การแจกแจงเอกรูปของภาคเศษส่วนของค่าเฉลี่ยของตัวหารเฉพาะกำลังเค

นายจิตรภาณุ ชัยรัตนกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562

UNIFORM DISTRIBUTION OF FRACTIONAL PARTS OF AVERAGE OF kTH POWER OF PRIME DIVISORS

Mr. Chitpanu Chairatanagul

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Mathematics
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2019

Thesis Title	Uniform Distribution of Fractional Parts of Average of kth Power of Prime Divisors
By	Mr. Chitpanu Chairatanagul
Field of Study	Mathematics
Thesis Advisor	Assistant Professor Keng Wiboonton, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

. . . . Dean of Faculty of Science

(Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

. . . . Chairperson

(Professor Yotsanan Meemark, Ph.D.)

. . . . Thesis Advisor

(Assistant Professor Keng Wiboonton, Ph.D.)

. . . . Examiner

(Associate Professor Tuangrat Chaichana, Ph.D.)

. . . . External Examiner

(Assistant Professor Aram Tangboonduangjit, Ph.D.)

จิตรภาณุ ชัยรัตนกุล : การแจกแจงเอกรูปของภาคเศษส่วนของค่าเฉลี่ยของตัวหารเฉพาะกำลังเค. (UNIFORM DISTRIBUTION OF FRACTIONAL PARTS OF AVERAGE OF kTH POWER OF PRIME DIVISORS) อ.ที่ปรึกษาวิทยานิพนธ์ : ผศ.ดร.เก่ง วิบูลย์ธัญญ์, 39 หน้า.

ในวิทยานิพนธ์ฉบับนี้ เราพิสูจน์ว่าสำหรับทุกจำนวนเต็มบวก k ที่มากกว่าหรือเท่ากับสอง ลำดับของภาคเศษส่วนของค่าเฉลี่ยเลขคณิตของกำลังอันดับ k ของตัวหารเฉพาะของ n ที่ถูกนับแบบมีหรือไม่มีภาวะซ้ำ โดยที่ n เป็นจำนวนเต็มบวกใด ๆ เป็นลำดับแจกแจงเอกรูปมอดุโลหนึ่ง

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต ....

สาขาวิชา คณิตศาสตร์ ลายมือชื่ออาจารย์ที่ปรึกษา ....

ปีการศึกษา 2562

6171924323 : MAJOR MATHEMATICS

KEYWORDS: ARITHMETIC FUNCTIONS / EXPONENTIAL SUMS / UNIFORM DISTRIBUTION

CHITPANU CHAIRATANAGUL : UNIFORM DISTRIBUTION OF FRACTIONAL PARTS OF AVERAGE OF kTH POWER OF PRIME DIVISORS. THESIS ADVISOR : ASST. PROF. KENG WIBOONTON, Ph.D., 39 pp.

In this thesis, we prove that for every positive integer k greater than or equal to two, the sequence of the fractional parts of the arithmetic mean of the kth power of prime divisors of n counted with or without multiplicity, as n runs through all positive integers, is uniformly distributed modulo one.

Department	Mathematics and Computer Science	Student's Signature	....
Field of Study	Mathematics	Advisor's Signature	....
Academic Year	2019		

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance and support of several people. Firstly, I would like to thank Assistant Professor Keng Wiboonton, Ph.D., my thesis advisor, for providing me his valuable advice, allowing me to establish this work, and supporting me during conducting the thesis and preparing the manuscript. I would also like to thank Professor Yotsanan Meemark, Ph.D., Associate Professor Tuangrat Chaichana, Ph.D. and Assistant Professor Aram Tangboonduangjit, Ph.D., my thesis committee, for their comments and suggestions to the thesis. I would like to thank the Professor Dr. Tab Nilanidhi Foundation for awarding me a partial scholarship. Special thanks are given to all my friends at the department. The time we spent together is one of the most memorable moments in my life. Finally, I would like to express my gratitude to my parents for their endless and invaluable love.

CONTENTS

Abstract in Thai iv

Abstract in English v

Acknowledgements vi

1 Introduction 1

2 Preliminaries 3

2.1 Equidistribution of sequences . . . 3

2.2 Some elementary results in analytic number theory . . . 6

2.2.1 Arithmetic functions, Dirichlet series and Euler products . . . 6

2.2.2 Summation formulae and some useful estimates . . . 10

2.2.3 Distribution of prime numbers . . . 18

2.2.4 Primes in arithmetic progressions . . . 22

2.2.5 Generalized Gauss sums . . . 25

3 Uniform distribution of arithmetic mean of kth power of prime divisors 27

3.1 Proof of Theorem 3.1 . . . 30

3.1.1 The exceptional set $\mathcal{E}$ . . . 30

3.1.2 The remaining n . . . 32

3.2 Proof of Theorem 3.2 . . . 35

References 37

Curriculum vitae 39

CHAPTER I

INTRODUCTION

Throughout this thesis, the letter p always denotes a prime number. We use the letter n to denote a positive integer unless otherwise stated. For any two integers a and b, the greatest common divisor of a and b is denoted by (a, b). The notation $\#A$ is used to denote the cardinality of a set A. If $g(x)$ is a non-negative function, we write $f(x) = O(g(x))$ to mean that there exists a constant $C > 0$ such that $|f(x)| \leq Cg(x)$ for all $x \geq x_0$ for some x_0. Alternatively, we may use the notation $f(x) \ll g(x)$, due to Vinogradov, if there is no main term. We write $f(x) = O_\alpha(g(x))$ or $f(x) \ll_\alpha g(x)$ to indicate that the constant C may depend on some parameter α. We write $f(x) = o(g(x))$ to mean that $f(x)/g(x) \to 0$ as x tends to its limit. As usual, we shall mean $x \to \infty$ unless otherwise specified.

Let $\omega(n)$ denote the number of distinct prime divisors of n. In 2005, Banks et al. [2] introduced the arithmetic mean of distinct prime divisors of n, which is defined for $n \geq 2$ by

$$\rho(n) = \frac{1}{\omega(n)} \sum_{p|n} p,$$

and set $\rho(1) = 1$. They studied the distributional properties of the sequence $(\rho(n))$ by showing that for sufficiently large N a bound for the exponential sum

$$\sum_{n=1}^{N} \mathbf{e}(a\rho(n)) \ll \frac{|a|N}{\log\log N} \tag{1.1}$$

holds for every integer $a \neq 0$. Here $\mathbf{e}(x) = \exp(2\pi i x)$ for all real numbers x. Moreover, they showed that for sufficiently large N the discrepancy $D(N)$ of the sequence $(\rho(n))$ is

$$D(N) \ll \frac{(\log\log\log N)^2}{\log\log N} \tag{1.2}$$

(see the definition of discrepancy in §2.1). Furthermore, they pointed out that the estimates (1.1) and (1.2) also hold for the function

$$\widetilde{\rho}(n) = \frac{1}{\Omega(n)} \sum_{\substack{p^a|n \\ a\geq 1}} p,$$

where $\Omega(n)$ denotes the number of prime divisors of n counted with multiplicity. Their results imply that the sequences $(\rho(n))$ and $(\widetilde{\rho}(n))$ are uniformly distributed.

The geometric mean of distinct prime divisors of n can be defined by

$$g(n) = \left(\prod_{p|n} p\right)^{1/\omega(n)}.$$

In 2006, Luca and Shparlinski [9] studied the uniform distribution of the sequences $(g(n))$, $(n^{1/\omega(n)})$ and $(n^{1/\Omega(n)})$ by proving that their discrepancies are all equal to $(\log N)^{-1+o(1)}$ as $N \to \infty$. Noting that these three sequences are all the same if n is square-free, and the function $n^{1/\Omega(n)}$ represents the geometric mean of prime divisors of n taken with multiplicity.

Let us define

$$\widetilde{h}(n) = \frac{\Omega(n)}{\sum_{\substack{p^a|n \\ a\geq 1}} 1/p} \quad \text{and} \quad h(n) = \frac{\omega(n)}{\sum_{p|n} 1/p},$$

which can be interpreted as the harmonic means of prime divisors of n taken with and without multiplicity, respectively. In 2009, Kátai and Luca [6] proved that the sequences $(h(n))$ and $(\widetilde{h}(n))$ are uniformly distributed. More generally, they proved that if $f(n)$ is an additive function such that there exist two positive constants c_1 and c_2 such that $f(p) < c_1/p$ and $0 < f(p^a) < c_2$ for all primes p and all positive integers a, then the sequences $(\omega(n)/f(n))$ and $(\Omega(n)/f(n))$ are uniformly distributed.

For every positive integer k, let $\rho_k(n)$ and $\widetilde{\rho}_k(n)$ be defined for $n \geq 2$ by

$$\rho_k(n) = \frac{1}{\omega(n)} \sum_{p|n} p^k \quad \text{and} \quad \widetilde{\rho}_k(n) = \frac{1}{\Omega(n)} \sum_{\substack{p^a|n \\ a\geq 1}} p^k,$$

and set $\rho_k(1) = 1 = \widetilde{\rho}_k(1)$. Thus $\widetilde{\rho}_k(n)$ and $\rho_k(n)$ represent the arithmetic means of the kth power of prime divisors of n taken with and without multiplicity, respectively. As defined above, we note that $\rho_1 = \rho$ and $\widetilde{\rho}_1 = \widetilde{\rho}$. The purpose of this thesis is to study the distributional properties of the sequences $(\rho_k(n))$ and $(\widetilde{\rho}_k(n))$ for all $k \geq 2$ by estimating their exponential sums and their discrepancies.

The rest of this thesis is structured as follows. Chapter 2 contains some materials which are necessary for us to develop our main results. The chapter is divided into two sections. Section 2.1 introduces the fundamental concept of uniform distribution of sequences of real numbers. Section 2.2 provides some terminology and elementary results in analytic number theory. Chapter 3 presents the main results of the thesis, namely the bounds for exponential sums and discrepancies of the sequences $(\rho_k(n))$ and $(\widetilde{\rho}_k(n))$ for $k \geq 2$.

CHAPTER II

PRELIMINARIES

2.1 Equidistribution of sequences

For any real number x, we denote by $[x]$ the unique integer such that $[x] \leq x < [x] + 1$, called the *integral part* of x. Let $\{x\} = x - [x]$, called the *fractional part* of x. Thus we may treat x as an element of the quotient group $\mathbb{R}/\mathbb{Z}$ by considering only its fractional part and ignoring its integral part.

Let (u_n) be a sequence of real numbers. For any positive integer N and any real number $0 \leq \alpha < 1$, let $A(N, \alpha)$ denote the number of positive integers n not exceeding N such that $0 \leq \{u_n\} \leq \alpha$. The sequence (u_n) is said to be *uniformly distributed* (modulo 1) or *equidistributed* (modulo 1) if

$$\lim_{N \to \infty} \frac{1}{N} A(N, \alpha) = \alpha \tag{2.1}$$

holds for all $\alpha \in [0, 1)$. Alternatively, we may represent (2.1) as

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[0,\alpha]}(\{u_n\}) = \int_0^1 \mathbb{1}_{[0,\alpha]}(u)\, \mathrm{d}u \tag{2.2}$$

for all $\alpha \in [0, 1)$. Here $\mathbb{1}_{[0,\alpha]}$ is the characteristic function of the closed interval $[0, \alpha]$, that is, $\mathbb{1}_{[0,\alpha]}(u) = 1$ if $u \in [0, \alpha]$ and 0 otherwise. The equation (2.2) leads us to establish the following theorem.

Theorem 2.1. *The following statements are equivalent:*

(i) The sequence (u_n) is uniformly distributed.

(ii) For every real- or complex-valued Riemann-integrable function f defined on $[0, 1]$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\{u_n\}) = \int_0^1 f(u)\, \mathrm{d}u. \tag{2.3}$$

(iii) The equation (2.3) holds for every real- or complex-valued continuous function f defined on $[0, 1]$.

In fact, the sequence (u_n) is uniformly distributed if and only if (2.3) holds for every Riemann-integrable function f defined on $\mathbb{R}$ with period 1.

Proof. The implication (ii) $\Rightarrow$ (iii) is trivial since every continuous function is Riemann-integrable. To prove that (iii) implies (i), we fix $0 \le \alpha < 1$, and let $\varepsilon > 0$ be arbitrary. Then there exist two continuous functions g_1 and g_2 such that $g_1 \le \mathbb{1}_{[0,\alpha]} \le g_2$ on $[0,1]$ and $\int_0^1 (g_2(u) - g_1(u))\,\mathrm{d}u < \varepsilon$. Thus we have

$$\begin{aligned}\int_0^1 \mathbb{1}_{[0,\alpha]}(u)\,\mathrm{d}u - \varepsilon \le \int_0^1 g_2(u)\,\mathrm{d}u - \varepsilon \le \int_0^1 g_1(u)\,\mathrm{d}u &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} g_1(\{u_n\}) \\ &\le \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{[0,\alpha]}(\{u_n\}).\end{aligned}$$

Similarly, we also have

$$\int_0^1 \mathbb{1}_{[0,\alpha]}(u)\,\mathrm{d}u + \varepsilon \ge \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{[0,\alpha]}(\{u_n\}).$$

Since ε can be arbitrarily small, we conclude that (u_n) is uniformly distributed by (2.2). To show that (i) implies (ii), suppose that f is Riemann-integrable on $[0,1]$, and let $\varepsilon > 0$ be arbitrary. Then there exist two step functions f^- and f^+ such that $f^- \le f \le f^+$ on $[0,1]$ (that is, f^- and f^+ are linear combinations of characteristic functions of closed subintervals of $[0,1]$) and $\int_0^1 (f^+(u) - f^-(u))\,\mathrm{d}u < \varepsilon$. Arguing as above, we yield

$$\int_0^1 f(u)\,\mathrm{d}u - \varepsilon \le \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\{u_n\}) \le \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\{u_n\}) \le \int_0^1 f(u)\,\mathrm{d}u + \varepsilon.$$

Therefore we obtain (2.3) since ε can be arbitrarily small. $\square$

The following criterion characterizes uniformly distributed sequences, due to Weyl in 1916 (see Theorem 4.1.9 in [12]).

Theorem 2.2 (Weyl's criterion)**.** *The sequence (u_n) is uniformly distributed if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{e}(au_n) = 0 \tag{2.4}$$

holds for all integers $a \ne 0$.

Historically, the first example of a uniform distributed sequence is the sequence $(n\alpha)$ for any irrational number α, called the *Kronecker sequence*, due to Kronecker who first proved in 1884 that this sequence is dense in the unit interval $[0,1]$. The uniform distribution of the Kronecker sequence was proved independently by Bohl, Sierpiński and Weyl in 1909–1910. We shall prove the following statement by using Weyl's criterion.

Corollary 2.3 (Bohl–Sierpiński–Weyl)**.** *The sequence* $(n\alpha)$ *is uniformly distributed if and only if* α *is irrational.*

Proof. If $\alpha = a/b$ for some integers a, b with $b \geq 1$, then $\sum_{n\leq N} \mathbf{e}(bn\alpha) = N$, and hence the sequence $(n\alpha)$ is not uniformly distributed since the limit (2.4) does not hold for b. Conversely, suppose that α is irrational. Then for any integer $k \neq 0$,

$$\left|\sum_{n=1}^{N} \mathbf{e}(kn\alpha)\right| = \left|\frac{\mathbf{e}(k\alpha)\,(1-\mathbf{e}(kN\alpha))}{1-\mathbf{e}(k\alpha)}\right| \leq \frac{2}{|1-\mathbf{e}(k\alpha)|} = \frac{1}{|\sin(\pi k\alpha)|} \leq \frac{1}{2\|k\alpha\|},$$

where $\|x\|$ denotes the distance from x to the nearest integer. Thus the limit (2.4) holds for all integers $k \neq 0$, and hence the sequence $(n\alpha)$ is uniformly distributed. □

The *discrepancy* of the sequence (u_n) is defined by

$$D(N) = \sup_{0\leq\alpha<1} \left|\frac{1}{N}A(N,\alpha) - \alpha\right|.$$

A bound for the discrepancy in terms of exponential sums is provided by the following theorem, due to Erdős and Turán in 1948 (see Theorem 2.5 in Chapter 2 of [8] or Theorem 11.4.8 in [11]).

Theorem 2.4 (The Erdős–Turán inequality)**.** *For any finite sequence of real numbers* $u_1, u_2, \ldots, u_N$*, and any positive integer* M*, there exists an absolute constant* $C > 0$ *such that*

$$D(N) \leq C\left(\frac{1}{M+1} + \sum_{a=1}^{M} \frac{1}{a}\left|\frac{1}{N}\sum_{n=1}^{N} \mathbf{e}(au_n)\right|\right).$$

The following theorem gives another criterion for uniform distributed sequences.

Theorem 2.5. *The sequence* (u_n) *is uniformly distributed if and only if* $D(N) \to 0$ *as* $N \to \infty$*.*

Proof. The sufficiency is obvious. To prove the necessity, we use the Erdős–Turán inequality to obtain

$$0 \leq \liminf_{N\to\infty} D(N) \leq \limsup_{N\to\infty} D(N) \ll \frac{1}{M+1} + \sum_{a=1}^{M} \frac{1}{a}\left|\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{e}(au_n)\right| \ll \frac{1}{M+1}$$

by noting that the absolute term vanishes by Weyl's criterion. Since the last term tends to 0 as $M \to \infty$, we conclude that $D(N) \to 0$ as $N \to \infty$. □

2.2 Some elementary results in analytic number theory

2.2.1 Arithmetic functions, Dirichlet series and Euler products

Analytic number theory normally deals with a *Dirichlet series* $\sum_{n=1}^{\infty} f(n)n^{-s}$, where $f(n)$ is an *arithmetic function*, a real- or complex-valued function whose domain is the set of natural numbers or some subset of the natural numbers, and s is a complex variable. Traditionally, we write $s = \sigma + it$ to represent its real and imaginary parts. Note that if the series $\sum_{n=1}^{\infty} |f(n)n^{-s}|$ does not converge everywhere or converge nowhere, then $\sigma_a \coloneqq \inf \{\sigma : \sum |f(n)n^{-s}| < \infty\}$ exists. We call σ_a the *abscissa of absolute convergence.* Also, we define $\sigma_a = -\infty$ and $\sigma_a = +\infty$ if the series $\sum |f(n)n^{-s}|$ converges everywhere and converges nowhere, respectively.

Several examples of arithmetic functions are given as follows:

- $\mathbf{id}(n) = n$ for all n, called the *identity function*,
- $\mathbf{1}(n) = 1$ for all n,
- $\delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases}$
- $\omega(n) = \sum\limits_{p \mid n} 1$, the number of distinct prime divisors of n,
- $\Omega(n) = \sum\limits_{\substack{p^a \mid n \\ a \geq 1}} 1$, the number of prime divisors of n counted with multiplicity,
- $\mu(n) = \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is square-free}, \\ 0 & \text{otherwise}, \end{cases}$
  known as the *Möbius mu function*,
- $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^{\times}$, known as *Euler's totient function*,
- $\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and some } k \geq 1, \\ 0 & \text{otherwise}, \end{cases}$
  known as the *von Mangoldt lambda function.*

An arithmetic function f is called *multiplicative* if $f(1) \neq 0$ and $f(mn) = f(m)f(n)$ whenever m and n are relatively prime. A multiplicative function f is called *completely multiplicative* if $f(mn) = f(m)f(n)$ for all m and n. The following theorem asserts that an absolutely convergent Dirichlet series of some multiplicative function can be expressed as an infinite product over primes (see Theorem 1.9 in [10]).

Theorem 2.6. *Let f be a multiplicative function such that $\sum_{n=1}^{\infty} f(n)n^{-s}$ converges absolutely for $\sigma > \sigma_a$, then*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^{2s}} + \cdots\right) \quad \textit{for } \sigma > \sigma_a. \tag{2.5}$$

Moreover, if f is completely multiplicative, then

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1} \quad \textit{for } \sigma > \sigma_a. \tag{2.6}$$

In each case, the product on the right is called the Euler product *of the Dirichlet series.*

Let us introduce the *Riemann zeta function*, which is defined for $\sigma > 1$ by

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

Since the function $\mathbf{1}$ is completely multiplicative and $\sum n^{-\sigma} < \infty$ for $\sigma > 1$, it follows from (2.6) that

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad \text{for } \sigma > 1. \tag{2.7}$$

This deduces that $\zeta(s) \neq 0$ for $\sigma > 1$ since the product on the right does not vanish.

Here we present two short proofs of infinitude of primes based on the relation (2.7).

First proof. This proof was first discovered by Euler in 1737. Treating s as a real number and then taking the limit $s \to 1^+$, the sum on the left tends to the harmonic series $\sum 1/n$ which is divergent, forcing the product on the right to be infinite.

Second proof. Putting $s = 2$ in (2.7), we obtain

$$\frac{\pi^2}{6} = \zeta(2) = \prod_p \left(1 - \frac{1}{p^2}\right)^{-1}.$$

Since π^2 is irrational (otherwise there exist positive integers a and b such that $\pi^2 = a/b$, and so π satisfies the algebraic equation $bx^2 - a = 0$, contradicting the transcendentality of π), the product on the right must have infinitely many factors.

Using (2.5) we also have

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) \quad \text{for } \sigma > 1 \tag{2.8}$$

by noting that μ is multiplicative and $\sum |\mu(n)| n^{-\sigma} \leq \sum n^{-\sigma} < \infty$ for $\sigma > 1$.

Theorem 2.7. *Let $F(s) = \sum f(n)n^{-s}$ and $G(s) = \sum g(n)n^{-s}$ be two Dirichlet series, and let $H(s) = \sum h(n)n^{-s}$ whose coefficients $h(n)$ are given by*

$$h(n) = \sum_{km=n} f(k)g(m). \tag{2.9}$$

Then $H(s)$ is absolutely convergent and $H(s) = F(s)G(s)$ for all s in the half-plane that $F(s)$ and $G(s)$ are both absolutely convergent.

Proof. We observe that

$$\begin{aligned} F(s)G(s) = \sum_{k=1}^{\infty} f(k)k^{-s} \sum_{m=1}^{\infty} g(m)m^{-s} &= \sum_{k=1}^{\infty}\sum_{m=1}^{\infty} f(k)g(m)(km)^{-s} \\ &= \sum_{n=1}^{\infty}\left(\sum_{km=n} f(k)g(m)\right) n^{-s} = H(s). \end{aligned}$$

The rearrangement of terms is justified by absolute convergences of $F(s)$ and $G(s)$. $\square$

Note that the equation (2.9) can be rewritten as

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

The function h is called the *Dirichlet convolution* of f and g, and denoted by $h = f * g$. By means of Theorem 2.7 together with (2.7) and (2.8), we obtain the identity

$$\sum_{d|n} \mu(d) = \delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases} \tag{2.10}$$

and we write $\mu * \mathbf{1} = \delta$. If we let $\mathcal{A}$ be the set of all arithmetic functions f with $f(1) \neq 0$, then $(\mathcal{A}, *)$ forms an abelian group with identity δ. Thus the equation (2.10) is equivalent to saying that the Möbius mu function and the function $\mathbf{1}$ are convolutional inverses of each other. This enables us to establish the *Möbius inversion formula*: $g = f * \mathbf{1}$ if and only if $f = g * \mu$ for any arithmetic functions f and g. By using (2.10), we find that

$$\varphi(n) = \sum_{\substack{m=1 \\ (m,n)=1}}^{n} 1 = \sum_{m=1}^{n} \delta((m,n)) = \sum_{m=1}^{n} \sum_{d|(m,n)} \mu(d) = \sum_{d|n} \mu(d) \sum_{q=1}^{n/d} 1 = \sum_{d|n} \mu(d)\frac{n}{d}.$$

This proves that $\varphi = \mu * \mathbf{id}$. Consequently, it can be easily derived that

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

Also, we obtain the identity $n = \sum_{d|n} \varphi(d)$ by the Möbius inversion formula. By using Theorem 2.6 and Theorem 2.7, we yield

$$\sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)} = \prod_p \frac{1-p^{-s}}{1-p^{1-s}} \quad \text{for } \sigma > 2.$$

Next, we show that any Dirichlet series has an *abscissa of convergence* σ_c with the property that the series converges for all s in the half-plane $\sigma > \sigma_c$, and diverges for all s in the half-plane $\sigma < \sigma_c$. Also, we let $\sigma_c = -\infty$ and $\sigma_c = +\infty$ if the series converges everywhere and converges nowhere, respectively. We shall prove the existence of such half-plane of convergence by means of the following lemma (see Lemma 2 in §11.6 of [1]).

Lemma 2.8. *Suppose that the partial sums of the Dirichlet series* $\sum f(n)n^{-s_0}$ *are bounded, namely* $\left|\sum_{n\le x} f(n)n^{-s_0}\right| \le M$ *for all* $x \ge 1$*. Then for all* s *with* $\sigma > \sigma_0$*,*

$$\left|\sum_{a<n\le b} f(n)n^{-s}\right| \le 2Ma^{\sigma_0-\sigma}\left(1 + \frac{|s-s_0|}{\sigma-\sigma_0}\right). \tag{2.11}$$

By letting $a \to \infty$ in (2.11), we see that $\sum f(n)n^{-s}$ converges for all s with $\sigma > \sigma_0$, and we let $\sigma_c = \inf\{\sigma : \sum f(n)n^{-s} < \infty\}$. The estimate (2.11) also implies that for an arbitrary constant $R > 0$ the Dirichlet series $\sum f(n)n^{-s}$ converges uniformly in the region $\mathcal{R} = \{s : \sigma \ge \sigma_0 \text{ and } |t - t_0| \le R|\sigma - \sigma_0|\}$ since the expression in the parentheses of (2.11) is bounded by $2 + R$ which is independent of s.

To obtain some more analytic properties of Dirichlet series, we require the following fact in complex analysis (see §5.2 in Chapter 2 of [13]).

Lemma 2.9. *Let* (f_n) *be a sequence of analytic functions on a region* Ω*. Suppose that* (f_n) *converges uniformly to a function* f *on every compact subset of* Ω*. Then* f *is analytic on* Ω*, and the sequences of derivatives* (f_n') *converges uniformly to* f' *on every compact subset of* Ω*.*

Applying Lemma 2.9 to the sequence of partial sums, we have the following theorem.

Theorem 2.10. *A Dirichlet series* $F(s) = \sum f(n)n^{-s}$ *is locally uniformly convergent and is analytic for all* s *in its half-plane of convergence* $\sigma > \sigma_c$*. The same statement also holds for its differentiated Dirichlet series* $F'(s) = -\sum f(n)(\log n)n^{-s}$*.*

Applying Theorem 2.10 to $f(n) = \mathbf{1}(n)$, we have

$$-\zeta'(s) = \sum_{n=1}^{\infty} \frac{\log n}{n^s} \quad \text{for } \sigma > 1. \tag{2.12}$$

Taking logarithms in (2.7) and then noting that $-\log(1-z) = \sum_{k=1}^{\infty} z^k/k$ for $|z| < 1$, we yield

$$\log \zeta(s) = -\sum_p \log\left(1 - \frac{1}{p^s}\right) = \sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{ks}} \quad \text{for } \sigma > 1. \tag{2.13}$$

Differentiating (2.13), by means of Theorem 2.10, gives

$$-\frac{\zeta'}{\zeta}(s) = \sum_p \sum_{k=1}^{\infty} \frac{\log p}{p^{ks}} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \quad \text{for } \sigma > 1. \tag{2.14}$$

By virtue of Theorem 2.7, (2.7), (2.12) and (2.14), we obtain the identity

$$\log n = \sum_{d|n} \Lambda(d). \tag{2.15}$$

2.2.2 Summation formulae and some useful estimates

We turn to introduce two basic tools in analytic number theory: Abel's summation formula and the Euler–Maclaurin summation formula. Several estimates of arithmetic functions that will be used in proving our main results are also provided in this subsection, and some of them can be derived from those summation formulae.

Theorem 2.11 (Abel's summation formula). *Let (a_n) be a sequence of real or complex numbers. Define*

$$A(x) = \sum_{n \le x} a_n.$$

Let f be a continuously differentiable function on the interval $[y, x]$ where x and y are real numbers with $0 < y < x$. Then

$$\sum_{y < n \le x} a_n f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(u)f'(u)\,\mathrm{d}u. \tag{2.16}$$

Proof. Let $M = [y]$ and $N = [x]$. The idea of the proof is that we can write

$$\begin{aligned}\sum_{y<n\le x} a_n f(n) &= \sum_{n=M+1}^{N} (A(n) - A(n-1))f(n) \\ &= A(N)f(N) - A(M)f(M) + \sum_{n=M+1}^{N} A(n-1)(f(n-1) - f(n)). \quad (2.17)\end{aligned}$$

We refer to this method as *partial summation* or *summation by parts*. Since $A(n-1)$ is constant in the interval $[n-1,n)$, the sum on the right of (2.17) becomes

$$-\sum_{n=M+1}^{N} A(n-1)\int_{n-1}^{n} f'(u)\,\mathrm{d}u = -\sum_{n=M+1}^{N}\int_{n-1}^{n} A(u)f'(u)\,\mathrm{d}u = -\int_{M}^{N} A(u)f'(u)\,\mathrm{d}u. \tag{2.18}$$

Since $A(u)$ is constant for all $u\in[N,x]$, the first term in (2.17) is

$$A(x)f(x) - \int_{N}^{x} A(u)f'(u)\,\mathrm{d}u. \tag{2.19}$$

Similarly, the second term in (2.17) is

$$A(y)f(y) - \int_{M}^{y} A(u)f'(u)\,\mathrm{d}u. \tag{2.20}$$

The assertion follows by substituting (2.18), (2.19) and (2.20) in (2.17). □

Instead of using partial summation, it is much more convenient to derive (2.16) by means of the *Riemann–Stieltjes integral* (see Appendix A in [10]) to express

$$\sum_{y<n\le x} a_n f(n) = \int_{y}^{x} f(u)\,\mathrm{d}A(u). \tag{2.21}$$

The equation (2.16) immediately follows by integrating (2.21) by parts. Note that if x and y in (2.21) are both integers, then we may choose any real number in $[y-1,y)$ to be the left endpoint, and any real number in $[x,x+1)$ to be the right endpoint without affecting the value of the integral. In many cases, it is useful to consider the integral from $y-\varepsilon$ to x where $\varepsilon>0$ is arbitrarily small. For convenience, we shall write $\int_{y^-}^{x}$ instead of $\lim_{\varepsilon\to0^+}\int_{y-\varepsilon}^{x}$.

Now, we suppose that f is continuously differentiable on the interval $[y,x]$, and we let $a_n=1$ for all n so that $A(x)=[x]$. Using (2.21), we obtain

$$\sum_{y<n\le x} f(n) = \int_{y}^{x} f(u)\,\mathrm{d}[u] = \int_{y}^{x} f(u)\,\mathrm{d}u - \int_{y}^{x} f(u)\,\mathrm{d}\{u\}$$

Integrating the last integral by parts, we have

$$\int_{y}^{x} f(u)\,\mathrm{d}\{u\} = f(x)\{x\} - f(y)\{y\} - \int_{y}^{x}\{u\}\,\mathrm{d}f(u).$$

Since $f\in C^1([y,x])$, the integral on the right is $\int_y^x \{u\}f'(u)\,\mathrm{d}u$. Therefore

$$\sum_{y<n\le x} f(n) = \int_{y}^{x} f(u)\,\mathrm{d}u - f(x)\{x\} + f(y)\{y\} + \int_{y}^{x}\{u\}f'(u)\,\mathrm{d}u. \tag{2.22}$$

The last integral can actually be repeatedly integrated by parts if f is continuously differentiable up to higher order. To generalize the formula (2.22), we introduce the sequence of polynomials $(B_n(x))_{n\geq 0}$ satisfying the following conditions:

$$B_0(x) = 1, \tag{C1}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}B_n(x) = nB_{n-1}(x) \quad \text{for } n \geq 1, \tag{C2}$$

$$\int_0^1 B_n(x)\,\mathrm{d}x = 0 \quad \text{for } n \geq 1. \tag{C3}$$

The polynomials $B_n(x)$ are called the *Bernoulli polynomials.* Alternatively, we may represent $B_n(x)$ by the equation

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n \quad \text{for } |z| < 2\pi.$$

By integrating (C3), we find that

$$B_n(0) = B_n(1) \quad \text{for } n \geq 2. \tag{2.23}$$

We define the *Bernoulli numbers* as $B_n = B_n(0)$ for all $n \geq 0$. It can be deduced from (C2) that

$$B_n(x) = \sum_{r=0}^{n} \binom{n}{r} B_r x^{n-r} \quad \text{for } n \geq 0. \tag{2.24}$$

Taking $x = 1$ in (2.24) and then using (2.23), we obtain

$$B_n = \sum_{r=0}^{n} \binom{n}{r} B_r \quad \text{for } n \geq 2. \tag{2.25}$$

The equation (2.25) provides a recursion formula for computing B_{n-1} in terms of B_0, B_1, ..., B_{n-2}. The first five Bernoulli polynomials are shown below:

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6},$$
$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \quad B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30}.$$

These polynomials can be computed by using (2.24) together with (2.25).

Theorem 2.12 (Euler–Maclaurin summation formula)**.** *Let k be any positive integer. Suppose that a function f is k-times continuously differentiable on the interval $[y, x]$*

where x and y are real numbers with $0 < y < x$. Then

$$\sum_{y<n\le x} f(n) = \int_y^x f(u)\,\mathrm{d}u + \sum_{j=1}^{k} \frac{(-1)^j}{j!}\Big(B_j(\{x\})f^{(j-1)}(x) - B_j(\{y\})f^{(j-1)}(y)\Big) \tag{2.26}$$
$$- \frac{(-1)^k}{k!}\int_y^x B_k(\{u\})f^{(k)}(u)\,\mathrm{d}u.$$

Proof. We prove by induction on k. The case $k = 1$ is exactly the same as (2.22). Suppose that $f \in C^{k+1}([y,x])$. Integrating the last integral of (2.26) by parts yields

$$\frac{1}{k+1}\left(B_{k+1}(\{x\})f^{(k)}(x) - B_{k+1}(\{y\})f^{(k)}(y) - \int_y^x B_{k+1}(\{u\})f^{(k+1)}(u)\,\mathrm{d}u\right),$$

which proves the inductive step. □

We now give an approximation formula for $n!$ by applying the Euler–Maclaurin summation formula to $f(u) = \log u$ with $y = 1$, $x = n$ and $k = 2$ so that

$$\log n! = n\log n - n + \frac{1}{2}\log n + C + \frac{1}{12n} - \frac{1}{2}\int_n^\infty B_2(\{u\})u^{-2}\,\mathrm{d}u, \tag{2.27}$$

where

$$C = \frac{11}{12} + \frac{1}{2}\int_1^\infty B_2(\{u\})u^{-2}\,\mathrm{d}u.$$

Exponentiating (2.27) gives

$$n! = \left(\frac{n}{e}\right)^n \sqrt{n}e^C\left(1 + O\left(\frac{1}{n}\right)\right). \tag{2.28}$$

Next, we show that $C = \log\sqrt{2\pi}$ by using the product formula for the sine function:

$$\frac{\sin \pi z}{\pi} = z\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right) \tag{2.29}$$

(see §3.2 in Chapter 5 of [13] for the derivation of this formula). Taking $z = 1/2$ in (2.29) yields

$$\frac{2}{\pi} = \prod_{n=1}^{\infty}\left(1 - \frac{1}{4n^2}\right), \tag{2.30}$$

known as the *Wallis product.* The right-hand side of (2.30) is equal to

$$\lim_{n\to\infty}\frac{(2n+1)(2n)!^2}{2^{4n}(n!)^4}. \tag{2.31}$$

Combining (2.30) and (2.31) and then using (2.28), we find that

$$\frac{2}{\pi} = \lim_{n\to\infty} \frac{2n+1}{2^{4n}} \cdot \frac{\left((2n)^{2n} e^{-2n} \sqrt{2n} e^{C}\right)^2}{\left(n^n e^{-n} \sqrt{n} e^{C}\right)^4} \cdot \left(1 + O\left(\frac{1}{n}\right)\right)^2 = 4e^{-2C},$$

which establishes *Stirling's formula*:

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\left(1 + O\left(\frac{1}{n}\right)\right). \tag{2.32}$$

Next, we shall use the Euler–Maclaurin summation formula to show that the Riemann zeta function can be extended meromorphically beyond the half-plane $\sigma > 1$.

Theorem 2.13. *Let $x \geq 1$, $\sigma > 0$ and $s \neq 1$. Then*

$$\zeta(s) = \sum_{n\leq x} \frac{1}{n^s} + \frac{x^{1-s}}{s-1} + \frac{\{x\}}{x^s} - s\int_x^\infty \frac{\{u\}}{u^{s+1}}\,\mathrm{d}u. \tag{2.33}$$

Proof. We write $\zeta(s) = \sum_{n\leq x} n^{-s} + \sum_{n>x} n^{-s}$ for $\sigma > 1$. The latter sum is

$$\int_x^\infty u^{-s}\,\mathrm{d}u + \{x\}x^{-s} - s\int_x^\infty \{u\}u^{-s-1}\,\mathrm{d}u$$

by (2.22). The former integral is $x^{1-s}/(s-1)$. The latter integral converges absolutely for $\sigma > 0$, and uniformly for $\sigma \geq \delta > 0$. Also, it is an analytic function of s for $\sigma > 0$. Thus the equation (2.33) holds for $\sigma > 0$ by analytic continuation. $\square$

Taking $x = 1$ in (2.33), we obtain

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{u\}}{u^{s+1}}\,\mathrm{d}u \quad \text{for } \sigma > 0. \tag{2.34}$$

By using (2.26), we can generalize (2.34) to all positive integers k as

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} + \sum_{j=2}^{k} \frac{B_j}{j!} \prod_{l=0}^{j-2}(s+l) - \frac{s(s+1)\cdots(s+k-1)}{k!} \int_1^\infty \frac{B_k(\{u\})}{u^{s+k}}\,\mathrm{d}u.$$

We now notice that the first term on the right has a simple pole at $s = 1$ and $(s-1)\zeta(s) \to 1$ as $s \to 1$. Also, the integral is analytic for $\sigma > 1-k$. We can let k be arbitrarily large so that the function ζ can be analytically continued into the entire complex plane. The following corollary gathers the facts discussed above.

Corollary 2.14. *The Riemann zeta function has an analytic continuation into the entire complex plane except for a simple pole at $s = 1$ with residue* 1.

In the remaining part of this subsection, we collect some estimates of arithmetic functions needed in the proof of several results given in Chapter 3.

Theorem 2.15. *For $x \geq 1$,*

(i) $\sum_{n \leq x} \frac{1}{n} = \log x + \gamma + O\left(\frac{1}{x}\right)$,

(ii) $\sum_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \zeta(s) + O(x^{-\sigma}) \quad$ *for $\sigma > 0$ and $s \neq 1$,*

(iii) $\sum_{n > x} \frac{1}{n^s} \ll x^{1-\sigma} \quad$ *for $\sigma > 1$,*

(iv) $\sum_{n \leq x} n^s = \frac{x^{s+1}}{s+1} + O(x^{\sigma}) \quad$ *for $\sigma > 0$,*

where γ is the Euler–Mascheroni constant *(also called* Euler's constant*), defined by*

$$\gamma = 1 - \int_1^{\infty} \{u\} u^{-2} \, \mathrm{d}u = 0.5772156649\ldots.$$

Proof. The estimates (ii) and (iii) are immediate consequences of Theorem 2.13. Taking $f(n) = 1/n$ in (2.22) yields

$$\sum_{n \leq x} \frac{1}{n} = 1 + \int_1^x \frac{\mathrm{d}u}{u} - \frac{\{x\}}{x} - \int_1^x \frac{\{u\}}{u^2} \, \mathrm{d}u. \tag{2.35}$$

The former integral is $\log x$. We express the latter integral as $\int_1^{\infty} - \int_x^{\infty}$. The integral $\int_x^{\infty} \{u\} u^{-2} \, \mathrm{d}u$ and the penultimate term of (2.35) are $\ll 1/x$. This proves (i). To prove (iv), we take $f(n) = n^s$ in (2.22) to give

$$\sum_{n \leq x} n^s = 1 + \int_1^x u^s \, \mathrm{d}u - \{x\} x^s + s \int_1^x \{u\} u^{s-1} \, \mathrm{d}u.$$

The assertion follows by noting that the former integral is $x^{s+1}/(s+1) + O(1)$, and the last two terms are $\ll x^{\sigma}$. $\square$

The following theorem is due to Mertens in 1874 (see Theorem 2.7 in [10]).

Theorem 2.16 (Mertens)**.** *For $x \geq 2$,*

(i) $\sum_{p \leq x} \frac{1}{p} = \log \log x + \beta + O\left(\frac{1}{\log x}\right)$,

(ii) $\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} = e^{\gamma} \log x + O(1)$,

where $\beta = \gamma - \sum_p \sum_{k=2}^{\infty} (kp^k)^{-1}$.

We now let $A(x) = \sum_{p \le x} 1/p$. Then, by Theorem 2.16(i), we may express $A(x) = \log\log x + \beta + R(x)$ where $R(x) \ll 1/\log x$. Thus we have

$$\sum_{p>x} \frac{1}{p\log p} = \int_x^\infty \frac{\mathrm{d}A(u)}{\log u} = \int_x^\infty \frac{\mathrm{d}\log\log u}{\log u} + \int_x^\infty \frac{\mathrm{d}R(u)}{\log u}.$$

The penultimate integral is $\int_x^\infty \left(u(\log u)^2\right)^{-1} \mathrm{d}u = 1/\log x$. The last integral is $-R(x)/\log x + \int_x^\infty \left(u(\log u)^2\right)^{-1} R(u)\,\mathrm{d}u \ll 1/(\log x)^2$. Therefore

$$\sum_{p>x} \frac{1}{p\log p} \ll \frac{1}{\log x}. \tag{2.36}$$

Letting $x \to \infty$ in (2.36), we see that the sum on the left approaches 0 and thus we obtain the following corollary.

Corollary 2.17. *The series $\sum_p (p\log p)^{-1}$ is convergent.*

The following theorem gives a minimal order for $\varphi(n)$ (see Theorem 2.9 in [10]).

Theorem 2.18. *For $n \ge 3$,*

$$\varphi(n) \ge \frac{n}{\log\log n}\left(e^{-\gamma} + O\left(\frac{1}{\log\log n}\right)\right).$$

Theorem 2.19. *For $x \ge 2$,*

$$\sum_{n\le x} \frac{n}{\varphi(n)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} x + O(\log x).$$

Proof. We first observe that $n/\varphi(n) = \sum_{d|n} \mu(d)^2/\varphi(d)$. To see this, we let $\mathcal{Q}$ denote the set of square-free integers, and suppose that n has a prime factorization $n = \prod_{p|n} p^{\alpha_p}$. We find that

$$\sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} = \sum_{\substack{d|n \\ d\in\mathcal{Q}}} \frac{1}{\varphi(d)} = \prod_{p|n}\left(1 + \frac{1}{p-1}\right) = \prod_{p|n} \frac{p^{\alpha_p}}{p^{\alpha_p}\left(1 - \dfrac{1}{p}\right)} = \frac{n}{\varphi(n)}.$$

Then for $x \ge 2$,

$$\begin{aligned}\sum_{n\le x} \frac{n}{\varphi(n)} = \sum_{n\le x}\sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} = \sum_{d\le x} \frac{\mu(d)^2}{\varphi(d)}\left[\frac{x}{d}\right] &= \sum_{d\le x} \frac{\mu(d)^2}{\varphi(d)}\left(\frac{x}{d} + O(1)\right) \\ &= x\sum_{d\le x} \frac{\mu(d)^2}{d\varphi(d)} + O\left(\sum_{\substack{d\le x \\ d\in\mathcal{Q}}} \frac{1}{\varphi(d)}\right).\end{aligned}$$

The last term is $\ll \prod_{p\leq x}(1-1/p)^{-1} \ll \log x$ by Theorem 2.16(ii). We now split the sum of the penultimate term as $\sum_{d=1}^{\infty} - \sum_{d>x}$. The latter sum is

$$\sum_{d>x} \frac{\mu(d)^2}{d\varphi(d)} \ll \sum_{d>x} \frac{1}{d\varphi(d)} \ll \sum_{d>x} \frac{\log\log d}{d^2} \ll \frac{\log\log x}{x}$$

by Theorem 2.18 and Theorem 2.15(iii). The former sum is

$$\sum_{d=1}^{\infty} \frac{\mu(d)^2}{d\varphi(d)} = \prod_p \left(1 + \frac{1}{p(p-1)}\right) = \prod_p \frac{p^6-1}{p(p^2-1)(p^3-1)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)}.$$

The proof is now complete. $\square$

For any arithmetic function f, the *variance* of f is defined as

$$\sigma_f^2 = \frac{1}{x}\sum_{n\leq x}(f(n)-\mu_f)^2,$$

where μ_f is the *mean value* of f, given by

$$\mu_f = \frac{1}{x}\sum_{n\leq x} f(n).$$

We now estimate the mean values and the variances of the functions ω and Ω. We first give an estimate for the mean value of ω by considering

$$\sum_{n\leq x}\omega(n) = \sum_{n\leq x}\sum_{p\mid n} 1 = \sum_{p\leq x}\left[\frac{x}{p}\right] = x\sum_{p\leq x}\frac{1}{p} + O(\pi(x)).$$

Here $\pi(x) = \sum_{p\leq x} 1$, the number of primes not exceeding x, called the *prime-counting function*. Using Theorem 2.16(i) and the crude bound $\pi(x) \leq \sum_{n\leq x} 1 \ll x$ (we shall give some better approximations to $\pi(x)$ in the next subsection), we obtain

$$\sum_{n\leq x}\omega(n) = x\log\log x + O(x). \tag{2.37}$$

The equation (2.37) can be interpreted probabilistically that if $n \leq x$ is chosen uniformly at random, then we would expect n to have approximately $\log\log x$ distinct prime divisors on average. The following theorem gives a bound for the variance of ω, due to Turán in 1934 (see Theorem 2.12 in [10]).

Theorem 2.20 (Turán)**.** *For $x \geq 3$,*

$$\sum_{n\leq x}(\omega(n)-\log\log x)^2 \ll x\log\log x. \tag{2.38}$$

We now consider

$$\sum_{n\le x}(\Omega(n)-\omega(n)) = \sum_{n\le x}\sum_{\substack{p^a|n\\ a\ge 2}} 1 = \sum_{p}\sum_{a\ge 2}\left[\frac{x}{p^a}\right] \ll x\sum_{p}\frac{1}{p(p-1)} \ll x.$$

Also, we have

$$\sum_{n\le x}(\Omega(n)-\omega(n))^2 = \sum_{p}\sum_{a\ge 2}(2a-1)\left[\frac{x}{p^a}\right] + \sum_{\substack{p_1,p_2\\ p_1\ne p_2}}\sum_{a,b\ge 2}\left[\frac{x}{p_1^a p_2^b}\right].$$

The former term on the right is $\ll x\sum_p(3p-1)/(p(p-1)^2) \ll x$. The latter term on the right is $\ll x\left(\sum_p 1/(p^2-p)\right)^2 \ll x$. Thus we have proved that

$$\sum_{n\le x}(\Omega(n)-\omega(n))^k \ll x \quad \text{if } k=1,2.$$

This deduces that (2.37) and (2.38) also hold if $\omega(n)$ is replaced by $\Omega(n)$. Consequently, we obtain the following statement, due to Hardy and Ramanujan in 1917.

Corollary 2.21 (Hardy–Ramanujan)**.** *For almost all $n \le x$,*

$$|\omega(n) - \log\log x| \ll (\log\log x)^{1/2+\varepsilon}$$

for every $\varepsilon > 0$. The same result also holds if $\omega(n)$ is replaced by $\Omega(n)$.

2.2.3 Distribution of prime numbers

Because of the randomness of primes, it seems useless to find any explicit formula for $\pi(x)$ which can be computed effectively. Instead, it is much more sensible to concern the asymptotic behaviour of $\pi(x)$ when x becomes large. Legendre conjectured in 1798 that $\pi(x) \approx x/(A\log x + B)$ for some constants A and B, and then refined in 1808 by proposing that $\pi(x) \approx x/(\log x - 1.08366)$ (we shall discuss later that Legendre was misled about his constant when $x \to \infty$). This suggests that

$$\frac{\pi(x)}{x/\log x} \to 1 \quad \text{as } x\to\infty. \tag{2.39}$$

It is usual to represent such behaviour by writing $f(x) \sim g(x)$ to indicate that $f(x)/g(x) \to 1$ as x tends to its limit. Thus we can reformulate (2.39) as

$$\pi(x) \sim \frac{x}{\log x}. \tag{2.40}$$

This statement is known as the *prime number theorem*. It was first proved independently by Hadamard and de la Vallée Poussin in 1896, based on the non-vanishing of $\zeta(s)$ on the line $\sigma = 1$ (see Chapter 3 of [14] for more details). A quantitative form of the prime number theorem is given by

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right). \tag{2.41}$$

Chebyshev was the first one who made an important contribution to proving the prime number theorem. In 1848, he showed that

$$\liminf_{x\to\infty} \frac{\pi(x)}{x/\log x} \le 1 \le \limsup_{x\to\infty} \frac{\pi(x)}{x/\log x}. \tag{2.42}$$

However, he was unable to prove the existence of the limit (2.39).

Proof of (2.42). Let $\limsup \pi(x)/(x/\log x) = a$, and let $\varepsilon > 0$ be arbitrary. Then there exists x_0 such that $\pi(x) \le (a+\varepsilon)x/\log x$ for all $x \ge x_0$. We see that

$$\sum_{p\le x} \frac{1}{p} = \int_{2^-}^{x} \frac{\mathrm{d}\pi(u)}{u} \le (a+\varepsilon)\int_{x_0}^{x} \frac{\mathrm{d}u}{u\log u} + O(1) = (a+\varepsilon)\log\log x + O_\varepsilon(1).$$

By Theorem 2.16(i) we obtain $a + \varepsilon \ge 1$, and so $a \ge 1$ since ε can be arbitrarily small. Similarly, we also have $\liminf \pi(x)/(x/\log x) \le 1$. □

In 1850, Chebyshev introduced

$$\vartheta(x) = \sum_{p\le x} \log p \quad \text{and} \quad \psi(x) = \sum_{n\le x} \Lambda(n).$$

He proved that for $x \ge 2$,

$$Ax + O(\log x) \le \psi(x) \le \frac{6}{5}Ax + O\left((\log x)^2\right),$$

where $A = \log(2^{1/2}3^{1/3}5^{1/5}30^{-1/30}) = 0.9212920229\ldots$.

Theorem 2.22 (Chebyshev)**.** *For $x \ge 2$, $\psi(x) \asymp x$.*

Here we write $f \asymp g$ to indicate that f and g have the same order of magnitude, i.e., both $f \ll g$ and $g \ll f$ hold. We next consider

$$\psi(x) - \vartheta(x) = \sum_{\substack{p^k\le x\\ k\ge 2}} \log p = \sum_{2\le k\le \log x/\log 2} \vartheta(x^{1/k}) \ll x^{1/2} + x^{1/3}\log x \ll x^{1/2}. \tag{2.43}$$

Also, we have
$$\pi(x)=\int_{2-}^{x}\frac{\mathrm{d}\vartheta(u)}{\log u}=\frac{\vartheta(x)}{\log x}+\int_{2-}^{x}\frac{\vartheta(u)}{u(\log u)^2}\,\mathrm{d}u.$$
The penultimate term is $\psi(x)/\log x+O(x^{1/2}/\log x)$ by (2.43). Since $\vartheta(u)\le\psi(u)\ll u$, the last integral is $\ll\int_{2-}^{x}(\log u)^{-2}\,\mathrm{d}u\ll x/(\log x)^2$. Therefore
$$\pi(x)=\frac{\psi(x)}{\log x}+O\left(\frac{x}{(\log x)^2}\right). \tag{2.44}$$
The following statement can be deduced from Theorem 2.22, (2.43) and (2.44).

Corollary 2.23. *For $x\ge 2$, $\vartheta(x)\asymp x$ and $\pi(x)\asymp x/\log x$.*

We now prove that the correct value of *Legendre's constant*, the number 1.08366, must be exactly 1. Suppose that there is a constant A such that
$$\pi(x)=\frac{x}{\log x-A}+o\left(\frac{x}{(\log x)^2}\right). \tag{2.45}$$
The equation (2.45) can be reformulated as
$$\pi(x)=\frac{x}{\log x}+(A+o(1))\frac{x}{(\log x)^2}. \tag{2.46}$$
Let us consider
$$\vartheta(x)=\int_{2-}^{x}\log u\,\mathrm{d}\pi(u)=\pi(x)\log x-\int_{2-}^{x}\frac{\pi(u)}{u}\,\mathrm{d}u.$$
By using (2.46), we find that the penultimate term is $x+(A+o(1))x/\log x$, and the last integral is
$$\int_{2-}^{x}\frac{\mathrm{d}u}{\log u}+(A+o(1))\int_{2-}^{x}\frac{\mathrm{d}u}{(\log u)^2}=(1+o(1))\frac{x}{\log x}.$$
Thus we obtain
$$\psi(x)=x+(A-1+o(1))\frac{x}{\log x}. \tag{2.47}$$
by means of (2.43). By using (2.47), we see that
$$\int_{2-}^{x}\frac{\psi(u)}{u^2}\,\mathrm{d}u=\int_{2-}^{x}\frac{\mathrm{d}u}{u}+(A-1+o(1))\int_{2-}^{x}\frac{\mathrm{d}u}{u\log u}=\log x+(A-1+o(1))\log\log x. \tag{2.48}$$
We now want to derive a formula for the leftmost integral in a different way. Consider
$$\sum_{n\le x}\log n=\sum_{n\le x}\sum_{d\mid n}\Lambda(d)=\sum_{d\le x}\Lambda(d)\left[\frac{x}{d}\right]=x\sum_{d\le x}\frac{\Lambda(d)}{d}+O(\psi(x)).$$

Note that the first equality is obtained by (2.15). The error term is $\ll x$ by Theorem 2.22. The leftmost sum is $x \log x - x + O(\log x)$ by using (2.22). Thus we have

$$\sum_{n \le x} \frac{\Lambda(n)}{n} = \log x + O(1). \tag{2.49}$$

The sum on the left is $\int_{2^-}^{x} u^{-1}\,\mathrm{d}\psi(u) = \int_{2^-}^{x} \psi(u)u^{-2}\,\mathrm{d}u + O(1)$. It follows that

$$\int_{2^-}^{x} \frac{\psi(u)}{u^2}\,\mathrm{d}u = \log x + O(1). \tag{2.50}$$

By comparing (2.48) with (2.50), we conclude that $A = 1$, correcting Legendre's constant.

In 1792–1973, Gauss observed that the density of primes in the neighborhood of x is approximately $1/\log x$. This led him to propose that a better approximation to $\pi(x)$ is

$$\pi(x) \sim \int_{2}^{x} \frac{\mathrm{d}u}{\log u} := \mathrm{li}(x). \tag{2.51}$$

We call the function $\mathrm{li}(x)$ the *logarithmic integral*. A sharper quantitative form of the prime number theorem is

$$\pi(x) = \mathrm{li}(x) + O\left(x \exp\left(-C\sqrt{\log x}\right)\right) \tag{2.52}$$

for some absolute constant $C > 0$ (see Theorem 6.9 in [10]). By integrating the integral in (2.51) by parts N times, we yield

$$\mathrm{li}(x) = \frac{x}{\log x}\left(\sum_{n=1}^{N} \frac{(n-1)!}{(\log x)^{n-1}} + O\left(\frac{1}{(\log x)^N}\right)\right).$$

This shows that the error term in (2.41) cannot be sharper than $O(x(\log x)^{-2})$ because $\mathrm{li}(x)$ contains the term $x(\log x)^{-2}$.

For every integer $k \ge 1$, we denote by $\pi_k(x)$ the number of positive integers $n \le x$ with $\omega(n) = k$. It can be shown that

$$\pi_k(x) \sim \frac{x(\log\log x)^{k-1}}{(k-1)!\log x}. \tag{2.53}$$

The same result also holds if $\omega(n)$ is replaced by $\Omega(n)$ (see Theorem 437 in [4]). Note that the asymptotic relation (2.53) in the case $k = 1$ is exactly the same as (2.40).

2.2.4 Primes in arithmetic progressions

The infinitude of primes in the arithmetic progression of the form $qk+a$ with $(a,q)=1$ and $k \geq 0$ was first studied by Dirichlet in 1837. He introduced the *Dirichlet characters* modulo q, which are defined as the extensions of group homomorphisms $\widetilde{\chi} : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$ to all $n \in \mathbb{Z}$ by setting

$$\chi(n) = \begin{cases} \widetilde{\chi}(n \bmod q) & \text{if } (n,q)=1, \\ 0 & \text{if } (n,q)>1. \end{cases}$$

As above, we see that the multiplicative group $(\mathbb{Z}/q\mathbb{Z})^\times$ of reduced residue classes modulo q has exactly $\varphi(q)$ Dirichlet characters, each of which is completely multiplicative and has period q. That is, $\chi(mn)=\chi(m)\chi(n)$ for all m,n and $\chi(n+q)=\chi(n)$ for all n. Also, we have the orthogonality relation

$$\frac{1}{\varphi(q)} \sum_{\chi \in \widehat{(\mathbb{Z}/q\mathbb{Z})^\times}} \overline{\chi(a)}\chi(n) = \begin{cases} 1 & \text{if } n \equiv a \pmod{q}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.54}$$

Here $\widehat{(\mathbb{Z}/q\mathbb{Z})^\times}$ denotes the set of all Dirichlet characters modulo q.

Let χ be any Dirichlet character modulo q. The *Dirichlet L-function* is defined for $\sigma > 1$ by

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

Since χ is completely multiplicative, it follows from (2.6) that

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \quad \text{for } \sigma > 1.$$

Let χ_0 denote the *principal* (or *trivial*) character, which is given by $\chi_0(n)=1$ if $(n,q)=1$ and 0 otherwise. Observe that

$$L(s,\chi_0) = \sum_{\substack{n=1 \\ (n,q)=1}}^{\infty} \frac{1}{n^s} = \zeta(s) \prod_{p \mid q} \left(1 - \frac{1}{p^s}\right) \quad \text{for } \sigma > 1.$$

Then, by Corollary 2.14, the function $L(s,\chi_0)$ is analytic for $\sigma > 0$ except for a simple pole at $s=1$ with residue $\varphi(q)/q$. For $\chi \neq \chi_0$, we have $\left|\sum_{n \leq x} \chi(n)\right| \leq \varphi(q)$ for all $x \geq 1$. Note that $\sum f(n)n^{-s}$ converges for $\sigma > 0$ if the partial sums $\sum_{n \leq x} f(n)$ are bounded (this can be proved easily by taking $s_0 = 0$ in Lemma 2.8 and then letting $a \to \infty$ in (2.11)). Therefore $L(s,\chi)$ is analytic for $\sigma > 0$ if $\chi \neq \chi_0$. By applying

Theorem 2.10 to $f(n) = \chi(n)$, we have

$$-L'(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)\log n}{n^s} \quad \text{for } \sigma > 1. \tag{2.55}$$

In fact, the equation (2.55) holds for $\sigma > 0$ if $\chi \neq \chi_0$.

Suppose that $(a,q) = 1$. The existence of infinitely many primes in the arithmetic progression $qk + a$ can be proved as a consequence of the asymptotic estimate

$$\sum_{\substack{n \le x \\ n \equiv a \pmod q}} \frac{\Lambda(n)}{n} = \frac{1}{\varphi(q)} \log x + O_q(1). \tag{2.56}$$

The sum on the left can be expressed as

$$\sum_{\substack{p \le x \\ p \equiv a \pmod q}} \frac{\log p}{p} + \sum_{\substack{p^k \le x \\ k \ge 2 \\ p \equiv a \pmod q}} \frac{\log p}{p^k}.$$

Note that the latter sum is $\le \sum_p \log p/(p^2 - p) \ll 1$. Thus we have

$$\sum_{\substack{p \le x \\ p \equiv a \pmod q}} \frac{\log p}{p} = \frac{1}{\varphi(q)} \log x + O_q(1).$$

This implies that there are infinitely many primes $p \equiv a \pmod q$ since $\log x \to \infty$ as $x \to \infty$. To prove (2.56), we use the orthogonality relation (2.54) to obtain

$$\sum_{\substack{n \le x \\ n \equiv a \pmod q}} \frac{\Lambda(n)}{n} = \frac{1}{\varphi(q)} \sum_{n \le x} \frac{\Lambda(n)}{n} \sum_{\chi} \overline{\chi(a)}\chi(n) = \frac{1}{\varphi(q)} \sum_{\chi} \overline{\chi(a)} \sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n}.$$

The contribution of the principal character is

$$\sum_{n \le x} \frac{\chi_0(n)\Lambda(n)}{n} = \sum_{\substack{n \le x \\ (n,q)=1}} \frac{\Lambda(n)}{n} = \sum_{n \le x} \frac{\Lambda(n)}{n} - \sum_{\substack{p^k \le x \\ p \mid q}} \frac{\log p}{p^k} = \log x + O_q(1)$$

by using (2.49) and noting that the last sum is $\le \sum_{p \mid q} \log p/(p-1) \ll_q 1$. It is therefore sufficient to show that

$$\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} \ll_\chi 1 \quad \text{if } \chi \neq \chi_0. \tag{2.57}$$

To prove (2.57), we suppose that $\chi \neq \chi_0$ and then consider

$$\sum_{n \le x} \frac{\chi(n)\log n}{n} = -L'(1,\chi) - \sum_{n>x} \frac{\chi(n)\log n}{n}. \tag{2.58}$$

Let $S(x) = \sum_{n \le x} \chi(n)$. The sum on the right of (2.58) is

$$\int_x^\infty \frac{\log u}{u}\,\mathrm{d}S(u) = -\frac{S(x)\log x}{x} - \int_x^\infty \frac{(1-\log u)S(u)}{u^2}\,\mathrm{d}u \ll_\chi \frac{\log x}{x}$$

by noting that $S(x) \ll_\chi 1$. By using (2.15), we have the sum on the left of (2.58) is

$$\sum_{n \le x} \frac{\chi(n)}{n} \sum_{d|n} \Lambda(d) = \sum_{qd \le x} \frac{\chi(qd)\Lambda(d)}{qd} = \sum_{d \le x} \frac{\chi(d)\Lambda(d)}{d} \sum_{q \le x/d} \frac{\chi(q)}{q}.$$

We express the rightmost sum as $L(1,\chi) - \sum_{q>x/d} \chi(q)/q$. Note that

$$\sum_{q>x/d} \frac{\chi(q)}{q} = \int_{x/d}^\infty \frac{\mathrm{d}S(u)}{u} = -\frac{S(x/d)}{x/d} + \int_{x/d}^\infty \frac{S(u)}{u^2}\,\mathrm{d}u \ll_\chi \frac{d}{x}.$$

Therefore

$$\sum_{n \le x} \frac{\chi(n)\log n}{n} = L(1,\chi)\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} + O_\chi\left(\frac{\psi(x)}{x}\right). \tag{2.59}$$

The error term is $\ll_\chi 1$ by Theorem 2.22. Combining (2.58) and (2.59) yield

$$L(1,\chi)\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} \ll_\chi 1 \quad \text{if } \chi \neq \chi_0. \tag{2.60}$$

We now notice that the factor $L(1,\chi)$ in (2.60) can be dropped to obtain (2.57) provided $L(1,\chi) \neq 0$ for $\chi \neq \chi_0$. Thus the following theorem plays an essential role in the proof of Dirichlet's theorem (see Theorem 4.9 in [10]).

Theorem 2.24 (Dirichlet)**.** *If χ is a non-principal Dirichlet character, then $L(1,\chi) \neq 0$.*

Let us introduce

$$\pi(x;q,a) = \sum_{\substack{p \le x \\ p \equiv a \pmod q}} 1 \quad \text{and} \quad \psi(x;q,a) = \sum_{\substack{n \le x \\ n \equiv a \pmod q}} \Lambda(n).$$

That is, the function $\pi(x;q,a)$ counts the number of primes not exceeding x in the arithmetic progression $qk + a$ where $k \ge 0$. The *prime number theorem for arithmetic progressions* asserts that if $(a,q) = 1$, then

$$\pi(x;q,a) \sim \frac{x}{\varphi(q)\log x} \quad \text{as } x \to \infty.$$

The following theorem represents the quantitative form of the prime number theorem for arithmetic progressions (see §11.3 in [10]).

Theorem 2.25 (The Siegel–Walfisz theorem). *For any $A > 0$ with $1 \leq q \leq (\log x)^A$ and $(a, q) = 1$,*

$$\psi(x; q, a) = \frac{x}{\varphi(q)} + O_A\left(x \exp\left(-C\sqrt{\log x}\right)\right)$$

for some absolute constant $C > 0$.

We also have the following theorem as a consequence of Theorem 2.25.

Theorem 2.26. *Under the assumption of Theorem 2.25, we have*

$$\pi(x; q, a) = \frac{\mathrm{li}(x)}{\varphi(q)} + O_A\left(x \exp\left(-C\sqrt{\log x}\right)\right)$$

for some absolute constant $C > 0$.

2.2.5 Generalized Gauss sums

A *generalized Gauss sum* is an exponential sum of the form

$$S_k(a, q) = \sum_{n=1}^{q} \mathbf{e}\left(\frac{an^k}{q}\right), \tag{2.61}$$

where $k \geq 2$, $q \geq 1$ and a is an integer relatively prime to q. Gauss showed that for any positive integer q,

$$\sum_{n=1}^{q} \mathbf{e}\left(\frac{n^2}{q}\right) = \frac{1 + i^{-q}}{1 + i^{-1}}\sqrt{q} = \begin{cases} (1+i)\sqrt{q} & \text{if } q \equiv 0 \pmod 4, \\ \sqrt{q} & \text{if } q \equiv 1 \pmod 4, \\ 0 & \text{if } q \equiv 2 \pmod 4, \\ i\sqrt{q} & \text{if } q \equiv 3 \pmod 4. \end{cases}$$

This shows that $S_2(1, q) \ll \sqrt{q}$. More generally, we have the following estimate which holds for all positive integers k and q (see Theorem 6 in [7]).

Theorem 2.27. *Let k and q be positive integers. If $(a, q) = 1$, then $S_k(a, q) \ll q^{1-1/k}$.*

The above estimate also holds if the sum in (2.61) runs through all positive integers n not exceeding q with $(n, q) = 1$. To see this, we consider

$$\sum_{\substack{n=1 \\ (n,q)=1}}^{q} \mathbf{e}\left(\frac{an^k}{q}\right) = \sum_{n=1}^{q} \mathbf{e}\left(\frac{an^k}{q}\right) \sum_{d|(n,q)} \mu(d) = \sum_{d|q} \mu(d) \sum_{\substack{n=1 \\ n\equiv 0 \pmod d}}^{q} \mathbf{e}\left(\frac{an^k}{q}\right)$$

by using (2.10). The rightmost sum is

$$\sum_{\substack{n=1 \\ n\equiv 0 \pmod{d}}}^{q} \mathbf{e}\left(\frac{an^k}{q}\right) = \sum_{m=1}^{q/d} \mathbf{e}\left(\frac{ad^k m^k}{q}\right). \tag{2.62}$$

For each prime divisor p of q, let l_p denote the unique integer such that $p^{l_p} \mid q$ but $p^{l_p+1} \nmid q$, and let $\alpha_p = \min\{k, l_p\}$. It is easy to see that $(ap^k, q) = p^{\alpha_p}$. Consequently, we have

$$(ad^k, q) = \prod_{p|d} p^{\alpha_p} := \mathcal{P}_d$$

for any square-free divisor d of q. If we let $b = ad^k/\mathcal{P}_d$ and $\widetilde{q} = q/\mathcal{P}_d$, then $(b, \widetilde{q}) = 1$ and the sum (2.62) becomes

$$\sum_{m=1}^{(\mathcal{P}_d/d)\widetilde{q}} \mathbf{e}\left(\frac{bm^k}{\widetilde{q}}\right) = \frac{\mathcal{P}_d}{d} \sum_{m=1}^{\widetilde{q}} \mathbf{e}\left(\frac{bm^k}{\widetilde{q}}\right).$$

The equality above follows from the fact that the function $e(\theta/q)$ has period q. Therefore

$$\sum_{\substack{n=1 \\ (n,q)=1}}^{q} \mathbf{e}\left(\frac{an^k}{q}\right) = \sum_{d|q} \frac{\mu(d)\mathcal{P}_d}{d} S_k(b, \widetilde{q}). \tag{2.63}$$

The sum on the right of (2.63) is

$$\ll q^{1-1/k} \sum_{d|q} \frac{\mu(d)}{d} (\mathcal{P}_d)^{1/k} = q^{1-1/k} \prod_{p|q} \left(1 - p^{-(1-\alpha_p/k)}\right) \ll q^{1-1/k}$$

by Theorem 2.27. This establishes the following result.

Corollary 2.28. *Let k and q be any positive integers. If $(a, q) = 1$, then*

$$\sum_{\substack{n=1 \\ (n,q)=1}}^{q} \mathbf{e}\left(\frac{an^k}{q}\right) \ll q^{1-1/k}.$$

Remark. By applying the equation (2.63) to the case $k = 1$, we obtain the identity

$$c_q(a) := \sum_{\substack{n=1 \\ (n,q)=1}}^{q} \mathbf{e}\left(\frac{an}{q}\right) = \mu(q),$$

provided $(a, q) = 1$. The function $c_q(a)$ is known as *Ramanujan's sum.*

CHAPTER III

UNIFORM DISTRIBUTION OF ARITHMETIC MEAN OF kTH POWER OF PRIME DIVISORS

In the previous chapter, we discussed several criteria for determining whether or not a given sequence is uniformly distributed. One may use Weyl's criterion (Theorem 2.2) to obtain the uniform distribution of the sequences $(\rho_k(n))$ and $(\widetilde{\rho}_k(n))$ for any fixed k by showing that for every integer $a \neq 0$ the asymptotic size of the associated exponential sums $\sum_{n \leq N} \mathbf{e}(au_k(n))$ is $o(N)$ as $N \to \infty$, where $u_k(n) = \rho_k(n)$ or $\widetilde{\rho}_k(n)$. In fact, we shall prove the following statement, based on the same techniques as described in [2].

Theorem 3.1. *Let $k \geq 2$. Suppose that N is sufficiently large. Then*

$$\sum_{n=1}^{N} \mathbf{e}(a\rho_k(n)) \ll \frac{|a|N}{(\log\log N)^{1/k}} \tag{3.1}$$

holds for every integer $a \neq 0$. The same estimate also holds if $\rho_k(n)$ is replaced by $\widetilde{\rho}_k(n)$.

Another criterion is to compute whether or not the discrepancy $D(N)$ of a given sequence approaches 0 as $N \to \infty$ (see Theorem 2.5). For any positive integer k, we denote by $D_k(N)$ and $\widetilde{D}_k(N)$ the discrepancies of the sequences $(\rho_k(n))$ and $(\widetilde{\rho}_k(n))$, respectively. The following theorem gives a bound for the discrepancies $D_k(N)$ and $\widetilde{D}_k(N)$.

Theorem 3.2. *Let $k \geq 2$. Suppose that N is sufficiently large. Then*

$$D_k(N) \ll \frac{\log\log\log N \log\log\log\log N}{(\sqrt{\log\log N})^{1+1/k}}. \tag{3.2}$$

The same estimate also holds if $D_k(N)$ is replaced by $\widetilde{D}_k(N)$.

By dividing both sides of (3.1) by N and then letting $N \to \infty$, or by taking $N \to \infty$ in (3.2), we conclude that the sequences $(\rho_k(n))$ and $(\widetilde{\rho}_k(n))$ are uniformly distributed for all $k \geq 2$, and hence for all positive integers k by the previous results (1.1) and (1.2) from [2].

To prove the preceding theorems, we require the following auxiliary lemmas.

Lemma 3.3. *Let $0 < \lambda < 1$. Suppose that $x \geq 1$ is a real number, and n is a positive integer with $(1-\lambda)x \leq n \leq (1+\lambda)x$. Then*

$$\frac{x^{n-1}}{(n-1)!} \ll_\alpha \frac{1}{\sqrt{x}} \exp(x) \exp\left(-\frac{(x-n)^2}{\alpha x}\right)$$

holds for any fixed real number $\alpha \geq 1/(\log 4 - 1)$.

Proof. Let $\xi = n/x$. The assumption together with Stirling's formula (2.32) give

$$\frac{x^{n-1}}{(n-1)!} \ll \frac{x^n}{n!} = \frac{1}{\sqrt{2\pi n}} \exp(n)\xi^{-n} \left(1 + O\left(\frac{1}{n}\right)\right)$$
$$\ll \frac{1}{\sqrt{x}} \exp(x) \exp\left(-x\xi \log \xi\right) \exp\left(-\frac{x}{2}(1-\xi)^2\right) \exp\left(-\frac{x}{2}(1-\xi^2)\right).$$

It can be shown that for any fixed real number $\alpha \geq 1/(\log 4 - 1)$, the function

$$f_\alpha(\xi) = \frac{1}{2}(\xi^2 - 1) - \left(\frac{1}{2} - \frac{1}{\alpha}\right)(\xi - 1)^2 - \xi \log \xi$$

defined for $\xi \in (0, 2)$ has a unique global maximum at $\xi = 1$ with $f_\alpha(1) = 0$. This implies that

$$\exp\left(-x\xi \log \xi\right) \exp\left(-\frac{x}{2}(1-\xi)^2\right) \exp\left(-\frac{x}{2}(1-\xi^2)\right) \ll_\alpha \exp\left(-\frac{x}{\alpha}(1-\xi)^2\right).$$

The proof is now complete. $\square$

Lemma 3.4. *Let $0 < \lambda < 1$ and $x \geq 2$. Then there exists a constant A depending on λ such that*

$$\sum_{n \leq x} \frac{n^{1-\lambda}}{\varphi(n)} = Ax^{1-\lambda} + O(1).$$

Proof. It can be derived from (2.16) that

$$\sum_{n \leq x} \frac{n^{1-\lambda}}{\varphi(n)} = x^{-\lambda} \sum_{n \leq x} \frac{n}{\varphi(n)} + \lambda \int_{1^-}^{x} \left(\sum_{n \leq u} \frac{n}{\varphi(n)}\right) \frac{1}{u^{1+\lambda}}\, \mathrm{d}u.$$

Using Theorem 2.19, we can express the right-hand side as

$$\alpha x^{1-\lambda} + O\left(\frac{\log x}{x^\lambda}\right) + \alpha\lambda \int_{1^-}^{x} u^{-\lambda}\, \mathrm{d}u + O\left(\int_{1^-}^{x} \frac{\log u}{u^{1+\lambda}}\, \mathrm{d}u\right),$$

where $\alpha = \zeta(2)\zeta(3)/\zeta(6)$. The penultimate integral is $x^{1-\lambda}/(1-\lambda) + O(1)$, and both error terms are $\ll 1$. Thus the sum in question is $\alpha x^{1-\lambda}/(1-\lambda) + O(1)$. $\square$

For any real number α, we define $\sigma_\alpha(n) = \sum_{d|n} d^\alpha$, called the *divisor function.*

Lemma 3.5. *Let $0 < \lambda < 1$, $x \geq 3$ and $\sqrt{x} \leq y \leq x$. Then*

$$\sum_{x \leq n \leq x+y} \frac{\sigma_{1-\lambda}(n)}{\varphi(n)} \ll \frac{y \log\log x}{(\sqrt{x})^{1+\lambda}}.$$

Proof. Let us consider

$$\begin{aligned}
\sum_{x\le n\le x+y}\frac{\sigma_{1-\lambda}(n)}{\varphi(n)} &= \sum_{x\le n\le x+y}\frac{1}{\varphi(n)}\sum_{d\mid n} d^{1-\lambda}\\
&\le \sum_{x\le n\le x+y}\frac{1}{\varphi(n)}\sum_{\substack{d\mid n\\ d\le\sqrt{2x}}}\left(d^{1-\lambda}+\left(\frac{n}{d}\right)^{1-\lambda}\right)\\
&= \sum_{d\le\sqrt{2x}} d^{1-\lambda}\sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{1}{\varphi(n)} + \sum_{d\le\sqrt{2x}} d^{-(1-\lambda)}\sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{n^{1-\lambda}}{\varphi(n)}.
\end{aligned} \tag{3.3}$$

Using Theorem 2.18, we have

$$\sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{1}{\varphi(n)} \ll \sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{\log\log n}{n} \ll \frac{\log\log x}{d}\sum_{x/d\le q\le (x+y)/d}\frac{1}{q}.$$

The last sum above is $\ll \log(1+y/x) \ll y/x$ by means of Theorem 2.15(i). Thus the former term in (3.3) is

$$\ll \frac{y\log\log x}{x}\sum_{d\le\sqrt{2x}} d^{-\lambda} \ll \frac{y\log\log x}{x}(\sqrt{x})^{1-\lambda} = \frac{y\log\log x}{(\sqrt{x})^{1+\lambda}},$$

where the second bound is obtained by Theorem 2.15(ii). Again, by Theorem 2.18, we yield

$$\sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{n^{1-\lambda}}{\varphi(n)} \ll \sum_{\substack{x\le n\le x+y\\ n\equiv 0 \pmod d}}\frac{\log\log n}{n^{\lambda}} \ll \frac{\log\log x}{d^{\lambda}}\sum_{x/d\le q\le (x+y)/d} q^{-\lambda}.$$

Using Theorem 2.15(ii), we have the last sum above is

$$\ll \left(\frac{x}{d}\right)^{1-\lambda}\left(\left(1+\frac{y}{x}\right)^{1-\lambda}-1\right) \ll \frac{y}{x^{\lambda}}\left(\frac{1}{d}\right)^{1-\lambda}.$$

Hence the latter term in (3.3) is

$$\ll \frac{y\log\log x}{x^{\lambda}}\sum_{d\le\sqrt{2x}} d^{-(2-\lambda)} \ll \frac{y\log\log x}{x^{\lambda}}(\sqrt{x})^{-(1-\lambda)} = \frac{y\log\log x}{(\sqrt{x})^{1+\lambda}}$$

by virtue of Theorem 2.15(ii). The proof is now complete. □

We are now in a position to prove Theorem 3.1 and Theorem 3.2.

3.1 Proof of Theorem 3.1

We proceed analogously to the proof of Theorem 1 in [2]. Let P_n denote the largest prime factor of $n \geq 2$ and set $P_1 = 1$. Let $Q = N^{1/u}$, where

$$u = \frac{2 \log\log\log N}{\log\log\log\log N}. \tag{3.4}$$

We define four sets of positive integers as follows:

- $\mathcal{E}_1 = \{n \leq N \,:\, P_n \leq Q\}$,
- $\mathcal{E}_2 = \left\{n \leq N \,:\, n \notin \mathcal{E}_1 \text{ and } {P_n}^2 \mid n\right\}$,
- $\mathcal{E}_3 = \{n \leq N \,:\, |\omega(n) - \log\log N| > \delta \log\log N\}$, where $\delta > 0$ is a sufficiently small absolute constant,
- $\mathcal{E}_4 = \{n \leq N/\log\log N\}$.

As we defined above, it can be easily deduced that if n is a positive integer not exceeding N with $n \notin \bigcup_{i=1}^{4} \mathcal{E}_i$, then the integer n can be uniquely written as $n = mP_n$. In this situation, we have $m < N/Q$ and $P_n \in \mathcal{L}_m$, where

$$L_m = \max\left\{Q, P_m, \frac{N}{m \log\log N}\right\} \quad \text{and} \quad \mathcal{L}_m = (L_m, N/m].$$

Next, we define

- $\mathcal{E}_5 = \left\{n \leq N \,:\, n \notin \bigcup_{i=1}^{4} \mathcal{E}_i \text{ and } L_m = Q\right\}$,
- $\mathcal{E}_6 = \left\{n \leq N \,:\, n \notin \bigcup_{i=1}^{5} \mathcal{E}_i \text{ and } L_m = P_m\right\}$.

Let $\mathcal{E} = \bigcup_{i=1}^{6} \mathcal{E}_i$. It was already shown in [2] that

$$\#\mathcal{E} \ll \frac{N}{\log\log N}. \tag{3.5}$$

However, we shall intentionally repeat the argument in the next subsection for the sake of completeness.

3.1.1 The exceptional set $\mathcal{E}$

To prove (3.5), it suffices to show that $\#\mathcal{E}_i \ll N/\log\log N$ holds for each $1 \leq i \leq 6$.

Let $\mathcal{S}(x, y)$ denote the set of positive integers not exceeding x all of whose primes factors are not exceeding y. That is, $\mathcal{S}(x, y) = \{n \leq x \,:\, P_n \leq y\}$. Such integers n are called *y-smooth*. Let $\Psi(x,y)$ denote the cardinality of the set $\mathcal{S}(x,y)$. In 1930, Dickman proved that for any fixed real number $u \geq 0$ there exists a function $\rho(u) > 0$ such that

$$\Psi(x, x^{1/u}) \sim x\rho(u) \quad \text{as } x \to \infty \tag{3.6}$$

(see Theorem 7.2 in [10]). The function $\rho(u)$ is known as the *Dickman–de Bruijn rho function*. It can be shown that

$$\rho(u) = u^{-u(1+o(1))} \quad \text{as } u \to \infty \tag{3.7}$$

(see [3] or [5] for more details). Applying (3.6) and (3.7), we obtain

$$\#\mathcal{E}_1 = \Psi(N, N^{1/u}) \ll Nu^{-u(1+o(1))} \ll \frac{N}{\log\log N},$$

where u is defined as (3.4). As for $\mathcal{E}_2$, it can be easily seen that

$$\#\mathcal{E}_2 \leq \sum_{p>Q} \left[\frac{N}{p^2}\right] \leq N \sum_{p>Q} \frac{1}{p^2} \leq N \sum_{n>Q} \frac{1}{n^2} \ll \frac{N}{Q} \ll \frac{N}{\log\log N},$$

where we have used Theorem 2.15(iii). By using Turán's estimate (2.38) we have

$$\#\mathcal{E}_3 \leq (\delta \log\log N)^{-2} \sum_{n \leq N} (\omega(n) - \log\log N)^2 \ll \frac{N}{\log\log N}.$$

It is clear from the definition that $\#\mathcal{E}_4 \ll N/\log\log N$. To estimate $\#\mathcal{E}_5$, we note that $N/(m \log\log N) \leq L_m = Q < N/m$, and hence

$$m \in \left[\frac{N}{Q\log\log N}, \frac{N}{Q}\right) := \mathscr{M}.$$

Let $\pi(\mathcal{L}_m)$ denote the number of primes in the interval $\mathcal{L}_m$. Since each $n \in \mathcal{E}_5$ can be uniquely factorized as mP_n for some prime $P_n \in \mathcal{L}_m = (Q, N/m]$ and some $m \in \mathscr{M}$, we yield

$$\#\mathcal{E}_5 \leq \sum_{m \in \mathscr{M}} \pi(\mathcal{L}_m) \leq \sum_{m \in \mathscr{M}} \pi\left(\frac{N}{m}\right) \ll \sum_{m \in \mathscr{M}} \frac{N}{m \log(N/m)},$$

where the last bound follows from Corollary 2.23. The last sum above is

$$\begin{aligned} &\leq \frac{N}{\log Q} \sum_{m \in \mathscr{M}} \frac{1}{m} \ll \frac{N}{\log Q}\left(\log\left(\frac{N}{Q}\right) - \log\left(\frac{N}{Q\log\log N}\right)\right) \\ &= \frac{N \log\log\log N}{\log Q} = \frac{2N(\log\log\log N)^2}{\log N \log\log\log\log N} \ll \frac{N}{\log\log N}, \end{aligned}$$

where we have used Theorem 2.15(i). As for $\mathcal{E}_6$, we observe that

$$P_m = L_m \geq \frac{N}{m \log\log N} \geq \frac{P_n}{\log\log N} > \frac{Q}{\log\log N}.$$

If we let $p_1 = P_m$ and $p_2 = P_n$, then we see that the set $\mathcal{E}_6$ contains positive integers $n \leq N$ composed of two distinct primes p_1 and p_2 such that $p_1 > Q/\log\log N$ and $p_1 < p_2 \leq p_1 \log\log N$. Fix p_1 and p_2. The number of positive integers not exceeding N which are divisible by p_1 and p_2 is $[N/(p_1p_2)] \ll N/(p_1p_2)$. Therefore

$$\#\mathcal{E}_6 \ll N \sum_{p_1 > Q/\log\log N} \frac{1}{p_1} \sum_{p_1 < p_2 \leq p_1 \log\log N} \frac{1}{p_2}.$$

Using Theorem 2.16(i), we have the inner sum on the right is

$$\log\log(p_1 \log\log N) - \log\log p_1 + O\left(\frac{1}{\log p_1}\right) \ll \frac{\log\log\log N}{\log p_1}.$$

Thus we obtain

$$\#\mathcal{E}_6 \ll N \log\log\log N \sum_{p > Q/\log\log N} \frac{1}{p \log p} \ll \frac{N \log\log\log N}{\log(Q/\log\log N)}$$

by using the estimate (2.36). The last term is $\ll N \log\log\log N/\log Q \ll N/\log\log N$, which proves (3.5).

3.1.2 The remaining n

We now define $\mathcal{N}$ to be the set of positive integers $n \leq N$ which do not belong to the exceptional set $\mathcal{E}$. The estimate (3.5) gives

$$\sum_{n=1}^{N} \mathbf{e}(a\rho_k(n)) = \sum_{n \in \mathcal{N}} \mathbf{e}(a\rho_k(n)) + O\left(\frac{N}{\log\log N}\right). \tag{3.8}$$

Moreover, we note that each $n \in \mathcal{N}$ can be uniquely written in the form $n = mP_n$, where

$$P_n \in \mathcal{L}_m = \left(\frac{N}{m \log\log N}, \frac{N}{m}\right].$$

Let $\mathcal{M}$ be the set of all acceptable values for m. That is, the set $\mathcal{M}$ contains an integer m for which $m \leq N$ and $n = mP_n$ for some $n \in \mathcal{N}$. For every integer $v \geq 1$, we define

$$\mathcal{N}_v = \{n \in \mathcal{N} : \omega(n) = v\} \quad \text{and} \quad \mathcal{M}_v = \{m \in \mathcal{M} : \omega(m) = v\}.$$

Let

$$h = (1-\delta)\log\log N \quad \text{and} \quad H = (1+\delta)\log\log N.$$

It can be seen that for $h \le v \le H$, there is a one-to-one correspondence between the element n of $\mathcal{N}_v$ and the element P_n of $\mathcal{L}_m$ for some $m \in \mathcal{M}_{v-1}$. Consequently, for $h \le v \le H$, we obtain

$$\#\mathcal{N}_v = \sum_{m \in \mathcal{M}_{v-1}} \pi(\mathcal{L}_m). \tag{3.9}$$

As discussed above, we have

$$\begin{aligned}
\sum_{n\in\mathcal{N}} \mathbf{e}(a\rho_k(n)) &= \sum_{h\le v\le H}\sum_{n\in\mathcal{N}_v} \mathbf{e}(a\rho_k(n)) = \sum_{h\le v\le H}\sum_{m\in\mathcal{M}_{v-1}}\sum_{P_n\in\mathcal{L}_m} \mathbf{e}(a\rho_k(mP_n)) \\
&= \sum_{h\le v\le H}\sum_{m\in\mathcal{M}_{v-1}} \mathbf{e}\left(\frac{a\rho_k(m)(v-1)}{v}\right)\sum_{P_n\in\mathcal{L}_m} \mathbf{e}\left(\frac{a{P_n}^k}{v}\right) \\
&\ll \sum_{h\le v\le H}\sum_{m\in\mathcal{M}_{v-1}}\sum_{P_n\in\mathcal{L}_m} \mathbf{e}\left(\frac{a{P_n}^k}{v}\right).
\end{aligned} \tag{3.10}$$

To estimate the innermost sum on the right of (3.10), we let $b = a/(a,v)$ and $v_a = v/(a,v)$. Note that $(b, v_a) = 1$ and for $\varepsilon > 0$,

$$v_a \le v \le H \le 2\log\log N \ll (\log Q)^\varepsilon \le \left(\log\left(\frac{N}{m\log\log N}\right)\right)^\varepsilon.$$

Applying Theorem 2.26 together with (2.52) and noting that $P_n \equiv d \pmod{v_a}$ implies ${P_n}^k \equiv d^k \pmod{v_a}$, we deduce that for $x \ge Q$,

$$\sum_{\substack{P_n\le x \\ P_n\equiv d \pmod{v_a}}} \mathbf{e}\left(\frac{b{P_n}^k}{v_a}\right) = \frac{\pi(x)}{\varphi(v_a)}\mathbf{e}\left(\frac{bd^k}{v_a}\right) + O\left(x\exp\left(-C\sqrt{\log x}\right)\right)$$

for some absolute constant $C > 0$. Summing over $d \in (\mathbb{Z}/v_a\mathbb{Z})^\times$, we find that

$$\sum_{P_n\le x} \mathbf{e}\left(\frac{b{P_n}^k}{v_a}\right) = \frac{\pi(x)}{\varphi(v_a)}\sum_{\substack{1\le d\le v_a \\ (d,v_a)=1}} \mathbf{e}\left(\frac{bd^k}{v_a}\right) + O\left(v_a x\exp\left(-C\sqrt{\log x}\right)\right).$$

Consequently, we have

$$\sum_{P_n\in\mathcal{L}_m} \mathbf{e}\left(\frac{b{P_n}^k}{v_a}\right) = \frac{\pi(\mathcal{L}_m)}{\varphi(v_a)}\sum_{\substack{1\le d\le v_a \\ (d,v_a)=1}} \mathbf{e}\left(\frac{bd^k}{v_a}\right) + O\left(\frac{v_a N}{m\exp\left(C\sqrt{\log Q}\right)}\right). \tag{3.11}$$

The sum on the right is $\ll (v_a)^{1-1/k}$ by using Corollary 2.28. Inserting (3.11) in (3.10) and using (3.9), we find that the right-hand side of (3.10) is

$$\ll \sum_{h\le v\le H} \frac{(v_a)^{1-1/k}}{\varphi(v_a)} \#\mathcal{N}_v + \frac{N}{\exp\left(C\sqrt{\log Q}\right)} \sum_{h\le v\le H} v \sum_{m<N} \frac{1}{m}. \tag{3.12}$$

The penultimate sum is $\ll (\log\log N)^2$ by Theorem 2.15(iv), and the last sum is $\ll \log N$ by Theorem 2.15(i). We now observe that

$$\varphi(v_a) = \frac{v}{(a,v)} \prod_{p\mid \frac{v}{(a,v)}} \left(1-\frac{1}{p}\right) \ge \frac{v}{(a,v)} \prod_{p\mid v} \left(1-\frac{1}{p}\right) = \frac{\varphi(v)}{(a,v)} \ge \frac{\varphi(v)}{|a|}. \tag{3.13}$$

Thus the former term in (3.12) is $\le |a| \sum_{h\le v\le H} \left(v^{1-1/k}/\varphi(v)\right) \#\mathcal{N}_v$, and the latter term in (3.12) is $\ll N\log N(\log\log N)^2 \exp\left(-C\sqrt{\log Q}\right) \ll N/\log\log N$. Using the trivial bound $\#\mathcal{N}_v \le \pi_v(N)$ together with the asymptotic estimate (2.53), we obtain

$$\sum_{n\in\mathcal{N}} \mathbf{e}(a\rho_k(n)) \ll \frac{|a|N}{\log N} \sum_{h\le v\le H} \frac{v^{1-1/k}}{\varphi(v)} \frac{(\log\log N)^{v-1}}{(v-1)!} + \frac{N}{\log\log N}. \tag{3.14}$$

By applying Lemma 3.3, we have

$$\frac{(\log\log N)^{v-1}}{(v-1)!} \ll \frac{\log N}{\sqrt{\log\log N}} \exp\left(-\frac{(v-\log\log N)^2}{3\log\log N}\right).$$

For every integer j, we let $\mathscr{I}_j$ be the closed interval defined by

$$\mathscr{I}_j = \left[\log\log N + (j-1/2)\sqrt{\log\log N}, \log\log N + (j+1/2)\sqrt{\log\log N}\right].$$

We find that each integer v in the interval $[h,H]$ is contained in the interval $\mathscr{I}_j$ for some integer j with $|j| \le \delta\sqrt{\log\log N}$. Hence the sum on the right of (3.14) is

$$\ll \frac{\log N}{\sqrt{\log\log N}} \sum_{|j|\le\delta\sqrt{\log\log N}} \exp\left(-\frac{j^2}{3}\right) \sum_{v\in\mathscr{I}_j} \frac{v^{1-1/k}}{\varphi(v)}.$$

The inner sum above is

$$\ll (\log\log N)^{1-1/k} \left(\left(1+\frac{1}{\sqrt{\log\log N}}\right)^{1-1/k} - 1\right) \ll (\log\log N)^{1/2-1/k}$$

by using Lemma 3.4. Note that $\sum_{j=-\infty}^{\infty} \exp(-j^2/3) \ll 1$. Thus we have

$$\sum_{h\le v\le H} \frac{v^{1-1/k}}{\varphi(v)} \frac{(\log\log N)^{v-1}}{(v-1)!} \ll \frac{\log N}{(\log\log N)^{1/k}}. \tag{3.15}$$

Combining together the estimates (3.8), (3.14) and (3.15), we obtain (3.1). If we replace $\rho_k(n)$ with $\widetilde{\rho}_k(n)$, then the function $\omega(n)$ in the proof will be replaced by $\Omega(n)$ and this will not affect the proof because the estimates (2.38) and (2.53) also hold for $\Omega(n)$.

3.2 Proof of Theorem 3.2

By the Erdős–Turán inequality, we find that for any positive integer M,

$$D_k(N) \ll \frac{1}{M} + \frac{1}{N}\sum_{a=1}^{M}\frac{1}{a}\left|\sum_{n=1}^{N}\mathbf{e}(a\rho_k(n))\right|. \tag{3.16}$$

Note that the replacement of $D_k(N)$ with $\widetilde{D}_k(N)$ does not affect the proof since the estimate (3.1) also holds for $\widetilde{\rho}_k(n)$. The latter term on the right of (3.16) is

$$\ll \frac{1}{N}\sum_{a=1}^{M}\frac{1}{a}\sum_{h\le v\le H}\frac{{v_a}^{1-1/k}}{\varphi(v_a)}\#\mathcal{N}_v + \frac{\log M}{\log\log N} \tag{3.17}$$

by (3.12) and Theorem 2.15(i). Recalling that $\varphi(v_a) \ge \varphi(v)/(a,v)$ from (3.13) and then interchanging the order of summation, we see that the former term of (3.17) is

$$\le \frac{1}{N}\sum_{h\le v\le H}\frac{v^{1-1/k}}{\varphi(v)}\#\mathcal{N}_v\sum_{a=1}^{M}\frac{1}{a}(a,v)^{1/k}. \tag{3.18}$$

The inner sum in (3.18) is

$$\sum_{d\mid v} d^{1/k}\sum_{\substack{1\le a\le M\\(a,v)=d}}\frac{1}{a} \le \sum_{d\mid v} d^{1/k}\sum_{\substack{1\le a\le M\\a\equiv 0 \pmod{d}}}\frac{1}{a} = \sum_{d\mid v}\left(\frac{1}{d}\right)^{1-1/k}\sum_{1\le b\le M/d}\frac{1}{b}.$$

The last sum is $\ll \log M$ by Theorem 2.15(i). Thus the right-hand side of (3.18) is

$$\ll \frac{\log M}{N}\sum_{h\le v\le H}\frac{\sigma_{1-1/k}(v)}{\varphi(v)}\#\mathcal{N}_v. \tag{3.19}$$

Arguing as in the proof of Theorem 3.1 and then applying Lemma 3.5, we find that the sum in (3.19) is

$$\ll \frac{N}{\sqrt{\log\log N}}\sum_{|j|\le\delta\sqrt{\log\log N}}\exp\left(-\frac{j^2}{3}\right)\sum_{v\in\mathscr{I}_j}\frac{\sigma_{1-1/k}(v)}{\varphi(v)} \ll \frac{N}{\sqrt{\log\log N}}\cdot\frac{\log\log\log\log N}{(\sqrt{\log\log N})^{1/k}}. \tag{3.20}$$

Combining together the estimates (3.16)–(3.20), we obtain

$$D_k(N) \ll \frac{1}{M} + \frac{\log M \log\log\log\log N}{(\sqrt{\log\log N})^{1+1/k}} + \frac{\log M}{\log\log N}.$$

The assertion follows by choosing M such that $M - 1 < \left(\sqrt{\log\log N}\right)^{1+1/k} \leq M$.

REFERENCES

[1] T. M. Apostol, *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1976.

[2] W. D. Banks, M. Z. Garaev, F. Luca, and I. E. Shparlinski, Uniform distribution of the fractional part of the average prime divisor, *Forum Math.* **17** (2005), no. 6, 885–901.

[3] A. Granville, Smooth numbers: computational number theory and beyond. In *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*, Math. Sci. Res. Inst. Publ., Vol. 44, Buhler, J. P. and Stevenhagen P., eds., Cambridge Univ. Press, Cambridge, pp. 267–323, 2008.

[4] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers. Sixth Edition*, Revised by D. R. Heath-Brown and J. H. Silverman, Oxford University Press, Oxford, 2008.

[5] A. Hildebrand and G. Tenenbaum, Integers without large prime factors, *J. Théor. Nombres Bordeaux.* **5** (1993), no. 2, 411–484.

[6] I. Kátai and F. Luca, Uniform distribution modulo 1 of the harmonic prime factor of an integer, *Unif. Distrib. Theory.* **4** (2009), no. 2, 115–132.

[7] N. M. Korobov, *Exponential Sums and Their Applications*, Springer Science+Business Media, Dordrecht, 1992.

[8] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley-Interscience, New York-London-Sydney, 1974.

[9] F. Luca and I. E. Shparlinski, Uniformity of distribution modulo 1 of the geometric mean prime divisor, *Bol. Soc. Mat. Mexicana (3).* **12** (2006), no. 2, 155–163.

[10] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge Studies in Advanced Mathematics, 97, Cambridge University Press, Cambridge, 2007.

[11] M. R. Murty, *Problems in Analytic Number Theory. Second Edition*, Graduate Texts in Mathematics, 206, Readings in Mathematics, Springer, New York, 2008.

[12] H. Niederreiter and A. Winterhof, *Applied Number Theory*. Springer, Cham, 2015.

[13] E. M. Stein and R. Shakarchi, *Complex Analysis*, Princeton Lectures in Analysis, 2, Princeton University Press, Princeton, NJ, 2003.

[14] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function. Second Edition*, Edited and with a preface by D. R. Heath-Brown, The Clarendon Press, Oxford University Press, New York, 1986.

CURRICULUM VITAE

Chitpanu Chairatanagul was born on November 5, 1995 in Bangkok, Thailand. He received his Bachelor of Science degree with First Class Honors after graduating the Honors Program in Mathematics from Chulalongkorn University, Thailand in 2018.