Rough Numbers and Variations on the Erdős--Kac Theorem

Kai Fan
Dartmouth College, steve.fan.gr@dartmouth.edu

ROUGH NUMBERS AND VARIATIONS ON THE ERDŐS–KAC THEOREM

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

in

Mathematics

by Kai (Steve) Fan

Guarini School of Graduate and Advanced Studies
Dartmouth College
Hanover, New Hampshire

October 1, 2023

Examining Committee:

Carl Pomerance, Chair

Paul Pollack

John Voight

Dana Williams

F. Jon Kull, Ph.D.
Dean of the Guarini School of Graduate and Advanced Studies

Abstract

The study of arithmetic functions, functions with domain $\mathbb{N}$ and codomain $\mathbb{C}$, has been a central topic in number theory. This work is dedicated to the study of the distribution of arithmetic functions of great interest in analytic and probabilistic number theory.

In the first part, we study the distribution of positive integers free of prime factors less than or equal to any given real number $y \geq 1$. Denoting by $\Phi(x, y)$ the count of these numbers up to any given $x \geq y$, we show, by a combination of analytic methods and sieves, that $\Phi(x, y) < 0.6x/\log y$ holds uniformly for all $3 \leq y \leq \sqrt{x}$, improving upon an earlier result of the author in the same range. We also prove numerically explicit estimates of the de Bruijn type for $\Phi(x, y)$ which are applicable in wide ranges.

In the second part, we turn to the topic of weighted Erdős–Kac theorems for general additive functions. Our results concern the distribution of additive functions $f(n)$ weighted by nonnegative multiplicative functions $\alpha(n)$ in a wide class. Building on the moment method of Granville, Soundararajan, Khan, Milinovich and Subedi, we establish uniform asymptotic formulas for the moments of $f(n)$ with a suitable growth rate. Our method also enables us to prove a qualitative result on the moments which extends a theorem of Delange and Halberstam on the moments of additive functions. As a consequence, we obtain a weighted analogue of the Kubilius–Shapiro theorem with simple and interesting applications to the Ramanujan tau function and Euler's totient function, the latter of which generalizes an old result of Erdős and Pomerance which shows that as an arithmetic function, the total number of prime factors of values of Euler's totient function satisfies a Gaussian law.

Preface

The present work summarizes my thesis research as a doctoral student at Dartmouth College. Explored herein are two intriguing topics from analytic and probabilistic number theory: the distribution of rough numbers and Erdős–Kac type theorems on the distribution of values of additive functions, both of which enjoy a rich history and find applications in other branches of number theory. Chapter 1 of this work provides historical backgrounds and motivations for the study of these two topics. The rest of the chapters are devoted to the main results and their proofs, with Chapter 2 focusing on explicit estimates for the number $\Phi(x, y)$ of y-rough numbers not exceeding x and Chapter 3 on weighted variants of the Erdős–Kac theorem. Also found in Chapter 3 are two applications to certain arithmetic functions of special interests.

My research on rough numbers was inspired by the problem of finding an explicit constant $C > 0$, as small as possible, for which the inequality $\Phi(x, y) \leq Cx/\log y$ holds for all $1 < y \leq x$, a problem proposed by Kevin Ford and communicated to me by my advisor Carl Pomerance. Back then I was in my third year of graduate study, having a hard time finding the right research topics and advisor for my thesis. I was hoping to work with Carl in analytic number theory, but I was aware that he had retired from his position. Nevertheless, we still kept in touch with each other and met occasionally to discuss number theory. During one of our meetings, Carl brought up the inequality on $\Phi(x, y)$ above that he heard from Ford, who observed that one could take $C = 2$ and wondered whether smaller constants were also permissible. This little problem immediately seized my attention, in that I had a vague feeling that I might be able to make progress based on what I read about sieve theory. After a weekend of investigation of this problem, I succeeded in solving it by combining the arithmetic large sieve with some explicit estimates on prime numbers. With Carl's help, I was able to simplify my proof and make it clearer and more compact than it originally was, and this soon led to the publication of my paper [27], which laid the foundation for my thesis research on this subject.

My interest in the distribution of additive functions, and especially in the Erdős–Kac type theorems, grew out of my independent study of this topic during my third year of graduate study. I was fortunate to obtain the firsthand knowledge about Selberg's central limit theorem for the Riemann zeta-function in an online graduate course taught by Kannan Soundararajan, which piqued my interest in exploring the limiting distribution of arithmetic functions. Then came an unexpected turn of events in the summer of 2022, which completely changed the path of my graduate research. It started with my encounter with the paper [38] on a weighted version of the Erdős–Kac theorem. After examining the paper, I arrived at the conclusion that their argument for the divisor functions could be generalized to treat a wide class of multiplicative functions, which led me to prove a few theorems, record them in a draft, and send it to Carl for feedback. Unaware of the dramatic event that would only unfold days after, I received an encouraging message from him and was very happy to hear that he was interested in my theorems. Soon we had a meeting during which I talked briefly about my results. When we were about to call it a day, Carl asked "How about you writing a thesis on rough numbers and Erdős–Kac?" Honestly, I was confused because at that point I had found no faculty member in the department to advisor my research on these topics. But then he continued "By working with me." In retrospect, it is hardly an exaggeration that the present work would not have existed had this event

not occurred.

I have fortunately received kind help and support from many wonderful people. I am most grateful to my thesis advisor Carl Pomerance for recognizing my talent and potential and kindly taking me as his student despite his retirement. I have been privileged to work closely with him and learn from him for the past year. Carl's considerable expertise in analytic number theory and unique perspective on problem solving have been influential in guiding me to become a qualified researcher.

In the same breath, I am deeply indebted to my parents for their constant love, encouragement, and support. It was their firm faith in me during the most difficult periods in my life that gave me the confidence and courage to persevere and work miracles along the unusual path without losing my way.

I am much obliged to the members of my thesis committee, Carl Pomerance, Paul Pollack, John Voight, and Dana Williams, for their helpful comments and suggestions from which this work has benefited greatly. My special thanks go to Pollack for his careful scrutiny of a draft of my paper on which the second part of this work is based and for his cheerful encouragement and constructive feedback. In addition, I would like to thank Eran Assaf for helpful discussions on holomorphic cusp forms related to my work.

Finally, I would like to express my gratitude to my cohort, Lizzie Buchanan, Richard Haburcak, Grant Molnar, and Alexander Wilson, and to Mits Kobayashi, for their invaluable help and support during my academic journey at Dartmouth. I surely will miss the joyful time we spent together as well as the fun math problems with which Grant used to delight me at the afternoon tea.

Hanover, New Hampshire, USA Kai (Steve) Fan

July 2023

Contents

Chapter 1

Introduction

In this chapter, we give a thorough introduction to the two themes into which we shall delve later: the distribution of rough numbers and the weighted Erdős–Kac theorems. Our intention is to provide a bird's-eye view of these themes but reserve the technical material to later chapters.

Starting in Section 1.1, we introduce the notion of y-rough numbers and its counting function $\Phi(x, y)$, followed by a discussion of some previous work on the asymptotic formulas and explicit estimates for $\Phi(x, y)$ and then a preview of the main results which will be restated formally and established later in Chapter 2. Moving on to the second theme in Section 1.2, we discuss the history of the celebrated Erdős–Kac theorem on the distribution of the number of distinct prime factors of a positive integer as well as some recent work on its weighted variants, and we provide a simple probabilistic heuristic for this theorem. Rather than formulate our main results here, we opt to present two intriguing applications of them following our discussion on the historical background while reserving the formal statements of our results and their proofs to Chapter 3.

Section 1.1

The Distribution of Rough Numbers

Let $x \geq y > 1$. Throughout this chapter, we shall always write $u = u(x, y) := \log x/\log y$, and the letters p and q will always denote primes. We say that a positive integer n is *y-rough* if all the prime divisors of n are greater than y. Let $\Phi(x, y)$ denote the number of y-rough numbers up to x. Explicitly, we have

$$\Phi(x, y) = \sum_{\substack{n \leq x \\ P^-(n) > y}} 1,$$

where $P^-(n)$ denotes the least prime divisor of n, with the convention that $P^-(1) = \infty$. When $1 \leq u \leq 2$, or equivalently when $\sqrt{x} \leq y \leq x$, we simply have $\Phi(x, y) = \pi(x) - \pi(y) + 1$, where $\pi(x)$ demotes the number of primes up to x. The function $\Phi(x, y)$ is closely related to the sieve of Eratosthenes, an ancient algorithm for finding primes, and $\Phi(x, y)$ has been extensively studied by mathematicians. A simple application of the inclusion-exclusion principle enables us to write

$$\Phi(x, y) = \sum_{d \mid P(y)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor, \tag{1.1.1}$$

where $\lfloor a \rfloor$ is the integer part of a for any $a \in \mathbb{R}$, μ is the Möbius function, and $P(y)$ denotes the product of primes up to y. If y is relatively small in comparison with x, say $y = x^{o(1)}$, the above formula can be used to obtain

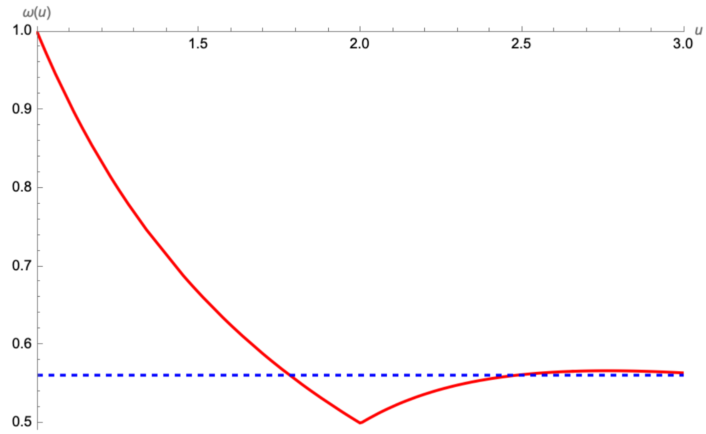
$$\Phi(x, y) \sim x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma} x}{\log y} \tag{1.1.2}$$

as $x \to \infty$, where $\gamma = 0.5772156...$ is the Euler–Mascheroni constant. This is also suggested by the heuristic based on the assumption that divisibility by a small prime

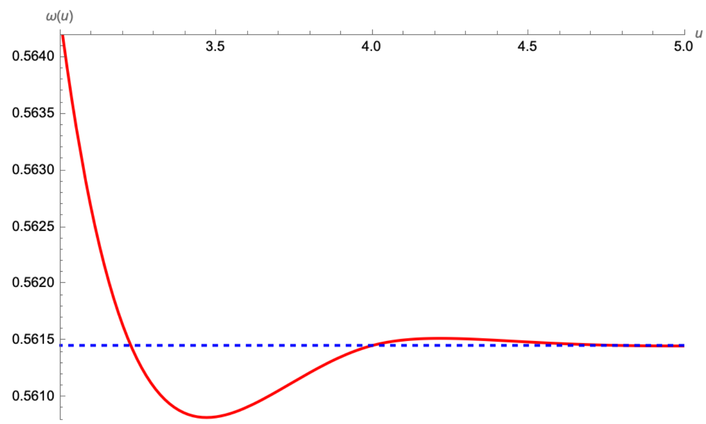
$p \le y$ and divisibility by a different prime $q \le y$ are close to being independent. However, it turns out that (1.1.2) does not hold uniformly, as already exemplified by the base case $1 \le u < 2$. The issue lies in the fact that the assumption on the independence of divisibility fails to hold for primes that are relatively large compared to x. For instance, if $y \in [\sqrt{x}, x]$ is large, and if $p_1, p_2, p_3 \in (y/8, y]$ are three distinct primes, whose existence is assured by Bertrand's postulate, then for a randomly chosen positive integer $n \le x$, the events $p_i \mid n$ $(1 \le i \le 3)$ are strongly correlated, in the sense that they cannot occur simultaneously. For this reason, the charming heuristic for (1.1.2) no longer makes sense when y is relatively large in comparison to x, and one would thus expect a heavy dependence of the asymptotic behavior of $\Phi(x, y)$ on the relation between x and y, or equivalently, on the values of u.

In 1937, Buchstab [7] showed that for any fixed $u > 1$, one has $\Phi(x, y) \sim \omega(u)x/\log y$ as $x \to \infty$, where $\omega(u)$ is defined to be the unique continuous solution to the delay differential equation $(u\omega(u))' = \omega(u-1)$ for $u \ge 2$, subject to the initial value condition $\omega(u) = 1/u$ for $u \in [1, 2]$. Comparing this result with the asymptotic formula obtained from (1.1.1), one would expect that $\omega(u) \to e^{-\gamma}$ as $u \to \infty$. Indeed, it can be shown [56, Corollary III.6.5] that $\omega(u) = e^{-\gamma} + O(u^{-u/2})$ for $u \ge 1$. Moreover, it is known that $\omega(u)$ oscillates above and below $e^{-\gamma}$ infinitely often. The following graphs generated by Mathematica provide a snapshot of the behavior of $\omega(u)$ on $[1, 7]$.

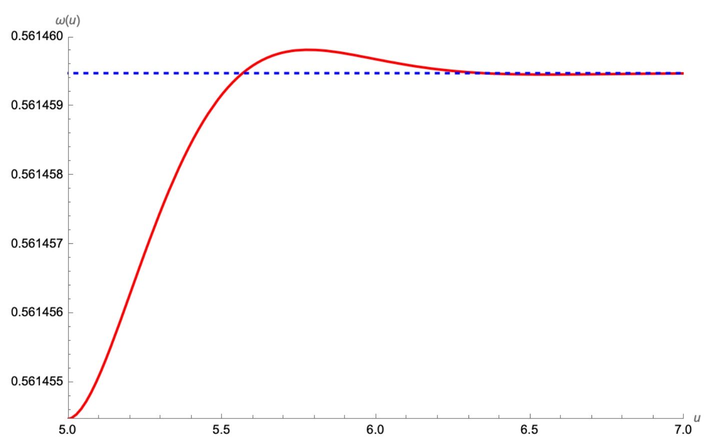
Figure 1.1: The Buchstab Function $\omega(u)$ on $[1, 7]$



(a) $u \in [1, 3]$



(b) $u \in [3, 5]$



(c) $u \in [5, 7]$

Buchstab's asymptotic formula can be proved easily based on the following identity

[56, Theorem III.6.3] named after him:

$$\Phi(x,y) = \Phi(x,z) + \sum_{y<p\le z}\sum_{v\ge 1}\Phi(x/p^v,p) \tag{1.1.3}$$

for any $z \in [y,x]$. The Buchstab function $\omega(u)$ then appears naturally in the iteration process, starting with $\Phi(x,y) \sim x/(u\log y)$ in the range $1 < u \le 2$. Since $1/2 \le \omega(u) \le 1$ for $u \in [1,\infty)$, Buchstab's asymptotic formula suggests that the relation $\Phi(x,y) \asymp x/\log y$ holds uniformly for $x \ge y > 1$. Thus, it is of interest to seek numerically explicit estimates for $\Phi(x,y)$ that are applicable in wide ranges. Confirming a conjecture of Ford, the author [27] showed that $\Phi(x,y) < x/\log y$ holds for all $x \ge y > 1$, which is essentially best possible when $x^{1-\epsilon} \le y \le \epsilon x$, where $\epsilon \in (0,1)$ is fixed. With a bit more effort, one can show, using the Buchstab identity (1.1.3), that

$$\Phi(x,y) = \frac{x}{\log y}\left(\omega(u) + O\left(\frac{1}{\log y}\right)\right) \tag{1.1.4}$$

uniformly for $2 \le y \le \sqrt{x}$ (see [56, Theorem III.6.4]).

In [10] de Bruijn provided a more precise approximation for $\Phi(x,y)$ than $\omega(u)x/\log y$. Let us fix some $y_0 \ge 2$ for the moment. Suppose that there exist a positive constant $C_0(y_0)$ and a positive decreasing function $R(z)$ defined on $[y_0,\infty)$, such that $R(z) \gg z^{-1}$, that $R(z) \downarrow 0$ as $z \to \infty$ and that for all $z \ge y_0$ we have

$$|\pi(z) - \mathrm{li}(z)| \le \frac{z}{\log z} R(z) \tag{1.1.5}$$

and

$$\int_z^\infty \frac{|\pi(t) - \mathrm{li}(t)|}{t^2}\,dt \le C_0(y_0)R(z), \tag{1.1.6}$$

where li(z) is the logarithmic integral defined by

$$\mathrm{li}(z) := \int_0^z \frac{dt}{\log t}.$$

The classical version of the Prime Number Theorem allows us to take $R(z) = \exp(-c\sqrt{\log z})$ for some suitable constant $c > 0$. Using the zero-free region of Korobov and Vinogradov for the Riemann zeta-function, we obtain $R(z) = \exp(-c'(\log z)^{3/5}(\log\log z)^{-1/5})$ for some absolute constant $c' > 0$. If the Riemann Hypothesis holds, then one can take $R(z) = c''z^{-1/2}\log^2 z$, where $c'' > 0$ is an absolute constant.

To state de Bruijn's result, we define

$$\mu_y(u) := \int_1^u y^{t-u}\omega(t)\,dt.$$

It is easy to see that $0 \le \mu_y(u)\log y \le 1 - y^{1-u}$ and that for every fixed $u \ge 1$, we have $\mu_y(u)\log y \to \omega(u)$ as $y \to \infty$. Precise expansions for $\mu_y(u)$ in terms of the powers of $\log y$ can be found in [56, Theorem III.6.18]. When $1 \le u \le 2$, the change of variable $t = \log v/\log y$ shows that

$$\mu_y(u)x = \int_1^u t^{-1}y^t\,dt = \int_y^x \frac{dv}{\log v} = \mathrm{li}(x) - \mathrm{li}(y).$$

Since $\Phi(x,y) = \pi(x) - \pi(y) + 1$ when $1 \le u \le 2$, (1.1.5) clearly implies that

$$\Phi(x,y) = \mu_y(u)x + (\pi(x) - \mathrm{li}(x)) - (\pi(y) - \mathrm{li}(y)) + 1 = \mu_y(u)x + O\left(\frac{xR(y)}{\log y}\right).$$

It can be shown using (1.1.5) and (1.1.6) that

$$\prod_{p\le y}\left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log y}\left(1 + O(R(y))\right).$$

Thus we have, equivalently,

$$\Phi(x,y) = \mu_y(u)e^{\gamma}x\log y\prod_{p\le y}\left(1-\frac{1}{p}\right) + O\left(\frac{xR(y)}{\log y}\right). \tag{1.1.7}$$

Essentially, de Bruijn [10] showed that this formula holds uniformly for $x \ge y \ge y_0$.

In Chapter 2 we shall prove several numerically explicit estimates for $\Phi(x,y)$. As one can see from Figure 1.1 above, the values of $\omega(u)$ for $u \ge 2$ indicate that better upper bounds for $\Phi(x,y)$ than $x/\log x$ should be expected in the narrower range $2 \le y \le \sqrt{x}$. In recent work jointly with Pomerance [28], the author showed that $\Phi(x,y) < 0.6x/\log y$ holds for all $3 \le y \le \sqrt{x}$. We shall present the proof of this result in Section 2.2.

It is also of interest to obtain numerically explicit versions of de Bruijn's formula (1.1.7). In Section 2.3 we shall show that for all $x \ge y \ge 2$, we have

$$\left|\Phi(x,y) - \mu_y(u)e^{\gamma}x\log y\prod_{p\le y}\left(1-\frac{1}{p}\right)\right| < 4.403611\frac{x}{(\log y)^{3/4}}\exp\left(-\sqrt{\frac{\log y}{6.315}}\right).$$

Moreover, if one assumes the validity of the Riemann Hypothesis, then

$$\left|\Phi(x,y) - \mu_y(u)e^{\gamma}x\log y\prod_{p\le y}\left(1-\frac{1}{p}\right)\right| < 0.449774\frac{x\log y}{\sqrt{y}}$$

holds for all $x \ge y \ge 11$.

Section 1.2

The Weighted Erdős–Kac Theorems

The celebrated Erdős–Kac theorem, first proved by Erdős and Kac [24] in 1940, states that if $\omega(n)$ denotes the number of distinct prime divisors of a positive integer n (not to be confused with the Buchstab function defined in Section 1.1), then

$$\lim_{x\to\infty}\frac{1}{x}\cdot\#\left\{n\le x\colon \frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}\le V\right\}=\Phi(V) \tag{1.2.1}$$

for any given $V\in\mathbb{R}$, where

$$\Phi(V):=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{V}e^{-v^2/2}\,dv$$

is the cumulative distribution function of the standard Gaussian distribution. This statistical result is a direct upgrade to an earlier theorem of Hardy and Ramanujan on the *normal order* of $\omega(n)$ (see [34] and [35, Theorem 431]), which asserts that given any $\epsilon>0$, the inequality $|\omega(n)-\log\log n|<\epsilon\log\log n$ holds for all but $o(x)$ values of $n\le x$. In fact, Erdős and Kac proved in the same paper a more general result in which the function $\omega(n)$ can be replaced by any strongly additive function f that is bounded on primes and has an unbounded "variance" $\sum_{p\le x}f(p)^2/p$. Recall that an arithmetic function $f\colon\mathbb{N}\to\mathbb{C}$ is said to be *additive* if $f(mn)=f(m)+f(n)$ for all positive integers $m,n\in\mathbb{N}$ with $\gcd(m,n)=1$. It is called *strongly additive* if it also satisfies the condition that $f(p^\nu)=f(p)$ for all prime powers p^ν. Thus, strongly additive functions are completely determined by their values at primes, which makes them a particularly nice subclass of additive functions.

Analogously, it can be shown that (1.2.1) remains true if one replaces $\omega(n)$ by its cousin $\Omega(n)$, which denotes the total number of prime factors of n, counting multiplicity. Indeed, this follows from the fact that

$$\sum_{n\le x}(\Omega(n)-\omega(n))=O(x). \tag{1.2.2}$$

In particular, given any $\epsilon > 0$ we have

$$\#\left\{n \le x \colon \Omega(n) - \omega(n) > \epsilon\sqrt{\log\log n}\right\} = O\left(\frac{x}{\epsilon\sqrt{\log\log x}}\right),$$

which is sufficient for deducing from the Erdős–Kac theorem that (1.2.1) also holds with $\Omega(n)$ in place of $\omega(n)$.

The Erdős–Kac theorem was first predicted by Kac. From a probabilistic point of view, one may model a positive integer $n \le x$ by a random variable $\mathbf{n}$ with the uniform probability distribution on $[1, x]$. For each prime $p \le x$, let $X_p(\mathbf{n})$ be a Bernoulli random variable which takes value 1 if $p \mid \mathbf{n}$ and 0 otherwise. Then $\text{Prob}(X_p(\mathbf{n}) = 1) = \lfloor x/p \rfloor / x = 1/p + O(1/x)$. It is clear that

$$\omega(\mathbf{n}) = \sum_{p \le x} X_p(\mathbf{n}).$$

The expectation of $\omega(\mathbf{n})$ is easily seen to be

$$\mathbb{E}(\omega(\mathbf{n})) = \sum_{p \le x} \mathbb{E}(X_p(\mathbf{n})) = \sum_{p \le x} \left(\frac{1}{p} + O\left(\frac{1}{x}\right)\right) = \log\log x + O(1).$$

by Mertens' second theorem [35, Theorem 427]. Assuming that the events $p \mid \mathbf{n}$ are mutually uncorrelated for distinct primes p, so that $\{X_p(\mathbf{n})\}_{p \le x}$ is a set of independent random variables, we see that the variance of $\omega(\mathbf{n})$ is

$$\text{Var}(\omega(\mathbf{n})) = \sum_{p \le x} \text{Var}(X_p(\mathbf{n})) = \sum_{p \le x} \left(\frac{1}{p}\left(1 - \frac{1}{p}\right) + O\left(\frac{1}{x}\right)\right) = \log\log x + O(1),$$

since

$$\text{Var}(X_p(\mathbf{n})) = \mathbb{E}(X_p(\mathbf{n})^2) - (\mathbb{E}(X_p(\mathbf{n})))^2 = \frac{1}{p}\left(1 - \frac{1}{p}\right) + O\left(\frac{1}{x}\right).$$

The central limit theorem for independent random variables then "implies" that as

$x \to \infty$, the distribution of

$$\frac{\omega(\mathbf{n}) - \log\log x}{\sqrt{\log\log x}} \tag{1.2.3}$$

approaches the standard Gaussian distribution. This heuristic for (1.2.1) resembles in spirit that given by Kac. However, as we have seen in Section 1.1, the events $p \mid \mathbf{n}$ are in fact far from being mutually uncorrelated, especially for primes p that are relatively large compared to x, and the effect of such correlation on the limiting distribution of (1.2.3) remains to be determined. Having obtained such an elegant heuristic, which was far from a rigorous proof, Kac gave a lecture at Princeton on the average number of prime factors of a random integer. Erdős, who was in the audience, soon interrupted and announced that he found a proof. This led to the publication of [24] on this subject by the two mathematicians, which opened the door to a new branch of mathematics now called "Probabilistic Number Theory".

The original proof of the Erdős–Kac theorem by Erdős and Kac used a combination of the central limit theorem and Brun's sieve and is quite complicated. Later, LeVeque [39, Theorem 1] introduced some modifications to this proof and obtained a quantitative version of (1.2.1) with a rate of convergence given by $O(\log\log\log x/\sqrt{\log\log x})$. A different proof, which is also quite involved, makes use of an asymptotic formula of Selberg [53] for $\pi_k(x)$ uniformly in the range $k \le \log\log x + V\sqrt{\log\log x}$ to estimate the number of natural numbers $n \le x$ with $\omega(n) \le \log\log x + V\sqrt{\log\log x}$, where $\pi_k(x)$ counts the number of natural numbers $n \le x$ with $\omega(n) = k$. A related approach was given by Rényi and Turán [49], who actually proved the stronger result, conjectured by LeVeque [39], that

$$\frac{1}{x} \cdot \# \left\{ n \le x \colon \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le V \right\} = \Phi(V) + O\left(\frac{1}{\sqrt{\log\log x}} \right) \tag{1.2.4}$$

holds uniformly for all $V \in \mathbb{R}$ and all $x \ge 3$, where the rate of convergence $O(1/\sqrt{\log\log x})$

is best possible in the sense that one cannot replace it by $o(1/\sqrt{\log\log x})$ without losing uniformity in V. The analytic approach of Rényi and Turán is rather deep. It requires, among other things, the asymptotics for $\pi_k(x)$ due to Erdős [22] and Sathe [51], analytic properties of the Riemann zeta-function on the line $\sigma = 1$, and the classical result from probability theory that a distribution is completely determined by its characteristic function. In order to obtain the optimal rate of convergence in (1.2.4), they also had to invoke the Berry–Esseen inequality from probability theory.

There is yet a third approach to proving the Erdős–Kac theorem (1.2.1). This approach, first suggested by Kac [37], is based on the fact that a Gaussian distribution is completely determined by its moments, which follows immediately from [4, Theorems 30.1, 30.2]. Hence, one can derive (1.2.1) by showing directly that for every $m \in \mathbb{N}$,

$$\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log x)^m = (\mu_m + o(1))(\log\log x)^{\frac{m}{2}} \tag{1.2.5}$$

as $x \to \infty$. Here μ_m is the mth moment of a standard Gaussian distribution given by

$$\mu_m = \begin{cases} m!\,/m!\,!\,, & \text{if } 2 \mid m, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$m!\,! := \prod_{k=0}^{\lfloor (m-1)/2 \rfloor} (m-2k)$$

for every $m \in \mathbb{N}$. It is easy to see by Mertens' theorem [35, Theorem 427] that the average of $\omega(n)$ for $n \le x$ is asymptotically $\log\log x$, which yields (1.2.5) in the case $m = 1$. Turán [59] proved an asymptotic formula in the case $m = 2$. Early proofs of (1.2.5) via the method of moments are due to Delange [12] in 1953 and Halberstam [31] in 1955, both of which are very complicated. Delange's proof relies on an asymptotic formula for the partial sum of the reciprocals of positive integers

n with $\omega(n) = k$, which is intimately related to $\pi_k(x)$. Later, he [13] provided an elementary proof of (1.2.5) for strongly additive functions, which is similar to but simpler than that of Halberstam. By exploiting an asymptotic formula for $\sum_{n\le x} z^{\omega(n)}$ with $z \in \mathbb{C}$, Delange [14] was also able to obtain an asymptotic expansion for the left-hand side of (1.2.4), improving upon the result of Rényi and Turán. Since $\pi_k(x)$ is precisely the coefficient of z^k in the partial sum of $z^{\omega(n)}$, his method is also related to earlier proofs of the Erdős–Kac theorem. On the other hand, Halberstam's proof was simplified and rendered more transparent by Billingsley [3] in 1969, who made further use of ideas and tools from probability theory. In 2007, Granville and Soundararajan [30] derived asymptotic formulas for the moments which hold uniformly in the range $m \le (\log\log x)^{1/3}$. Their method is so flexible that it can also be modified to study the distribution of values of additive functions in a rather general sieve-theoretic framework.

More generally, one can study the distribution of values of $\omega(n)$ weighted by certain nonnegative multiplicative functions $\alpha(n)$. Recall that an arithmetic function $\alpha\colon \mathbb{N} \to \mathbb{C}$ is said to be *multiplicative* if $\alpha(1) = 1$ and $\alpha(mn) = \alpha(m)\alpha(n)$ for all positive integers $m, n \in \mathbb{N}$ with $\gcd(m,n) = 1$. For instance, Elliott [21] showed, based on the Landau–Selberg–Delange method, that

$$\lim_{x\to\infty} \left(\sum_{n\le x} d(n)^c\right)^{-1} \sum_{\substack{n\le x \\ \omega(n)\le 2^c \log\log x + V\sqrt{2^c \log\log x}}} d(n)^c = \Phi(V) \tag{1.2.6}$$

for any given $c \in \mathbb{R}$ and $V \in \mathbb{R}$, where $d(n)$ denotes the number of positive divisors of n. Take the case $c = 1$, for example. For "normal" numbers $n \le x$ with about $\log\log x$ prime factors, $d(n)$ is near to $(\log x)^{\log 2}$, but as is well-known and easy to see, on average $d(n)$ is more closely modeled by $\log x$. This mismatch occurs because the average of $d(n)$ is skewed by rare values of n with $d(n)$ abnormally large. For instance, given any $\epsilon \in (0,1)$, we have [35, Theorem 317] $d(n) > 2^{(1-\epsilon)\log n/\log\log n}$ for any large *primorial* n, i.e., any large positive integer n which is the product of the first k primes for some $k \in \mathbb{N}$. Elliott's theorem quantifies this mismatch, so that in particular, numbers n most influential to the average of $d(n)$ have about $2\log\log x$ prime factors. And in fact, there is a Gaussian distribution with variance $\sqrt{2\log\log x}$. It is this type of theorem that we refer to as a weighted Erdős–Kac theorem. At issue here is what weights, like $d(n)^c$, can be handled. Of course, one can also consider other additive functions than $\omega(n)$.

Building on the method of Granville and Soundararajan, Khan, Milinovich and Subedi [38] recently proved

$$\lim_{x\to\infty}\left(\sum_{n\le x} d_k(n)\right)^{-1} \sum_{\substack{n\le x\\ \omega(n)\le k\log\log x + V\sqrt{k\log\log x}}} d_k(n) = \Phi(V)$$

for any given $k \in \mathbb{N}$ and $V \in \mathbb{R}$, where

$$d_k(n) := \#\left\{(a_1, ..., a_k) \in \mathbb{N}^k \colon a_1 \cdots a_k = n\right\}$$

is the k-fold divisor function. Weighted versions of the Erdős–Kac theorem with general nonnegative multiplicative weight functions $\alpha(n)$ have also been obtained by Elboim and Gorodetsky [18] and Tenenbaum [57, 58]. Elboim and Gorodetsky showed, by using a generalization of Billingsley's argument and a mean-value estimate due to de la Bretèche and Tenenbaum [11, Theorem 2.1], that if there exist absolute constants $A, \theta > 0$, $d > -1$ and $r \in (0,2)$, such that

$$\sum_{p\le x} \frac{\alpha(p)\log p}{p^d} = \theta x + O\left(\frac{x}{(\log x)^A}\right)$$

and such that $\alpha(p^\nu) = O((rp^d)^\nu)$ for all prime powers p^ν, then we have

$$\lim_{x\to\infty} \left(\sum_{n\le x} \alpha(n)\right)^{-1} \sum_{\substack{n\le x \\ \Omega(n)\le \theta\log\log x + V\sqrt{\theta\log\log x}}} \alpha(n) = \Phi(V)$$

for any given $V \in \mathbb{R}$ (see the first part of [18, Theorem 1.1]). This powerful result, which can be shown to hold with $\omega(n)$ in place of $\Omega(n)$ by the same argument, clearly includes the theorem of Elliott and that of Khan, Milinovich and Subedi as special cases. On the other hand, the theorem of Elboim and Gorodetsky follows from an even more general and technical result of Tenenbaum [57, Corollary 2.5], which we will not state here. Tenenbaum's proof utilizes characteristic functions and his effective mean-value estimates for a wide class of multiplicative functions, and it provides effective estimates for the rate of convergence for the distribution functions in consideration.

In Chapter 3 we shall generalize the method used by Granville, Soundararajan, Khan, Milinovich and Subedi to study the distribution of additive functions $f(n)$ weighted by nonnegative multiplicative functions $\alpha(n)$ in a wide class $\mathcal{M}^*$, which will be defined in Section 3.1. Our work is the first to apply this method to prove weighted Erdős–Kac theorems with general additive functions and multiplicative weights. We obtained uniform estimates for moments of strength comparable to that of the original estimate of Granville and Soundararajan. In particular, we showed that under certain conditions, the distribution of $f(n)$ with respect to the natural probability measure induced by $\alpha(n)$ is approximately Gaussian, which generalizes the result of Elboim and Gorodetsky [18] and that of Delange and Halberstam [15]. For technical reasons, we defer the formulation of our theorems until Section 3.2. Instead, we give here two interesting applications to the distribution of arithmetic functions of special interests, which were only studied previously by different methods.

Let $\tau(n)$ be the Ramanujan tau function, whose definition and properties can be found in Section 3.1. The following weighted Erdős–Kac theorem concerning $|\tau(n)|$ was first obtained by Elliott [19] in 2012 using ideas from probability theory.

Theorem 1.2.1. *Let*

$$A_\tau(x) := \sum_{\substack{p\le x\\ \tau(p)\neq 0}} \tau(p)^2 p^{-12} \log|\tau(p)p^{-11/2}|,$$

$$B_\tau(x) := \sum_{\substack{p\le x\\ \tau(p)\neq 0}} \tau(p)^2 p^{-12} \left(\log|\tau(p)p^{-11/2}|\right)^2.$$

Then we have

$$\lim_{x\to\infty}\left(\sum_{n\le x}\tau(n)^2 n^{-11}\right)^{-1} \sum_{\substack{n\le x\\ |\tau(n)|n^{-11/2}\le\exp\left(A_\tau(x)+V\sqrt{B_\tau(x)}\right)}} \tau(n)^2 n^{-11} = \Phi(V) \qquad (1.2.7)$$

for every fixed $V\in\mathbb{R}$.

We remark that the condition $\tau(p)$ in the definitions of $A_\tau(x)$ and $B_\tau(x)$ may be dropped, since $t^2\log|t|\to 0$ as $t\to 0$. More generally, Theorem 1.2.1 holds with $\tau(n)$ replaced by the Fourier coefficients of any elliptic holomorphic new form of weight at least 2 (see [20, Theorem 1]). Furthermore, using the Sato–Tate conjecture for non-CM holomorphic modular forms of weights at least 2, established by Barnet-Lamb, Geraghty, Harris and Taylor [1], we may estimate $A_\tau(x)$ to be $(1/4+o(1))\log\log x$ and replace $B_\tau(x)$ by $((\pi^2/12-5/8)\log\log x)^{1/2}$.

Along with (1.2.6), Elliott [21] also showed

$$\lim_{x\to\infty}\left(\sum_{n\le x} d(n)^2\right)^{-1} \sum_{\substack{n\le x\\ d(n)\le\exp\left(4\log 2\log\log x+V\sqrt{4(\log 2)^2\log\log x}\right)}} d(n)^2 = \Phi(V). \qquad (1.2.8)$$

If we view the Dirichlet series of $d(n)$

$$\sum_{n\geq 1}\frac{d(n)}{n^s}=\zeta(s)^2=\prod_p\left(1-2p^{-s}+p^{-2s}\right)^{-1}$$

as having an Euler product of degree 2, with $\zeta(s)^3$ its symmetric square and

$$\sum_{n\geq 1}\frac{d(n)^2}{n^s}=\frac{\zeta(s)^4}{\zeta(2s)}$$

the analogue of the corresponding Rankin–Selberg L-function, then (1.2.8) may be viewed as a limiting case of the aforementioned generalization of Theorem 1.2.1 for the coefficients of an Eisenstein series of weight 1 with respect to the modular group (see [20]).

Our first application is a simple proof of Theorem 1.2.1 as a corollary of our general theorems, which we present in Section 3.10.

Our second application concerns the distribution of the number of prime factors of values of Euler's totient function $\varphi(n)$, which is defined explicitly by

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

Recall that $\Omega(n)$ denotes the total number of prime factors of n, counting multiplicity. In [25] (with the proof of a lemma later corrected in [23]), Erdős and Pomerance proved that for every fixed $V\in\mathbb{R}$, we have

$$\lim_{x\to\infty}\frac{1}{x}\cdot\#\left\{n\leq x\colon\Omega(\varphi(n))\leq\frac{1}{2}(\log\log x)^2+V\frac{(\log\log x)^{3/2}}{\sqrt{3}}\right\}=\Phi(V). \tag{1.2.9}$$

Recently, Wang, Wei, Yan and Yi [60, Theorem 1.3] showed that the above holds with $\Phi(V)$ replaced by $\delta(S)\Phi(V)$ if we let n run over those positive integers whose largest prime factors lie in a given subset S of primes of a positive relative natural density $\delta(S)$.

It is also natural to explore weighted variants of (1.2.9) with multiplicative weights other than 1. In Section 3.11 we give a simple and straightforward extension by showing that if the multiplicative weight $\alpha(n)$ in the class $\mathcal{M}^*$ also satisfies the condition that $\alpha(p)$ is close to βp^{σ_0-1} for "almost all" primes, where $\beta, \sigma_0 > 0$ are absolute constants, then the distribution of $\Omega(\varphi(n))$ weighted by $\alpha(n)$ is approximately Gaussian with mean $\beta(\log\log x)^2/2$ and variance $\beta(\log\log x)^3/3$. The following result is a special case of this.

Theorem 1.2.2. *Given any $\kappa > 0$ and $c, V \in \mathbb{R}$, we have*

$$\lim_{x\to\infty} \left(\sum_{n\le x} d_\kappa(n)^c\right)^{-1} \sum_{\substack{n\le x \\ \Omega(\varphi(n))\le \kappa^c(\log\log x)^2/2+V\sqrt{\kappa^c(\log\log x)^3/3}}} d_\kappa(n)^c = \Phi(V). \qquad (1.2.10)$$

And the same holds if $d_\kappa(n)$ is replaced by $\kappa^{\omega(n)}$ or $\kappa^{\Omega(n)}$, where in the latter case, one has to assume $\kappa^c < 2$.

It is worth mentioning that the condition that $\alpha(p)$ is close to βp^{σ_0-1} for "almost all" primes can be relaxed if one is content with abstract expressions of the means and variances such as $A_\tau(x)$ and $B_\tau(x)$ in Theorem 1.2.1. For instance, it can be shown that $\Omega(\varphi(n))$ still possesses a Gaussian distribution if $\alpha(p)$ is bounded above and bounded away from 0 for "almost all" primes. Less straightforward generalizations will require information about the distribution of values of $\alpha(p)$, or the more tractable function $\alpha(n)\Lambda(n)$, in arithmetic progressions, where $\Lambda(n)$ is the von Mangoldt function. We hope to return to this problem in future research.

Chapter 2

The Distribution of Rough Numbers

In this chapter we study the distribution of rough numbers, numbers which are free of small prime factors. In Section 2.1, we give precise formulations of our results previewed in Section 1.1. Sections 2.2 and 2.3 are devoted to the proofs of these results.

Before embarking on our study of rough numbers, we extend the definition of $\omega(u)$ by setting $\omega(u) = 0$ for all $u < 1$, so that $\omega(u)$ satisfies the original delay differential equation for all $u \in \mathbb{R} \setminus \{1, 2\}$. It is clear that $\omega(u)$ has a jump discontinuity at $u = 1$, but its right derivative at $u = 1$ exists. On the other hand, despite the fact that $\omega(u)$ is only continuous but not differentiable at $u = 2$, both the left and right derivatives of $\omega(u)$ at $u = 2$ exist. Thus, if we write $\omega'(1)$ and $\omega'(2)$ for the right derivatives of $\omega(u)$ at $u = 1$ and $u = 2$, respectively, then we have $(u\omega(u))' = \omega(u-1)$ for all $u \in \mathbb{R}$. And we shall adopt this convention throughout the chapter.

Section 2.1

Main Results

Our first result is the inequality $\Phi(x,y) < 0.6x/\log y$ proved in [28], which improves upon the inequality $\Phi(x,y) < x/\log y$ in the range $y \le \sqrt{x}$. More precisely, we have the following theorem.

Theorem 2.1.1. *For all $3 \le y \le \sqrt{x}$, we have $\Phi(x,y) < 0.6x/\log y$. The same inequality holds when $2 \le y \le \sqrt{x}$ and $x \ge 10$.*

Theorem 2.1.1 provides a fairly good upper bound for $\Phi(x,y)$ in the range $2 \le y \le \sqrt{x}$, especially considering that the absolute maximum of $\omega(u)$ over $[2,\infty)$ is given by $M_0 = 0.5671432...$, attained at the unique critical point $u = 2.7632228...$ of the function $(\log(u-1)+1)u^{-1}$ on $[2,3]$. The proof of Theorem 2.1.1 will be given in the next section.

In Section 2.3 we shall derive an explicit version of (1.1.7), which will then be applied to obtain numerically explicit estimates with suitable y_0 and $R(y)$. Our main results are summarized in the following theorem.

Theorem 2.1.2. *For all $x \ge y \ge 2$, we have*

$$\left|\Phi(x,y) - \mu_y(u)e^{\gamma}x\log y\prod_{p\le y}\left(1-\frac{1}{p}\right)\right| < 4.403611\frac{x}{(\log y)^{3/4}}\exp\left(-\sqrt{\frac{\log y}{6.315}}\right).$$

Conditionally on the Riemann Hypothesis, we have

$$\left|\Phi(x,y) - \mu_y(u)e^{\gamma}x\log y\prod_{p\le y}\left(1-\frac{1}{p}\right)\right| < 0.449774\frac{x\log y}{\sqrt{y}}$$

for all $x \ge y \ge 11$.

In view of the asymptotic formula

$$\prod_{p\le y}\left(1-\frac{1}{p}\right)\sim\frac{e^{-\gamma}}{\log y},$$

it is natural to obtain an approximation of $\Phi(x,y)$ by the simpler $\mu_y(u)x$, which is sometimes more convenient to use. The following consequence of Theorem 2.1.2 provides approximations of this type.

Corollary 2.1.3. *For all $x \ge y \ge 2$, we have*

$$|\Phi(x,y)-\mu_y(u)x|< 4.434084\frac{x}{(\log y)^{3/4}}\exp\left(-\sqrt{\frac{\log y}{6.315}}\right).$$

Conditionally on the Riemann Hypothesis, we have

$$|\Phi(x,y)-\mu_y(u)x|< 0.460680\frac{x\log y}{\sqrt{y}}$$

for all $x \ge y \ge 11$.

Section 2.2

The 0.6 Inequality

This section is devoted to the proof of Theorem 2.1.1. The tools which we shall use are numerically explicit estimates of primes, the inclusion-exclusion principle, and a numerically explicit version of the upper bound in Selberg's sieve. The main idea of the proof may be summarized as follows. For small numerical values of y, the desired inequality follows by a careful application of the inclusion-exclusion principle. The case where u is large is then settled by applying our explicit version of Selberg's upper bound sieve. After this we are left with the case where u is small. Starting with the case $2 \le u < 3$, which can be handled easily by conventional analytic approaches, we iterate based on Buchstab's identity to complete the proof for all small values of u in consideration.

2.2.1. A Prime Lemma

Let $\pi(x)$ denote the number of primes $p \le x$. Recall

$$\mathrm{li}(x) := \int_0^x \frac{dt}{\log t},$$

where the principal value is taken for the singularity at $t = 1$. There is a long history in trying to find the first point when $\pi(x) \ge \mathrm{li}(x)$, which we now know is beyond 10^{19}. We prove a lemma based on what is currently known.

Lemma 2.2.1. *Let* $\beta_0 = 2.3 \times 10^{-8}$. *For* $x \ge 2$, *we have* $\pi(x) < (1 + \beta_0)\,\mathrm{li}(x)$.

Proof. The result is true for $x \le 10$, so assume $x \ge 10$. Consider the Chebyshev function

$$\theta(x) := \sum_{p \le x} \log p.$$

We use [40, Prop. 2.1], which depends strongly on extensive calculations of Büthe [8, 9] and Platt [45]. This result asserts in part that $\theta(x) \le x - .05\sqrt{x}$ for $1427 \le x \le 10^{19}$ and for larger x, $\theta(x) < (1 + \beta_0)x$. One easily checks that $\theta(x) < x$ for $x < 1427$, so we have

$$\theta(x) < (1 + \beta_0)x, \quad x > 0.$$

By partial summation, we have

$$\begin{aligned}
\pi(x) &= \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t(\log t)^2}\,dt \\
&< \frac{(1 + \beta_0)x}{\log x} + \int_2^{10} \frac{\theta(t)}{t(\log t)^2}\,dt + (1 + \beta_0)\int_{10}^x \frac{dt}{(\log t)^2}.
\end{aligned}$$

Since $\int dt/(\log t)^2 = -t/\log t + \mathrm{li}(t)$, we have

$$\begin{aligned}\pi(x) &< (1+\beta_0)\,\mathrm{li}(x) + \int_2^{10} \frac{\theta(t)}{t(\log t)^2}\,dt + (1+\beta_0)(10/\log 10 - \mathrm{li}(10)) \\ &< (1+\beta_0)\,\mathrm{li}(x) - .144. \end{aligned} \tag{2.2.1}$$

This gives the lemma. □

After checking for $x \le 10$, we remark that an immediate corollary of (2.2.1) is the inequality

$$\pi(x) - k < (1+\beta_0)(\mathrm{li}(x) - k), \quad 2 \le k \le \pi(x),\ k \le 10^7. \tag{2.2.2}$$

2.2.2. Presieving: Inclusion–Exclusion Revisited

For small values of $y \ge 2$, we can do a complete inclusion–exclusion to compute $\Phi(x, y)$. Let $P(y)$ denote the product of the primes $p \le y$. We have

$$\Phi(x, y) = \sum_{d|P(y)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor. \tag{2.2.3}$$

As a consequence, we have

$$\Phi(x, y) \le \sum_{d|P(y)} \mu(d)\frac{x}{d} + \sum_{\substack{d|P(y)\\ \mu(d)=1}} 1 = x \prod_{p \le y}\left(1 - \frac{1}{p}\right) + 2^{\pi(y)-1}. \tag{2.2.4}$$

We illustrate how this elementary inequality can be used in the case when $\pi(y) = 5$, that is, $11 \le y < 13$. Then the product in (2.2.4) is $16/77 < .207793$. The remainder term in (2.2.4) is 16. And we have

$$\Phi(x, y) < .207793x + 16 < .6x/\log 13$$

when $x \geq 613$. There remains the problem of dealing with smaller values of x, which we address momentarily. We apply this method for $y < 71$.

Table 2.1: Small y.

y interval	x bound	max
$[2, 3)$	22	.61035
$[3, 5)$	51	.57940
$[5, 7)$	96	.55598
$[7, 11)$	370	.56634
$[11, 13)$	613	.55424
$[13, 17)$	1603	.56085
$[17, 19)$	2753	.54854
$[19, 23)$	6296	.55124
$[23, 29)$	17539	.55806
$[29, 31)$	30519	.55253
$[31, 37)$	76932	.55707
$[37, 41)$	1.6×10^5	.55955
$[41, 43)$	2.9×10^5	.55648
$[43, 47)$	5.9×10^5	.55369
$[47, 53)$	1.4×10^6	.55972
$[53, 59)$	3.0×10^6	.55650
$[59, 61)$	5.4×10^6	.55743
$[61, 67)$	1.2×10^7	.55685
$[67, 71)$	2.4×10^7	.55641

Pertaining to Table 2.1, for x beyond the "x bound" and y in the given interval, we have $\Phi(x, y) < .6x/\log y$. The column "max" in Table 2.1 is the supremum of $\Phi(x, y)/(x/\log y)$ for y in the given interval and $x \geq y^2$ with x below the x bound. The max statistic was computed by creating a table of the integers up to the x bound with a prime factor $\leq y$, taking the complement of this set in the set of all integers up to the x bound, and then computing $(j \log p)/n$ where n is the jth member of the set and p is the upper bound of the y interval. The max of these numbers is recorded as the max statistic. The computation was done by Mathematica.

As one can see, for $y \geq 3$ the max statistic in Table 2.1 is below .6. However, for the interval $[2, 3)$ it is above .6. One can compute that it is $< .6$ once $x \geq 10$.

This method can be extended to larger values of y, but the x bound becomes prohibitively large. With a goal of keeping the x bound smaller than 3×10^7, we can extend a version of inclusion-exclusion to $y < 241$ as follows.

First, we "pre-sieve" with the primes 2, 3, and 5. For any $x \geq 0$ the number of integers $n \leq x$ with $\gcd(n, 30) = 1$ is $(4/15)x + r$, where $|r| \leq 14/15$, as can be easily verified by looking at values of $x \in [0, 30]$. We change the definition of $P(y)$ to be the product of the primes in $(5, y]$. Then for $y \geq 5$, we have

$$\Phi(x, y) \leq \frac{4}{15} \sum_{d|P(y)} \mu(d) \frac{x}{d} + \frac{14}{15} 2^{\pi(y)-3}.$$

However, it is better to use the Bonferroni inequalities in the form

$$\Phi(x, y) \leq \frac{4}{15} \sum_{j \leq 4} \sum_{\substack{d|P(y) \\ \nu(d)=j}} (-1)^j \frac{x}{d} + \sum_{i=0}^{4} \binom{\pi(y) - 3}{i} = xs(y) + b(y),$$

say, where $\nu(d)$ is the number of distinct prime factors of d. (We remark that the expression $b(y)$ could be replaced with $\frac{14}{15} b(y)$.) The inner sums in $s(y)$ can be computed easily using Newton's identities, and we see that

$$\Phi(x, y) \leq .6x/\log y \quad \text{for} \quad x > b(y)/(.6/\log y - s(y)).$$

We have verified that this x bound is smaller than 30,000,000 for $y < 241$ and we have verified that $\Phi(x, y) < .6x/\log y$ for x up to this bound and $y < 241$.

This completes the proof of Theorem 2.1.1 for $y < 241$.

2.2.3. Large u: Selberg's Sieve

In this section we prove Theorem 2.1.1 in the case that $u = \log x/\log y \geq 7.5$ and $y \geq 241$. Our principal tool is a numerically explicit form of Selberg's sieve.

Let $\mathcal{A}$ be a set of positive integers $a \le x$ and with $|\mathcal{A}| \approx X$. Let $\mathcal{P} = \mathcal{P}(y)$ be a set of primes $p \le y$. For each $p \in \mathcal{P}$ we have a collection of $\alpha(p)$ residue classes mod p, where $\alpha(p) < p$. Let $P = P(y)$ denote the product of the members of $\mathcal{P}$. Let g be the multiplicative function defined for numbers $d \mid P$ where $g(p) = \alpha(p)/p$ when $p \in \mathcal{P}$. We let

$$V := \prod_{p\in\mathcal{P}}(1 - g(p)) = \prod_{p\in\mathcal{P}}\left(1 - \frac{\alpha(p)}{p}\right).$$

We define $r_d(\mathcal{A})$ via the equation

$$\sum_{\substack{a\in\mathcal{A}\\ d|a}} 1 = g(d)X + r_d(\mathcal{A}).$$

The thought is that $r_d(\mathcal{A})$ should be small. We are interested in $S(\mathcal{A}, \mathcal{P})$, the number of those $a \in \mathcal{A}$ such that a is coprime to P.

We will use Selberg's sieve as given in [29, Theorem 7.1]. This involves an auxiliary parameter $D < X$ which can be freely chosen. Let h be the multiplicative function supported on divisors of P such that $h(p) = g(p)/(1 - g(p))$. In particular if each $\alpha(p) = 1$, then each $g(p) = 1/p$ and $h(p) = 1/(p - 1)$, so $h(d) = 1/\varphi(d)$ for $d \mid P$, where φ is Euler's function. Henceforth we will make this assumption (that each $\alpha(p) = 1$). Let

$$J = J_D = \sum_{\substack{d|P\\ d<\sqrt{D}}} h(d), \quad R = R_D = \sum_{\substack{d|P\\ d<D}} \tau_3(d)|r_d(\mathcal{A})|,$$

where $\tau_3(n) = d_3(n)$ is the number of ordered factorizations $n = abc$ with $a, b, c \in \mathbb{N}$. Selberg's sieve gives in this situation that

$$S(\mathcal{A}, \mathcal{P}) \le X/J + R. \tag{2.2.5}$$

Note that if $D \geq P^2$, then

$$J = \sum_{d \mid P} h(d) = \prod_{p \in \mathcal{P}} (1 + h(p)) = \prod_{p \in \mathcal{P}} (1 - g(p))^{-1} = V^{-1},$$

so that $X/J = XV$. This is terrific, but if D is so large, the remainder term R in (2.2.5) is also large, making the estimate useless. So, the trick is to choose D judiciously so that R is under control with J being near to V^{-1}.

Consider the case when each $|r_d(\mathcal{A})| \leq r$ for a constant r. In this situation the following lemma is useful.

Lemma 2.2.2. *Suppose that $|r_d(\mathcal{A})| \leq r$ for all $d < D$ with $d \mid P(y)$. For $y \geq 241$, we have*

$$R \leq r \sum_{\substack{d<D \\ d \mid P(y)}} \tau_3(d) \leq rD(\log y)^2 \prod_{\substack{p \leq y \\ p \notin \mathcal{P}}} \left(1 + \frac{2}{p}\right)^{-1}.$$

Proof. Let $\tau(n) = d(n)$ be the number of positive divisors of n. Note that

$$\sum_{d \mid P(y)} \frac{\tau(d)}{d} = \prod_{p \in \mathcal{P}} \left(1 + \frac{2}{p}\right) = \prod_{p \leq y} \left(1 + \frac{2}{p}\right) \prod_{\substack{p \leq y \\ p \notin \mathcal{P}}} \left(1 + \frac{2}{p}\right)^{-1}.$$

One can show that for $y \geq 241$ the first product on the right is smaller than $.95(\log y)^2$, but we will only use the "cleaner" bound $(\log y)^2$ (which holds when $y \geq 53$). Thus,

$$\begin{aligned} \sum_{\substack{d<D \\ d \mid P(y)}} \tau_3(d) &= \sum_{\substack{d<D \\ d \mid P(y)}} \sum_{j \mid d} \tau(j) \leq \sum_{\substack{j<D \\ j \mid P(y)}} \tau(j) \sum_{\substack{d<D/j \\ d \mid P(y)}} 1 \\ &< D \sum_{\substack{j<D \\ j \mid P(y)}} \frac{\tau(j)}{j} < D(\log y)^2 \prod_{\substack{p \leq y \\ p \notin \mathcal{P}}} \left(1 + \frac{2}{p}\right)^{-1}. \end{aligned}$$

This completes the proof. □

To get a lower bound for J in (2.2.5) we proceed as in [29, Section 7.4]. Recall that we are assuming each $\alpha(p) = 1$ and so $h(d) = 1/\varphi(d)$ for $d \mid P$.

Let

$$I = \sum_{\substack{d \geq \sqrt{D} \\ d \mid P}} \frac{1}{\varphi(d)},$$

so that $I + J = V^{-1}$. Hence

$$J = V^{-1} - I = V^{-1}(1 - IV), \tag{2.2.6}$$

so we want an upper bound for IV. Let ε be arbitrary with $\varepsilon > 0$. We have

$$I < D^{-\varepsilon} \sum_{d \mid P} \frac{d^{2\varepsilon}}{\varphi(d)} = D^{-\varepsilon} \prod_{p \in \mathcal{P}} \left(1 + \frac{p^{2\varepsilon}}{p-1}\right),$$

and so, assuming each $\alpha(p) = 1$,

$$IV < D^{-\varepsilon} \prod_{p \in \mathcal{P}} \left(1 + \frac{p^{2\varepsilon} - 1}{p}\right) =: f(D, \mathcal{P}, \varepsilon). \tag{2.2.7}$$

In particular, if $y \geq 241$ and each $|r_d(\mathcal{A})| \leq r$, then

$$S(\mathcal{A}, \mathcal{P}) \leq XV \left(1 - f(D, \mathcal{P}, \varepsilon)\right)^{-1} + rD(\log y)^2 \prod_{\substack{p \leq y \\ p \notin \mathcal{P}}} \left(1 + \frac{2}{p}\right)^{-1}. \tag{2.2.8}$$

We shall choose D so that the remainder term is small in comparison to XV, and once D is chosen, we shall choose ε so as to minimize $f(D, \mathcal{P}, \varepsilon)$.

The case when $y \leq 500{,}000$ ***and*** $u \geq 7.5$***.*** We wish to apply (2.2.8) to estimate $\Phi(x, y)$ when $u \geq 7.5$, that is, when $x \geq y^{7.5}$. We have a few choices for $\mathcal{A}$ and $\mathcal{P}$. The most natural choice is that $\mathcal{A}$ is the set of all integers $\leq x$, $X = x$, and $\mathcal{P}$ is the set of all primes $\leq y$. In this case, each $|r_d(\mathcal{A})| \leq 1$, so that we can take $r = 1$

in (2.2.8). Instead we choose (as in the last section) $\mathcal{A}$ as the set of all integers $\leq x$ that are coprime to 30 and we choose $\mathcal{P}$ as the set of primes p with $7 \leq p \leq y$. Then $X = 4x/15$ and one can check that each $|r_d(\mathcal{A})| \leq 14/15$, so we can take $r = 14/15$ in (2.2.8). Also,

$$\prod_{\substack{p \leq y \\ p \notin \mathcal{P}}} \left(1 + \frac{2}{p}\right)^{-1} = \frac{3}{14},$$

when $y \geq 5$. With this choice of $\mathcal{A}$ and $\mathcal{P}$, (2.2.8) becomes

$$\Phi(x,y) \leq XV \left(1 - D^{-\varepsilon} \prod_{7 \leq p \leq y} \left(1 + \frac{p^{2\varepsilon} - 1}{p}\right)\right)^{-1} + \frac{1}{5} D (\log y)^2, \tag{2.2.9}$$

when $y \geq 241$.

Our "target" for $\Phi(x,y)$ is $.6x/\log y$. We choose D here so that our estimate for the remainder term is 1% of the target, namely $.006x/\log y$. Thus, in light of Lemma 2.2.2, we choose

$$D = .03x/(\log y)^3.$$

We have verified that for every value of $y \leq 500{,}000$ and $x \geq y^{7.5}$ that the right side of (2.2.9) is smaller than $.6x/\log y$. Note that to verify this, if p, q are consecutive primes with $241 \leq p < q$, then $S(\mathcal{A}, \mathcal{P})$ is constant for $p \leq y < q$, and so it suffices to show the right side of (2.2.9) is smaller than $.6x/\log q$. Further, it suffices to take $x = p^{7.5}$, since as x increases beyond this point with $\mathcal{P}$ and ε fixed, the expression $f(D, \mathcal{P}, \varepsilon)$ decreases. For smaller values of y in the range, we used Mathematica to choose the optimal choice of ε. For larger values, we let ε be a judicious constant over a long interval. As an example, we chose $\varepsilon = .085$ in the top half of the range.

The case when $y \geq 500{,}000$ ***and*** $u \geq 7.5$***.*** As in the discussion above we have a few choices to make, namely for the quantities D and ε. First, we choose $x = y^{7.5}$, since the case $x \geq y^{7.5}$ follows from the proof of the case of equality. We choose D as before, namely $.03x/(\log y)^3$. We also choose

$$\varepsilon = 1/\log y.$$

Our goal is to prove a small upper bound for $f(D, \mathcal{P}, \varepsilon)$ given in (2.2.7). We have

$$f(D, \mathcal{P}, \varepsilon) < D^{-\varepsilon} \exp\left(\sum_{7 \leq p \leq y} \frac{p^{2\varepsilon} - 1}{p} \right).$$

We treat the two sums separately. First, by Rosser–Schoenfeld [50, Theorems 9, 20], one can show that

$$-\sum_{p \leq y} \frac{1}{p} < -\log\log y - .26$$

for all $y \geq 2$, so that

$$-\sum_{7 \leq p \leq y} \frac{1}{p} < -\log\log y - .26 + 31/30 \tag{2.2.10}$$

for $y \geq 7$. For the second sum we have

$$\sum_{7 \leq p \leq y} p^{2\varepsilon - 1} = 7^{2\varepsilon-1} + (\pi(y) - 4)y^{2\varepsilon-1} + \int_{11}^{y} (1 - 2\varepsilon)(\pi(t) - 4)t^{2\varepsilon-2}\, dt.$$

At this point we use (2.2.2) , so that

$$\begin{aligned}
\frac{1}{1+\beta_0} \sum_{11 \leq p \leq y} p^{2\varepsilon-1} &< (\mathrm{li}(y) - 4)y^{2\varepsilon-1} + \int_{11}^{y} (1-2\varepsilon)(\mathrm{li}(t) - 4)t^{2\varepsilon-2}\, dt \\
&= (\mathrm{li}(y) - 4)y^{2\varepsilon-1} - (\mathrm{li}(t) - 4)t^{2\varepsilon-1}\Big|_{11}^{y} + \int_{11}^{y} \frac{t^{2\varepsilon-1}}{\log t}\, dt \\
&= (\mathrm{li}(11) - 4)11^{2\varepsilon-1} + \mathrm{li}(t^{2\varepsilon})\Big|_{11}^{y} \\
&= (\mathrm{li}(11) - 4)11^{2\varepsilon-1} + \mathrm{li}(y^{2\varepsilon}) - \mathrm{li}(11^{2\varepsilon}),
\end{aligned}$$

and so

$$\frac{1}{1+\beta_0}\sum_{7\le p\le y} p^{2\varepsilon-1} < 7^{2\varepsilon-1} + (\mathrm{li}(11)-4)11^{2\varepsilon-1} + \mathrm{li}(y^{2\varepsilon}) - \mathrm{li}(11^{2\varepsilon}). \tag{2.2.11}$$

There are a few things to notice, but we will not need them. For example, $\mathrm{li}(y^{2\varepsilon}) = \mathrm{li}(e^2)$ and $\mathrm{li}(11^{2\varepsilon}) \approx \log(11^{2\varepsilon}-1)+\gamma$.

Let $S(y)$ be the sum of the right side of (2.2.10) and $1+\beta_0$ times the right side of (2.2.11). Then

$$f(D,\mathcal{P},\varepsilon) < D^{-\varepsilon}e^{S(y)}.$$

The expression XV in (2.2.9) is

$$x\prod_{p\le y}\left(1-\frac{1}{p}\right).$$

We know from [40] that this product is $< e^{-\gamma}/\log y$ for $y \le 2\times 10^9$, and for larger values of y, it follows from [17, Theorem 5.9] (which proof follows from [17, Theorem 4.2] or [6, Corollary 11.2]) that it is $< (1+2.1\times 10^{-5})e^{-\gamma}/\log y$. We have

$$\begin{aligned}\Phi(x,y) &\le XV\left(1-f(D,\mathcal{P},\varepsilon)\right)^{-1} + \frac{1}{5}D(\log y)^2 \\ &< \left(1+2.1\times 10^{-5}\right)\frac{x}{e^{\gamma}\log y}\left(1-D^{-\varepsilon}e^{S(y)}\right)^{-1} + \frac{.006x}{\log y}.\end{aligned} \tag{2.2.12}$$

We have verified that $(1-D^{-\varepsilon}e^{S(y)})^{-1}$ is decreasing in y, and that at $y = 500{,}000$ it is smaller than 1.057. Thus, (2.2.12) implies that

$$\Phi(x,y) < (1+2.1\times 10^{-5})\frac{1.057x}{e^{\gamma}\log y} + \frac{.006x}{\log y} < \frac{.5995x}{\log y}.$$

This concludes the case of $u \ge 7.5$.

2.2.4. The Case $2 \le u < 3$: Analytic Methods

In this section we prove that $\Phi(x,y) < .575x/\log y$ when $u \in [2,3)$, that is, when $y^2 \le x < y^3$, subject to the constraint $y \ge 241$.

For small values of y, we calculate the maximum of $\Phi(x,y)/(x/\log y)$ for $y^2 \le x < y^3$ directly, as we did in Section 2.2.2 when we checked below the x bounds in Table 2.1 and the bound 3×10^7. We have done this for $241 \le y \le 1100$, and in this range we have

$$\Phi(x,y) < .56404\frac{x}{\log y}, \quad y^2 \le x < y^3, \quad 241 \le y \le 1100.$$

Suppose now that $y > 1100$ and $y^2 \le x < y^3$. We have

$$\Phi(x,y) = \pi(x) - \pi(y) + 1 + \sum_{y<p\le x^{1/2}} (\pi(x/p) - \pi(p) + 1). \tag{2.2.13}$$

Indeed, if n is counted by $\Phi(x,y)$, then n has at most 2 prime factors (counted with multiplicity), so $n = 1$, n is a prime in $(y,x]$ or $n = pq$, where p, q are primes with $y < p \le q \le x/p$.

Let p_j denote the jth prime. Note that

$$\sum_{p\le t} \pi(p) = \sum_{j\le \pi(t)} j = \frac{1}{2}\pi(t)^2 + \frac{1}{2}\pi(t).$$

Thus,

$$\sum_{y<p\le x^{1/2}} (\pi(p) - 1) = \frac{1}{2}\pi(x^{1/2})^2 - \frac{1}{2}\pi(x^{1/2}) - \frac{1}{2}\pi(y)^2 + \frac{1}{2}\pi(y),$$

and so

$$\Phi(x,y) = \pi(x) - M(x,y) + \sum_{y<p\le x^{1/2}} \pi(x/p), \tag{2.2.14}$$

where

$$M(x,y) := \frac{1}{2}\pi(x^{1/2})^2 - \frac{1}{2}\pi(x^{1/2}) - \frac{1}{2}\pi(y)^2 + \frac{3}{2}\pi(y) - 1.$$

We use Lemma 2.2.1 on various terms in (2.2.14). In particular, we have (assuming $y \geq 5$)

$$\Phi(x,y) < (1+\beta_0)\,\mathrm{li}(x) + \sum_{y<p\leq x^{1/2}} (1+\beta_0)\,\mathrm{li}(x/p) - M(x,y). \tag{2.2.15}$$

Via partial summation, we have

$$\begin{aligned}\sum_{y<p\leq x^{1/2}} \mathrm{li}(x/p) = x^{1/2}\,\mathrm{li}(x^{1/2}) \sum_{y<p\leq x^{1/2}} \frac{1}{p} \\ - \int_y^{x^{1/2}} \left(\mathrm{li}(x/t) - \frac{x/t}{\log(x/t)}\right) \sum_{y<p\leq t} \frac{1}{p}\,dt.\end{aligned} \tag{2.2.16}$$

For $1100 \leq t \leq 10^4$ we have checked numerically using Mathematica that

$$0 < \sum_{p\leq t} \frac{1}{p} - \log\log t - B < .00624,$$

where $B = .261497\ldots$ is the Meissel–Mertens constant. Further, for $10^4 \leq t \leq 10^6$,

$$0 < \sum_{p\leq t} \frac{1}{p} - \log\log t - B < .00161.$$

(The lower bounds here follow as well from [50, Theorem 20].) It thus follows for $1100 \leq y \leq 10^4$ that

$$\sum_{y<p\leq x^{1/2}} \frac{1}{p} < \log\frac{\log(x^{1/2})}{\log y} + \beta_1, \qquad \sum_{y<p\leq t} \frac{1}{p} > \log\frac{\log t}{\log y} - \beta_1, \tag{2.2.17}$$

where $\beta_1 = .00624$. Now suppose that $y \geq 10^4$. Using [17, Eq. (5.7)] and the value

4.4916 for "η_3" from [6, Table 15], we have that

$$\left|\sum_{p\le t}\frac{1}{p} - \log\log t - B\right| < 1.9036/(\log t)^3,\ t \ge 10^6.$$

Thus, (2.2.17) continues to hold for $y \ge 10^4$ with .00624 improved to .00322. We thus have from (2.2.16)

$$\sum_{y<p\le x^{1/2}} \operatorname{li}(x/p) < x^{1/2}\operatorname{li}(x^{1/2})\left(\log\frac{\log(x^{1/2})}{\log y} + \beta_1\right) - \int_y^{x^{1/2}}\left(\operatorname{li}(x/t) - \frac{x/t}{\log(x/t)}\right)\left(\log\frac{\log t}{\log y} - \beta_1\right)dt. \tag{2.2.18}$$

Let $R(t) = (1+\beta_0)\operatorname{li}(t)/(t/\log t)$, so that $R(t) \to 1+\beta_0$ as $t\to\infty$. We write the first term on the right side of (2.2.15) as

$$\frac{x}{u\log y}R(x) = \frac{R(y^u)}{u}\frac{x}{\log y},$$

and note that the first term on the right of (2.2.18) is less than

$$R(y^{u/2})\frac{2}{u}(\log(u/2)+\beta_1)\frac{x}{\log y}.$$

For the expression $\frac{1}{2}\pi(x^{1/2})^2 - \frac{1}{2}\pi(x^{1/2})$ in $M(x,y)$ we use the inequality $\pi(t) > t/\log t + t/(\log t)^2$ when $t \ge 599$, which follows from [2, Lemma 3.4] and a calculation (also see [17, Corollary 5.2]). Further, we use $\pi(y) \le R(y)y/\log y$ for the rest of $M(x,y)$.

For $1100 \le y \le 10^4$, we take $\beta_1 = .00624$. Using these estimates and numerical integration for the integral in (2.2.18) we find that

$$\Phi(x,y) < .575\frac{x}{\log y},\quad 1100 \le y \le 10^4,\quad y^2 \le x < y^3.$$

For $y \geq 10^4$, we take $\beta_1 = .00322$. Observe that

$$\mathrm{li}(t) = \frac{t}{\log t} + \frac{t}{\log^2 t} + \frac{2t}{\log^3 t} + 6\int_0^t \frac{dv}{(\log v)^4} > \frac{t}{\log t} + \frac{t}{\log^2 t} + \frac{2t}{\log^3 t}$$

for $t \geq 28.5$. Applying this inequality and (2.2.17) to the integral in (2.2.16), we see that this integral is bounded below by

$$\int_y^{x^{1/2}} \frac{x/t}{\log^2(x/t)} \left(\log \frac{\log t}{\log y} - \beta_1\right) dt + 2\int_y^{x^{1/2}} \frac{x/t}{\log^3(x/t)} \left(\log \frac{\log t}{\log y} - \beta_1\right) dt. \tag{2.2.19}$$

Since

$$\int_y^{x^{1/2}} \frac{x/t}{(\log(x/t))^2}\, dt = \left.\frac{x}{\log(x/t)}\right|_{t=y}^{x^{1/2}} = \left(\frac{2}{u} - \frac{1}{u-1}\right)\frac{x}{\log y},$$

and

$$\begin{aligned}\int_y^{x^{1/2}} \frac{x/t}{(\log(x/t))^2} \log \frac{\log t}{\log y}\, dt &= x\left.\left(\frac{1}{\log(x/t)} \log\frac{\log t}{\log y} + \frac{1}{\log x}\log\frac{\log(x/t)}{\log t}\right)\right|_{t=y}^{x^{1/2}} \\ &= \left(\frac{\log(u/2)}{u/2} - \frac{\log(u-1)}{u}\right)\frac{x}{\log y}\end{aligned}$$

by partial integration, the first integral in (2.2.19) is equal to $A(u)x/\log y$, where

$$A(u) := \frac{\log(u/2)}{u/2} - \frac{\log(u-1)}{u} - \beta_1\left(\frac{2}{u} - \frac{1}{u-1}\right).$$

Similarly, we observe that

$$\int_y^{x^{1/2}} \frac{x/t}{(\log(x/t))^3}\, dt = \left.\frac{x}{2\log^2(x/t)}\right|_{t=y}^{x^{1/2}} = \left(\frac{4}{u^2} - \frac{1}{(u-1)^2}\right)\frac{x}{2\log^2 y}$$

and that

$$\begin{aligned}&\int_y^{x^{1/2}} \frac{x/t}{(\log(x/t))^3} \log \frac{\log t}{\log y}\, dt \\ &= \frac{x}{2}\left(\frac{1}{\log^2(x/t)} \log \frac{\log t}{\log y} + \frac{1}{\log^2 x} \log \frac{\log(x/t)}{\log t} - \frac{1}{(\log x)(\log(x/t))}\right)\Bigg|_{t=y}^{x^{1/2}} \\ &= \left(\frac{4}{u^2} \log \frac{u}{2} - \frac{\log(u-1)+2}{u^2} + \frac{1}{u(u-1)}\right) \frac{x}{2\log^2 y}.\end{aligned}$$

Hence, the second integral in (2.2.19) is equal to $B(u)x/(2\log^2 y)$, where

$$B(u) := \frac{4}{u^2} \log \frac{u}{2} - \frac{\log(u-1)+2}{u^2} + \frac{1}{u(u-1)} - \beta_1\left(\frac{4}{u^2} - \frac{1}{(u-1)^2}\right).$$

Using these identities for the integrals in (2.2.19) and estimating the other terms as before, we verify that

$$\Phi(x,y) < .572\frac{x}{\log y}, \quad y > 10^4, \quad y^2 \le x < y^3.$$

2.2.5. Iteration and Completion of the Proof of Theorem 2.1.1

Suppose k is a positive integer and we have shown that

$$\Phi(x,y) \le c_k \frac{x}{\log y} \tag{2.2.20}$$

for all $y \ge 241$ and $u = \log x/\log y \in [2,k)$. We can try to find some c_{k+1} not much larger than c_k such that

$$\Phi(x,y) \le c_{k+1} \frac{x}{\log y}$$

for $y \geq 241$ and $u < k+1$. We start with c_3, which by the results of the previous section we can take as .56404 when $241 \leq y < 1100$ and as .575 when $y \geq 1100$. In this section we attempt to find c_k for $k \leq 8$ such that $c_8 < .6$. It would then follow from Section 2.2.3 that $\Phi(x,y) < .6x/\log y$ for all $u \geq 2$ and $y \geq 241$.

Suppose that (2.2.20) holds and that y is such that $x^{1/(k+1)} < y \leq x^{1/k}$. We have

$$\Phi(x,y) = \Phi(x, x^{1/k}) + \sum_{y<p\leq x^{1/k}} \Phi(x/p, p^-), \tag{2.2.21}$$

where p^- can be taken to be any real number in $(p-1,p)$. Indeed the sum counts all $n \leq x$ with least prime factor $p \in (y, x^{1/k}]$, and $\Phi(x, x^{1/k})$ counts all $n \leq x$ with least prime factor $> x^{1/k}$. As we have seen, it suffices to deal with the case when $y = q_0^-$ for some prime q_0.

Note that if (2.2.20) holds, then it also holds for $y = x^{1/k}$. Indeed, if y is a prime, then $\Phi(x,y) = \Phi(x, y+\epsilon)$ for all $0 < \epsilon < 1$, and in this case $\Phi(x,y) \leq c_k x/\log(y+\epsilon)$, by hypothesis. Letting $\epsilon \to 0$ shows we have $\Phi(x,y) \leq c_k x/\log y$ as well. If y is not prime, then for all sufficiently small $\epsilon > 0$, we again have $\Phi(x,y) = \Phi(x, y+\epsilon)$ and the same proof works.

Thus, we have (2.2.20) holding for all of the terms on the right side of (2.2.21). This implies that

$$\Phi(x, q_0^-) \leq c_k x \left(\frac{1}{\log(x^{1/k})} + \sum_{q_0 \leq p \leq x^{1/k}} \frac{1}{p \log p} \right). \tag{2.2.22}$$

We expect that the parenthetical expression here is about the same as $1/\log q_0$, so let us try to quantify this. Let

$$\epsilon_k(q_0) = \max\left\{ \frac{-1}{\log q_0} + \frac{1}{\log(x^{1/k})} + \sum_{q_0 \leq p \leq x^{1/k}} \frac{1}{p\log p} : y^k < x \leq y^{k+1} \right\}.$$

Let q_1 be the largest prime $\leq x^{1/k}$, so that

$$\epsilon_k(q_0) = \max\left\{\frac{-1}{\log q_0} + \frac{1}{\log q_1} + \sum_{q_0 \leq p \leq q_1} \frac{1}{p \log p} : q_0 \leq q_1 \leq q_0^{1+1/k}\right\}.$$

It follows from (2.2.22) that

$$\Phi(x, y) = \Phi(x, q_0^-) \leq c_k x \left(\frac{1}{\log q_0} + \epsilon_k(q_0)\right) = \frac{c_k x}{\log y}(1 + \epsilon_k(q_0)\log q_0).$$

Note that as k grows, $\epsilon_k(q_0)$ is non-increasing since the max is over a smaller set of primes q_1. Thus, we have the inequality

$$\Phi(x, q_0^-) \leq c_3(1 + \epsilon_3(q_0)\log q_0)^j \frac{x}{\log y}, \quad x^{1/3} < q_0 \leq x^{1/(3+j)}. \tag{2.2.23}$$

Thus, we would like

$$c_3(1 + \epsilon_3(q_0)\log q_0)^5 < .6 \tag{2.2.24}$$

We have checked (2.2.24) numerically for primes $q_0 < 1000$ and it holds for $q_0 \geq 241$.

This leaves the case of primes > 1000. We have the identity

$$\begin{aligned}\sum_{q_0 \leq p \leq q_1} & \frac{1}{p \log p} \\ &= \frac{-\theta(q_0^-)}{q_0(\log q_0)^2} + \frac{\theta(q_1)}{q_1(\log q_1)^2} + \int_{q_0}^{q_1} \theta(t)\left(\frac{1}{t^2(\log t)^2} + \frac{2}{t^2(\log t)^3}\right) dt,\end{aligned}$$

via partial summation, where θ is again Chebyshev's function. First assume that $q_1 < 10^{19}$. Then using [9, Theorem 2], we have $\theta(t) \leq t$, so that

$$\sum_{q_0 \leq p \leq q_1} \frac{1}{p \log p} < \frac{q_0 - \theta(q_0^-)}{q_0(\log q_0)^2} + \frac{1}{\log q_0} - \frac{1}{\log q_1}.$$

We also have [8], [9] that $q_0 - \theta(q_0^-) < 1.95\sqrt{q_0}$, so that one can verify that

$$\epsilon_3(q_0) < \frac{1.95}{\sqrt{q_0}(\log q_0)^2}$$

and so (2.2.24) holds for $q_0 > 1000$. It remains to consider the cases when $q_1 > 10^{19}$, which implies $q_0 > 10^{14}$. Here we use $|\theta(t) - t| < 3.965t/(\log t)^2$, which is from [17, Theorem 4.2] or [6, Corollary 11.2]. This shows that (2.2.24) holds here as well, completing the proof of Theorem 2.1.1.

Section 2.3

Numerically Explicit Versions of de Bruijn's Estimate

The purpose of this section is to prove Theorem 2.1.2 and Corollary 2.1.3. The key to the proofs is an explicit version of (1.1.7) of generic nature, which we shall develop in Section 2.3.3. For our applications in Section 2.3.4, we shall also need numerically explicit lower bounds for $\Phi(x, y)$ for y in a suitable, wide range, which will be the focus of Sections 2.3.1 and 2.3.2.

2.3.1. Lower Bounds for $\Phi(x, y)$

Before moving on to the derivation of Theorem 2.1.2, we prove a clean lower bound for $\Phi(x, y)$ which is applicable in a wide range. This lower bound, which is interesting in itself, will be used in the proof of Theorem 2.1.2 and Corollary 2.1.3 in Section 2.3.4. We start by proving the following result, which provides a numerically explicit lower bound for the implicit constant in the error term in (1.1.4). It is worth noting that our method can easily be adapted to yield a numerically explicit upper bound as well.

Proposition 2.3.1. *Define* $\Delta(x, y)$ *by*

$$\Phi(x, y) = \frac{x}{\log y}\left(\omega(u) + \frac{\Delta(x, y)}{\log y}\right)$$

for $2 \leq y \leq \sqrt{x}$. *Let* $y_0 = 602$. *For every positive integer* $k \geq 3$, *we define*

$$\Delta_k^- = \Delta_k^-(y_0) := \inf\left\{\min(\Delta(x, y), 0)\colon y \geq y_0 \text{ and } 2 \leq u < k\right\}.$$

Then $\Delta_3^- > -0.563528$, $\Delta_4^- > -0.887161$, *and* $\Delta_k^- > -0.955421$ *for all* $k \geq 5$.

Proof. Let $y_1 := 2{,}278{,}383$. Suppose first that $y \geq y_1$ and set

$$G(v) := \sum_{x^{1/v} < p \leq \sqrt{x}} \frac{1}{p}$$

for $2 \leq v \leq u$. By [17, Theorem 5.6][1], we have

$$\left|G(v) - \log\frac{v}{2}\right| \leq \frac{c_1}{\log^2 y} \tag{2.3.1}$$

for all $y \geq y_1$, where $c_1 = 0.4/\log y_1$. We shall also make use of the following inequality [17, Corollary 5.2][2]:

$$\frac{z}{\log z}\left(1 + \frac{c_3}{\log z}\right) \leq \pi(z) \leq \frac{z}{\log z}\left(1 + \frac{c_2}{\log z}\right), \tag{2.3.2}$$

[1]In [6] it is claimed that the proof of [17, Theorem 4.2] is incorrect due to the application of an incorrect zero density estimate of Ramaŕe [48, Theorem 1.1]. In a footnote on p. 2299 of the same paper, the authors state that the bounds asserted in [17] are likely affected for this reason. However, since they also give a correct proof of [17, Theorem 4.2] (see [6, Corollary 11.2]), one verifies easily that the proof of [17, Theorem 5.6], which relies only on [17, Theorem 4.2], partial summation, and numerical computation, remains valid.

[2]For the same reason mentioned above, it is reasonable to suspect that the bounds given in [17, Corollary 5.2] are also affected. However, one can verify these bounds without much difficulty. Indeed, (5.2) of [17, Corollary 5.2] is superseded by [50, Corollary 1], while (5.3) and (5.4) of [17, Corollary 5.2] follow from [2, Lemmas 3.2–3.4] and direct calculations.

where $c_2 = 1+2.53816/\log y_1$ and $c_3 = 1+2/\log y_1$. We start with the range $2 \le u \le 3$. In this range, we have

$$\begin{aligned}\Phi(x,y) &= \#\{n \le x \colon P^-(n) > y \text{ and } \Omega(n) \le 2\}\\ &= \pi(x) - \pi(y) + 1 + \sum_{y<p\le\sqrt{x}} \sum_{p \le q \le x/p} 1\\ &= \pi(x) - \pi(y) + 1 + \sum_{y<p\le\sqrt{x}} (\pi(x/p) - \pi(p) + 1),\end{aligned}$$

where $\Omega(n)$ denotes the total number of prime factors of n, with multiplicity counted. Since

$$\sum_{y<p\le\sqrt{x}} (\pi(p)-1) = \sum_{\pi(y)<j\le\pi(\sqrt{x})} (j-1) = \frac{\pi(\sqrt{x})(\pi(\sqrt{x})-1)}{2} - \frac{\pi(y)(\pi(y)-1)}{2},$$

we see that

$$\pi(x) - \pi(y) + 1 - \sum_{y<p\le\sqrt{x}} (\pi(p)-1) > \pi(x) - \frac{\pi(\sqrt{x})^2}{2} + \frac{\pi(\sqrt{x})}{2}.$$

It follows from (2.3.2) that

$$\Phi(x,y) > \frac{x}{\log x}\left(1+\frac{c_3}{\log x}\right) - \frac{x}{2\log^2\sqrt{x}}\left(1+\frac{c_2}{\log\sqrt{x}}\right)^2 + \frac{\sqrt{x}}{2\log\sqrt{x}} + \sum_{y<p\le\sqrt{x}} \pi(x/p). \tag{2.3.3}$$

To handle the sum in (2.3.3), we appeal to (2.3.2) again to arrive at

$$\sum_{y<p\le\sqrt{x}} \pi(x/p) \ge \sum_{y<p\le\sqrt{x}} \left(\frac{x}{p\log(x/p)} + \frac{c_3 x}{p\log^2(x/p)}\right).$$

By partial summation we see that

$$\sum_{y<p\le\sqrt{x}} \frac{1}{p\log(x/p)} = \frac{1}{\log x}\int_{2^-}^{u} \frac{v}{v-1}\, dG(v) = \frac{1}{\log y}\left(\frac{G(u)}{u-1} + \frac{1}{u}\int_{2^-}^{u} \frac{G(v)}{(v-1)^2}\, dv\right)$$

From (2.3.1) it follows that

$$\frac{G(u)}{u-1} \ge \frac{1}{u-1}\left(\log\frac{u}{2} - \frac{c_1}{\log^2 y}\right),$$

and

$$\begin{aligned}\int_{2^-}^{u} \frac{G(v)}{(v-1)^2}\, dv &\ge \int_{2}^{u} \frac{1}{(v-1)^2}\left(\log\frac{v}{2} - \frac{c_1}{\log^2 y}\right) dv\\ &= -\frac{1}{u-1}\log\frac{u}{2} + \int_{2}^{u}\frac{1}{v(v-1)}\, dv - \frac{c_1}{\log^2 y}\left(1-\frac{1}{u-1}\right)\\ &= -\frac{u}{u-1}\log\frac{u}{2} + \log(u-1) - \frac{c_1}{\log^2 y}\left(1-\frac{1}{u-1}\right).\end{aligned}$$

Hence

$$\begin{aligned}\sum_{y<p\le\sqrt{x}} \frac{x}{p\log(x/p)} &\ge \frac{x}{\log y}\left(\frac{\log(u-1)}{u} - \frac{2c_1}{u\log^2 y}\right)\\ &= \frac{x}{\log y}\left(\omega(u) - \frac{2c_1}{u\log^2 y}\right) - \frac{x}{\log x}. \qquad (2.3.4)\end{aligned}$$

Similarly, we have

$$\sum_{y<p\le\sqrt{x}} \frac{1}{p\log^2(x/p)} = \frac{1}{\log^2 x}\int_{2^-}^{u}\left(\frac{v}{v-1}\right)^2 dG(v) = \frac{1}{\log^2 x}\left(\frac{G(u)u^2}{(u-1)^2} + 2\int_{2^-}^{u}\frac{vG(v)}{(v-1)^3}\, dv\right).$$

By (2.3.1) we have

$$\frac{G(u)u^2}{(u-1)^2} \ge \frac{u^2}{(u-1)^2}\left(\log\frac{u}{2} - \frac{c_1}{\log^2 y}\right),$$

and

$$\int_{2^-}^{u} \frac{vG(v)}{(v-1)^3}\,dv \geq \int_{2}^{u} \frac{v}{(v-1)^3}\left(\log\frac{v}{2} - \frac{c_1}{\log^2 y}\right) dv.$$

Since

$$\begin{aligned}\int_{2}^{u} \frac{v}{(v-1)^3}\log\frac{v}{2}\,dv &= -\left(\frac{1}{u-1} + \frac{1}{2(u-1)^2}\right)\log\frac{u}{2} + \int_{2}^{u}\left(\frac{1}{v-1} + \frac{1}{2(v-1)^2}\right)\frac{dv}{v}\\ &= -\frac{2u-1}{2(u-1)^2}\log\frac{u}{2} + \frac{1}{2}\int_{2}^{u}\left(\frac{1}{(v-1)^2} + \frac{1}{v(v-1)}\right) dv\\ &= -\frac{u^2}{2(u-1)^2}\log\frac{u}{2} + \frac{1}{2}\left(\log(u-1) + 1 - \frac{1}{u-1}\right)\end{aligned}$$

and

$$\int_{2}^{u} \frac{v}{(v-1)^3}\,dv = -\frac{2u-1}{2(u-1)^2} + \frac{3}{2},$$

we have

$$\sum_{y<p\leq\sqrt{x}} \frac{x}{p\log^2(x/p)} \geq \frac{x}{\log^2 x}\left(\log(u-1) + \frac{u-2}{u-1} - \frac{4c_1}{\log^2 y}\right). \tag{2.3.5}$$

Inserting (2.3.4) and (2.3.5) into (2.3.3) yields

$$\Delta(x,y) \geq g(u) - \frac{2c_1}{u\log y} + \frac{\log y}{uy^{3/2}} - \frac{1}{u^2}\left(2 - c_3 + \frac{4c_1c_3}{\log^2 y} + \frac{8c_2}{u\log y} + \frac{8c_2^2}{u^2\log^2 y}\right),$$

where

$$g(u) := \frac{c_3}{u^2}\left(\log(u-1) + \frac{u-2}{u-1}\right).$$

Using Mathematica we find that $\Delta_3^- > -0.301223$ when $y \geq y_1$.

Now we proceed to bound Δ_k^- for $k \geq 4$ recursively when $y \geq y_1$. Let $k \geq 3$ be arbitrary. It is easily seen that the following variant of Buchstab's identity (1.1.3) holds for any $z \in [y, x]$:

$$\Phi(x, y) = \Phi(x, z) + \sum_{y<p\leq z} \Phi(x/p, p^-), \tag{2.3.6}$$

where p^- can be taken to be any real number in $(p-1, p)$. For $3 \leq k \leq u < k+1$ and $y \geq y_1$, we obtain by taking $z = x^{1/3}$ that

$$\Phi(x, y) = \Phi\left(x, x^{1/3}\right) + \sum_{y<p\leq x^{1/3}} \Phi(x/p, p^-). \tag{2.3.7}$$

We have already shown that

$$\Phi\left(x, x^{1/3}\right) \geq \frac{x}{\log x^{1/3}} \left(\omega\left(\frac{\log x}{\log x^{1/3}}\right) + \frac{\Delta_3^-}{\log x^{1/3}}\right) = \frac{3x}{\log y}\left(\frac{\omega(3)}{u} + \frac{3\Delta_3^-}{u^2 \log y}\right). \tag{2.3.8}$$

Note that $2 < \log(x/p)/\log(p^-) < k$. Thus, we have

$$\Phi(x/p, p^-) \geq \frac{x}{p\log(p^-)}\left(\omega\left(\frac{\log(x/p)}{\log(p^-)}\right) + \frac{\Delta_k^-}{\log(p^-)}\right).$$

Since $\omega(u)$ is continuous on $[1, \infty)$, it follows from (2.3.7) and (2.3.8) that

$$\Phi(x, y) \geq \frac{3x}{\log y}\left(\frac{\omega(3)}{u} + \frac{3\Delta_3^-}{u^2\log y}\right) + \sum_{y<p\leq x^{1/3}} \frac{x}{p\log p}\left(\omega\left(\frac{\log x}{\log p} - 1\right) + \frac{\Delta_k^-}{\log p}\right). \tag{2.3.9}$$

By partial summation we see that

$$\sum_{y<p\leq x^{1/3}} \frac{1}{p\log^2 p} < \int_y^\infty \frac{1}{t\log^2 t}\, d\pi(t) = -\frac{\pi(y)}{y\log^2 y} + \int_y^\infty \frac{\log t + 2}{t^2\log^3 t}\pi(t)\, dt,$$

which, by (2.3.2), is

$$
\begin{aligned}
&< -\frac{1}{\log^3 y}\left(1+\frac{c_3}{\log y}\right)+\int_y^\infty \frac{\log t+2}{t\log^4 t}\left(1+\frac{c_2}{\log t}\right)\,dt\\
&= -\frac{1}{\log^3 y}\left(1+\frac{c_3}{\log y}\right)+\frac{1}{2\log^2 y}+\frac{c_2+2}{3\log^3 y}+\frac{c_2}{2\log^4 y}\\
&= \frac{1}{\log^2 y}\left(\frac{1}{2}+\left(\frac{c_2}{3}-1\right)\frac{1}{\log y}+\left(\frac{c_2}{2}-c_3\right)\frac{1}{\log^2 y}\right)\\
&< \frac{1}{\log^2 y}\left(\frac{1}{2}+\left(\frac{c_2}{3}-1\right)\frac{1}{\log y}\right).
\end{aligned}
$$

Hence

$$
\sum_{y<p\le x^{1/3}} \frac{\Delta_k^- x}{p\log^2 p} \ge \frac{\Delta_k^- x}{\log^2 y}\left(\frac{1}{2}+\left(\frac{c_2}{3}-1\right)\frac{1}{\log y}\right). \tag{2.3.10}
$$

On the other hand, we have

$$
\begin{aligned}
\sum_{y<p\le x^{1/3}} \frac{1}{p\log p}\omega\left(\frac{\log x}{\log p}-1\right) &= \frac{1}{\log x}\int_{3^-}^{u} v\omega(v-1)\,dG(v)\\
&= \frac{1}{\log x}\left(\int_3^u \omega(v-1)\,dv+\int_{3^-}^u v\omega(v-1)\,d\left(G(v)-\log\frac{v}{2}\right)\right).
\end{aligned}
$$

Observe that

$$
\int_3^u \omega(v-1)\,dv = u\omega(u)-3\omega(3)
$$

and that

$$
\begin{aligned}
\int_{3^-}^u v\omega(v-1)\,d\left(G(v)-\log\frac{v}{2}\right) &= u\omega(u-1)\left(G(v)-\log\frac{v}{2}\right)-3\omega(2)\left(G(3)-\log\frac{3}{2}\right)\\
&\quad -\int_{3^-}^u \left(G(v)-\log\frac{v}{2}\right)\,d(v\omega(v-1)).
\end{aligned}
$$

By [56, (6.23), p. 562] and [56, Theorems III.5.7 & III.6.6], we have, for all $v \ge 3$, that

$$\frac{d}{dv}(v\omega(v-1)) = \omega(v-2) + \omega'(v-1) \geq \frac{1}{2} - \rho(v-1) \geq \frac{1}{2} - \rho(2) = \log 2 - \frac{1}{2},$$
$$\frac{d}{dv}(v\omega(v-1)) \leq 1 + \rho(v-1) \leq 1 + \rho(2) = 2 - \log 2,$$

where ρ is the Dickman-de Bruijn function defined to be the unique continuous solution to the delay differential equation $t\rho'(t) + \rho(t-1) = 0$ for $t \geq 1$, subject to the initial value condition $\rho(t) = 1$ for $0 \leq t \leq 1$. Moreover, we have

$$\lim_{v\to 3^-} \frac{d}{dv}(v\omega(v-1)) = \lim_{v\to 3^-} (\omega(v-2) + \omega'(v-1)) = -\frac{1}{4}.$$

It follows by (2.3.1) that

$$\int_{3^-}^{u} \left(G(v) - \log\frac{v}{2}\right) d(v\omega(v-1)) \leq \frac{c_1}{\log^2 y}\left(u\omega(u-1) - 3\omega(2)\right).$$

Thus we have

$$\int_{3^-}^{u} v\omega(v-1)\, d\left(G(v) - \log\frac{v}{2}\right) \geq -\frac{2c_1 u\omega(u-1)}{\log^2 y} \geq -\frac{2c_1 M_0 u}{\log^2 y},$$

where $M_0 = 0.5671432....$ Hence we have shown that

$$\sum_{y<p\leq x^{1/3}} \frac{x}{p\log p}\omega\left(\frac{\log x}{\log p} - 1\right) \geq \frac{x}{\log y}\left(\omega(u) - \frac{3\omega(3)}{u} - \frac{2c_1 M_0}{\log^2 y}\right). \tag{2.3.11}$$

Combining (2.3.9), (2.3.10) and (2.3.11), we deduce that

$$\Delta(x,y) \geq \frac{9\Delta_3^-}{u^2} + \frac{\Delta_k^-}{2} - \frac{1}{\log y}\left(2c_1 M_0 - \left(\frac{c_2}{3} - 1\right)\Delta_k^-\right)$$

for $k \le u < k+1$. Therefore, $\Delta_{k+1}^- \ge \min(\Delta_k^-, a_k^-)$ for all $k \ge 3$, where

$$a_k^- := \frac{9\Delta_3^-}{k^2} + \frac{\Delta_k^-}{2} - \frac{1}{\log y_1} \cdot \max\left(2c_1 M_0 - \left(\frac{c_2}{3} - 1\right)\Delta_k^-, 0\right).$$

Consequently, we have $\Delta_4^- > -0.451835$ and $\Delta_k^- > -0.480075$ for all $k \ge 5$.

Suppose now that $602 \le y \le y_1$. By [50, Theorem 20] we can replace (2.3.1) with

$$\left|G(v) - \log\frac{v}{2}\right| \le \frac{d_1}{\sqrt{y}\log y},$$

where $d_1 = 2$. Moreover, (2.3.2) remains true if we replace c_2 and c_3 by $d_2 = 1.2762$ and $d_3 = 1$, respectively, according to [17, Corollary 5.2]. With these changes, we run the same argument used to handle the case $y \ge y_1$ and get

$$\Delta(x, y) > g(u) - \frac{2d_1}{u\sqrt{y}} + \frac{\log y}{uy^{3/2}} - \frac{1}{u^2}\left(2 - d_3 + \frac{4d_1 d_3}{\sqrt{y}\log y} + \frac{8d_2}{u\log y} + \frac{8d_2^2}{u^2\log^2 y}\right).$$

when $2 \le u \le 3$ and

$$\Delta(x, y) \ge \frac{9\Delta_3^-}{u^2} + \frac{\Delta_k^-}{2} - \frac{1}{\log y}\left(\frac{2d_1 M_0 \log y}{\sqrt{y}} - \left(\frac{d_2}{3} - 1\right)\Delta_k^-\right)$$

when $3 \le k \le u < k+1$, so that we can take

$$a_k^- = \frac{9\Delta_3^-}{k^2} + \frac{\Delta_k^-}{2} - \frac{1}{\log y_0} \cdot \max\left(\frac{2d_1 M_0 \log y_0}{\sqrt{y_0}} - \left(\frac{d_2}{3} - 1\right)\Delta_k^-, 0\right).$$

As a consequence, we have $\Delta_3^- > -0.563528$, $\Delta_4^- > -0.887161$ and $\Delta_k^- > -0.955421$ for all $k \ge 5$. This completes the proof of the proposition. $\square$

2.3.2. The Inequality $\Phi(x,y) > 0.4x/\log y$

Proposition 2.3.1 allows us to show that the clean inequality $\Phi(x,y) > 0.4x/\log y$ holds for all $7 \le y \le x^{2/3}$. In addition to Proposition 2.3.1, we also need a numerical lower bound for $\omega(u)$ on $[3,\infty)$.

Lemma 2.3.2. *We have* $\omega(u) > 0.549307$ *for all* $u \ge 3$*.*

Proof. Consider first the case $u \in [3,4]$. Since $(t\omega(t))' = \omega(t-1)$ for $t \ge 2$ and $\omega(t) = (\log(t-1)+1)/t$ for $t \in [2,3]$, we have

$$\omega(u) = \frac{1}{u}\left(\log 2 + 1 + \int_3^u \frac{\log(t-2)+1}{t-1}\,dt\right)$$

for $u \in [3,4]$. Note that $u\omega'(u) = \omega(u-1) - \omega(u) = S(u)/u$, where

$$S(u) := \frac{u(\log(u-2)+1)}{u-1} - \log 2 - 1 - \int_3^u \frac{\log(t-2)+1}{t-1}\,dt.$$

Since

$$\begin{aligned} S'(u) &= \frac{1}{u-1}\left(\log(u-2) + 1 + \frac{u}{u-2} - \frac{u(\log(u-2)+1)}{u-1} - (\log(u-2)+1)\right) \\ &= \frac{u(1-(u-2)\log(u-2))}{(u-2)(u-1)^2}, \end{aligned}$$

we know that $S(u)$ is strictly increasing on $[3,u_1]$ and strictly decreasing on $[u_1,4]$, where $u_1 = 3.7632228...$ is the unique solution to the equation $(u-2)\log(u-2) = 1$. But $S(3) = 1/2 - \log 2 < 0$ and

$$S(4) = \frac{\log 2 + 1}{3} - \int_3^4 \frac{\log(t-2)+1}{t-1}\,dt > 0.$$

It follows that $S(u)$ has a unique zero $u_2 \in [3,4]$. The numerical value of u_2 is given by $u_2 = 3.4697488...$, according to Mathematica. Hence $S(u) < 0$ for $u \in [3,u_2)$

and $S(u) > 0$ for $u \in (u_2, 4]$. The same is true for $\omega'(u)$, which implies that $\omega(u)$ is strictly decreasing on $[3, u_2]$ and strictly increasing on $[u_2, 4]$. Thus, $\omega(u) \geq \omega(u_2) = 0.5608228...$ for $u \in [3, 4]$.

Consider now the case $u \in [4, \infty)$. It is known [36] that $\omega(t)$ satisfies

$$|\omega(t) - e^{-\gamma}| \leq \frac{\rho(t-1)}{t}$$

for all $t \geq 1$. Since $\rho(t)$ is strictly decreasing on $[4, \infty)$, we have $\omega(u) \geq e^{-\gamma} - \rho(3)/4$ for all $u \geq 4$. To find the value of $\rho(3)$, we use $t\rho'(t) + \rho(t-1) = 0$ for $t \geq 1$ and $\rho(t) = 1 - \log t$ for $t \in [1, 2]$ to obtain

$$\rho(u) = 1 - \log 2 - \int_2^u \frac{1 - \log(t-1)}{t} \, dt$$

for $u \in [2, 3]$. It follows that

$$\omega(u) \geq e^{-\gamma} - \frac{1}{4}\left(1 - \log 2 - \int_2^3 \frac{1 - \log(t-1)}{t} \, dt\right) = 0.5493073...$$

for all $u \geq 4$. We have therefore shown that $\omega(u) > 0.549307$ for all $u \geq 3$. □

We are now ready to prove the asserted inequality $\Phi(x, y) > 0.4x/\log y$.

Theorem 2.3.3. *We have $\Phi(x, y) > 0.4x/\log y$ for all $7 \leq y \leq x^{2/3}$.*

Proof. In the range $\max(7, x^{2/5}) \leq y \leq x^{2/3}$, we have trivially $\Phi(x, y) \geq \pi(x) - \pi(y) + 1$. By [17, Corollary 5.2] we have

$$\begin{aligned}
\pi(x) - \pi(y) &\geq \frac{x}{\log x}\left(1 + \frac{1}{\log x}\right) - \frac{y}{\log y}\left(1 + \frac{1.2762}{\log y}\right) \\
&= \left(\frac{1}{u}\left(1 + \frac{1}{\log x}\right) - \frac{y}{x}\left(1 + \frac{1.2762u}{\log x}\right)\right)\frac{x}{\log y} \\
&> \left(\frac{2}{5}\left(1 + \frac{1}{\log x}\right) - \frac{1}{x^{1/3}}\left(1 + \frac{3.1905}{\log x}\right)\right)\frac{x}{\log y} > 0.4\frac{x}{\log y}
\end{aligned}$$

whenever $x \geq 41{,}217$. Furthermore, we have verified $\Phi(x, y) > 0.4x/\log y$ for $\max(7, x^{2/5}) \leq y \leq x^{2/3}$ with $x \leq 41{,}217$ using Mathematica. Hence, $\Phi(x, y) > 0.4x/\log y$ holds in the range $\max(7, x^{2/5}) \leq y \leq x^{2/3}$.

Consider now the case $\max(x^{1/3}, 7) \leq y \leq x^{2/5}$. Following the proof of Proposition 2.3.1, we have

$$\begin{aligned} \Phi(x, y) &= \pi(x) - \pi(y) + 1 + \sum_{y<p\leq\sqrt{x}} (\pi(x/p) - \pi(p) + 1) \\ &= \pi(x) - M(x, y) + \sum_{y<p\leq x^{1/2}} \pi(x/p), \end{aligned} \tag{2.3.12}$$

where

$$M(x, y) := \frac{1}{2}\pi\left(\sqrt{x}\right)^2 - \frac{1}{2}\pi\left(\sqrt{x}\right) - \frac{1}{2}\pi(y)^2 + \frac{3}{2}\pi(y) - 1.$$

To handle the sum in (2.3.12), we appeal to Theorem 5 and its corollary from [50] to arrive at

$$G(v) - \log\frac{v}{2} > -\frac{1}{2\log^2\sqrt{x}} - \frac{1}{\log^2 y} \geq -\frac{33}{25\log^2 y}$$

in the range $\max(x^{1/3}, 7) \leq y \leq x^{2/5}$. By [50, Corollary 1] we have

$$\sum_{y<p\leq x^{1/2}} \pi(x/p) > x \sum_{y<p\leq x^{1/2}} \frac{1}{p\log(x/p)} = \frac{x}{\log x}\int_{2^-}^{u} \frac{v}{v-1}\, dG(v),$$

provided that $x \geq 289$. The right-hand side of the above can be estimated in the same way as in the proof of Proposition 2.3.1, so we obtain

$$\sum_{y<p\leq\sqrt{x}} \pi(x/p) > \frac{x}{\log y}\left(\omega(u) - \frac{66}{25u\log^2 y}\right) - \frac{x}{\log x}.$$

On the other hand, we see by [17, Corollary 5.2] and [50, Corollary 2] that

$$\pi(x) - M(x,y) > \pi(x) - \frac{1}{2}\pi\left(\sqrt{x}\right)^2 \geq \frac{x}{\log x}\left(1 + \frac{1}{\log x}\right) - \frac{25x}{8\log^2 x} = \frac{x}{\log x} - \frac{17x}{8\log^2 x}.$$

for $x \geq 114^2$. Collecting the estimates above and using the inequality $\omega(u) \geq \omega(5/2) = 2(\ln(3/2)+1)/5$ for $u \in [5/2, 3]$, we find that

$$\Phi(x,y) > \frac{\omega(5/2)x}{\log y} - \frac{17x}{8\log^2 x} - \frac{66x}{25u\log^3 y} \geq \frac{\omega(5/2)x}{\log y} - \frac{17x}{50\log^2 y} - \frac{132x}{125\log^3 y} > 0.4\frac{x}{\log y}$$

for all $\max(46, x^{1/3}) \leq y \leq x^{2/5}$. For $x^{1/3} \leq y \leq x^{2/5}$ with $7 \leq y \leq 46$, we have verified the inequality $\Phi(x,y) > 0.4x/\log y$ directly through numerical computation.

Next, we consider the range $7 \leq y < x^{1/3}$. By Proposition 2.3.1 and Lemma 2.3.2 we have

$$\Phi(x,y) > \frac{x}{\log y}\left(0.549307 - \frac{0.955421}{\log y}\right) > 0.4\frac{x}{\log y},$$

provided that $y \geq 602$. To deal with the range $7 \leq y \leq \min(x^{1/3}, 602)$, we follow the inclusion-exclusion technique used in Section 2.2.2. For any integer $n \geq 1$, let $\nu(n)$ denote the number of distinct prime factors of n as before. We start by "pre-sieving" with the primes 2, 3, and 5: for any $x \geq 1$ the number of integers $n \leq x$ with $\gcd(n, 30) = 1$ is $(4/15)x + r_x$, where $|r_x| \leq 14/15$. Let $P_5(y)$ be the product of the primes in $(5, y]$. Then we have by the Bonferroni inequalities that

$$\Phi(x,y) \geq \sum_{\substack{d|P_5(y)\\ \nu(d)\leq 3}} \mu(d)\left(\frac{4}{15}\cdot\frac{x}{d} + r_{x/d}\right) \geq a(y)x - b(y),$$

where

$$a(y) := \frac{4}{15} \sum_{\substack{d|P_5(y) \\ \nu(d)\le 3}} \frac{\mu(d)}{d} = \frac{4}{15} \sum_{j=0}^{3} (-1)^j \sum_{\substack{d|P_5(y) \\ \nu(d)=j}} \frac{1}{d},$$

$$b(y) := \frac{14}{15} \sum_{j=0}^{3} \binom{\pi(y)-3}{j}.$$

By Newton's identities, the inner sum in the definition of $a(y)$ can be represented in terms of the power sums of $1/p$ over all primes $5 < p \le y$. Thus, we have $\Phi(x,y) > 0.4x/\log y$ whenever $a(y) > 0.4/\log y$ and $x > b(y)/(a(y) - 0.4/\log y)$. Using Mathematica, we find that the inequality $\Phi(x,y) > 0.4x/\log y$ holds for $7 \le y \le 602$ and $x \ge 13{,}160{,}748$. Finally, we have verified the inequality $\Phi(x,y) > 0.4x/\log y$ directly for $7 \le y \le x^{1/3}$ with $x \le 13{,}160{,}748$ by numerical calculations, completing the proof of our theorem. □

Remark 2.3.1. Note that for $y \in [5,7)$ we have

$$\Phi(x,y) \ge \frac{4}{15}x - \frac{14}{15} > 0.4\frac{x}{\log 5} \ge 0.4\frac{x}{\log y},$$

provided that $x \ge 52$. Combined with Theorem 2.3.3 and numerical examination of the case $11 \le x \le 52$, this implies that the inequality $\Phi(x,y) > 0.4x/\log y$ holds in the slightly larger range $5 \le y \le x^{2/3}$ if one assumes $x \ge 41$.

2.3.3. The General Approach

To prove Theorem 2.1.2, we shall first develop an explicit version of (1.1.7) with a general $R(y)$, following [10], where $R(y)$ is a positive decreasing function satisfying the same conditions described in the introduction. Suppose that $y_0 \ge 3$. For each

$z \geq 2$, put

$$Q(z) := \prod_{p \leq z} \left(1 - \frac{1}{p}\right).$$

We start by estimating $Q(y)$ for $y \geq y_0$. Using a Stieltjes integral, we may write

$$\log \frac{Q(z)}{Q(y)} = \int_y^z \log\left(1 - t^{-1}\right) \, d\,\mathrm{li}(y) + \int_y^z \log\left(1 - t^{-1}\right) \, d(\pi(y) - \mathrm{li}(t)), \tag{2.3.13}$$

where $z \geq y \geq y_0$. The first integral on the right-hand side of the above is equal to

$$\int_y^z \log\left(1 - t^{-1}\right) \frac{dt}{\log t} = -\log \frac{\log z}{\log y} + \int_y^z \left(t^{-1} + \log\left(1 - t^{-1}\right)\right) \frac{dt}{\log t}.$$

Since

$$-\frac{1}{2t(t-1)} < t^{-1} + \log\left(1 - t^{-1}\right) < 0$$

for all $t \geq y_0$, we have

$$-\frac{1}{2} \int_y^\infty \frac{dt}{t(t-1)\log t} < \int_y^z \left(t^{-1} + \log\left(1 - t^{-1}\right)\right) \frac{dt}{\log t} < 0.$$

But a change of variable shows that

$$\int_y^\infty \frac{dt}{t(t-1)\log t} = \int_1^\infty \frac{dt}{t(y^t - 1)} \leq \frac{1}{y-1} \int_1^\infty \frac{dt}{t^2} = \frac{1}{y-1},$$

where we have used the inequality $y^t - 1 \geq (y-1)t$ for $t \geq 1$ and $y \geq y_0$. It follows that

$$-\frac{1}{2(y-1)} \leq \int_y^z \log\left(1 - t^{-1}\right) \, d\,\mathrm{li}(y) + \log \frac{\log z}{\log y} < 0. \tag{2.3.14}$$

Now we estimate the second integral on the right-hand side of (2.3.13). By (1.1.5) and partial integration we have

$$\left|\int_y^z \log\left(1-t^{-1}\right)\, d(\pi(y)-\mathrm{li}(t))\right| \le \log\left(1-y^{-1}\right)^{-1}\frac{y}{\log y}R(y)+\log\left(1-z^{-1}\right)^{-1}\frac{z}{\log z}R(z)$$
$$+\int_y^z \frac{|\pi(t)-\mathrm{li}(t)|}{t(t-1)}\,dt.$$

Using (1.1.6) we see that

$$\int_y^z \frac{|\pi(t)-\mathrm{li}(t)|}{t(t-1)}\,dt \le \frac{C_0(y_0)y_0}{y_0-1}R(y).$$

It is clear that the function

$$\log\left(1-t^{-1}\right)^{-1}\frac{t}{\log t}=\frac{1}{\log t}\sum_{n=0}^{\infty}\frac{t^{-n}}{n+1}$$

is strictly decreasing for $t\in(1,\infty)$. Since $R(t)$ is decreasing on $[y_0,\infty)$, we find that

$$\left|\int_y^z \log\left(1-t^{-1}\right)\, d(\pi(y)-\mathrm{li}(t))\right| \le \left(2\log\left(1-y_0^{-1}\right)^{-1}\frac{y_0}{\log y_0}+\frac{C_0(y_0)y_0}{y_0-1}\right)R(y).$$

Combining this inequality with (2.3.13) and (2.3.14) yields

$$-C_2(y_0)R(y)\le \log\frac{Q(z)}{Q(y)}+\log\frac{\log z}{\log y}\le C_1(y_0)R(y) \tag{2.3.15}$$

for $z\ge y\ge y_0$, where

$$C_1(y_0)=2\log\left(1-y_0^{-1}\right)^{-1}\frac{y_0}{\log y_0}+\frac{C_0(y_0)y_0}{y_0-1},$$
$$C_2(y_0)=C_1(y_0)+\sup_{t\ge y_0}\frac{1}{2(t-1)R(t)}.$$

Exponentiating (2.3.15) we obtain

$$-C_4(y_0)R(y) \le \frac{Q(z)\log z}{Q(y)\log y} - 1 \le C_3(y_0)R(y) \tag{2.3.16}$$

for $z \ge y \ge y_0$, where

$$\begin{aligned} C_3(y_0) &= \sup_{t \ge y_0} \frac{\exp(C_1(y_0)R(t)) - 1}{R(t)} = \frac{\exp(C_1(y_0)R(y_0)) - 1}{R(y_0)}, \\ C_4(y_0) &= \sup_{t \ge y_0} \frac{1 - \exp(-C_2(y_0)R(t))}{R(t)} = C_2(y_0). \end{aligned}$$

As a consequence, we have by letting $z \to \infty$ in (2.3.16) and using the fact that $Q(z)\log z \to e^{-\gamma}$ as $z \to \infty$, that

$$e^{\gamma}\log y(1 - C_4(y_0)R(y)) \le \frac{1}{Q(y)} \le e^{\gamma}\log y(1 + C_3(y_0)R(y)). \tag{2.3.17}$$

Similarly, we derive from (2.3.15) that

$$\frac{e^{-\gamma}}{\log y}(1 - C_6(y_0)R(y)) \le Q(y) \le \frac{e^{-\gamma}}{\log y}(1 + C_5(y_0)R(y)) \tag{2.3.18}$$

for $y \ge y_0$, where

$$\begin{aligned} C_5(y_0) &= \sup_{t \ge y_0} \frac{\exp(C_2(y_0)R(t)) - 1}{R(t)} = \frac{\exp(C_2(y_0)R(y_0)) - 1}{R(y_0)}, \\ C_6(y_0) &= \sup_{t \ge y_0} \frac{1 - \exp(-C_1(y_0)R(t))}{R(t)} = C_1(y_0). \end{aligned}$$

For $x \ge y \ge 2$, we define

$$\psi(x, y) := \frac{\Phi(x, y)}{xQ(y)}.$$

We then need to estimate $\eta(x, y) = \psi(x, y) - \lambda(x, y)$, where $\lambda(x, y) := e^{\gamma}\mu_y(u)\log y$. For $1 \le u \le 2$ this can be done straightforward. Indeed, we have $\Phi(x, y) = \pi(x) -$

$\pi(y)+1$ and $\omega(u)=1/u$ when $1 \le u \le 2$, so that

$$\eta(x,y) = \frac{\pi(x)-\pi(y)+1}{xQ(y)} - e^{\gamma}\log y \int_1^u t^{-1}y^{t-u}\,dt.$$

Note that

$$\left|\pi(x)-\pi(y)-x\int_1^u t^{-1}y^{t-u}\,dt\right| = \left|\pi(x)-\pi(y)-\int_y^x \frac{dt}{\log t}\right| \le \left(\frac{x}{\log x}+\frac{y}{\log y}\right)R(y).$$

From (2.3.17) it follows that $|\eta(x,y)|\le e^{\gamma}\alpha_y(u)R(y)$ for $y \ge y_0$ and $u \in [1,2]$, where

$$\alpha_y(u) := \frac{\log y}{y^u R(y)} + C_3(y_0)\left(\frac{\log y}{y^u} + \log y\int_1^u t^{-1}y^{t-u}\,dt\right) + (1+C_3(y_0)R(y))\left(\frac{1}{u}+y^{1-u}\right).$$

Integration by parts enables us to write

$$\log y \int_1^u t^{-1}y^{t-u}\,dt = \frac{1}{u} - y^{1-u} + \int_1^u t^{-2}y^{t-u}\,dt$$

for $y \ge y_0$. Hence $|\eta(x,y)|\le e^{\gamma}\eta_1(y)R(y)$ for $y \ge y_0$ and $u \in [1,2]$, where

$$\eta_1(y) := \sup_{t\ge y}\frac{\log t}{tR(t)} + \max_{u\in[1,2]}\left(C_3(y_0)I_y(u) + (1+C_3(y_0)R(y))\left(\frac{1}{u}+y^{1-u}\right)\right) \quad (2.3.19)$$

with

$$I_y(u) := \frac{1}{u} + \int_1^u t^{-2}y^{t-u}\,dt. \quad (2.3.20)$$

We remark that $I_y(u)$ is strictly decreasing on $[1,2]$ and hence satisfies $I_y(u) < 1$ for $u \in (1,2]$, since its derivative is

$$I_y'(u) = -\int_1^u t^{-2}y^{t-u}\log y\,dt < 0.$$

Thus, (2.3.19) simplifies to

$$\eta_1(y) = \sup_{t\geq y} \frac{\log t}{tR(t)} + C_3(y_0) + 2(1 + C_3(y_0)R(y)). \tag{2.3.21}$$

Suppose now that $y \geq y_0$ and $u \geq 2$. From (2.3.6) it follows that

$$\psi(x, y) = \psi(x, z)\frac{Q(z)}{Q(y)} + \sum_{y<p\leq z} \psi(x/p, p^-) \cdot \frac{1}{p} \prod_{y<q<p} \left(1 - \frac{1}{q}\right), \tag{2.3.22}$$

where $z \geq y \geq y_0$. Put $h := \log z/\log y \geq 1$ and

$$H_y(v) := \sum_{y<p\leq y^v} \frac{1}{p} \prod_{y<q<p} \left(1 - \frac{1}{q}\right) \tag{2.3.23}$$

for $v \geq 1$. Then we have $H_y(v) = 1 - Q(y^v)/Q(y)$. By partial summation, we see that (2.3.22) becomes

$$\psi(x, y) = \psi(y^u, y^h)(1 - H_y(h)) + \int_1^h \psi(y^{u-v}, (y^v)^-)\, dH_y(v). \tag{2.3.24}$$

By (2.3.16) we have

$$|H_y(v) - 1 + v^{-1}| \leq C_7(y_0)R(y),$$

where $C_7(y_0) := \max(C_3(y_0), C_4(y_0))$. Thus, one can think of $1 - v^{-1}$ as a smooth approximation to $H_y(v)$. Since we also expect $\lambda(x, y)$ to be a smooth approximation to $\psi(x, y)$, in view of (2.3.24) it is reasonable to expect

$$E_1(h; y, u) := \lambda(y^u, y) - \lambda(y^u, y^h)h^{-1} - \int_1^h \lambda(y^{u-v}, y^v)v^{-2}\, dv$$

to be small in size as a function of y. This can be easily verified when $1 \leq h \leq u/2$.

Following de Bruijn [10], we have

$$\frac{\partial}{\partial h} E_1(h; y, u) = -h^{-1} \cdot \frac{\partial}{\partial h} \lambda(y^u, y^h) + h^{-2}\lambda(y^u, y) - h^{-2}\lambda(y^{u-h}, y^h). \tag{2.3.25}$$

Since

$$\frac{\lambda(y^u, y^h)}{e^\gamma \log y} = h \int_1^{u/h} y^{ht-u}\omega(t)\, dt,$$

we find

$$\frac{\partial}{\partial h}\left(\frac{\lambda(y^u, y^h)}{e^\gamma \log y}\right) = \int_1^{u/h} y^{ht-u}\omega(t)\, dt + h\left(\log y \int_1^{u/h} y^{ht-u}(t\omega(t))\, dt - uh^{-2}\omega(uh^{-1})\right).$$

Recall that $(t\omega(t))' = \omega(t-1)$ for $t \in \mathbb{R}$ with the obvious extension $\omega(t) = 0$ for $t < 1$. It follows that

$$\begin{aligned}
\log y \int_1^{u/h} y^{ht-u}(t\omega(t))\, dt &= h^{-1}y^{ht-u}(t\omega(t))\Big|_1^{u/h} - h^{-1}\int_1^{u/h} y^{ht-u}\omega(t-1)\, dt \\
&= uh^{-2}\omega(uh^{-1}) - h^{-1}y^{h-u} - h^{-1}y^h \int_1^{u/h-1} y^{ht-u}\omega(t)\, dt \\
&= uh^{-2}\omega(uh^{-1}) - h^{-1}y^{h-u} - \left(h^2 e^\gamma \log y\right)^{-1} \lambda(y^{u-h}, y^h).
\end{aligned}$$

Hence we have

$$\begin{aligned}
\frac{\partial}{\partial h}\lambda(y^u, y^h) &= e^\gamma \log y \left(\int_1^{u/h} y^{ht-u}\omega(t)\, dt - y^{h-u}\right) - h^{-1}\lambda(y^{u-h}, y^h) \\
&= h^{-1}\lambda(y^u, y^h) - e^\gamma y^{h-u}\log y - h^{-1}\lambda(y^{u-h}, y^h).
\end{aligned}$$

Inserting this in (2.3.25) yields

$$\frac{\partial}{\partial h} E_1(h; y, u) = h^{-1}e^\gamma y^{h-u}\log y.$$

Integrating both sides with respect to h and using the initial value condition $E_1(1; y, u) = 0$, we obtain

$$E_1(h; y, u) = e^{\gamma} \log y \int_1^h t^{-1} y^{t-u} \, dt < e^{\gamma} y^{h-u}. \qquad (2.3.26)$$

In what follows, we shall always suppose that $1 \le h \le u/2$. Following de Bruijn [10], we proceed to show that

$$E_3(h; y, u) := \lambda(y^u, y) - \lambda(y^u, y^h)(1 - H(h)) - \int_1^h \lambda(y^{u-v}, y^v) \, dH(h)$$

is small in size as a function of y. This is intuitive, because

$$\lambda(y^u, y^h) h^{-1} - \int_1^h \lambda(y^{u-v}, y^v) v^{-2} \, dv,$$

which is a good approximation to $\lambda(y^u, y)$ as we have already demonstrated, can be thought of as a smooth approximation to

$$\lambda(y^u, y^h)(1 - H(h)) - \int_1^h \lambda(y^{u-v}, y^v) \, dH(h).$$

Moreover, we have by (2.3.24) that

$$\eta(x, y) = \eta(y^u, y^h)(1 - H_y(h)) + \int_1^h \eta(y^{u-v}, (y^v)^-) \, dH_y(v) - E_3(h; y, u), \qquad (2.3.27)$$

which will later be used to estimate $\eta(x, y)$. To estimate $E_3(h; y, u)$, let us write $E_3(h; y, u) = E_1(h; y, u) + E_2(h; y, u)$, where

$$E_2(h; y, u) := -\int_1^h \lambda(y^{u-v}, y^v) \, d\left(H(v) - 1 + v^{-1}\right) + (H(h) - 1 + h^{-1}) \lambda(y^u, y^h).$$

Then we expect $E_2(h; y, u)$ to be small in size as a function of y. Using (2.3.23) and the observation that $H(1) = 0$, we have

$$|E_2(h; y, u)| \leq \left(\left|\lambda(y^u, y^h) - \lambda(y^{u-h}, y^h)\right| + \int_1^h \left| \frac{\partial}{\partial v} \lambda(y^{u-v}, y^v) \right| dv \right) C_7(y_0) R(y). \tag{2.3.28}$$

Note that

$$\begin{aligned} \frac{\lambda(y^u, y^h) - \lambda(y^{u-h}, y^h)}{he^\gamma \log y} &= \int_1^{u/h} y^{ht-u} \omega(t)\, dt - \int_2^{u/h} y^{ht-u} \omega(t-1)\, dt \\ &= \int_1^2 y^{ht-u} \omega(t)\, dt + \int_2^{u/h} y^{ht-u} (\omega(t) - \omega(t-1))\, dt \\ &= \int_1^2 t^{-1} y^{ht-u}\, dt - \int_2^{u/h} y^{ht-u} t\omega'(t)\, dt. \end{aligned}$$

By Theorems III.5.7 and III.6.6 in [56] we have

$$|\omega'(t)| \leq \rho(t) \leq \frac{1}{\Gamma(t+1)} \tag{2.3.29}$$

for all $t \geq 1$. It follows that

$$\left|\lambda(y^u, y^h) - \lambda(y^{u-h}, y^h)\right| \leq he^\gamma \log y \left(\int_1^2 t^{-1} y^{ht-u}\, dt + \int_2^{u/h} y^{ht-u} t\rho(t)\, dt \right). \tag{2.3.30}$$

This inequality will later be used in conjunction with the formulas

$$h \log y \int_1^2 t^{-1} y^{ht-u}\, dt = \frac{y^{2h-u}}{2} - y^{h-u} + \int_1^2 t^{-2} y^{ht-u}\, dt \tag{2.3.31}$$

and

$$h\log y\int_2^{u/h} y^{ht-u}t\rho(t)\,dt = uh^{-1}\rho(uh^{-1}) - 2\rho(2)y^{2h-u} - \int_2^{u/h} y^{ht-u}(t\rho(t))'\,dt$$
$$\le uh^{-1}\rho(uh^{-1}) - 2\rho(2)y^{2h-u} + \int_2^{u/h} y^{ht-u}\rho(t-1)\,dt. \tag{2.3.32}$$

On the other hand, we have

$$\frac{\lambda(y^{u-v},y^v)}{e^\gamma\log y} = v\int_2^{u/v} y^{vt-u}\omega(t-1)\,dt,$$

which implies that

$$\frac{\partial}{\partial v}\left(\frac{\lambda(y^{u-v},y^v)}{e^\gamma\log y}\right) = \int_2^{u/v} y^{vt-u}(1+tv\log y)\omega(t-1)\,dt - uv^{-1}\omega(uv^{-1}-1).$$

By partial integration, the right side of the above is easily seen to be

$$-2y^{2v-u} - \int_2^{u/v} y^{vt-u}t\omega'(t-1)\,dt.$$

Hence, we arrive at

$$\int_1^h\left|\frac{\partial}{\partial v}\lambda(y^{u-v},y^v)\right|dv \le e^\gamma\log y\left(2\int_1^h y^{2v-u}\,dv + \int_1^h\int_2^{u/v} y^{vt-u}t|\omega'(t-1)|\,dtdv\right).$$

Furthermore, we have by Fubini's theorem that

$$\int_1^h\int_2^{u/v} y^{vt-u}t|\omega'(t-1)|\,dtdv = \int_2^{u/h}\int_1^h y^{vt-u}t|\omega'(t-1)|\,dv\,dt + \int_{u/h}^u\int_1^{u/t} y^{vt-u}t|\omega'(t-1)|\,dv\,dt,$$

the right side of which is easily seen to be

$$\frac{1}{\log y}\left(\int_2^{u/h} y^{ht-u}|\omega'(t-1)|\,dt + \int_{u/h}^{u} |\omega'(t-1)|\,dt - \int_2^{u} y^{t-u}|\omega'(t-1)|\,dt\right).$$

It follows that

$$\int_1^h \left|\frac{\partial}{\partial v}\lambda(y^{u-v}, y^v)\right| dv < e^\gamma\left(y^{2h-u} + \int_2^{u/h} y^{ht-u}|\omega'(t-1)|\,dt + \int_{u/h}^{u} |\omega'(t-1)|\,dt\right). \tag{2.3.33}$$

This estimate together with (2.3.30) will lead us to a good estimate for $E_2(h; y, u)$.

Now we derive estimates for $E_3(h; y, u)$ that suit our needs. Suppose that $k \le u < k+1$ and take $h = h_k = u/k$, where $k \ge 2$ is a positive integer. We first consider the case $k = 2$. In view of (2.3.31), we see that (2.3.30) simplifies to

$$\left|\lambda\left(y^u, y^{h_2}\right) - \lambda\left(y^{u-h_2}, y^{h_2}\right)\right| < e^\gamma\left(\frac{1}{2} + \int_1^2 t^{-2} y_0^{t-2}\,dt\right) = e^\gamma I_{y_0}(2)$$

for $y \ge y_0$ (see (2.3.20) for the definition of $I_{y_0}(2)$). By (2.3.33) we have

$$\int_1^{h_2} \left|\frac{\partial}{\partial v}\lambda(y^{u-v}, y^v)\right| dv \le e^\gamma\left(1 + \int_2^3 |\omega'(t-1)|\,dt\right) = \frac{3e^\gamma}{2},$$

since $\omega'(t) = -1/t^2$ for $t \in [1, 2)$. Combining these estimates with (2.3.26) and (2.3.28), we obtain $E_3(h_2; y, u) \le e^\gamma \xi_2(y_0) R(y)$ for $y \ge y_0$ and $2 \le u < 3$, where

$$\xi_2(y_0) := \max_{t \ge y_0} \frac{1}{tR(t)} + C_7(y_0)\left(I_{y_0}(2) + \frac{3}{2}\right).$$

Now we handle the case $k \ge 3$. From (2.3.29)–(2.3.32) it follows that

$$\left|\lambda\left(y^u, y^{h_k}\right) - \lambda\left(y^{u-h_k}, y^{h_k}\right)\right| < e^\gamma\left(\frac{1}{\Gamma(k)} + \left(2\log 2 - \frac{3}{2}\right) y^{2-k} + \int_1^2 t^{-2} y^{t-k}\,dt \right.$$
$$\left. + \int_2^3 y^{t-k}(1 - \log(t-1))\,dt + \int_3^k y^{t-k}\frac{dt}{\Gamma(t)}\right),$$

where we have used the fact that $\rho(t) = 1 - \log t$ for $t \in [1, 2]$. By (2.3.29) and (2.3.33) we have

$$\int_1^{h_k} \left| \frac{\partial}{\partial v} \lambda(y^{u-v}, y^v) \right| dv \le e^\gamma \left(y^{2-k} + \int_2^3 y^{t-k} \frac{dt}{(t-1)^2} + \int_3^k y^{t-k} \frac{dt}{\Gamma(t)} + \int_k^{k+1} \frac{dt}{\Gamma(t)} \right).$$

Together with (2.3.26) and (2.3.28), these inequalities imply that $E_3(h_k; y, u) \le e^\gamma \xi_k(y_0) R(y)$ for $y \ge y_0$ and $3 \le k \le u < k+1$, where

$$\begin{aligned}\xi_k(y_0) := &\left(\max_{t \ge y_0} \frac{1}{tR(t)} \right) y_0^{2-k} + C_7(y_0) \left(\frac{1}{(k-1)!} + \int_k^{k+1} \frac{dt}{\Gamma(t)} + \left(2\log 2 - \frac{1}{2} \right) y_0^{2-k} \right. \\ &\left. + \int_1^2 t^{-2} y_0^{t-k}\, dt + \int_2^3 y_0^{t-k} \left(1 - \log(t-1) + \frac{1}{(t-1)^2} \right) dt + 2 \int_3^k y_0^{t-k} \frac{dt}{\Gamma(t)} \right).\end{aligned}$$

As a direct corollary, we obtain

$$\begin{aligned}\sum_{k=2}^{\infty} \xi_k(y_0) = &\frac{y_0}{y_0 - 1} \max_{t \ge y_0} \frac{1}{tR(t)} + C_7(y_0) \left(e - \frac{1}{2} + \int_3^\infty \frac{dt}{\Gamma(t)} + \frac{1}{y_0 - 1} \left(2\log 2 - 1 \right.\right. \\ &\left.\left. + y_0 I_{y_0}(2) + \int_2^3 y_0^{t-2} \left(1 - \log(t-1) + \frac{1}{(t-1)^2} \right) dt + 2 \int_3^\infty y_0^{\{t\}} \frac{dt}{\Gamma(t)} \right) \right),\end{aligned}$$

where we have applied partial summation to derive

$$\begin{aligned}\sum_{k=3}^{\infty} \int_3^k y_0^{t-k} \frac{dt}{\Gamma(t)} &= \left(\sum_{k=3}^{\infty} y_0^{-k} \right) \int_3^\infty y_0^t \frac{dt}{\Gamma(t)} - \int_3^\infty \left(\sum_{3 \le k \le t} y_0^{-k} \right) y_0^t \frac{dt}{\Gamma(t)} \\ &= \frac{y_0^{-3}}{1 - y_0^{-1}} \int_3^\infty y_0^t \frac{dt}{\Gamma(t)} - \int_3^\infty \frac{y_0^{t-3}(1 - y_0^{-\lfloor t \rfloor + 2})}{1 - y_0^{-1}} \cdot \frac{dt}{\Gamma(t)} \\ &= \frac{1}{y_0 - 1} \int_3^\infty y_0^{\{t\}} \frac{dt}{\Gamma(t)}.\end{aligned}$$

For computational purposes, we can transform the last integral above by observing that

$$\int_3^\infty y_0^{\{t\}} \frac{dt}{\Gamma(t)} = \int_0^1 \left(\sum_{n=0}^{\infty} \frac{1}{(t+2)\cdots(t+2+n)} \right) y_0^t \frac{dt}{\Gamma(t+2)}.$$

Let

$$\gamma(s, z) := \int_0^z v^{s-1} e^{-v}\, dv$$

be the lower incomplete gamma function, where $s \in \mathbb{C}$ with $\Re(s) > 0$ and $z \geq 0$. It is well known that

$$\gamma(s,z) = z^s e^{-z} \sum_{n=0}^{\infty} \frac{z^n}{s(s+1)\cdots(s+n)},$$

from which it follows that

$$\sum_{n=0}^{\infty} \frac{1}{(t+2)\cdots(t+2+n)} = \gamma(t+2,1)e.$$

Thus we obtain

$$\begin{aligned}\sum_{k=2}^{\infty} \xi_k(y_0) = {} & \frac{y_0}{y_0-1} \max_{t \geq y_0} \frac{1}{tR(t)} \\ & + C_7(y_0) \left(e - \frac{1}{2} + \int_3^{\infty} \frac{dt}{\Gamma(t)} + \frac{1}{y_0-1} \left(2\log 2 - 1 + y_0 I_{y_0}(2) \right.\right. \\ & \left.\left. + \int_2^3 y_0^{t-2} \left(1 - \log(t-1) + \frac{1}{(t-1)^2} \right) dt + 2e \int_0^1 y_0^t \frac{\gamma(t+2,1)}{\Gamma(t+2)}\, dt \right) \right).\end{aligned} \tag{2.3.34}$$

In Mathematica, the function $\gamma(t+2,1)$ can be evaluated by "Gamma[t+2,0,1]".

Finally, we are ready to estimate $\eta(x,y)$. Let

$$\eta_k(y) := \frac{1}{e^{\gamma} R(y)} \sup_{\substack{u \in [k,k+1) \\ t \geq y}} |\eta(t^u, t)|$$

for $k \geq 1$ and $y \geq y_0$, where the value of $\eta_1(y)$ is provided by (2.3.21). Using (2.3.27) and the estimates for $E_3(h_k; y, u)$ with $y \geq y_0$ and $2 \leq k \leq u < k+1$, we find

$$\eta_k(y) \leq \eta_{k-1}(y) + \xi_k(y_0)$$

for all $k \geq 2$ and $y \geq y_0$, from which we derive

$$\eta_k(y) \leq \eta_1(y) + \sum_{\ell=2}^{k} \xi_\ell(y_0)$$

for all $k \geq 1$ and $y \geq y_0$. Since $\eta_1(y)$ is decreasing on $[y_0, \infty)$, we have therefore shown that

$$|\eta(x,y)| \leq e^{\gamma}\left(\eta_1(y_0) + \sum_{k=2}^{\infty} \xi_k(y_0)\right) R(y) \tag{2.3.35}$$

for all $y \geq y_0$, where the infinite sum can be evaluated using (2.3.34). To derive an explicit version of de Bruijn's result (1.1.7), we observe that (2.3.18), (2.3.35) and [50, Theorem 23] imply that $Q(y)|\eta(x,y)| \leq C_8(y_0)R(y)/\log y$ for all $y \geq y_0$, where

$$C_8(y_0) := \beta(y_0)\left(\eta_1(y_0) + \sum_{k=2}^{\infty} \xi_k(y_0)\right)$$

with

$$\beta(y_0) := \begin{cases} 1, & \text{if } 3 \leq y_0 < 10^8, \\ \exp(C_2(y_0)R(y_0)), & \text{if } y_0 \geq 10^8. \end{cases}$$

Hence, it follows that

$$\left|\Phi(x,y) - \mu_y(u)e^{\gamma}x\log y \prod_{p\leq y}\left(1 - \frac{1}{p}\right)\right| < \frac{C_8(y_0)xR(y)}{\log y} \tag{2.3.36}$$

for all $y \geq y_0$.

2.3.4. Deduction of Theorem 2.1.2 and Corollary 2.1.3

Now we apply (2.3.36) to obtain explicit estimates for $\Phi(x,y)$ with specific choices of $R(y)$. Unconditionally, it has been shown [43, Corollary 2] that

$$|\pi(z) - \mathrm{li}(z)| \leq 0.2593\frac{z}{(\log z)^{3/4}}\exp\left(-\sqrt{\frac{\log z}{6.315}}\right)$$

for all $z \geq 229$. With $y_0 \geq 229$, the function

$$R(z) = 0.2593(\log z)^{1/4} \exp\left(-\sqrt{\frac{\log z}{6.315}}\right)$$

is strictly decreasing on $[y_0, \infty)$ and satisfies (1.1.5) and (1.1.6) with

$$C_0(y_0) = 2\sqrt{\frac{6.315}{\log y_0}},$$

since

$$\begin{aligned}
\int_z^\infty \frac{1}{t(\log t)^{3/4}} \exp\left(-\sqrt{\frac{\log t}{6.315}}\right) dt &= 2\int_{\sqrt{\log z}}^\infty \frac{1}{\sqrt{t}} \exp\left(-\frac{t}{\sqrt{6.315}}\right) dt \\
&< \frac{2}{(\log z)^{1/4}} \int_{\sqrt{\log z}}^\infty \exp\left(-\frac{t}{\sqrt{6.315}}\right) dt \\
&= \frac{2\sqrt{6.315}}{(\log z)^{1/4}} \exp\left(-\sqrt{\frac{\log z}{6.315}}\right)
\end{aligned}$$

for $z \geq y_0$. Numerical computation by Mathematica allows us to conclude that

$$\left|\Phi(x,y) - \mu_y(u)e^\gamma x \log y \prod_{p \leq y}\left(1 - \frac{1}{p}\right)\right| < 4.403611 \frac{x}{(\log y)^{3/4}} \exp\left(-\sqrt{\frac{\log y}{6.315}}\right) \tag{2.3.37}$$

for all $x \geq y \geq 229$. Suppose now that $2 \leq y < 229$. Using the inequalities $\Phi(x,y) < x/\log y$ [27, Theorem], $\prod_{p \leq y}(1 - 1/p) < e^{-\gamma}/\log y$ [50, Theorem 23] and $0 \leq \mu_y(u) < 1/\log y$, we have, for all $2 \leq y < 229$,

$$\left|\Phi(x,y) - \mu_y(u)e^\gamma x \log y \prod_{p \leq y}\left(1 - \frac{1}{p}\right)\right| < \frac{2x}{\log y} < 4.403611 \frac{x}{(\log y)^{3/4}} \exp\left(-\sqrt{\frac{\log y}{6.315}}\right).$$

Combining this with (2.3.37) proves the first half of Theorem 2.1.2.

Under the assumption of the Riemann Hypothesis, it is known [52, Corollary 1] that

$$|\pi(z) - \mathrm{li}(z)| < \frac{1}{8\pi}\sqrt{z}\log z$$

for all $z \geq 2{,}657$. With $y_0 = 2{,}657$ and $R(z) = \log^2 z/(8\pi\sqrt{z})$, we have

$$\int_z^\infty \frac{|\pi(t) - \mathrm{li}(t)|}{t^2}\,dt \leq \frac{1}{8\pi}\int_z^\infty \frac{\log t}{t^{3/2}}\,dt = \frac{\log z + 2}{4\pi\sqrt{z}} \leq C_0(y_0)R(z)$$

for $z \geq y_0$, where

$$C_0(y_0) = \frac{2(\log y_0 + 2)}{\log^2 y_0}.$$

Therefore, we conclude by (2.3.36) and numerical calculations that

$$\left|\Phi(x,y) - \mu_y(u)e^\gamma x\log y\prod_{p\leq y}\left(1 - \frac{1}{p}\right)\right| < 0.184563\frac{x\log y}{\sqrt{y}} \tag{2.3.38}$$

for all $x \geq y \geq 2{,}657$. The values of relevant constants are recorded in the table below.

Table 2.2: Numerical Constants

constants	unconditional estimates		conditional estimates	
y_0	229	10^8	2,657	10^8
$R(y_0)$	.156576	.097363	.047992	.001351
$C_0(y_0)$	2.156096	1.171019	.317985	.120362
$C_1(y_0)$	2.534430	1.279593	.571800	.228936
$C_2(y_0)$	2.548436	1.279593	.575723	.228940
$C_3(y_0)$	3.110976	1.362717	.579718	.228971
$C_4(y_0)$	2.548436	1.279593	.575723	.228940
$C_5(y_0)$	3.131827	1.362717	.583750	.228975
$C_6(y_0)$	2.534430	1.279593	.571800	.228936
$C_7(y_0)$	3.110976	1.362717	.579718	.228971
$C_8(y_0)$	16.982691	9.079975	4.638553	2.967998
$\eta_1(y_0)$	6.236726	3.628074	2.697198	2.229726
$\sum_{k=2}^\infty \xi_k(y_0)$	10.745960	4.388310	1.941356	.737355

To complete the proof of the second half of Theorem 2.1.2, it remains to deal with the case $11 \le y \le 2{,}657$. For simplicity of notation we set

$$D(x,y) := \Phi(x,y) - \mu_y(u) e^{\gamma} x \log y \prod_{p \le y} \left(1 - \frac{1}{p}\right).$$

Using Mathematica we find that

$$M := \max_{11 \le z \le 2{,}657} \frac{\mathrm{li}(z) - \pi(z)}{\sqrt{z}\log z} < 0.259141,$$
$$m := \min_{11 \le z \le 2{,}657} e^{\gamma} \log z \prod_{p \le z} \left(1 - \frac{1}{p}\right) > 0.876248.$$

If $\sqrt{x} \le y < x$, then

$$\Phi(x,y) = \mu_y(u)x + (\pi(x) - \mathrm{li}(x)) - (\pi(y) - \mathrm{li}(y)) + 1.$$

Note that $x \le y^2 < 10^8$. Since $\pi(z) < \mathrm{li}(z)$ for $2 \le z \le 10^8$ by [50, Theorem 16] and

$$\prod_{p \le z} \left(1 - \frac{1}{p}\right) < \frac{e^{-\gamma}}{\log z}$$

for $0 < z \le 10^8$ by [50, Theorem 23], we have

$$\begin{aligned} |D(x,y)| &< (1-m)\left(1 - y^{-1}\right) \frac{x}{\log y} + M\sqrt{x}\log x + 1 \\ &\le \left((1-m)\left(1-y^{-1}\right) + M\frac{\log^2 y}{\sqrt{y}} + \frac{\log y}{y}\right) \frac{x}{\log y}, \end{aligned} \tag{2.3.39}$$

where we have used the fact that $\log x/\sqrt{x}$ is strictly decreasing on $[e^2, \infty)$. Numerical computation shows that the right side of (2.3.39) is $< 0.449774 x \log y/\sqrt{y}$ for $11 \le y \le 2{,}657$. Suppose now that $11 \le y \le \sqrt{x}$. By Theorem 2.1.1, Theorem 2.3.3 and

[50, Theorem 23] we have, for $11 \le y \le 2{,}657$,

$$D(x,y) \le \left(0.6 - \frac{m}{2}\left(1 - y^{-1}\right)\right)\frac{x}{\log y} < 0.449774\frac{x\log y}{\sqrt{y}},$$
$$D(x,y) > (0.4 - M_0)\frac{x}{\log y} > -0.449774\frac{x\log y}{\sqrt{y}}.$$

This settles the case $11 \le y \le 2{,}657$ and completes the proof of Theorem 2.1.2.

The proof of Corollary 2.1.3 is similar, and we shall only sketch it. When $y \ge y_0$, where $y_0 = 229$ for the unconditional estimate and $y_0 = 2{,}657$ for the conditional estimate, we have by the triangle inequality that

$$|\Phi(x,y) - \mu_y(u)x| < |D(x,y)| + \left|1 - e^{\gamma}\log y\prod_{p\le y}\left(1 - \frac{1}{p}\right)\right|\frac{x}{\log y}.$$

Then we bound $|D(x,y)|$ using the values of $C_8(y_0)$ listed in Table 2.2. To estimate the second term, we use (2.3.18) when $y \ge 10^8$ and the inequality

$$m(y) < e^{\gamma}\log y\prod_{p\le y}\left(1 - \frac{1}{p}\right) < 1$$

when $y_0 \le y \le 10^8$, where $m(y)$ is given by

$$m(y) := \begin{cases} 0.983296, & \text{if } 229 \le y \le 2{,}657, \\ 0.996426, & \text{if } 2{,}657 \le y < 210{,}000, \\ 0.999643, & \text{if } 210{,}000 \le y \le 10^8, \end{cases}$$

according to [50, Theorem 23] and Mathematica. This leads to the asserted bounds for $y \ge y_0$. Suppose now that $y \le y_0$. In this case, the proof of the unconditional bound is exactly the same as that of the unconditional bound in Theorem 2.1.2. As for the conditional bound, we argue in the same way as in the proof of Theorem 2.1.2

to get

$$|\Phi(x,y) - \mu_y(u)x| \leq \left(M\frac{\log^2 y}{\sqrt{y}} + \frac{\log y}{y} \right) \frac{x}{\log y}$$

when $\sqrt{x} \leq y < x$ and

$$|\Phi(x,y) - \mu_y(u)x| \leq \left(0.6 - \frac{1}{2}\left(1 - y^{-1}\right) \right) \frac{x}{\log y},$$
$$|\Phi(x,y) - \mu_y(u)x| > (0.4 - M_0)\frac{x}{\log y},$$

when $11 \leq y \leq \sqrt{x}$. Together, these inequalities yield the asserted conditional bound.

Remark 2.3.2. The bounds in Theorem 2.1.2 and its corrollary may be improved. For example, the numerical values of the sum $\sum_{k=2}^{\infty} \xi_k(y_0)$ may be reduced by keeping ρ (or even $|\omega'|$) in all of the relevant integrals, but of course the computational complexity is expected to increase as a cost. In addition, our method would allow an extension of the range $x \geq y \geq 11$ in the second half of Theorem 2.1.2 to the entire range $x \geq y \geq 2$ if we argue with $y_0 = 2{,}657$ replaced by some smaller value and enlarge the constant 0.449774.

Chapter 3

The Weighted Erdős–Kac Theorems

In this chapter we study the distribution of additive functions weighted by nonnegative multiplicative functions. The main purpose is to establish weighted versions of the Erdős–Kac theorem by generalizing the method of moments of Granville, Soundararajan, Khan, Milinovich and Subedi. Compared to previous approaches to computing moments, this sieve-theoretic approach allows one to identify easily the main term in the asymptotic of the mth moment and obtain asymptotic formulas uniformly in a wide range of m. And we shall take great advantage of these benefits in our treatment as well. As a result, we are able to obtain, without much complication, results which are applicable to a wide class of multiplicative functions, and in particular, imply the theorem of Elboim and Gorodetsky.

In Section 3.1, we reveal the class $\mathcal{M}^*$ of nonnegative multiplicative functions which we alluded to in Section 1.1 and discuss some interesting examples. In Section 3.2, we introduce additional definitions and notation and state our main results in a coherent way. Proofs of these results will be presented in detail in Sections 3.3–3.9. In Sections 3.10 and 3.11, respectively, we describe how Corollary 1.2.1 and 1.2.2 can be derived from our results. We close this chapter with a brief discussion on related problems and possible generalizations.

Before delving into the technical material, we introduce some commonly used terminologies and notation in analysis and number theory that will also be adopted throughout this chapter. Given any real or complex valued functions $f(x)$ and $g(x)$ with a common domain $\mathcal{D} \subseteq \mathbb{R}$, we shall use Landau's big-O notation $f(x) = O(g(x))$ or Vinogradov's notation $f(x) \ll g(x)$ to mean that there exists an absolute constant $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all $x \in \mathcal{D}$. Likewise, we shall use the notation $f(x) \gg g(x)$ interchangeably with $g(x) = O(f(x))$. If $f(x) = O(g(x))$ and $g(x) = O(f(x))$ hold simultaneously, then we adopt the short-hand notation $f(x) \asymp g(x)$. If $\mathcal{D}$ contains a neighborhood of ∞ and $f(x)/g(x) \to 0$ as $x \to \infty$, then we write $f(x) = o(g(x))$. Similarly, we write $f(x) \sim g(x)$ if $f(x)/g(x) \to 1$ as $x \to \infty$. We shall occasionally make use of the function $\epsilon_{a,b}$ defined by

$$\epsilon_{a,b} := \begin{cases} 0, & \text{if } a = b, \\ 1, & \text{otherwise,} \end{cases}$$

for any $a, b \in \mathbb{R}$. Equivalently, $\epsilon_{a,b} = 1 - \delta_{a,b}$, where $\delta_{a,b}$ is the Kronecker delta function.

Throughout, the letter p always denotes a prime, and we write $\pi(x)$ for the prime-counting function, namely, $\pi(x) = \sum_{p \leq x} 1$. For any $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ for the integer part of x and $\lceil x \rceil$ for the least integer $\geq x$. For every $n \in \mathbb{N}$, we denote by $P^-(n)$ and $P^+(n)$ the least and the greatest prime factor of n, respectively, with the convention that $P^-(1) = \infty$ and $P^+(1) = 1$. We say that $n \in \mathbb{Z} \setminus \{0\}$ is *squareful*, *square-full*, or *powerful* if for any prime $p \mid n$, one has $p^2 \mid n$. Given any prime power p^ν, the relation $p^\nu \parallel n$ means that $p^\nu \mid n$ but $p^{\nu+1} \nmid n$. Thus n is *squarefree* if every prime divisor p of n satisfies $p \parallel n$. In addition, we denote by R_n the *radical* or *squarefree kernel* of n, i.e.,

$$R_n := \mathrm{rad}(n) = \prod_{p|n} p.$$

Finally, we write

$$\binom{m}{m_1, ..., m_k} := \frac{m!}{m_1! \cdots m_k!}$$

for the multinomial coefficient of shape $(m_1, ..., m_k)$ of size $m = m_1 + \cdots + m_k$.

Section 3.1

The Class $\mathcal{M}^*$

The weight functions $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ that we shall consider throughout the chapter form a nice subclass $\mathcal{M}^*$ of nonnegative multiplicative functions, nice in the sense that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0, 1)$ and $r \in (0, 1)$, such that the following conditions hold:

$$\text{(i)} \quad \alpha(p^\nu) \ll p^{(\varrho_0+\sigma_0-1)\nu}, \tag{3.1.1}$$

$$\text{(ii)} \quad \sum_{p \leq x} \frac{\alpha(p) \log p}{p^{\sigma_0 - 1}} = \beta x + O\left(\frac{x}{(\log x)^{A_0}}\right), \tag{3.1.2}$$

$$\text{(iii)} \quad {\sum_p}' \left(\frac{\alpha(p)^2}{p^{2(r+\sigma_0-1)}} + \sum_{\nu \geq 2} \frac{\alpha(p^\nu)}{p^{(r+\sigma_0-1)\nu}}\right) < \infty, \tag{3.1.3}$$

$$\text{(iv)} \quad \sum_{\nu \geq 1} \frac{\nu \alpha(p^\nu)}{p^{\sigma_0 \nu}} \ll \frac{(\log\log(p+1))^{\vartheta_0}}{p}, \tag{3.1.4}$$

where the restricted sum ${\sum_p}'$ is over all but finitely many primes p. Note that due to the restricted sum in condition (iii), we can ignore the primes for which $\sum_{\nu \geq 2} \alpha(p^\nu)/p^{(r+\sigma_0-1)\nu} = \infty$. It is not hard to verify that $\mathcal{M}^*$ is closed under Dirichlet convolution. In particular, condition (ii) can be viewed as a weighted version of the Prime Number Theorem. As we shall see in the next section, conditions (i)–(iii) enable us to obtain, in general, an almost optimal estimate for the partial sum of $\alpha(n)$, thanks to [11, Theorem 2.1], which is also one of the key ingredients in the argument of Elboim and Gorodetsky [18]. Condition (iv) essentially speaks about the growth rate of the local factors of the Dirichlet series $F(s) = \sum_{n\geq 1} \alpha(n)n^{-s}$ at $s = \sigma_0$. More precisely, if we denote by

$$F_\alpha(s;p) := \sum_{\nu\geq 0} \frac{\alpha(p^\nu)}{p^{\nu s}}$$

the local factor of $F(s)$ at p, then condition (iv) is equivalent to

$$\frac{F'_\alpha(\sigma_0;p)}{\log p} \ll \frac{(\log\log(p+1))^{\vartheta_0}}{p},$$

where $F'_\alpha(\sigma_0;p)$ is the derivative of $F_\alpha(s;p)$ with respect to s evaluated at $s = \sigma_0$. Like conditions (ii) and (iii), this condition places a holistic constraint on the growth of $\alpha(p^\nu)$, and it is one of the types that we expect to hold for many multiplicative functions of interest. It may be worth noting that $\mathcal{M}^*$ properly contains the subclass of multiplicative functions considered by Elboim and Gorodetsky [18]. A simple example which falls into $\mathcal{M}^*$ but is not covered by the theorem of Elboim and Gorodetsky is the multiplicative function $\alpha(n)$ defined by $\alpha(p) = 1$ for all primes p and $\alpha(p^\nu) = p^{\nu/3}$ for all prime powers p^ν with $\nu \geq 2$.

Some familiar multiplicative functions which belong to $\mathcal{M}^*$ are: the power function n^λ, the cth power of the κ-fold divisor function $d_\kappa(n)^c$, the sum-of-divisors function $\sigma_\lambda(n)$, Euler's totient function $\varphi(n)$, the functions $\kappa_1^{\omega(n)}$ and $\kappa_2^{\Omega(n)}$, the characteristic function $\mu(n)^2$ of square-free numbers, the function $r_2(n)/4$, and the function which counts the number of positive divisors of n representable as a sum of two integral squares, where $c \in \mathbb{R}$, $\lambda > -1$, $\kappa, \kappa_1 > 0$, $\kappa_2 \in (0,2)$, $\mu(n)$ is the Möbius function, and $r_2(n) := \#\{(a,b) \in \mathbb{Z}^2 : n = a^2 + b^2\}$.

Perhaps a less obvious example is $\rho_g(n)$, which denotes the number of zeros of a nonconstant irreducible polynomial $g \in \mathbb{Z}[x]$ in $\mathbb{Z}/n\mathbb{Z}$. The Chinese remainder theorem implies that ρ_g is multiplicative. For this particular example one can take $A_0 = \beta = \sigma_0 = 1$, $\vartheta_0 = \varrho_0 = 0$, and $r \in (1/2, 1)$ to be any positive number. The interested reader is referred to [22, Lemmas 3,7] and [33, Lemma 1] for more details.

We conclude this section with another interesting example, which arises from the theory modular forms. Consider $\alpha(n) = \tau(n)^2$, where $\tau(n)$ is the Ramanujan τ-function, which may be defined as the nth Fourier coefficient of the modular discriminant $\Delta(z)$, i.e.,

$$\Delta(z) := q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_{n=1}^{\infty}\tau(n)q^n.and$$

Ramanujan [47] conjectured that $\tau(n)$ is multiplicative, that $\tau(p^{\nu+1}) = \tau(p)\tau(p^{\nu}) - p^{11}\tau(p^{\nu-1})$ for all primes p and all $\nu \in \mathbb{N}$, and that $|\tau(p)| \leq 2p^{11/2}$ for all primes p. As he pointed out, these conjectures would imply that $|\tau(n)| \leq n^{11/2}d(n)$. The first two conjectures were proved by Mordell [41] in 1917, and the third one was proved by Deligne [16] in 1974 as a consequence of his proof of the Weil conjectures for algebraic varieties over finite fields. In addition, it can be shown [42] that the Dirichlet series $\sum_{n\geq 1}\tau(n)^2 n^{-s-11}$ has the Hoheisel Property. We say that a Dirichlet series $F(s) = \sum_{n\geq 1} a_n n^{-s}$ has the Hoheisel Property if (a) $F(s)$ possesses the explicit formula

$$\sum_{p\leq x} a_p \log p = x - \sum_{|\gamma|\leq T}\frac{x^{\rho}}{\rho} + O\left(\frac{x(\log Tx)^2}{T}\right),$$

where $0 < T \leq \sqrt{x}$ and the sum on the right-hand side runs over all the zeros $\rho = \beta + i\gamma$ of $F(s)$ with $\beta \geq 0$ and $|\gamma| \leq T$, (b) $F(s)$ has a zero-free region $\sigma \geq 1 - c_0/\log(|t|+2)$ for some absolute constant $c_0 > 0$, (c) the number of zeros $\rho = \beta + i\gamma$ of $F(s)$ with $\beta \geq \sigma$ and $|\gamma| \leq T$ is $\ll T^{c_1(1-\sigma)}$ uniformly for all $1/2 \leq \sigma \leq 1$ and all sufficiently large T, where $c_1 > 0$ is an absolute constant, and (d) the number of zeros

$\rho = \beta + i\gamma$ of $F(s)$ with $\beta \geq 0$ and $|\gamma| \leq T$ is $\ll T \log T$ as $T \to \infty$. In particular, (a), (b) and (d) are sufficient for establishing the following analogue of the Prime Number Theorem for $\tau(n)^2/n^{11}$:

$$\sum_{p \leq x} \frac{\tau(p)^2}{p^{11}} \log p = x + O\left(x \exp\left(-c_0\sqrt{\log x}\right)\right)$$

with some absolute constant $c_0 > 0$. From these properties of $\tau(n)$ it follows that $\alpha(n) = \tau(n)^2$ satisfies conditions (i)–(iv) with any fixed $A_0 > 0$, $\beta = 1$, $\sigma_0 = 12$, $\vartheta_0 = 0$, and any fixed $\varrho_0 \in (0,1)$ and $r \in (1/2, 1)$.

Section 3.2

Main Results

Let $\alpha(n)$ be a multiplicative function in the subclass $\mathcal{M}^*$ defined in the previous section, and let

$$S(x) = S_\alpha(x) := \sum_{n \leq x} \alpha(n)$$

be the partial sum of $\alpha(n)$ over $n \leq x$. For any additive function $f\colon \mathbb{N} \to \mathbb{R}$, we may define

$$A(x) = A_{\alpha,f}(x) := \sum_{p \leq x} \alpha(p) \frac{f(p)}{p^{\sigma_0}},$$
$$B(x) = B_{\alpha,f}(x) := \sum_{p \leq x} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}}.$$

If we model $n \leq x$ by a random variable $\mathbf{n}$ defined on the sample space $\mathbb{N} \cap [1, x]$ having a probability distribution with respect to the natural probability measure induced by α, that is to say, $\text{Prob}(\mathbf{n} = k) = \alpha(k)/S(x)$ for every $k \in \mathbb{N} \cap [1, x]$, then one may hope that $f(\mathbf{n})$ also obeys a certain distribution law with respect to the same probability measure under suitable conditions. We shall show, by estimating the weighted mth moment defined by

$$M(x;m) = M_{\alpha,f}(x;m) := S(x)^{-1} \sum_{n \le x} \alpha(n)(f(n) - A(x))^m$$

for every $m \in \mathbb{N}$, that for certain additive functions f, the distribution of $f(n)$ is approximately Gaussian with mean $A(x)$ and variance $B(x)$. More precisely, the limiting distribution of the normalization $(f(n) - A(x))/\sqrt{B(x)}$ of $f(n)$ is standard Gaussian. To state our results in a coherent manner, we set $\chi_m := (1 + (-1)^m)/2$, the characteristic function of even integers, and

$$C_m := \frac{m!}{2^{m/2}\Gamma(m/2+1)}$$

for all $m \in \mathbb{N}$, where Γ is the Gamma function. One quickly notes that $C_m = \mu_m = (m-1)!!$ for m even. Since the numbers C_m play a nonnegligible role in the error terms of our uniform estimates for $M(x;m)$, we find it more convenient to use C_m in place of μ_m. Our first result is the following theorem.

Theorem 3.2.1. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \le M$ for all primes p, where $M > 0$ is an absolute constant. Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\ge 0}$ be a multiplicative function, and suppose that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \ge 0$, $\varrho_0 \in [0,1)$ and $r \in (0,1)$, such that $\alpha(n)$ satisfies the conditions* (i)–(iv).

(a) If $\beta = 1$ and $0 < h_0 < (3/2)^{2/3}$ is arbitrary, and if $B(x) \to \infty$ as $x \to \infty$, then we have

$$M(x;m) = C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{M m^{\frac{3}{2}}}{\sqrt{B(x)}} \right) \right)$$

uniformly for all sufficiently large x and all $1 \le m \le h_0(B(x)/M^2)^{1/3}$.

(b) If $\beta \neq 1$ and if $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$, then we have

$$M(x;m) = C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{Mm^{\frac{3}{2}} \log\log\log x}{\sqrt{B(x)}} \right) \right)$$

uniformly for all sufficiently large x and all $1 \leq m \ll B(x)^{1/3}/(\log\log\log x)^{2/3}$.

The implicit constants in the error terms of both asymptotic formulas above depend at most on the explicit and implicit constants in the hypotheses except for M.

Remark 3.2.1. It may be worth pointing out that as in Theorem 3.2.1, the implicit constants in the estimates appearing in the rest of the paper depend at most on the explicit and implicit constants in the hypotheses unless stated otherwise.

In the case where $\alpha(n) = 1$ and $f(n) = \omega(n)$, we recover [30, Theorem 1] with a slightly more flexible range $1 \leq m \leq h_0(\log\log x)^{1/3}$ compared to the original range $1 \leq m \leq (\log\log x)^{1/3}$. Though Theorem 3.2.1 is formulated for strongly additive functions, similar things can be said about the additive functions whose values at prime powers do not grow too fast and are hence not expected to contribute very much. A simple example of such functions is $\Omega(n)$. Since $\Omega(p^\nu) = \nu$ for all p^ν, one can show, by establishing (1.2.2), that $\Omega(n)$ does not differ from its cousin $\omega(n)$ very much for "most" values of n, and so they are expected to have the same distribution. More generally, we shall prove the following variant of Theorem 3.2.1 for additive functions. For simplicity's sake, we shall focus on a subclass of the multiplicative functions in $\mathcal{M}^*$.

Theorem 3.2.2. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be an additive function such that $f(p^\nu) = O(\nu^\kappa)$ for all prime powers p^ν, where $\kappa \geq 0$ is an absolute constant. Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function, and suppose that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0, 1/2)$ and $\lambda \in (0, 2^{1-2\varrho_0})$, such that $\alpha(n)$ satisfies* (3.1.2), (3.1.4), *and the condition that $\alpha(p^\nu) = O((\lambda p^{\varrho_0+\sigma_0-1})^\nu)$ for all prime powers p^ν.*

(a) If $\beta = 1$ and $0 < h_0 < (3/2)^{2/3}$ is arbitrary, and if $B(x) \to \infty$ as $x \to \infty$, then we have

$$M(x;m) = C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{M m^{\kappa + \frac{3}{2}}}{\sqrt{B(x)}} \right) \right)$$

uniformly for all sufficiently large x and all $m \in \mathbb{N}$ satisfying $m \leq h_0 (B(x)/M^2)^{1/3}$ and $m \ll B(x)^{1/(2\kappa+3)}$, where $M > 0$ is an absolute constant for which $|f(p)| \leq M$ holds for all primes p.

(b) If $\beta \neq 1$ and if $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$, then we have

$$M(x;m) = C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{M m^{\frac{3}{2}} (\log\log\log x + m^{\kappa})}{\sqrt{B(x)}} \right) \right)$$

uniformly for all sufficiently large x and all

$$1 \leq m \ll \min\left(B(x)^{1/(2\kappa+3)}, \frac{B(x)^{1/3}}{(\log\log\log x)^{2/3}} \right).$$

The implicit constants in the error terms of both asymptotic formulas above depend at most on the explicit and implicit constants in the hypotheses except for M.

Theorem 3.2.2 clearly implies the first part of [18, Theorem 1.1] if we set $f(n) = \Omega(n)$, $\kappa = 1$, and $\vartheta_0 = \varrho_0 = 0$. It is easy to see that if $\alpha(p^\nu) = O((\lambda p^{r_0 + \sigma_0 - 1})^\nu)$ for all prime powers p^ν, where $\sigma_0 > 0$, $r_0 \in [0, 1/2)$ and $\lambda \in (0, 2^{1-2r_0})$ are absolute constants, then conditions (i) and (iii) are automatically fulfilled with any fixed $\max(r_0 + \log_2 \lambda, 0) \leq \varrho_0 < 1$, $r_0 + \max(1/2, \log_2 \lambda) < r < 1$, and of course the same parameter σ_0. Indeed, we shall derive Theorem 3.2.2 as a corollary of Theorem 3.2.1.

Let $g \in \mathbb{Z}[x]$ be a nonconstant irreducible polynomial. As in Section 3.1, let $\rho_g(n)$ denote the number of zeros of g in $\mathbb{Z}/n\mathbb{Z}$. More generally, if $g \in \mathbb{Q}[x]$ is a nonconstant irreducible polynomial, we may extend the definition above by setting $\rho_g(n) = 0$ if $\gcd(n, c_g) > 1$, where $c_g \in \mathbb{N}$ is the least positive integer such that $c_g g(x) \in \mathbb{Z}[x]$, and insisting that $\rho_g(n)$ be the number of zeros of $g(x)$ (or equivalently, $c_g g(x)$) in $\mathbb{Z}/n\mathbb{Z}$ when $\gcd(n, c_g) = 1$. Extended this way with the convention that $\rho_g(1) = 1$, the function $\rho_g(n)$ is still a multiplicative function of n. It is known [33, Lemma 1] that ρ_g is bounded on prime powers and that

$$\sum_{p \le x} \frac{\rho_g(p)}{p} = \log\log x + M_{\rho_g} + O\left(\frac{1}{\log x}\right).$$

Given a strongly additive function $f\colon \mathbb{N} \to \mathbb{R}$, we define

$$A_{f,g}(x) := \sum_{p \le x} \rho_g(p) \frac{f(p)}{p},$$
$$B_{f,g}(x) := \sum_{p \le x} \rho_g(p) \frac{f(p)^2}{p}.$$

For simplicity's sake, suppose that $g(\mathbb{N}) \subseteq \mathbb{N}$. In the case $g \in \mathbb{Z}[x]$, Halberstam [32, Theorem 3] showed that if $B_{f,g}(x) \to \infty$ as $x \to \infty$, and if $f(p) = o\left(\sqrt{B_{f,g}(p)}\right)$[1], then

$$\frac{1}{x} \sum_{n \le x} (f(g(n)) - A_{f,g}(x))^m = (\mu_m + o(1)) B_{f,g}(x)^{\frac{m}{2}}$$

for every fixed $m \in \mathbb{N}$. Under the stronger condition $f(p) = O(1)$, Theorem 3.2.1 leads to a weighted version of this result in the case $g(n) = n$. As for the remaining cases we have the following theorem.

Theorem 3.2.3. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \le M$ for all primes p, where $M > 0$ is an absolute constant, and let $g \in \mathbb{Q}[x]$ be a nonconstant*

[1] Halberstam [32] wrote that for $g(x) = x$ this pair of conditions contain the condition that $f(p) = o((\log p)^\epsilon)$ for every given $\epsilon > 0$. However, this claim is incorrect, as noted by Prof. Pomerance. In fact, a simple counterexample may be constructed as follows. Let $\mathcal{P}$ be an arbitrary infinite subset of odd primes such that $\sum_{p \in \mathcal{P}} 1/p < \infty$, and put $\mathcal{P}(x) := \mathcal{P} \cap [3, x]$. Define $f(p) = \sqrt{\log\log p}$ for $p \in \mathcal{P}$ and $f(p) = 1$ for $p \notin \mathcal{P}$. From partial summation it follows that $\sum_{p \in \mathcal{P}(x)} f(p)^2/p = o(\log\log x)$. Then one sees readily that $f(p) = o((\log p)^\epsilon)$ for any given $\epsilon > 0$, while $f(p) \sim \sqrt{B(p)}$ for large $p \in \mathcal{P}$.

irreducible polynomial such that $g(0) \neq 0$ *and* $g(\mathbb{N}) \subseteq \mathbb{N}$*. Let* $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ *be a multiplicative function, and suppose that there exist absolute constants* $A_0, \beta, \sigma_0 > 0$*,* $\vartheta_0 \geq 0$*,* $\varrho_0 \in [0,1)$ *and* $r \in (0,1)$*, such that* $\alpha(n)$ *satisfies the conditions* (i)–(iv)*. Furthermore, suppose that there exists an absolute constant* $B_0 > 0$ *and a function* $\delta(x) \in (0,1]$*, such that*

$$\Delta_\alpha(x;q,a) := \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \alpha(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leq x \\ \gcd(n,q)=1}} \alpha(n) = O\left(\frac{S(x)}{\varphi(q)(\log x)^{B_0}}\right) \tag{3.2.1}$$

uniformly for all sufficiently large x*, all* $q \in \mathbb{N} \cap [1, x^{\delta(x)}]$*, and all* $a \in \mathbb{Z}$ *coprime to* q*. If* $0 < h_0 < (3/2)^{2/3}$ *is arbitrary, and if* $\delta(x)^2 B_{f,g}(x) \to \infty$ *as* $x \to \infty$*, then we have*

$$S(x)^{-1} \sum_{n \leq x} \alpha(n) \left(f(g(n)) - A_{f,g}(x)\right)^m = C_m B_{f,g}(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{M m^{\frac{3}{2}}}{\delta(x)\sqrt{B_{f,g}(x)}}\right)\right)$$

uniformly for all sufficiently large x *and all* $m \in \mathbb{N}$ *satisfying* $1 \leq m \leq h_0(B_{f,g}(x)/M^2)^{1/3}$ *and* $m \ll (\delta(x)^2 B_{f,g}(x))^{1/3}$*, where the implicit constant in the error term depends at most on the explicit and implicit constants in the hypotheses except for* M*.*

Roughly speaking, (3.2.1) can be viewed as a condition of the *Siegel–Walfisz type* which ensures that $\alpha(n)$ is well distributed among the reduced residue classes $a \,(\mathrm{mod}\, q)$ for all q in a reasonably wide range. A classical example of $\alpha(n)$ that satisfies all of the conditions in Theorem 3.2.3 is $d_k(n)$, where $k \in \mathbb{N}$. In this case it is known [5, 44] that one can actually take $\delta(x)$ to be a constant depending on k and $\Delta_\alpha(x;q,a) = O(x^{1-\epsilon}/\varphi(q))$ for some constant $\epsilon \in (0,1)$. Anther interesting example is $r_2(n)/4$ for which one may take $\delta(x) = 2/3 - \epsilon$ and any fixed $B_0 > 0$ [5].

We shall only sketch the proof of Theorem 3.2.3, since it is similar to, and in fact, much easier than that of Theorem 3.2.1. The argument used in the proof may also be modified to study the joint distribution of $f(n+h_1)$ and $f(n+h_2)$ with any fixed integers $h_1 \neq h_2$.

It is not hard to see that the condition $f(p) = O(1)$ in Theorem 3.2.1 can be relaxed, especially when we do not pursue uniformity in m in the asymptotics for the mth moment. For instance, in the case $\alpha(n) \equiv 1$ Delange and Halberstam showed [15, Theorem 1] that if $f\colon \mathbb{N} \to \mathbb{R}$ is a strongly additive function such that $B(x) \to \infty$ as $x \to \infty$, $f(p) = O(\sqrt{B(p)})$ for all primes p, and

$$\sum_{\substack{p \le x \\ |f(p)| > \epsilon\sqrt{B(x)}}} \frac{f(p)^2}{p} = o(B(x)) \tag{3.2.2}$$

for any given $\epsilon > 0$, then

$$\frac{1}{x} \sum_{n \le x} (f(n) - A(x))^m = (\mu_m + o(1)) B(x)^{\frac{m}{2}}$$

for every fixed $m \in \mathbb{N}$. This result implies at once the Kubilius–Shapiro theorem [54, Theorem A] under the additional assumption $f(p) = O(\sqrt{B(p)})$. On the other hand, Delange and Halberstam noted that their result no longer holds if this additional assumption is removed, which incidentally exposes the limitation of the method of moments compared to the method evolved by Erdős and Kac. Nevertheless, it will be clear in the sequel that the proof of Theorem 3.2.1 makes it possible for us to obtain the following natural extension of the result of Delange and Halberstam.

Theorem 3.2.4. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function. Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\ge 0}$ be a multiplicative function, and suppose that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \ge 0$, $\varrho_0 \in [0, 1)$ and $r \in (0, 1)$, such that $\alpha(n)$ satisfies the conditions* (i)–(iv)*. Define*

$$B^*(x) := \begin{cases} B(x), & \text{if } \beta = 1, \\ B(x)/(\log\log\log x)^2, & \text{if } \beta \neq 1, \end{cases}$$

and suppose $B^*(x) \to \infty$ *as* $x \to \infty$. *If there exists an absolute constant* $K > 0$ *such that* $f(n) = o(\sqrt{B(x)})$ *for all squarefree* $n \in [1, x]$ *composed of prime factors* p *satisfying* $|f(p)| > K\sqrt{B^*(x)}$, *and if*

$$\sum_{\substack{p \le x \\ |f(p)| > \epsilon\sqrt{B^*(x)}}} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}} = o(B^*(x))$$

for any given $\epsilon > 0$, *then* $M(x; m) = (\mu_m + o(1))B(x)^{\frac{m}{2}}$ *for every fixed* $m \in \mathbb{N}$.

Note that the theorem of Delange and Halberstam [15, Theorem 1] corresponds to the case $\alpha \equiv 1$. The proof of Theorem 3.2.4, which we shall only outline, is based on the proofs of Theorem 3.2.1 and [15, Theorem 1]. We shall also obtain as a corollary the following analogue of the Kubilius–Shapiro theorem [54, Theorems A, C].

Corollary 3.2.5. *Under the notation and hypotheses in Theorem 3.2.4, we have*

$$\lim_{x \to \infty} S(x)^{-1} \sum_{\substack{n \le x \\ f(n) \le A(x) + V\sqrt{B(x)}}} \alpha(n) = \Phi(V)$$

for any given $V \in \mathbb{R}$. *The same is true if* f *is merely additive.*

It is clear that Theorem 3.2.4 implies Corollary 3.2.5 when f is strongly additive. To handle the more general case where f is merely additive, we shall prove a weighted version of [54, Theorem B] which shows that when it comes to the distribution problem, there is no essential difference between strongly additive functions and general additive functions, and thus the distribution of an additive function f is determined solely by its values at primes.

Before embarking on the proofs, we describe briefly the main steps in the proof of the uniform estimates for moments. The starting point is the approximation to moments used by Granville, Soundararajan, Khan, Milinovich and Subedi. Though the underlying idea is the same, we need a more complicated version of this approximation (see Lemma 3.4.1) due to the more general nature of our multiplicative weight functions $\alpha(n)$. To utilize it, we first need to develop an asymptotic formula for the mean value of $\alpha(n)$ with $n \leq x$ restricted to any squarefree integer $a \in \mathbb{N} \cap [1, x]$ (see Lemma 3.3.3). An important feature of this formula is that it holds uniformly for all squarefree integer $a \in \mathbb{N} \cap [1, x]$, which is key to both applying the moment approximation and making the moment estimates uniform. This formula will serve as the substitute for the one concerning $d_k(n)$ used by Khan, Milinovich and Subedi. Unlike the proof given by Khan, Milinovich and Subedi, which is based on Peron's formula and makes use of the special property $d_k(mn) \leq d_k(m)d_k(n)$ for all $m, n \in \mathbb{N}$, our proof uses the mean value estimate for $\alpha(n)$ given by [11, Theorem 2.1] and is completely elementary. This is done in the next section.

After applying the moment approximation, we find that the estimation of the main contribution can be worked out as in [30] and [38]. It is the estimation of the error terms that is more involved in our case. In particular, the estimation of the error term in the moment approximation supplied by Lemma 3.4.1 in Section 3.4 requires separate treatments according as $\beta = 1$ or $\beta \neq 1$. Besides, since the error term in our asymptotic formula for the mean value of $\alpha(n)$ over $a \mid n$ provided by Lemma 3.3.3 in Section 3.3 is weaker than what one can obtain for the special weight $d_k(n)$ by complex analytic approaches, we need to handle the case $\beta \in (0, 1)$ with some special care and tailor the selection of parameters accordingly in order to minimize the error terms. With these new technical complications being taken care of, we obtain the desired uniform estimates for moments stated in Theorems 3.2.1 and 3.2.2.

Remark 3.2.2. The condition that $f(p) = o((\log p)^\epsilon)$ for any given $\epsilon > 0$, mentioned by Halberstam [32], does not imply (3.2.2) in general. To see this, assume for the moment that there exists an infinite subset $\mathcal{P}$ of primes such that

$$s_{\mathcal{P}}(x) := \sum_{p \in \mathcal{P} \cap [17,x]} \frac{1}{p} = \frac{\log\log x}{\log\log\log x} + c + o(1) \tag{3.2.3}$$

for sufficiently large x, where $c \in \mathbb{R}$ is an absolute constant. Define

$$f(p) = (\log p)^{1/(2 \log\log\log p)}$$

for $p \in \mathcal{P}$ and $f(p) = 1$ for $p \notin \mathcal{P}$. Clearly, $f(p) = o((\log p)^{\epsilon})$ for any given $\epsilon > 0$. It is easily seen by partial summation that

$$\sum_{p \in \mathcal{P} \cap [17,x]} \frac{f(p)^2}{p} = \int_{17^-}^{x} (\log t)^{1/\log\log\log t}\, ds_{\mathcal{P}}(t) = (1 + o(1))(\log x)^{1/\log\log\log x},$$

which implies that

$$B(x) = \sum_{p \in \mathcal{P} \cap [17,x]} \frac{f(p)^2}{p} + O(\log\log x) = (1 + o(1))(\log x)^{1/\log\log\log x}.$$

Take $y = x^{1/\log\log x}$ and $\epsilon = 1/2$. Since

$$\log\log y = \log\log x - \log\log\log x,$$
$$\log\log\log y = \left(1 + O\left(\frac{1}{\log\log x}\right)\right) \log\log\log x,$$

we have

$$(\log y)^{1/\log\log\log y} = \exp\left(\frac{\log\log x}{\log\log\log x} - 1 + O\left(\frac{1}{\log\log\log x}\right)\right).$$

It follows that

$$f(p)^2 > (\log y)^{1/\log\log\log y} > \frac{1}{3}(\log x)^{1/\log\log\log x} > \epsilon^2 B(x)$$

for $p \in \mathcal{P} \cap (y, x]$ when x is sufficiently large. Hence, we have

$$\sum_{\substack{p \le x \\ |f(p)| > \epsilon\sqrt{B(x)}}} \frac{f(p)^2}{p} \ge \sum_{p \in \mathcal{P} \cap (y,x]} \frac{f(p)^2}{p} > \frac{1}{2}(\log x)^{1/\log\log\log x} > \frac{1}{3}B(x).$$

It remains to construct a set $\mathcal{P}$ with the desired property (3.2.3). Note first that $\sum_{p \le x} 1/p = \log\log x + O(1)$ grows slightly faster than our target

$$u(x) := \frac{\log\log x}{\log\log\log x},$$

according to Mertens' second theorem [35, Theorem 427]. Moreover, if $p < p'$ are large consecutive primes, then $u(p') - u(p) = o(1/\log p)$, by Bertrand's postulate. Let 17 be the first prime in $\mathcal{P}$. Suppose that we have already selected for $\mathcal{P}$ the primes up to q, where q is prime. We put the next prime q' in $\mathcal{P}$ if $s_{\mathcal{P}}(q) < u(q)$ and leave it out of $\mathcal{P}$ otherwise. Then the running sum $s_{\mathcal{P}}(x)$ changes by at most $1/q$ as x moves from q to q', while the target $u(x)$ changes by at most $o(1/\log q)$ as x moves from q to q'. Thus, the difference $s_{\mathcal{P}}(x) - u(x)$ can be kept within $o(1/\log x)$. In particular, (3.2.3) holds for $\mathcal{P}$ with $c = 0$.

Section 3.3

Mean Values of Multiplicative Functions

Without loss of generality, we may assume $A_0 \in (0,1)$ in the sequel. In addition, we shall also make use of the asymptotic formula

$$\sum_{p \le x} \frac{\alpha(p)}{p^{\sigma_0}} = \beta \log\log x + M_\alpha + O\left((\log x)^{-A_0}\right) \tag{3.3.1}$$

with some constant $M_\alpha \in \mathbb{R}$, which follows immediately from (3.1.2) via partial summation. In view of our assumption that $f(p) = O(1)$, this formula implies trivially that $B(x) \ll \log\log x$. Moreover, if we define, for every prime p,

$$\psi_0(p) := \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}},$$

then we infer from (3.1.1), (3.1.3) and (3.1.4) that

$$\psi_0(p) \ll \frac{(\log\log(p+1))^{\vartheta_0}}{p}$$

and that $\sum_p \psi_0(p) < \infty$.

Lemma 3.3.1. *Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\ge 0}$ be a multiplicative function satisfying* (3.1.1) *and* (3.1.4) *with some $\sigma_0, \vartheta_0 > 0$ and $\varrho_0 \in [0,1)$. Fix $h \in \mathbb{R}$, $\epsilon_0 \in (0,1)$ and $c_0 \in [1, \epsilon_0^{-1})$, and define*

$$I_{\alpha,h}(x;a) := \sum_{\substack{q \le x \\ R_q = a}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q}\right)^h,$$

where $a \in \mathbb{N} \cap [1,x]$ is squarefree. Then there exists a constant $\delta_0 > 0$ such that uniformly for all sufficiently large x, any $\delta \in [\delta_0 \log\log x/\log x, 1]$, and any squarefree

$a \in \mathbb{N} \cap [1,x]$ *with* $\omega(a) \le (1-\varrho_0)\epsilon_0\delta^{-1}$, *we have*

$$I_{\alpha,h}(x;a) = \left(\tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(\omega(a))}}{\log x}\left(\frac{1}{x^{c_0\delta\omega(a)}} + \frac{\epsilon_{h,0}L(a)\log P^+(a)}{a}\right)\right)\right)(\log x)^h,$$

where

$$\tilde{\lambda}_\alpha(a) := \prod_{p|a}\sum_{\nu=1}^{\infty}\frac{\alpha(p^\nu)}{p^{\sigma_0\nu}},$$
$$L(a) := \prod_{p|a}(\log\log(p+1))^{\vartheta_0}.$$

Proof. Let $\delta \in (0,1]$ and fix $c_1 \in (c_0, \epsilon_0^{-1})$. Put $\delta_1 := (1-\varrho_0)^{-1}c_1\delta$ and $y := x^{k\delta_1}$. For any squarefree $a = p_1\cdots p_k \in \mathbb{N}\cap[1,x]$ with $p_1 < ... < p_k \le x$ and $k \le (1-\varrho_0)\epsilon_0\delta^{-1}$, we have $k\delta_1 \le c_1\epsilon_0 < 1$ and

$$I_{\alpha,h}(x;a) = \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le x\\ \nu_1,\dots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^h.$$

On the one hand, we see that

$$\begin{aligned}
&\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\dots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^h\\
&= \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\dots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}(\log 3x)^h\left(1 + O\left(\frac{\epsilon_{h,0}}{\log 3x}\sum_{i=1}^{k}\nu_i\log p_i\right)\right)\\
&= \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\dots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}(\log x)^h + O\left(\frac{2^{O(k)}\epsilon_{h,0}L(a)\log p_k}{a}(\log x)^{h-1}\right),
\end{aligned}$$

by (3.1.4). From (3.1.1) it follows that

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\dots,\nu_k\ge 1}} \frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}} = \tilde{\lambda}_\alpha(a) + O\left(2^{O(k)}\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}> y\\ \nu_1,\dots,\nu_k\ge 1}} \frac{1}{p_1^{(1-\varrho_0)\nu_1}\cdots p_k^{(1-\varrho_0)\nu_k}}\right).$$

The sum in the error term above may be split into two sums according as $p_2^{\nu_2}\cdots p_k^{\nu_k}\le y$ or $p_2^{\nu_2}\cdots p_k^{\nu_k}>y$. In the first sum we must have $p_1^{\nu_1}>y/(p_2^{\nu_2}\cdots p_k^{\nu_k})$. Thus summing over ν_1 and then over $\nu_2,...,\nu_k$, we see that the first sum is

$$\ll \frac{1}{y^{1-\varrho_0}}\sum_{\substack{p_2^{\nu_2}\cdots p_k^{\nu_k}\le y\\ \nu_2,\dots,\nu_k\ge 1}} 1 \le \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k\delta}(\log p_2)\cdots(\log p_k)}.$$

The second sum is simply

$$\sum_{\substack{p_2^{\nu_2}\cdots p_k^{\nu_k}> y\\ \nu_2,\dots,\nu_k\ge 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2}\cdots p_k^{(1-\varrho_0)\nu_k}}\sum_{\nu_1\ge 1}\frac{1}{p_1^{(1-\varrho_0)\nu_1}} \ll \sum_{\substack{p_2^{\nu_2}\cdots p_k^{\nu_k}> y\\ \nu_2,\dots,\nu_k\ge 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2}\cdots p_k^{(1-\varrho_0)\nu_k}}.$$

It follows that

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}> y\\ \nu_1,\dots,\nu_k\ge 1}} \frac{1}{p_1^{(1-\varrho_0)\nu_1}\cdots p_k^{(1-\varrho_0)\nu_k}} \ll \sum_{\substack{p_2^{\nu_2}\cdots p_k^{\nu_k}> y\\ \nu_2,\dots,\nu_k\ge 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2}\cdots p_k^{(1-\varrho_0)\nu_k}} + \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k\delta}(\log p_2)\cdots(\log p_k)}.$$

Repeating this argument, we obtain

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}> y\\ \nu_1,\dots,\nu_k\ge 1}} \frac{1}{p_1^{(1-\varrho_0)\nu_1}\cdots p_k^{(1-\varrho_0)\nu_k}} \le \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k\delta}(\log p_2)\cdots(\log p_k)},$$

from which we deduce

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\ldots,\nu_k\ge 1}} \frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}} = \tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1k\delta}(\log p_2)\cdots(\log p_k)}\right). \tag{3.3.2}$$

On the other hand, we have

$$\begin{aligned}\sum_{x_1<p^\nu\le x_2} \frac{\alpha(p^\nu)}{p^{\sigma_0\nu}}\left(\log\frac{3x_2}{p^\nu}\right)^h &\ll \sum_{\substack{\log_p x_1<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}}\left(\log\frac{3x_2}{p^\nu}\right)^h\\ &= -\int_{\log_p x_1}^{\log_p x_2}\left(\log\frac{3x_2}{p^t}\right)^h d\left(\sum_{\substack{t<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}}\frac{1}{p^{(1-\varrho_0)t}}\right)\end{aligned}$$

uniformly for all primes p and all $0<x_1\le x_2$. Using integration by parts, we see that the integral above is equal to

$$-(\log(3x_2/x_1))^h \sum_{\substack{\log_p x_1<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}} - \int_{\log_p x_1}^{\log_p x_2}\left(\sum_{\substack{t<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}}\frac{1}{p^{(1-\varrho_0)t}}\right) d\left(\log\frac{3x_2}{p^t}\right)^h.$$

Since

$$\sum_{\substack{t<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}}\frac{1}{p^{(1-\varrho_0)t}} < \frac{1}{p^{(1-\varrho_0)(\lfloor t\rfloor+1)}}\cdot\frac{p^{1-\varrho_0}}{p^{1-\varrho_0}-1} \ll \frac{1}{p^{(1-\varrho_0)t}},$$

we have

$$(\log(3x_2/x_1))^h \sum_{\substack{\log_p x_1<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}} \ll \frac{(\log(3x_2/x_1))^h}{x_1^{1-\varrho_0}}$$

and

$$\int_{\log_p x_1}^{\log_p x_2}\left(\sum_{\substack{t<\nu\le\log_p x_2\\ \nu\in\mathbb{Z}}}\frac{1}{p^{(1-\varrho_0)t}}\right)d\left(\log\frac{3x_2}{p^t}\right)^h \ll \epsilon_{h,0}\log p\int_{\log_p x_1}^{\log_p x_2}\frac{1}{p^{(1-\varrho_0)t}}\left(\log\frac{3x_2}{p^t}\right)^{h-1}dt$$
$$=\frac{\epsilon_{h,0}}{(3x_2)^{1-\varrho_0}}\int_{\log 3}^{\log(3x_2/x_1)}t^{h-1}e^{(1-\varrho_0)t}\,dt$$
$$\ll\frac{\epsilon_{h,0}}{(3x_2)^{1-\varrho_0}}(\log(3x_2/x_1))^{h-1}\left(\frac{3x_2}{x_1}\right)^{1-\varrho_0}$$
$$=\frac{\epsilon_{h,0}(\log(3x_2/x_1))^{h-1}}{x_1^{1-\varrho_0}}.$$

Hence, it follows that

$$\sum_{x_1<p^\nu\le x_2}\frac{\alpha(p^\nu)}{p^{\sigma_0\nu}}\left(\log\frac{3x_2}{p^\nu}\right)^h\ll\frac{(\log(3x_2/x_1))^h}{x_1^{1-\varrho_0}}\tag{3.3.3}$$

uniformly for all primes p and all $0<x_1\le x_2$. This inequality implies immediately

$$\sum_{\substack{y<p_1^{\nu_1}\cdots p_k^{\nu_k}\le x\\ \nu_1,\dots,\nu_k\ge1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^h\le\frac{2^{O(k)}(\log x)^h}{y^{1-\varrho_0}}\sum_{\substack{p_2^{\nu_2}\cdots p_k^{\nu_k}\le x\\ \nu_2,\dots,\nu_k\ge1}}1$$
$$\le\frac{2^{O(k)}(\log x)^{k+h-1}}{x^{c_1k\delta}(\log p_2)\cdots(\log p_k)}.$$

Lemma 3.3.1 now follows upon combining the above with (3.3.2) and taking $\delta_0=1/(c_1-c_0)$ with the range $\delta\ge\delta_0\log\log x/\log x$ in mind. □

Let $\alpha\colon\mathbb{N}\to\mathbb{R}_{\ge0}$ be a multiplicative function as in Theorem 3.2.1 with $A_0\in(0,1)$. Suppose first that (3.1.3) holds with the restricted sum $\sum_p'$ replaced by the full sum $\sum_p$. For $\sigma_0=1$ De la Bretèche and Tenenbaum [11, Theorem 2.1] showed

$$\sum_{n\le x}\alpha(n)=\frac{1}{\Gamma(\beta)}\prod_p\left(1-\frac1p\right)^\beta\sum_{\nu=0}^\infty\frac{\alpha(p^\nu)}{p^\nu}x(\log x)^{\beta-1}\left(1+O\left(\frac{1}{(\log x)^{A_0}}\right)\right),$$

where the implicit constant in the error term depends at most on the explicit and implicit constants in the hypotheses. For the general case where $\sigma_0 > 0$ is arbitrary, it is easy to show, by applying the above to $\alpha(n)/n^{\sigma_0 - 1}$ and employing partial summation as in the proof of [18, Corollary 3.3], that

$$S(x) = \sum_{n \le x} \alpha(n) = \lambda_\alpha x^{\sigma_0} (\log x)^{\beta - 1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}}\right)\right), \tag{3.3.4}$$

where

$$\lambda_\alpha := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_p \left(1 - \frac{1}{p}\right)^\beta \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}}. \tag{3.3.5}$$

Suppose now that (3.1.3) holds with the restricted sum $\sum_p'$ being the sum $\sum_{p > Q_0}$, where $Q_0 \ge 1$ is some absolute constant. Let $P_0 := \prod_{p \le Q_0} p$ and $\mathbf{1}_{P_0}(n)$ the indicator function of the set $\{n \in \mathbb{N} \colon \gcd(n, P_0) = 1\}$. Then $\alpha(n)\mathbf{1}_{P_0}(n)$ is a nonnegative multiplicative function satisfying (3.1.1)–(3.1.4) with the sum $\sum_p'$ in (3.1.3) replaced by the full sum $\sum_p$. In particular, (3.3.4) is applicable to $\alpha(n)\mathbf{1}_{P_0}(n)$. Thus, we obtain

$$\sum_{\substack{n \le x \\ \gcd(n, P_0) = 1}} \alpha(n) = \lambda_\alpha(P_0) x^{\sigma_0} (\log 3x)^{\beta - 1} \left(1 + O\left(\frac{1}{(\log 3x)^{A_0}}\right)\right), \tag{3.3.6}$$

where

$$\lambda_\alpha(P_0) := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_{p \le Q_0} \left(1 - \frac{1}{p}\right)^\beta \cdot \prod_{p > Q_0} \left(1 - \frac{1}{p}\right)^\beta \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}}.$$

Examining the proof of Lemma 3.3.1, we find that for every given $h \in \mathbb{R}$,

$$\sum_{\substack{q \le x \\ R_q \mid P_0}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q}\right)^h = \prod_{p \le Q_0} \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} (\log x)^h \left(1 + O\left(\frac{1}{\log x}\right)\right)$$

for all sufficiently large x. Combining this with (3.3.6) gives

$$S(x) = \sum_{\substack{q \le x \\ R_q | P_0}} \alpha(q) \sum_{\substack{n' \le x/q \\ \gcd(n', P_0) = 1}} \alpha(n') = \lambda_\alpha x^{\sigma_0} (\log x)^{\beta - 1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}} \right) \right),$$

which is the same as (3.3.4).

For our applications, we will need an asymptotic formula for

$$S(x; a) = S_\alpha(x; a) := \sum_{\substack{n \le x \\ \gcd(n, a) = 1}} \alpha(n)$$

uniform in $a \in \mathbb{N} \cap [1, x]$. One may be tempted to apply (3.3.4) to the function $\alpha(n)\mathbf{1}_a(n)$, where $\mathbf{1}_a(n)$ is the indicator function of the set $\{n \in \mathbb{N}\colon \gcd(n, a) = 1\}$. However, it is not immediately clear whether the implied constant in the error term obtained via this naive approach is independent of $a \in \mathbb{N} \cap [1, x]$. Fortunately, the following lemma provides the desired estimate for $S(x; a)$ under the hypotheses (i)–(iv).

Lemma 3.3.2. *Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\ge 0}$ be a multiplicative function satisfying* (3.1.1)–(3.1.4) *with some $A_0 \in (0, 1)$, $\beta, \sigma_0 > 0$, $\vartheta_0 \ge 0$, $\varrho_0 \in [0, 1)$ and $r \in (0, 1)$. Then we have*

$$S(x; a) = \lambda_\alpha(a) x^{\sigma_0} (\log x)^{\beta - 1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}} \right) \right)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$, where

$$\lambda_\alpha(a) := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_{p | a} \left(1 - \frac{1}{p} \right)^\beta \cdot \prod_{p \nmid a} \left(1 - \frac{1}{p} \right)^\beta \sum_{\nu = 0}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}},$$

The implicit constant in the error term depends at most on the explicit and implicit constants in the hypotheses.

Proof. Let $a \in \mathbb{N} \cap [1, x]$. For simplicity of notation, we write $\sum^a$ for sums in which the indices take values coprime to a. As we have demonstrated above, there is no loss of generality by assuming that $\sigma_0 = 1$ and that (3.1.3) holds with the restricted sum $\sum_p'$ replaced by the full sum $\sum_p$. We start by determining the relation between $\lambda_\alpha(a)$ and λ_α. Note that condition (iv) implies that $\alpha(p) \ll (\log\log(p+1))^{\vartheta_0}$, from which it follows that

$$\sum_{p|a} \alpha(p) \log p \ll (\log\log x)^{\vartheta_0} \log a \leq (\log\log x)^{\vartheta_0} \log x.$$

By (3.1.2) we have

$$\sum_{p\leq x}{}^{a} \alpha(p) \log p = \beta x + O\left(\frac{x}{(\log x)^{A_0}}\right),$$

from which we deduce

$$\sum_{p\leq x}{}^{a} \frac{\alpha(p)}{p} = \beta \log\log x + M_\alpha + O\left((\log x)^{-A_0}\right).$$

Combining the above with (3.3.1), we see that

$$\begin{aligned}
\lambda_\alpha(a) &= \lambda_\alpha \prod_{p|a} \left(\sum_{\nu=0}^{\infty} \frac{\alpha(p^\nu)}{p^\nu}\right)^{-1} \\
&= \lambda_\alpha \exp\left(-\sum_{p|a} \frac{\alpha(p)}{p} + O(1)\right) \\
&= \lambda_\alpha \exp\left(-\sum_{p\leq x} \frac{\alpha(p)}{p} + \sum_{p\leq x}{}^{a} \frac{\alpha(p)}{p} + O(1)\right) \asymp \lambda_\alpha \asymp 1.
\end{aligned}$$

This assures us that there is no need to differentiate $\lambda_\alpha(a)$ and λ_α in the error terms.

Next, we connect $S(x;a)$ with

$$T(x;a) = T_\alpha(x;a) := \sideset{}{^a}\sum_{n\le x} \frac{\alpha(n)}{n}.$$

It is clear from (3.3.4) that $S(x;a) \le S(x;1) \ll x(\log x)^{\beta-1}$ and $T(x;a) \le T(x,1) \ll (\log x)^\beta$. Moreover, it is shown in the the proof of [11, Theorem 2.1] that

$$T(x;1) = \left(1 + O\left(\frac{1}{\log x}\right)\right) \frac{\lambda_\alpha}{\beta} (\log x)^\beta. \tag{3.3.7}$$

Following the proof of [11, Theorem 2.1], we find

$$\begin{aligned}
S(x;a)\log x &= \sideset{}{^a}\sum_{n\le x} \alpha(n)\log n + \sideset{}{^a}\sum_{n\le x} \alpha(n) \log\frac{x}{n} \\
&= \sideset{}{^a}\sum_{k\le x} \alpha(k) \sideset{}{^a}\sum_{\substack{p^\nu \le x/k \\ p\nmid k}} \alpha(p^\nu)\log p^\nu + \int_{1^-}^{x} \frac{S(t;a)}{t}\, dt \\
&= \sideset{}{^a}\sum_{k\le x} \alpha(k) \sideset{}{^a}\sum_{p\le x/k} \alpha(p)\log p + O\left(\sum_{k\le x}\alpha(k) \sum_{\substack{p\le x/k \\ p\mid k}} \alpha(p)\log p\right) \\
&\quad + O\left(\sum_{k\le x}\alpha(k) \sum_{\substack{p^\nu\le x/k \\ \nu\ge 2}} \alpha(p^\nu)\log p^\nu\right) + O\left(x(\log x)^{\beta-1}\right) \\
&= \beta x T(x,a) - \sideset{}{^a}\sum_{k\le x}\alpha(k) \sum_{\substack{p\le x/k \\ p\mid a}} \alpha(p)\log p + O\left(x \sum_{k\le x} \frac{\alpha(k)}{k(\log(3x/k))^{A_0}}\right) \\
&\quad + O\left(x(\log x)^{\beta-1}\right),
\end{aligned} \tag{3.3.8}$$

since (3.1.3) and (3.3.4) imply that

$$
\begin{aligned}
&\sum_{k\le x}\alpha(k)\sum_{\substack{p\le x/k\\ p|k}}\alpha(p)\log p+\sum_{k\le x}\alpha(k)\sum_{\substack{p^\nu\le x/k\\ \nu\ge 2}}\alpha(p^\nu)\log p^\nu\\
&=\sum_{k\le x}\alpha(k)\sum_{\substack{p^\nu\le x/k\\ \nu\ge 2}}\big(\alpha(p)\alpha(p^{\nu-1})\log p^{\nu-1}+\alpha(p^\nu)\log p^\nu\big)\\
&\ll x^{(r+1)/2}\sum_{k\le x}\frac{\alpha(k)}{k^{(r+1)/2}}\sum_p\sum_{\nu\ge 2}\left(\frac{\alpha(p)}{p^r}\cdot\frac{\alpha(p^{\nu-1})}{p^{r(\nu-1)}}+\frac{\alpha(p^\nu)}{p^{r\nu}}\right)\\
&\ll x^{(r+1)/2}\sum_{k\le x}\frac{\alpha(k)}{k^{(r+1)/2}}\sum_p\left(\frac{\alpha(p)^2}{p^{2r}}+\sum_{\nu\ge 2}\frac{\alpha(p^\nu)}{p^{r\nu}}\right)\\
&\ll x^{(r+1)/2}\sum_{k\le x}\frac{\alpha(k)}{k^{(r+1)/2}}\\
&\ll x(\log x)^{\beta-1}.
\end{aligned}
$$

By partial summation we have

$$
\begin{aligned}
\sum_{k\le x}\frac{\alpha(k)}{k(\log(3x/k))^{A_0}}&=\frac{S(x)}{x(\log 3)^{A_0}}+\int_{1^-}^{x}\frac{\log(3x/t)-A_0}{t^2(\log(3x/t))^{A_0+1}}S(t)\,dt\\
&\ll(\log x)^{\beta-1}+\int_1^x\frac{(\log 3t)^{\beta-1}}{t(\log(3x/t))^{A_0}}\,dt\\
&=(\log x)^{\beta-1}+\int_0^{\log x}\frac{(\log 3+t)^{\beta-1}}{(\log 3x-t)^{A_0}}\,dt\\
&\ll(\log x)^{\beta-1}+\frac{1}{(\log x)^{A_0}}\int_0^{(\log x)/2}(\log 3+t)^{\beta-1}\,dt\\
&\quad+(\log x)^{\beta-1}\int_{(\log x)/2}^{\log x}\frac{1}{(\log 3x-t)^{A_0}}\,dt\\
&\ll(\log x)^{\beta-A_0}. \qquad (3.3.9)
\end{aligned}
$$

Let $x_1 := x/(\log x)^2$. For $k \le x_1$ we see that

$$\sum_{\substack{p \le x/k \\ p \mid a}} \alpha(p) \log p \ll (\log\log x)^{\vartheta_0} \log a \ll \frac{x}{k(\log(x/k))^{A_0}},$$

so that

$$\sum_{k \le x_1}{}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p \mid a}} \alpha(p) \log p \ll x \sum_{k \le x} \frac{\alpha(k)}{k(\log(3x/k))^{A_0}} \ll x(\log x)^{\beta - A_0}. \qquad (3.3.10)$$

On the other hand, we have by (3.1.2) that

$$\sum_{x_1 < k \le x}{}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p \mid a}} \alpha(p) \log p \ll x \sum_{x_1 < k \le x} \frac{\alpha(k)}{k} \ll x\left((\log x)^{\beta} - (\log x_1)^{\beta} + O((\log x)^{\beta-1})\right)$$

$$\ll x(\log x)^{\beta-1} \log\log x, \qquad (3.3.11)$$

where we have used (3.3.7) to estimate the sum over k and the mean value theorem to get

$$(\log x)^{\beta} - (\log x_1)^{\beta} = \beta \xi^{\beta-1} \log \frac{x}{x_1} \ll (\log x)^{\beta-1} \log\log x$$

for some $\xi \in (\log x_1, \log x)$. Combining (3.3.10) with (3.3.11), we obtain

$$\sum_{k \le x}{}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p \mid a}} \alpha(p) \log p \ll x(\log x)^{\beta - A_0}.$$

Inserting this and (3.3.9) into (3.3.8) yields

$$S(x;a) = \frac{\beta x}{\log x} T(x;a) + O\left(x(\log x)^{\beta-1-A_0}\right) \qquad (3.3.12)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$.

To estimate $T(x;a)$, we repeat the argument above with $\alpha(n)$ replaced by $\alpha(n)/n$. From (3.1.2) it follows that

$$\sum_{p\le x}\frac{\alpha(p)}{p}\log p=\beta\log x+O\left((\log x)^{1-A_0}\right). \tag{3.3.13}$$

Setting

$$U(x;a):=\sideset{}{^a}\sum_{n\le x}\frac{\alpha(n)}{n}\log\frac{x}{n}=\int_{1^-}^{x}\frac{T(t;a)}{t}\,dt, \tag{3.3.14}$$

we have

$$\begin{aligned}
T(x;a)\log x &= \sideset{}{^a}\sum_{n\le x}\frac{\alpha(n)}{n}\log n+\sideset{}{^a}\sum_{n\le x}\frac{\alpha(n)}{n}\log\frac{x}{n}\\
&= \sideset{}{^a}\sum_{k\le x}\frac{\alpha(k)}{k}\sideset{}{^a}\sum_{\substack{p^\nu\le x/k\\ p\nmid k}}\frac{\alpha(p^\nu)}{p^\nu}\log p^\nu+U(x;a)\\
&= \sideset{}{^a}\sum_{k\le x}\frac{\alpha(k)}{k}\sideset{}{^a}\sum_{p\le x/k}\frac{\alpha(p)}{p}\log p+O\left(\sideset{}{^a}\sum_{k\le x}\frac{\alpha(k)}{k}\sum_{\substack{p\le x/k\\ p\mid k}}\frac{\alpha(p)}{p}\log p\right)\\
&\quad+O\left(\sideset{}{^a}\sum_{k\le x}\frac{\alpha(k)}{k}\sum_{\substack{p^\nu\le x/k\\ \nu\ge 2}}\frac{\alpha(p^\nu)}{p^\nu}\log p^\nu\right)+U(x;a)\\
&= (\beta+1)U(x;a)-\sideset{}{^a}\sum_{k\le x}\frac{\alpha(k)}{k}\sum_{\substack{p\le x/k\\ p\mid a}}\frac{\alpha(p)}{p}\log p+O\left((\log x)^{1-A_0}T(x;a)\right),
\end{aligned}$$

since

$$\begin{aligned}
&\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}\sum_{\substack{p\le x/k\\ p\mid k}}\frac{\alpha(p)}{p}\log p+\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}\sum_{\substack{p^\nu\le x/k\\ \nu\ge 2}}\frac{\alpha(p^\nu)}{p^\nu}\log p^\nu\\
&=\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}\sum_{\substack{p^\nu\le x/k\\ \nu\ge 2}}\frac{\alpha(p)\alpha(p^{\nu-1})\log p^{\nu-1}+\alpha(p^\nu)\log p^\nu}{p^\nu}\\
&\ll\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}\sum_{p}\left(\frac{\alpha(p)^2}{p^2}\log p+\sum_{\nu\ge 2}\frac{\alpha(p^\nu)}{p^\nu}\log p^\nu\right)\\
&\ll\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}=T(x;a).
\end{aligned}$$

In view of (3.3.13), we have

$$\begin{aligned}
\sum_{\substack{p\le x/k\\ p\mid a}}\frac{\alpha(p)}{p}\log p&\le\sum_{p\le(\log x)^2}\frac{\alpha(p)}{p}\log p+\sum_{\substack{(\log x)^2<p\le x\\ p\mid a}}\frac{\alpha(p)}{p}\log p\\
&\ll\log\log x+(\log\log x)^{\vartheta_0}\sum_{\substack{(\log x)^2<p\le x\\ p\mid a}}\frac{\log p}{p}\\
&\ll\log\log x+(\log\log x)^{\vartheta_0}\omega(a)\frac{\log\log x}{(\log x)^2}\ll\log\log x,
\end{aligned}$$

so that

$$\sum_{k\le x}{}^{a}\frac{\alpha(k)}{k}\sum_{\substack{p\le x/k\\ p\mid a}}\frac{\alpha(p)}{p}\log p\ll(\log\log x)T(x;a).$$

It follows that

$$T(x;a)\log x=(\beta+1)U(x;a)+O\left((\log x)^{1-A_0}T(x;a)\right).$$

Hence, there exists a function $\epsilon(x;a)$ such that $\epsilon(x;a)=O((\log x)^{-A_0})$ and

$$T(x;a)=\frac{1}{1-\epsilon(x;a)}\cdot\frac{\beta+1}{\log x}U(x;a)\tag{3.3.15}$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$.

Finally, we estimate $U(x; a)$ and $T(x; a)$ by following the proof of [55, Theorem A]. For $y \geq 2$ and $a \in \mathbb{N} \cap [1, y]$, let

$$V(y; a) := \log\left(\frac{\beta + 1}{(\log y)^{\beta+1}} U(y; a)\right).$$

In light of (3.3.14) and (3.3.15), we have

$$\begin{aligned}
\frac{d}{dy} V(y; a) &= -\frac{\beta + 1}{y \log y} + \frac{1}{U(y; a)} \cdot \frac{d}{dy} U(y; a) \\
&= -\frac{\beta + 1}{y \log y} + \frac{T(y; a)}{U(y; a) y} \\
&= \frac{\beta + 1}{y \log y} \cdot \frac{\epsilon(y; a)}{1 - \epsilon(y; a)} \ll \frac{1}{y (\log y)^{A_0 + 1}}
\end{aligned}$$

uniformly for all sufficiently large y and all $a \in \mathbb{N} \cap [1, y]$, which implies that

$$V_a := \int_2^\infty \frac{d}{dy} V(y; a)\, dy < \infty.$$

Since

$$V(x; a) - V(2; a) = V_a - \int_x^\infty \frac{d}{dy} V(y; a)\, dy = V_a + O\left((\log x)^{-A_0}\right)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$, it follows that

$$\frac{\beta + 1}{(\log x)^{\beta+1}} U(x; a) = \exp(V(x; a)) = \exp(V_a + V(2; a)) \left(1 + O\left((\log x)^{-A_0}\right)\right).$$

Combining this estimate with (3.3.15), we infer

$$T(x; a) = \exp(V_a + V(2; a))(\log x)^\beta \left(1 + O\left((\log x)^{-A_0}\right)\right) \tag{3.3.16}$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$. The leading coefficient can be made explicit by arguing as in the proof of [55, Theorem A]. Alternatively, we can also take advantage of (3.3.6). Fixing $a \in \mathbb{N}$, we have by (3.3.6) with $\sigma_0 = 1$ that

$$T(x;a) = \frac{\lambda_\alpha(a)}{\beta}(\log x)^\beta \left(1 + O\left((\log x)^{-A_0}\right)\right)$$

for all sufficiently large x. Comparing this with (3.3.16) shows that $\exp(V_a + V(2;a)) = \lambda_\alpha(a)/\beta$. Carrying this back into (3.3.16), we obtain

$$T(x;a) = \frac{\lambda_\alpha(a)}{\beta}(\log x)^\beta \left(1 + O\left((\log x)^{-A_0}\right)\right)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$. Inserting the above into (3.3.12) completes the proof Lemma 3.3.2. □

The next result, which is key to the computation of moments, is a direct corollary of Lemmas 3.3.1 and 3.3.2.

Lemma 3.3.3. *Let $\alpha: \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function satisfying* (3.1.1)–(3.1.4) *with some $A_0 \in (0,1)$, $\beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0,1)$ and $r \in (0,1)$. Fix $\epsilon_0 \in (0,1)$. Then there exist constants $\delta_0 > 0$ and $Q_0 \geq 2$, such that uniformly for all sufficiently large x, any $\delta \in [\delta_0 \log\log x/\log x, 1]$, and any square-free $a \in \mathbb{N} \cap [1, x]$ with $\omega(a) \leq (1 - \rho_0)\epsilon_0 \delta^{-1}$, $P^-(a) > Q_0$ and $P^+(a) \leq x^\delta$, we have*

$$\sum_{\substack{n \leq x \\ a|n}} \alpha(n) = \lambda_\alpha \left(F(\sigma_0, a) + O\left(\frac{2^{O(\omega(a))} L(a)}{a}\left(\frac{1}{(\log x)^{A_0}} + \frac{\epsilon_{\beta,1} \log P^+(a)}{\log x}\right)\right)\right) x^{\sigma_0} (\log x)^{\beta - 1},$$

where $L(a)$ is defined as in Lemma 3.3.1,

$$F(\sigma_0, a) := \prod_{p|a}\left(1 - \left(\sum_{\nu=0}^{\infty} \alpha(p^\nu) p^{-\sigma_0 \nu}\right)^{-1}\right),$$

and λ_α is defined by (3.3.5).

Proof. Suppose that $\delta_0 > 0$ is a constant for which Lemma 3.3.1 holds when $c_0 = 1$ and $h \in \{\beta - 1, \beta - 1 - A_0\}$. Let $Q_0 \geq 2$ be such that

$$\sum_{\nu=1}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \leq \frac{1}{2}$$

for all $p > Q_0$. Then we have

$$F(\sigma_0, p) = \sum_{\nu=1}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} + O\left(\left(\sum_{\nu=1}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}}\right)^2\right) = \frac{\alpha(p)}{p^{\sigma_0}} + O\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right) \quad (3.3.17)$$

for all $p > Q_0$. For any square-free integer $a \in [1, x]$ with $\omega(a) \leq (1 - \varrho_0)\epsilon_0 \delta^{-1}$, $P^-(a) > Q_0$ and $P^+(a) \leq x^\delta$, we have by Lemma 3.3.2 that

$$\sum_{\substack{n \leq x \\ \gcd(n,a)=1}} \alpha(n) = \lambda_\alpha(a) x^{\sigma_0} (\log 3x)^{\beta - 1} \left(1 + O\left(\frac{1}{(\log 3x)^{A_0}}\right)\right). \quad (3.3.18)$$

Note that

$$\sum_{\substack{n \leq x \\ a | n}} \alpha(n) = \sum_{\substack{q \leq x \\ R_q = a}} \alpha(q) \sum_{\substack{n' \leq x/q \\ \gcd(n',a)=1}} \alpha(n').$$

By (3.3.18), the main term of the inner sum contributes

$$\lambda_\alpha(a) x^{\sigma_0} \sum_{\substack{q \leq x \\ R_q = a}} \alpha(q) \left(\log \frac{3x}{q}\right)^{\beta - 1},$$

which, by Lemma 3.3.1, is equal to

$$\lambda_\alpha(a) x^{\sigma_0} \left(\tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(\omega(a))}}{\log x}\left(\frac{1}{x^{\delta\omega(a)}} + \frac{\epsilon_{\beta,1} L(a) \log P^+(a)}{a}\right)\right)\right) (\log x)^{\beta - 1}$$
$$= \lambda_\alpha \left(F(\sigma_0, a) + O\left(\frac{2^{O(\omega(a))}}{\log x}\left(\frac{1}{a} + \frac{\epsilon_{\beta,1} L(a) \log P^+(a)}{a}\right)\right)\right) x^{\sigma_0} (\log x)^{\beta - 1},$$

since $a \leq x^{\delta\omega(a)}$. Analogously, the contribution from the error term of the inner sum is

$$\begin{aligned} &\ll \lambda_\alpha \left(F(\sigma_0, a) + \frac{2^{O(\omega(a))} L(a) \log P^+(a)}{a \log x} \right) x^{\sigma_0} (\log x)^{\beta-1-A_0} \\ &\ll \frac{\lambda_\alpha 2^{O(\omega(a))} L(a)}{a} x^{\sigma_0} (\log x)^{\beta-1-A_0}, \end{aligned}$$

where we have used the estimate $F(\sigma_0, a) \ll 2^{O(\omega(a))} L(a)/a$, which follows directly from (3.1.4) and (3.3.17). Combining these estimates completes the proof of Lemma 3.3.3. □

Remark 3.3.1. We point out that the lower bound Q_0 for $\omega(a)$ in the lemma above is by and large an artificial thing, whose value is insignificant for our applications. However, we need it because (3.3.17) may not hold for small primes. As we shall see later, having such a lower bound also frees us from dealing with minor contributions from small primes.

Section 3.4

Computing Moments

By rescaling the strongly additive function f in Theorem 3.2.1, we may assume, without loss of generality, that $|f(p)| \leq 1$ for all primes p. Note that $0 \leq F(\sigma_0, p) < 1$ for all primes p. For every p we define $f_p \colon \mathbb{N} \to \mathbb{R}$ by

$$f_p(n) := \begin{cases} f(p)(1 - F(\sigma_0, p)), & \text{if } p \mid n, \\ -f(p) F(\sigma_0, p), & \text{otherwise.} \end{cases}$$

Given any $q \in \mathbb{N}$ we may also extend f_p via complete multiplicativity by setting

$$f_q(n) := \prod_{p^\nu \| q} f_p(n)^\nu.$$

It is clear that $|f_q(n)| \leq 1$. The following result provides an approximation of the moments of f in terms of those of f_p.

Lemma 3.4.1. *Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function satisfying* (3.1.1)–(3.1.4) *with some $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0,1)$ and $r \in (0,1)$. Let $f\colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \leq 1$ for all primes p. Then there exists a constant $Q_0 \geq 2$, such that*

$$\sum_{n \leq y} \alpha(n)(f(n) - A(x))^m = \sum_{n \leq y} \alpha(n) \left(\sum_{Q_0 < p \leq z} f_p(n) \right)^m + O\left(E(y,z,w;m)\right)$$

holds uniformly for all sufficiently large $x \geq z$, any $y \geq 1$, and all $m \in \mathbb{N}$, where

$$E(y,z,w;m) := \sum_{\substack{a+b+c=m \\ 0 \leq a < m \\ b,c \geq 0}} \binom{m}{a,b,c} 2^{O(m-a)} \left(\log(v+2)\right)^c \sum_{n \leq y} \alpha(n) \left| \sum_{Q_0 < p \leq z} f_p(n) \right|^a \omega(n;z,w)^b,$$

$v := \log x / \log z$, $w := x^{1/\log(v+2)}$, and

$$\omega(n;z,w) := \sum_{\substack{z < p \leq w \\ p \mid n}} 1.$$

Proof. Let $Q_0 \geq 2$ be a constant for which (3.3.17) holds. Suppose that $z > Q_0$ is sufficiently large. By (3.1.4), (3.3.17) and the fact that $\sum_p \psi_0(p) < \infty$, we find

$$\sum_{Q_0 < p \leq x} f(p) F(\sigma_0, p) = A(x) + O(1).$$

We compute

$$\begin{aligned} f(n) - A(x) &= \sum_{\substack{p|n \\ p>Q_0}} f(p) - \sum_{Q_0<p\le x} f(p)F(\sigma_0,p) + O(1) \\ &= \sum_{\substack{Q_0<p\le z \\ p|n}} f(p) + \sum_{\substack{p>z \\ p|n}} f(p) - \sum_{Q_0<p\le z} f(p)F(\sigma_0,p) - \sum_{z<p\le x} f(p)F(\sigma_0,p) + O(1) \\ &= \sum_{Q_0<p\le z} f_p(n) + \sum_{\substack{p>z \\ p|n}} f(p) - \sum_{z<p\le x} f(p)F(\sigma_0,p) + O(1). \end{aligned}$$

By (3.3.1) we have

$$\left|\sum_{z<p\le x} f(p)F(\sigma_0,p)\right| \le \sum_{z<p\le x} \frac{\alpha(p)}{p^{\sigma_0}} + O(1) \le \beta \log(v+2) + O(1).$$

Since

$$\sum_{\substack{p>z \\ p|n}} |f(p)| \le \sum_{\substack{z<p\le x \\ p|n}} 1 < \omega(n;z,w) + \log(v+2),$$

it follows that

$$f(n) - A(x) = \sum_{Q_0<p\le z} f_p(n) + O\left(\omega(n;z,w) + \log(v+2)\right).$$

We have therefore proved

$$\sum_{n\le y} \alpha(n)(f(n)-A(x))^m = \sum_{n\le y} \alpha(n) \left(\sum_{Q_0<p\le z} f_p(n) + O\left(\omega(n;z,w) + \log(v+2)\right) \right)^m.$$

Opening the mth power on the right-hand side by means of the multinomial theorem completes the proof of Lemma 3.4.1. □

Let $z = x^{1/v}$ and $w = x^{1/\log(v+2)}$ be as in Lemma 3.4.1, where $v \ge 1$ is a function of x and m to be chosen later. Fix $\epsilon_0 \in (0,1)$ and $\eta_0 \in (0,1]$, and suppose that

$y \in [x^{\eta_0}, x]$. Under the hypotheses in Theorem 3.2.1, we seek to estimate the weighted moments

$$\sum_{n\le y} \alpha(n) \left(\sum_{Q_0<p\le z} f_p(n) \right)^m,$$

appearing in Lemma 3.4.1. Expanding out the mth power we see that

$$\sum_{n\le y} \alpha(n) \left(\sum_{Q_0<p\le z} f_p(n) \right)^m = \sum_{Q_0<p_1,\ldots,p_m\le z} \sum_{n\le y} \alpha(n) f_{p_1\cdots p_m}(n). \tag{3.4.1}$$

This suggests that we study the sum

$$\sum_{n\le y} \alpha(n) f_q(n)$$

for $q \in \mathbb{N}$ with $\omega(q) \le m$, $P^-(q) > Q_0$ and $P^+(q) \le z$. A key observation is that $f_q(n) = f_q(\gcd(n, R_q))$. From this we deduce

$$\sum_{n\le y} \alpha(n) f_q(n) = \sum_{a|R_q} f_q(a) \sum_{\substack{n\le y\\ \gcd(n,R_q)=a}} \alpha(n) = \sum_{ab|R_q} f_q(a)\mu(b) \sum_{\substack{n\le y\\ ab|n}} \alpha(n).$$

Note that $\log y/\log z \in [\eta_0 v, v]$. By Lemma 3.3.3, there exists a constant $v_0 > 0$, independent of Q_0 and η_0, such that

$$\sum_{n\le y} \alpha(n) f_q(n) = \lambda_\alpha \left(G(\sigma_0, q) + O\left(2^{O(m)} E_y(q)\right)\right) y^{\sigma_0} (\log y)^{\beta-1} \tag{3.4.2}$$

holds uniformly for all sufficiently large x, any $y \in [x^{\eta_0}, x]$ and $v \in [\eta_0^{-1}, v_0 \log x/\log\log x]$, and all $m \le (1-\varrho_0)\epsilon_0 \log y/\log z$, where

$$\begin{aligned} G(\sigma_0, q) &:= \sum_{ab \mid R_q} f_q(a)\mu(b)F(\sigma_0, ab), \\ E_y(q) &:= \sum_{ab \mid R_q} \frac{|f_q(a)|L(ab)}{ab} \left(\frac{1}{(\log y)^{A_0}} + \frac{\epsilon_{\beta,1} \log P^+(ab)}{\log y} \right). \end{aligned}$$

Combining (3.4.2) with (3.4.1) gives

$$\sum_{n \le y} \alpha(n) \left(\sum_{Q_0 < p \le z} f_p(n) \right)^m = \lambda_\alpha \left(G(z) + O\left(2^{O(m)} D(y,z)\right)\right) y^{\sigma_0} (\log y)^{\beta - 1}, \quad (3.4.3)$$

where

$$\begin{aligned} G(z) &:= \sum_{Q_0 < p_1, \ldots, p_m \le z} G(\sigma_0, p_1 \cdots p_m), \\ D(y,z) &:= \sum_{Q_0 < p_1, \ldots, p_m \le z} E_y(p_1 \cdots p_m). \end{aligned}$$

Section 3.5

Estimation of $G(z)$ and $D(y,z)$

It is easy to see that $G(\sigma_0,q)$ is multiplicative as a function of q. Indeed, given any $q_1,q_2\in\mathbb{N}$ with $\gcd(q_1,q_2)=1$, we have

$$\begin{aligned} G(\sigma_0,q_1)G(\sigma_0,q_2) &= \sum_{\substack{a_1b_1|R_{q_1}\\ a_2b_2|R_{q_2}}} f_{q_1}(a_1)f_{q_2}(a_2)\mu(b_1)\mu(b_2)F(\sigma_0,a_1b_1)F(\sigma_0,a_2b_2)\\ &= \sum_{\substack{a_1b_1|R_{q_1}\\ a_2b_2|R_{q_2}}} f_{q_1}(a_1a_2)f_{q_2}(a_1a_2)\mu(b_1b_2)F(\sigma_0,a_1a_2b_1b_2)\\ &= \sum_{\substack{a_1b_1|R_{q_1}\\ a_2b_2|R_{q_2}}} f_{q_1q_2}(a_1a_2)\mu(b_1b_2)F(\sigma_0,a_1a_2b_1b_2)\\ &= \sum_{ab|R_{q_1q_2}} f_{q_1q_2}(a)\mu(b)F(\sigma_0,ab)=G(\sigma_0,q_1q_2). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} G(\sigma_0,p^\nu) &= f_{p^\nu}(1)+f_{p^\nu}(p)F(\sigma_0,p)-f_{p^\nu}(1)F(\sigma_0,p)\\ &= (-f(p)F(\sigma_0,p))^\nu+(f(p)(1-F(\sigma_0,p)))^\nu F(\sigma_0,p)-(-f(p)F(\sigma_0,p))^\nu F(\sigma_0,p)\\ &= f(p)^\nu F(\sigma_0,p)(1-F(\sigma_0,p))\left((-1)^\nu F(\sigma_0,p)^{\nu-1}+(1-F(\sigma_0,p))^{\nu-1}\right) \end{aligned}$$

for all prime powers p^ν. Note that $G(\sigma_0,p)=0$, $|G(\sigma_0,p^\nu)|\le 1/4$, and $G(\sigma_0,p^\nu)\ge 0$ when $2\mid\nu$. In addition, we have by (3.3.17) that

$$G(\sigma_0,p^2)=f(p)^2F(\sigma_0,p)(1-F(\sigma_0,p))=\alpha(p)\frac{f(p)^2}{p^{\sigma_0}}+O\left(\psi_0(p)+\frac{\alpha(p)^2}{p^{2\sigma_0}}\right) \quad (3.5.1)$$

and that

$$|G(\sigma_0, p^\nu)| \le |f(p)|^\nu F(\sigma_0, p) \le \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} + O\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right) \qquad (3.5.2)$$

for all p^ν with $p > Q_0$ and $\nu \ge 2$.

Now we proceed to estimate $G(z)$ in the main term of (3.4.3). Recall that $y \in [x^{\eta_0}, x]$ and $z = x^{1/v}$. We shall suppose in this section that $1 \le m \le \min(v, h_0 B(x)^{1/3})$, $\log(v+2) = o(B(x))$, and $m\log(v+2) \ll B(x)$, where $0 < h_0 < (3/2)^{2/3}$ is any given constant, and obtain a uniform treatment for $G(z)$ and $D(y,z)$ under this more general assumption. Since $G(\sigma_0, q)$ is multiplicative in q and $G(\sigma_0, p) = 0$ for all $p > Q_0$, we have

$$G(z) = \sum_{\substack{Q_0 < p_1, \ldots, p_m \le z \\ p_1 \cdots p_m \text{ square-full}}} G(\sigma_0, p_1 \cdots p_m). \qquad (3.5.3)$$

When $2 \mid m$, the main contribution arises from

$$\frac{m!}{(m/2)!\,2^{m/2}} \sum_{\substack{Q_0 < p_1, \ldots, p_{m/2} \le z \\ p_1, \ldots, p_{m/2} \text{ distinct}}} G(\sigma_0, p_1^2 \cdots p_{m/2}^2) = C_m \sum_{\substack{Q_0 < p_1, \ldots, p_{m/2} \le z \\ p_1, \ldots, p_{m/2} \text{ distinct}}} \prod_{i=1}^{m/2} G(\sigma_0, p_i^2), \qquad (3.5.4)$$

since the number of ways to partition a set of m elements into $m/2$ two-element equivalence classes is

$$\frac{m!}{(m/2)!\,2^{m/2}} = \frac{m!}{m!\,!} = C_m.$$

The sum on the right-hand side of (3.5.4) can be rewritten as

$$\sum_{\substack{Q_0 < p_1, \ldots, p_{m/2-1} \le z \\ p_1, \ldots, p_{m/2-1} \text{ distinct}}} \prod_{i=1}^{m/2-1} G(\sigma_0, p_i^2) \sum_{\substack{Q_0 < p_{m/2} \le z \\ p_{m/2} \ne p_1, \ldots, p_{m/2-1}}} G(\sigma_0, p_{m/2}^2).$$

By (3.5.1) and (3.3.1), the inner sum over $p_{m/2}$ is equal to

$$\sum_{Q_0<p\le z} G(\sigma_0,p^2) - \sum_{i=1}^{m/2-1} G(\sigma_0,p_i^2) = B(z) + O\left(\sum_{Q_0<p\le q_N} \frac{\alpha(p)}{p^{\sigma_0}}\right) = B(z) + O\left(\log\log(m+2)\right),$$

where $N = m/2 + \pi(Q_0)$ and q_N is the Nth prime. Repeating this argument we obtain

$$\sum_{\substack{Q_0<p_1,\dots,p_{m/2}\le z\\ p_1,\dots,p_{m/2}\ \text{distinct}}} \prod_{i=1}^{m/2} G(\sigma_0,p_i^2) = \left(B(z) + O\left(\log\log(m+2)\right)\right)^{m/2}.$$

But

$$B(x) - B(z) = \sum_{z<p\le x} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} \le \beta\log(v+2) + O(1).$$

Hence when m is even, the main contribution to $G(z)$ is given by

$$C_m\left(B(x) + O(\log(v+2))\right)^{m/2} = C_m B(x)^{\frac{m}{2}}\left(1 + O\left(mB(x)^{-1}\log(v+2)\right)\right).$$

The remaining contribution to $G(z)$ comes from

$$\sum_{s<m/2}\ \sum_{Q_0<p_1<\dots<p_s\le z}\ \sum_{\substack{k_1+\dots+k_s=m\\ k_1,\dots,k_s\ge 2}} \binom{m}{k_1,\dots,k_s} \prod_{i=1}^{s} G(\sigma_0,p_i^{k_i}). \tag{3.5.5}$$

Since (3.5.5) vanishes when $m\le 2$, we may suppose $m\ge 3$. By (3.5.2) we see that

$$\prod_{i=1}^{s}\left|G(\sigma_0,p_i^{k_i})\right| \le \prod_{i=1}^{s}\left(\alpha(p_i)\frac{f(p_i)^2}{p_i^{\sigma_0}} + O\left(\psi_0(p_i) + \frac{\alpha(p_i)^2}{p_i^{2\sigma_0}}\right)\right).$$

Thus, we have

$$\sum_{Q_0<p_1<\dots<p_s\le z}\prod_{i=1}^{s}\left|G(\sigma_0,p_i^{k_i})\right| \le \frac{1}{s!}(B(x)+O(1))^s = \frac{1}{s!}B(x)^s\left(1+O\left(sB(x)^{-1}\right)\right) \ll \frac{B(x)^s}{s!}.$$

Since

$$\sum_{\substack{k_1+\cdots+k_s=m\\ k_1,\ldots,k_s\geq 2}} \binom{m}{k_1,\ldots,k_s} \leq \frac{m!}{2^s} \sum_{\substack{k_1+\cdots+k_s=m\\ k_1,\ldots,k_s\geq 2}} 1 = \frac{m!}{2^s}\binom{m-s-1}{s-1},$$

(3.5.5) is

$$\ll m! \sum_{s<m/2} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} B(x)^s.$$

To estimate the sum above, we put $m_1 := \lfloor (m-1)/2 \rfloor$ and observe that

$$\begin{aligned}\sum_{s<m/2} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} B(x)^s &= B(x)^{m_1} \sum_{s\leq m_1} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} B(x)^{s-m_1}\\ &\leq B(x)^{m_1} m^{-3m_1} \sum_{s\leq m_1} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s},\end{aligned}$$

where we have used the assumption that $B(x) \geq m^3/h_0^3$ with some $0 < h_0 < (3/2)^{2/3}$.

Let

$$e_m := \begin{cases} 1, & \text{if } 2 \mid m,\\ 1/2, & \text{otherwise.}\end{cases}$$

Then $m_1 = m/2 - e_m$. Note that

$$\begin{aligned} m^{-3m_1} \sum_{s\leq m/4} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s} &\leq m^{-3m_1} \sum_{s\leq m/4} \frac{1}{s!\,(s-1)!} \left(\frac{9}{4}\right)^{m_1-s} \left(\frac{m^4}{2}\right)^{s}\\ &\ll m^{-3m_1} \left(\frac{9}{4}\right)^{m_1} \left(\frac{m^4}{2}\right)^{m/4} \ll \frac{C_m}{m!} m^{3e_m},\end{aligned}$$

since

$$\frac{C_m}{m!} = \frac{1}{2^{m/2}\Gamma(m/2+1)} \asymp m^{-\frac{m+1}{2}} e^{\frac{m}{2}}$$

by Stirling's formula. Next, we have

$$\begin{aligned} m^{-3m_1} \sum_{m/4<s\le m/3} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s} &\le 2^{O(m)} m^{-3m_1} \sum_{m/4<s\le m/3} \frac{1}{s!\,(s-1)!}\left(\frac{m^4}{2}\right)^s \\ &\le \frac{2^{O(m)} m^{-3m_1}}{m^{m/2}} \sum_{m/4<s\le m/3} \left(\frac{m^4}{2}\right)^s \\ &\le \frac{2^{O(m)} m^{-3m_1}}{m^{m/2}} m^{4m/3} \\ &= 2^{O(m)} m^{-2m/3+3e_m} \ll \frac{C_m}{m!} m^{3e_m}. \end{aligned}$$

Finally, we observe that

$$\begin{aligned} &m^{-3m_1} \sum_{m/3<s\le m_1} \frac{1}{s!\,2^s}\binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s} \\ &= m^{-3m_1} \sum_{m/3<s\le m_1} \frac{1}{s!\,2^s}\binom{m-s-1}{m-2s} h_0^{3(m_1-s)} m^{3s} \\ &\le m^{-3m_1} \sum_{m/3<s\le m_1} \frac{1}{s!\,2^s}(m-s)^{m-2s} h_0^{3(m_1-s)} m^{3s} \\ &\le \frac{m^{-3m_1}}{m_1!} \sum_{m/3<s\le m_1} \frac{m_1!}{s!\,2^s}\left(\frac{2m}{3}\right)^{m-2s} h_0^{3(m_1-s)} m^{3s} \\ &\le \frac{m^{-3m_1}}{m_1!} \sum_{m/3<s\le m_1} \frac{1}{2^s}\left(\frac{m}{2}\right)^{m_1-s}\left(\frac{2m}{3}\right)^{m-2s} h_0^{3(m_1-s)} m^{3s} \\ &\ll \frac{m^{m-2m_1}}{m_1!\,2^{m/2}} \sum_{m/3<s\le m_1} \left(\frac{2h_0^{3/2}}{3}\right)^{m-2s} \ll \frac{C_m}{m!} m^{3e_m}. \end{aligned}$$

Collecting the estimates above, we see that the contribution to $G(z)$ from (3.5.5) is

$$\ll C_m m^{3e_m} B(x)^{m_1} = C_m B(x)^{\frac{m}{2}} \left(\frac{m^3}{B(x)}\right)^{e_m} \le C_m B(x)^{\frac{m}{2}} \frac{m^{\frac{3}{2}}}{\sqrt{B(x)}}.$$

We can therefore conclude that

$$G(z) = C_m B(x)^{\frac{m}{2}} \left(\chi_m \left(1 + O\left(\frac{m \log(v+2)}{B(x)} \right) \right) + O\left(\frac{m^{\frac{3}{2}}}{\sqrt{B(x)}} \right) \right). \tag{3.5.6}$$

Next, we estimate $D(y,z)$ in the error term of (3.4.3). By definition, we have

$$D(y,z) = \sum_{s \le m} \sum_{Q_0 < p_1 < \ldots < p_s \le z} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} E_y\left(p_1^{k_1} \cdots p_s^{k_s}\right).$$

Let

$$H(\sigma_0, q) := \sum_{ab | R_q} \frac{|f_q(a)| L(ab)}{ab}.$$

Then $H(\sigma_0, q)$ is multiplicative in q. Moreover, we have

$$E_y(q) \le H(\sigma_0, q) \left(\frac{1}{(\log y)^{A_0}} + \frac{\epsilon_{\beta,1} \log P^+(q)}{\log y} \right).$$

It follows that $D(y,z) \le D_1(y,z) + \epsilon_{\beta,1} D_2(y,z)$, where

$$D_1(y,z) := \frac{1}{(\log y)^{A_0}} \sum_{s \le m} \sum_{Q_0 < p_1 < \ldots < p_s \le z} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \prod_{i=1}^{s} H(\sigma_0, p_i^{k_i}),$$

$$D_2(y,z) := \frac{1}{\log y} \sum_{s \le m} \sum_{Q_0 < p_1 < \ldots < p_s \le z} \log p_s \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \prod_{i=1}^{s} H(\sigma_0, p_i^{k_i}).$$

By Mertens' theorems [35, Theorems 425, 427] we have, for any $t \ge 3$, that

$$\sum_{p \le t} \frac{(\log\log(p+1))^{\vartheta_0}}{p} = \frac{1}{\vartheta_0 + 1} (\log\log t)^{\vartheta_0 + 1} + O(1) \tag{3.5.7}$$

and that

$$\sum_{p\le t}\frac{(\log\log(p+1))^{\vartheta_0}\log p}{p}=\left(1+O\left(\frac{1}{\log t}+\frac{\vartheta_0}{\log\log t}\right)\right)(\log\log t)^{\vartheta_0}\log t. \quad (3.5.8)$$

Furthermore, let

$$T_n(t):=\sum_{k=0}^{n}\left\{{n \atop k}\right\}t^k$$

denote the nth Touchard polynomial, where

$$\left\{{n \atop k}\right\}:=\frac{1}{k!}\sum_{\substack{n_1+\cdots+n_k=n\\ n_1,\dots,n_k\in\mathbb{N}}}\binom{n}{n_1,\dots,n_k}$$

is the kth Stirling number of the second kind of size n. The sequence $\{T_n(t)\}_{n=0}^{\infty}$ of the Touchard polynomials is known to satisfy the recurrence relation

$$T_{n+1}(t)=t\sum_{i=0}^{n}\binom{n}{i}T_i(t),$$

from which one verifies readily by induction that

$$T_n(t)\le\left(t+\frac{n-1}{2}\right)^n \quad (3.5.9)$$

for all $n\ge 1$ and $t\ge 0$. Since

$$\begin{aligned}H(\sigma_0,p^\nu)&=|f(p)|F(\sigma_0,p)\left(1+\frac{L(p)}{p}\right)+\frac{|f(p)|L(p)}{p}(1-F(\sigma_0,p))\\&=|f(p)|\left(F(\sigma_0,p)+\frac{(\log\log(p+1))^{\vartheta_0}}{p}\right)\end{aligned}$$

for any prime powers p^ν with $p>Q_0$, we obtain, from (3.3.17), (3.5.7), (3.5.8) and

(3.5.9), that

$$D_1(y,z) \le \frac{2^{O(m)}}{(\log x)^{A_0}} \sum_{s\le m} \frac{1}{s!} (\log\log z)^{s(\vartheta_0+1)} \sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s\in\mathbb{N}}} \binom{m}{k_1, \ldots, k_s}$$
$$\le \frac{2^{O(m)}}{(\log x)^{A_0}} T_m\left((\log\log z)^{\vartheta_0+1}\right) \le \frac{2^{O(m)}}{(\log x)^{A_0}} (\log\log x)^{m(\vartheta_0+1)},$$

and that

$$D_2(y,z) \le \frac{2^{O(m)}\log z}{\log x} \sum_{s\le m} \frac{1}{(s-1)!} (\log\log z)^{s(\vartheta_0+1)-1} \sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s\in\mathbb{N}}} \binom{m}{k_1, \ldots, k_s}$$
$$\le \frac{2^{O(m)}}{v\log\log z} T_m\left((\log\log z)^{\vartheta_0+1}\right) \le \frac{2^{O(m)}}{v} (\log\log x)^{m(\vartheta_0+1)-1}.$$

Hence, we conclude that

$$D(y,z) \le 2^{O(m)} (\log\log x)^{m(\vartheta_0+1)-1} \left(\frac{\log\log x}{(\log x)^{A_0}} + \frac{\epsilon_{\beta,1}}{v} \right). \tag{3.5.10}$$

Section 3.6

Estimation of $E(y, z, w; m)$

In this section, we seek to bound the function $E(y, z, w; m)$ introduced in Lemma 3.4.1 under the assumptions in Theorem 3.2.1. We start with the case $\beta = 1$. Suppose that $1 \le m \le h_0 B(x)^{1/3}$, where $0 < h_0 < (3/2)^{2/3}$ is any given constant. Recall that $y \in [x^{\eta_0}, x]$, $z = x^{1/v}$ and $w = x^{1/\log(v+2)}$. With the choice $v = (1-\varrho_0)^{-1}\epsilon_0^{-1}\eta_0^{-1}m$, we clearly have $v \in [\eta_0^{-1}, v_0 \log x/\log\log x]$ and $m \le (1-\varrho_0)\epsilon_0 \log y/\log z$. Inputting

(3.5.6) and (3.5.10) into (3.4.3), we obtain

$$\sum_{n\le y}\alpha(n)\left(\sum_{Q_0<p\le z} f_p(n)\right)^m = \lambda_\alpha C_m B(x)^{\frac{m}{2}}\left(\chi_m + O\left(\frac{m^{\frac{3}{2}}}{\sqrt{B(x)}}\right)\right) y^{\sigma_0}. \qquad (3.6.1)$$

The key lies in the estimation of the sum

$$\sum_{n\le y}\alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a \omega(n;z,w)^b. \qquad (3.6.2)$$

In the present case, we may simply use the trivial bound $\omega(n;z,w) \ll v \ll m$, so that (3.6.2) is bounded above by

$$2^{O(b)} m^b \sum_{n\le y}\alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a.$$

It is clear that we can use (3.6.1) to handle the sum above. If a is even, then this sum is $\ll \lambda_\alpha C_a B(x)^{\frac{a}{2}} y^{\sigma_0}$; if a is odd, then it is

$$\begin{aligned}&\le \left(\sum_{n\le y}\alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^{a-1}\right)^{1/2}\left(\sum_{n\le y}\alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^{a+1}\right)^{1/2}\\ &\ll \lambda_\alpha\sqrt{C_{a-1}C_{a+1}}B(x)^{\frac{a}{2}}y^{\sigma_0}\end{aligned}$$

by the Cauchy–Schwarz inequality. The sequence $\{C_\ell\}_{\ell=1}^{\infty}$ is strictly increasing, which can be easily seen from the identity

$$\frac{C_{\ell+1}}{C_\ell} = \frac{\ell+1}{\sqrt{2}}\cdot\frac{\Gamma(\ell/2+1)}{\Gamma((\ell+1)/2+1)} = \sqrt{2}\cdot\frac{\Gamma(\ell/2+1)}{\Gamma((\ell/2+1/2)}$$

and the fact that $\Gamma(y)$ is strictly increasing on $[3/2, \infty)$. Moreover, we have by Stirling's formula that

$$\frac{C_\ell}{C_{\ell+1}} \ll \frac{1}{\ell+1} \cdot \frac{((\ell+1)/2)^{\ell/2+1} e^{-(\ell+1)/2}}{(\ell/2)^{(\ell+1)/2} e^{-\ell/2}} \ll \frac{1}{\sqrt{\ell+1}},$$

which implies that

$$C_a \leq 2^{O(m-a)} C_m \sqrt{\frac{a!}{m!}} \leq 2^{O(m-a)} C_m \sqrt{\frac{a^a}{m^m}} \leq \frac{2^{O(m-a)} C_m}{(\sqrt{m})^{m-a}}$$

for all $0 \leq a \leq m$. Hence, (3.6.2) is bounded above by

$$\frac{2^{O(m-a)} \lambda_\alpha C_m m^b}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} y^{\sigma_0} \leq 2^{O(m-a)} \lambda_\alpha C_m \left(\sqrt{m}\right)^{m-a} B(x)^{\frac{a}{2}} y^{\sigma_0}.$$

Inputting this inequality into the definition of $E(y, z, w; m)$, we conclude that

$$E(y, z, w; m) \leq \lambda_\alpha C_m y^{\sigma_0} \sum_{a=0}^{m-1} \binom{m}{a} B(x)^{\frac{a}{2}} \left(O\left(\sqrt{m}\right)\right)^{m-a} \ll \lambda_\alpha C_m m^{\frac{3}{2}} B(x)^{\frac{m-1}{2}} y^{\sigma_0}. \tag{3.6.3}$$

Now we consider the case $\beta \neq 1$. Suppose that $1 \leq m \ll B(x)^{1/3}/(\log\log\log x)^{2/3}$ and that $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$. In this case we take $v = (\log\log x)^{m(\vartheta_0+2)}$, so that $v \in [2\eta_0^{-1}, v_0 \log x/\log\log x]$ and $m \leq (1-\varrho_0)\epsilon_0 \log t/\log z$ for any $t \in [x^{\eta_0/2}, x]$ when x is sufficiently large. Inserting (3.5.6) and (3.5.10) into (3.4.3) leads to

$$\sum_{n \leq t} \alpha(n) \left(\sum_{Q_0 < p \leq z} f_p(n) \right)^m = \lambda_\alpha C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{m^{\frac{3}{2}}}{\sqrt{B(x)}} \right) \right) t^{\sigma_0} (\log t)^{\beta-1} \tag{3.6.4}$$

uniformly for all $t \in [x^{\eta_0/2}, x]$. Again, we need to estimate (3.6.2) uniformly for

$y \in [x^{\eta_0}, x]$. Note that (3.6.2) can be rewritten as

$$\sum_{k=1}^{b} \sum_{z<p_1<\ldots<p_k\le w} \sum_{\substack{l_1+\cdots+l_k=b\\ l_1,\ldots,l_k\ge 1}} \binom{b}{l_1,\ldots,l_k} \sum_{\substack{n\le y\\ p_1\cdots p_k|n}} \alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a .$$

Observe that

$$\begin{aligned}\sum_{\substack{n\le y\\ p_1\cdots p_k|n}} \alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a &= \sum_{\substack{q\le y\\ R_q=p_1\cdots p_k}} \alpha(q) \sum_{\substack{n'\le y/q\\ \gcd(n',q)=1}} \alpha(n')\left|\sum_{Q_0<p\le z} f_p(n')\right|^a\\ &\le \sum_{\substack{q\le y\\ R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\le y/q} \alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a ,\end{aligned}$$

since $p_1,\ldots,p_k>p$. If $q=p_1^{\nu_1}\cdots p_k^{\nu_k}>\sqrt{y}$ with given $z<p_1<\ldots<p_k\le w$, then we have the trivial estimate

$$\sum_{n\le y/q} \alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a \le \pi(z)^a \sum_{n\le 3y/q} \alpha(n) \ll \lambda_\alpha \pi(z)^a \left(\frac{y}{q}\right)^{\sigma_0}\left(\log\frac{3y}{q}\right)^{\beta-1}$$

by (3.3.4) and the fact that $|f_p(n)|\le 1$. By the proof of Lemma 3.3.1, and particularly by (3.3.3), we find that

$$\begin{aligned}\sum_{\substack{\sqrt{y}<q\le y\\ R_q=p_1\cdots p_k}} \frac{\alpha(q)}{q^{\sigma_0}}\left(\log\frac{3y}{q}\right)^{\beta-1} &= \sum_{\substack{\sqrt{y}<p_1^{\nu_1}\cdots p_k^{\nu_k}\le y\\ \nu_1,\ldots,\nu_k\ge 1}} \frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3y}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^{\beta-1}\\ &\ll \frac{2^{O(k)}(\log y)^{k+\beta-2}}{\left(\sqrt{y}\right)^{1-\varrho_0}},\end{aligned}$$

from which it follows that

$$\sum_{\substack{\sqrt{y}<q\le y\\ R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\le y/q} \alpha(n)\left|\sum_{Q_0<p\le z} f_p(n)\right|^a \ll \lambda_\alpha \pi(z)^a \frac{2^{O(k)}y^{\sigma_0}(\log y)^{k+\beta-2}}{\left(\sqrt{y}\right)^{1-\varrho_0}}.$$

Summing the above over all $z < p_1 < ... < p_k \le w$ yields immediately

$$\sum_{z<p_1<...<p_k\le w} \sum_{\substack{\sqrt{y}<q\le y \\ R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\le y/q} \alpha(n) \left| \sum_{Q_0<p\le z} f_p(n) \right|^a$$
$$\le \lambda_\alpha \pi(z)^a \pi(w)^k \frac{2^{O(k)} y^{\sigma_0} (\log y)^{k+\beta-2}}{k! \left(\sqrt{y}\right)^{1-\varrho_0}} \le \frac{\lambda_\alpha y^{\sigma_0} (\log y)^{\beta-1}}{k! \left(\sqrt[3]{y}\right)^{1-\varrho_0}} \tag{3.6.5}$$

for sufficiently large x, since $y \in [x^{\eta_0}, x]$, $a + k \le m \ll (\log\log x)^{1/3}/(\log\log\log x)^{2/3}$, and

$$\pi(z)^a \pi(w)^k \le \left(\frac{w}{\log w} + O\left(\frac{w}{(\log w)^2} \right) \right)^m \ll \left(\frac{w}{\log w} \right)^m \le \frac{x^{1/\log\log\log x} (m \log\log\log x)^m}{(\log x)^m}.$$

If $q = p_1^{\nu_1} \cdots p_k^{\nu_k} \le \sqrt{y}$, then $x^{\eta_0/2} \le \sqrt{y} \le y/q \le y \le x$. Thus, we can apply (3.6.4) with $t = y/q$ to handle

$$\sum_{n\le y/q} \alpha(n) \left| \sum_{Q_0<p\le z} f_p(n) \right|^a.$$

If a is even, then this sum is

$$\ll \lambda_\alpha C_a B(x)^{\frac{a}{2}} \left(\frac{y}{q} \right)^{\sigma_0} \left(\log \frac{y}{q} \right)^{\beta-1} \le \frac{2^{O(m-a)} \lambda_\alpha C_m}{\left(\sqrt{m}\right)^{m-a}} B(x)^{\frac{a}{2}} \left(\frac{y}{q} \right)^{\sigma_0} \left(\log \frac{y}{q} \right)^{\beta-1};$$

if a is odd, then it is

$$\le \left(\sum_{n\le y/q} \alpha(n) \left| \sum_{Q_0<p\le z} f_p(n) \right|^{a-1} \right)^{1/2} \left(\sum_{n\le y/q} \alpha(n) \left| \sum_{Q_0<p\le z} f_p(n) \right|^{a+1} \right)^{1/2}$$
$$\ll \lambda_\alpha \sqrt{C_{a-1} C_{a+1}} B(x)^{\frac{a}{2}} \left(\frac{y}{q} \right)^{\sigma_0} \left(\log \frac{y}{q} \right)^{\beta-1}$$
$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m}{\left(\sqrt{m}\right)^{m-a}} B(x)^{\frac{a}{2}} \left(\frac{y}{q} \right)^{\sigma_0} \left(\log \frac{y}{q} \right)^{\beta-1}$$

by Cauchy–Schwarz. It follows that

$$\begin{aligned}
&\sum_{\substack{q\le\sqrt{y}\\ R_q=p_1\cdots p_k}}\alpha(q)\sum_{n\le y/q}\alpha(n)\left|\sum_{Q_0<p\le z}f_p(n)\right|^a\\
&\le\frac{2^{O(m-a)}\lambda_\alpha C_m y^{\sigma_0}}{(\sqrt{m})^{m-a}}B(x)^{\frac{a}{2}}\sum_{\substack{q\le\sqrt{y}\\ R_q=p_1\cdots p_k}}\frac{\alpha(q)}{q^{\sigma_0}}\left(\log\frac{y}{q}\right)^{\beta-1}\\
&\le\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{\frac{a}{2}}y^{\sigma_0}(\log y)^{\beta-1}\prod_{i=1}^{k}\sum_{\nu=1}^{\infty}\frac{\alpha(p_i^\nu)}{p_i^{\sigma_0\nu}}\\
&=\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{\frac{a}{2}}y^{\sigma_0}(\log y)^{\beta-1}\prod_{i=1}^{k}\left(\frac{\alpha(p_i)}{p_i^{\sigma_0}}+\psi_0(p_i)\right)
\end{aligned}$$

for all $0\le a<m$. Since (3.3.1) implies that

$$\sum_{z<p_1<\ldots<p_k\le w}\prod_{i=1}^{k}\left(\frac{\alpha(p_i)}{p_i^{\sigma_0}}+\psi_0(p_i)\right)\le\frac{1}{k!}\left(\sum_{z<p\le w}\left(\frac{\alpha(p)}{p^{\sigma_0}}+\psi_0(p)\right)\right)^k\le\frac{2^{O(k)}}{k!}(\log v)^k,$$

we obtain

$$\begin{aligned}
&\sum_{z<p_1<\ldots<p_k\le w}\ \sum_{\substack{q\le\sqrt{y}\\ R_q=p_1\cdots p_k}}\alpha(q)\sum_{n\le y/q}\alpha(n)\left|\sum_{Q_0<p\le z}f_p(n)\right|^a\\
&\le\frac{2^{O(m-a)}\lambda_\alpha C_m}{k!\,(\sqrt{m})^{m-a}}(\log v)^k B(x)^{\frac{a}{2}}y^{\sigma_0}(\log y)^{\beta-1}.
\end{aligned}\tag{3.6.6}$$

Combining (3.6.6) with (3.6.5) and extending the inner sum over q to the entire range, we conclude that

$$\begin{aligned}
&\sum_{z<p_1<\ldots<p_k\le w}\ \sum_{\substack{q\le y\\ R_q=p_1\cdots p_k}}\alpha(q)\sum_{n\le y/q}\alpha(n)\left|\sum_{Q_0<p\le z}f_p(n)\right|^a\\
&\le\frac{2^{O(m-a)}\lambda_\alpha C_m}{k!\,(\sqrt{m})^{m-a}}(\log v)^k B(x)^{\frac{a}{2}}y^{\sigma_0}(\log y)^{\beta-1}.
\end{aligned}$$

Hence, (3.6.2) is bounded above by

$$\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} y^{\sigma_0} (\log y)^{\beta-1} \sum_{k=1}^{b} \frac{(\log v)^k}{k!} \sum_{\substack{l_1+\cdots+l_k=b \\ l_1,\ldots,l_k \geq 1}} \binom{b}{l_1, \ldots, l_k}$$
$$= \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} y^{\sigma_0} (\log y)^{\beta-1} \sum_{k=1}^{b} \left\{ {b \atop k} \right\} (\log v)^k$$
$$\leq \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} T_b(\log v) y^{\sigma_0} (\log y)^{\beta-1}.$$

It follows by (3.5.9) that the above does not exceed

$$\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} (\log v)^b y^{\sigma_0} (\log y)^{\beta-1},$$

where we have used the observation that $\log v > m \log\log\log x > m \geq b$. In other words, we have shown that

$$\sum_{n\leq y} \alpha(n) \left| \sum_{Q_0<p\leq z} f_p(n) \right|^a \omega(n; z, w)^b \leq \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{\frac{a}{2}} (\log v)^b y^{\sigma_0} (\log y)^{\beta-1}.$$

Inputting this inequality into the definition of $E(y, z, w; m)$, we conclude that

$$E(y, z, w; m) \leq \lambda_\alpha C_m y^{\sigma_0} (\log y)^{\beta-1} \sum_{a=0}^{m-1} \binom{m}{a} B(x)^{\frac{a}{2}} \left(O\left(\frac{\log v}{\sqrt{m}} \right) \right)^{m-a}$$
$$\ll \lambda_\alpha C_m \sqrt{m} (\log v) B(x)^{\frac{m-1}{2}} y^{\sigma_0} (\log y)^{\beta-1}$$
$$\ll \lambda_\alpha C_m m^{\frac{3}{2}} (\log\log\log x) B(x)^{\frac{m-1}{2}} y^{\sigma_0} (\log y)^{\beta-1}. \tag{3.6.7}$$

Section 3.7

Deduction of Theorems 3.2.1 and 3.2.2

Theorem 3.2.1 now follows immediately upon combining (3.6.1) and (3.6.4) with (3.6.3) and (3.6.7) and invoking Lemma 3.4.1 and (3.3.4). In fact, we have shown that the same asymptotic formulas which hold for $M(x;m)$ also hold for

$$S(y)^{-1}\sum_{n\le y}\alpha(n)(f(n)-A(x))^m \tag{3.7.1}$$

uniformly in the range $y\in[x^{\eta_0},x]$, where $\eta_0\in(0,1]$ is any fixed constant.

Now we prove Theorem 3.2.2. Recall that under the hypotheses in Theorem 3.2.2, the multiplicative function $\alpha(n)$ satisfies conditions (i)–(iv). We shall again suppose $A_0\in(0,1)$ throughout the proof. Define the strongly additive function $\tilde{f}\colon\mathbb{N}\to\mathbb{R}$, called the *strongly additive contraction of* f, by $\tilde{f}(p)=f(p)$ for all primes p. Then

$$\sum_{n\le x}\alpha(n)(f(n)-A(x))^m=\sum_{k=0}^{m}\binom{m}{k}\sum_{n\le x}\alpha(n)\left(\tilde{f}(n)-A(x)\right)^k\left(f(n)-\tilde{f}(n)\right)^{m-k} \tag{3.7.2}$$

for every $m\in\mathbb{N}$. The term corresponding to $k=m$ can be estimated directly using Theorem 3.2.1. Hence, it remains to deal with

$$\sum_{n\le x}\alpha(n)\left(\tilde{f}(n)-A(x)\right)^k\left(f(n)-\tilde{f}(n)\right)^l \tag{3.7.3}$$

for $0 \le k < m$ and $l = m - k$. Note that

$$\left|\sum_{n\le x}\alpha(n)\left(\tilde f(n)-A(x)\right)^k\left(f(n)-\tilde f(n)\right)^l\right|$$
$$\le \sum_{n\le x}\alpha(n)\left|\tilde f(n)-A(x)\right|^k\left|\sum_{p^\nu\|n,\nu\ge 2}(f(p^\nu)-f(p))\right|^l$$
$$\le \sum_{p_1,\dots,p_l\le\sqrt{x}}\ \sum_{\substack{p_1^{\nu_1},\dots,p_l^{\nu_l}\le x\\ \nu_1,\dots,\nu_l\ge 2}}|f(p_1^{\nu_1})-f(p_1)|\cdots|f(p_l^{\nu_l})-f(p_l)|\sum_{\substack{n\le x\\ p_1^{\nu_1},\dots,p_l^{\nu_l}\|n}}\alpha(n)\left|\tilde f(n)-A(x)\right|^k.$$

Since $f(p^\nu) = O(\nu^\kappa)$ for all p^ν, the last expression above does not exceed

$$2^{O(l)}\sum_{s\le l}\ \sum_{p_1<\dots<p_s\le\sqrt{x}}\ \sum_{\substack{l_1+\dots+l_s=l\\ l_1,\dots,l_s\in\mathbb{N}}}\binom{l}{l_1,\dots,l_s}\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le x\\ \nu_1,\dots,\nu_s\ge 2}}\nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\sum_{\substack{n\le x\\ p_1^{\nu_1},\dots,p_s^{\nu_s}\|n}}\alpha(n)\left|\tilde f(n)-A(x)\right|^k.$$

If we write $n = p_1^{\nu_1}\cdots p_s^{\nu_s}n'$ with $\gcd(n', p_1\cdots p_s) = 1$, then it is clear that

$$\left|\tilde f(n)-A(x)\right|^k = \left|\tilde f(n')-A(x)+\sum_{i=1}^s f(p_i)\right|^k \le \sum_{a=0}^k\binom{k}{a}\left|\tilde f(n')-A(x)\right|^a\left|\sum_{i=1}^s f(p_i)\right|^{k-a}.$$

Thus, the innermost sum of $\alpha(n)|\tilde f(n)-A(x)|^k$ is

$$\le \alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})\sum_{a=0}^k\binom{k}{a}\left|\sum_{i=1}^s f(p_i)\right|^{k-a}\sum_{n\le x/\left(p_1^{\nu_1}\cdots p_s^{\nu_s}\right)}\alpha(n)\left|\tilde f(n)-A(x)\right|^a,\quad (3.7.4)$$

where we have dropped the superscript of n for simplicity of notation. Since the right-hand side of the above clearly vanishes if $p_1\cdots p_s > \sqrt{x}$, we may assume $p_1\cdots p_s \le \sqrt{x}$ instead. Let $\lambda' := 1-\varrho_0-\log_2\lambda > \rho_0$, and choose a constant $\max(1/2, \sqrt{\varrho_0/\lambda'}) < \delta_0 < 1$, so that $1-\varrho_0+\delta_0^2\lambda' > 1$. Let $x_s := x/(p_1\cdots p_s)$ and $y_s := x_s^{\delta_0}$. Then $x_s \ge \sqrt{x} \ge p_1\cdots p_s$. If $p_1^{\nu_1}\cdots p_s^{\nu_s} > p_1\cdots p_s y_s$ with given $p_1 < \dots < p_s$, then we use the trivial estimate

$$\sum_{n\le x/\left(p_1^{\nu_1}\cdots p_s^{\nu_s}\right)} \alpha(n)\left|\tilde{f}(n)-A(x)\right|^a \ll 2^{O(a)}(\log x)^a \sum_{n\le x/\left(p_1^{\nu_1}\cdots p_s^{\nu_s}\right)} \alpha(n)$$

$$\ll \lambda_\alpha 2^{O(a)}(\log x)^a \left(\frac{x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\sigma_0}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}.$$

Thus, (3.7.4) is

$$\ll \frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}\lambda_\alpha 2^{O(k)}x^{\sigma_0}(\log x)^k.$$

Since $\alpha(p^\nu)=O((\lambda p^{\varrho_0+\sigma_0-1})^\nu)$ for all p^ν, we have

$$\sum_{\substack{p_1\cdots p_s y_s<p_1^{\nu_1}\cdots p_s^{\nu_s}\le x\\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}$$

$$\le 2^{O(l)}\sum_{\substack{p_1\cdots p_s y_s<p_1^{\nu_1}\cdots p_s^{\nu_s}\le x\\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}$$

$$\le \frac{2^{O(l)}}{(p_1\cdots p_s)^{1-\varrho_0}}\sum_{\substack{y_s<p_1^{\nu_1}\cdots p_s^{\nu_s}\le x_s\\ \nu_1,\ldots,\nu_s\ge 1}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\left(\log\frac{3x_s}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}.$$

It is not hard to see that the proof of (3.3.3) also gives

$$\sum_{z_1<p^\nu\le z_2}\left(\frac{\lambda}{p^{1-\varrho_0}}\right)^{\nu}\left(\log\frac{3z_2}{p^\nu}\right)^{\beta-1} \ll \frac{(\log(3z_2/z_1))^{\beta-1}}{z_1^{1-\varrho_0-\log_p\lambda}}$$

uniformly for all primes p and all $0 < z_1 \le z_2$. Thus, we have

$$\begin{aligned}
&\sum_{\substack{y_s < p_1^{\nu_1}\cdots p_s^{\nu_s} \le x_s \\ \nu_1,\dots,\nu_s \ge 1}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\left(\log\frac{3x_s}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1} \\
&\le 2^{O(l)}(\log x)^{\kappa l}\sum_{\substack{y_s < p_1^{\nu_1}\cdots p_s^{\nu_s} \le x_s \\ \nu_1,\dots,\nu_s \ge 1}}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\left(\log\frac{3x_s}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1} \\
&\le 2^{O(l)}(\log x)^{\kappa l}\sum_{\substack{p_2^{\nu_2}\cdots p_s^{\nu_s} \le x_s \\ \nu_2,\dots,\nu_s \ge 1}}\left(\frac{\lambda}{p_2^{\log_{p_1}\lambda}}\right)^{\nu_2}\cdots\left(\frac{\lambda}{p_s^{\log_{p_1}\lambda}}\right)^{\nu_s}\frac{(\log(3x_s/y_s))^{\beta-1}}{y_s^{1-\varrho_0-\log_{p_1}\lambda}} \\
&\le \frac{2^{O(l)}(\log x)^{(\kappa+1)m+\beta-2}}{x^{\delta_0(1-\delta_0)\lambda'/2}(p_1\cdots p_s)^{\delta_0^2\lambda'}} \le \frac{2^{O(l)}(\log x)^{\beta-1}}{x^{(1-\delta_0)\lambda'/5}(p_1\cdots p_s)^{\delta_0^2\lambda'}},
\end{aligned}$$

where the penultimate inequality follows from the previous line together with the observations that $p_i^{\log p_1 \lambda} > \lambda$ for all $2 \le i \le s$, that $x^{(1+\delta_0)/2} \ge (p_1\cdots p_s)^{1+\delta_0}$, and that

$$y_s^{1-\varrho_0-\log_{p_1}\lambda} \ge y_s^{\lambda'} = \left(x^{(1-\delta_0)/2}\cdot\frac{x^{(1+\delta_0)/2}}{p_1\cdots p_s}\right)^{\delta_0\lambda'} \ge x^{\delta_0(1-\delta_0)\lambda'/2}(p_1\cdots p_s)^{\delta_0^2\lambda'}.$$

It follows that

$$\begin{aligned}
&\sum_{\substack{p_1\cdots p_s y_s < p_1^{\nu_1}\cdots p_s^{\nu_s} \le x \\ \nu_1,\dots,\nu_s \ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1} \\
&\le \frac{2^{O(l)}}{(p_1\cdots p_s)^{1-\varrho_0+\delta_0^2\lambda'}}x^{-(1-\delta_0)\lambda'/5}(\log x)^{\beta-1},
\end{aligned}$$

from which we deduce that

$$\begin{aligned}
&\sum_{p_1<\ldots<p_s\leq\sqrt{x}}\ \sum_{\substack{p_1\cdots p_s y_s<p_1^{\nu_1}\cdots p_s^{\nu_s}\leq x\\ \nu_1,\ldots,\nu_s\geq 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s} \sum_{\substack{n\leq x\\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)\left|\tilde{f}(n)-A(x)\right|^k\\
&\leq 2^{O(m)}\lambda_\alpha x^{\sigma_0-(1-\delta_0)\lambda'/5}(\log x)^{k+\beta-1}\sum_{p_1<\ldots<p_s\leq\sqrt{x}}\frac{1}{(p_1\cdots p_s)^{1-\varrho_0+\delta_0^2\lambda'}}\\
&\leq \frac{1}{s!}\lambda_\alpha x^{\sigma_0-(1-\delta_0)\lambda'/6}(\log x)^{\beta-1}. \qquad (3.7.5)
\end{aligned}$$

On the other hand, if $p_1^{\nu_1}\cdots p_s^{\nu_s}\leq p_1\cdots p_s y_s$, then $x^{(1-\delta_0)/2}\leq x/(p_1^{\nu_1}\cdots p_s^{\nu_s})\leq x$. Thus, we can apply the asymptotic formulas for (3.7.1) with $\eta_0=(1-\delta_0)/2$ and $y=x/(p_1^{\nu_1}\cdots p_s^{\nu_s})$, in conjunction with the Cauchy–Schwarz inequality, to estimate the inner sum in (3.7.4). As a consequence, we have

$$\sum_{n\leq x/\left(p_1^{\nu_1}\cdots p_s^{\nu_s}\right)}\alpha(n)\left|\tilde{f}(n)-A(x)\right|^a \ll \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{\frac{a}{2}}\left(\frac{x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\sigma_0}\left(\log\frac{x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}.$$

Inserting this into (3.7.4) shows that the sum

$$\sum_{\substack{n\leq x\\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}}\alpha(n)\left|\tilde{f}(n)-A(x)\right|^k$$

is

$$\begin{aligned}
&\leq \frac{2^{O(m-k)}\lambda_\alpha C_m}{(\sqrt{m})^{m-k}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}\left(\sqrt{B(x)}+O\left(\frac{1}{\sqrt{m}}\sum_{i=1}^{s}|f(p_i)|\right)\right)^k x^{\sigma_0}(\log x)^{\beta-1}\\
&=\frac{2^{O(l)}\lambda_\alpha C_m}{m^{l/2}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}B(x)^{\frac{k}{2}}\left(1+O\left(\sqrt{\frac{m}{B(x)}}\right)\right)x^{\sigma_0}(\log x)^{\beta-1}\\
&\leq \frac{2^{O(l)}\lambda_\alpha C_m}{m^{l/2}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}B(x)^{\frac{k}{2}}x^{\sigma_0}(\log x)^{\beta-1}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le p_1\cdots p_s y_s\\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}\\
&\le 2^{O(l)}\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le p_1\cdots p_s y_s\\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\\
&\le \frac{2^{O(l)}}{(p_1\cdots p_s)^{1-\varrho_0}}\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le y_s\\ \nu_1,\ldots,\nu_s\ge 1}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}\\
&\le \frac{2^{O(l)}}{(p_1\cdots p_s)^{1-\varrho_0}}\prod_{i=1}^{s}\mathrm{Li}_{-\lceil\kappa l_i\rceil}\left(\lambda/p_i^{1-\varrho_0}\right),
\end{aligned}$$

where

$$\mathrm{Li}_{-\ell}(\zeta) := \sum_{n=1}^{\infty} n^{\ell}\zeta^{n}$$

is the polylogarithm function of order $-\ell$ and complex argument ζ with $|\zeta|< 1$, where $\ell \ge 0$ is any integer. For example, $\mathrm{Li}_0(\zeta) = \zeta/(1-\zeta)$ and $\mathrm{Li}_{-1}(\zeta) = \zeta/(1-\zeta)^2$. The function $\mathrm{Li}_{-\ell}(\zeta)$ can be expressed in terms of the Eulerian polynomial $A_\ell(\zeta)$:

$$\mathrm{Li}_{-\ell}(\zeta) = \frac{\zeta A_\ell(\zeta)}{(1-\zeta)^{\ell+1}},$$

where

$$A_\ell(\zeta) := \sum_{j=0}^{\ell}\left\langle {\ell \atop j}\right\rangle \zeta^{j}$$

is the ℓth Eulerian polynomial, and

$$\left\langle {\ell \atop j}\right\rangle := \sum_{a=0}^{j}(-1)^a\binom{\ell+1}{a}(j+1-a)^{\ell}$$

is the jth Eulerian number of size ℓ. Combinatorially, it is known that, for every

$\ell \geq 1$,

$$\left\langle \begin{matrix} \ell \\ j \end{matrix} \right\rangle = \#\{\tau \in S_\ell \colon \tau \text{ has exactly } j \text{ ascents}\},$$

where S_ℓ is the set of all permutations of $\{1, ..., \ell\}$. Using this combinatorial intepretation one finds that $A_\ell(1) = \#S_\ell = \ell!$. Since $l_1 + \cdots + l_s = l \leq m$, we have

$$\prod_{i=1}^{s} \mathrm{Li}_{-\lceil \kappa l_i \rceil} \left(\lambda / p_i^{1-\varrho_0}\right) \leq \frac{2^{O(l)} \lceil \kappa l_1 \rceil! \cdots \lceil \kappa l_s \rceil!}{(p_1 \cdots p_s)^{1-\varrho_0}} = \frac{2^{O(l)} \left(l_1^{l_1} \cdots l_s^{l_s}\right)^{\kappa}}{(p_1 \cdots p_s)^{1-\varrho_0}} \leq \frac{2^{O(l)} m^{\kappa l}}{(p_1 \cdots p_s)^{1-\varrho_0}},$$

by Stirling's formula. Hence, we obtain

$$\sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq p_1 \cdots p_s y_s \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s})}{p_1^{\sigma_0 \nu_1} \cdots p_s^{\sigma_0 \nu_s}} \leq \frac{2^{O(l)} m^{\kappa l}}{(p_1 \cdots p_s)^{2(1-\varrho_0)}}.$$

It follows that

$$\begin{aligned} &\sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq p_1 \cdots p_s y_s \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \sum_{\substack{n \leq x \\ p_1^{\nu_1}, \ldots, p_s^{\nu_s} \| n}} \alpha(n) \left| \tilde{f}(n) - A(x) \right|^k \\ &\leq \frac{2^{O(l)} \lambda_\alpha C_m m^{\kappa l}}{m^{l/2} (p_1 \cdots p_s)^{2(1-\varrho_0)}} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1}. \end{aligned}$$

Summing the above over $p_1 < ... < p_s \leq \sqrt{x}$, we arrive at

$$\begin{aligned} &\sum_{p_1 < \ldots < p_s \leq \sqrt{x}} \sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq p_1 \cdots p_s y_s \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \sum_{\substack{n \leq x \\ p_1^{\nu_1}, \ldots, p_s^{\nu_s} \| n}} \alpha(n) \left| \tilde{f}(n) - A(x) \right|^k \\ &\leq 2^{O(l)} \lambda_\alpha C_m m^{(\kappa - 1/2) l} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1} \sum_{p_1 < \ldots < p_s \leq \sqrt{x}} \frac{1}{(p_1 \cdots p_s)^{2(1-\varrho_0)}} \\ &\leq \frac{2^{O(l)}}{s!} \lambda_\alpha C_m m^{(\kappa - 1/2) l} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1}, \end{aligned}$$

since $\varrho_0 \in [0, 1/2)$. Combining this estimate with (3.7.5), we obtain

$$\sum_{p_1 < \ldots < p_s \le \sqrt{x}} \sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \le x \\ \nu_1, \ldots, \nu_s \ge 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \sum_{\substack{n \le x \\ p_1^{\nu_1}, \ldots, p_s^{\nu_s} \| n}} \alpha(n) \left| \tilde{f}(n) - A(x) \right|^k$$
$$\le \frac{2^{O(l)}}{s!} \lambda_\alpha C_m m^{(\kappa - 1/2)l} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1}.$$

Therefore, (3.7.3) is bounded above by

$$2^{O(l)} \lambda_\alpha C_m m^{(\kappa - 1/2)l} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1} \sum_{s \le l} \frac{1}{s!} \sum_{\substack{l_1 + \cdots + l_s = l \\ l_1, \ldots, l_s \in \mathbb{N}}} \binom{l}{l_1, \ldots, l_s}$$
$$\le 2^{O(l)} \lambda_\alpha C_m m^{(\kappa - 1/2)l} B(x)^{\frac{k}{2}} x^{\sigma_0} (\log x)^{\beta - 1} T_l(1)$$
$$\le 2^{O(l)} C_m m^{(\kappa + 1/2)l} B(x)^{\frac{k}{2}} S(x),$$

which allows us to conclude that

$$\sum_{k=0}^{m-1} \binom{m}{k} \sum_{n \le x} \alpha(n) \left(\tilde{f}(n) - A(x) \right)^k \left(f(n) - \tilde{f}(n) \right)^{m-k} \ll C_m m^{\kappa + \frac{3}{2}} B(x)^{\frac{m-1}{2}} S(x),$$

provided that in addition, we also have $1 \le m \ll B(x)^{1/(2\kappa+3)}$. Inserting the above estimate and the estimate for the term corresponding to $k = m$ into (3.7.2) completes the proof of Theorem 3.2.2.

Section 3.8

Proof of Theorem 3.2.3 (sketch)

Now we outline the proof of Theorem 3.2.3. The first step is to redefine $f_q(n)$ introduced in Section 3.4. Again, let us suppose that $A_0 \in (0, 1)$ and that $|f(p)| \le 1$ for all primes p. For every $q \in \mathbb{N}$ we define

$$\bar{F}(\sigma_0, q) := \prod_{p|q}(1 - F(\sigma_0, p)),$$
$$\widetilde{F}(\sigma_0, q) := \frac{\rho_g(q)}{\varphi(q)}\bar{F}(\sigma_0, q).$$

For each prime p we put

$$f_p(n) := \begin{cases} f(p)(1 - \widetilde{F}(\sigma_0, p)), & \text{if } p \mid n, \\ -f(p)\widetilde{F}(\sigma_0, p), & \text{otherwise.} \end{cases}$$

And as before, we set

$$f_q(n) := \prod_{p^\nu \| q} f_p(n)^\nu$$

for any $q \in \mathbb{N}$. In addition, let $c_g \in \mathbb{N}$ be the least positive integer such that $c_g g(x) \in \mathbb{Z}[x]$, and let $Q_0 > c_g|g(0)| \geq 1$ be such that (3.3.17) holds. Then for each $q \in \mathbb{N}$ with $P^-(q) > Q_0$ we have $\mathcal{Z}_q(g) \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ and $\rho_g(q) = \#\mathcal{Z}_q(g)$, where $\mathcal{Z}_q(g)$ denotes the zero locus of g in $\mathbb{Z}/q\mathbb{Z}$. In particular, we have $0 \leq \rho_g(q) \leq \varphi(q)$, which implies that $0 \leq \widetilde{F}(\sigma_0, q) \leq 1$ and that $|f_q(n)| \leq 1$ for all $n \in \mathbb{N}$.

Next, we need an analogue of Lemma 3.4.1. Let x be sufficiently large, and set $z := x^{\delta(x)/m} > Q_0$, so and $v := \log x / \log z = m/\delta(x)$. Then we have

$$\sum_{Q_0 < p \leq x} f(p)\widetilde{F}(\sigma_0, p) = A_{f,g}(x) + O(1)$$

by (3.1.1), (3.3.17), and the facts that ρ_g is bounded on prime powers and that

$\sum_p \psi_0(p) < \infty$. It is easily seen that

$$f(g(n)) - A_{f,g}(x) = \sum_{Q_0 < p \le z} f_p(g(n)) + \sum_{\substack{p>z \\ p|g(n)}} f(p) - \sum_{z<p\le x} f(p)\widetilde{F}(\sigma_0, p) + O(1).$$

Note that

$$\sum_{z<p\le x} f(p)\widetilde{F}(\sigma_0,p) = A_{f,g}(x) - A_{f,g}(z) + O(1) \ll \log\left(\frac{m}{\delta(x)} + 1\right).$$

Since $1 \le g(n) \ll n^{d_g}$ uniformly for all $n \in \mathbb{N}$, where $d_g := \deg g \ge 1$, we have

$$\sum_{\substack{p>z \\ p|g(n)}} f(p) \ll \frac{m}{\delta(x)}.$$

It follows that

$$\sum_{n\le x} \alpha(n)(f(g(n)) - A_{f,g}(x))^m = \sum_{n\le x} \alpha(n) \left(\sum_{Q_0<p\le z} f_p(g(n))\right)^m + O(E_g(x;m)), \quad (3.8.1)$$

where

$$E_g(x;m) := \sum_{k=0}^{m-1} \binom{m}{k} 2^{O(m-k)} \left(m\delta(x)^{-1}\right)^{m-k} \sum_{n\le x} \alpha(n) \left|\sum_{p\le z} f_p(g(n))\right|^k.$$

Now we turn to the estimation of

$$\sum_{n\le x} \alpha(n) \left(\sum_{Q_0<p\le z} f_p(g(n))\right)^m = \sum_{Q_0<p_1,\ldots,p_m\le z} \sum_{n\le x} \alpha(n) f_{p_1\cdots p_m}(g(n)). \quad (3.8.2)$$

Let $q \in \mathbb{N}$ with $\omega(q) \leq m$, $P^-(q) > Q_0$ and $P^+(q) \leq z$. Then we have

$$\sum_{n\leq x} \alpha(n) f_q(g(n)) = \sum_{ab|R_q} f_q(a)\mu(b) \sum_{\substack{n\leq x \\ ab|g(n)}} \alpha(n) = \sum_{ab|R_q} f_q(a)\mu(b) \sum_{c\in \mathcal{Z}_{ab}(g)} \sum_{\substack{n\leq x \\ n\equiv c \,(\mathrm{mod}\, ab)}} \alpha(n).$$

Recall that $\mathcal{Z}_q(g) \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$. Thus in place of Lemma 3.3.3, we need to input in our analysis the information about the distribution of values of $\alpha(n)$ with n restricted to reduced residue classes. By (3.2.1) and Lemma 3.3.2, the innermost sum is equal to

$$\begin{aligned}
&\frac{1}{\varphi(ab)} \sum_{\substack{n\leq x \\ \gcd(n,ab)=1}} \alpha(n) + O\left(\frac{S(x)}{\varphi(ab)(\log x)^{B_0}}\right) \\
&= \frac{1}{\varphi(ab)} \lambda_\alpha(ab) x^{\sigma_0} (\log x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}}\right)\right) + O\left(\frac{1}{\varphi(ab)} \lambda_\alpha x^{\sigma_0} (\log x)^{\beta-1-B_0}\right) \\
&= \frac{1}{\varphi(ab)} \lambda_\alpha \bar{F}(\sigma_0, ab) x^{\sigma_0} (\log x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}}\right)\right) + O\left(\frac{1}{\varphi(ab)} \lambda_\alpha x^{\sigma_0} (\log x)^{\beta-1-B_0}\right) \\
&= \frac{1}{\varphi(ab)} \lambda_\alpha x^{\sigma_0} (\log x)^{\beta-1} \left(\bar{F}(\sigma_0, ab) + O\left(\frac{1}{(\log x)^{A_1}}\right)\right),
\end{aligned}$$

where $A_1 := \min(A_0, B_0)$. It follows that

$$\sum_{n\leq x} \alpha(n) f_q(g(n)) = \lambda_\alpha \left(\widetilde{G}_1(\sigma_0, q) + O\left(\frac{\widetilde{G}_2(\sigma_0, q)}{(\log x)^{A_1}}\right)\right) x^{\sigma_0} (\log x)^{\beta-1},$$

where

$$\begin{aligned}
\widetilde{G}_1(\sigma_0, q) &:= \sum_{ab|R_q} f_q(a)\mu(b) \widetilde{F}(\sigma_0 . ab), \\
\widetilde{G}_2(\sigma_0, q) &:= \sum_{ab|R_q} \frac{\rho_g(ab)}{\varphi(ab)} |f_q(a)|.
\end{aligned}$$

It is clear that $\widetilde{G}_1$ and $\widetilde{G}_2$ are both multiplicative in q. Easy calculation shows that

$$\widetilde{G}_1(\sigma_0, p^\nu) = f(p)^\nu \widetilde{F}(\sigma_0, p) \left(1 - \widetilde{F}(\sigma_0, p)\right) \left((-1)^\nu \widetilde{F}(\sigma_0, p)^{\nu-1} + (1 - \widetilde{F}(\sigma_0, p))^{\nu-1}\right)$$

for any prime power p^ν. In particular, we have $\widetilde{G}_1(\sigma_0, p) = 0$, $|\widetilde{G}_1(\sigma_0, p^\nu)| \leq 1/4$, and $\widetilde{G}_1(\sigma_0, p^\nu) \geq 0$ when $2 \mid \nu$. Moreover, we have that

$$\widetilde{G}_1(\sigma_0, p^2) = f(p)^2 \widetilde{F}(\sigma_0, p)\left(1 - \widetilde{F}(\sigma_0, p)\right) = \rho_g(p)\frac{f(p)^2}{p} + O\left(\frac{F(\sigma_0, p)}{p} + \frac{1}{p^2}\right),$$

and that

$$|\widetilde{G}_1(\sigma_0, p^\nu)| \leq |f(p)|^\nu \widetilde{F}(\sigma_0, p) \leq \rho_g(p)\frac{f(p)^2}{\varphi(p)} = \rho_g(p)\frac{f(p)^2}{p} + O\left(\frac{1}{p^2}\right)$$

for all p^ν with $p > Q_0$ and $\nu \geq 2$. Thus, one may view ρ_g as the multiplicative weight instead in the estimates above. These observations allow us to conclude, by arguing as in Section 3.5, that the contribution to (3.8.2) from $\widetilde{G}_1$ is

$$\lambda_\alpha C_m B(x)^{\frac{m}{2}}\left(\chi_m\left(1 + O\left(\frac{m\log(m/\delta(x)+2)}{B_{f,g}(x)}\right)\right) + O\left(\frac{m^{\frac{3}{2}}}{\sqrt{B_{f,g}(x)}}\right)\right)x^{\sigma_0}(\log x)^{\beta-1},$$

while the contribution from $\widetilde{G}_2$ is $\ll \lambda_\alpha 2^{O(m)} x^{\sigma_0}(\log x)^{\beta-1-A_0}(\log\log x)^m$. Now that we have the estimate for (3.8.2), we can bound $E_g(x;m)$ as before by combining it with the Cauchy–Schwarz inequality. Hence, we have

$$E_g(x;m) \ll \frac{\lambda_\alpha C_m m^{\frac{3}{2}}}{\delta(x)\sqrt{B_{f,g}(x)}} x^{\sigma_0}(\log x)^{\beta-1}.$$

Carrying these estimates back in (3.8.1) completes the proof of Theorem 3.2.3.

Section 3.9

Proofs of Theorem 3.2.4 and Corollary 3.2.5 (sketch)

Now we outline the proof of Theorem 3.2.4, which borrows the ideas from the proofs of Theorem 3.2.1 and [15, Theorem 1] with proper modifications. Let $0 < \epsilon < \min(1, K)$, and take $z := x^{1/v}$ and

$$w := \begin{cases} x^{1/\log(v+2)}, & \text{if } \beta = 1, \\ x^{1/(\epsilon \log(v+2))}, & \text{if } \beta \neq 1, \end{cases}$$

where we recall that $v \asymp m$ when $\beta = 1$ and $v = (\log\log x)^{m(\vartheta_0+2)}$ when $\beta \neq 1$ as chosen in Section 3.6. Having made these choices, we have $\epsilon \log(v+2) \to \infty$ as $x \to \infty$ in the case $\beta \neq 1$. Let

$$\mathcal{P}_\epsilon^-(x) := \left\{p \leq x \colon |f(p)| \leq \epsilon\sqrt{B^*(x)}\right\},$$
$$\mathcal{P}_\epsilon^+(x) := \left\{p \leq x \colon \epsilon\sqrt{B^*(x)} < |f(p)| \leq K\sqrt{B^*(x)}\right\},$$
$$\mathcal{P}_\infty(x) := \left\{p \leq x \colon |f(p)| > K\sqrt{B^*(x)}\right\},$$

and put $\mathcal{P}_K(x) := \mathcal{P}_\epsilon^-(x) \cup \mathcal{P}_\epsilon^+(x)$. We consider the strongly additive function

$$f_\epsilon(n;x) := \sum_{\substack{p|n \\ p \in \mathcal{P}_\epsilon^-(x)}} f(p) + \epsilon_{\beta,1} \sum_{\substack{p|n \\ p \in \mathcal{P}_\epsilon^+(x) \cap (z,x]}} f(p) + \sum_{\substack{p|n \\ p \in \mathcal{P}_\infty(x)}} f(p),$$

where we recall that $\epsilon_{\beta,1}$ takes value 0 if $\beta = 1$ and 1 otherwise, and define

$$A_\epsilon(x) := \sum_{p\in\mathcal{P}_\epsilon^-(x)} \alpha(p)\frac{f(p)}{p^{\sigma_0}},$$
$$B_\epsilon(x) := \sum_{p\in\mathcal{P}_\epsilon^-(x)} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}}.$$

By hypothesis,

$$B(x) - B_\epsilon(x) = \sum_{\substack{p\le x\\ |f(p)|>\epsilon\sqrt{B^*(x)}}} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} = o(B^*(x)),$$

and so

$$|A_\epsilon(x) - A(x)| \le \frac{1}{\epsilon\sqrt{B^*(x)}} \sum_{\substack{p\le x\\ |f(p)|>\epsilon\sqrt{B^*(x)}}} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} = o\left(\epsilon^{-1}\sqrt{B^*(x)}\right).$$

We expect that the distribution of $f_\epsilon(n;x)$ is close to being Gaussian with mean $A(x)$ and variance $B(x)$ when x gets sufficiently large. In what follows, we shall restrict our attention to the case $\beta \ne 1$, since the opposite case $\beta = 1$ is not only similar but also easier. Looking back at the proof of Lemma 3.4.1, we find, for sufficiently large x, that

$$\sum_{p\in\mathcal{P}_\epsilon^-(x)\cap(Q_0,x]} f(p)F(\sigma_0,p) = A_\epsilon(x) + O\left(\epsilon\sqrt{B^*(x)}\right) = A(x) + O\left(\epsilon\sqrt{B^*(x)}\right),$$

so that

$$f_\epsilon(n;x) - A(x) = \sum_{p\in\mathcal{P}_\epsilon^-(x)\cap(Q_0,z]} f_p(n) + \sum_{\substack{p|n\\ p\in\mathcal{P}_K(x)\cap(z,w]}} f(p) + O\left(\epsilon\sqrt{B(x)}\right), \quad (3.9.1)$$

where we have used the hypothesis that $f(n) = o(\sqrt{B(x)})$ for all $n \le x$ whose prime factors p satisfy $|f(p)| > K\sqrt{B^*(x)}$. This leads to an analogue of Lemma 3.4.1 in which the second sum above plays the same role as $\omega(n; z, w)$. In analogy to Lemma 3.4.1, we deduce from (3.9.1) that for every fixed $m \in \mathbb{N}$, one has

$$\sum_{n \le y} \alpha(n)(f_\epsilon(n; x) - A(x))^m = \sum_{n \le y} \alpha(n) \left(\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) \right)^m + O\left(E_\epsilon(y, z, w; m)\right), \tag{3.9.2}$$

uniformly for all sufficiently large x and any $y \ge 1$, where

$$E_\epsilon(y, z, w; m) := \sum_{\substack{a+b+c=m \\ 0 \le a < m \\ b, c \ge 0}} \binom{m}{a, b, c} \left(\epsilon\sqrt{B(x)}\right)^c \sum_{n \le y} \alpha(n) \left| \sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) \right|^a \omega_f(n; z, w)^b,$$

and

$$\omega_f(n; z, w) := \sum_{\substack{p | n \\ p \in \mathcal{P}_K(x) \cap (z, w]}} |f(p)|.$$

To estimate the right-hand side of (3.9.2), one only needs to recycle the arguments used in the proof of Theorem 3.2.1 and make slight modifications. For instance, the estimation of

$$\sum_{n \le y} \alpha(n) \left(\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) \right)^m$$

is essentially the same as that of (3.4.1) given in Sections 3.4 and 3.5, except that we use the inequality $|f(p)| \le \epsilon\sqrt{B^*(x)}$ for $p \in \mathcal{P}_\epsilon^-(x)$ in place of the bound $f(p) = O(1)$ throughout the argument. This way, we see that

$$G(\sigma_0, p^2) = \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} + O\left(\epsilon^2 B^*(z)\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right)\right)$$

and that

$$|G(\sigma_0, p^\nu)| \leq \left(\epsilon\sqrt{B^*(z)}\right)^{\nu-2}\left(\alpha(p)\frac{f(p)^2}{p^{\sigma_0}} + O\left(\epsilon^2 B^*(z)\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right)\right)\right)$$

for all $p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]$ and $\nu \geq 2$. Using these two estimates in place of (3.5.1) and (3.5.2) and following the argument in Section 3.5, we obtain

$$\begin{aligned}\sum_{n\leq y}\alpha(n)\left(\sum_{p\in\mathcal{P}_\epsilon^-(x)\cap(Q_0,z]} f_p(n)\right)^m &= \lambda_\alpha\left(\mu_m + O\left(\frac{\epsilon\log v}{\log\log\log x}\right)\right)B(x)^{\frac{m}{2}}y^{\sigma_0}(\log y)^{\beta-1} \\ &= \lambda_\alpha(\mu_m + O(\epsilon))B(x)^{\frac{m}{2}}y^{\sigma_0}(\log y)^{\beta-1} \qquad (3.9.3)\end{aligned}$$

uniformly for $y \in [x^{\eta_0}, x]$, where $\eta_0 \in (0,1]$ is any given constant. On the other hand, the estimation of $E_\epsilon(y, z, w; m)$ reduces to that of

$$\sum_{n\leq y}\alpha(n)\left|\sum_{p\in\mathcal{P}_\epsilon^-(x)\cap(Q_0,z]} f_p(n)\right|^a \omega_f(n; z, w)^b.$$

The argument is essentially the same as that of (3.6.2) in the case $\beta \neq 1$ given in Section 3.6. The only difference is that we now make use of the estimates that $f(p) \leq K\sqrt{B^*(x)}$ for all $p \in \mathcal{P}_K(x)$ and that

$$\sum_{p\in\mathcal{P}_K(x)\cap(z,w]}\alpha(p)\frac{|f(p)|^\nu}{p^{\sigma_0}} \ll \epsilon B(x)^{\frac{\nu}{2}}$$

for all $\nu \geq 1$, in place of the estimates that $f(p) = O(1)$ and that

$$\sum_{z<p\leq w}\alpha(p)\frac{|f(p)|^\nu}{p^{\sigma_0}} = O(\log v) = O(\log\log\log x),$$

respectively. The second estimate can be easily seen by considering $p \in \mathcal{P}_\epsilon^-(x)$ and $p \in \mathcal{P}_\epsilon^+(x)$ separately. Indeed, one derives it by adding up the inequalities

$$\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (z,w]} \alpha(p) \frac{|f(p)|^\nu}{p^{\sigma_0}} \ll (\epsilon B^*(x))^{\frac{\nu}{2}} \sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (z,x]} \frac{\alpha(p)}{p^{\sigma_0}} \ll (\epsilon B^*(x))^{\frac{\nu}{2}} \log v \ll \epsilon B(x)^{\frac{\nu}{2}}$$

and

$$\begin{aligned} \sum_{p \in \mathcal{P}_\epsilon^+(x) \cap (z,w]} \alpha(p) \frac{|f(p)|^\nu}{p^{\sigma_0}} &\ll (B^*(x))^{\frac{\nu-1}{2}} \sum_{p \in \mathcal{P}_\epsilon^+(x) \cap (z,x]} \alpha(p) \frac{|f(p)|}{p^{\sigma_0}} \\ &\ll \epsilon^{-1} (B^*(x))^{\frac{\nu}{2}-1} \sum_{\substack{p \leq x \\ |f(p)| > \epsilon\sqrt{B^*(x)}}} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}} \\ &= o\left(\epsilon^{-1} B^*(x)^{\frac{\nu}{2}}\right) \ll \epsilon B(x)^{\frac{\nu}{2}}. \end{aligned}$$

One shows in this way that $E_\epsilon(y,z,w;m) = O(\epsilon \lambda_\alpha B(x)^{\frac{m}{2}} y^{\sigma_0} (\log y)^{\beta-1})$. Inserting this estimate and (3.9.3) in (3.9.2) and taking $y = x$ yields

$$S(x)^{-1} \sum_{n \leq x} \alpha(n) \left(f_\epsilon(n;x) - A(x)\right)^m = (\mu_m + O(\epsilon)) B(x)^{\frac{m}{2}}$$

for every fixed $m \in \mathbb{N}$ and all sufficiently large x, where the implied constant in the error term is independent of ϵ.

To complete the proof of Theorem 3.2.4 for the case $\beta \neq 1$, it is sufficient to show

$$S(x)^{-1} \sum_{n \leq x} \alpha(n) |f(n) - f_\epsilon(n;x)|^m = O\left(\epsilon B^*(x)^{\frac{m}{2}}\right) \tag{3.9.4}$$

for every given $\epsilon \in (0,1)$ and $m \in \mathbb{N}$, where the implicit constant in the error term is independent of ϵ. Since the case where m is odd follows from the case where m is even by Cauchy–Schwarz, we need only to consider the latter case. The proof of this case is largely the same as that of [15, Lemma 2], except for the slight complication in the possible case $\beta \in (0,1)$. When m is even, we have

$$S(x)^{-1}\sum_{n\le x}\alpha(n)|f(n)-f_\epsilon(n;x)|^m = S(x)^{-1}\sum_{n\le x}\alpha(n)\sum_{\substack{p_1,\dots,p_m|n\\ p_1,\dots,p_m\in\mathcal{P}_\epsilon^+(x)\cap[2,z]}} f(p_1)\cdots f(p_m),$$

which, after grouping terms according to the distinct primes among $p_1, \dots, p_m$, becomes

$$S(x)^{-1}\sum_{s\le m}\sum_{\substack{p_1<\dots<p_s\le z\\ p_1,\dots,p_s\in\mathcal{P}_\epsilon^+(x)}}\sum_{\substack{k_1+\dots+k_s=m\\ k_1,\dots,k_s\in\mathbb{N}}}\binom{m}{k_1,\dots,k_s} f(p_1)^{k_1}\cdots f(p_s)^{k_s}\sum_{\substack{n\le x\\ p_1\cdots p_s|n}}\alpha(n). \quad (3.9.5)$$

By (3.3.4) we have

$$\sum_{\substack{n\le x\\ p_1\cdots p_s|n}}\alpha(n) = \sum_{\substack{q\le x\\ R_q=p_1\cdots p_s}}\alpha(q)\sum_{\substack{n'\le x/q\\ \gcd(n',q)=1}}\alpha(n') \ll \lambda_\alpha x^{\sigma_0}\sum_{\substack{q\le x\\ R_q=p_1\cdots p_s}}\frac{\alpha(q)}{q^{\sigma_0}}\left(\log\frac{3x}{q}\right)^{\beta-1}.$$

Appealing to (3.3.3) we derive

$$\begin{aligned}\sum_{\substack{q\le x\\ R_q=p_1\cdots p_s}}\frac{\alpha(q)}{q^{\sigma_0}}\left(\log\frac{3x}{q}\right)^{\beta-1} &\ll (\log x)^{\beta-1}\sum_{\substack{q\le\sqrt{x}\\ R_q=p_1\cdots p_s}}\frac{\alpha(q)}{q^{\sigma_0}} + \sum_{\substack{\sqrt{x}<q\le x\\ R_q=p_1\cdots p_s}}\frac{\alpha(q)}{q^{\sigma_0}}\left(\log\frac{3x}{q}\right)^{\beta-1}\\ &\ll (\log x)^{\beta-1}\prod_{i=1}^{s}\sum_{\nu=1}^{\infty}\frac{\alpha(p_i^\nu)}{p_i^{\sigma_0\nu}} + \frac{(\log x)^{s+\beta-2}}{(\sqrt{x})^{1-\varrho_0}}\\ &= (\log x)^{\beta-1}\prod_{i=1}^{s}\left(\frac{\alpha(p_i)}{p_i^{\sigma_0}}+\psi_0(p_i)\right) + \frac{(\log x)^{s+\beta-2}}{(\sqrt{x})^{1-\varrho_0}}.\end{aligned}$$

These estimates together with (3.3.4) imply that (3.9.5) is $\ll \Sigma_1+\Sigma_2$, where

$$\Sigma_1 := \sum_{s\le m}\sum_{\substack{p_1<\dots<p_s\le z\\ p_1,\dots,p_s\in\mathcal{P}_\epsilon^+(x)}}\sum_{\substack{k_1+\dots+k_s=m\\ k_1,\dots,k_s\in\mathbb{N}}}\binom{m}{k_1,\dots,k_s}\left|f(p_1)^{k_1}\cdots f(p_s)^{k_s}\right|\prod_{i=1}^{s}\left(\frac{\alpha(p_i)}{p_i^{\sigma_0}}+\psi_0(p_i)\right),$$

$$\Sigma_2 := \frac{(\log x)^{m-1}}{(\sqrt{x})^{1-\varrho_0}}\sum_{s\le m}\sum_{\substack{p_1<\dots<p_s\le z\\ p_1,\dots,p_s\in\mathcal{P}_\epsilon^+(x)}}\sum_{\substack{k_1+\dots+k_s=m\\ k_1,\dots,k_s\in\mathbb{N}}}\binom{m}{k_1,\dots,k_s}\left|f(p_1)^{k_1}\cdots f(p_s)^{k_s}\right|.$$

Since $f(p) \leq K\sqrt{B^*(x)}$ for all $p \in \mathcal{P}_\epsilon^+(x)$, we have

$$\Sigma_2 \ll \frac{(\log x)^{m-1}}{(\sqrt{x})^{1-\varrho_0}}\pi(z)^m B^*(x)^{\frac{m}{2}} = o\left(B^*(x)^{\frac{m}{2}}\right) \ll \epsilon B^*(x)^{\frac{m}{2}}.$$

To bound Σ_1, we observe

$$\left|f(p_1)^{k_1}\cdots f(p_s)^{k_s}\right| \ll B^*(x)^{\frac{m-s}{2}}|f(p_1)\cdots f(p_s)|.$$

Thus, we have

$$\begin{aligned}\Sigma_1 &\leq \sum_{s\leq m} B^*(x)^{\frac{m-s}{2}}\frac{1}{s!}\left(\sum_{\substack{p\leq z\\ p\in\mathcal{P}_\epsilon^+(x)}}\left(\alpha(p)\frac{|f(p)|}{p^{\sigma_0}}+\psi_0(p)\right)\right)^s \sum_{\substack{k_1+\cdots+k_s=m\\ k_1,\ldots,k_s\in\mathbb{N}}}\binom{m}{k_1,\ldots,k_s}\\ &= \sum_{s\leq m} B^*(x)^{\frac{m-s}{2}}\frac{1}{s!}\left(o\left(\epsilon^{-1}\sqrt{B^*(x)}\right)\right)^s \sum_{\substack{k_1+\cdots+k_s=m\\ k_1,\ldots,k_s\in\mathbb{N}}}\binom{m}{k_1,\ldots,k_s} \ll \epsilon B^*(x)^{\frac{m}{2}}.\end{aligned}$$

Combining these estimates completes the proof of (3.9.4) in the case $\beta \neq 1$.

As we mentioned in Section 3.2, Corollary 3.2.5 is an immediate consequence of Theorem 3.2.4 when f is strongly additive. The transition to the general additive case is then accomplished by applying the following analogue of [54, Theorem B]. And this is the only place where we need to make use of characteristic functions.

Lemma 3.9.1. *Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function, and suppose that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0,1)$ and $r \in (0,1)$, such that $\alpha(n)$ satisfies the conditions* (i)–(iv). *Let $f\colon \mathbb{N} \to \mathbb{R}$ be an additive function, and denote by $\tilde{f}$ the strongly additive contraction of f. Suppose that $B(x) \to \infty$ as $x \to \infty$. Then $X_N(n) := (f(n) - A(N))/\sqrt{B(N)}$ possesses a limiting distribution function with respect to the natural probability measure induced by α if and only if $\widetilde{X}_N(n) := (\tilde{f}(n) - A(N))/\sqrt{B(N)}$ does, in which case they share the same limiting* *distribution function.*

Proof. As before, we shall assume $A_0 \in (0,1)$. For each $N \in \mathbb{N}$, the distribution functions of $X_N(n)$ and $\widetilde{X}_N(n)$ are given by

$$\Phi_N(V) = S(N)^{-1} \sum_{\substack{n \le N \\ X_N \le V}} \alpha(n),$$

$$\widetilde{\Phi}_N(V) = S(N)^{-1} \sum_{\substack{n \le N \\ \widetilde{X}_N \le V}} \alpha(n),$$

respectively. We have to show that $\Phi_N(V)$ converges weakly to a distribution function as $N \to \infty$ if and only if $\widetilde{\Phi}_N(V)$ does, in which case they converge weakly to the same distribution function. Note that the characteristic functions of $X_N(n)$ and $\widetilde{X}_N(n)$ are

$$\varphi_N(t) = S(N)^{-1} \sum_{n \le N} \alpha(n) e^{itX_N(n)},$$

$$\widetilde{\varphi}_N(t) = S(N)^{-1} \sum_{n \le N} \alpha(n) e^{it\widetilde{X}_N(n)},$$

respectively. By Lévy's continuity theorem [56, Theorem III.2.6], it suffices to show

$$\lim_{N \to \infty} (\varphi_N(t) - \widetilde{\varphi}_N(t)) = 0 \tag{3.9.6}$$

for any given $t \in \mathbb{R}$. To prove this, let us fix $t \in \mathbb{R}$ and let $\epsilon \in (0, 1/(2|t|+1))$ be arbitrary. Denote by $J_\epsilon(N)$ the greatest integer not exceeding $\sqrt{N}$ such that the inequality $|f(n)| \le \epsilon\sqrt{B(N)}$ holds for all $1 \le n \le J_\epsilon(N)$. Since $B(N) \nearrow \infty$ as

$N \to \infty$, we have $J_\epsilon(N) \nearrow \infty$ as $N \to \infty$. By (3.3.4) we have

$$\begin{aligned}
|\varphi_N(t) - \widetilde{\varphi}_N(t)| &\le S(N)^{-1} \sum_{n \le N} \alpha(n) \left| \exp\left(it \frac{f(n) - \tilde{f}(n)}{\sqrt{B(N)}} \right) - 1 \right| \\
&= S(N)^{-1} \sum_{\substack{a \le N \\ a \text{ squareful}}} \alpha(a) \left| \exp\left(it \frac{f(a) - f(R_a)}{\sqrt{B(N)}} \right) - 1 \right| \sum_{\substack{b \le N/a \\ b \text{ squarefree} \\ \gcd(b,a)=1}} \alpha(b) \\
&\ll S(N)^{-1} \lambda_\alpha N^{\sigma_0} \sum_{\substack{a \le N \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a} \right)^{\beta - 1} \left| \exp\left(it \frac{f(a) - f(R_a)}{\sqrt{B(N)}} \right) - 1 \right|.
\end{aligned}$$

From (3.1.1) and (3.1.3) it follows that

$$\sum_{\substack{a=1 \\ a \text{ squareful}}}^{\infty} \frac{\alpha(a)}{a^s} = \prod_p \left(1 + \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\nu s}} \right)$$

is absolutely convergent for $s \in \mathbb{C}$ with $\Re(s) > \max(\varrho_0, r) + \sigma_0 - 1$. Thus

$$c(\delta) := \sum_{\substack{a=1 \\ a \text{ squareful}}}^{\infty} \frac{\alpha(a)}{a^{\sigma_0 - \delta}} = \prod_p \left(1 + \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\nu(\sigma_0 - \delta)}} \right) < \infty$$

for any $\delta < 1 - \max(\varrho_0, r)$. Since

$$\left| it \frac{f(a) - f(R_a)}{\sqrt{B(N)}} \right| \le 2\epsilon |t| < 1$$

for all $a \le J_\epsilon(N)$, this implies

$$\sum_{\substack{a \le J_\epsilon(N) \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a} \right)^{\beta - 1} \left| \exp\left(it \frac{f(a) - f(R_a)}{\sqrt{B(N)}} \right) - 1 \right| \ll \epsilon |t| (\log N)^{\beta - 1}.$$

Now fix $0 < \delta < 1 - \max(\varrho_0, r)$. By partial summation we have

$$\sum_{\substack{a \le x \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} = c(0) - \int_x^\infty \frac{1}{t^\delta}\, d\left(\sum_{\substack{a \le t \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0 - \delta}}\right) = c(0) + o\left(x^{-\delta}\right)$$

when x is sufficiently large. It follows that

$$\begin{aligned}
&\sum_{\substack{J_\epsilon(N) < a \le N \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a}\right)^{\beta-1} \left|\exp\left(it \frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right| \\
&\le 2 \sum_{\substack{J_\epsilon(N) < a \le N \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a}\right)^{\beta-1} \\
&= 2 \int_{J_\epsilon(N)}^N \left(\log \frac{3N}{t}\right)^{\beta-1} d\left(\sum_{\substack{a \le t \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}}\right) \\
&= o\left(N^{-\delta}\right) + o\left((\log N)^{\beta-1} J_\epsilon(N)^{-\delta}\right) + o\left(\int_{J_\epsilon(N)}^N t^{-1-\delta} \left(\log \frac{3N}{t}\right)^{\beta-2} dt\right)
\end{aligned}$$

for sufficiently large N. By a change of variable we see that

$$\begin{aligned}
\int_{J_\epsilon(N)}^N t^{-1-\delta} \left(\log \frac{3N}{t}\right)^{\beta-2} dt &= (3N)^{-\delta} \int_{\log 3}^{\log(3N/J_\epsilon(N))} e^{\delta t} t^{\beta-2}\, dt \\
&\ll (3N)^{-\delta} \left(\frac{3N}{J_\epsilon(N)}\right)^\delta \left(\log \frac{3N}{J_\epsilon(N)}\right)^{\beta-2} \\
&\ll (\log N)^{\beta-2} J_\epsilon(N)^{-\delta}.
\end{aligned}$$

Hence, we have

$$\sum_{\substack{J_\epsilon(N) < a \le N \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a}\right)^{\beta-1} \left|\exp\left(it \frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right| = o\left((\log N)^{\beta-1} J_\epsilon(N)^{-\delta}\right).$$

for sufficiently large N. Gathering the estimates above, we obtain

$$\varphi_N(t) - \widetilde{\varphi}_N(t) \ll \epsilon|t| + o\left(J_\epsilon(N)^{-\delta}\right)$$

for sufficiently large N, where the implicit constants are independent of t, ϵ and N. From this estimate we infer that

$$\limsup_{N\to\infty} |\varphi_N(t) - \widetilde{\varphi}_N(t)| = O(\epsilon|t|),$$

where the implicit constant is independent of t and ϵ. Since $\epsilon \in (0, 1/(2|t|+1))$ is arbitrary, we obtain (3.9.6) as desired. □

Section 3.10

An Application to the Ramanujan τ-function

Let $\tau(n)$ be the Ramanujan τ-function. The goal of this section is to prove Theorem 1.2.1. In fact, we shall show that this result follows from Corollary 3.2.5 in combination with Lemma 3.9.1 and [19, Lemma 7] without difficulty. In comparison to Elliott's probabilistic approach, our approach enables us to get around some of the complications resulting from the analysis of $\tau(n)$.

To illustrate this, let $\alpha(n) = \tau(n)^2/n^{11}$, and define the additive function $f(n)$ by $f(p^\nu) = \log\sqrt{\alpha(p^\nu)}$ if $\alpha(p^\nu) \neq 0$ and $f(p^\nu) = 0$ otherwise, where p^ν is any prime power. It is easy to verify, using the facts about $\tau(n)$ discussed in Section 3.1, that $\alpha(n)$ satisfies conditions (i)–(iv) with any fixed $A_0 > 0$, $\beta = 1$, $\sigma_0 = 1$, $\vartheta_0 = 0$, and any fixed $\varrho_0 \in (0,1)$ and $r \in (1/2, 1)$. Moreover, we have $\alpha(n) \leq d(n)^2$ by Deligne's bound. To prove (1.2.7), it suffices to demonstrate that the limiting distribution of $(f(n) - A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by

α is the standard Gaussian distribution, where it is clear that $A(x) = A_\tau(x)$ and $B(x) = B_\tau(x)$.

Let us consider the strongly additive function $f_0(n)$ defined by $f_0(p) = \log\sqrt{\alpha(p)}$ if $p \notin E_0$ and $f_0(p) = 0$ otherwise, where $E_0 := \{p > 2\colon \alpha(p) \le \exp(-2\sqrt[3]{\log\log p})\}$. Denote by $A_0(x)$ and $B_0(x)$ the expected mean and variance of $f_0(n)$ weighted by $\alpha(n)$, respectively. It can be shown [19, Lemma 7] that $B(x) \asymp \log\log x$. Since the inequality $t|\log t| \le \sqrt{t}$ holds for all $t \in [0,1]$, we have

$$\sum_{\substack{p\le x\\ p\in E_0}} \alpha(p)\frac{|f(p)|}{p} \le \sum_{\substack{p\le x\\ p\in E_0}} \frac{\sqrt{\alpha(p)}}{p} \le \sum_{p>2}\frac{1}{p}\exp\left(-\sqrt[3]{\log\log p}\right) < \infty.$$

It follows that $A_0(x) = A(x) + O(1)$. A similar argument shows that $B_0(x) = B(x) + O(1) \asymp \log\log x$. Thus, $f_0(p) = O(B_0(p)^{1/3})$ for all p, which shows that $f_0(n)$ satisfies the hypotheses in Corollary 3.2.5. Hence, the limiting distribution of $(f_0(n) - A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by α is the standard Gaussian distribution.

To complete our argument, let $\widetilde{f}$ be the strongly additive contraction of f. Then $f_0(n) \ge \widetilde{f}(n)$ for all $n \in \mathbb{N}$. Moreover, Deligne's bound and the fact that $\tau(n) \in \mathbb{Z}$ for all $n \in \mathbb{N}$ imply that $-(11\log p)/2 \le f(p) \le \log 2$ whenever $\alpha(p) \ne 0$. Since

$$\sum_{\nu\ge1}\frac{\alpha(p^\nu)}{p^\nu} \le \frac{\alpha(p)}{p} + \sum_{\nu\ge2}\frac{(\nu+1)^2}{p^\nu} = \frac{\alpha(p)}{p} + O\left(\frac{1}{p^2}\right),$$

we have

$$\begin{aligned} S(x)^{-1}\sum_{n\le x}\alpha(n)\left(f_0(n)-\widetilde{f}(n)\right) &= S(x)^{-1}\sum_{\substack{p\le x\\ p\in E_0}}|f(p)|\sum_{\substack{n\le x\\ p|n}}\alpha(n)\\ &= S(x)^{-1}\sum_{\substack{p\le x\\ p\in E_0}}|f(p)|\sum_{\nu\ge 1}\alpha(p^\nu)\sum_{\substack{n'\le x/p^\nu\\ p\nmid n'}}\alpha(n')\\ &\ll S(x)^{-1}\lambda_\alpha x\sum_{\substack{p\le x\\ p\in E_0}}|f(p)|\sum_{\nu\ge 1}\frac{\alpha(p^\nu)}{p^\nu}\\ &\ll \sum_{\substack{p\le x\\ p\in E_0}}\alpha(p)\frac{|f(p)|}{p}+O\left(\sum_{p>2}\frac{\log p}{p^2}\right)\ll 1. \end{aligned}$$

This estimate is sufficient for us to conclude that the limiting distribution of $(\widetilde{f}(n)-A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by α is also the standard Gaussian distribution. By Lemma 3.9.1, the same is true for $(f(n)-A(x))/\sqrt{B(x)}$.

Remark 3.10.1. The above argument can be easily modified to yield similar results for the Fourier coefficients of elliptic holomorphic newforms of weight at least 2. The reader is referred to [20] for examples of such results.

Section 3.11

The Number of Prime Factors of $\varphi(n)$

Recall that for each $n\in\mathbb{N}$, $\Omega(n)$ denotes the number of prime factors of n, counting multiplicity, and that Euler's totient function $\varphi(n)$ may be defined explicitly by

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

We are interested in the distribution of $\Omega(\varphi(n))$ weighted by certain multiplicative functions. And the goal of this chapter is to prove Theorem 1.2.2. As promised, we shall actually prove the following more general result.

Theorem 3.11.1. *Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function, and suppose that there exist absolute constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0,1)$ and $r \in (0,1)$, such that $\alpha(n)$ satisfies the conditions* (i)–(iv). *Furthermore, suppose that $\alpha(p) \sim \beta p^{\sigma_0 - 1}$ for all but a subset E of primes p, where $\#(E \cap [2,x]) = o(x(\log\log x)^{2-\vartheta_0}/(\log x)^3)$ as $x \to \infty$. Then*

$$\lim_{x\to\infty} S(x)^{-1} \sum_{\substack{n\leq x \\ \Omega(\varphi(n))\leq \beta(\log\log x)^2/2+V\sqrt{\beta(\log\log x)^3/3}}} \alpha(n) = \Phi(V) \tag{3.11.1}$$

for every $V \in \mathbb{R}$.

Proof. Let us first determine the weighted mean $A(x)$ and variance $B(x)$ of the additive function $f(n) = \Omega(\varphi(n))$. The unweighted mean and variance of $\Omega(\varphi(n))$ are provided by Lemmas 2.3 and 2.4 from [25]:

$$\sum_{p\leq x} \frac{\Omega(p-1)}{p} = \frac{1}{2}(\log\log x)^2 + O(\log\log x), \tag{3.11.2}$$

$$\sum_{p\leq x} \frac{\Omega(p-1)^2}{p} = \frac{1}{3}(\log\log x)^3 + O\left((\log\log x)^2\right). \tag{3.11.3}$$

Since $\alpha(p) \sim \beta p^{\sigma_0 - 1}$ for all $p \notin E$, it follows from (3.1.4) and (3.11.2) that

$$\begin{aligned}
A(x) &= \sum_{p \le x} \alpha(p) \frac{\Omega(p-1)}{p^{\sigma_0}} \\
&= \sum_{\substack{p \le x \\ p \notin E}} \alpha(p) \frac{\Omega(p-1)}{p^{\sigma_0}} + O\left(\sum_{\substack{p \le x \\ p \in E}} \frac{\log p}{p} (\log\log p)^{\vartheta_0} \right) \\
&= \beta \sum_{p \le x} \frac{\Omega(p-1)}{p} + o\left(\sum_{p \le x} \frac{\Omega(p-1)}{p} \right) + O\left(\sum_{\substack{p \le x \\ p \in E}} \frac{\log p}{p} (\log\log p)^{\vartheta_0} \right) \\
&= (1 + o(1)) \frac{\beta}{2} (\log\log x)^2 + O\left(\sum_{\substack{2 < p \le x \\ p \in E}} \frac{\log p}{p} (\log\log p)^{\vartheta_0} \right).
\end{aligned}$$

Put $E(x) := \#(E \cap [2, x])$. By hypothesis, we have $E(x) = o(x(\log\log x)^{2-\vartheta_0}/(\log x)^3)$ as $x \to \infty$. It is easily seen by partial summation that

$$\begin{aligned}
\sum_{\substack{2 < p \le x \\ p \in E}} \frac{\log p}{p} (\log\log(p+1))^{\vartheta_0} &\ll E(x) \frac{\log x}{x} (\log\log x)^{\vartheta_0} + \int_{3^-}^{x} E(t) \frac{\log t}{t^2} (\log\log t)^{\vartheta_0} \, dt \\
&= O(1).
\end{aligned}$$

Hence, we obtain

$$A(x) = (1 + o(1)) \frac{\beta}{2} (\log\log x)^2. \tag{3.11.4}$$

Similarly, since

$$\begin{aligned}
\sum_{\substack{2 < p \le x \\ p \in E}} \frac{(\log p)^2}{p} (\log\log(p+1))^{\vartheta_0} &\ll E(x) \frac{(\log x)^2}{x} (\log\log x)^{\vartheta_0} + \int_{3^-}^{x} E(t) \frac{(\log t)^2}{t^2} (\log\log t)^{\vartheta_0} \, dt \\
&= o\left((\log\log x)^3 \right),
\end{aligned}$$

we have by (3.11.3) that

$$B(x) = \sum_{p \le x} \alpha(p) \frac{\Omega(p-1)^2}{p^{\sigma_0}} = (1 + o(1)) \frac{\beta}{3} (\log\log x)^3. \qquad (3.11.5)$$

Next, we estimate the tail of the weighted variance of $\Omega(\varphi(n))$ over the primes p for which $\Omega(p-1)$ are large. It is known [26, Corollary 1] that

$$\#\{n \le x \colon \Omega(n) \ge T\} \ll 2^{-T} T^4 x \log x \qquad (3.11.6)$$

uniformly for all $x \ge 3$ and $T \ge 1$. We have by (3.1.4)

$$\sum_{\substack{p \le x \\ \Omega(p-1) > T}} \alpha(p) \frac{\Omega(p-1)^2}{p^{\sigma_0}} \ll (\log\log x)^{\vartheta_0} \sum_{\substack{n \le x \\ \Omega(n) > T}} \frac{\Omega(n)^2}{n}$$

uniformly for all $x \ge 3$ and $T \ge 1$. The sum on the right-hand side can be easily shown to be $O(2^{-T} T^4 (\log x)^4)$ by using (3.11.6) and partial summation. The details were worked out in the proof of [25, Theorem 3.1]. It follows that

$$\sum_{\substack{p \le x \\ \Omega(p-1) > T}} \alpha(p) \frac{\Omega(p-1)^2}{p^{\sigma_0}} \ll 2^{-T} T^4 (\log x)^4 (\log\log x)^{\vartheta_0}.$$

Taking $T = 8 \log\log x$, we have

$$\sum_{\substack{p \le x \\ \Omega(p-1) > 8 \log\log x}} \alpha(p) \frac{\Omega(p-1)^2}{p^{\sigma_0}} \ll (\log x)^{-3/2} \qquad (3.11.7)$$

for all $x \ge 3$.

Although we are tempted to apply Corollary 3.2.5 directly to $\Omega(\varphi(n))$, it is not legitimate to do so, because the value of $\Omega(p-1)$ can be as large as $\log p$. To circumvent this problem, we consider instead the strongly additive function $f_0(n)$

defined by $f_0(p) = \Omega(p-1)$ if $\Omega(p-1) \le 9 \log\log 2p$ and $f_0(p) = 0$ otherwise. Denote by $A_0(x)$ and $B_0(x)$ the expected mean and variance of $f_0(n)$ weighted by $\alpha(n)$, respectively. In order to show that $A_0(x)$ and $B_0(x)$ are close to $A(x)$ and $B(x)$, we need only an upper bound of the correct magnitude for the unweighted variance of $\Omega(p-1)$. By [25, Lemmas 2.1, 2.2] we have

$$\sum_{p \le x} \Omega(p-1) = \frac{x \log\log x}{\log x} + O\left(\frac{x}{\log x}\right),$$
$$\sum_{p \le x} \Omega(p-1)^2 = \frac{x(\log\log x)^2}{\log x} + O\left(\frac{x \log\log x}{\log x}\right).$$

Now simple calculation shows

$$\frac{1}{\pi(x)} \sum_{p \le x} \left(\Omega(p-1) - \log\log x\right)^2 = O(\log\log x).$$

Hence, we have

$$\begin{aligned} &\frac{1}{\pi(x)} \sum_{p \le x} \left(\Omega(p-1) - \log\log 2p\right)^2 \\ &\ll \frac{1}{\pi(x)} \sum_{p \le \sqrt{x}} (\log p)^2 + \frac{1}{\pi(x)} \sum_{\sqrt{x} < p \le x} \left(\Omega(p-1) - \log\log x + O(1)\right)^2 \\ &\ll \log\log x, \end{aligned}$$

which is precisely what we need. (Halberstam [33, Theorem 3] showed that

$$\frac{1}{\pi(x)} \sum_{p \le x} \left(\omega(p-1) - \log\log x\right)^2 = (1 + o(1)) \log\log x.$$

His method may be adapted to yield the same asymptotic formula with $\Omega(p-1)$ in place of $\omega(p-1)$, but the upper bound that we just derived is sufficient for our purposes.) From this estimate we deduce at once

$$\sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1)^2 \ll \frac{x \log\log x}{\log x}.$$

By partial summation we obtain

$$\sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{p} \ll (\log\log x)^2, \tag{3.11.8}$$

$$\sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{p \log\log 2p} \ll \log\log x, \tag{3.11.9}$$

$$\sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{\log\log 2p} \ll \frac{x}{\log x}. \tag{3.11.10}$$

Thus, it follows by (3.11.8) that

$$\begin{aligned} B(x) - B_0(x) &= \sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \alpha(p) \frac{\Omega(p-1)^2}{p^{\sigma_0}} \\ &\ll \sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{p} + o\left((\log\log x)^3\right) \\ &= o\left((\log\log x)^3\right), \end{aligned}$$

where the error $o((\log\log x)^3)$ on the second line arises from the contribution from the primes in E. Analogously, we have by (3.11.9) that

$$\begin{aligned} A(x) - A_0(x) &= \sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \alpha(p) \frac{\Omega(p-1)}{p^{\sigma_0}} \\ &\ll \sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{p \log\log 2p} + O(1) \\ &\ll \log\log x. \end{aligned} \tag{3.11.11}$$

Applying Corollary 3.2.5 to $f_0(n)$, we hence conclude that the distribution of $f_0(n)$ is approximately Gaussian with mean $\beta(\log\log x)^2/2$ and variance $\beta(\log\log x)^3/3$.

According to Lemma 3.9.1, it suffices to show that the distribution of the additive contraction $\tilde{f}(n)$ of $f(n) = \Omega(\varphi(n))$ is approximately Gaussian with mean $\beta(\log\log x)^2/2$ and variance $\beta(\log\log x)^3/3$. To this end, we show that $\tilde{f}(n)$ and $f_0(n)$ are close on average. It is clear that

$$\tilde{f}(n) - f_0(n) = \sum_{\substack{p | n \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1). \tag{3.11.12}$$

Note that

$$\begin{aligned} & S(x)^{-1} \sum_{n \le x} \alpha(n) \sum_{\substack{p | n \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \\ &= S(x)^{-1} \sum_{\substack{p \le x \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \sum_{\substack{n \le x \\ p | n}} \alpha(n) \\ &= S(x)^{-1} \sum_{\substack{p^\nu \le x, \nu \ge 1 \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \alpha(p^\nu) \sum_{\substack{n' \le x/p^\nu \\ p \nmid n}} \alpha(n') \\ &\ll (\log x)^{1-\beta} \sum_{\substack{p^\nu \le x, \nu \ge 1 \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x}{p^\nu} \right)^{\beta - 1}. \end{aligned}$$

To proceed, we split the last sum above into sums over the ranges $p^\nu \le \sqrt{x}$, $\sqrt{x} < p^\nu \le x$ with $p \le (\sqrt{x})^{1-\varrho_0}$, and $\sqrt{x} < p^\nu \le x$ with $p > (\sqrt{x})^{1-\varrho_0}$. The first sum is

$$\begin{aligned}\sum_{\substack{p^\nu \le \sqrt{x}, \nu \ge 1 \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x}{p^\nu}\right)^{\beta-1} &\ll (\log x)^{\beta-1} \sum_{\substack{p \le \sqrt{x} \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \sum_{\nu \ge 1} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \\ &= (\log x)^{\beta-1} \sum_{\substack{p \le \sqrt{x} \\ \Omega(p-1) > 9 \log\log 2p}} \alpha(p) \frac{\Omega(p-1)}{p^{\sigma_0}} \\ &\quad + (\log x)^{\beta-1} \sum_p \log p \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \\ &\ll (\log x)^{\beta-1} \log\log x,\end{aligned}$$

by (3.1.3) and (3.11.11). The second sum is

$$\begin{aligned}&\sum_{\substack{\sqrt{x} < p^\nu \le x, \nu \ge 1 \\ p \le (\sqrt{x})^{1-\varrho_0} \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x}{p^\nu}\right)^{\beta-1} \\ &= \sum_{\substack{p \le (\sqrt{x})^{1-\varrho_0} \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \sum_{\log_p \sqrt{x} < \nu \le \log_p x} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x}{p^\nu}\right)^{\beta-1} \\ &\ll \frac{(\log x)^{\beta-1}}{(\sqrt{x})^{1-\varrho_0}} \sum_{\substack{p \le (\sqrt{x})^{1-\varrho_0} \\ \Omega(p-1) > 9 \log\log 2p}} \Omega(p-1) \\ &\ll \frac{(\log x)^{\beta-1}}{(\sqrt{x})^{1-\varrho_0}} \sum_{\substack{p \le (\sqrt{x})^{1-\varrho_0} \\ \Omega(p-1) > 9 \log\log 2p}} \frac{\Omega(p-1)^2}{\log\log 2p} \\ &\ll (\log x)^{\beta-2},\end{aligned}$$

by (3.3.3) and (3.11.10). Finally, by (3.1.4) and (3.11.7) (with $\alpha(p) = p^{\sigma_0 - 1}$ for all

p), we see that the third sum is

$$\begin{aligned}
&\sum_{\substack{\sqrt{x}<p^\nu\le x,\nu\ge 1\\ p>(\sqrt{x})^{1-\varrho_0}\\ \Omega(p-1)>9\log\log 2p}} \Omega(p-1)\frac{\alpha(p^\nu)}{p^{\sigma_0\nu}}\left(\log\frac{3x}{p^\nu}\right)^{\beta-1}\\
&\ll (\log x)^{\max(\beta-1,0)}(\log\log x)^{\vartheta_0} \sum_{\substack{p\le x\\ \Omega(p-1)>8\log\log x}} \frac{\Omega(p-1)}{p}\\
&\ll (\log x)^{\max(\beta-1,0)}(\log\log x)^{\vartheta_0-1} \sum_{\substack{p\le x\\ \Omega(p-1)>8\log\log x}} \frac{\Omega(p-1)^2}{p}\\
&\ll (\log\log x)^{\vartheta_0-1}(\log x)^{\max(\beta-1,0)-3/2}\\
&\ll (\log x)^{\beta-1}.
\end{aligned}$$

Combining these estimates, we obtain

$$S(x)^{-1}\sum_{n\le x}\alpha(n)\sum_{\substack{p\mid n\\ \Omega(p-1)>9\log\log 2p}}\Omega(p-1)\ll \log\log x,$$

which, together with (3.11.12), implies that

$$S(x)^{-1}\sum_{n\le x}\alpha(n)\left(\tilde{f}(n)-f_0(n)\right)\ll \log\log x = o\left((\log\log x)^{3/2}\right).$$

This allows us to conclude that just like $f_0(n)$, the distribution of $\tilde{f}(n)$ is also approximately Gaussian with mean $\beta(\log\log x)^2/2$ and variance $\beta(\log\log x)^3/3$. And so the same can be said about $f(n)$. □

Remark 3.11.1. It is possible to adapt the proof of [25, Theorem 3.2] to obtain an analogue of Corollary 3.11.1 for $\omega(\varphi(n))$. This essentially requires a weighted version of the Turán–Kubilius inequality. It is not hard to show, by modifying the proof of the classical Turán–Kubilius inequality, that if $f\colon \mathbb{N}\to\mathbb{C}$ is an additive function, and if $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a multiplicaitve function satisfying the conditions (i)–(iv), then

$$S(x)^{-1} \sum_{n \leq x} \left| f(n) - A^{\#}(x) \right|^2 \ll B^{\#}(x) \tag{3.11.13}$$

for all $x \geq 1$, where

$$\begin{aligned} A^{\#}(x) &:= \sum_{p^\nu \leq x} \alpha(p^\nu) \bar{F}(\sigma_0, p) \frac{f(p^\nu)}{p^{\sigma_0 \nu}} \left(1 - \frac{\log p^\nu}{\log 3x}\right)^{\beta - 1}, \\ B^{\#}(x) &:= \sum_{p^\nu \leq x} \alpha(p^\nu) \frac{|f(p^\nu)|^2}{p^{\sigma_0 \nu}} \left(1 - \frac{\log p^\nu}{\log 3x}\right)^{\beta - 1}. \end{aligned}$$

In the case $\beta \geq 1$, the factor $(1 - \log p^\nu / \log 3x)^{\beta - 1}$ in the expressions of $A^{\#}(x)$ and $B^{\#}(x)$ above may be removed. But in the case $\beta \in (0, 1)$, this factor is $\asymp (\log x)^{1-\beta}$ when p^ν is close to x, so some care needs to be taken in practice. To ensure satisfactory estimates, one may suppose further in this case that $\alpha(p^\nu)/p^{\sigma_0 \nu}$ decays suitably fast as p^ν grows, so that the tails of $A^{\#}(x)$ and $B^{\#}(x)$ only contribute negligible amounts. For instance, one may assume that $\alpha(n)$ satisfies the same conditions as in Theorem 3.2.2, in addition to the hypothesis that $\alpha(p) \sim \beta p^{\sigma_0 - 1}$ for all but a subset E of primes p, where $\#(E \cap [2, x]) = o(x (\log \log x)^{2 - \vartheta_0} / (\log x)^3)$ as $x \to \infty$. Then one may show, by arguing as in the proof of [25, Theorem 3.2] and employ Corollary 1.2.2 and (3.11.13), that (1.2.10) also holds for every $V \in \mathbb{R}$ with $\omega(\varphi(n))$ in place of $\Omega(\varphi(n))$. The interested reader can fill in the required details without much difficulty.

Section 3.12

Concluding Remarks

Although in the present work we only focused on the subclass $\mathcal{M}^*$ of multiplicative functions, it is also of interest to consider weight functions $\alpha(n)$ which satisfy certain

Landau–Selberg–Delange type conditions. Given more information about $\alpha(n)$ and its associated Dirichlet series $F(s) = \sum_{n=1}^{\infty} \alpha(n) n^{-s}$, better results are obtainable in some circumstances. Below we give a brief description of the method in the special case where $F(s)$ is close to an integral power of the Riemann zeta-function $\zeta(s)$.

For a complex number $s \in \mathbb{C}$, we write $\sigma = \Re(s)$ and $t = \Im(s)$. Let $\alpha \colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function whose Dirichlet series $F(s) = \sum_{n=1}^{\infty} \alpha(n) n^{-s}$ is absolutely convergent for $s \in \mathbb{C}$ with $\sigma > \sigma_0$, where $\sigma_0 > 0$ is an absolute constant. Suppose that there exist absolute constants $\beta \in \mathbb{N}$, $0 < \theta_0 < \sigma_0$, $B > 0$, and $0 < \delta < 1$, such that $H_\beta(s) := F(s)\zeta(s - \sigma_0 + 1)^{-\beta}$ has an analytic continuation in the half plane $\sigma \geq \theta_0$ with

$$\lim_{s \to \sigma_0} F(s)(s - \sigma_0)^\beta > 0,$$

and such that $|H_\beta(s)| \leq B(1 + |t|)^{1-\delta}$ for all $s \in \mathbb{C}$ with $\sigma \geq \theta_0$. It is clear that $F(s)$ has (absolute) abscissa of convergence σ_0. Adapting the argument used in the proof of [38, Lemma 2.1] or [56, Theorem II.5.2], one can show that there exists some constant $\epsilon_0 > 0$ such that

$$S(x) = \frac{1}{\sigma_0} \operatorname{Res}_{s=\sigma_0} \left(\frac{F(s) x^s}{s - \sigma_0 + 1} \right) - x^{\sigma_0} (\log x)^{\beta - 1} \sum_{k=1}^{\beta-1} \sum_{j=0}^{k-1} c_{j,k} \frac{\mu_j(\beta)}{(\log x)^k} + O\left(B x^\theta\right) \tag{3.12.1}$$

uniformly for all $x \geq 3$ and $\theta \in (\sigma_0 - \epsilon_0, \sigma_0)$, where

$$\mu_k(\beta) := \frac{1}{k!} \cdot \frac{d^k}{ds^k} \left(\frac{F(s)(s - \sigma_0)^\beta}{s - \sigma_0 + 1} \right) \Bigg|_{s=\sigma_0},$$
$$c_{j,k} := \frac{(-1)^{k-j}(\sigma_0 - 1)}{(\beta - k - 1)!\, \sigma_0^{k-j+1}},$$

and the implicit constant in the error term depends at most on $\beta, \sigma_0, \theta_0, \delta, \epsilon_0$. Notably, one gains an asymptotic for $S(x)$ with a power-saving error term uniformly in B, in contrast to what is provided by (3.3.4). Furthermore, suppose that there exists an absolute constant $\lambda > 0$ such that $\alpha(p^\nu) = O\left((\lambda p^{\sigma_0 - 1})^\nu\right)$ for all prime powers p^ν. Let

$$F(s,a) := \prod_{p|a} \left(1 - \left(\sum_{\nu=0}^{\infty} \alpha(p^\nu) p^{-\nu s} \right)^{-1} \right)$$

for $s \in \mathbb{C}$ with $\sigma \geq \theta_0$ and squarefree $a \in \mathbb{N}$. When $s = \sigma_0$, this definition coincides with the one introduced in Lemma 3.3.3. As in the proof of Lemma 3.3.3, it is not hard to show that

$$F(s,p) = \frac{\alpha(p)}{p^s} + O\left(\frac{\alpha(p)^2}{p^{2\sigma}} + \frac{1}{p^{2(\sigma - \sigma_0 + 1)}} \right) \tag{3.12.2}$$

for all $s \in \mathbb{C}$ with $\sigma \geq \theta_0$ and all sufficiently large p. In addition, we observe that

$$\begin{aligned} \sum_{\substack{n=1 \\ a|n}}^{\infty} \frac{\alpha(n)}{n^s} &= \left(\prod_{p\nmid a} \sum_{\nu=0}^{\infty} \alpha(p^\nu) p^{-\nu s} \right) \left(\prod_{p|a} \sum_{\nu=1}^{\infty} \alpha(p^\nu) p^{-\nu s} \right) \\ &= F(s) \left(\prod_{p|a} \sum_{\nu=0}^{\infty} \alpha(p^\nu) p^{-\nu s} \right)^{-1} \left(\prod_{p|a} \sum_{\nu=1}^{\infty} \alpha(p^\nu) p^{-\nu s} \right) = F(s)F(s,a) \end{aligned}$$

for $s \in \mathbb{C}$ with $\sigma > \sigma_0$ and squarefree $a \in \mathbb{N}$. Applying (3.12.1) to the above Dirichlet series expansion of $F(s)F(s,a)$ and using (3.12.2) to obtain upper bounds for $H_\beta(s)F(s,a)$ uniformly in $\sigma \geq \theta_0$, we see that there exist constants $\epsilon \in (0,1)$, $Q_0 \geq 2$, and $d_{j,k} \in \mathbb{R}$, where $0 \leq j < k < \beta$, such that

$$\begin{aligned} \sum_{\substack{n \leq x \\ a|n}} \alpha(n) = &\frac{\mu_0(\beta) F(\sigma_0, a)}{(\beta - 1)!\, \sigma_0} x^{\sigma_0} (\log x)^{\beta - 1} + x^{\sigma_0} (\log x)^{\beta-1} \sum_{k=1}^{\beta-1} \sum_{j=0}^{k} d_{j,k} \frac{F^{(j)}(\sigma_0, a)}{(\log x)^k} \\ &+ O\left(B 2^{O(\omega(a))} a^{\sigma_0 - 1} (x/a)^\theta \right) \end{aligned} \tag{3.12.3}$$

uniformly for all $x \geq 3$, $\theta \in (\sigma_0 - \epsilon, \sigma_0)$ and square-free $a \in \mathbb{N}$ with $P^-(a) > Q_0$, where $F^{(j)}(\sigma_0, a)$ is the jth order derivative of $F(s,a)$ with respect to s evaluated at

$s = \sigma_0$. Again, one may compare this result with Lemma 3.3.3.

Now, if $f\colon \mathbb{N} \to \mathbb{R}$ is a strongly additive function with $|f(p)| \leq M$ for all primes p, where $M > 0$ is an absolute constant, and if $0 < h_0 < (3/2)^{2/3}$ is fixed but arbitrary, then we obtain, by using (3.12.3) as a substitute for Lemma 3.3.3 and arguing as before with the adoption of the technique used in [38, Section 4.2], that

$$M(x;m) = C_m B(x)^{\frac{m}{2}} \left(\chi_m + O\left(\frac{m^{\frac{3}{2}}}{\sqrt{B(x)}} \right) \right)$$

uniformly for all sufficiently large x and all $1 \leq m \leq h_0(B(x)/M^2)^{1/3}$, provided that $B(x) \to \infty$ as $x \to \infty$. Analogously, let $f\colon \mathbb{N} \to \mathbb{R}$ is strongly additive such that $f(p) = O(\sqrt{B(p)})$ for all primes p, $B(x) \to \infty$ as $x \to \infty$, and

$$\sum_{\substack{p \leq x \\ |f(p)| > \epsilon\sqrt{B(x)}}} \alpha(p) \frac{f(p)^2}{p} = o(B(x))$$

for any given $\epsilon > 0$. Then $M(x;m) = (\mu_m + o(1))B(x)^{\frac{m}{2}}$ for every fixed $m \in \mathbb{N}$. These results supplement Theorems 3.2.1 and 3.2.4. It may be worth pointing out that in the proofs of these results one can simply take $z = x^{1/v}$ with v being a suitable constant multiple of m. We invite the reader to fill in the details.

One of the key ingredients in the proof of Theorem 3.2.1 is an asymptotic formula for

$$\sum_{\substack{n \leq x \\ d|n}} \alpha(n),$$

which is provided by Lemma 3.3.3. More generally, let $\mathcal{A}(x) = \{a_n\}_{n \leq x}$ be a non-decreasing sequence of positive integers, and suppose that

$$\mathcal{A}_{d,\alpha}(x) := \sum_{\substack{n \leq x \\ d|a_n}} \alpha(n) = \rho(d)S(x) + r_d(x) \tag{3.12.4}$$

for square-free integers $d \in \mathbb{N}$, where $\rho\colon \mathbb{N} \to [0,1]$ is a multiplicative function, and $r_d(x)$ is a remainder term which is expected to be small for all d or small on average over d. Here, $\rho(d)$ can be viewed as the density of the set $\{n \in \mathbb{N}\colon d \mid a_n\}$ with respect to the probability measure induced by α. In this sieve-theoretic setting one can derive, without much difficulty, an analogue of [30, Proposition 4]. It may be of interest to determine if such an analogue can be used to obtain general weighted Erdős–Kac theorems for various interesting sequences $\{a_n\}$ studied relatively recently, including $g(p_n)$, $\varphi(n)$, the Carmichael function $\lambda(n)$, and the aliquot sum $s(n) := \sigma(n) - n$, where $g \in \mathbb{Z}[x]$ is an irreducible polynomial, p_n is the nth prime, and $\lambda(n)$ denotes the exponent of $(\mathbb{Z}/n\mathbb{Z})^\times$ (see [33], [25, 23] and [46]). Besides, the same approach may also be adapted to prove results of weighted Erdős–Kac type for short intervals as well as in the function field setting. We will explore these and other related problems in future research.

Bibliography

[1] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy II*, Publ. Res. Inst. Math. Sci. **47** (2011), no. 1, 29–98. MR 2827723

[2] D. Berkane and P. Dusart, *On a constant related to the prime counting function*, Mediterr. J. Math. **13** (2016), no. 3, 929–938. MR 3513147

[3] P. Billingsley, *On the central limit theorem for the prime divisor functions*, Amer. Math. Monthly **76** (1969), 132–139. MR 0242222

[4] ———, *Probability and measure*, 3rd. ed., Wiley Ser. Probab. Stat., Wiley-Intersci. Publ., John Wiley & Sons, Inc., New York, 1995. MR 1324786

[5] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Arith. **156** (1986), no. 3–4, 203–251. MR 0834613

[6] S. Broadbent, H. Kadiri, A. Lumley, N. Ng, and K. Wilk, *Sharper bounds for the Chebyshev function $\theta(x)$*, Math. Comp. **90** (2021), no. 331, 2281–2315. MR 4280302

[7] A. A. Buchstab, *Asymptotic estimates of a general number-theoretic function (Russian, German summary)*, Mat. Sb. **44** (1937), 1239–1246.

[8] J. Büthe, *Estimating $\pi(x)$ and related functions under partial RH assumptions*, Math. Comp. **85** (2016), no. 301, 2483–2498. MR 3511289

[9] ———, *An analytic method for bounding $\psi(x)$*, Math. Comp. **87** (2018), no. 312, 1991–2009. MR 3787399

[10] N. G. de Bruijn, *On the number of uncancelled elements in the sieve of Eratosthenes*, Nederl. Akad. Wetensch., Proc. **53** (1950), 803–812. MR 0035785

[11] R. de la Bretèche and G. Tenenbaum, *Remarks on the Selberg–Delange method*, Acta Arith. **200** (2021), no. 4, 349–369. MR 4331908

[12] H. Delange, *Sur le nombre des diviseurs premiers de n*, C. R. Acad. Sci. (Paris) **237** (1953), 542–544. MR 0057282

[13] ———, *Sur un théorème d'Erdős et Kac*, Acad. Roy. Belg. Bull. Cl. Sci. **42** (1956), no. 5, 130–144. MR 0076800

[14] ———, *Sur le nombre des diviseurs premiers de n*, Acta Arith. **7** (1962), 191–215. MR 0136588

[15] H. Delange and H. Halberstam, *A note on additive functions*, Pacific J. Math. **7** (1957), 1551–1556. MR 0092820

[16] P. Deligne, *La conjecture de Weil. I.*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307. MR 0340258

[17] P. Dusart, *Explicit estimates for some functions over primes*, Ramanujan J. **45** (2018), no. 1, 227–251. MR 3745073

[18] D. Elboim and O. Gorodetsky, *Multiplicative arithmetic functions and the Ewens measure*, preprint (2022), to appear in Israel J. Math., `arXiv:1909.00601`.

[19] P. D. T. A. Elliott, *A central limit theorem for Ramanujan's tau function*, Ramanujan J. **29** (2012), no. 1–3, 145–161. MR 2994095

[20] ______, *Central limit theorems for classical cusp forms*, Ramanujan J. **36** (2015), no. 1–2, 81–98. MR 3296711

[21] ______, *Corrigendum: central limit theorems for classical cusp forms*, Ramanujan J. **36** (2015), no. 1–2, 99–102. MR 3296712

[22] P. Erdős, *On the integers having exactly k prime factors*, Ann. of Math. **49** (1948), no. 2, 53–66. MR 0023279

[23] P. Erdős, A. Granville, C. Pomerance, and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, Analytic Number Theory, Proc. Conf. in honor of Paul T. Bateman, B. C. Berndt et al. (eds.), Progr. Math., vol. 85, Birkhäuser, Boston, 1990. MR 1084181

[24] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. **62** (1940), 738–742. MR 0002374

[25] P. Erdős and C. Pomerance, *On the normal number of prime factors of $\varphi(n)$*, Rocky Mountain J. Math. **15** (1985), no. 2, 343–352. MR 0823246

[26] P. Erdős and A. Sárközy, *On the number of prime factors of integers*, Acta Sci. Math. (Szeged) **42** (1980), no. 3–4, 237–246. MR 0603312

[27] K. (S.) Fan, *An inequality for the distribution of numbers free of small prime factors*, Integers **22** (2022), no. A26, 12 pp. MR 4390731

[28] K. (S.) Fan and C. Pomerance, *An inequality related to the sieve of Eratosthenes*, J. Number Theory **254** (2024), 169–183.

[29] J. Friedlander and H. Iwaniec, *Opera de cribro*, Amer. Math. Soc. Colloq. Publ., vol. 57, Amer. Math. Soc., Providence, RI., 2010. MR 2647984

[30] A. Granville and K. Soundararajan, *Sieving and the Erdős–Kac theorem*, Equidistribution in Number Theory, An Introduction, A. Granville and Z. Rudnick (eds.), NATO Sci. Ser. II Math. Phys. Chem., vol. 237, Springer, Dordrecht, 2007. MR 2290492

[31] H. Halberstam, *On the distribution of additive number-theoretic functions*, J. London Math. Soc. **30** (1955), 43–53. MR 0066406

[32] ———, *On the distribution of additive number-theoretic functions (II)*, J. London Math. Soc. **31** (1956), 1–14. MR 0073626

[33] ———, *On the distribution of additive number-theoretic functions (III)*, J. London Math. Soc. **31** (1956), 14–27. MR 0073627

[34] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number* n, Quart. J. Pure. Appl. Math. **48** (1917), 76–92. MR 2280878

[35] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6th. ed., Revised by D. R. Heath-Brown and J. H. Silverman; With a forward by A. J. Wiles, Oxford Univ. Press, Oxford, 2008. MR 2445243

[36] W. B. Jurkat and H.-E. Richert, *An improvement of Selberg's sieve method. I*, Acta Arith. **11** (1965), 217–410. MR 0202680

[37] M. Kac, *Probability methods in some problems of analysis and number theory*, Bull. Amer. Math. Soc. **55** (1949), 641–665. MR 0031504

[38] R. Khan, M. B. Milinovich, and U. Subedi, *A weighted version of the Erdős–Kac theorem*, J. Number Theory **239** (2022), 1–20. MR 4434484

[39] Wm. J. LeVeque, *On the size of certain number-theoretic functions*, Trans. Amer. Math. Soc. **66** (1949), 440–463. MR 0030993

[40] J. D. Lichtman and C. Pomerance, *Explicit estimates for the distribution of numbers free of large prime factors*, J. Number Theory **183** (2018), 1–23. MR 3715225

[41] L. J. Mordell, *On Mr. Ramanujan's empirical expansions of modular functions*, Proc. Camb. Philos. Soc. **19** (1917), 17–124.

[42] C. J. Moreno, *The Hoheisel phenomenon for generalized Dirichlet series*, Proc. Amer. Math. Soc. **40** (1973), 47–51. MR 0327682

[43] M. J. Mossinghoff and T. S. Trudgian, *Nonnegative trigonometric polynomials and a zero-free region for the Riemann zeta-function*, J. Number Theory **157** (2015), 329–349. MR 3373245

[44] D. T. Nguyen, *Generalized divisor functions in arithmetic progressions: I*, J. Number Theory **227** (2021), 30–93. MR 4254563

[45] D. J. Platt, *Isolating some non-trivial zeros of zeta*, Math. Comp. **86** (2017), no. 307, 2449–2467. MR 3647966

[46] P. Pollack and L. Troupe, *Sums of proper divisors follow the Erdős–Kac law*, Proc. Amer. Math. Soc. **151** (2023), no. 3, 977–988. MR 4531632

[47] S. Ramanujan, *On certain arithmetical functions*, Trans. Camb. Philos. Soc. **22** (1916), no. 9, 159–184. MR 2280861

[48] O. Ramaŕe, *An explicit density estimate for Dirichlet L-series*, Math. Comp. **85** (2016), no. 297, 325–356. MR 3404452

[49] A. Rényi and P. Turán, *On a theorem of Erdös–Kac*, Acta Arith. **4** (1958), 71–84. MR 0096629

[50] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94. MR 0137689

[51] L. G. Sathe, *On a problem of Hardy on the distribution of integers having a given number of prime factors II*, J. Indian Math. Soc. (N.S.) **17** (1953), 83–141. MR 0058632

[52] L. Schoenfeld, *Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II*, Math. Comp. **30** (1976), no. 134, 337–360. MR 0457374

[53] A. Selberg, *Note on a paper by L. G. Sathe*, J. Indian Math. Soc. (N.S.) **18** (1954), 83–87. MR 0067143

[54] H. N. Shapiro, *Distribution functions of additive arithmetic functions*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 426–430. MR 0079609

[55] J. M. Song, *Sums of nonnegative multiplicative functions over integers without large prime factors I*, Acta Arith. **97** (2001), no. 4, 329–351. MR 1823551

[56] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd. ed., Grad. Stud. Math., vol. 163, Amer. Math. Soc., Providence, RI., 2015. MR 3363366

[57] ———, *Moyennes effectives de fonctions multiplicatives complexes*, Ramanujan J. **44** (2017), no. 3, 641–701. MR 3723446

[58] ———, *Correction to: moyennes effectives de fonctions multiplicatives complexes*, Ramanujan J. **53** (2020), no. 1, 243–244. MR 4148467

[59] P. Turán, *On a theorem of Hardy and Ramanujan*, Ramanujan J. **9** (1934), no. 4, 274–276. MR 1574877

[60] B. Wang, Z. Wei, P. Yan, and S. Yi, *Generalizations of the Erdős-Kac theorem and the prime number theorem*, to appear in Commun. Math. Stat., arXiv:2303.05803.