{d}t \leqslant \int f(t) |A(it)|^2 \, \mathrm{d}t = \sum_m \sum_n a_m \overline{a_n} \int f(t) \left(\frac{m}{n}\right)^{it} \mathrm{d}t.$$

For $m = n$ (the diagonal terms) we get

$$\int f(t) \, \mathrm{d}t = T + X.$$

For $m \neq n$ (the off-diagonal terms) we get by partial integration

$$\int f(t) \left(\frac{m}{n}\right)^{it} \mathrm{d}t = \left(\log \frac{m}{n}\right)^{-2} \int_{-X}^{T+X} f'(t) \, \mathrm{d} \left(\frac{m}{n}\right)^{it} \ll X^{-1} \left(\log \frac{m}{n}\right)^{-2}.$$

Hence

$$\sum_{m \neq n} a_m \overline{a_n} \int f(t) \left(\frac{m}{n}\right)^{it} \mathrm{d}t \ll X^{-1} \sum_{m \neq n} |a_m a_n| \left(\log \frac{m}{n}\right)^{-2}$$

$$\leqslant X^{-1} \sum_n |a_n|^2 \sum_{1 \leqslant m < n} \left(\frac{m}{n-m}\right)^2 \ll X^{-1} H.$$

Adding the diagonal contribution we see that the integral (13.12) is estimated from above by

$$TG + O(XG + X^{-1}H). \tag{13.14}$$

Similarly we get (13.14) as a lower bound for the integral (13.12). Hence, choosing $X = (H/G)^{1/2}$ we get the formula (13.12). □

Estimating H by GN^2 we get

COROLLARY 13.2. *For $T > 0$ we have*

$$\int_0^T |A(it)|^2 \, \mathrm{d}t = \big(T + O(N)\big) G \tag{13.15}$$

where the implied constant is absolute.

This result shows that G is the mean value of $|A(it)|^2$ on the segment $0 < t < T$, provided T is somewhat larger than the length of the polynomial.

Sometimes we need estimates for sums $A(s)$ whose coefficients a_n depend on s in a steady fashion. By (13.15) we derive

COROLLARY 13.3. *Suppose for $T \leqslant t \leqslant 2T$ and $n \leqslant N$ that we have*

$$|a_n(t)| \leqslant a_n, \qquad t|a_n'(t)| \leqslant a_n. \tag{13.16}$$

Then

$$\int_T^{2T} \left| \sum_{n \leqslant N} a_n(t) n^{it} \right|^2 \mathrm{d}t \leqslant \big(2T + O(N)\big) \sum_{n \leqslant N} |a_n|^2. \tag{13.17}$$

PROOF. Apply the Cauchy-Schwarz inequality to

$$\left| \sum_n a_n(t) n^{it} \right| \leqslant \left| \sum_n a_n(T) n^{it} \right| + \int_T^{2T} \left| \sum_n a_n'(\tau) n^{it} \right| \mathrm{d}\tau.$$ □

COROLLARY 13.4. *For $T \geqslant 2$ we have*

$$\int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^2 \mathrm{d}t = T \log T + O\big(T (\log T)^{\frac{1}{2}}\big). \tag{13.18}$$

PROOF. By the approximation (6.5) for $s = \frac{1}{2} + it$ with $\frac{1}{2}T < t < T$ we have

$$\zeta(s) = \sum_{n \leqslant T} n^{-s} + O(T^{-\frac{1}{2}}).$$

Hence

$$|\zeta(s)|^2 = \Big|\sum_{n\leqslant T} n^{-s}\Big|^2 + O(1)$$

and (13.12) shows that

$$\int_{T/2}^{T} \Big|\zeta\Big(\frac{1}{2}+it\Big)\Big|^2 \,\mathrm{d}t = \frac{T}{2}\sum_{n\leqslant T}\frac{1}{n} + O\left(T + \Big(\sum_{n\leqslant T}\frac{1}{n}\Big)^{1/2}\Big(\sum_{n\leqslant T} n\Big)^{1/2}\right)$$
$$= \frac{T}{2}\log T + O\big(T(\log T)^{\frac{1}{2}}\big).$$

This implies the formula (13.18). □

Corollary 13.4 shows that the average value of $|\zeta(\frac{1}{2}+it)|^2$ is of size $\log t$ for $t\geqslant 2$. There are more precise results of the form

(13.19) $$\int_0^T \Big|\zeta\Big(\frac{1}{2}+it\Big)\Big|^2 \,\mathrm{d}t = T(\log T + 2\gamma - 1 + \log 2\pi) + O(T^{\theta+\varepsilon})$$

where $\gamma = 0.57\ldots$ is the Euler constant and $\frac{1}{4}\leqslant\theta\leqslant\frac{1}{2}$. For example (13.19) is known to hold with $\theta = 7/22$. There is a conjecture that (13.19) holds with $\theta = 1/4$. In 1926, A. E. Ingham evaluated the fourth power moment

(13.20) $$\int_0^T \Big|\zeta\Big(\frac{1}{2}+it\Big)\Big|^4 \,\mathrm{d}t = \frac{T}{2\pi^2}(\log T)^4 + O\big(T(\log T)^3\big).$$

Therefore the average value of $|\zeta(\frac{1}{2}+it)|^4$ is of size $(\log t)^4$, which is not compatible with the average value of $|\zeta(\frac{1}{2}+it)|^2$.

3. Large Values of Dirichlet Polynomials

The question is: how often does $A(s)$ take large values on a set of points $s = s_r$ which are well-spaced? To address this question, first we are going to establish an estimate for $|A(s_r)|$ in terms of an integral of $|A(s)|$ over a short vertical segment centered at $s = s_r$. In fact we consider sums of type

(13.21) $$A_r = \sum_{1\leqslant n\leqslant N} a_n f_r(n) n^{it_r}, \qquad N\geqslant 2,$$

where $f_r(n)$ are smooth functions which have relatively small derivatives.

Lemma 13.5. *Suppose $f_r(x)$ is of $\mathcal{C}^2$-class on $1\leqslant x\leqslant N$ with*

(13.22) $$x^a|f_r^{(a)}(x)|\leqslant 2, \qquad a = 0,\ 1,\ 2.$$

Then

(13.23) $$|A_r| \ll (\log N)\int_{-\infty}^{\infty} |A(it_r - it)|\big(|t|+1\big)^{-2}\,\mathrm{d}t$$

where the implied constant is absolute.

PROOF. One can extend $\frac{1}{4}f_r(x)$ to be supported in $[\frac{1}{2}, 2N]$ and satisfy (13.22). Then the Mellin transform of $f_r(x)$ satisfies

$$\hat{f}_r(it) = \int_0^\infty f_r(x)x^{it-1}\,\mathrm{d}x \ll \big(|t|+1\big)^{-2}\log N.$$

By Mellin's inversion, for every $1 \leqslant n \leqslant N$ we have

$$f_r(n) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}_r(it)n^{-it}\,\mathrm{d}t.$$

Hence

$$A_r = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}_r(it)A(it_r - it)\,\mathrm{d}t.$$

This yields (13.23). □

Next, by Cauchy-Schwarz inequality (13.23) yields

$$|A_r|^2 \ll (\log N)^2 \int_{-\infty}^{\infty} \big|A(it_r - it)\big|^2\big(|t|+1\big)^{-2}\,\mathrm{d}t. \tag{13.24}$$

Let $\mathcal{T} = \{t_1, \ldots, t_R\}$ be a set of points in the segment $0 < t < T$ with

$$|t_r - t_{r'}| \geqslant 1, \qquad \text{if } r \neq r'. \tag{13.25}$$

Summing (13.24) over $r = 1, \ldots, R$ we get

$$\begin{aligned}\sum_{r=1}^{R} |A_r|^2 &\ll (\log N)^2 \int_{-\infty}^{\infty} |A(it)|^2 \sum_{r=1}^{R} \big(|t-t_r|+1\big)^{-2}\,\mathrm{d}t \\ &\ll (\log N)^2 \int_{-\infty}^{\infty} |A(it)|^2 \left(1 + \frac{|t|}{T}\right)^{-2}\,\mathrm{d}t.\end{aligned}$$

Now applying Corollary 13.2 we conclude the following

THEOREM 13.6. *Let $\mathcal{T} = \{t_1, \ldots, t_R\}$ be a set of points in the segment $0 < t < T$ such that* (13.25) *is true, and for every r let $f_r(x)$ be a function satisfying* (13.22). *Then for any complex numbers a_n we have*

$$\sum_{r=1}^{R} \left|\sum_{n \leqslant N} a_n f_r(n) n^{it_r}\right|^2 \ll (T+N)G(\log N)^2 \tag{13.26}$$

where the implied constant is absolute.

COROLLARY 13.7. *Let $\mathcal{S} = \{s_1, \ldots, s_R\}$ be a set of points with*

$$s_r = \sigma_r + it_r, \qquad \sigma_r \geqslant 0,\ 0 < t_r < T,$$

which are well-spaced, *i.e., $|t_r - t_{r'}| \geqslant 1$ for $r \neq r'$. Then*

$$\sum_{r=1}^{R} |A(s_r)|^2 \ll (T+N)G(\log N)^2 \tag{13.27}$$

where the implied constant is absolute.

Suppose that for every s_r we have

$$|A(s_r)| \geqslant V. \tag{13.28}$$

Then (13.27) shows that (13.28) cannot happen too often;

$$R \ll (T+N)GV^{-2}(\log N)^2. \tag{13.29}$$

http://dx.doi.org/10.1090/ulect/062/14

CHAPTER 14

Zeros off the Critical Line

All the zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ in the strip $0 \leqslant \operatorname{Re} s \leqslant 1$ are on the critical line $\operatorname{Re} s = \frac{1}{2}$. Although this statement (the Riemann Hypothesis) seems to be out of reach for some time, we can prove that almost all zeros are arbitrarily close to the critical line.

THEOREM 14.1. *Let $\frac{1}{2} \leqslant \alpha \leqslant 1$, $T \geqslant 3$ and $N(\alpha, T)$ denote the number of zeros $\rho = \beta + i\gamma$ with*

$$\alpha \leqslant \beta < 1, \qquad 0 < \gamma \leqslant T. \tag{14.1}$$

We have

$$N(\alpha, T) \ll T^{4\alpha(1-\alpha)}(\log T)^{15}. \tag{14.2}$$

In 1920, F. Carlson [**Car21**] proved (14.2) (up to the logarithmic factor). Our arguments are a bit different, but they are based on the same principles.

Let $s = \sigma + it$ be in the rectangle $\alpha \leqslant \sigma \leqslant 1$, $T < t \leqslant 2T$. We begin by the approximation (see (6.5))

$$\zeta(s) = \sum_{n<T} n^{-s} + O(T^{-\alpha}) \tag{14.3}$$

where the implied constant is absolute. It is expected that the partial sums

$$M(s) = \sum_{1 \leqslant m \leqslant M} \mu(m) m^{-s} \tag{14.4}$$

yield good approximations to

$$\frac{1}{\zeta(s)} = \sum_{m=1}^{\infty} \mu(m) m^{-s}$$

for s to the right of the critical line. Having this fact in mind we are going to investigate the "mollified" zeta function $\zeta(s)M(s)$, which should be close to 1. For s in the rectangle $\sigma \geqslant \alpha$, $T < t \leqslant 2T$, we have

$$M(s) \ll M^{1-\alpha} \log T, \qquad \text{if } 1 \leqslant M \leqslant T. \tag{14.5}$$

Hence

$$\begin{aligned}\zeta(s)M(s) &= \Big(\sum_{1 \leqslant n \leqslant T} n^{-s}\Big)\Big(\sum_{1 \leqslant m \leqslant M} \mu(m) m^{-s}\Big) + O(M^{1-\alpha}T^{-\alpha}\log T)\\ &= 1 + \sum_{M < l \leqslant MT} a_l l^{-s} + O(M^{1-\alpha}T^{-\alpha}\log T) \end{aligned} \tag{14.6}$$

with

$$a_l = \sum_{\substack{mn=l\\ m\leqslant M,\, n\leqslant T}} \mu(m), \qquad |a_l| \leqslant \tau(l).$$

According to a "folklore philosophy," a sum of oscillating terms (in unbiased fashion) should get cancellation which yields a bound with saving factor being about the square root of the number of terms. Applying this to the polynomial

$$A(s) = \sum_{M<l\leqslant MT} a_l l^{-s} \tag{14.7}$$

in (14.6) we would get $\zeta(s)M(s) = 1 + O(M^{\frac{1}{2}-\alpha+\varepsilon})$. This implies $\zeta(s) \neq 0$ for $\sigma \geqslant \alpha$ and any $\alpha > \frac{1}{2}$ (the Riemann hypothesis), which is too good to hope for in the near future. A good news is that one can establish almost the same things on average over $s = \sigma + it$ in the rectangle $\alpha \leqslant \sigma \leqslant 1$, $T \leqslant t \leqslant 2T$.

For technical reasons we split the polynomial $A(s)$ into partial sums $A_L(s)$ with the extra restriction $L < l \leqslant 2L$. We need at most $2\log T$ such partial sums with $M \leqslant L \leqslant MT$ to cover the range $M < l \leqslant MT$. We get

$$\zeta(s)M(s) = 1 + \sum_L A_L(s) + O\big((\log T)^{-2}\big) \tag{14.8}$$

where $M \leqslant L \leqslant MT$ with $1 \leqslant M \leqslant T(\log T)^{-6}$. Hence, if $s = \rho$ is a zero of $\zeta(s)$ in the relevant rectangle, then

$$|A_L(\rho)| > (3\log T)^{-1} \tag{14.9}$$

for some L. In other words, at least one of the polynomials $A_L(s)$ assumes a relatively large value at $s = \rho$. For this reason we call $A_L(s)$ the zero detector polynomial. On the other hand (13.29) shows that (14.9) cannot happen too often, specifically the number R_L of zeros ρ detected by (14.9) satisfies

$$R_L \ll (T+L)\left(\sum_{L<l\leqslant 2L} |a_l|^2 l^{-2\alpha}\right)(\log T)^5.$$

Here the extra factor $\log T$ is introduced to account for the clusters of zeros which are not well-spaced; see (8.18). Hence

$$R_L \ll (T+L)L^{1-2\alpha}(\log T)^8 \ll \big(TM^{1-2\alpha} + (MT)^{2-2\alpha}\big)(\log T)^8.$$

Choosing $M = T^{2\alpha-1}(\log T)^{-6}$, this gives $R_L \ll T^{4\alpha(1-\alpha)}(\log T)^{14}$, and (14.2) holds.

Actually the above choice of M requires $M \geqslant 1$, which means $\alpha \geqslant \frac{1}{2} + 3(\log\log T)/\log T$. However, for α smaller, the estimate (14.2) is obvious; $N(\alpha, T) \leqslant N(T) \ll T\log T$.

Corollary 14.2. *Let $\Phi(T) \to \infty$ as $T \to \infty$. Then almost all zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ are in the region*

$$\left|\beta - \frac{1}{2}\right| < \Phi(T)(\log\log T)/\log T. \tag{14.10}$$

In 1942 A. Selberg proved that $N(\alpha, T) \ll (\alpha - \frac{1}{2})^{-1}T$ if $\alpha > \frac{1}{2}$, the implied constant being absolute. Hence, almost all zeros satisfy $|\beta - \frac{1}{2}| < \Phi(T)/\log T$.

http://dx.doi.org/10.1090/ulect/062/15

CHAPTER 15

Zeros on the Critical Line

Recall that $N(T) = N(0, T)$ denotes the number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ in the rectangle $0 < \beta < 1$, $0 < \gamma \leqslant T$, and that

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi e} + O(\log T)$$

for $T \geqslant 2$. Denote by $N_0(T)$ the number of zeros with $\beta = \frac{1}{2}$ and $0 < \gamma \leqslant T$. We call these zeros the *critical zeros*. By numerical computations with bare hands, B. Riemann showed that the first critical zero is

$$\rho_1 = \frac{1}{2} + i\gamma_1, \qquad \text{with} \qquad \gamma_1 = 14.13\ldots$$

The Riemann Hypothesis asserts that all non-real zeros are critical;

$$N_0(T) = N(T).$$

Fifty five years later, G. H. Hardy proved that there are infinitely many critical zeros, and in 1921 he and J. E. Littlewood [**HL21**] got a very good lower bound for $N_0(T)$:

THEOREM 15.1. *For $T \geqslant 15$ we have*

$$N_0(T) \gg T. \tag{15.1}$$

In this section we shall prove this bound, which will be then improved in Part 2 of these lectures.

The idea is to count the sign change of the normalized zeta function on the critical line;

$$Z(u) = \frac{H(\frac{1}{2} + iu)}{|H(\frac{1}{2} + iu)|} \zeta\left(\frac{1}{2} + iu\right). \tag{15.2}$$

Here $H(s)$ denotes the local zeta function at the infinite place, precisely

$$H(s) = \frac{1}{2} s(1 - s)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right). \tag{15.3}$$

Recall the functional equation

$$H(s)\zeta(s) = H(1 - s)\zeta(1 - s). \tag{15.4}$$

This implies that $Z(u)$ is real and even for $u \in \mathbb{R}$. Therefore, when $Z(u)$ changes sign, it yields a critical zero $\rho = \frac{1}{2} + i\gamma$ in between, because $H(\frac{1}{2} + iu)$ does not vanish. Therefore we want to show that the sign change of $Z(u)$ occurs quite often.

To this end we compare the two integrals

$$I(t) = \int_t^{t+\Delta} Z(u)\,\mathrm{d}u \tag{15.5}$$

$$J(t) = \int_t^{t+\Delta} |Z(u)|\,\mathrm{d}u \tag{15.6}$$

in the range $T \leqslant t \leqslant 2T$, where Δ is a sufficiently large absolute constant to be chosen later. We need an upper bound for $|I(t)|$ and a lower bound for $J(t)$ on average over a subset $\mathcal{T} \subset [T, 2T]$.

We begin by estimating $J(t)$, because it is quite easy. We have

$$\begin{aligned} J(t) &= \int_t^{t+\Delta} \left|\zeta\left(\frac{1}{2}+iu\right)\right| \mathrm{d}u \geqslant \left|\int_t^{t+\Delta} \zeta\left(\frac{1}{2}+iu\right) \mathrm{d}u\right| \\ &\geqslant \Delta - \left|\int_t^{t+\Delta} \left(\zeta\left(\frac{1}{2}+iu\right) - 1\right) \mathrm{d}u\right| \\ &= \Delta - \left|\int_t^{t+\Delta} \left(\sum_{1<n\leqslant T} n^{-\frac{1}{2}-iu}\right) \mathrm{d}u\right| + O(\Delta T^{-\frac{1}{2}}) \\ &= \Delta - \left|\sum_{1<n\leqslant T} \frac{1-n^{-i\Delta}}{\log n} n^{-\frac{1}{2}-it}\right| + O(\Delta T^{-\frac{1}{2}}) \end{aligned}$$

by the approximation (12.3). Next, by (13.15) we get

$$\int_T^{2T} \left|\sum_{1<n\leqslant T} \frac{1-n^{-i\Delta}}{\log n} n^{-\frac{1}{2}-it}\right|^2 \mathrm{d}t \ll T.$$

Combining the above estimates we derive using the Cauchy-Schwarz inequality that

$$\int_{\mathcal{T}} J(t)\,\mathrm{d}t > \Delta|\mathcal{T}| + O\big(|\mathcal{T}|^{\frac{1}{2}} T^{\frac{1}{2}} + \Delta|\mathcal{T}|T^{-\frac{1}{2}}\big) \tag{15.7}$$

where $|\mathcal{T}|$ is the measure of $\mathcal{T}$ and the implied constant is absolute.

An upper bound for $|I(t)|$ is harder to get. We shall prove

Lemma 15.2. *For $\Delta \geqslant 1$ and $T \geqslant \Delta^6$ we have*

$$\int_T^{2T} |I(t)|^2\,\mathrm{d}t \ll \Delta T \tag{15.8}$$

where the implied constant is absolute.

Assuming Lemma 15.2 we can complete the proof of Theorem 15.1 as follows. First by Cauchy-Schwarz inequality we get

$$\int_{\mathcal{T}} |I(t)|\,\mathrm{d}t \ll \big(\Delta|\mathcal{T}|T\big)^{\frac{1}{2}} \tag{15.9}$$

where the implied constant is absolute. Let $\mathcal{T}$ be the subset of $[T, 2T]$ for which

$$|I(t)| = J(t). \tag{15.10}$$

This is the set of t's such that $Z(u)$ does not change sign in the interval $(t,\, t+\Delta)$. Since

$$\int_{\mathcal{T}} |I(t)|\,\mathrm{d}t = \int_{\mathcal{T}} J(t)\,\mathrm{d}t, \tag{15.11}$$

we deduce by comparing the bounds (15.7) and (15.9) that the measure of the set $\mathcal{T}$ satisfies

$$|\mathcal{T}| \ll \Delta^{-1} T \tag{15.12}$$

where the implied constant is absolute. If Δ is sufficiently large we get $|\mathcal{T}| \leqslant \frac{1}{2}T$. Hence the set $\mathcal{S} = [T,\, 2T] \setminus \mathcal{T}$ of points t with

$$|I(t)| < J(t) \tag{15.13}$$

has measure $|\mathcal{S}| \geqslant \frac{1}{2}T$. Such set contains a sequence $\{t_1, \ldots, t_R\}$ of Δ-spaced points of length $R \geqslant T/2\Delta$. For every t_r, the function $Z(u)$ must change sign in the segment $t_r < u < t_r + \Delta$, hence there is a critical zero $\rho_r = \frac{1}{2} + i\gamma_r$ with $t_r < \gamma_r < t_r + \Delta$. Therefore the number of critical zeros $\rho = \frac{1}{2} + i\gamma$ with $T < \gamma < 2T$ is at least $T/2\Delta - 1$. This proves (15.1) by changing $2T$ to T.

It remains to prove Lemma 15.2. By the convexity bound $\zeta(s) \ll |s|^{\frac{1}{4}}$ on the critical line $\operatorname{Re} s = \frac{1}{2}$, we arrange the integral (15.8) as follows

$$\begin{aligned}
\int_T^{2T} |I(t)|^2\,\mathrm{d}t &= \int_T^{2T} \left| \int_0^{\Delta} Z(t+u)\,\mathrm{d}u \right|^2 \mathrm{d}t \\
&= \int_0^{\Delta}\int_0^{\Delta}\int_T^{2T} Z(t+u_1)\overline{Z}(t+u_2)\,\mathrm{d}t\,\mathrm{d}u_1\,\mathrm{d}u_2 \\
&= \int_0^{\Delta}\int_0^{\Delta}\int_T^{2T} Z(t)\overline{Z}(t+u_2-u_1)\,\mathrm{d}t\,\mathrm{d}u_1\,\mathrm{d}u_2 + O(\Delta^3 T^{\frac{1}{2}}) \\
&= \int_{-\Delta}^{\Delta} (\Delta - |u|) \int_T^{2T} Z(t)\overline{Z}(t+u)\,\mathrm{d}t\,\mathrm{d}u + O(\Delta^3 T^{\frac{1}{2}}).
\end{aligned}$$

Recall that $Z(t)$ is the normalized zeta function given by (15.2). Here the H-factors can be evaluated quite precisely using Stirling's formula, giving

$$\frac{H(\frac{1}{2}+it)\overline{H}(\frac{1}{2}+it+iu)}{|H(\frac{1}{2}+it)\overline{H}(\frac{1}{2}+it+iu)|} = \left(\frac{2\pi}{t}\right)^{iu/2} \left(1 + O\left(\frac{u^2+1}{T}\right)\right)$$

for $T < t < 2T$ and $|u| \leqslant \Delta$, the implied constant being absolute. Hence using the approximation (12.3) and the convexity bound $\zeta(s) \ll |s|^{\frac{1}{4}}$ we obtain

$$Z(t)\overline{Z}(t+u) = \sum\sum_{1\leqslant m,n\leqslant T} (mn)^{-\frac{1}{2}} \left(\frac{m}{n}\right)^{it} \left(\frac{2\pi m^2}{t}\right)^{iu/2} + O(\Delta^2 T^{-\frac{1}{2}})$$

and

$$\int_T^{2T} |I(t)|^2\,\mathrm{d}t = \sum\sum_{1\leqslant m,n\leqslant T} \frac{c(m,n)}{\sqrt{mn}} + O(\Delta^4 T^{\frac{1}{2}})$$

where

$$c(m,n) = \int_{-\Delta}^{\Delta} (\Delta - |u|) \int_T^{2T} \left(\frac{m}{n}\right)^{it} \left(\frac{2\pi m^2}{t}\right)^{iu/2} \mathrm{d}t\,\mathrm{d}u. \tag{15.14}$$

In the following lines we assume without mention that m, n run over positive integers $\leqslant T$.

The integration in u gives

$$\begin{aligned}\int_{-\Delta}^{\Delta}(\Delta-|u|)y^{iu/2}\,\mathrm{d}u &= \Delta^2\int_{-1}^{1}(1-|u|)y^{iu\Delta/2}\,\mathrm{d}u\\ &= 2\Delta^2\int_0^1(1-u)\cos\left(\frac{\Delta u}{2}\log y\right)\,du\\ &= \Delta^2\chi\left(\frac{\Delta}{4}\log y\right)\end{aligned}$$

where $\chi(x)=(\sin x/x)^2$. Hence (15.14) becomes

$$c(m,n)=\Delta^2\int_T^{2T}\left(\frac{m}{n}\right)^{it}\chi\left(\frac{\Delta}{4}\log\frac{2\pi m^2}{t}\right)\,\mathrm{d}t.$$

For the diagonal terms $m=n$ we get

$$c(m,m)=\Delta^2\int_T^{2T}\chi\left(\frac{\Delta}{4}\log\frac{2\pi m^2}{t}\right)\,\mathrm{d}t.$$

Hence the contribution of the diagonal terms is equal to

$$\begin{aligned}\sum_{1\leqslant m\leqslant T}\frac{c(m,m)}{m} &= \Delta^2\sum_{1\leqslant m\leqslant T}m\int_{T/m^2}^{2T/m^2}\chi\left(\frac{\Delta}{4}\log\frac{2\pi}{t}\right)\,\mathrm{d}t\\ &\leqslant 2\Delta^2\int_1^T x\int_{T/2x^2}^{2T/x^2}\chi\left(\frac{\Delta}{4}\log\frac{2\pi}{t}\right)\,\mathrm{d}t\,\mathrm{d}x\\ &= \Delta^2\int_{T/2}^{2T}\left(\int_{1/t}^{T/t}\chi\left(\frac{\Delta}{4}\log(2\pi y)\right)\frac{\mathrm{d}y}{y}\right)\,\mathrm{d}t\\ &\leqslant \Delta^2\int_{T/2}^{2T}\left(\int_0^{\infty}\chi\left(\frac{\Delta}{4}\log(2\pi y)\right)\frac{\mathrm{d}y}{y}\right)\,\mathrm{d}t\\ &= \frac{3}{2}\Delta^2 T\int_{-\infty}^{\infty}\chi\left(\frac{\Delta}{4}z\right)\,\mathrm{d}z = 6\pi\Delta T\end{aligned}$$

which is admissible for (15.8). In the off-diagonal terms $m\neq n$ we integrate by parts, getting

$$\begin{aligned}i\,c(m,n)\log\frac{m}{n} &= \Delta^2\int_T^{2T}\chi\left(\frac{\Delta}{4}\log\frac{2\pi m^2}{t}\right)\,\mathrm{d}\left(\frac{m}{n}\right)^{it}\\ &\ll \Delta^2\chi\left(\frac{\Delta}{4}\log\frac{\pi m^2}{T}\right)+\Delta^2\chi\left(\frac{\Delta}{4}\log\frac{2\pi m^2}{T}\right)\\ &\qquad+\Delta^3\int_T^{2T}\left|\chi'\left(\frac{\Delta}{4}\log\frac{2\pi m^2}{t}\right)\right|\frac{\mathrm{d}t}{t}.\end{aligned}$$

Since $\chi(x)\ll\min(1,\,x^{-1})$, the first two terms are bounded by $\Delta c(m)$ with

$$c(m)=\min\left(\Delta,\,\left|\log\frac{\pi m^2}{T}\right|^{-1}+\left|\log\frac{2\pi m^2}{T}\right|^{-1}\right).$$

Similarly the third term is bounded by

$$\Delta^2 \int_{\Delta \log(2\pi m^2/T)}^{\Delta \log(\pi m^2/T)} \left|\chi'\left(\frac{z}{4}\right)\right| \,\mathrm{d}z \ll \Delta c(m),$$

because $\chi'(x) \ll \min(1,\, x^{-2})$. Hence we have

$$c(m,n) \ll \Delta c(m) \left|\log \frac{m}{n}\right|^{-1}.$$

Since

$$\sum_{n \neq m} n^{-\frac{1}{2}} \left|\log \frac{m}{n}\right|^{-1} \ll T^{\frac{1}{2}} \log T$$

and

$$\sum_{m} m^{-\frac{1}{2}} c(m) \ll T^{\frac{1}{2}} (\log T)^{-1} + \Delta T^{\frac{1}{4}},$$

we conclude that the contribution of the off-diagonal terms is

$$\sum\sum_{m \neq n} \frac{c(m,n)}{\sqrt{mn}} \ll \Delta T^{\frac{1}{2}} (\log T)\big(T^{\frac{1}{2}} (\log T)^{-1} + \Delta T^{\frac{1}{4}}\big) \ll \Delta T.$$

This completes the proof of Lemma 15.2 and Theorem 15.1.

Part 2

The Critical Zeros after Levinson

http://dx.doi.org/10.1090/ulect/062/16

CHAPTER 16

Introduction

The bound (15.1) of Hardy and Littlewood was improved by A. Selberg [**Sel42**], giving

$$N_0(T) \gg T \log T, \qquad T \geqslant 15. \tag{16.1}$$

This means that a positive proportion of zeros of $\zeta(s)$ rest on the line $\operatorname{Re} s = \frac{1}{2}$. The implied constant in (16.1) can be determined, however it will be very small. Selberg's arguments follow these of Hardy-Littlewood in principle, but he also introduced a number of innovations in the construction of the mollifier. Briefly speaking, Selberg's mollifier is the square of the Dirichlet polynomial which is a smooth truncation of $\zeta(s)^{-1/2}$. On the critical line, the smoothing (rather than the sharp cut as in Chapter 14) is powerful: it produces the gain of factor $\log T$ in (16.1). This idea is indispensable in many current works on L-functions.

A different approach for showing the bound (16.1) was proposed by N. Levinson [**Lev74**]. Although Levinson also applies a smoothed mollifier, he does not follow the Hardy-Littlewood-Selberg set up. He makes use of the functional equation with great finesse. Levinson's approach seems to be a gamble, because it is not clear up front that at the end the numerical constants are good enough to yield a positive lower bound for $N_0(T)$. In the Hardy-Littlewood-Selberg set-up there is no risk of getting a negative result. However, the risk taken by Levinson turned out to be rewarding. It yields a large proportion of the critical zeros. Specifically, Levinson succeeded to show that at least 1/3 of the zeros are critical; precisely

$$N_0(T) > \kappa N(T), \qquad \kappa = 0.3420. \tag{16.2}$$

Subsequently, B. Conrey [**Con89**] introduced several innovations and new inputs from estimates for exponential sums (spectral theory of Kloosterman sums) to show that at least 2/5 of the zeros are critical; precisely (16.2) holds with $\kappa = 0.4088$. The best bound obtained so far is $\kappa = 0.4128$ due to S. Feng [**Fen12**].

Our presentation of Levinson's method (including Conrey's innovations) is divided into two groups. In the first group (Chapters 17-22) we establish a general lower bound (16.2) where κ is expressed by an integral of a certain Dirichlet polynomial, see (22.8). Here the principles of the method are exposed. The second group (Chapters 24-28) is devoted to the evaluation of the relevant integral, which we handle differently than Levinson. In Chapters 23 and 29 we derive computer-friendly formulas for special choices of crop functions and provide numerical values.

http://dx.doi.org/10.1090/ulect/062/17

CHAPTER 17

Detecting Critical Zeros

Put

$$H(s) = \frac{1}{2}s(1-s)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right), \tag{17.1}$$

$$\xi(s) = H(s)\zeta(s) = \frac{1}{2}s(1-s)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s). \tag{17.2}$$

The method of Levinson begins by writing the functional equation

$$\xi(s) = \xi(1-s) \tag{17.3}$$

in the following form

$$Y(s)\xi(s) = H(s)G(s) + H(1-s)G(1-s), \tag{17.4}$$

where $Y(s)$ is a simple function satisfying $Y(s) = Y(1-s)$ (a factor which slightly modifies $H(s)$), and $G(s)$ is a natural relative of $\zeta(s)$. For example, (17.4) holds with $Y(s) = 2$ and $G(s) = \zeta(s)$. However this obvious choice yields poor results. Of course, $G(s)$ is not uniquely defined by the expression (17.4). Many interesting results come out of (17.4) when $G(s)$ is judiciously chosen. The original choice by Levinson is (see the motivation in [Co])

$$G(s) = \zeta(s) + \lambda\zeta'(s) \tag{17.5}$$

where λ is a real number at our disposal. Then (17.4) holds with

$$Y(s) = 2 - \lambda\left(\frac{H'}{H}(s) + \frac{H'}{H}(1-s)\right). \tag{17.6}$$

This is quite simple asymptotically (see Lemma 18.1). Indeed $Y(s) \sim -\lambda\log|s|$ as $|s| \to \infty$ in vertical strips. Hence $Y(s)$ does not vanish for large $|s|$.

The key feature of the equation (17.4) is that

$$Y(s)\xi(s) = 2\,\mathrm{Re}\,H(s)G(s), \qquad \text{if } \mathrm{Re}\,s = \frac{1}{2}. \tag{17.7}$$

Hence the critical zeros of $\zeta(s)$ are just the points $s = \rho$ at which

$$\mathrm{Re}\,H(s)G(s) = 0, \qquad \mathrm{Re}\,s = \frac{1}{2}, \tag{17.8}$$

except for a few zeros of $Y(s)$. Equivalently, these are the points on the critical line $\mathrm{Re}\,s = \frac{1}{2}$ with $G(s) = 0$, or $G(s) \neq 0$ and

$$\arg H(s)G(s) \equiv \frac{\pi}{2} \pmod{\pi}. \tag{17.9}$$

Therefore the problem reduces to the question of how often (17.9) occurs on the line $\mathrm{Re}\,s = \frac{1}{2}$.

http://dx.doi.org/10.1090/ulect/062/18

CHAPTER 18

Conrey's Construction

J. B. Conrey [**Con85**] introduced more derivatives to $G(s)$. He takes

$$2H(s)G(s) = \xi(s) + \sum_{k \text{ odd}} g_k \xi^{(k)}(s), \tag{18.1}$$

where g_k are real numbers at our disposal, almost all vanishing. Note that $G(s)$ given by (18.1) is holomorphic in $\mathbb{C}$ except for a simple pole at $s = 1$. By the functional equation (17.3) it follows that $\xi^{(k)}(s) = (-1)^k \xi^{(k)}(1-s)$, so (18.1) yields

$$\xi(s) = H(s)G(s) + H(1-s)G(1-s). \tag{18.2}$$

Moreover, for k odd $\xi^{(k)}(s)$ is purely imaginary on the line $\operatorname{Re} s = \frac{1}{2}$, so (18.1) yields

$$\operatorname{Re} H(s)G(s) = \xi(s), \qquad \text{if } \operatorname{Re} s = \frac{1}{2}. \tag{18.3}$$

Hence the critical zeros of $\zeta(s)$ are exactly the points on the line $\operatorname{Re} s = \frac{1}{2}$ for which either $G(s) = 0$, or $G(s) \neq 0$ and

$$\arg H(s)G(s) \equiv \frac{\pi}{2} \pmod{\pi}. \tag{18.4}$$

The derivatives of $\xi(s)$ can be expressed by derivatives of $\zeta(s)$;

$$\xi^{(k)}(s) = \sum_{0 \leqslant j \leqslant k} \binom{k}{j} H^{(k-j)}(s) \zeta^{(j)}(s).$$

Hence (18.1) becomes

$$2G(s) = \zeta(s) + \sum_{\substack{0 \leqslant j \leqslant k \\ k \text{ odd}}} \sum g_k \binom{k}{j} \frac{H^{(k-j)}}{H}(s) \zeta^{(j)}(s). \tag{18.5}$$

We shall simplify $G(s)$ by obtaining very strong approximations in the rectangle

$$s = \sigma + it, \quad \frac{1}{3} \leqslant \sigma \leqslant A, \quad T \leqslant t \leqslant 2T, \tag{18.6}$$

with $A \geqslant 3$ and $T \geqslant 2A$.

Lemma 18.1. *For s in the rectangle* (18.6) *and $m \geqslant 0$ we have*

$$H^{(m)}(s) = H(s) \left(\frac{1}{2} \log \frac{s}{2\pi} \right)^m \left(1 + O\left(\frac{1}{T} \right) \right) \tag{18.7}$$

where the implied constant depends only on m and A.

Proof. It follows from Stirling's formula. □

LEMMA 18.2. *For s in the rectangle* (18.6) *we have*

$$\zeta^{(j)}(s) = \sum_{l \leqslant T} (-\log l)^j l^{-s} + O(T^{-\frac{1}{4}}), \tag{18.8}$$

where the implied constant depends only on j and A.

PROOF. It follows from (6.3) and (2.18). □

Lemmas 18.1 and 18.2 yield

$$H(s)^{-1}\xi^{(k)}(s) = \sum_{l \leqslant T} \left(\frac{1}{2}\log\frac{s}{2\pi} - \log l\right)^k l^{-s} + O(T^{-\frac{1}{4}}). \tag{18.9}$$

Inserting (18.9) into (18.1) we obtain a clear expression for Conrey's function

$$G(s) = \sum_{l \leqslant T} Q\left(\frac{\log l}{\log T} + \delta(s)\right) l^{-s} + O(T^{-\frac{1}{4}}), \tag{18.10}$$

where

$$\delta(s) = \frac{\log(2\pi T/s)}{2\log T} \ll \frac{1}{\log T} \tag{18.11}$$

for s in the rectangle (18.6), and $Q(x)$ is the polynomial

$$Q(x) = \frac{1}{2} + \frac{1}{2}\sum_{k \text{ odd}} g_k (\log T)^k \left(\frac{1}{2} - x\right)^k. \tag{18.12}$$

Since g_k are real numbers, the absence of terms for k even translates to the symmetry equation

$$Q(x) + Q(1-x) = 1. \tag{18.13}$$

Conversely, any real polynomial $Q(x)$ which satisfies (18.13) can be written in the form (18.12) with real coefficients g_k.

The small perturbation by $\delta(s)$ in (18.10) is not essential, it can be isolated by Taylor's expansion. We write (18.10) in the form

$$G(s) = L(s) + \delta(s)L_1(s) + L_2(s) + O(T^{-\frac{1}{4}}), \tag{18.14}$$

where

$$L(s) = \sum_{l \leqslant T} Q\left(\frac{\log l}{\log T}\right) l^{-s}, \tag{18.15}$$

$$L_1(s) = \sum_{l \leqslant T} Q'\left(\frac{\log l}{\log T}\right) l^{-s}, \tag{18.16}$$

$$L_2(s) = \sum_{l \leqslant T} \delta_l(s) l^{-s} \tag{18.17}$$

with

$$\delta_l(s) = Q\left(\frac{\log l}{\log T} + \delta(s)\right) - Q\left(\frac{\log l}{\log T}\right) - \delta(s)Q'\left(\frac{\log l}{\log T}\right). \tag{18.18}$$

Note that the coefficients of $L_2(s)$ are very small; indeed we have

$$\delta_l(s),\ s\delta_l'(s) \ll (\log T)^{-2}. \tag{18.19}$$

http://dx.doi.org/10.1090/ulect/062/19

CHAPTER 19

The Argument Variations

Let $T \geqslant 3$ and fix U with

$$T(\log T)^{-\frac{1}{4}} \leqslant U \leqslant T. \tag{19.1}$$

Let $N_{01}(T, U)$ denote the number of zeros of $\zeta(s)$ counted *without* multiplicity in the segment

$$\rho = \frac{1}{2} + i\gamma, \qquad T \leqslant \gamma \leqslant T + U. \tag{19.2}$$

Some of these zeros can be also the zeros of $G(s)$, which we are going to pass around. Let $\mathcal{C}$ be the segment $s = \frac{1}{2} + it$, $T \leqslant t \leqslant T + U$ with small circular dents to the right centered at the common critical zeros of $\zeta(s)$ and $G(s)$ as illustrated in Figure 19.1. We assume that the circular dents are so small that $\zeta(s)G(s)$ does not

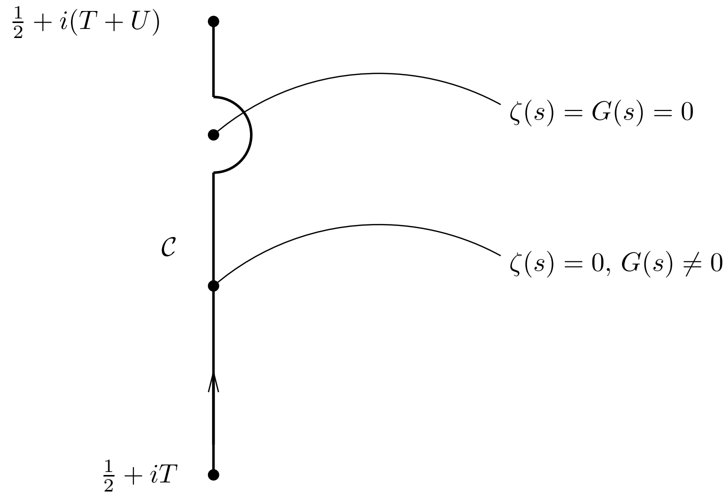

FIGURE 19.1

vanish in the dent, except at the center. The function $f(s) = H(s)G(s)$ does not vanish on $\mathcal{C}$, so as s runs through $\mathcal{C}$ starting from $\frac{1}{2} + iT$ and ending at $\frac{1}{2} + i(T+U)$, the argument of $f(s)$ changes continuously. If the change is of π along a section of $\mathcal{C}$, then this section contains a point s with

$$\arg f(s) \equiv \frac{\pi}{2} \pmod{\pi}.$$

This point is either a zero of $\zeta(s)$, or it is on the boundary of a circular dent centered at a common zero of $\zeta(s)$ and $G(s)$, in which case it also accounts for a zero of $\zeta(s)$.

Therefore we have the following lower bound for $N_{01}(T,U)$:

$$N_{01}(T,U) \geqslant \frac{1}{\pi}\Delta_{\mathcal{C}} \arg H(s)G(s) - 1, \tag{19.3}$$

where $\Delta_{\mathcal{C}} \arg$ stands for the continuous variation of the argument along $\mathcal{C}$, starting at $s = \frac{1}{2} + iT$ and ending at $s = \frac{1}{2} + i(T+U)$.

The factor $H(s)$ alone changes its argument quickly, so in reality it supplies almost all the critical zeros. Indeed, by Stirling's formula one shows that

$$\frac{1}{\pi}\Delta_{\mathcal{C}} \arg H(s) = \frac{U}{2\pi}\log\frac{T}{2\pi} + O(U) = N(T,U) + O(U), \tag{19.4}$$

where $N(T,U)$ is the number of all zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ with $0 < \beta < 1$ and $T \leqslant \gamma \leqslant T+U$ counted *with* multiplicity. Hence (19.3) becomes

$$N_{01}(T,U) \geqslant N(T,U) + \frac{1}{\pi}\Delta_{\mathcal{C}} \arg G(s) + O(U). \tag{19.5}$$

Next, let $\mathcal{R}$ be the closed rectangle that has $\mathcal{C}$ (with the small circular dents) as its left side, and that has a right side that is sufficiently far (the segment $\operatorname{Re} s = A$ with A a large constant). The orientation of the boundary $\partial\mathcal{R}$ is inherited from $\mathcal{C}$, so we move along $\partial\mathcal{R}$ clockwise (the orientation is "negative," according to the standard terminology). The argument variation of $G(s)$ along $\partial\mathcal{R}$ is then equal to

$$\Delta_{\partial\mathcal{R}} \arg G(s) = -2\pi N_G(\mathcal{R}) \tag{19.6}$$

where $N_G(\mathcal{R})$ denotes the number of zeros of $G(s)$ inside $\mathcal{R}$ counted *with* multiplicity (we assume that $G(s)$ does not vanish on $\partial\mathcal{R}$).

We shall show that the variation of the argument of $G(s)$ along the horizontal sides of $\mathcal{R}$ and over the vertical side of $\mathcal{R}$ on the far right is small. The method is similar to that applied to $\zeta(s)$ in Section 9, but there are some technical differences because $G(s)$ has no Euler product.

First we estimate the argument variation along the horizontal side $\mathcal{C}_1 = \{s = \sigma + iT\,;\ \frac{1}{2} \leqslant \sigma \leqslant A\}$. We have

$$\Delta_{\mathcal{C}_1} \arg G(s) = \int_{\frac{1}{2}+iT}^{A+iT} \operatorname{Im}\frac{G'}{G}(s)\,\mathrm{d}s. \tag{19.7}$$

It is clear by (18.1) that $(s-1)G(s)$ is entire of order one. Actually, it satisfies

$$(s-1)G(s) \ll (|s|+3)^{\mu|s|} \qquad \text{for all } s \in \mathbb{C},$$

where μ is a positive constant.

Put $s_0 = A + iT$ and $f(z) = (z + s_0 - 1)G(z + s_0)$. By the above estimate, it follows that $f(z)$ satisfies (8.21) for all $z \in \mathbb{C}$ with some constants $b, c \ll \log T$.

Now we assume that the polynomial $Q(x)$ satisfies

$$Q(0) = 1. \tag{19.8}$$

Then (18.14) yields

$$G(s_0) = 1 + O\Bigg(\sum_{2\leqslant l\leqslant T} l^{-A}\Bigg) + O(1/\log T) > \frac{1}{2}$$

if A is sufficiently large. Hence $f(0) = (s_0-1)G(s_0)$ satisfies (8.20). By Theorem 8.9 we derive

$$\frac{G'}{G}(s) = \sum_{\rho}(s-\rho)^{-1} + O(\log T) \qquad \text{if } |s-s_0| \leqslant A, \tag{19.9}$$

where ρ runs over the zeros of $G(s)$ in the circle $|s-s_0| \leqslant 2A$. Introducing (19.9) into (19.7) we conclude that

$$\Delta_{\mathcal{C}_1} \arg G(s) \ll \log T, \tag{19.10}$$

because the argument variation of $s-\rho$ is at most π and the number of ρ's in (19.9) is $O(\log T)$. Similarly we estimate the argument variation along the horizontal side $\mathcal{C}_2 = \{s = \sigma + i(T+U)\,;\ \frac{1}{2} \leqslant \sigma \leqslant A\}$, getting

$$\Delta_{\mathcal{C}_2} \arg G(s) \ll \log T. \tag{19.11}$$

Now we are going to estimate the argument variation along the vertical side $\mathcal{C}_3 = \{s = A + it\,;\ T \leqslant t \leqslant T+U\}$. For $s \in \mathcal{C}_3$ we have

$$G(s) = L(s) + O(1/\log T),$$

where $L(s)$ is the Dirichlet polynomial (18.15). Moreover, we have

$$G'(s) = L'(s) + O(1/\log T).$$

Since $L(s) \asymp 1$ we get (see the arguments following (13.8))

$$\frac{G'}{G}(s) = \frac{L'}{L}(s) + O(1/\log T) = \sum_{q=2}^{\infty} c_q q^{-s} + O(1/\log T)$$

with $c_q \ll q$. This gives

$$\begin{aligned}\Delta_{\mathcal{C}_3} \arg G(s) &= \int_{A+iT}^{A+i(T+U)} \operatorname{Im} \frac{G'}{G}(s)\,\mathrm{d}s \\ &\ll \sum_{q=2}^{\infty} |c_q|(\log q)^{-1} q^{-A} + U/\log T.\end{aligned}$$

Hence

$$\Delta_{\mathcal{C}_3} \arg G(s) \ll U/\log T. \tag{19.12}$$

Subtracting the contributions (19.10), (19.11), and (19.12) from (19.6) we get

$$\Delta_{\mathcal{C}} \arg G(s) = -2\pi N_G(\mathcal{R}) + O(U/\log T). \tag{19.13}$$

Finally, inserting this into (19.5) we get

$$N_{01}(T,U) \geqslant N(T,U) - 2N_G(\mathcal{R}) + O(U). \tag{19.14}$$

http://dx.doi.org/10.1090/ulect/062/20

CHAPTER 20

Attaching a Mollifier

Having performed the argument variations, we arrived in (19.14) to a problem of counting zeros, not the critical zeros of $\zeta(s)$ but instead the zeros of $G(s)$ in the rectangle $\mathcal{R}$. Moreover, we need an upper bound for $N_G(\mathcal{R})$ to get a lower bound for $N_{01}(T,U)$. The new task is easier, because it can be reduced to estimates for the relevant analytic functions. However, one cannot guarantee success; if the bound for $N_G(\mathcal{R})$ exceeds one half of $N(T,U)$, we end up with a negative lower bound for $N_{01}(T,U)$. Therefore our treatment of $N_G(\mathcal{R})$ must not be wasteful.

Clearly $N_G(\mathcal{R})$ can only increase if we replace $G(s)$ by

$$F(s) = G(s)M(s) \tag{20.1}$$

where $M(s)$ is any regular function in $\mathcal{R}$. This extra factor $M(s)$ may add zeros, but hopefully not too many. On the other hand, $M(s)$ is designed to dampen the extra large values of $G(s)$ so the product $F(s) = G(s)M(s)$ is expected to be smaller than $G(s)$ and our intent to use classical methods of counting zeros by estimating a contour integral becomes more promising. Naturally $M(s)$ is called "mollifier." We shall consider a specific choice of $M(s)$ which will be given by a Dirichlet polynomial

$$M(s) = \sum_{m \leqslant X} c(m) m^{-s} \tag{20.2}$$

with coefficients

$$c(m) \ll m. \tag{20.3}$$

An important condition is that

$$c(1) = 1. \tag{20.4}$$

This implies that $\log M(s)$ is given by an absolutely convergent Dirichlet series

$$\log M(s) = \sum_{q=2}^{\infty} \lambda(q) q^{-s}, \qquad \text{if } \operatorname{Re} s \geqslant A, \tag{20.5}$$

where A is a large constant (see the arguments in Section 1 of Chapter 13). We have

$$N_G(\mathcal{R}) \leqslant N_F(\mathcal{R}). \tag{20.6}$$

Next we expand $\mathcal{R}$ to a slightly larger rectangle $\mathcal{D}$ so the zeros of $F(s)$ in $\mathcal{R}$ have an ample distance from the left side of $\mathcal{D}$. Specifically we move the left side $\mathcal{C}$ of $\mathcal{R}$ to the segment $\mathcal{C}_a = \{s = a + it\,;\, T \leqslant t \leqslant T+U\}$ with $\frac{1}{3} \leqslant a < \frac{1}{2}$. Then we have

$$n(T,U) + N_F(\mathcal{R}) \leqslant \frac{1}{\frac{1}{2} - a} \sum_{\rho \in \mathcal{D}} \operatorname{dist}(\rho), \tag{20.7}$$

where ρ runs over the zeros of $F(s)$ in $\mathcal{D}$ *with* multiplicity and $\mathrm{dist}(\rho)$ denotes the distance of ρ to the left side of $\mathcal{D}$. Here, $n(T,U)$ is the number of common zeros of $\zeta(s)$ and $G(s)$ on the segment $s = \frac{1}{2} + it$, $T \leqslant t \leqslant T+U$ (the centers of the small circular dents in $\mathcal{C}$) counted *with* the multiplicity as it appears in $G(s)$. By (20.7), (20.6), and (19.14), we get

$$N_{01}(T,U) - 2n(T,U) \geqslant N(T,U) - \frac{2}{\frac{1}{2}-a}\sum_{\rho\in\mathcal{D}} \mathrm{dist}(\rho) + O(U). \tag{20.8}$$

Let $N_{00}(T,U)$ denote the number of zeros of $\zeta(s)$ on the segment $s = \frac{1}{2} + it$, $T \leqslant t \leqslant T+U$, which are not zeros of $G(s)$, counted *without* multiplicity. Therefore $N_{00}(T,U) \geqslant N_{01}(T,U) - n(T,U)$, and (20.8) yields

$$N_{00}(T,U) \geqslant N(T,U) - \frac{2}{\frac{1}{2}-a}\sum_{\rho\in\mathcal{D}} \mathrm{dist}(\rho) + O(U). \tag{20.9}$$

Recall that $N(T,U)$ is the number of all zeros of $\zeta(s)$ in the box $0 < \sigma < 1$, $T \leqslant t \leqslant T+U$, counted with multiplicity;

$$N(T,U) = \frac{U}{2\pi}\log T + O(U). \tag{20.10}$$

It remains to estimate the sum of $\mathrm{dist}(\rho)$ over ρ in $\mathcal{D}$.

http://dx.doi.org/10.1090/ulect/062/21

CHAPTER 21

The Littlewood Lemma

We evaluate the sum of $\operatorname{dist}(\rho)$ in (20.9) by the following formula

LEMMA 21.1 (J. E. Littlewood, 1924). *Let $F(s)$ be a holomorphic function in a rectangle $\mathcal{D}$ with sides parallel to the axis, not vanishing on the sides. Then*

$$\sum_{\rho\in\mathcal{D}} \operatorname{dist}(\rho) = \frac{1}{2\pi i}\int_{\partial\mathcal{D}} \log F(s)\,\mathrm{d}s. \tag{21.1}$$

Here the zeros ρ of $F(s)$ are counted *with* multiplicity and $\log F(s)$ is a continuous branch of the logarithm on $\partial\mathcal{D}$, that is

$$\log F(s) = \log|F(s)| + i\arg F(s), \tag{21.2}$$

where the argument is defined by continuous variation starting with any fixed value at a chosen point on $\partial\mathcal{D}$ and going through $\partial\mathcal{D}$ in the negative direction (clockwise).

PROOF. We can assume that all the zeros $\rho = \beta+i\gamma$ of $F(s)$ in $\mathcal{D}$ have different heights γ. This can be accomplished by small perturbations of the zeros and by the continuity of both sides of (21.1) with respect to such perturbations. Removing from $\mathcal{D}$ the horizontal segments which connect the zeros with the left side of $\mathcal{D}$ we get a simply connected domain $\mathcal{D}'$ free of zeros of $F(s)$. Therefore there exists a holomorphic branch of $\log F(s)$ in $\mathcal{D}'$. Remove from $\mathcal{D}$ rounded strips $\mathcal{D}_\rho$ of width ε centered along the removed segments, as in Figure 21.1. Let $\mathcal{C}$ be the boundary

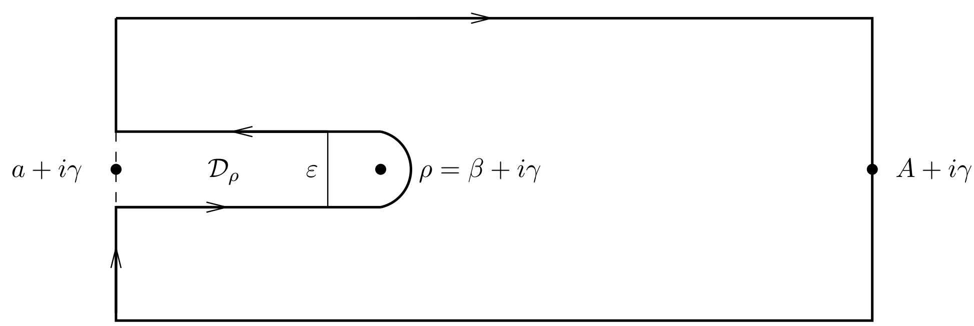

FIGURE 21.1

of $\mathcal{D}\setminus\bigcup_\rho \mathcal{D}_\rho$, negatively oriented. By Cauchy's Theorem

$$\frac{1}{2\pi i}\int_{\mathcal{C}} \log F(s)\,\mathrm{d}s = 0.$$

We have $\mathcal{C} = \partial D \cup \bigcup_\rho \partial\mathcal{D}_\rho$ up to short segments at the left ends of $\mathcal{D}_\rho$. Hence

$$\frac{1}{2\pi i}\int_{\partial\mathcal{D}} \log F(s)\,\mathrm{d}s + \sum_\rho \frac{1}{2\pi i}\int_{\partial\mathcal{D}_\rho} \log F(s)\,\mathrm{d}s = O(\varepsilon).$$

For every $\rho = \beta + i\gamma$ we have

$$\frac{1}{2\pi i}\int_{\partial\mathcal{D}_\rho} \log F(s)\,\mathrm{d}s = \frac{1}{2\pi i}\int_a^{\beta-\varepsilon} \big(\log F(\sigma+i\gamma-i\varepsilon) - \log F(\sigma+i\gamma+i\varepsilon)\big)\,\mathrm{d}\sigma$$
$$+ O\left(\varepsilon\log\frac{1}{\varepsilon}\right),$$

where the error term is obtained by estimating trivially the contribution of the integral along the circular ending of $\mathcal{D}_\rho$.

Writing $F(s) = (s-\rho)F_1(s)$ in the strip $\mathcal{D}_\rho$, where $F_1(s)$ does not vanish, we get

$$\begin{aligned}\log F(\sigma+i\gamma-i\varepsilon) &- \log F(\sigma+i\gamma+i\varepsilon)\\ &= \log(\sigma-\beta-i\varepsilon) - \log(\sigma-\beta+i\varepsilon) + O(\varepsilon)\\ &= i\arg(\sigma-\beta-i\varepsilon) - i\arg(\sigma-\beta+i\varepsilon) + O(\varepsilon)\\ &= -2\pi i + O\big(\varepsilon/(\beta-\sigma)\big).\end{aligned}$$

Integrating this over σ with $a \leqslant \sigma \leqslant \beta - \varepsilon$ we get

$$\frac{1}{2\pi i}\int_{\partial\mathcal{D}_\rho} \log F(s)\,\mathrm{d}s = a - \beta + O\left(\varepsilon\log\frac{1}{\varepsilon}\right),$$

and

$$\frac{1}{2\pi i}\int_{\partial\mathcal{D}} \log F(s)\,\mathrm{d}s = \sum_\rho(\beta-a) + O\left(\varepsilon\log\frac{1}{\varepsilon}\right).$$

Finally, letting $\varepsilon \to 0$ we complete the proof of (21.1). □

Note that the left side of (21.1) can be written as

$$\sum_{\rho\in\mathcal{D}} \operatorname{dist}(\rho) = \int_a^A n(\sigma)\,\mathrm{d}\sigma \tag{21.3}$$

where $n(\sigma)$ denotes the number of zeros $\rho = \beta + i\gamma \in \mathcal{D}$ with $\beta \geqslant \sigma$.

Assuming $\mathcal{D} = \{s = \sigma + it\,;\, a \leqslant \sigma \leqslant A,\, T \leqslant t \leqslant T+U\}$ we can write (21.1) as follows:

$$\sum_{\rho\in\mathcal{D}} \operatorname{dist}(\rho) = \frac{1}{2\pi}\int_T^{T+U} \big(\log F(a+it) - \log F(A+it)\big)\,\mathrm{d}t$$
$$+ \frac{1}{2\pi i}\int_a^A \Big(\log F\big(\sigma+i(T+U)\big) - \log F(\sigma+iT)\Big)\,\mathrm{d}\sigma.$$

This is a real number, so we have

$$\sum_{\rho\in\mathcal{D}} \operatorname{dist}(\rho) = \frac{1}{2\pi}\int_T^{T+U} \Big(\log\big|F(a+it)\big| - \log\big|F(A+it)\big|\Big)\,\mathrm{d}t$$
$$+ \frac{1}{2\pi}\int_a^A \Big(\arg F\big(\sigma+i(T+U)\big) - \arg F(\sigma+iT)\Big)\,\mathrm{d}\sigma. \tag{21.4}$$

http://dx.doi.org/10.1090/ulect/062/22

CHAPTER 22

The Principal Inequality

It remains to evaluate the sum of $\mathrm{dist}(\rho)$ in the inequality (20.9). To this end we apply the Littlewood formula (21.4) for the product $F(s) = G(s)M(s)$. We are going to show that the integrals of $\log|F(A+it)|$, $\arg F(\sigma+iT)$, $\arg(F(\sigma+i(T+U)))$ yield small contributions, so that (21.4) simplifies to

$$\sum_{\rho\in\mathcal{D}} \mathrm{dist}(\rho) = \frac{1}{2\pi}\int_T^{T+U} \log\bigl|F(a+it)\bigr|\,\mathrm{d}t + O(U/\log T). \tag{22.1}$$

Actually, one can establish (22.1) with the much better error term $O(\log T)$, but we do not need a result stronger than (22.1).

First we deal with the horizontal integrations of $\arg F(\sigma+iT)$ and $\arg F(\sigma + i(T+U))$. We have

$$\arg F(A+iT) - \arg F(\sigma+iT) = \int_{\sigma+iT}^{A+iT} \mathrm{Im}\,\frac{F'}{F}(s)\,\mathrm{d}s.$$

Integrating in σ over the segment $a \leqslant \sigma \leqslant A$, we get

$$(A-a)\arg F(A+iT) - \int_a^A \arg F(\sigma+iT)\,\mathrm{d}\sigma = \int_a^A (\sigma-a)\,\mathrm{Im}\,\frac{F'}{F}(s)\,\mathrm{d}s.$$

We have already handled a similar integral in (19.7) for the function $G(s)$ by using the expansion (19.9). However, the mollifier $M(s)$ also satisfies this expansion; see Corollary 8.10. Therefore, by the same arguments, we get

$$(A-a)\arg F(A+iT) - \int_a^A \arg F(\sigma+iT)\,\mathrm{d}\sigma = O(\log T).$$

Replacing T by $T+U$ and subtracting we get

$$\int_a^A \Bigl(\arg F\bigl(\sigma+i(T+U)\bigr) - \arg F(\sigma+iT)\Bigr)\,\mathrm{d}\sigma = (a-A)\Delta_{\mathcal{C}_3}\arg F(s) + O(\log T).$$

We have already handled $\Delta_{\mathcal{C}_3}\arg G(s)$ getting the bound (19.12). However, the mollifier $M(s)$ also satisfies the expansion (13.8), so the same bound (19.12) holds for $\Delta_{\mathcal{C}_3}\arg F(s)$, and we get

$$\int_a^A \Bigl(\arg F\bigl(\sigma+i(T+U)\bigr) - \arg F(\sigma+iT)\Bigr)\,\mathrm{d}\sigma \ll U/\log T. \tag{22.2}$$

Next we deal with the vertical integration of

$$\log\bigl|F(A+it)\bigr| = \log\bigl|G(A+it)\bigr| + \log\bigl|M(A+it)\bigr|.$$

By the Dirichlet series expansion (20.5) we get

$$\int_T^{T+U} \log\bigl|M(A+it)\bigr|\,\mathrm{d}t = \mathrm{Re}\sum_{q=2}^{\infty}\lambda(q)q^{-A-iT}\frac{q^{-iU}-1}{i\log q}x \ll 1.$$

Moreover, $G(s) = L(s)(1+O(1/\log T))$ where $L(s)$ is the Dirichlet polynomial (18.15), so by similar arguments we get

$$\int_T^{T+U} \log\bigl|G(A+it)\bigr|\,\mathrm{d}t \ll 1 + U/\log T.$$

This completes the proof of (22.1).

Inserting (22.1) into (20.9) we obtain

$$N_{00}(T,U) \geqslant N(T,U) - \frac{1}{\pi(\frac{1}{2}-a)}\int_T^{T+U}\log\bigl|F(a+it)\bigr|\,\mathrm{d}t + O\left(\frac{U}{(\frac{1}{2}-a)\log T}\right),$$

where $F(s) = G(s)M(s)$ and $\frac{1}{3} < a < \frac{1}{2}$. We shall apply this for a close to $\frac{1}{2}$, specifically

$$a = \frac{1}{2} - \frac{R}{\log T} \tag{22.3}$$

with R a positive constant to be chosen for best results. We get

$$N_{00}(T,U) \geqslant N(T,U)\left\{1 - \frac{2}{R}l(R) + O\left(\frac{1}{\log T}\right)\right\} \tag{22.4}$$

where

$$l(R) = \frac{1}{U}\int_T^{T+U}\log\bigl|F(a+it)\bigr|\,\mathrm{d}t. \tag{22.5}$$

Next, by concavity of the logarithm, we infer that

$$l(R) \leqslant \log I(R) \tag{22.6}$$

where

$$I(R) = \frac{1}{U}\int_T^{T+U}\bigl|F(a+it)\bigr|\,\mathrm{d}t. \tag{22.7}$$

Note that $I(R)$ depends on T, U, but if T is sufficiently large we shall give a strong estimate for $I(R)$ which does not depend on T, U. For reference we conclude the results of Chapters 16-22 by the following principal inequality of the Levinson-Conrey method.

THEOREM 22.1 (Levinson-Conrey). *Let $N(T,U)$ be the number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ with $0 < \beta < 1$, $T < \gamma \leqslant T+U$, counted with multiplicity. Let $N_{00}(T,U)$ be the number of these zeros with $\beta = \frac{1}{2}$, counted without multiplicity, which are not zeros of $G(s)$ (see* (18.1) *and* (18.14)*). If T is large and $T(\log T)^{-\frac{1}{4}} \leqslant U \leqslant T$, then*

$$N_{00}(T,U) \geqslant N(T,U)\left\{1 - \frac{2}{R}\log I(R) + O\left(\frac{1}{\log T}\right)\right\} \tag{22.8}$$

where $I(R)$ is the absolute mean value (22.7) *of $F(s) = G(s)M(s)$ on $\mathrm{Re}\, s = a$, and $M(s)$ is any Dirichlet polynomial with coefficients $c(1) = 1$, $|c(m)| \leqslant 1$* for $m \leqslant T$.

http://dx.doi.org/10.1090/ulect/062/23

CHAPTER 23

Positive Proportion of the Critical Zeros

To complete a lower bound for $N_{00}(T,U)$ we need an upper bound for the integral $I(R)$. Estimating $I(R)$ is a hard part of the Levinson-Conrey method. Since this problem is interesting on its own, we postpone the work to the last three chapters, where we establish results more general than needed. For now we assume that we have a bound

$$\frac{1}{U}\int_T^{T+U} \big|F(a+it)\big|\,\mathrm{d}t \leqslant c(R)+o(1), \tag{23.1}$$

where $c(R)>1$. Hence (22.8) yields

$$N_{00}(T,U) \geqslant \big(\kappa+o(1)\big)N(T,U) \tag{23.2}$$

with

$$\kappa = 1-\frac{2}{R}\log c(R). \tag{23.3}$$

The original choice of $G(s)$ by Levinson is given by (in the setting (18.1))

$$2H(s)G(s) = \xi(s)+\frac{2}{\log T}\xi'(s) \tag{23.4}$$

This can be approximated by a nice Dirichlet polynomial;

$$\begin{aligned} G(s) &= \left(\frac{1}{2}+\frac{H'}{H}(s)(\log T)^{-1}\right)\zeta(s)+\zeta'(s)(\log T)^{-1} \\ &= \big(1-\delta(s)\big)\zeta(s)+\zeta'(s)(\log T)^{-1}+O(T^{-\frac{1}{2}}) \\ &= \sum_{l\leqslant T} Q\left(\frac{\log l}{\log T}\right)l^{-s}+O\left(\frac{|\zeta(s)|}{\log T}+\frac{1}{\sqrt{T}}\right) \end{aligned} \tag{23.5}$$

where $Q(x)$ is the linear polynomial

$$Q(x) = 1-x. \tag{23.6}$$

For the mollifier, Levinson takes (recall (22.3))

$$M(s) = \sum_{m\leqslant T^{\frac{1}{2}}} \mu(m)P\left(\frac{\log m}{\log T}\right)m^{a-\frac{1}{2}-s}, \tag{23.7}$$

where $P(y)$ is the linear polynomial

$$P(y) = 1-2y. \tag{23.8}$$

For this particular choice of $F(s)=G(s)M(s)$, the results in Chapter 25 yield the asymptotic formula

$$\frac{1}{U}\int_T^{T+U} \big|F(a+it)\big|^2\,\mathrm{d}t \sim C(R) \tag{23.9}$$

as $T \to \infty$, where

$$C(R) = \frac{e^{2R} - 1 - 2R}{2R^3} + \frac{e^{2R} - 1}{24R} + \frac{1}{R} - \frac{R}{12} + \frac{7}{12}. \tag{23.10}$$

Of course (23.9) agrees with the result of Levinson.

By Cauchy-Schwarz inequality we get (23.1) with

$$c(R) = C(R)^{\frac{1}{2}}. \tag{23.11}$$

Hence (23.2) holds with

$$\kappa = 1 - \frac{1}{R} \log C(R). \tag{23.12}$$

Levinson computes this for $R = 1.3$, getting $C(R) = 2.32\ldots$, and

$$\kappa = 0.34\ldots \tag{23.13}$$

Note that in the case of $G(s)$ given as a linear combination of $\zeta(s)$ and $\zeta'(s)$ we have $N_{00}(T, U)$ which counts only the simple zeros of $\zeta(s)$. This feature of Levinson's method was observed by A. Selberg and D. R. Heath-Brown [**HB79**] Indeed, if $s = \rho$ is a zero of order $\geqslant 2$, then it is a zero of $G(s)$ which is not counted by $N_{00}(T, U)$. Thus, we conclude the following

THEOREM 23.1. *At least 34% of all the zeros of $\zeta(s)$ are simple and are on the critical line.*

http://dx.doi.org/10.1090/ulect/062/24

CHAPTER 24

The First Moment of Dirichlet Polynomials

We consider the Dirichlet polynomials

$$A(s)=\sum_{n\leqslant N}a_nn^{-s} \tag{24.1}$$

which are relevant to the integral (22.7) in the Levinson-Conrey principal inequality (22.8). The coefficients a_n will be specialized gradually in later sections. Here we give quick results which will be used for estimating lower order contributions.

We want to estimate the mean-value of $|A(s)|$ over a segment of the critical line:

$$I(T,U)=\frac{1}{U}\int_T^{T+U}\left|A\left(\frac{1}{2}+it\right)\right|\,\mathrm{d}t, \tag{24.2}$$

for polynomials $A(s)$ which factor into two polynomials, each of length strictly smaller than T;

$$A(s)=\left(\sum_{l<T}b_ll^{-s}\right)\left(\sum_{m<T}c_mm^{-s}\right). \tag{24.3}$$

By the Cauchy-Schwarz inequality and Theorem 13.1 we get

$$U^2I(T,U)^2\ll\left\{U\sum|b_l|^2l^{-1}+\left(\sum|b_l|^2l^{-1}\right)^{\frac{1}{2}}\left(\sum|b_l|^2l\right)^{\frac{1}{2}}\right\}$$
$$\left\{U\sum|c_m|^2m^{-1}+\left(\sum|c_m|^2m^{-1}\right)^{\frac{1}{2}}\left(\sum|c_m|^2m\right)^{\frac{1}{2}}\right\}.$$

Suppose the coefficients b_l satisfy the bound

$$|b_l|\leqslant\left(\frac{\log l}{\log T}\right)^r, \tag{24.4}$$

with some $r\geqslant 0$. Then

$$\sum_{l<T}|b_l|^2l^{-1}\ll\frac{\log T}{r+1},\qquad\sum_{l<T}|b_l|^2l\ll T^2$$

where the implied constants are absolute. Moreover, suppose that the coefficients c_m satisfy the bound

$$|c_m|\leqslant 1. \tag{24.5}$$

By the above estimates we get

$$U^2I(T,U)^2\ll\left\{U\frac{\log T}{r+1}+T\left(\frac{\log T}{r+1}\right)^{\frac{1}{2}}\right\}\left\{U\log T+T(\log T)^{\frac{1}{2}}\right\}.$$

Hence

LEMMA 24.1. *If* $T(\log T)^{-\frac{1}{2}} \leqslant U \leqslant T$ *and* b_l, c_m *satisfy* (24.4) *and* (24.5), *then*

$$I(T, U) \ll (r+1)^{-\frac{1}{4}} \log T, \tag{24.6}$$

where the implied constant is absolute.

http://dx.doi.org/10.1090/ulect/062/25

CHAPTER 25

The Second Moment of Dirichlet Polynomials

We need an upper bound for the integral (24.2) much better than (24.6). By the Cauchy-Schwarz inequality,

$$I(T,U)^2 \leqslant I_2(T,U), \tag{25.1}$$

where

$$I_2(T,U) = \frac{1}{U}\int_T^{T+U} \left|A\left(\frac{1}{2}+it\right)\right|^2 \mathrm{d}t. \tag{25.2}$$

Our goal is to establish an asymptotic formula for $I_2(T,U)$. To this end it suffices to establish separately an upper bound and a lower bound which are asymptotically equal. We shall show only the upper bound, because the arguments for the lower bound are almost identical. To be fair, we only need the upper bound.

We begin by smoothing the integration as described in Appendix A. Let $\Phi(t)$ be the function given by (A.12). Recall that $\Phi(t)$ majorizes the characteristic function of the interval $[T, T+U]$, and its Fourier transform $\hat{\Phi}(y)$ satisfies

$$\hat{\Phi}(0) = U + V \tag{25.3}$$

$$\hat{\Phi}(y) \ll U \exp(-2\sqrt{\pi|y|V}) \tag{25.4}$$

$$\hat{\Phi}'(y) \ll TU \exp(-2\sqrt{\pi|y|V}), \tag{25.5}$$

where V is at our disposal, subject to $0 < V < U < T$. We get

$$UI_2(T,U) \leqslant \int \Phi(t)\left|A\left(\frac{1}{2}+it\right)\right|^2 \mathrm{d}t = I(\Phi), \tag{25.6}$$

say. Opening the square and integrating we get

$$I(\Phi) = \sum_{n_1}\sum_{n_2} \frac{a_{n_1}\overline{a_{n_2}}}{\sqrt{n_1 n_2}}\hat{\Phi}\left(\frac{1}{2\pi}\log\frac{n_1}{n_2}\right). \tag{25.7}$$

Here we pull out the contribution of the diagonal terms $n_1 = n_2$;

$$I_0(\Phi) = \hat{\Phi}(0)\sum_n |a_n|^2 n^{-1}, \tag{25.8}$$

because they are distinctly different from the other terms $n_1 \neq n_2$. We shall evaluate $I_0(\Phi)$ asymptotically in Chapter 26. Then in Chapters 27-28 we shall show that the remaining contribution

$$I^*(\Phi) = \sum\sum_{n_1 \neq n_2} \frac{a_{n_1}\overline{a_{n_2}}}{\sqrt{n_1 n_2}}\hat{\Phi}\left(\frac{1}{2\pi}\log\frac{n_1}{n_2}\right) \tag{25.9}$$

is negligible.

The bounds (25.4) and (25.5) for the Fourier transform are pretty good, nevertheless they are not sufficient for estimating $I^*(\Phi)$. Some arithmetical properties of the coefficients a_n will be necessary.

http://dx.doi.org/10.1090/ulect/062/26

CHAPTER 26

The Diagonal Terms

We assume that a_n are given in the following form

$$a_n = \sum_{\substack{lm=n \\ l,m<T}} \mu(m) Z(\gamma_m, \gamma_l), \tag{26.1}$$

where we put

$$\gamma_c = \frac{\log c}{\log T} \tag{26.2}$$

for any $c \geqslant 1$, and $T \geqslant 3$ is fixed. Moreover $Z(x, y)$ is a function on the unit square $0 \leqslant x, y < 1$ such that

$$Z(x, y) \text{ is continuous} \tag{26.3}$$

$$|Z(x, y)| \leqslant (1 - x)(1 - y). \tag{26.4}$$

We also assume that $Z(x, y)$ is of $\mathcal{C}^3$-class with bounded partial derivatives, except for x or y in a finite set, say $\mathcal{W}$. Specifically,

$$\left| \frac{\partial^{\alpha+\beta} Z}{\partial x^\alpha \partial y^\beta} \right| \leqslant 1, \quad \text{if } x \notin \mathcal{W},\ y \notin \mathcal{W}, \tag{26.5}$$

for $0 < \alpha + \beta \leqslant 3$. In addition, $\partial Z / \partial x$ is continuous in y for every $x \notin \mathcal{W}$ and $\partial Z / \partial y$ is continuous in x for every $y \notin \mathcal{W}$.

Sequences of the convolution type (26.1) also appear in other areas of number theory, particularly in sieve theory, so we are going to evaluate the sum

$$E = \sum_n a_n^2 n^{-1} \tag{26.6}$$

in more generality than what is required for evaluating the second moment $I_2(T, U)$.

Proposition 26.1. *Suppose that $Z(x, y)$ satisfies* (26.3), (26.4), *and* (26.5). *Then*

$$E = \mathcal{E} + O\big((\log\log T)^{15} / \log T\big), \tag{26.7}$$

where

$$\mathcal{E} = \int_0^1 \int_0^1 \mathcal{D} Z(x, y)\, \mathrm{d}x\, \mathrm{d}y, \tag{26.8}$$

and $\mathcal{D}$ is the differential operator given by

$$\mathcal{D} Z = \left(\frac{\partial Z}{\partial x} \right)^2 + 2Z \frac{\partial^2 Z}{\partial x \partial y} + \left(\frac{\partial Z}{\partial y} \right)^2. \tag{26.9}$$

The implied constant in (26.7) *depends only on the number of points in $\mathcal{W}$.*

One can express the integral (26.8) in terms of partial derivatives of the first order.

PROPOSITION 26.2. *For $Z(x,y)$ satisfying* (26.3), (26.4), *and* (26.5), *we have*

$$\mathcal{E} = Z(0,0)^2 + \int_0^1\!\!\int_0^1 \big(\partial Z(x,y)\big)^2 \,\mathrm{d}x\,\mathrm{d}y, \tag{26.10}$$

where

$$\partial Z = \frac{\partial Z}{\partial x} - \frac{\partial Z}{\partial y}. \tag{26.11}$$

PROOF. If $x \notin \mathcal{W}$ and $y \notin \mathcal{W}$, then we have

$$(\partial Z)^2 = \mathcal{D}Z + 2\frac{\partial Z}{\partial x}\frac{\partial Z}{\partial y} - 2Z\frac{\partial^2 Z}{\partial x \partial y}.$$

Let $y \notin \mathcal{W}$. Integrating by parts in the x-variable we get (because $\partial Z/\partial x$ is continuous in x)

$$\int_0^1 Z(x,y)\frac{\partial^2}{\partial x\partial y}Z(x,y)\,\mathrm{d}x = -Z(0,y)\frac{\partial Z}{\partial y}(0,y) - \int_0^1 \frac{\partial Z}{\partial x}\frac{\partial Z}{\partial y}(x,y)\,\mathrm{d}x.$$

Then integrating this in the y-variable we find that (because the set of singular points (x,y) has measure zero)

$$\begin{aligned}\int_0^1\!\!\int_0^1 \big((\partial Z)^2 - \mathcal{D}Z\big)\,\mathrm{d}x\,\mathrm{d}y &= 2\int_0^1 Z(0,y)\frac{\partial Z}{\partial y}(0,y)\,\mathrm{d}y\\ &= \int_0^1 \frac{\partial}{\partial y}\big(Z(0,y)^2\big)\,\mathrm{d}y = -Z(0,0)^2,\end{aligned}$$

which completes the proof. □

Note that if $Z(x,y)$ is real, then $\mathcal{E} \geqslant Z(0,0)^2$. This is not surprising, because $a_n^2 \geqslant 0$ and $a_1 = Z(0,0)$.

One can easily derive from Propositions 26.1 and 26.2 a more general result for sums of type

$$\sum_n a_n b_n n^{-1},$$

where $B = (b_n)$ is another sequence of type (26.1), say

$$b_n = \sum_{\substack{lm=n\\ l,m<T}} \mu(m)V(x,y). \tag{26.12}$$

THEOREM 26.3. *Suppose $Z(x,y)$ and $V(x,y)$ satisfy* (26.3), (26.4), *and* (26.5). *Then*

$$\sum_n a_n b_n n^{-1} = Z(0,0)V(0,0) + \int_0^1\!\!\int_0^1 \partial Z(x,y)\partial V(x,y)\,\mathrm{d}x\,\mathrm{d}y + O\big((\log\log T)^{15}/\log T\big), \tag{26.13}$$

where the implied constant depends only on $|\mathcal{W}|$.

PROOF. It follows from the identity $4ab = (a+b)^2 - (a-b)^2$ and applying the results for the sequences $a_n + b_n$, $a_n - b_n$ with test functions $Z+V$, $Z-V$, respectively. □

The operator ∂ has many cute properties. First of all it is linear and it acts as differentiation;

$$\partial VZ = V\partial Z + Z\partial V.$$

If $V(x,y) = V(x+y)$, then $\partial V = 0$. Therefore ∂ commutes with multiplication by functions which depend only on $x+y$. Hence, Theorem 26.3 implies

COROLLARY 26.4. *Let $F(z)$ be a function of $\mathcal{C}^3$-class on $0 \leqslant z \leqslant 1$. Suppose $Z(x,y)$ and $V(x,y)$ satisfy* (26.3), (26.4), *and* (26.5). *Then*

$$\sum_n a_n b_n F(\gamma_n) n^{-1} = F(0)Z(0,0)V(0,0) + \int_0^1\int_0^1 F(x+y)\partial Z(x,y)\partial V(x,y)\,\mathrm{d}x\,\mathrm{d}y + O\big((\log\log T)^{15}/\log T\big), \tag{26.14}$$

where the implied constant depends only on $|\mathcal{W}|$ and F.

Now we proceed to the proof of Proposition 26.1. Opening a_n^2 we arrange the sum (26.6) as follows:

$$E = \sum_{\substack{lm=l'm' \\ l,m,l',m'<T}} \mu(m)\mu(m')(lm)^{-1} Z(\gamma_m,\gamma_l)Z(\gamma_{m'},\gamma_{l'}).$$

The equation $lm = l'm'$ implies $m = ad$, $m' = bd$ with $(a,b) = 1$ and $l = bc$, $l' = ac$. Hence E becomes

$$E = \sum\sum\sum\sum_{\substack{a,b,c,d \\ \max(a,b)\max(c,d)<T}} \mu(d)\frac{\mu(abd)}{abcd} Z(\gamma_{ad},\gamma_{bc})Z(\gamma_{bd},\gamma_{ac}).$$

Next we are going to reduce the range of a,b. Choose

$$\Delta = \Delta(T) = \exp(\log\log T)^3. \tag{26.15}$$

By the Prime Number Theorem in the form (see (11.7))

$$\sum_{\Delta<a\leqslant x} \frac{\mu(a)}{a} \ll \exp(-\lambda\sqrt{\log\Delta}) \ll (\log T)^{-2012} \tag{26.16}$$

and by partial summation we can remove $a > \Delta$, $b > \Delta$ in E up to a small error term;

$$E = \sum\sum\sum\sum_{\substack{a,b,c,d \\ \max(a,b)\max(c,d)<T \\ \max(a,b)<\Delta}} \mu(d)\frac{\mu(abd)}{abcd} Z(\gamma_{ad},\gamma_{bc})Z(\gamma_{bd},\gamma_{ac}) + O\big((\log T)^{-2000}\big).$$

Now the variables

$$u = \gamma_a = \frac{\log a}{\log T} < \frac{\log\Delta}{\log T}$$
$$v = \gamma_b = \frac{\log b}{\log T} < \frac{\log\Delta}{\log T}$$

are quite small, so it is useful to apply Taylor's expansion at $(x, y) = (\gamma_d, \gamma_c)$;

$$Z(x+u, y+v)Z(x+v, y+u) = ZZ + (u+v)Z\left(\frac{\partial Z}{\partial x} + \frac{\partial Z}{\partial y}\right) + uv\mathcal{D}Z$$
$$+ \frac{u^2+v^2}{2}\left(Z\frac{\partial^2 Z}{\partial x^2} + 2\frac{\partial Z}{\partial x}\cdot\frac{\partial Z}{\partial y} + Z\frac{\partial^2 Z}{\partial y^2}\right) + O(u^3+v^3).$$

This approximation by Taylor's expansion is valid if x and y stay away from the singularity points in the set $\mathcal{W}$. Specifically, we can use this formula if x and y are distanced from $\mathcal{W}$ by at least $\log\Delta/\log T$. For $x = \gamma_d$ and $y = \gamma_c$ this means that c and d are not in the forbidden segments $[\Delta^{-1}T^w, \Delta T^w]$ for every $w \in \mathcal{W}$. However, for technical reasons, we want to avoid the extended segments $[\Delta^{-2}T^w, \Delta T^w]$. We call (c, d) an *exceptional pair* if either c or d is in some $[\Delta^{-2}T^w, \Delta T^w]$. The remaining pairs (c, d) are called *regular*. Accordingly, we split the sum E into

$$E = E_0 + E_{11} + O\big((\log T)^{-2000}\big), \tag{26.17}$$

where E_0 runs over the exceptional pairs and E_{11} runs over the regular pairs.

First we are going to evaluate

$$E_{11} = \sum_{\substack{\max(a,b)\max(c,d)<T \\ \max(a,b)<\Delta,\ (c,d) \text{ regular}}}\sum\sum\sum \mu(d)\frac{\mu(abd)}{abcd}Z(\gamma_{ad}, \gamma_{bc})Z(\gamma_{bd}, \gamma_{ac}). \tag{26.18}$$

We begin by applying the above Taylor expansion at $(x, y) = (\gamma_d, \gamma_c)$ with $(u, v) = (\gamma_a, \gamma_b)$. The contribution of the error term $O(u^3 + v^3)$ to E_{11} is estimated by $O((\log\Delta)^5/\log T)$. Next we can extend the range of c, d from $\max(c, d) < T/\max(a, b)$ to $\max(c, d) < T$, up to the same error term $O((\log\Delta)^5/\log T)$. Indeed, in the added range we have either $T\Delta^{-1} < c < T$ or $T\Delta^{-1} < d < T$. In both cases

$$Z(\gamma_d, \gamma_c) \ll \frac{\log\Delta}{\log T} \tag{26.19}$$

by the condition (26.6) (this condition cannot be dispensed!). Hence the contribution from the added range is trivially bounded by

$$\sum_{a,b<\Delta}\sum \frac{1}{ab}(\log\Delta)(\log T)\left(\frac{\log\Delta}{\log T}\right)^2 \ll \frac{(\log\Delta)^5}{\log T}. \tag{26.20}$$

Now we are left with

$$E_{11} = \sum_{\substack{a,b<\Delta,\ c,d<T \\ (c,d) \text{ regular}}}\sum\sum\sum \mu(d)\frac{\mu(abd)}{abcd}\left\{Z^2(\gamma_d, \gamma_c) + \gamma_{ab}\left(\frac{\partial Z}{\partial x} + \frac{\partial Z}{\partial y}\right)(\gamma_d, \gamma_c)\right.$$
$$\left. + \gamma_a\gamma_b\mathcal{D}Z(\gamma_d, \gamma_c) + \gamma_a^2(\cdots) + \gamma_b^2(\cdots)\right\} + O\big((\log\Delta)^5/\log T\big).$$

Here the key observation is that the variables γ_a, γ_b are separated from γ_d, γ_c. Using the estimate

$$\sum_{\substack{a<\Delta \\ (a,k)=1}} \frac{\mu(a)}{a} \ll \exp(-\lambda\sqrt{\log\Delta}), \tag{26.21}$$

which holds uniformly in $k < T$ with an absolute constant $\lambda > 0$ (it can be derived from (11.7)), we are left with

$$E_{11} = \sum_{\substack{a,b<\Delta,\ c,d<T \\ (c,d) \text{ regular}}}\sum\sum\sum \mu(d)\frac{\mu(abd)}{abcd}\gamma_a\gamma_b \mathcal{D}Z(\gamma_d,\gamma_c) + O\left(\frac{(\log\Delta)^5}{\log T}\right).$$

Next we reduce the range $c, d < T$ to

$$ad < T, \qquad bc < T$$

by applying the same arguments which allowed us to relax the condition $\max(a,b)\max(c,d) < T$. Having done this we now apply the approximation

$$\mathcal{D}Z(\gamma_d,\gamma_c) = \mathcal{D}Z(\gamma_{ad},\gamma_{bc}) + O\left(\frac{\log\Delta}{\log T}\right)$$

which follows by the mean-value theorem for $\mathcal{D}Z(x,y)$. We get

$$E_{11} = \sum_{\substack{a,b<\Delta,\ ad,bc<T \\ (c,d) \text{ regular}}}\sum\sum\sum \mu(d)\frac{\mu(abd)}{abcd}\gamma_a\gamma_b \mathcal{D}Z(\gamma_{ad},\gamma_{bc}) + O\left(\frac{(\log\Delta)^5}{\log T}\right).$$

Note that for (c,d) regular, we get $l = bc$, $m = ad$ off the segments $[\Delta^{-1}T^w, \Delta T^w]$ for every $w \in \mathcal{W}$. Therefore we call such pairs (l,m) regular. Now, we can replace the condition that (c,d) is regular by the condition that (l,m) is regular up to the error term (26.20). Next we remove the conditions $a, b < \Delta$ by the same arguments which allowed us to install them using the PNT. Considering $l = bc$, $m = ad$ as single variables, we write the result in the following form

$$E_{11} = (\log T)^{-2} \sum_{\substack{l,m<T \\ (l,m) \text{ regular}}}\sum \frac{\Lambda(l,m)}{lm}\mathcal{D}Z(\gamma_m,\gamma_l) + O\left(\frac{(\log\Delta)^5}{\log T}\right),$$

where

$$\begin{aligned}\Lambda(l,m) &= \sum_{a|m}\sum_{b|l}\mu(bm)\mu\left(\frac{m}{a}\right)(\log a)(\log b) \\ &= \Lambda(m)\sum_{b|l}\mu(bm)\log b = -\mu(m)\Lambda(m)\Lambda\big(l/(l,m^\infty)\big).\end{aligned}$$

Hence

$$E_{11} = (\log T)^{-2} \sum_{\substack{l,m<T \\ (l,m) \text{ regular}}}\sum \frac{\Lambda(l)\Lambda(m)}{lm}\mathcal{D}Z(\gamma_m,\gamma_l) + O\left(\frac{(\log\Delta)^5}{\log T}\right).$$

Here the restriction to (l,m) regular can be easily removed because l, m are prime powers. After that, applying the PNT we obtain (26.7) for E_{11} by partial summation.

Now we go to estimating E_0. Since the long Taylor expansion cannot hold, we shall use shorter expansions which produce weaker approximations, but sufficient because the number of exceptional pairs (c,d) is relatively small. We make further splitting

$$E_0 = E_{10} + E_{01} + E_{00}.$$

Here E_{10} runs over the pairs (c,d) with c being off the segments $[\Delta^{-2}T^w, \Delta T^w]$ and d being in the union of these segments. The sum E_{01} is defined analogously.

Finally E_{00} runs over the pairs (c, d) with c and d in the union of $[\Delta^{-2}T^w, \Delta T^w]$, $w \in \mathcal{W}$.

First we estimate E_{10}. Since in the relevant range $Z(x, y)$ is of $\mathcal{C}^2$-class in y, we have

$$Z(x+u, y+v) = Z(x+u, y) + v\frac{\partial Z}{\partial y}(x+u, y) + O(v^2).$$

We also have

$$\frac{\partial Z}{\partial y}(x+u, y) = \frac{\partial Z}{\partial y}(x, y) + O(u)$$

because $\frac{\partial Z}{\partial y}(x, y)$ is continuous in x. Hence

$$Z(x+u, y+v) = Z(x+u, y) + v\frac{\partial Z}{\partial y}(x, y) + O\big(v(u+v)\big).$$

The same approximation holds with u, v interchanged. Hence we get

$$Z(x+u, y+v)Z(x+v, y+u) = Z(x+u, y)Z(x+v, y) + (u+v)Z(x, y)\frac{\partial Z}{\partial y}(x, y) + O(u^2 + v^2)$$

by applying $Z(x+u, y) = Z(x, y) + O(u)$ and $Z(x+v, y) = Z(x, y) + O(v)$. We also get

$$Z(x+u, y)Z(x+v, y) = Z(x, y)\big(Z(x+u, y) + Z(x+v, y) - Z(x, y)\big) + O(uv).$$

Together we obtain the desired expansion

$$\begin{aligned} &Z(x+u, y+v)Z(x+v, y+u) \\ &\quad= Z(x, y)\left(Z(x+u, y) + Z(x+v, y) - Z(x, y) + (u+v)\frac{\partial Z}{\partial y}(x, y)\right) \\ &\qquad\qquad + O(u^2 + v^2) \\ &\qquad\qquad = Z^2(x, y) + O\big((u+v)|Z(x, y)| + u^2 + v^2\big). \end{aligned}$$

The contribution of the error term $O(u^2 + v^2)$ to E_{10} is then estimated by $O((\log \Delta)^5/\log T)$. Moreover, we can extend the range of c, d from $\max(c, d) < T/\max(a, b)$ to $\max(c, d) < T$ by arguments which were applied to E_{11} using (26.4). Having done this, the variables $u = \gamma_a$, $v = \gamma_b$ run freely over $a, b < \Delta$. Since the individual terms on the right side of the above expansion depend on u or v, but not on both, we get negligible contributions by applying (26.21). We conclude that

$$E_{10} \ll (\log \Delta)^5/\log T.$$

This estimate holds for E_{01} by the same arguments.

Finally, we estimate E_{00} quickly starting from

$$Z(x+u, y+v)Z(x+v, y+u) = Z^2(x, y) + O(u+v).$$

The contribution of the error term $O(u + v)$ to E_{00} is then estimated by $O((\log \Delta)^5/\log T)$. Then, dealing with the main term $Z^2(x, y)$ we can extend the range of c, d from $\max(c, d) < T/\max(a, b)$ to $\max(c, d) < T$ as we did so three times before in the context of E_{11}, E_{10}, and E_{01}. Now the variables $u = \gamma_a$, $v = \gamma_b$ run freely over $a, b < \Delta$, producing a negligible contribution of $Z^2(x, y)$ to E_{00} by applying (26.21). We conclude that

$$E_{00} \ll (\log \Delta)^5/\log T.$$

Gathering the above estimates we complete the proof of Proposition 26.1.

Remarks. If the crop function $Z(x,y)$ had continuous derivatives to sufficiently large order, then the above arguments would be technically simpler and shorter. However, we considered $Z(x,y)$ subject to less demanding conditions so the results can be applied in the future to a larger variety of functions.

http://dx.doi.org/10.1090/ulect/062/27

CHAPTER 27

The Off-diagonal Terms

These are the terms in (25.9). Our goal is to show that their contribution

$$I^*(\Phi) = \sum_{n_1 \neq n_2}\sum \frac{a_{n_1}\overline{a_{n_2}}}{\sqrt{n_1 n_2}} \hat{\Phi}\left(\frac{1}{2\pi}\log\frac{n_1}{n_2}\right) \tag{27.1}$$

is negligible. In general, such a claim would be false, but for our special coefficients (26.1), and with some restrictions on $Z(x,y)$, we shall prove

$$I^*(\Phi) \ll U(\log T)^{-\frac{1}{2}} \tag{27.2}$$

if $T(\log T)^{-\frac{1}{4}} \leqslant 2V \leqslant U \leqslant T$. We shall introduce the relevant restrictions on $Z(x,y)$ gradually when required by the arguments.

We begin with the basic requirements (26.3) and (26.4). Therefore our coefficients a_n are supported on $1 \leqslant n \leqslant T^2$ and are bounded by a smoothly cropped divisor function;

$$a_n \ll \sum_{\substack{lm=n\\ l,m<T}} \left(1 - \frac{\log l}{\log T}\right)\left(1 - \frac{\log m}{\log T}\right) \leqslant \tau(n). \tag{27.3}$$

Let $I_h(\Phi)$ denote the partial sum of (27.1) with $n_1 - n_2 = h$, so

$$I^*(\Phi) = \sum_{h\neq 0} I_h(\Phi). \tag{27.4}$$

Because $\hat{\Phi}(y)$ decays rapidly (see (25.4)), it is easy to estimate $I_h(\Phi)$ for large $|h|$. Indeed, all terms of $I^*(\Phi)$ with $|\log(n_1/n_2)| > V(2\log T)^2$ contribute at most $O(UT^{-\frac{1}{2}})$ by the trivial estimation

$$\sum_{\substack{1\leqslant n_1,n_2<T\\ V|\log(n_1/n_2)|>(2\log T)^2}}\sum \tau(n_1)\tau(n_2)(n_1n_2)^{-\frac{1}{2}} \exp\left(-\sqrt{2V\left|\log(n_1/n_2)\right|}\right) \ll T^{-\frac{1}{2}},$$

which is much smaller than the desired bound (27.2). The remaining terms have $|\log(n_1/n_2)| < V^{-1}(2\log T)^2 < \log 2$, so $\frac{1}{2} < \frac{n_1}{n_2} < 2$ and

$$|h| \leqslant 2n\left|\log\frac{n_1}{n_2}\right| < \frac{8n}{V}(\log T)^2 < \frac{8}{V}(T\log T)^2,$$

where $n = \min(n_1, n_2)$. Hence the remaining terms are for

$$n \gg |h|V(\log T)^{-2} \geqslant V(\log T)^{-2},$$

and they are relatively close to the diagonal. Due to these properties, we shall be able to make several cosmetical modifications in $I_h(\Phi)$.

First we write

$$\log\frac{n_1}{n_2}=\log\left(1+\frac{h}{n_2}\right)=\frac{h}{n_2}+O\left(\frac{h^2}{n_1n_2}\right),$$

and by the mean-value theorem,

$$\hat\Phi\left(\frac{1}{2\pi}\log\frac{n_1}{n_2}\right)=\hat\Phi\left(\frac{h}{2\pi n_2}\right)+O\left(\frac{h^2}{n_1n_2}TU\exp\left(-\sqrt{V|h|/n}\right)\right).$$

Summing over $n_1\neq n_2$ we see that the error term in the above approximation contributes at most

$$TU\sum_{n_1}\sum_{n_2}|a_{n_1}a_{n_2}|(n_1n_2)^{-\frac{3}{2}}h^2\exp\left(-\sqrt{V|h|/n}\right)$$
$$\ll TU\sum_n\frac{|a_n|^2}{n^3}\sum_h h^2\exp\left(-\sqrt{V|h|/n}\right)\ll\frac{TU}{V^3}\sum_n|a_n|^2.$$

By $|a_n|\ll\tau(n)$ one would get $\sum|a_n|^2\ll T^2(\log T)^3$, which is not good enough. However, using (27.3) we get

$$\sum_n|a_n|^2\ll T^2(\log T)^{-2}. \tag{27.5}$$

This yields the bound $U(T/V)^3(\log T)^{-2}$, which is smaller than the desired (27.2).

We prove (27.5) by elementary means as follows:

$$\sum_n|a_n|^2$$
$$\ll\sum\sum_{l_1m_1=l_2m_2}\left(1-\frac{\log l_1}{\log T}\right)\left(1-\frac{\log m_1}{\log T}\right)\left(1-\frac{\log l_2}{\log T}\right)\left(1-\frac{\log m_2}{\log T}\right)$$
$$\leqslant\sum_l\sum_m\tau(lm)\left(1-\frac{\log l}{\log T}\right)^2\left(1-\frac{\log m}{\log T}\right)^2$$
$$\leqslant\left(\sum_{l<T}\tau(l)\left(1-\frac{\log l}{\log T}\right)^2\right)^2,$$

and

$$\sum_{l<T}\tau(l)\left(1-\frac{\log l}{\log T}\right)^2\leqslant 2(\log T)^{-2}\sum_{a<\sqrt{T}}\sum_{b<T/a}\left(\log\frac{T}{ab}\right)^2$$
$$\ll T(\log T)^{-2}\sum_{a<\sqrt{T}}a^{-1}\ll T(\log T)^{-1}.$$

This yields (27.5).

We have proved that $\hat\Phi(\frac{1}{2\pi}\log\frac{n_1}{n_2})$ in (27.1) can be replaced by $\hat\Phi(h/2\pi n_2)$. Similarly we can replace $\sqrt{n_1n_2}$ in (27.1) by n_2 because of the approximation

$$\frac{1}{\sqrt{n_1n_2}}=\frac{1}{n_2}+O\left(\frac{|h|}{n_1n_2}\right).$$

Indeed the error term $O(|h|/n_1n_2)$ contributes at most

$$U\sum_n |a_n|^2 n^{-2}\sum_n |h|\exp\big(-\sqrt{V|h|/n}\big)$$
$$\ll UV^{-2}\sum_n |a_n|^2 \ll U(T/V)^2(\log T)^{-2},$$

which is smaller than the desired (27.2). Applying the above approximations we can write (27.4) in the slightly modified form

$$I^*(\Phi) = \sum_{h\neq 0}\mathcal{J}_h(\Phi) + O(U/\log T), \tag{27.6}$$

where

$$\mathcal{J}_h(\Phi) = \sum\sum_{n_1-n_2=h} a_{n_1}a_{n_2}n_2^{-1}\hat{\Phi}(h/2\pi n_2). \tag{27.7}$$

Next we exploit the convolution shape (26.1) of the coefficients a_n to make further modifications in $\mathcal{J}_h(\Phi)$. For notational simplicity we consider $Z(x,y)$ as a continuous function in $x \geqslant 0$, $y \geqslant 0$ which vanishes if $x \geqslant 1$ or $y \geqslant 1$. The extended function $Z(x,y)$ has bounded partial derivatives that are piecewise continuous. Moreover, we extend the function γ_t for all $t > 0$ by

$$\gamma_t = \max\left(0,\ \frac{\log t}{\log T}\right).$$

Obviously, the modified function γ_t is continuous and has derivative $\mathrm{d}\gamma_t/\mathrm{d}t \leqslant (t\log T)^{-1}$. The above modifications do not affect (27.7), which becomes

$$\mathcal{J}_h(\Phi) = \sum\sum_{l_1m_1-l_2m_2=h} \mu(m_1)\mu(m_2)(l_2m_2)^{-1}Z(\gamma_{m_1},\gamma_{l_1})Z(\gamma_{m_2},\gamma_{l_2})\hat{\Phi}(h/2\pi l_2m_2).$$

Since we think of $|h|$ being relatively small, it appears that l_1 is close to l_2m_2/m_1 so we can replace γ_{l_1} by $\gamma_{l_2m_2/m_1}$ in $Z(\gamma_{m_1},\gamma_{l_1})$ with a small error term. Precisely we have $l_1 = l_2m_2/m_1 + h/m_1$ and $\gamma_{l_1} = \gamma_{l_2m_2/m_1} + O(|h|/l_1m_1\log T)$. Hence, by the mean-value theorem,

$$Z(\gamma_{m_1},\gamma_{l_1}) - Z(\gamma_{m_1},\gamma_{l_2m_2/m_1}) \ll (1-\gamma_{m_1})|h|/l_1m_1\log T.$$

The error term in the above approximation contributes to $\mathcal{J}_h(\Phi)$ at most

$$\frac{|h|U}{\log T}\sum\sum_{\substack{l_1m_1-l_2m_2=h\\ l_1,l_2,m_1,m_2<T}} (1-\gamma_{m_1})(1-\gamma_{m_2})(1-\gamma_{l_2})(l_1l_2m_1m_2)^{-1}\exp\big(-\sqrt{|h|V/n}\big),$$

where $n = \min(l_1m_1, l_2m_2)$. This expression is similar to the one which we have encountered when replacing $\sqrt{n_1n_2}$ by n_2. Previously we have used the bound (27.3) for a_{n_1} and a_{n_2}. In the current situation the bound (27.3) for a_{n_2} is the same, but for a_{n_1} it is slightly different since the factor $1-\log l_1/\log T$ is replaced by $1/\log T$. This replacement makes no difference in the following estimates, consequently we get again a bound which is smaller than the desired (27.2). Therefore we can write (27.6) in the form

$$I^*(\Phi) = \sum_{h\neq 0}\mathcal{K}_h(\Phi) + O(U/\log T), \tag{27.8}$$

where

$$\text{(27.9)}\quad \mathcal{K}_h(\Phi) = \sum\sum_{l_1m_1-l_2m_2=h} \mu(m_1)\mu(m_2)(l_2m_2)^{-1}Z(\gamma_{m_1},\gamma_{\frac{l_2m_2}{m_1}})Z(\gamma_{m_2},\gamma_{l_2})\hat{\Phi}(h/2\pi l_2m_2).$$

Recall that l_1, l_2, m_1, m_2 in (27.9) are positive integers, and since $Z(x,y)$ vanishes if $x \geqslant 1$ or $y \geqslant 1$, it follows that $m_1, m_2, l_2, l_2m_2/m_1 < T$.

Remarks. The above arguments seem to be delicate and the resulting approximations turn out to be barely sufficient. However we could apply crude estimates if we were willing to accept the slightly stronger condition that a_n are supported on $n \leqslant T^2(\log T)^{-2012}$. But we have not imposed such condition exclusively for learning extra features. The reader can find that the first degree vanishing of the function $Z(x,y)$ at $x=1$ and $y=1$ in the construction of a_n cannot be dispensed.

So far we only performed cosmetical modifications in the off-diagonal terms; the more essential transformations (harmonic analysis of a_n) are yet to be made. What we have accomplished in $\mathcal{K}_h(\Phi)$ is that the variable l_1 disappeared in the summation terms. Therefore, the equation $l_1m_1 - l_2m_2 = h$ can be interpreted as the congruence $l_2m_2 \equiv -h \pmod{m_1}$. Putting $m_1 = da_1$, $m_2 = da_2$ with $(a_1,a_2)=1$, this congruence reduces to $l_2 \equiv -k\overline{a_2} \pmod{a_1}$, where $k = h/d$ and $\overline{a_2}$ denotes the multiplicative inverse of a_2 modulo a_1. The sum (27.9) becomes

$$\mathcal{K}_h(\Phi) = \sum_{dk=h}\mu(d)\sum_{a_1}\sum_{a_2}\frac{\mu(da_1a_2)}{da_1a_2}S_{dk}(a_1,a_2), \tag{27.10}$$

where

$$S_{dk}(a_1,a_2) = \sum_{l\equiv -k\overline{a_2} \pmod{a_1}} W(l/a_1) \tag{27.11}$$

and

$$W(x) = x^{-1}Z(\gamma_{da_1},\gamma_{xa_2})Z(\gamma_{da_2},\gamma_{xa_1})\hat{\Phi}(k/2\pi xa_1a_2). \tag{27.12}$$

Remember that $W(x)$ depends also on d, k, a_1, a_2, but we omit these variables for notational simplicity.

Next we replace the summation over l in (27.11) by integration (l is a positive integer). Precisely, we apply the Euler-Maclaurin formula

$$\sum_{l\equiv\alpha \pmod a} F(l/a) = \int F(x)\,\mathrm{d}x + \int \psi\left(x - \frac{\alpha}{a}\right)F'(x)\,\mathrm{d}x$$

where $\psi(x) = x - [x] - \frac{1}{2}$ is the saw function. This formula holds for any compactly supported function F which is continuous and has a piecewise continuous and bounded derivative. In our case, it yields

$$S_{dk}(a_1,a_2) = \int_0^T W(x)\,\mathrm{d}x + \int_0^T \psi\left(x + \frac{k\overline{a_2}}{a_1}\right)W'(x)\,\mathrm{d}x. \tag{27.13}$$

Actually the integration segment is shorter, $0 < x < T/\max(a_1,a_2)$. The two parts of (27.13) behave distinctly, so it makes sense to split $\mathcal{K}_h(\Phi)$ accordingly;

$$\mathcal{K}_h(\Phi) = \mathcal{L}_h(\Phi) + \mathcal{M}_h(\Phi), \tag{27.14}$$

where

$$\mathcal{L}_h(\Phi) = \sum_{dk=h} \mu(d) \sum_{a_1} \sum_{a_2} \frac{\mu(da_1a_2)}{da_1a_2} \int_0^T W(x)\,\mathrm{d}x, \tag{27.15}$$

and

$$\mathcal{M}_h(\Phi) = \sum_{dk=h} \mu(d) \sum_{a_1} \sum_{a_2} \frac{\mu(da_1a_2)}{da_1a_2} \int_0^T \psi\left(x + \frac{k\overline{a_2}}{a_1}\right) W'(x)\,\mathrm{d}x. \tag{27.16}$$

Then (27.8) becomes

$$I^*(\Phi) = \mathcal{L}(\Phi) + \mathcal{M}(\Phi) + O(U/\log T), \tag{27.17}$$

where

$$\mathcal{L}(\Phi) = \sum_{h\neq 0} \mathcal{L}_h(\Phi), \qquad \mathcal{M}(\Phi) = \sum_{h\neq 0} \mathcal{M}_h(\Phi). \tag{27.18}$$

Estimation of $\mathcal{L}(\Phi)$. This part makes use of the sign changes of the Möbius factor $\mu(da_1a_2)$ in an essential fashion. Introducing (27.12) into the integral in (27.15) and changing the variable of integration $x \to x/a_1a_2$ we get

$$\mathcal{L}(\Phi) = \int_0^\infty \frac{\rho(x)}{x} \sum_d \mu(d) \sum_{a_1} \sum_{a_2} \frac{\mu(da_1a_2)}{da_1a_2} Z(\gamma_{da_1}, \gamma_{x/a_1}) Z(\gamma_{da_2}, \gamma_{x/a_2})\,\mathrm{d}x$$

where

$$\rho(x) = \sum_{k\neq 0} \hat{\Phi}(k/2\pi x) = -\hat{\Phi}(0) + 2\pi x \sum_r \Phi(2\pi x r)$$

by Poisson's summation formula. The sum over $k \neq 0$ yields

$$\rho(x) \ll U \sum_{k=1}^{\infty} \exp\left(-\sqrt{2kV/x}\right) \ll xUV^{-1}$$

and the sum over r is bounded by $x^{-1}U$, because $xr \asymp U$, so it yields $\rho(x) \ll U$. Together, we get

$$\rho(x) \ll \min(x, V)\, UV^{-1}.$$

Let $\Delta = \Delta(T) = \exp(\log\log T)^3$ as in Section 26. By the Prime Number Theorem we find that the contribution of terms $a_1 \geqslant \Delta$ or $a_2 \geqslant \Delta$ in $\mathcal{L}(\Phi)$ is bounded by

$$\sum_{d<T} \frac{\log T}{d} \exp\left(-\lambda\sqrt{\Delta}\right) \int_0^{T^2} |\rho(x)| \frac{\mathrm{d}x}{x} \ll T(\log T)^{-2012}.$$

The remaining terms with $a_1, a_2 < \Delta$ contribute at most

$$(\log T)(\log \Delta)^2 \int_0^{\Delta T} |\rho(x)|(1-\gamma_{x/\Delta})^2 \frac{\mathrm{d}x}{x} \ll T\frac{(\log\Delta)^5}{\log T}$$

because $Z(x,y) \ll 1-y$ if $0 \leqslant y \leqslant 1$. Adding these two estimates, we obtain

$$\mathcal{L}(\Phi) \ll T(\log T)^{-1}(\log\log T)^{15}. \tag{27.19}$$

Estimation of $\mathcal{M}(\Phi)$. This part requires an additional restriction on the support of $Z(x,y)$ in the x variable which controls the length of the mollifier. We assume that the crop function $Z(x,y)$ is continuous in $x \geqslant 0$, $y \geqslant 0$, has support on $0 \leqslant x \leqslant \frac{1}{2}$, $0 \leqslant y \leqslant 1$, and it has bounded partial derivatives, piecewise continuous.

This assumption implies

$$Z(x,y) \ll 1-2x \qquad \text{if } 0 \leqslant x \leqslant \frac{1}{2}. \tag{27.20}$$

Having this property, we can estimate $\mathcal{M}(\Phi)$ quite easily. We start from

$$|\mathcal{M}(\Phi)| \leqslant \sum_k \sum_d \sum_{\substack{da_1<\sqrt{T}\\ da_2<\sqrt{T}}}\sum (da_1a_2)^{-1}\int_0^T |W'(x)|\,\mathrm{d}x,$$

where k, d, a_1, a_2 are positive integers and $W(x)$ is given by (27.12). Next we write a series of estimates;

$$\begin{aligned}
Z(\gamma_{da_1},\gamma_{xa_2}) &\ll 1-2\gamma_{da_1},\\
\frac{\mathrm{d}}{\mathrm{d}x}Z(\gamma_{da_1},\gamma_{xa_2}) &\ll \frac{1-2\gamma_{da_1}}{x\log T},\\
\hat{\Phi}(k/2\pi x a_1a_2) &\ll U\exp\big(-\sqrt{2kV/xa_1a_2}\big),\\
\frac{\mathrm{d}}{\mathrm{d}x}\hat{\Phi}(k/2\pi x a_1a_2) &\ll \frac{kTU}{x^2a_1a_2}\exp\big(-\sqrt{2kV/xa_1a_2}\big).
\end{aligned}$$

Hence

$$W'(x) \ll (1-2\gamma_{da_1})(1-2\gamma_{da_2})\left(1+\frac{kT}{xa_1a_2}\right)\frac{U}{x^2}\exp\big(-\sqrt{2kV/xa_1a_2}\big)$$

and

$$\int_0^T |W'(x)|\,\mathrm{d}x \ll TUV^{-2}(1-2\gamma_{da_1})(1-2\gamma_{da_2})a_1a_2k^{-1}\exp\big(-\sqrt{k/T}\big).$$

This yields

$$\mathcal{M}(\Phi) \ll TUV^{-2}(\log T)\sum_{d<\sqrt{T}} d^{-1}\left(\sum_{ad<\sqrt{T}}(1-2\gamma_{da})\right)^2.$$

Here we have

$$\sum_{ad<\sqrt{T}}(1-2\gamma_{da}) \ll \sqrt{T}/d\log T,$$

so

$$\mathcal{M}(\Phi) \ll UV^{-2}T^2(\log T)^{-1}. \tag{27.21}$$

Finally, adding (27.19) and (27.21) we conclude the proof of (27.2).

Notes about further improvements. Recall that the acceptable (negligible) bound (27.2) for the contribution of the off-diagonal terms is established for the crop function $Z(x,y)$ which is continuous in $x \geqslant 0$, $y \geqslant 0$, supported on $0 \leqslant x \leqslant \frac{1}{2}$, $0 \leqslant y \leqslant 1$, and it has piecewise continuous, bounded partial derivatives $\partial Z/\partial x$, $\partial^2 Z/\partial x\partial y$, and $\partial Z/\partial y$. Although the first part $\mathcal{L}(\Phi)$ of $I^*(\Phi)$ is fine for $Z(x,y)$ supported on $0 \leqslant x, y \leqslant 1$, the additional restriction of the support of $Z(x,y)$ in the x-variable is needed only for our method of estimating the second part $\mathcal{M}(\Phi)$. The supporting segment $0 \leqslant x \leqslant \frac{1}{2}$ could be extended slightly by exploiting the sign changes of $\mu(da_1a_2)\psi(x+k\overline{a_2}/a_1)$, which produce cancellations in the sums over a_1, a_2, and k. This depends on estimates for sums of Kloosterman sums which can be borrowed from the spectral theory of automorphic forms. Brian Conrey [**Con89**] succeeded (in his own settings) to get results for $Z(x,y)$ supported on $0 \leqslant x \leqslant \frac{4}{7}$, $0 \leqslant y \leqslant 1$. His arguments go far beyond the scope of these lecture notes.

http://dx.doi.org/10.1090/ulect/062/28

CHAPTER 28

Conclusion

Gathering the results of Chapters 25-27 (take $V = T(\log T)^{-\frac{1}{4}}$) we conclude

THEOREM 28.1. *Let $Z(x, y)$ be a continuous function in $x, y \geqslant 0$ with bounded and piecewise continuous partial derivatives of order up to* 3. *Moreover, $\partial Z/\partial x$ is continuous in y, and $\partial Z/\partial y$ is continuous in x. Suppose $Z(x, y)$ is supported on $0 \leqslant x \leqslant \frac{1}{2}$, $0 \leqslant y \leqslant 1$. Let $A(s)$ be the Dirichlet polynomial with coefficients a_n given by* (26.1). *Then*

$$\frac{1}{U}\int_T^{T+U}\left|A\left(\frac{1}{2}+it\right)\right|^2 \mathrm{d}t = \mathcal{E} + O\big((\log T)^{-\frac{1}{8}}\big) \tag{28.1}$$

if $T(\log T)^{-\frac{1}{8}} \leqslant U \leqslant T$, where $\mathcal{E}$ is given by (26.10).

The restriction of the support of $Z(x, y)$ in the x-variable to the segment $0 \leqslant x \leqslant \frac{1}{2}$ is crucial in our proof of Theorem 28.1. However, it is reasonable to expect that it can be relaxed substantially (see relevant investigations by D. Farmer [**Far93**]).

CONJECTURE (Long Mollifier). *The formula* (28.1) *holds true if $Z(x, y)$ is supported on $0 \leqslant x, y \leqslant 1$.*

We have now all the ingredients for completing the lower bound (16.2) for the number of critical zeros. As described in Chapter 23, it remains to prove (23.1) for the product $F(s) = G(s)M(s)$, where $G(s)$ is a linear combination of $\zeta(s)$ and $\zeta'(s)$ given by (23.4) and $M(s)$ is the mollifier given by the Dirichlet polynomial (23.7). The linear polynomials $P(x) = 1 - 2x$, $Q(y) = 1 - y$ which determine $M(s)$ and $G(s)$ are the original choice of Levinson, but they could be different.

We are going to derive (23.1) in greater generality. We take $G(s)$ given by (18.1) and $M(s)$ given by (23.7) with $P(x)$ quite arbitrary. For $G(s)$ we have a nice approximation (18.14) by Dirichlet polynomials $L(s)$, $L_1(s)$, $L_2(s)$ of length T. Actually $L_2(s)$ is not a Dirichlet polynomial in a strict sense, because its coefficients $\delta_l(s)$ depend on s; see (18.18). However, they are relatively small; see (18.19). So by Corollary 13.3

$$\int_T^{2T} |L_2(a+it)|^2 \,\mathrm{d}t \ll T(\log T)^{-3}.$$

Moreover, one gets directly from (13.15) that

$$\int_T^{2T} |M(a+it)|^2 \,\mathrm{d}t \ll T\log T.$$

Hence, by the Cauchy-Schwarz inequality we get the bound

$$\int_T^{2T} |L_2M(a+it)| \,\mathrm{d}t \ll T(\log T)^{-1},$$

which is good enough to be an insignificant contribution. Now, by the approximations (18.14) and (18.11), we get

$$
\begin{aligned}
(28.2)\qquad \int_T^{T+U} |F(a+it)|\,\mathrm{d}t = \int_T^{T+U} |LM(a+it)|\,\mathrm{d}t \\
+ O\left((\log T)^{-1} \int_T^{T+U} |L_1 M(a+it)|\,\mathrm{d}t + T(\log T)^{-1} \right).
\end{aligned}
$$

We are going to show that the integral of $|L_1 M|$ also gives an insignificant contribution. If one splitted the product $L_1 M$ by the Cauchy-Schwarz inequality and applied the mean values for $|L_1|^2$ and $|M|^2$, then one would get the estimate

$$\int_T^{2T} |L_1 M(a+it)|\,\mathrm{d}t \ll T \log T,$$

which is too weak (just by a constant). We can do better by applying Theorem 28.1. Unfortunately, (28.1) is not applicable directly for $A_1(s) = L_1 M(s - \frac{1}{2} + a)$ because the corresponding crop function

$$Z(x,y) = P(x) Q'(y) e^{Ry}$$

may fail to be continuous at $y = 1$. For example, our primary choice $Q(y) = \max(0, 1-y)$, $y \geqslant 0$, does not qualify. Fortunately, we can go around this problem in general by using the following decomposition

$$Q'(y) = (1 - y^r) Q'(y) + y^r Q'(y),$$

where r is a positive integer at our disposal. Accordingly, $A_1(s) = A_2(s) + A_3(s)$. In the first term, $A_2(s)$, the resulting crop function

$$Z(x,y) = P(x)(1-y^r) Q'(y) e^{Ry}$$

does satisfy the conditions of Theorem 28.1 (after rescaling by a factor of r to make the partial derivatives bounded uniformly in r), giving

$$(28.3)\qquad \int_T^{T+U} \left| A_2\left(\frac{1}{2} + it\right) \right| \mathrm{d}t \ll rU.$$

The second term, $A_3(s)$, is the product of the mollifier $M(s)$ against the Dirichlet polynomial

$$\sum_{l<T} b_l l^{-s} = \sum_{l<T} \left(\frac{\log l}{\log T}\right)^r Q'\left(\frac{\log l}{\log T}\right) l^{-s}$$

whose coefficients $b_l = \gamma_l^r Q'(\gamma_l)$ satisfy (24.4). Hence by Lemma 24.1 we get

$$(28.4)\qquad \int_T^{T+U} \left| A_3\left(\frac{1}{2} + it\right) \right| \mathrm{d}t \ll r^{-\frac{1}{4}} U \log T.$$

Adding up (28.3) and (28.4) we get (take $r \asymp (\log T)^{\frac{4}{5}}$)

$$(28.5)\qquad \int_T^{T+U} |L_1 M(a+it)|\,\mathrm{d}t \ll U (\log T)^{\frac{4}{5}}.$$

Hence (28.2) reduces to

$$(28.6)\qquad \int_T^{T+U} |F(a+it)|\,\mathrm{d}t = \int_T^{T+U} |LM(a+it)|\,\mathrm{d}t + O\big(U(\log T)^{-\frac{1}{5}}\big).$$

Here the Dirichlet polynomial $A(s) = LM(s - \frac{1}{2} + a)$ has coefficients a_n given by (26.1) with the crop function

$$Z(x, y) = P(x)Q(y)e^{Ry}, \tag{28.7}$$

which satisfies all the conditions of Theorem 28.1. Therefore (28.1) holds with $\mathcal{E} = C(R)$, where

$$C(R) = 1 + \int_0^1 \int_0^1 \big(\partial Z(x, y)\big)^2 \, \mathrm{d}x \, \mathrm{d}y. \tag{28.8}$$

This completes the proof of (23.1) with $c(R) = C(R)^{\frac{1}{2}}$ in full generality of $G(s)$.

http://dx.doi.org/10.1090/ulect/062/29

CHAPTER 29

Computations and the Optimal Mollifier

We end Part 2 of these lectures by computing $C(R)$ for various crop functions $Z(x, y)$ of type (28.7). First we assume that $P(x)$, $Q(y)$ are continuous functions in $x, y \geqslant 0$ with

$$P(0) = Q(0) = 1$$
$$P(x) = Q(y) = 0 \qquad \text{if } x, y \geqslant 1$$

and that the derivatives $P'(x)$, $Q'(y)$ are bounded and piecewise continuous. The additional property

$$Q(y) + Q(1-y) = 1 \qquad \text{if } 0 \leqslant y \leqslant 1$$

emerges from the Conrey construction of $G(s)$, but it is not required for the forthcoming analysis.

By (28.8), we obtain

$$C(R) = 1 + \int_0^1 \int_0^1 \Big(P'(x)Q(y)e^{Ry} - P(x)\big(Q(y)e^{Ry}\big)'\Big)^2 \,\mathrm{d}x\,\mathrm{d}y.$$

Opening the square we find that the cross terms yield

$$-2\left(\int_0^1 P'(x)P(x)\,\mathrm{d}x\right)\left(\int_0^1 \big(Q(y)e^{Ry}\big)'Q(y)e^{Ry}\,\mathrm{d}y\right) = -\frac{1}{2}P(0)^2Q(0)^2$$

and

$$C(R) = \frac{1}{2} + A\int_0^1 P'(x)^2\,\mathrm{d}x + A_1\int_0^1 P(x)^2\,\mathrm{d}x, \tag{29.1}$$

with

$$A = \int_0^1 \big(Q(y)e^{Ry}\big)^2\,\mathrm{d}y, \qquad A_1 = \int_0^1 \Big(\big(Q(y)e^{Ry}\big)'\Big)^2\,\mathrm{d}y. \tag{29.2}$$

Example (the original choice of Levinson). Take

$$Q(y) = \max(0, 1-y). \tag{29.3}$$

Then

$$A = \frac{1}{2R^2}\left(\frac{e^{2R}-1}{2R} - 1 - R\right), \tag{29.4}$$

$$A_1 = \frac{1}{2}\left(\frac{e^{2R}-1}{2R} + 1 - R\right) = 1 + AR^2. \tag{29.5}$$

Moreover, if $0 < \theta \leqslant 1$ and

$$P(x) = \max(0, 1 - x/\theta), \tag{29.6}$$

then

$$\int_0^1 \big(P'(x)\big)^2\,\mathrm{d}x = \frac{1}{\theta}, \qquad \int_0^1 P(x)^2\,\mathrm{d}x = \frac{\theta}{3}. \tag{29.7}$$

For this choice, (29.1) yields

$$C(R) = \frac{1}{2\theta R^2}\left(\frac{e^{2R}-1}{2R} + R - 1\right) + \frac{\theta}{6}\left(\frac{e^{2R}-1}{2R} + 1 - R\right) + \frac{1}{2}. \tag{29.8}$$

In particular, if $\theta = \frac{1}{2}$, we get (23.10) and complete the proof of Theorem 23.1.

Question. Given $R > 0$ and $Q(y)$, what is the best mollifying function $P(x)$, i.e., the continuous function which minimizes $C(R)$, subject to

$$P(0) = 1 \qquad \text{and} \qquad P(x) = 0, \quad \text{if } x \geqslant \theta? \tag{29.9}$$

We are going to find the optimal $P(x)$ by variational calculus. To this end, consider any smooth test function $g(x)$ on $0 \leqslant x \leqslant \theta$ with $g(0) = g(\theta) = 0$. Changing $P(x)$ to $P(x) + \varepsilon g(x)$ in (29.1) we get

$$\begin{aligned}\frac{1}{2} + A\int_0^\theta \big(P'(x) + \varepsilon g'(x)\big)^2\,\mathrm{d}x + A_1\int_0^\theta \big(P(x) + \varepsilon g(x)\big)^2\,\mathrm{d}x \\ = C(R) + 2\varepsilon A\int_0^\theta P'(x)g'(x)\,\mathrm{d}x + 2\varepsilon A_1\int_0^\theta P(x)g(x)\,\mathrm{d}x \\ + \varepsilon^2 A\int_0^\theta \big(g'(x)\big)^2\,\mathrm{d}x + \varepsilon^2 A_1\int_0^\theta g(x)^2\,\mathrm{d}x \\ = C(R) + 2\varepsilon\int_0^\theta \big(A_1P(x) - AP''(x)\big)g(x)\,\mathrm{d}x + O(\varepsilon^2)\end{aligned}$$

because, by partial integration and the boundary conditions, we have

$$\int_0^\theta P'(x)g'(x)\,\mathrm{d}x = -\int_0^\theta P''(x)g(x)\,\mathrm{d}x.$$

Since ε and $g(x)$ are arbitrary (ε is positive or negative), it follows that the minimizing function $P(x)$ must satisfy the Euler-Lagrange differential equation

$$AP''(x) = A_1P(x). \tag{29.10}$$

There are two linearly independent solutions $e^{\lambda x}$, $e^{-\lambda x}$ with

$$\lambda = (A_1/A)^{\frac{1}{2}}. \tag{29.11}$$

By the boundary conditions (29.9), we find the unique solution

$$P(x) = \sinh\lambda(\theta - x)/\sinh\lambda\theta. \tag{29.12}$$

Integrating by parts we find by (29.10) that

$$A\int_0^\theta P'(x)^2\,\mathrm{d}x = A\int_0^\theta P'(x)\,\mathrm{d}P(x) = -AP'(0) - A_1\int_0^\theta P(x)^2\,\mathrm{d}x.$$

Hence (29.1) for the optimal function $P(x)$ becomes $C(R) = \frac{1}{2} - AP'(0)$. Since $P'(0) = -\lambda\cosh\lambda\theta/\sinh\lambda\theta$, we conclude that for the optimal function (29.12) we have

$$C(R) = \frac{1}{2} + \lambda A\frac{\cosh\lambda\theta}{\sinh\lambda\theta} = \frac{1}{2} + (AA_1)^{\frac{1}{2}}\frac{\cosh\lambda\theta}{\sinh\lambda\theta}, \tag{29.13}$$

where A, A_1 are given by (29.2) and $\lambda = (A_1/A)^{\frac{1}{2}}$.

In particular, if $P(x)$ is optimal for $Q(y) = \max(0, 1-y)$, then

$$C(R) = \frac{1}{2} + \big(A(1+AR^2)\big)^{\frac{1}{2}} \frac{\cosh \lambda\theta}{\sinh \lambda\theta}, \tag{29.14}$$

with $\lambda = A^{-\frac{1}{2}}(1+AR^2)^{\frac{1}{2}}$ and $A = A(R)$ given by (29.4).

Epilogue. Assuming the Long Mollifier Conjecture, the above formulas hold for $\theta = 1$, so the optimal mollifier is attained for

$$P(x) = \sinh \lambda(1-x)/\sinh \lambda. \tag{29.15}$$

For $Q(y) = \max(0, 1-y)$, the numerical computations reveal that the best result is obtained for R near $3/4$ (the Levinson-Littlewood shift parameter). For $R = 3/4$ we get

$$A = 0.507667597, \qquad \lambda = 1.591317958, \qquad C(R) = 1.377774263$$

and the proportion of simple critical zeros of $\zeta(s)$ is at least

$$\kappa = 1 - \frac{1}{R}\log C(R) = 0.572707541. \tag{29.16}$$

Therefore one may say;

The Riemann Hypothesis is more likely to be true than not!

http://dx.doi.org/10.1090/ulect/062/30

APPENDIX A

Smooth Bump Functions

We begin by examining the function

$$f(x) = \begin{cases} \exp\big(-1/x(1-x)\big) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}. \tag{A.1}$$

Therefore $f(x)$ is smooth, supported on $0 \leqslant x \leqslant 1$, and it satisfies

$$f(x) = f(1-x). \tag{A.2}$$

LEMMA A.1. *For $0 < x \leqslant \frac{1}{2}$ we have*

$$|f^{(n)}(x)| \leqslant n! x^{-n} \exp(-1/2x), \quad n \geqslant 1. \tag{A.3}$$

PROOF. By Cauchy's formula

$$f^{(n)}(x) = \frac{n!}{2\pi i} \int_{|z|=x} f(z+x) z^{-n-1} \,\mathrm{d}z$$

we get

$$|f^{(n)}(x)| \leqslant n! x^{-n} \max_{|z|=x} |f(z+x)|.$$

We have

$$\begin{aligned} |f(z+x)| &= \exp\left(-\operatorname{Re}\left(\frac{1}{z+x} + \frac{1}{1-z-x}\right)\right) \\ &\leqslant \exp\left(-\operatorname{Re}\frac{1}{z+x}\right) = \exp(-1/2x) \end{aligned}$$

for every z with $|z| = x$. Indeed, writing $z = xe^{i\theta}$ we need to see that

$$\operatorname{Re}\frac{1}{e^{i\theta}+1} = \frac{1+\cos\theta}{(1+\cos\theta)^2 + \sin^2\theta} = \frac{1}{2}.$$ □

COROLLARY A.2. *For $n \geqslant 1$ we have*

$$|f^{(n)}(x)| \leqslant n! \left(\frac{2n}{e}\right)^n. \tag{A.4}$$

PROOF. The maximum of the upper bound (A.3) is attained at $x = 1/2n$. □

Put

$$C = \int_0^1 f(x)\,\mathrm{d}x; \tag{A.5}$$

an absolute constant, and

$$F(y) = \frac{1}{C} \int_{-\infty}^{y} f(x)\,\mathrm{d}x. \tag{A.6}$$

Clearly $F(y)$ is a smooth funtion with graph as in Figure A.1. That is, $F(y) = 0$ if $y \leqslant 0$, $0 \leqslant F(y) \leqslant 1$ if $0 \leqslant y \leqslant 1$, and $F(y) = 1$ if $y \geqslant 1$.

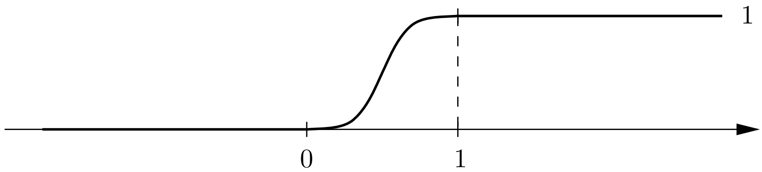

FIGURE A.1

Using $F(y)$ we construct the function

$$G(y) = F\left(\frac{y+V}{V}\right) - F\left(\frac{y-U}{V}\right), \tag{A.7}$$

where $0 < V < U$. Therefore $G(y)$ is a smooth function on $\mathbb{R}$ whose graph is seen in Figure A.2.

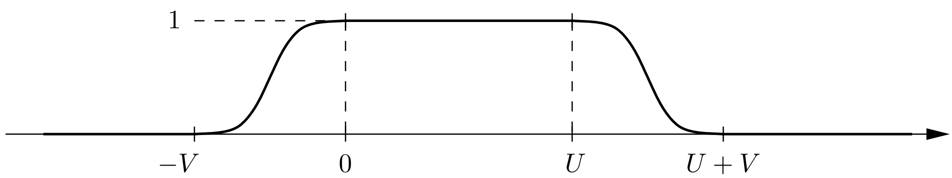

FIGURE A.2

The Fourier transform of $G(y)$ is equal to

$$\begin{aligned}\hat{G}(x) &= \int G(y)\,\mathrm{e}(xy)\,\mathrm{d}y = \frac{-1}{2\pi i x}\int G'(y)\,\mathrm{e}(xy)\,\mathrm{d}y \\ &= \frac{-1}{2\pi i x V C}\int \left(f\left(\frac{y+V}{V}\right) - f\left(\frac{y-U}{V}\right)\right)\mathrm{d}y \\ &= \frac{-1}{2\pi i x C}\big(\mathrm{e}(-xV) - \mathrm{e}(xU)\big)\hat{f}(xV).\end{aligned}$$

We arrange this in the form

$$\hat{G}(x) = e^{\pi i x (U-V)}\frac{\sin \pi x(U+V)}{\pi x C}\hat{f}(xV). \tag{A.8}$$

For $x = 0$ this gives

$$\hat{G}(0) = U + V. \tag{A.9}$$

Next we estimate $\hat{G}(x)$ for any x. First, by (A.4) we get

$$\begin{aligned}\hat{f}(x) &= \left(\frac{-1}{2\pi i x}\right)^n \int f^{(n)}(y)\,\mathrm{e}(xy)\,\mathrm{d}y \\ &\ll n!\left(\frac{2n}{e}\right)^n (2\pi|x|)^{-n} \ll \sqrt{n}\left(\frac{n^2}{\pi e^2|x|}\right)^n\end{aligned}$$

for any $n \geqslant 1$. Choosing $n = 1 + [\sqrt{\pi|x|}]$ we get

LEMMA A.3. *For any $x \in \mathbb{R}$ we have*

$$\hat{f}(x) \ll \big(1 + |x|\big)e^{-2\sqrt{\pi|x|}}. \tag{A.10}$$

Applying (A.10) to (A.8) we get

COROLLARY A.4. *For any $y \in \mathbb{R}$ we have*

$$\hat{G}(y) \ll Ue^{-2\sqrt{\pi|y|V}}, \tag{A.11}$$

where the implied constant is absolute.

Finally, we take

$$\Phi(t) = G(t - T) \tag{A.12}$$

This is a smooth function whose graph is seen in Figure A.3.

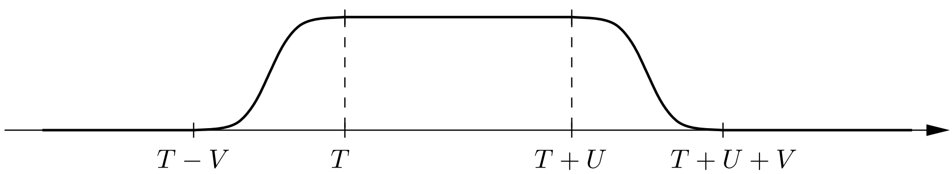

FIGURE A.3

COROLLARY A.5. *Let $\Phi(t)$ be as above with $0 < V < U < T$. Then the Fourier transform of $\Phi(t)$ satisfies*

$$\hat{\Phi}(y) = \mathrm{e}(yT)\hat{G}(y) \ll Ue^{-2\sqrt{\pi|y|V}}. \tag{A.13}$$

In particular,

$$\hat{\Phi}(0) = U + V. \tag{A.14}$$

Moreover,

$$\hat{\Phi}'(y) \ll TUe^{-2\sqrt{\pi|y|V}}, \tag{A.15}$$

where the implied constant is absolute.

http://dx.doi.org/10.1090/ulect/062/31

APPENDIX B

The Gamma Function

Here are a few basic properties of the gamma function:

$$\Gamma(s) = \int_0^\infty e^{-y} y^{s-1}\,\mathrm{d}y, \qquad \text{if } \sigma > 0,$$
$$= \int_1^\infty e^{-y} y^{s-1}\,\mathrm{d}y + \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(s+m)}, \qquad \text{for all } s.$$

This shows that $\Gamma(s)$ is meromorphic in the whole s-plane with only simple poles at $s = 0, -1, -2, -3, \ldots$, and

$$\operatorname*{res}_{s=-m} \Gamma(s) = \frac{(-1)^m}{m!}, \qquad m = 0, 1, 2, \ldots$$

The recurrence formula:

$$\Gamma(s+1) = s\Gamma(s).$$

The functional equation:

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

The duplication formula:

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

The Weierstrass product:

$$s\Gamma(s) = e^{-\gamma s} \prod_{m=1}^{\infty} \left(1 + \frac{s}{m}\right)^{-1} e^{\frac{s}{m}}.$$

This shows that $\Gamma(s)$ has no zeros, so $1/\Gamma(s)$ is entire.

Stirling's approximate formulas:

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \frac{1}{2} \log 2\pi + O\left(\frac{1}{|s|}\right)$$
$$\Gamma(s) = \left(\frac{2\pi}{s}\right)^{\frac{1}{2}} \left(\frac{s}{e}\right)^s \left\{1 + O\left(\frac{1}{|s|}\right)\right\}$$
$$\psi(s) = \big(\log \Gamma(s)\big)' = \frac{\Gamma'}{\Gamma}(s) = \log s + O\left(\frac{1}{|s|}\right).$$

These formulas hold in the sector $|\arg s| \leqslant \pi - \varepsilon$, where the implied constants depend only on ε.

In vertical strips we have:

$$\Gamma(\sigma + it) = (2\pi)^{\frac{1}{2}} (it)^{\sigma - \frac{1}{2}} \left(\frac{t}{e}\right)^{it} e^{-\frac{\pi}{2} t} \left\{1 + O\left(\frac{1}{t}\right)\right\}$$

$$\left|\Gamma(\sigma + it)\right| = (2\pi)^{\frac{1}{2}} t^{\sigma - \frac{1}{2}} e^{-\frac{\pi}{2} t} \left\{1 + O\left(\frac{1}{t}\right)\right\}$$

it $t > 0$, $\sigma' < \sigma < \sigma''$, where the implied constants depend only on σ', σ''.

Bibliography

[Car21] F. Carlson, *Über die Nullstellen der Dirichletschen Reihen und der Riemannschen ζ-Funktion.*, Ark. Mat. Astron. Fys. **15** (1921), no. 20, 28.

[Con85] J. B. Conrey, *On the distribution of the zeros of the Riemann zeta-function*, Topics in Analytic Number Theory (Austin, Tex., 1982), Univ. Texas Press, Austin, TX, 1985, pp. 28–41.

[Con89] ———, *More than two fifths of the zeros of the Riemann zeta function are on the critical line*, J. Reine Angew. Math. **399** (1989), 1–26.

[Far93] D. W. Farmer, *Long mollifiers of the Riemann zeta-function*, Mathematika **40** (1993), no. 1, 71–87.

[Fen12] S. Feng, *Zeros of the Riemann zeta function on the critical line*, J. Number Theory **132** (2012), no. 4, 511–542.

[HB79] D. R. Heath-Brown, *Simple zeros of the Riemann zeta-function on the critical line*, Bull. London Math. Soc. **11** (1979), 17–18.

[HL21] G. H. Hardy and J. E. Littlewood, *The zeros of Riemann's zeta-function on the critical line*, Math. Z. **10** (1921), no. 3-4, 283–317.

[Lev74] N. Levinson, *More than one third of zeros of Riemann's zeta-function are on $\sigma = 1/2$*, Advances in Math. **13** (1974), 383–436.

[Rie59] B. Riemann, *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsber. Akad. Berlin (1859), 671–680.

[Sel42] A. Selberg, *On the zeros of Riemann's zeta-function on the critical line*, Arch. Math. Naturvid. **45** (1942), no. 9, 101–114.

Index

The Riemann zeta function was introduced by L. Euler (1737) in connection with questions about the distribution of prime numbers. Later, B. Riemann (1859) derived deeper results about the prime numbers by considering the zeta function in the complex variable. The famous Riemann Hypothesis, asserting that all of the non-trivial zeros of zeta are on a critical line in the complex plane, is one of the most important unsolved problems in modern mathematics.

The present book consists of two parts. The first part covers classical material about the zeros of the Riemann zeta function with applications to the distribution of prime numbers, including those made by Riemann himself, F. Carlson, and Hardy–Littlewood. The second part gives a complete presentation of Levinson's method for zeros on the critical line, which allows one to prove, in particular, that more than one-third of non-trivial zeros of zeta are on the critical line. This approach and some results concerning integrals of Dirichlet polynomials are new. There are also technical lemmas which can be useful in a broader context.

For additional information and updates on this book, visit

www.ams.org/bookpages/ulect-62

ISBN 978-1-4704-1851-9

9 781470 418519

ULECT/62

AMS *on the* Web
www.ams.org