Université de Montréal

Anatomy of Smooth Integers

par

Marzieh Mehdizadeh

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en mathématiques

July 24, 2017

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée

Anatomy of Smooth Integers

présentée par

Marzieh Mehdizadeh

a été évaluée par un jury composé des personnes suivantes :

Matilde Lalin
(président-rapporteur)

Andrew Granville
(directeur de recherche)

Dimitris Koukoulopoulos
(co-directeur)

Pavel Winternitz
(membre du jury)

Greg Martin
(examinateur externe)

(représentant du doyen de la FES)

Thèse acceptée le:

Date d'acceptation

SOMMAIRE

Dans le premier chapitre de cette thèse, nous passons en revue les outils de la théorie analytique des nombres qui seront utiles pour la suite. Nous faisons aussi un survol des entiers $y-$friables, c'est-à-dire des entiers dont chaque facteur premier est plus petit ou égal à y.
Au deuxième chapitre, nous présenterons des problèmes classiques de la théorie des nombres probabiliste et donnerons un bref historique d'une classe de fonctions arithmétiques sur un espace probabilisé.
Le problème de Erdős sur la table de multiplication demande quel est le nombre d'entiers distincts apparaissant dans la table de multiplication $N \times N$. L'ordre de grandeur de cette quantité a été déterminé par Kevin Ford (2008). Dans le chapitre 3 de cette thèse, nous étudions le nombre d'ensembles $y-$friables de la table de multiplication $N \times N$. Plus concrètement, nous nous concentrons sur le changement du comportement de la fonction $A(x, y)$ par rapport au domaine de y, où $A(x, y)$ est une fonction qui compte le nombre d'entiers $y-$ friables distincts et inférieurs à x qui peuvent être représentés comme le produit de deux entiers $y-$ friables inférieurs à $\sqrt{x}$.
Dans le quatrième chapitre, nous prouvons un théorème de Erdős-Kac modifié pour l'ensemble des entiers $y-$ friables. Si $\omega(n)$ est le nombre de facteurs premiers distincts de n, nous prouvons que la distribution de $\omega(n)$ est gaussienne pour un certain domaine de y en utilisant la méthode des moments.

Mots clés: Théorie des nombres analytiques, théorie des nombres probabiliste, méthode des moments, entiers $y-$friables, problème d'Erdős sur la table de multiplication, théorie de Erdős-Kac.

SUMMARY

The object of the first chapter of this thesis is to review the materials and tools in analytic number theory which are used in following chapters. We also give a survey on the development concerning the number of $y-$smooth integers, which are integers free of prime factors greater than y.

In the second chapter, we shall give a brief history about a class of arithmetical functions on a probability space and we discuss on some well-known problems in probabilistic number theory.

We present two results in analytic and probabilistic number theory.

The Erdős multiplication table problem asks what is the number of distinct integers appearing in the $N \times N$ multiplication table. The order of magnitude of this quantity was determined by Kevin Ford (2008). In chapter 3 of this thesis, we study the number of $y-$smooth entries of the $N \times N$ multiplication. More concretely, we focus on the change of behaviour of the function $A(x,y)$ in different ranges of y, where $A(x,y)$ is a function that counts the number of distinct $y-$smooth integers less than x which can be represented as the product of two $y-$smooth integers less than $\sqrt{x}$.

In Chapter 4, we prove an Erdős-Kac type of theorem for the set of $y-$smooth integers. If $\omega(n)$ is the number of distinct prime factors of n, we prove that the distribution of $\omega(n)$ is Gaussian for a certain range of y using method of moments.

Keywords: Analytic number theory, probabilistic number theory, method of moments, $y-$smooth integers, Erdős multiplication table problem, Erdős-Kac theorem.

CONTENTS

REMERCIEMENTS

Firstly, I would like to express my sincere gratitude to my main supervisor Prof. Andrew Granville for the continuous support of my Ph.D study, for his patience, motivation, and immense knowledge. I really appreciate his guidance and encouragement in all the time of research and writing of this thesis. I could not have imagined completing my Ph.D without his persistent help.

Very special thanks to my co-supervisor Prof. Dimitris Koukoulopoulos for his insights and suggestions that greatly assisted my research.

Besides my supervisors, I would like to thank the rest of my thesis committee: Prof. Matilde Lalin, Prof. Greg Martin, Prof. Pavel Winternits and Prof. Maksym Radziwill for carefully reading this thesis, and for their corrections and comments.

I would like to thank my fellow doctoral students in Département de mathématiques - Université de Montréal, for their feedback, cooperation and of course friendship: Farzad Aryan, Mohammad Bardestani, Lyosha Klurman, Kevin Henriot, Francois Charette, Dimitri Dias and Crystel Bujold.

I also want to thank Prof. Adam Harper, Prof. Sary Drappeau and Prof. Bruno Martin. It was fantastic to have the opportunity to discuss with them about my research.

Last but not the least, I would like to thank my family for supporting me spiritually throughout writing this thesis and my life in general.

Chapter 1

INTRODUCTION AND PRELIMINARIES

1.1. Notation

Analytic number theory involves estimating various quantities which are not easy to calculate precisely. Here we present some notation that are used frequently for bounding (estimating) functions in this thesis.
We write $f(x) = O(g(x))$ or $f(x) \ll g(x)$ if there exists an absolute constant C such that

$$\Big| f(x) \Big| \leq Cg(x) \quad .$$

Here the inequality holds either for all x for which the functions are defined, or for all sufficiently large x (i.e. all x larger than some fixed constant), that will be clear in context.
The notation $f(x) \asymp g(x)$ (f is of order g), means that

$$f(x) \ll g(x) \quad \text{and} \quad g(x) \ll f(x).$$

If $g(x) \neq 0$, we write $f(x) = o(g(x))$ if

$$\frac{f(x)}{g(x)} \to 0 \quad \text{as} \quad x \to \infty.$$

Also, for $g(x) \neq 0$ we write $f(x) \sim g(x)$ as $x \to \infty$, if

$$\frac{f(x)}{g(x)} \to 1 \quad \text{as} \quad x \to \infty.$$

Finally, we will use the following notations:
$\mathbb{R}, \mathbb{N}, \mathbb{C}$ the set of real, natural and complex numbers respectively.
$\lfloor x \rfloor$ the greatest integer $\leq x$.

$\lceil x \rceil$ the smallest integer $\geq x$.
(a,b) the greatest common divisor of a and b.
$Re(s)$ the real part of $s \in \mathbb{C}$.

1.2. Prime number theorem

The Prime Number Theorem concerns estimating the number of prime integers up to x, namely

$$\pi(x) := \sum_{p \leq x} 1.$$

This theorem states

$$\pi(x) \sim \frac{x}{\log x} \quad \text{as} \quad x \to \infty.$$

This means that the probability that a random integer less than x is prime is about $\frac{1}{\log x}$.

Some well-known proofs of the prime number theorem are due to Atle Selberg and Paul Erdős (1949), and a simple proof is a result of Newman [27] (1980). The most common proofs are based on reformulating the problem in terms of better-behaved prime counting functions that have smoother behaviour than $\pi(x)$ and give an equivalent result. Here we define two of such functions:
The *first Chebyshev function* $\psi(x)$, defined by

$$\psi(x) := \sum_{\substack{p^k \leq x \\ p \text{ } prime}} \log p = \sum_{n \leq x} \Lambda(n).$$

where $\Lambda(n)$ is the Von Mangolt function, namely

$$\Lambda(n) := \begin{cases} \log p, & \text{if } n = p^k \text{ and } k \geq 1 \\ 0, & \text{otherwise.} \end{cases}$$

The *second Chebyshev function* is defined by

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

We are now in the position to state the Prime Number Theorem in three equivalent classical forms,

Theorem 1.2.1. *(Prime number theorem) There is a constant $c > 0$ such that*

$$\psi(x) = x + O\left(\frac{x}{\exp(c\sqrt{\log x})}\right),$$

$$\theta(x) = x + O\left(\frac{x}{\exp(c\sqrt{\log x})}\right),$$

and

$$\pi(x) = Li(x) + O\left(\frac{x}{\exp(c\sqrt{\log x})}\right),$$

for $x \geq 2$.

The function $Li(x)$ is the logarithmic integral

$$Li(x) := \int_2^x \frac{du}{\log u}.$$

Using integration by parts K times, one can arrive at

$$Li(x) = x\sum_{k=1}^{K} \frac{(k-1)!}{(\log x)^k} + O_K\left(\frac{x}{(\log x)^K}\right).$$

Thus, we can deduce that

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

1.2.1. Mertens' estimates

A second type of estimates below the level of the PNT (Prime Number Theorem) are estimates for certain weighted sums over primes. These estimates are very strong with small error terms, but they are not strong enough to imply the prime number theorem.

Theorem 1.2.2. *(Mertens' estimates) We have*

(1)

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1),$$

(2)

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

(3)

$$\sum_{p\le x}\frac{1}{p}=\log\log x+A+O\left(\frac{1}{\log x}\right),$$

and

(4)

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

where A is a constant and $\gamma = 0.5772\dots$ is the Euler's constant defined by

$$\gamma := \lim_{n\to\infty}\left(-\ln n+\sum_{k=1}^{n}\frac{1}{k}\right).$$

1.3. Euler products and Riemann's zeta function

Definition 1.3.1. *(Arithmetic and multiplicative functions) An arithmetic function is a function defined from $\mathbb{N}$ to $\mathbb{C}$, and a multiplicative function is an arithmetic function such that*

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

It is completely multiplicative if

$$f(mn)=f(m)f(n) \quad \forall m,n\in\mathbb{N}.$$

The series defined by

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

is called the *Dirichlet series* associated with the function f, where $f(n)$ is an arithmetic function and s is a complex variable denoted by $s=\sigma+it$.
The most famous Dirichlet series is the *Riemann zeta function* $\zeta(s)$, defined by

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

considered as a function on the complex variable s and continued with a simple pole at $s=1$, occupies a central role in analytic number theory. It also has a very important property that various Dirichlet series can be expressed in terms of it.
We have a representation of any Dirichlet series associated with a multiplicative function as an infinite product over primes, called the *Euler product*. If $F(s)$ is a Dirichlet series, the

Euler product of $F(s)$ is

$$\prod_p \left(1 + \sum_{m=1}^{\infty} \frac{f(p^m)}{p^{ms}}\right). \tag{1.3.1}$$

We have

Theorem 1.3.1. *(Euler product identity) Let f be a multiplicative function with the Dirichlet series F.*

$i)$: If $F(s)$ converges absolutely at some point s, then the infinite product (1.3.1) converges absolutely and is equal to $F(s)$.

$ii)$: The Dirichlet series $F(s)$ converges if and only if

$$\sum_{p^m} \frac{|f(p^m)|}{p^{ms}} < \infty.$$

A famous example of the Euler product is the *Riemann's zeta function* $\zeta(s)$, represented as follows

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(1 + \sum_{m=1}^{\infty} \frac{1}{p^{ms}}\right) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad (\sigma > 1).$$

1.4. Perron's Formula.

Here we introduce Perron's formula. This formula plays a fundamental role in proof of the Prime Number Theorem and estimating $y-$smooth integers to be defined later.

Let a_n be an arithmetic function, and

$$F(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

be a Dirichlet series and uniformly convergent for $Re(s) > \sigma$. Now we introduce the *normalised summatory function*

$$A(x) := \sum_{n<x} a_n + \frac{1}{2} a_x,$$

where $a_x = 0$ if x is not an integer. Perron's formula states

$$A(x) = \frac{1}{2\pi i} \int_{\kappa - i\infty}^{\kappa + i\infty} F(s) \frac{x^s}{s} ds, \tag{1.4.1}$$

where $x > 0$ and $\kappa > \max(0,\sigma)$).

The proof of Perron's formula relies on the following Laplace inversion formula, achieved by applying the residue theorem in complex analysis:

$$\frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{x^s}{s}ds=\begin{cases}1 & \text{if} \quad x\geq 1,\\ 1/2 & \text{if} \quad x=1,\\ 0 & \text{if} \quad 0<x<1.\end{cases} \tag{1.4.2}$$

A better version of Perron's formula used in applications gives an explicit bound for the contribution from the domain $|\tau|>T$ to the integral.

Theorem 1.4.1. *For $T\geq 1$, we have*

$$A(x)=\frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}F(s)\frac{x^s}{s}ds+O_k\left(x^\kappa\sum_{n=1}^{\infty}\frac{|a_n|}{n^\kappa\left(1+T\Big|\log(x/n)\Big|\right)}\right),$$

where $\kappa>\max(0,\sigma)$ and $x\geq 1$.

1.5. $y-$SMOOTH INTEGERS

1.5.1. Rankin's bound

An integer n is said to be $y-$smooth or $y-$friable if none of its prime factors are greater than y. There are many publications in recent decades about $y-$smooth numbers. We first define the set of $y-$smooth integers up to x as follows:

$$S(x,y):=\{1\leq n\leq x: P(n)\leq y\},$$

where $P(n)$ denotes the largest prime factor of n, with the convention that $P(1)=1$. We set

$$\Psi(x,y):=\Big|S(x,y)\Big|.$$

A simple bound for $\Psi(x,y)$ can be obtained by Rankin's method. The main idea of Rankin's

method is that for a multiplicative function $f(n)$, we have the following explicit upper bound for $f(n)$

$$\begin{aligned}\sum_{n\leq x}f(n)\leq\sum_{n\leq x}f(n)\left(\frac{x}{n}\right)^\sigma&\leq x^\sigma\sum_{n=1}^{\infty}\frac{f(n)}{n^\sigma}\\ &=x^\sigma\prod_p\left(1+\frac{f(p)}{p^\sigma}+\frac{f(p^2)}{p^{2\sigma}}+\dots\right).\end{aligned} \tag{1.5.1}$$

Now by the definition of $y-$smooth integers and using the Rankin's bound, we get

$$\Psi(x,y) = \sum_{\substack{n\le x \\ P(n)\le y}} 1 \le \sum_{\substack{n\ge 1 \\ P(n)\le y}} \left(\frac{x}{n}\right)^{\sigma} = x^{\sigma} \sum_{\substack{n\ge 1 \\ P(n)\le y}} \frac{1}{n^{\sigma}}, \tag{1.5.2}$$

where σ is a real positive number. The last sum can be bounded by

$$\sum_{\substack{n\ge 1 \\ P(n)\le y}} \frac{1}{n^{\sigma}} = \prod_{p\le y} \left(1 - \frac{1}{p^{\sigma}}\right)^{-1} \ll \exp\left\{\sum_{p\le y} \frac{1}{p^{\sigma}}\right\}, \tag{1.5.3}$$

since by the Taylor expansion of the logarithm, we have

$$\begin{aligned} \log \prod_{p\le y} \left(1 - \frac{1}{p^{\sigma}}\right)^{-1} &= -\sum_{p\le y} \log\left(1 - \frac{1}{p^{\sigma}}\right) \\ &= \sum_{p\le y} \sum_{m=1}^{\infty} \frac{1}{mp^{m\sigma}} = \sum_{p\le y} \frac{1}{p^{\sigma}} + O(1). \end{aligned}$$

Now by using the Taylor expansion of the exponential function and by the assumption that $\sigma \ge 1 - \frac{1}{\log y}$, we get

$$\exp\left\{\sum_{p\le y} \frac{1}{p^{\sigma}}\right\} = \exp\left\{\sum_{p\le y} \frac{1}{p}\left(\frac{1}{p^{\sigma-1}}\right)\right\} = \exp\left\{\sum_{p\le y} \frac{1}{p} + O\left((1-\sigma)\sum_{p\le y} \frac{\log p}{p}\right)\right\}.$$

Taking $\sigma = 1 - \frac{1}{\log y}$, gives

$$\exp\left\{\sum_{p\le y} \frac{1}{p^{\sigma}}\right\} \ll \log y. \tag{1.5.4}$$

Combining (1.5.3) with (1.5.4), one can arrive at the upper bound

$$\sum_{\substack{n\ge 1 \\ P(n)\le y}} \frac{1}{n^{\sigma}} \ll \log y. \tag{1.5.5}$$

Finally, substituting (1.5.5) into (1.5.2), gives that

$$\Psi(x,y) \ll xe^{-u} \log y.$$

where u is defined as

$$u := \frac{\log x}{\log y}.$$

By a more complex method, one can remove $\log y$ in the upper bound (see [31, Theorem 1, III.5]).

1.5.2. Sieve methods

There are various other methods that have been developed by several authors for evaluating $\Psi(x,y)$. Sieve methods deal with estimates for the number of elements in a finite set A that are not divisible by any prime p from some set $\mathcal{P}$ of primes. One heuristic estimate of $\Psi(x,y)$ can be obtained by the sieve estimate

$$\#\{n \leq x : (n,p) = 1 \quad \forall p \in \mathcal{P}\} \asymp x \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right), \tag{1.5.6}$$

which holds for any set of primes $\mathcal{P}$ in $[1,x^{1/2-\epsilon})$ (It can be found in [20].) So, by (1.5.6), one can predict an approximation of $\Psi(x,y)$ as follows

$$\Psi(x,y) \asymp x \prod_{y<p\leq x} \left(1 - \frac{1}{p}\right) = x \frac{\prod_{p\leq x}(1-p^{-1})}{\prod_{p\leq y}(1-p^{-1})}.$$

Now by Mertens' formula we get

$$\Psi(x,y) \asymp \frac{x}{u},$$

where $u := \frac{\log x}{\log y}$. But we will see that the order of magnitude of $\Psi(x,y)$ decreases exponentially in u. The reason for this difference is that (1.5.6) is true by independent assumptions which are not satisfied if $\mathcal{P} \cap [\sqrt{x},x] \neq \emptyset$. For example if $p \in [\sqrt{x}, x]$ divides n, then n is not divisible by any other prime in $[\sqrt{x}, x]$.

Here we find an estimate for $\Psi(x,y)$ where y takes large values compared to x. We use induction on $\lfloor u \rfloor = \lfloor \frac{\log x}{\log y} \rfloor$. If $\lfloor u \rfloor = 0$, then we can write

$$\Psi(x,y) = \lfloor x \rfloor - \# \{n \leq x : \exists p \in (y,x] \quad \text{such that} \quad p|n\}, \tag{1.5.7}$$

but if $u < 1$, then $y > x$, and trivially we have

$$\Psi(x,y) = \lfloor x \rfloor = x + O(1).$$

If $\lfloor u \rfloor = 1$, then we have $x^{1/2} < y \leq x$. Thus,

$$\Psi(x,y) = \lfloor x \rfloor - \sum_{y<p\le x} \#\{n \le x : p|n\} = \lfloor x \rfloor - \sum_{y<p\le x} \lfloor \frac{x}{p} \rfloor.$$

Thus, by using Mertens' estimate, we obtain

$$\Psi(x,y) = x(1-\log u) + O\left(\frac{x}{\log x}\right). \tag{1.5.8}$$

If $\lfloor u \rfloor = 2$, then $x^{1/3} < y \le x^{1/2}$, and we can write

$$\begin{aligned}\Psi(x,y) &= \Psi(x,\sqrt{x}) - \#\left\{n \le x : y < P(n) \le \sqrt{x}\right\} \\ &= \Psi(x,\sqrt{x}) - \sum_{y<p\le\sqrt{x}} \Psi(x/p,p).\end{aligned} \tag{1.5.9}$$

For all primes p in $(y,\sqrt{x}]$ we have $\sqrt{x/p} < p \le x/p$ as soon as $x^{1/3} < y \le x^{1/2}$. So by using the result in (1.5.8) and Mertens' estimate, we obtain

$$\Psi(x,y) = x(1-\log 2) - \sum_{y<p\le\sqrt{x}} \frac{x}{p}\left(1-\log\left(\frac{\log(x/p)}{\log p}\right)\right) + O\left(\frac{x}{\log x}\right). \tag{1.5.10}$$

Now by applying the Prime Number Theorem to the last sum, one can arrive at the following estimate:

$$\begin{aligned}\Psi(x,y) &= x(1-\log 2) + \int_y^{\sqrt{x}} \left(1-\log\left(\frac{\log x/t}{\log t}\right)\right)\frac{dt}{t\log t} + O\left(\frac{x}{\log x}\right), \\ &\sim x\rho(u),\end{aligned} \tag{1.5.11}$$

where

$$\rho(u) := 1 - \log u + \int_2^u \log(v-1)\frac{dv}{v} \qquad 2 \le u < 3.$$

This argument gives us the formula

$$\Psi(x,y) \sim x\rho(u) \qquad x^{\epsilon} < \ y \le x, \tag{1.5.12}$$

where ρ is defined by initial condition $\rho(u) = 1$ when $0 \le u \le 1$ and the general form

$$\rho(u) = \rho(k) - \int_k^u \rho(v-1)\frac{dv}{v} \quad (k \le u \le k+1). \tag{1.5.13}$$

The function $\rho(u)$ is a continuous function, called the Dickman function (1930), and satisfies the following *differential equation* obtained by differentiating (1.5.13)

$$u\rho^{'}(u) + \rho(u-1) = 0 \qquad (u \geq 1),$$

with the initial condition $\rho(u) = 1$ when $0 \leq u \leq 1$.

Theorem 1.5.1. *(Dickman's function properties) Dickman's function $\rho(u)$ has the following properties*

(1)

$$u\rho(u) = \int_{u-1}^{u} \rho(v)dv \quad (u \geq 1),$$

(2)

$$\rho(u) > 0 \quad (u > 0),$$

(3)

$$\rho^{'}(u) < 0 \quad (u > 1),$$

(4)

$$\rho(u) \leq 1/\Gamma(u+1) \quad (u \geq 0),$$

where

$$\Gamma(z) := \int_{0}^{\infty} x^{z-1}e^{-x}dx \quad z \in \mathbb{C}.$$

The above argument can be deduced from Buchstab's identity by applying induction on $\lfloor u \rfloor$.

Theorem 1.5.2. *(Buchstab's identity) For $x \geq 1$, $z \geq y > 0$, we have*

$$\Psi(x,y) = \Psi(x,z) - \sum_{y<p\leq z} \Psi(x/p,p).$$

The problem of finding an estimate for $\Psi(x,y)$ sounds more complicated and hence interesting when y is less than any small power of x. The question is; does (1.5.12) hold for smaller values of y? In (1951), de Bruijn [7] obtained a uniform estimate in the form of (1.5.12) in a wider range. The result states that

$$\Psi(x,y) = x\rho(u)\left\{1 + O\left(\frac{\log(u+1)}{\log y}\right)\right\}, \tag{1.5.14}$$

holds uniformly in the range

$$1 \le u \le (\log x)^{3/5-\epsilon}, \quad \text{that is,} \quad \exp\left\{(\log x)^{5/8+\epsilon}\right\} \le y \le x. \tag{1.5.15}$$

The range in (1.5.14) was significantly improved by Hildebrand (1986). Indeed, he obtained the largest range in which $\Psi(x,y) \sim x\rho(u)$ holds. Here we state the result in the form of a theorem from [23].

Theorem 1.5.3. *(Hildebrand) For any fixed $\epsilon > 0$, the relation (1.5.14) holds uniformly in the range*

$$y \ge 2, \qquad 1 \le u \le \exp\left\{(\log y)^{3/5-\epsilon}\right\}, \quad \textit{that is,} \quad y \ge \exp\left\{(\log\log x)^{5/3+\epsilon}\right\}.$$

The function $\Psi(x,y)$ behaves quite differently when y is small compared to x. In this case, $\Psi(x,y)$ is approximately equal to the volume of the $\pi(y)-$dimensional simplex defined by

$$t_i \ge 0 \quad (i = 1,..,\pi(y)) \quad \sum_{i=1}^{\pi(y)} t_i \log p_i \le \log x.$$

A change of variable gives us the volume of this $\pi(y)-$dimensional complex as

$$\#\left\{(\nu_1,...\nu_{\pi(y)}) : \nu_1 \ge 0,..., \nu_{\pi(y)} \ge 0 : \quad \sum_{i=1}^{\pi(y)} \nu_i \le 1\right\} \prod_{i=1}^{\pi(y)} \frac{\log x}{\log p_i} = \frac{1}{\pi(y)!} \prod_{i=1}^{\pi(y)} \frac{\log x}{\log p_i}.$$

In 1969, Ennola [11], gave the above idea and obtained the following sharp estimate of $\Psi(x,y)$ for small values of y compared to $\log x$.

Theorem 1.5.4. *(Ennola) Uniformly for $2 \le y \le \sqrt{\log x \log\log x}$, we have that*

$$\Psi(x,y) = \frac{1}{\pi(y)!} \prod_{p \le y} \left(\frac{\log x}{\log p}\right) \left\{1 + O\left(\frac{y^2}{\log x \log y}\right)\right\}.$$

1.5.3. Saddle point method

As we mentioned in previous section, the Rankin's method is a simple method for approximating $\Psi(x,y)$. This method is based on the inequality

$$\Psi(x,y) \leq \sum_{\substack{n\geq 1\\ P(n)\leq y}} (x/n)^{\sigma} = x^{\sigma} \prod_{p\leq y} \left(1-p^{-\sigma}\right)^{-1} \quad (\sigma > 0), \tag{1.5.16}$$

where σ will be chosen optimally such that the minimum on the right-hand side will be attained. With $s = \sigma + i\tau$, set

$$\zeta(s,y) = \prod_{p\leq y} \left(1-p^{-s}\right)^{-1}.$$

Also, we define

$$\phi(s,y) := \log\zeta(s,y) \quad (\sigma > 0),$$

and $\phi_k(s,y)$ to be the kth partial derivative of $\phi(s,y)$ with respect to s. Thus,

$$\phi_1(s,y) = -\sum_{p\leq y} \frac{\log p}{p^s - 1}.$$

By the Rankin's upper bound in (1.5.16), we have

$$\Psi(x,y) \leq \inf_{\sigma>0} x^{\sigma}\zeta(\sigma,y). \tag{1.5.17}$$

This infimum will be attained at the point $\sigma = \alpha = \alpha(x,y)$, which is the unique solution of the following equation:

$$\phi_1(\alpha,y) + \log x = 0, \quad \text{that is,} \quad \sum_{p\leq y} \frac{\log p}{p^{\alpha} - 1} = \log x. \tag{1.5.18}$$

In 1986, Hildebrand and Tenenbaum [25] used the saddle point method to develop the old result of de Bruijn for the range $x \geq y \geq 2$, and proved the following result.

Theorem 1.5.5. *(Hildebrand-Tenenbaum) Uniformly for $x \geq y \geq 2$, we have*

$$\Psi(x,y) = \frac{x^{\alpha}\zeta(\alpha,y)}{\alpha\sqrt{2\pi\phi_2(\alpha,y)}} \left\{1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right\}.$$

The strategy of their proof is based on Perron's integral (1.4.2).
For $n \leq x$, we have $x/n \geq 1$, and one can use the first case of Perron's formula in (1.4.2), and write

$$\begin{aligned}\Psi(x,y) &= \sum_{\substack{n\leq x\\ P(n)\leq y}} 1 = \sum_{\substack{n\geq 1\\ P(n)\leq y}} \int_{Re(s)=\alpha} \frac{(x/n)^s}{s} ds + O(1) \\ &= \int_{Re(s)=\alpha} \left(\sum_{\substack{n\geq 1\\ P(n)\leq y}} \frac{1}{n^s} \right) \frac{x^s}{s} ds + O(1) \\ &= \int_{Re(s)=\alpha} \zeta(s,y) \frac{x^s}{s} ds + O(1).\end{aligned} \tag{1.5.19}$$

Hildebrand and Tenenbaum showed that the main contribution to the integral above comes from a small neighbourhood around α, where α is the optimization point in Rankin's bound. They arrived at the following approximation

$$\Psi(x,y) = \frac{1}{2i\pi} \int_{\alpha-i/\log y}^{\alpha+i/\log y} \zeta(s,y) \frac{x^s}{s} ds + Error, \tag{1.5.20}$$

with

$$Error = \left(x^\alpha \zeta(\alpha,y) \left(Y(\epsilon)^{-1} + \exp\left\{ -cu(\log 2u)^{-2} \right\} \right) \right),$$

where c is a positive constant and $Y(\epsilon) = \exp\left\{ (\log y)^{3/2-\epsilon} \right\}$, for $0 \leq \epsilon \leq 1$. Also, developing the Taylor series on x^s/s and $\log\left(\frac{\zeta(s,y)}{\zeta(\alpha,y)}\right)$ gives

$$\begin{aligned}\frac{1}{2i\pi} \int_{\alpha-i/\log y}^{\alpha+i/\log y} \zeta(s,y) \frac{x^s}{s} ds &= \frac{x^\alpha}{2\alpha\pi} \zeta(\alpha,y) \int_{-1/\log y}^{1/\log y} e^{-t^2\phi_2/2} \left(1 - \frac{it}{\alpha} - \frac{it^3}{6}\phi_3 + Error \right) dt \\ &= \frac{x^\alpha}{\alpha} \zeta(\alpha,y) \frac{1}{\sqrt{2\pi\phi_2}} \left(1 + Error \right),\end{aligned} \tag{1.5.21}$$

since

$$\int_{-\infty}^{+\infty} e^{-t^2\phi_2/2} dt = \sqrt{\frac{2\pi}{\phi_2}}.$$

Thus, they could deduce the estimate in Theorem 1.5.5.

Also, by using the sharp form of the Prime Number Theorem for the summand in (1.5.18), one can arrive at the following estimate for α.

$$\alpha(x,y) = \frac{\log(1+y/\log x)}{\log y}\left\{1 + O\left(\frac{\log\log(y+1)}{\log y}\right)\right\}, \tag{1.5.22}$$

uniformly in $x \geq y \geq 2$.

There is another estimate for $\Psi(x,y)$ due to Saias in [28]. He used the saddle point method to prove the following result with a good error term. Let

$$\Lambda(x,y) := \begin{cases} x\int_0^\infty \rho(u - \frac{\log t}{\log y})\frac{d[t]}{t} & x \notin \mathbb{N} \\ (\frac{1}{2})\left(\Lambda(x-0,y) + \Lambda(x+0,y)\right) & x \in \mathbb{N}. \end{cases}$$

The estimate

$$\Psi(x,y) = \Lambda(x,y)\left(1 + O_\varepsilon\left(\exp\{-(\log y)^{\frac{3}{5}-\varepsilon}\}\right)\right) \tag{1.5.23}$$

holds in the range (1.5.15).

1.5.4. Local behaviour of $\Psi(x,y)$

Here we shall see that how the behaviour of $\Psi(x,y)$ changes when x is replaced by cx, where $1 \leq c \leq y$. Hildebrand and Tenenbaum deduced the following result from Theorem 1.5.5 by changing the path of integration and replacing the saddle point $\alpha(x,y)$ by $\alpha^{'}(x,y) := \alpha(cx,y)$.

Theorem 1.5.6. *(Hildebrand, Tenenbaum) We have, uniformly for $x \geq y \geq 2$ and $1 \leq c \leq y$,*

$$\Psi(cx,y) = \Psi(x,y)c^{\alpha(x,y)}\left(1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right).$$

In 2005, Tenenbaum and de la Breteche studied the local behaviour of $\Psi(x,y)$ by estimating the general form

$$\Psi_m(x,y) := \sum_{\substack{n\in S(x,y)\\(n,m)=1}} 1.$$

They used the saddle point method and established the following result that provides an estimate for $\Psi_m(x/d,y)$.

Theorem 1.5.7. *(Tenenbaum, de la Breteche)There exist constants b_1, b_2 and a function $b = b(x,y,d,m)$ satisfying $b_1 \leq b \leq b_2, (x \geq y \geq 2, d \geq 1, m \geq 1)$, such that, uniformly for $0 \leq \delta \leq 1/2$, $x \geq y \geq 2$, $p(m) \leq y$, $w(m) \ll \frac{\sqrt{y}}{(\log y)^\delta}$ and $1 \leq d \leq \frac{x}{y}$, we have*

$$\Psi_m\left(\frac{x}{d}, y\right) = \{1 + O(h_m)\}\left(1 - \frac{t^2}{u^2 + \bar{u}^2}\right)^{b\bar{u}} g_m(\alpha)\frac{\Psi(x,y)}{d^\alpha}, \tag{1.5.24}$$

where $\bar{u} := \min\{u, \frac{y}{\log y}\}$, $t := \frac{\log d}{\log y}$, $h_m \asymp h_1 \asymp \frac{1}{u_y} + \frac{t}{u}$, $u_y := \bar{u} + \frac{\log y}{\log(u+2)}$, and

$$g_m(\alpha) := \prod_{p|m}\left(1 - p^{-\alpha}\right) = \sum_{d|m}\frac{\mu(d)}{d^\alpha}.$$

In this thesis, the case $m = 1$ and $d \leq y$ will be of special interest, and by simplifying (1.5.24) (using the Taylor expansion of logarithmic and exponential functions), one can arrive at an estimate as follows.

Corollary 1.5.1. *Let $0 < \epsilon < 1$ and $1 \leq d \leq y$, then the estimate*

$$\Psi(x/d,y) = \left\{1 + O\left(\frac{1}{u_y} + \frac{t}{u} + \frac{t^2\bar{u}}{u^2 + \bar{u}^2}\right)\right\}\frac{\Psi(x,y)}{d^\alpha}$$

holds uniformly for $x \geq y \geq 2$.

1.5.5. Ultra-smooth integers

$y-$ultra-smooth or $y-$power-smooth integers are defined as integers whose canonical decomposition is free of prime powers exceeding y. For example, $720 = (2^4 3^2 5^1)$ is $5-$smooth but is not $5-$power-smooth (because there are several prime powers greater than 5, like $3^2 = 9 \nleq 5$. It is $16-$power-smooth since its greatest prime factor power is $2^4 = 16$.
Let

$$v_p = v_p(y) := \left\lfloor \frac{\log y}{\log p} \right\rfloor,$$

where p is a prime factor of n and $p \leq y$. So, v_p is the largest possible exponent of a prime factor of a $y-$ultra-smooth integer. We note that

$$U(x,y) := \{n \leq x : p^v || n \Rightarrow v \leq v_p\},$$

and set

$$\Upsilon(x,y) := |U(x,y)|.$$

Moreover, we define

$$N_y := e^{\psi(y)},$$

where $\psi(y) := \sum_{p\leq y} v_p \log p$, is the Chebychev's function. We can say that the integers counted in $\Upsilon(x,y)$ are divisors of N_y, therefore

$$\Upsilon(x,y) = \tau(N_y) = \prod_{p\leq y}(1+v_p).$$

Tenenbaum [32] (2015) proved that the number of $y-$ultra-smooth integers is close to the number of $y-$smooth integers when y is large compared to $\log x$.

Theorem 1.5.8. *(Tenenbaum) Let $\epsilon > 0$. For $x \geq y \geq 2$, we have*

$$\Upsilon(x,y) = \Psi(x,y)\left\{1 + O\left(\frac{u\log 2u}{\sqrt{y}\log y}\right)\right\} \quad x \geq y \geq (\log x)^{2+\epsilon}.$$

For smaller values of y the problem gets more complicated. However Tenenbaum obtained the following

$$\Upsilon(x,y) = x^{\beta} Z(\beta,y) G(\beta\sqrt{\sigma_2})\left\{1 + O\left(\frac{1}{u}\right)\right\}, \quad 2\log x < \psi(y) \ll (\log x)^3,$$

where

$$Z(\beta,y) := \prod_{p\leq y} \frac{1-p^{(1-\nu_{p+1})s}}{1-p^{-s}} \quad Re(s) > 0,$$

is the Dirichlet series associated to the counting function $\Upsilon(x,y)$, and $\beta = \beta(x,y)$ is the saddle point relevant the the Perron's integral for $\Upsilon(x,y)$ which a unique solution of the equation

$$\phi_1(\beta,y) := \frac{-Z'(s,y)}{Z(s,y)} = \log x,$$

and

$$\sigma_j := (-1)^{j-1}\frac{d^{j-1}\phi_1(s,y)}{d\sigma_{j-1}} \quad \text{and} \quad G(z) := e^{-z^2/2}\Phi(z) \quad (z \in \mathbb{R}),$$

where

$$\Phi(z) := \frac{1}{\sqrt{2\pi}}\int_z^{\infty} e^{-t^2/2} dt \quad (z \in \mathbb{R}).$$

The next corollary exhibits the behaviours of $\Upsilon(x,y)$ on either side of the threshold $y \approx (\log x)^2$.

Corollary 1.5.2. *(Tenenbaum) As $x \to \infty$, we have*

$$\Upsilon(x,y) \sim \Psi(x,y), \quad \textit{when} \quad \frac{y}{(\log x)^2} \to \infty,$$

and

$$\Upsilon(x,y) = o(\Psi(x,y)), \quad \textit{when} \quad \frac{y}{(\log x)^2} \to 0.$$

1.6. A history of the Erdős multiplication table problem

We know that $N \times N$ multiplication table with N^2 entries is a symmetric matrix such that most entries appear twice. Now the question is that how many distinct entries appear in this multiplication table? Let $A(N)$ denote the number of distinct entries in a $N \times N$ multiplication table. For example $A(5) = 14$ and $A(10) = 42$. Now the problem is to see how the behaviour of the function $A(N)$ changes with N^2.

In 1955 and 1960, Erdős studied this problem in two papers. In 1955, Erdős [14] could show that

$$\frac{A(N)}{N^2} \to 0 \quad as \quad N \to \infty,$$

and conjectured that $A(N)$ is approximately of the shape

$$A(N) = \frac{N^2}{(\log N)^c},$$

where c is a constant.

In 1960, Erdős [12] indicated the value of c as follows

$$c = 1 - \frac{1 + \log\log 2}{\log 2} = 0.08607\ldots,$$

and proved

$$A(N) = \frac{N^2}{(\log N)^{c+o(1)}} \quad \text{as} \quad N \to \infty.$$

In [21], Hall and Tenenbaum improved the upper bound of $A(N)$ as follows

$$A(N) \leq \frac{c_1 N^2}{(\log N)^c \sqrt{\log\log N}}, \tag{1.6.1}$$

In 2008, Ford showed that $A(N)$ is of the order of magnitude

$$\frac{N^2}{(\log N)^c(\log\log N)^{3/2}}.$$

Let $H(x,y,z)$ be the number of integers $n \le x$ having a divisor in the interval $(y,z]$, for all x, y and z. More formally

$$H(x,y,z) := \#\{n \le x : \exists\, d|n, \quad y < d \le z\}.$$

In 1935, Besicovitch [4] studied the quantities of $H(x,y,z)$ and proved that

$$\liminf_{y\to\infty} \varepsilon(y,2y) = 0,$$

where

$$\varepsilon(y,z) := \lim_{x\to\infty} \frac{H(x,y,z)}{x}.$$

Let $A(x)$ be the number of integers less than x that can be represented as the product of two integers less than $\sqrt{x}$. It is easy to see that the bounds for $A(x)$ are intimately connected with bounds for the function $H(x,y,2y)$ via this inequality:

$$H\left(\frac{x}{2},\frac{\sqrt{x}}{4},\frac{\sqrt{x}}{2}\right) \le A(x) \le \sum_{k\ge 0} H\left(\frac{x}{2^k},\frac{\sqrt{x}}{2^{k+1}},\frac{\sqrt{x}}{2^k}\right). \tag{1.6.2}$$

Hence, studying $A(x)$ boils down to understanding $H(x, y, 2y)$, which is slightly easier to study. Ford [16] proved

$$H(x,y,2y) \asymp \frac{x}{(\log x)^{\delta}(\log\log x)^{3/2}},$$

and from (1.6.2) he subsequently deduced the following estimate

$$A(x) \asymp \frac{x}{(\log x)^{\delta}(\log\log x)^{3/2}},$$

But we still do not know the asymptotic estimate for $A(x)$. It is worth mentioning that Koukoulopoulos in his Ph.D thesis [26] extended the multiplication table problem to the higher dimensional table.

Now we pose a question about the multiplication table for $y-$smooth integers; For a large real number x, how many distinct $y-$smooth integers up to x can be written as the product of two $y-$smooth integers less than $\sqrt{x}$?

The main subject of this thesis (in Section 3) is to study the behaviour of the following function in different ranges of y

$$A(x,y) := \Big|S(\sqrt{x},y) \cdot S(\sqrt{x},y)\Big|.$$

We define $H(x,y;z,2z)$ as the number of all $y-$smooth integers having at least one divisor in the given interval $(z,2z]$. By a simple argument one can arrive at the following inequalities,

$$H\left(\frac{x}{2},y;\frac{\sqrt{x}}{4},\frac{\sqrt{x}}{2}\right) \leq A(x,y) \leq \sum_{k\geq 0} H\left(\frac{x}{2^k},y;\frac{\sqrt{x}}{2^{k+1}},\frac{\sqrt{x}}{2^k}\right),$$

where $x \geq y > 2$. One would be tempted to estimate $A(x,y)$ by obtaining an upper and a lower bounds for $H(x,y;z,2z)$, but our main goal in Section 3 is to understand the behaviour of the function $A(x,y)$ directly instead of considering $H(x,y;z,2z)$.

It is good to explain a connection between estimating $A(x,y)$ and sum-product problem in additive combinatorics. For any non-empty subset A of integers, the *sum-set* and *product-set* of A are defined as

$$A \cdot A = \{a_1a_2 : a_i \in A\}, \qquad A + A = \{a_1 + a_2 : a_i \in A\}.$$

A famous conjecture of Erdős and Szemerédi states that the sum-set and product-set of a finite set of integers cannot both be small, more formally

$$\max\{|A \cdot A|, |A + A|\} \gg_\varepsilon |A|^{2-\varepsilon}.$$

With this connection to the sum-product problem, Banks and Covert [3], by invoking combinatorial tools, have considered the behaviour of $A(x^2,y) = |S(x,y) \cdot S(x,y)|$ in different ranges of y, particularly for the cases when y is relatively small or large.

For small values of y compared to $\log x$, they state the following result

Theorem 1.6.1. *(Banks, Covert) Suppose that $y \geq 2$ and $y = o(\log x)$. Then*

$$\Big|S(x,y) \cdot S(x,y)\Big| = \Psi(x,y)^{1+o(1)}.$$

For large values of y, they could show that the value of $\Big|S(x,y).S(x,y)\Big|$ is large

Theorem 1.6.2. *(Banks,Covert) Let $y/(\log x) \to \infty$. Then*

$$\left|S(x,y) \cdot S(x,y)\right| = \Psi(x,y)^{2+o(1)}.$$

For the values of y near $\log x$ they prove an estimate as follows

Theorem 1.6.3. *(Banks, Covert) Suppose that $y = \kappa \log x$, where $\kappa > 0$ is fixed. Then*

$$\left|S(x,y) \cdot S(x,y)\right| = \Psi(x,y)^{\alpha_\kappa + o(1)},$$

where

$$\alpha_\kappa = \frac{2\log(1+\kappa/2) + \kappa \log(1+2/\kappa)}{\log(1+\kappa) + \kappa\log(1+1/\kappa)}.$$

Chapter 2

PROBABILISTIC NUMBER THEORY

2.1. Additive functions in a probability space

There are various fields of application in probabilistic number theory. One of these fields is the theory of finding the distribution of arithmetic functions. Probabilistic methods study the *normal order* of an arithmetic function, and are based on considering the arithmetic function as a random variable and excluding set of integers with zero density and studying the more normal behaviour of the function elsewhere.

In the classical and more recent research the distribution of values of a function in number theory is reduced to consideration of the sum

$$A_N = \frac{1}{N} \sum_{n=1}^{N} f(n),$$

which is the mean value of the function $f(n)$ on the sample set $\{1,2,...,N\}$, and obtaining an approximation for it in terms of a function of N. However the values of the function $f(n)$ oscillate around the mean value within a very wide range.

In this section, we will discuss the Turan-Kubilius inequality which is a helpful tool to prove the results about the normal order of an arithmetic function. Also, we introduce the method of moments in probability. This is a well-known method used by many number theorists. We first need some definitions.

Definition 2.1.1. *(Additive function) The arithmetic function $f(n)$ is additive if, for co-prime integers u and v, we have*

$$f(uv) = f(u) + f(v).$$

The arithmetic function is said to be totally additive if, for any integers u and v,

$$f(uv) = f(u) + f(v),$$

By the definition above, we have the following representation for an additive function:

$$f(n) = \sum_{p^r || n} f(p^r). \tag{2.1.1}$$

Definition 2.1.2. *(Strongly additive function) An additive function $f(n)$ is called strongly additive if*

$$f(p^r) = f(p),$$

where r is any positive integer and p is prime.

Therefore (2.1.1) reduces to

$$f(n) = \sum_{p|n} f(p)$$

for a strongly additive function, where p is prime.

Examples: The most common example for additive function is the function $\omega(n)$, defined as the number of distinct prime factors of n. The function $\omega(n)$ is also a strongly additive function. An example of a totally additive function is the function $\Omega(n)$, defined as the number of prime factors of n, counted multiplicity.

Definition 2.1.3. *An arithmetic function f has normal order g if g is an arithmetic function such that, for any $\epsilon > 0$, we have*

$$\Big|f(n) - g(n)\Big| \leq \epsilon \Big|g(n)\Big|,$$

on a set of integers $n \in \mathbb{N}$ of density 1. In other words we say

$$f(n) = (1 + o(1))\, g(n) \quad \textit{almost everywhere.}$$

Definition 2.1.4. *(Distribution of a function) The arithmetic function $f(n)$ is said to have distribution $F(x)$ if*

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{\substack{n\leq N\\ f(n)<x}} 1 = F(x). \tag{2.1.2}$$

We consider the probability space $S_N = (\Omega_N, \mathcal{F}_N, P_N)$, where $\Omega_N = \{1,2,...,N\}$, $\mathcal{F}_N$ is the set of all subsets of Ω_N, and P_N is the uniform measure on $\mathcal{F}_N$, then the arithmetic function $f(n)$ restricted to Ω_N is a random variable on S_N. Using this assumption we can say that (2.1.2) is equivalent to

$$\lim_{N\to\infty} P_N(f(n) < x) = F(x).$$

If f is a random variable and an arithmetic function on the sample space $\{1,2,\dots,N\}$, the expectation and the variance of f with respect to the discrete uniform measure on $\{1,2,\dots,N\}$, are defined by

$$\mathbb{E}_N[f] := \frac{1}{N} \sum_{1\leq n\leq N} f(n),$$

and

$$\mathbb{V}_N[f] := \frac{1}{N} \sum_{1\leq n\leq N} \{f(n) - \mathbb{E}_N(f)\}^2.$$

Moreover, the *kth* moment of the random variable $f(n)$ is defined as

$$\mathbb{E}_N[f^k] := \frac{1}{N} \sum_{1\leq n\leq N} f^k(n).$$

Theorem 2.1.1. *(Chebyshev's inequality) Let X be a real-valued random variable with expectation $\mathbb{E}(X)$ and variance $\mathbb{V}(X)$. Then, for any $a \in \mathbb{R}$ we have*

$$P\left(\left|X - \mathbb{E}(X)\right| \geq a\right) \leq \frac{\mathbb{V}(X)}{a^2}.$$

2.2. Turan-Kubilius inequality for the variance

The *Chebyshev's inequality* provides a good tool to study the normal behaviour of an additive function. The first result of this kind was given by Hardy and Ramanujan (1917) and prove by Turan(1934). The theorem states that for almost all n up to N, the values of $\omega(n)$ and $\Omega(n)$ are asymptotically $\log\log N$. In fact they proved that

$$\mathbb{E}_N(\omega(n)) = \mathbb{E}_N(\Omega(n)) \sim \log\log N,$$

and

$$\mathbb{V}_N(\omega(n)) = \mathbb{V}_N(\Omega(n)) \sim \log\log N,$$

By Chebyshev's inequality, one can arrive at

$$P_N\left(\Big|\omega(n) - \log\log N\Big| \geq (\log\log N)^{3/4}\right) \leq \frac{1}{\sqrt{\log\log N}}.$$

Turan and Kubilius (1956) gave the same result for a wider class of additive functions.

Theorem 2.2.1. *(Turan-Kubilius inequality) Let f be an complex valued additive function and $N \in \mathbb{N}$. Then*

$$\mathbb{E}_N\left(\Big|f - \mathbb{E}_N f\Big|^2\right) \ll \sum_{p^v \leq N} \frac{\Big|f(p^v)\Big|^2}{p^v}. \tag{2.2.1}$$

Setting

$$A_N = \sum_{p^v \leq N} \frac{f(p^v)}{p}, \quad B_N^2 = \sum_{p^v \leq N} \frac{f^2(p^v)}{p},$$

by Chebyshev's inequality and using (2.2.1), we get

$$P_N\left(\Big|f(n) - A_N\Big| \geq \epsilon(N) B_N\right) \leq \frac{1}{\epsilon^2(N)}. \tag{2.2.2}$$

Now if $B_N = o(A_N)$, then we can select $\epsilon(N) \to \infty$, for example $\epsilon(N) = \frac{A_N}{(B_N)^{1/2}}$ which by (2.2.2) gives that $f(n) \sim A_N$ for almost all $n \leq N$.

In 1982, the Turan-Kubilius inequality on the sample space $y-$smooth integers was proved by Alladi [1] in a wide range of y. Later on, in 1993, Xuan [33] showed the Turan-Kubilius inequality for a wider range of y. Finally, in 2005, Tenenbaum and de la Breteche [8] gave another proof by using the saddle point method and proved the inequality for the whole range of y. Posing the quantities

$$A_f(x,y) := \sum_{p^v \in S(x,y)} \frac{f(p^v)}{p^{v\alpha}} \left(1 - \frac{1}{p^v}\right), \quad B_f^2(x,y) := \sum_{p^v \in S(x,y)} \frac{|f(p^v)|^2}{p^{v\alpha}} \left(1 - \frac{1}{p^v}\right),$$

and

$$V_f(x,y) := \frac{1}{\Psi(x,y)} \sum_{n \in S(x,y)} \left|f(n) - A_f(x,y)\right|^2,$$

they showed the following,

Theorem 2.2.2. *(Turan-Kubilius inequality for smooth integers) There exists an absolute constant C such that the inequality*

$$V_f(x,y) \le C B_f^2(x,y)$$

holds for every additive function f, for all x and y such that $x \ge y \ge 2$.

In particular, they proved this inequality for the function $\omega_t(n)$, defined as the number of all distinct prime factors of n which are less than t, where $x \ge y \ge t \ge 2$. By defining

$$M(t) := \sum_{p \le t} \frac{1}{p^\alpha},$$

they showed that

$$A_{\omega_t}(x,y) = B_f^2(x,y) = M(t) + O(1),$$

and proved the following

Corollary 2.2.1. *(Tenenbaum, de la Breteche) Uniformly for $x \ge y \ge t \ge 2$ and $h > 0$, we have*

$$\sum_{n \in S(x,y)} \left|\omega_t(n) - M(t)\right|^2 \ll \Psi(x,y) M(t). \tag{2.2.3}$$

By combining the above with Chebyshev's inequality, we have

$$\sum_{\substack{n \in S(x,y) \\ |\omega_t(n) - M(t)| > h\sqrt{M(t)}}} 1 \ll \frac{\Psi(x,y)}{h^2}. \tag{2.2.4}$$

2.3. Central limit theorem and the method of moments

Here we introduce a function of great importance in the development of probabilistic number theory, which leads to the *central limit theorem* that we shall see later.

Definition 2.3.1. *(Standard normal distribution) The function defined by*

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

is called the normal density function; its integral

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$

is the standard normal distribution function.

Now we are ready to state the central limit theorem. The central limit theorem is a special case of the law of large numbers in probability theory.

Theorem 2.3.1. *(Central Limit Thorem) If $\{X_k\}$ be a sequence of mutually independent random variables with the same probability distribution (i.i.d), suppose that $\mu = \mathbb{E}[X_k]$ and $\sigma^2 = Var(X_k)$ exist and let $S_N = X_1 + ... + X_N$. Then for every real fixed β,*

$$\mathcal{P}_N \left\{ \frac{S_N - N\mu}{\sigma\sqrt{N}} < \beta \right\} \to \Phi(\beta).$$

Theorem 2.3.2. *(Lindeberg-Feller theorem) If $X_1, X_2, \ldots, X_n$ are independent, uniformly bounded random variables with mean 0 and finite variances σ_i, and if $\sum_{i=1}^{n} \sigma_i^2$ diverges, then the distribution of $\frac{\sum_{i=1}^{n} X_i}{(\sum_{i=1}^{n} \sigma_i^2)^{1/2}}$ converges to the normal distribution function.*

Now we introduce a useful sufficient condition for a distribution to be determined by its moments (details can be found in [10]). In fact, we will see that if the moments of our real-valued random variables are very close to those of the standard normal, then the distribution of our random variables is close to the normal distribution.

Theorem 2.3.3. *(General method of moments) Suppose that $\int x^k dF_n(x)$ (kth moment of the random variable associated with F_n) has a limit μ_k for each k, and*

$$\lim_{k\to\infty} \sup \mu_{2k}^{1/2k}/2k < \infty.$$

Then F_n converges weakly (converges in distribution) to the unique distribution with these moments.

In the following lemma we will see that the *kth* moments of the normal distribution do not grow rapidly as k increases.

Lemma 2.3.1. *(Normal moments) For $k \in \mathbb{N}$, let*

$$m_k := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^k e^{-x^2/2} dx$$

be the kth moment of the normal distribution. Then if k is odd, then $m_k = 0$, and if k is even, then

$$m_k = \frac{k!}{2^{k/2}(k/2)!}.$$

Corollary 2.3.1. *(Method of moments for normal distribution) Let $\{X_n\}$ be a sequence of real valued random variables, and suppose that for each $k \in \mathbb{N}$, we have*

$$\mathbb{E}(X_n^k) \to m_k \quad as \quad n \to \infty,$$

where m_k are the moments of the standard normal distribution. Then we have convergence in distribution

$$P(X_n \leq x) \to \Phi(x) \quad as \quad n \to \infty.$$

2.3.1. Erdős-Kac theorem

One application of the central limit theorem on additive functions is the following result obtained by Erdős and Kac [13] (1940).

Theorem 2.3.4. *(Erdős-Kac theorem) Let f be a strongly additive function. Suppose that $|f(p)| \leq 1$ for all primes p, and*

$$\sum_{p \leq N} \frac{f(p)^2}{p} \to \infty \quad (N \to \infty).$$

Then

$$P_N \left(\frac{f - \mathbb{E}_N(f)}{\sqrt{\mathbb{E}_N \left| f - \mathbb{E}_N(f) \right|^2}} \leq x \right) \to \Phi(x) \quad (N \to \infty).$$

The conditions that $|f(p)| \leq 1$ and that f is strongly additive can be weakened, but not removed. The condition that $\sum_{p \leq y} \frac{f(p)^2}{p} \to \infty$, or in other words that the variance of f tends to infinity, is important to achieve a normal limit.

In particular, if $f(n) = \omega(n)$, they showed that

$$\frac{\omega(n) - \log \log n}{\sqrt{\log \log n}}$$

is asymptotically normally distributed, more formally for any real number x

$$\lim_{N\to\infty} P_N \left\{ \left| \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \right| < x \right\} = \Phi(x). \tag{2.3.1}$$

For each $\epsilon > 0$, we can easily show that

$$\lim_{N\to\infty} P_N \left\{ \left| \frac{\log\log n - \log\log N}{(\log\log N)^{1/2}} \right| > \epsilon \right\} = 0$$

So (2.3.1) is equivalent to the following statement which we prove it in this section.

$$\lim_{N\to\infty} P_N \left\{ \left| \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \right| < x \right\} = \Phi(x). \tag{2.3.2}$$

The Erdős-Kac Theorem has been studied by several mathematicians using different methods. For instance, Erdős and Kac in their original proof of (2.3.1) used complicated sieve methods. But an interesting proof of this theorem is Billingsley's result (1969) who used the method of moments and gave an easy demonstration of this theorem. Here we briefly explain the key steps of Billingsley's result, because we shall use the same method to prove a similar result for $y-$smooth integers.

Let P_N denote the uniform distribution on the set $\{1,2,...,N\}$, and $\lim_{N\to\infty} P_N$ exists. Moreover, let A_p be the set of integers divisible by the prime p. Then we can say that $\lim_{N\to\infty} P_N(A_p) = \frac{1}{p}$, and if $p \neq q$, we have

$$\lim_{N\to\infty} P_N(A_p \cap A_q) = \frac{1}{pq} = \lim_{N\to\infty} P_N(A_p)P_N(A_q).$$

Thus, the events are independent.

Now we are ready to prove the Erdős-Kac theorem in 5 steps:

Step (1): Let $n \leq N$ and $\mathbb{1}_{p|n}(n) = 1$, if p divides n and $= 0$, otherwise. Then, we can represent $\omega(n)$ as follows:

$$\omega(n) = \sum_{p \text{ prime}} \mathbb{1}_{p|n}(n). \tag{2.3.3}$$

We shall compare the indicator function $\mathbb{1}_{p|n}(n)$ with some independent random variables X_p's such that for each prime p:

$$P(X_p = 1) = 1/p \quad \text{and} \quad P(X_p = 0) = 1 - 1/p,$$

and the mean and variance of $S := \sum_{p\le N} X_p$ are

$$\mathbb{E}_N(S) = \sum_{p\le N} 1/p \quad \text{and} \quad \mathbb{V}_N(S) = \sum_{p\le N} 1/p(1-1/p)$$

respectively. By the Mertens' estimate, we have

$$\mathbb{E}_N(S) = \mathbb{V}_N(S) = \log\log N + O(1).$$

Step (2): The key step of the proof is to show that it is unaffected if we replace $\omega(n)$ by $\omega_\alpha(n)$, defined by

$$\omega_\alpha(n) = \sum_{p\le \alpha_N} \mathbb{1}_{p|n},$$

where

$$\alpha_N := N^{1/\log\log N}.$$

By using the Mertens' estimate, one can easily show that

$$\left(\sum_{\alpha_N\le p\le N} 1/p\right) /(\log\log N)^{1/2} \to 0 \quad \text{as} \quad N\to\infty.$$

Thus,

$$\sum_{\alpha_N\le p\le N} \frac{1}{p} = o(\sqrt{\log\log N}). \tag{2.3.4}$$

We have

$$\mathbb{E}_N(\omega(n)) = \mathbb{E}_N(\omega_\alpha(n)) + \mathbb{E}_N\left(\sum_{\alpha_N\le p\le N} \mathbb{1}_{p|n}\right). \tag{2.3.5}$$

Now using (2.3.4), we get

$$\mathbb{E}_N\left(\sum_{\alpha_N\le p\le N} \mathbb{1}_{p|n}\right) = \sum_{\alpha_N\le p\le N} \frac{1}{N}\left\lfloor\frac{N}{p}\right\rfloor \le \sum_{\alpha_N\le p\le N} 1/p = o(\sqrt{\log\log N}).$$

We deduce that it is enough to prove the Erdős-Kac theorem for ω_α instead of ω, in other words we will show that

$$\lim_{N\to\infty} P_N\left\{\left|\frac{\omega_\alpha(n)-\log\log N}{\sqrt{\log\log N}}\right|<x\right\}=\Phi(x). \tag{2.3.6}$$

Step (3): We define

$$S_\alpha := \sum_{p\leq\alpha_N} X_p.$$

If $b_N=\mathbb{E}_N(S_\alpha)$ and $c_N^2=\mathbb{V}_N(S_\alpha)$, By Step (2), we obtain

$$b_N=c_N^2=\log\log N+o\left((\log\log N)^{1/2}\right).$$

This means that S_α and ω_α have a same mean value and variance.

Step (4): To complete the proof, it suffices to show

$$P_N\left(\frac{\omega_\alpha(n)-b_N}{c_N}\leq x\right)\to\Phi(x) \quad \text{as} \quad N\to\infty.$$

By Theorem (2.3.2), we have

$$\lim_{N\to\infty} P_N\left(\frac{S_\alpha-b_N}{c_N}<x\right)=\Phi(x),$$

and since $|X_p|\leq 1$, the moments of S_α are bounded. Therefore, its moments tend to the moments of the normal distribution. Consequently,

$$\lim_{n\to\infty}\mathbb{E}_N\left[\left(\frac{S_\alpha-b_N}{c_N}\right)^r\right]=m_r \quad r=1{,}2,\ldots,$$

where m_r is the rth moment of the normal distribution.

Step (5): If $p_1,\ldots,p_k$ are all the primes satisfying $p_1<\cdots<p_k\leq\alpha_N$. By using the definition of X_{p_i} $(1\leq i\leq k)$ and $\mathbb{1}_{p_i|n}$, we have

$$\mathbb{E}_N\left[X_{p_1}...X_{p_k}\right]=\frac{1}{p_1\ldots p_k},$$

and

$$\mathbb{E}_N\left[\mathbb{1}_{p_1|n}...\mathbb{1}_{p_k|n}\right] = \frac{1}{N}\left\lfloor\frac{N}{p_1\ldots p_k}\right\rfloor,$$

Thus,

$$\mathbb{E}_N\left[X_{p_1}...X_{p_k}\right] - \mathbb{E}_N\left[\mathbb{1}_{p_1|n}\ldots\mathbb{1}_{p_k|n}\right] \leq 1/N \tag{2.3.7}$$

We now compute the difference of the rth moment of two random variables S_α and ω_α. By using the multinomial theorem and (2.3.7), we have

$$\left|\mathbb{E}_N(S_\alpha^r) - \mathbb{E}_N(\omega_\alpha^r)\right| \leq \sum_{k=1}^{r}\sum_{r_i}\frac{r!}{r_1!...r_k!}\sum_{p_1,...,p_k\leq\alpha_N}\left(\mathbb{E}_N\left[X_{p_1}^{r_1}...X_{p_k}^{r_k}\right] - \mathbb{E}_N\left[\mathbb{1}_{p_1|n}^{r_1}...\mathbb{1}_{p_k|n}^{r_k}\right]\right), \tag{2.3.8}$$

where $\sum_{r_i}$ is over k-tuples $(r_1,\ldots,r_k)$ of positive integers with $r_1+\ldots r_k = r$. So by the multinomial theorem and (2.3.7), we have

$$\left|\mathbb{E}_N(S_\alpha^r) - \mathbb{E}_N(\omega_\alpha^r)\right| \leq \frac{1}{N}\left(\sum_{p\leq\alpha_N}1\right)^r \leq \frac{\alpha_N^r}{N} \to 0 \quad \text{as} \quad N\to\infty. \tag{2.3.9}$$

Now by applying the binomial theorem and using (2.3.9), we easily get

$$\begin{aligned}\left|\mathbb{E}_N\left(S_\alpha - b_N\right)^r - \mathbb{E}_N\left(\omega_\alpha - b_N\right)^r\right| &\leq \sum_{k=0}^{r}\binom{r}{k}\frac{\alpha_N^k}{N}b_N^{r-k}\\ &= \frac{1}{N}\left(\alpha_N + b_N\right)^k \to 0 \quad (N\to\infty).\end{aligned} \tag{2.3.10}$$

This is more than enough to conclude that

$$\mathbb{E}_N\left[\left(\frac{\omega_\alpha - b_N}{c_N}\right)^r\right] \to m_r \quad \text{as} \quad N\to\infty,$$

and the desired result follows from (2.3.3).

Motivated by the Erdős-Kac theorem, Alladi [2], Hensley [22] and Hildebrand [24] studied a type of this theorem for $y-$smooth integers. The problem gets more complicated, since the behaviour of $\omega(n)$ changes in different ranges of y.

First, Hensley proved an analogue of the Erdős-Kac theorem for $y-$smooth integers using a *local limit theorem* i.e., an estimate for

$$\sum_{\substack{n\in S(x,y)\\ \Omega(n)=k}} 1,$$

where u lies in the range

$$(\log y)^{1/3} \leq u \leq \frac{\sqrt{y}}{2\log y}.$$

In this direction, Alladi (1987) obtained an analogue of the Erdős-Kac theorem for the additive function $\Omega(n)$ in the range

$$\exp\left(\log\log x\right)^{\frac{5}{3}+\epsilon} < y \leq x, \tag{2.3.11}$$

where ϵ is arbitrarily small. He used the saddle point method to show that

$$\mathbb{E}\left[\left(\frac{\Omega(n)-\eta(x,y)}{\sqrt{\theta(x,y)}}\right)^k\right] \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} u^k e^{-u^2/2}du \quad \text{as} \quad x\to\infty,$$

where $\eta(x,y)$ and $\theta(x,y)$ are the mean value and the variance of $\Omega(n)$, in the range (2.3.11), where $n \in S(x,y)$.

In the same year, Hildebrand , by considering the characteristic function of $\Omega(n)$, could prove that the characteristic function of $\Omega(n)$ tends to the characteristic function of a Gaussian distribution as x tends to infinity. More formally, for $-\pi \leq \theta \leq \pi$, we have

$$\frac{1}{\Psi(x,y)}\sum_{n\in S(x,y)} e^{i\theta\Omega(n)} = e^{i\theta M-\theta^2 V/2}\left\{1+O\left(\frac{1}{\sqrt{\bar{u}}}+\frac{|\theta|^3u^3}{\bar{u}^2}\right)\right\} + O_\epsilon\left(\exp\left(-(\log y)^{\frac{3}{2}-\epsilon}\right)\right),$$

uniformly in the range

$$u \geq (\log y)^{20}, \tag{2.3.12}$$

where $M = M(x,y)$ and $V = V(x,y)$ are defined as the mean value and the variance of $\Omega(n)$ in range (2.3.12). Then, he concluded his main theorem which says that for large values of x the distribution of $\Omega(n)$ in the range (2.3.12) follows a Guassian distribution.

Motivated by these results, we studied the same problem by the method of moments. Although this method is very interesting and simple it is not strong enough to cover the whole range of y in Erdős-Kac problem. However, using this method in Section 4, we will prove an analogue of Erdős-Kac problem for $y-$smooth integers in a small range of y.

Chapter 3

THE Y-SMOOTH MULTIPLICATION TABLE

3.1. Introduction

The *multiplication table problem* involves estimating

$$A(x) := \#\{ab : a,b \leq \sqrt{x}, \quad \text{and} \quad a,b \in \mathbb{N}\}.$$

This interesting question, posed by Erdős, has been studied by many authors. Erdős in [14], showed that for all $\varepsilon > 0$, we have

$$\frac{x}{(\log x)^{\delta+\epsilon}} \leq A(x) \leq \frac{x}{(\log x)^{\delta-\varepsilon}} \quad (x \to \infty), \tag{3.1.1}$$

where

$$\delta = 1 - \frac{1 + \log\log 2}{\log 2} = 0.0860\ldots. \tag{3.1.2}$$

The best estimate of $A(x)$ is a result due to Kevin Ford [16]. He proved the following estimate, that significantly improved the order of magnitude of $A(x)$ as follows

$$A(x) \asymp \frac{x}{(\log x)^{\delta}(\log\log x)^{3/2}}. \tag{3.1.3}$$

Notation: In this chapter, we use the notation $f(x) \asymp g(x)$ if both $f(x) \ll g(x)$ and $g(x) \ll f(x)$ hold, where $f(x) \ll g(x)$ or $f(x) = O(g(x))$ interchangeably to mean that $|f(x)| \leq cg(x)$ holds with some constant c for all x in a range which will normally be clear from the context. Also, the notation $f(x) \sim g(x)$ means that $f(x)/g(x) \to 1$ as $x \to \infty$, and $f(x) = o(g(x))$ means that $f(x)/g(x) \to 0$ as $x \to \infty$.

Also, u is defined as

$$u := \frac{\log x}{\log y} \qquad x \geq y \geq 2,$$

and we let $\log_k x$ denote the k-fold iterated logarithm, defined by $\log_1 x := \log x$ and $\log_k x = \log\log_{k-1} x$, for $k > 1$.

Motivated by this background, in this paper we investigate the multiplication table problem for smooth integers. The set of $y-$smooth numbers, is defined by

$$S(x,y) := \{n \leq x : P(n) \leq y\},$$

where $P(n)$ denotes the largest prime factor of an integer $n \geq 2$, and $P(1) = 1$. Set

$$\Psi(x,y) := \Big|S(x,y)\Big|.$$

Our main aim in this work is to study

$$A(x,y) := \#\{ab : a,b \in S(\sqrt{x},y)\}.$$

Hence computing $A(x,y)$ is equivalent to estimating the size of $S(\sqrt{x},y) \cdot S(\sqrt{x},y)$.
A simple approximation of $\Psi(x,y)$ proved by Canfield, Erdős and Pomerance [6] states that for a fixed $\epsilon > 0$, we have

$$\Psi(x,y) = xu^{-u(1+o(1))} \quad \text{as} \quad u \to \infty, \tag{3.1.4}$$

for $u \leq y^{1-\epsilon}$, that is $y \geq (\log x)^{1+\epsilon}$.
By estimate (3.1.4), one can see that for u large (or y small), the value of $\Psi(x,y)$ is small. It counts the integers having large number of prime factors. Since in this case every n has a lot of small prime factors, we can find a and b such that $n = ab$ and $a,b \leq \sqrt{x}$.

If u is small (which means that y is large), then by (3.1.4), one can deduce that the value of $\Psi(x,y)$ is large compared to x. In this case, $S(x,y)$ contains integers with large prime factors and we expect the size of $S(\sqrt{x},y) \cdot S(\sqrt{x},y)$ to be small.

It is good to mention that by a connection to sum-product problem, Banks and Covert [3] by invoking combinatorial tools, have considered the behaviour of $A(x^2,y) = \Big|S(x,y) \cdot S(x,y)\Big|$ in different ranges of y, particularly for the cases when y is relatively small or large. (See Section 1 of this thesis.)

Here we present a simple idea to prove that $A(x,y)$ has a same size as $\Psi(x,y)$ when y is small compared to $\log x$. Let $n \leq \frac{x}{y}$ be a $y-$smooth number. If $n \leq \sqrt{x}$ then trivially we have $n \in A(x,y)$. Thus, we assume that $\sqrt{x} \leq n$. Let $p_1 \leq p_2 \leq \cdots \leq p_k$ be prime factors of n. Consider the following sequence obtained by prime factors of n:

$$n_0 = 1, \qquad n_j = \prod_{i=1}^{j} p_i, \qquad 1 \leq j \leq k.$$

Since $n \geq \sqrt{x}$ then there exists a unique integer s, with $0 \leq s < k$ such that $n_s < \sqrt{x} \leq n_{s+1}$. Each prime factor of n is less than y, therefore

$$n_s \leq \sqrt{x} \leq n_{s+1} \leq n_s y.$$

Set $d = n_s$, then

$$\frac{\sqrt{x}}{y} \leq d \leq \sqrt{x}.$$

Since $n \leq x/y$, then we easily conclude that

$$\frac{n}{d} \leq \sqrt{x}.$$

Therefore,

$$\Psi(x/y,y) \leq A(x,y) \leq \Psi(x,y),$$

and by a simple argument one can deduce that as $x,y \to \infty$ then $\Psi(x/y,y) \sim \Psi(x,y)$ when $y = o(\log x)$, (see Lemma 3.2.1). This argument leads us to state the following theorem.

Theorem 3.1.1. *If $y = o(\log x)$ then we have*

$$A(x,y) \sim \Psi(x,y) \quad \textit{as} \quad x,y \to \infty.$$

The problem gets harder, and hence, more interesting when y takes larger values compared to $\log x$. We shall prove the following theorem for small values of y compared to x.

Theorem 3.1.2. *We have*

$$A(x,y) \sim \Psi(x,y) \quad \textit{as} \quad x,y \to \infty,$$

when u and y satisfy the range

$$\frac{u \log u}{(\log y \log_2 y \log_3 y)^2} \to \infty, \quad \textit{which implies,} \quad y \leq \exp\left\{\frac{(\log x)^{1/3}}{(\log_2 x)^{1/3+\epsilon}}\right\}, \tag{3.1.5}$$

for $\epsilon > 0$ arbitrarily small.

Theorem 3.1.2 is proved in Section 3. The proof relies on some probabilistic arguments and recent estimates for $\Psi(x/p,y)$ where p is a prime factor of n.

If y takes values very close to x, which implies u is small compared to $\log\log y$, then we will show the following theorem.

Theorem 3.1.3. *Let $\epsilon > 0$ is arbitrarily small, then we have*

$$A(x,y) = o(\Psi(x,y)) \quad as \quad x,y \to \infty,$$

where u and y satisfying the range

$$u < (L-\epsilon)\log_2 y, \quad \textit{which implies}, \quad y \geq \exp\left\{\frac{\log x}{(L-\epsilon)\log_2 x}\right\}, \tag{3.1.6}$$

where $L := \frac{1-\log 2}{\log 2}$.

Theorem 3.1.3 is proved in Section 4, by applying an Erdős' idea [12], suitably modified for $y-$smooth integers.

In what follows, we will give a heuristic argument that predicts the behaviour of $A(x,y)$ in ranges (3.1.5) and (3.1.6).

We define the function $\tau(n;A,B)$ to be the number of all divisors of n in the interval $(A,B]$. In other words.

$$\tau(n;A,B) := \#\{d : d|n \Rightarrow A < d \leq B\}.$$

Let $n \in S((1-\eta)x,y)$ be a square-free number with k prime factors, where $\eta \to 0$ as $x \to \infty$. Assume that the set

$$D(n) := \{\log d : d|n\}$$

is uniformly distributed in the interval $[0, \log n]$. So

$$P\left(d \in (A,B)\right) := \tau(n)\frac{\log B - \log A}{\log n}, \tag{3.1.7}$$

where the sample space is defined by

$$S := \{n \leq x : \omega(n) = k\},$$

and n being chosen uniformly at random. Then by this assumption, the expected value of the function $\tau(n,(1-\eta)\sqrt{x},\sqrt{x})$ should be

$$\mathbb{E}\left[\tau(n,(1-\eta)\sqrt{x},\sqrt{x})\right] = \frac{2^k \log(1/(1-\eta))}{\log\sqrt{x}} \asymp \frac{2^k}{u\log y}. \tag{3.1.8}$$

Alladi and Hildebrand in [2] and [24] showed that the normal number of prime factors of $y-$smooth integers is very close to its expected value $u + \log_2 y$ in different ranges of y. Hence, from (3.1.8), we deduce that

$$\mathbb{E}\left[\tau(n,(1-\eta)\sqrt{x},\sqrt{x})\right] \asymp \frac{2^{u+\log_2 y}}{\log y}.$$

If $2^{u+\log_2 y}/\log y \to \infty$, then we expect that n will have a divisor d in the interval $((1-\eta)\sqrt{x},\sqrt{x}]$.

We know $n \leq (1-\eta)x$. Thus, $n/d \leq \sqrt{x}$, and we can deduce that $n \in A(x,y)$, this means that

$$\Psi((1-\eta)x,y) \leq A(x,y).$$

Trivially $A(x,y) \leq \Psi(x,y)$. So by this argument, we obtain

$$A(x,y) \sim \Psi(x,y),$$

when $\eta \to 0$ as $x \to \infty$.

On the other hand, if $2^{u+\log_2 y}/\log y \to 0$, then we expect that none of integers in $S((1-\eta)x,y)$ have a divisor in $((1-\eta)\sqrt{x}, \sqrt{x}]$ (except a set of measure 0), this means that

$$A(x,y) = o(\Psi(x,y)) \quad \text{as} \quad x,y \to \infty.$$

This heuristic gives an evidence for the following conjecture:

Conjecture 3.1.1. *If $L := \frac{1-\log 2}{\log 2}$, then we have the following dichotomy*

(1) : If $u - L\log_2 y \to +\infty$, which implies

$$y \leq \exp\left\{\frac{\log x}{L\log_2 x}\right\},$$

Then, we have

$$A(x,y) \sim \Psi(x,y) \quad \text{as} \quad x,y \to \infty.$$

(2) : If $u - L\log_2 y \to -\infty$*, which implies that for small* $\epsilon > 0$

$$y \geq \exp\left\{\frac{\log x}{(L-\epsilon)\log_2 x}\right\},$$

Then, we have

$$A(x,y) = o(\Psi(x,y)) \quad as \quad x,y \to \infty.$$

Theorem 3.1.2 and Theorem 3.1.3 are in the direction of the first case and the second case of Conjecture (3.1.1) respectively, but the claimed ranges in the conjecture are stronger than the claimed ranges in Theorem 3.1.2 and Theorem 3.1.3, and the reason stems from uniformity assumption about $D(n)$.

3.2. Preliminaries

In this section, we review some results used in the proof of our main theorems. We first fix some notation. In this chapter $\rho(u)$ is the Dickman-de Bruijn function, as we defined in the introduction. By [18, 3.9] we have the following estimate for $\rho(u)$

$$\rho(u) = \left(\frac{e+o(1)}{u\log u}\right)^u \quad \text{as} \quad u \to \infty. \tag{3.2.1}$$

Theorem 3.2.1 (Hildebrand [23])**.** *The estimate*

$$\Psi(x,y) = x\rho(u)\left(1 + O_\epsilon\left(\frac{\log(u+1)}{\log y}\right)\right) \tag{3.2.2}$$

holds uniformly in the range

$$x \geq 3, \qquad 1 \leq u \leq \frac{\log x}{(\log_2 x)^{\frac{5}{3}+\epsilon}}, \quad \textit{that is,} \quad y \geq \exp\left((\log_2 x)^{\frac{5}{3}+\epsilon}\right), \tag{3.2.3}$$

where ϵ *is any fixed positive number.*

Combining (3.2.2) with the asymptotic formula (3.2.1), one can arrive at the following simple corollary

Corollary 3.2.1. *We have*

$$\Psi(x,y) = xu^{-(u+o(u))},$$

as y *and* u *tend to infinity, uniformly in the range* (3.2.3)*, for any fixed* $\epsilon > 0$*.*

We will apply this estimate in the proof of Theorem 3.1.3. However this estimate of $\Psi(x,y)$ is not very sharp for large values of u, for which the saddle point method is more effective.

Let $\alpha := \alpha(x,y)$ be a real number satisfying

$$\sum_{p\leq y}\frac{\log p}{p^{\alpha}-1}=\log x. \tag{3.2.4}$$

One can show that α is unique. This function will play an essential role in this work, so we briefly recall some fundamental facts of this function that are used frequently. By [9, Lemma 3.1] we have the following estimates for α.

$$\alpha(x,y)=\frac{\log\left(1+y/\log x\right)}{\log y}\left\{1+O\left(\frac{\log_2 y}{\log y}\right)\right\} \qquad x\geq y\geq 2. \tag{3.2.5}$$

For any $\epsilon>0$, we have the particular cases

$$\alpha(x,y)=1-\frac{\xi(u)}{\log y}+O\left(\frac{1}{L_{\epsilon}(y)}+\frac{1}{u(\log y)^2}\right) \qquad \text{if } \ y\geq(\log x)^{1+\epsilon}, \tag{3.2.6}$$

where

$$L_{\epsilon}(y)=\exp\left\{(\log y)^{3/5-\epsilon}\right\}, \tag{3.2.7}$$

and $\xi(t)$ is the unique real non-zero root of the equation

$$e^{\xi(t)}=1+t\xi(t). \tag{3.2.8}$$

Also for small values of y, we have

$$\alpha(x,y)=\frac{\log(1+\frac{y}{\log x})}{\log y}\left\{1+O\left(\frac{1}{\log y}\right)\right\} \ \text{ if } \ 2\leq y\leq(\log x)^2. \tag{3.2.9}$$

We now turn to another ingredient related to the behaviour of $\Psi(x,y)$. The following estimate is a special case of a general result of de La Breteche and Tenenbaum [9, Theorem 2.4].

Theorem 3.2.2. *If $d\leq y$, then uniformly for $x\geq y\geq 2$ we have*

$$\Psi(x/d,y)=\left\{1+O\left(\frac{1}{u}+\frac{\log y}{y}\right)\right\}\frac{\Psi(x,y)}{d^{\alpha}}. \tag{3.2.10}$$

We can deduce the following lemma by Theorem 3.2.2 which completes the proof of Theorem 3.1.1

Lemma 3.2.1. *If* $y \geq 2$ *and* $y = o(\log x)$*, then we have*

$$\Psi(x/y, y) \sim \Psi(x,y) \quad as \quad x \to \infty. \tag{3.2.11}$$

Proof. (i): Let $y \geq (\log_2 x)^2$ and $y = o(\log x)$. By applying (3.2.10), if $d = y$, we obtain

$$\Psi(x/y,y) = \frac{\Psi(x,y)}{y^{\alpha}} \left\{ 1 + O\left(\frac{\log y}{y}\right) \right\}. \tag{3.2.12}$$

By combination of the above estimate along with (3.2.9), we get

$$\Psi(x/y,y) = \frac{\Psi(x,y)}{\left(1 + \frac{y}{\log x}\right)^{1+O\left(\frac{1}{\log y}\right)}} \left\{ 1 + O\left(\frac{\log y}{y}\right) \right\}. \tag{3.2.13}$$

We remark again that $y = o(\log x)$, so we obtain

$$\frac{1}{(1 + y/\log x)^{1+O(1/\log y)}} \to 1 \qquad \text{when} \qquad x \to \infty.$$

Also, we have

$$\frac{\log y}{y} \to 0 \quad \text{when} \quad x \to \infty,$$

since $y \geq (\log_2 x)^2$. Thus, by (3.2.13), we conclude

$$\frac{\Psi(x/y,y)}{\Psi(x,y)} \to 1 \quad \text{when} \quad x \to \infty.$$

(ii) : Let $2 \leq y \leq (\log_2 x)^2$, then by recalling Ennola's theorem 1.5.4, we get

$$\begin{aligned} \Psi(x/y,y) &= \frac{1}{\pi(y)!} \prod_{p \leq y} \frac{\log x/y}{\log p} \left\{ 1 + O\left(\frac{y^2}{\log x \log y}\right) \right\} \\ &= \frac{1}{\pi(y)!} \prod_{p \leq y} \frac{\log x}{\log p} \prod_{p \leq y} \left(1 - \frac{\log y}{\log x}\right) \left\{ 1 + O\left(\frac{y^2}{\log x \log y}\right) \right\} \\ &= \Psi(x,y) \left(1 + O\left(\pi(y) \frac{\log y}{\log x}\right)\right) \\ &= \Psi(x,y) \left(1 + O\left(\frac{y}{\log x}\right)\right), \end{aligned} \tag{3.2.14}$$

which gives that

$$\Psi(x/y,y) \sim \Psi(x,y) \quad \text{as} \quad x \to \infty,$$

and this completes the proof. □

Finally, we define

$$\theta(x,y,z) := \#\{n \le x : p|n \Rightarrow z \le p \le y\}.$$

This function has been studied extensively in the literature. Namely Friedlander [17] and Saias [29, 30] gave several estimates for $\theta(x,y,z)$ in different ranges. The following theorem is due to Saias [30, Theorem 5] which is used in Section 4.

Theorem 3.2.3. *There exists a constant $c > 0$ such that for $x \ge y \ge z \ge 2$ we have*

$$\theta(x,y,z) \le c\frac{\Psi(x,y)}{\log z}. \tag{3.2.15}$$

3.3. Proof of Theorem 3.1.2

We begin this section by setting some notation. Let η be defined by

$$\eta := \frac{1}{\log_3 y},$$

and set

$$N := \left\lfloor \frac{\log_2 y - \log \eta}{\log 2} + 2 \right\rfloor, \tag{3.3.1}$$

which play an essential role in process of the proof.

The idea of the proof of Theorem 3.1.2 is a combination of some probabilistic and combinatorial techniques. Before going through the details, we give a sketch of proof here.

The first step of proving Theorem 3.1.2 is to study the number of all prime factors of n in the *narrow intervals*

$$J_i := \left[(1-\kappa)y^{1-\frac{1}{2^i}}, y^{1-\frac{1}{2^i}}\right], \qquad 1 \le i \le N,$$

of *multiplicative length* $(1-\kappa)^{-1}$, where κ is defined as

$$\kappa := \frac{\eta}{2N}. \tag{3.3.2}$$

Also, we define the *tail interval*

$$J_\infty := [(1-\kappa)y, y].$$

Let $\omega_i(n)$ be the number of prime factors of n in J_i for each $i \in \{1,2,\ldots,N,\infty\}$, more formally

$$\omega_i(n) := \#\{p|n : p \in J_i\}. \tag{3.3.3}$$

We define $\mu_i(x,y)$ to be the expectation of $\omega_i(n)$, defined by

$$\mu_i(x,y) := \frac{1}{\Psi(x,y)} \sum_{n\in S(x,y)} \omega_i(n), \tag{3.3.4}$$

In Proposition 3.3.1, we will prove that for almost all $y-$smooth integers the value of $\omega_i(n)$ exceeds $\mu_i(x,y)/2$. We establish this by applying the Chebyshev's inequality

$$\frac{\#\{n \in S(x,y) : \omega_i(n) \le \mu_i(x,y)/2\}}{\Psi(x,y)} \le \frac{4\sigma_i^2(x,y)}{\mu_i^2(x,y)}, \tag{3.3.5}$$

where

$$\sigma_i^2(x,y) := \frac{1}{\Psi(x,y)} \sum_{n\in S(x,y)} \left(\omega_i(n) - \mu_i(x,y)\right)^2, \tag{3.3.6}$$

is the variance of $\omega_i(n)$ and $i \in \{1,2,\ldots,N,\infty\}$. We will conclude that there is at least one prime factor p_i in each J_i for $1 \le i \le N$ and N prime factors $q_1,\ldots,q_N$ in J_∞. Then by using the product of these prime factors in Corollary 3.3.1, we will find a divisor D_j of n such that

$$(1-\kappa)^N y^{N-j/2^N} \le D_j \le y^{N-j/2^N},$$

for an integer j in $\{0,1,\ldots 2^N-1\}$.
Then, we fix an integer n in $S((1-\eta)x,y)$, and by defining $m := \frac{n}{\prod_{i=1}^N p_i q_i}$, we will easily show that there is a divisor d_j of n, such that

$$\frac{\sqrt{n}}{y^N} y^{j/2^N} < d_j < \frac{\sqrt{n}}{y^N} y^{(j+1)/2^N}.$$

Multiplying D_j and d_j and using the definitions of η, κ and N, gives a new divisor d of n that helps us to write n as the product of two divisors less than $\sqrt{x}$.

Before stating technical lemmas we get an estimate for the expected value of $\omega_i(n)$ for all $1 \le i \le N$ and $i = \infty$. By changing the order of summation in (3.3.4), we can easily see that

$$\mu_i(x,y) = \sum_{p\in J_i} \frac{\Psi(x/p,y)}{\Psi(x,y)}. \tag{3.3.7}$$

By (3.2.10), we have the following estimate

$$\mu_i(x,y) = \sum_{p\in J_i} \frac{1}{p^{\alpha}} \left(1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right), \tag{3.3.8}$$

for all $1 \le i \le N$ and $x \ge y \ge 2$. Also, we obtain the following estimate for $\mu_i(x/q,y)$, where q is a prime divisor of n.

$$\mu_i(x/q,y) = \sum_{p\in J_i} \frac{1}{p^{\alpha_q}} \left\{1 + O\left(\frac{1}{u_q} + \frac{\log y}{y}\right)\right\}, \tag{3.3.9}$$

where $u_q := u - \log q/\log y$. By substitution we obtain $x/q = y^{u_q}$. Set the saddle point $\alpha_q := \alpha(x/q,y)$, defined as the unique real number satisfying in

$$\sum_{p\le y} \frac{\log p}{p^{\alpha_q} - 1} = \log(x/q). \tag{3.3.10}$$

We are ready to prove the following lemma that shows the difference between $\mu_i(x/q,y)$ and $\mu_i(x,y)$ is small.

Lemma 3.3.1. *Let q be a prime divisor of $n \in S(x,y)$, then we have*

$$\left|\mu_i(x/q,y) - \mu_i(x,y)\right| \ll \frac{\mu_i(x,y)}{u}.$$

Proof. We use the estimate

$$0 < -\alpha'(u) := -\frac{d\alpha(u)}{du} \asymp \frac{\bar{u}}{u^2 \log y}, \tag{3.3.11}$$

established in [25, formula 6.6], where $\bar{u} := \min\{u, \frac{y}{\log y}\}$. By (3.3.11), we deduce

$$\left|\alpha'(u)\right| \ll \frac{1}{u\log y}. \tag{3.3.12}$$

Then applying (3.3.12), gives that

$$\begin{aligned} \alpha - \alpha_q \le \int_{u_q}^{u} \left|\alpha'(v)\right| dv &\ll \int_{u}^{u_q} \frac{dv}{v\log y} \\ &= \frac{1}{\log y} \log\left(\frac{u}{u_q}\right) \asymp \frac{\log q}{\log y \log x}. \end{aligned} \tag{3.3.13}$$

By expanding $\mu_i(x/q,y) - \mu_i(x,y)$ and using (3.3.7) and (3.3.9), we get

$$\begin{aligned}\left|\mu_i(x/q,y) - \mu_i(x,y)\right| &= \Big|\sum_{p\in J_i}\left(\frac{\Psi(x/pq,y)}{\Psi(x/q,y)} - \frac{\Psi(x/p,y)}{\Psi(x,y)}\right)\Big| \\ &\leq \sum_{p\in J_i}\frac{1}{p^{\alpha}}\left\{\left|p^{\alpha-\alpha_q} - 1\right| + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right\}.\end{aligned} \tag{3.3.14}$$

By the Taylor expansion of the exponential function and invoking (3.3.13) we obtain

$$\exp\{(\alpha - \alpha_q)\log p\} - 1 \ll \frac{\log p \log q}{\log y \log x}. \tag{3.3.15}$$

We recall that $p,q \leq y$ for $1 \leq i \leq N$ and $i = \infty$. From this we infer that

$$\left|p^{\alpha-\alpha_q} - 1\right| \ll \frac{1}{u},$$

this finishes the proof. □

In the following lemma we shall find an upper bound for $\sigma_i^2(x,y)$ (defined in (3.3.6)) for each $i \in \{1,2,\ldots,N,\infty\}$.

Lemma 3.3.2. *We have*

$$\sigma_i^2(x,y) \ll \mu_i(x,y) + \mu_i^2(x,y)/u,$$

where $i \in \{1,2,\ldots,N,\infty\}$.

Proof. By the definition of $\sigma_i^2(x,y)$ in (3.3.6), we have

$$\sigma_i^2(x,y) = \frac{1}{\Psi(x,y)}\sum_{n\in S(x,y)}\left[\omega_i^2(n) - 2\mu_i(x,y)\omega_i(n) + \mu_i^2(x,y)\right].$$

Using the definition of $\omega_i(n)$ in (3.3.3), gives

$$\sum_{n\in S(x,y)}\omega_i(n) = \sum_{n\in S(x,y)}\sum_{p\in J_i}\mathbb{1}_{p|n} = \sum_{p\in J_i}\Psi(x/p,y),$$

where the indicator function $\mathbb{1}_{p|n}$ is 1 or 0 according to the prime p divides n or not. By the definition of $\mu_i(x,y)$ in (3.3.7), one can deduce that

$$\sum_{n\in S(x,y)}\omega_i(n) = \Psi(x,y)\mu_i(x,y).$$

By applying (3.3.7) and the equation above, we obtain

$$\begin{aligned}\Psi(x,y)\sigma_i^2(x,y) &= \sum_{n\in S(x,y)} \Big[\omega_i^2(n) - 2\mu_i(x,y)\omega_i(n) + \mu_i^2(x,y)\Big] \\ &= \sum_{n\in S(x,y)} \omega_i^2(n) - 2\Psi(x,y)\mu_i^2(x,y) + \psi(x,y)\mu_i^2(x,y) \\ &= \left(\sum_{\substack{p,q\in J_j \\ p\neq q}} \Psi(x/pq,y)\right) - \Psi(x,y)\mu_i^2(x,y) + \sum_{p\in J_i} \Psi(x/p,y) \\ &:= S_1 + S_2,\end{aligned}$$

where $S_1 := \sum_{\substack{p,q\in J_j \\ p\neq q}} \Psi(x/pq,y) - \Psi(x,y)\mu_i^2(x,y)$ and $S_2 := \sum_{p\in J_i} \Psi(x/p,y)$. We next find an upper bound for each S_i. We first consider S_1, by using (3.3.7) we can get

$$\sum_{\substack{p,q\in J_i \\ p\neq q}} \Psi(x/pq,y) - \Psi(x,y)\mu_i^2(x,y) \leq \sum_{p\in J_i} \Psi(x/p,y)\left(\mu_i(x/p,y) - \mu_i(x,y)\right). \tag{3.3.16}$$

By Lemma 3.3.1 and using (3.3.16), we obtain the following upper bound for S_1

$$S_1 \leq C\frac{\Psi(x,y)\mu_i^2(x,y)}{u}, \tag{3.3.17}$$

where C is a positive constant. It remains to estimate S_2, from (3.3.7) we have

$$S_2 = \Psi(x,y)\mu_i(x,y).$$

By substituting the upper bounds for S_1 and S_2, we get

$$\sigma_i^2(x,y) = \frac{S_1 + S_2}{\Psi(x,y)} \ll \left(\mu_i(x,y) + \frac{\mu_i^2(x,y)}{u}\right),$$

and the proof is complete. □

Now we give an order of magnitude for $\mu_i(x,y)$, where $i \in \{1,2,\ldots,N,\infty\}$

Lemma 3.3.3. *We have*

$$\mu_i(x,y) \asymp \kappa\frac{Y^{1-\frac{1}{2^i}}}{\log y},$$

where $i \in \{1,2,\ldots,N,\infty\}$*, and*

$$Y := y^{1-\alpha}.$$

Proof. By the definition of each J_i, we obtain the following simple inequalities

$$\frac{1}{y^{\alpha(1-1/2^i)}}\#\{p\in J_i\}\le\sum_{p\in J_i}\frac{1}{p^\alpha}\le\frac{1}{(1-\kappa)y^{\alpha(1-1/2^i)}}\#\{p\in J_i\}. \tag{3.3.18}$$

By applying the prime number theorem, we obtain

$$\begin{aligned}\#\{p:p\in J_i\}&=\pi(y^{1-1/2^i})-\pi((1-\kappa)y^{1-1/2^i})\\&=\frac{y^{1-1/2^i}}{\log\left(y^{1-1/2^i}\right)}-\frac{(1-\kappa)y^{1-1/2^i}}{\log\left((1-\kappa)y^{1-1/2^i}\right)}+O\left(\frac{y^{1-1/2^i}}{\log^2 y}\right)\\&=\frac{y^{1-1/2^i}}{(1-1/2^i)\log y}-\frac{(1-\kappa)y^{1-1/2^i}}{(1-1/2^i)\log y}\left(1+O\left(\frac{\log(1-\kappa)}{\log y}\right)\right)\\&=\frac{\kappa y^{1-1/2^i}}{(1-1/2^i)\log y}(1+o(1)),\end{aligned} \tag{3.3.19}$$

The last equality is true, since the given values of κ and N in (3.3.2) and (3.3.1) imply

$$\kappa\asymp 1/(\log_2 y\log_3 y). \tag{3.3.20}$$

By substituting (3.3.19) in (3.3.18) we have

$$\mu_i(x,y)\asymp\kappa\frac{Y^{1-1/2^i}}{\log y}, \tag{3.3.21}$$

□

By the above lemmas, we are now ready for proving the following proposition.

Proposition 3.3.1. *If u and y satisfy in range given in*(3.1.5)*, we have*

$$\#\left\{n\in S(x,y):\omega_i(n)>\frac{\mu_i(x,y)}{2}\quad\forall i\in\{1,\ldots,N,\infty\}\right\}\sim\Psi(x,y)\quad as\quad x,y\to\infty,$$

Proof. By the Chebyshev's inequality in (3.3.5) and using the upper bound for $\sigma_i^2(x,y)$ in lemma (3.3.2), we get

$$\#\left\{n\in S(x,y):\omega_i(n)\le\frac{\mu_i(x,y)}{2}\right\}\ll\Psi(x,y)\left(\frac{1}{\mu_i(x,y)}+\frac{1}{u}\right).$$

By the above inequality, we obtain an upper bound for the following set

$$\begin{aligned} M &:= \#\left\{ n \in S(x,y) : \exists i \in \{1,\dots,N,\infty\} \quad \text{such that } \ \omega_i(n) \le \frac{\mu_i(x,y)}{2} \right\} \\ &\ll \Psi(x,y) \left[\frac{1}{\mu_\infty(x,y)} + \frac{N}{u} + \sum_{i=1}^{N} \frac{1}{\mu_i(x,y)} \right]. \end{aligned} \tag{3.3.22}$$

Our main task that finishes the proof is to find a range such that $M/\Psi(x,y)$ tends to 0. By using Lemma 3.3.3 and substituting the order of magnitude of $\mu_i(x,y)$ in (3.3.22), we get

$$M \ll \Psi(x,y) \left[\frac{\log y}{\kappa Y} + \frac{N}{u} + \frac{\log y}{\kappa} \sum_{i=1}^{N} \frac{1}{Y^{1-1/2^i}} \right]. \tag{3.3.23}$$

In what follows, we find a lower bound for Y in two different ranges of y

(**i**) : If $y \le (\log x)^2$, then by (3.2.9) $\alpha \le 1/2 + o(1)$ as $y \to \infty$. Therefore,

$$Y \ge y^{1/2-o(1)} \ge y^{1/3}.$$

By substituting this lower bound in (3.3.23) and using the precise value of N in (3.3.1), we have

$$\begin{aligned} M &\ll \Psi(x,y) \left[\frac{\log y}{\kappa y^{1/3}} + \frac{N}{u} + \frac{\log y}{\kappa y^{1/3}} \sum_{i=1}^{N} y^{1/3(2^i)} \right] \\ &\ll \Psi(x,y) \left[\frac{\log_2 y}{u} + \frac{y^{1/6}\log y}{\kappa y^{1/3}} \left(1 + O\left(N y^{-1/12}\right)\right) \right] \\ &\ll \Psi(x,y) \frac{\log y}{\kappa y^{1/6}}, \end{aligned} \tag{3.3.24}$$

By using the asymptotic value of κ in (3.3.20), we obtain

$$M \ll \Psi(x,y) \frac{\log y \log_2 y \log_3 y}{y^{1/6}},$$

and clearly we have

$$M = o(\Psi(x,y)) \quad \text{as} \quad x,y \to \infty,$$

this finishes the proof for the case $y \le (\log x)^2$.

(**ii**) : If $y \geq (\log x)^2$, by applying (3.2.6), we have

$$1-\alpha = \frac{\xi(u)}{\log y} + O\left(\frac{1}{L_\epsilon(y)} + \frac{1}{u(\log y)^2}\right). \tag{3.3.25}$$

Using [31, Lemma 8.1], we have the following estimate of ξ

$$\xi(t) = \log(t\log t) + O\left(\frac{\log_2 t}{\log t}\right) \quad \text{if} \quad t > 3.$$

Therefore,

$$1-\alpha = \frac{\log(u\log u)}{\log y} + O\left(\frac{\log_2 u}{\log y \log u}\right),$$

Thus, we get

$$\begin{aligned} Y &= u\log u\left[1 + O\left(\frac{\log_2 u}{\log u}\right)\right] \\ &\asymp u\log u. \end{aligned} \tag{3.3.26}$$

By combining the above with the estimate in (3.3.26), and using the value of N in (3.3.1), we get

$$\begin{aligned} M &\ll \Psi(x,y)\left[\frac{\log y}{\kappa u\log u} + \frac{N}{u} + \frac{\log y}{\kappa u \log u}\sum_{i=1}^{N}(u\log u)^{1/2^i}\right] \\ &\ll \Psi(x,y)\left[\frac{N}{u} + \frac{\log y}{\kappa u\log u}\left((u\log u)^{1/2} + (u\log u)^{1/2^2} + ... + (u\log u)^{1/2^N}\right)\right] \\ &\ll \Psi(x,y)\left[\frac{N}{u} + \frac{\log y}{\kappa(u\log u)^{1/2}}\left(1 + O\left(N(u\log u)^{-1/4}\right)\right)\right] \\ &\ll \Psi(x,y)\left[\frac{\log_2 y}{u} + \frac{\log y}{\kappa(u\log u)^{1/2}}\right], \end{aligned} \tag{3.3.27}$$

By using the order of κ in (3.3.20), one can arrive at the following upper bound of M

$$M \ll \Psi(x,y)\frac{\log y\log_2 y\log_3 y}{(u\log u)^{1/2}}. \tag{3.3.28}$$

So there exists a constant c such that for all $i \in \{1,\dots,N,\infty\}$, we have

$$\#\{n \in S(x,y) : \omega_i(n) > \mu_i(x,y)/2 \quad \forall i\} \geq \Psi(x,y)\left(1 - c\frac{\log y\log_2 y\log_3 y}{(u\log u)^{1/2}}\right), \tag{3.3.29}$$

and this finishes the proof by letting

$$\frac{u \log u}{(\log y \log_2 y \log_3 y)^2} \to \infty.$$

$\square$

Corollary 3.3.1. *If x and y satisfy the range (3.1.5), then almost all n in $S(x,y)$ are divisible by at least one prime factor p_i in J_i, and N prime factors $q_1,\ldots,q_N$ in J_∞. Moreover, the product $\prod_{i=1}^N p_i q_i$ has a divisor D_j in each of intervals $[(1-\kappa)^N y^{N-j/2^N}, y^{N-j/2^N}]$, where $j \in \{0,1,\ldots,2^N-1\}$.*

PROOF. The first part of Corollary is a direct conclusion of Proposition 3.3.1.
For the second part, let n be a $y-$smooth integer satisfying the first part of Corollary. We fix the following divisor of n

$$D := \prod_{i=1}^N p_i q_i,$$

where $p_i \in J_i$ and $q_1,\ldots,q_N \in J_\infty$.
Let j be an arbitrary integer in $\{0,1,\ldots,2^N-1\}$. Moreover, we define

$$a_0 := N - \sum_{i=1}^N a_i,$$

where a_i's get the values 0 or 1 such that

$$\sum_{i=1}^N \frac{a_i}{2^i} = j/2^N. \tag{3.3.30}$$

We now define the divisor of D_j of D with the following form

$$D_j := \prod_{i=1}^N p_i^{a_i} \prod_{i=1}^{a_0} q_i,$$

By using the bounds of p_is and q_is, one can get the following bounds for D_j.

$$(1-\kappa)^N y^{N-\sum_{i=1}^N a_i/2^i} \le D_j \le y^{N-\sum_{i=1}^N a_i/2^i},$$

By using (3.3.30), we have

$$(1-\kappa)^N y^{N-j/2^N} \le D_j \le y^{N-j/2^N},$$

and this finishes our proof. $\square$

We are ready now to prove Theorem 3.1.2.

Proof of Theorem 3.1.2. Let $n \leq (1-\eta)x$ be a $y-$smooth integer with at least one prime factor p_i in each J_i , where $i = 1,..,N$, and N prime divisors $q_1,q_2,...,q_N$ in J_∞. Set

$$m := \frac{n}{\prod_{i=1}^{N} p_i q_i}.$$

By this definition, we get

$$\frac{n}{\prod_{i=1}^{N} p_i q_i} \geq \frac{n}{y^{2N}} > \sqrt{n},$$

when $4N \leq u$. Thus,

$$m > \sqrt{n}.$$

Let $\{r_v\}$ be the increasing sequence of prime factors of m and set $d_v = r_1...r_v$.
Clearly, m has at least one divisor bigger than $\frac{\sqrt{n}}{y^N}$. We suppose that l is the smallest integer such that $d_l \geq \frac{\sqrt{n}}{y^N}$, and evidently we have

$$d_{l-1} \leq \frac{\sqrt{n}}{y^N},$$

So, we arrive at the following bounds for d_l

$$\frac{\sqrt{n}}{y^N} \leq d_l \leq y d_{l-1} \leq \frac{\sqrt{n}}{y^{N-1}}, \tag{3.3.31}$$

We pick $k \in \{0,1,2...,2^N - 1\}$ such that

$$\frac{\sqrt{n}}{y^N} y^{k/2^N} \leq d_l \leq \frac{\sqrt{n}}{y^N} y^{(k+1)/2^N}. \tag{3.3.32}$$

By the second part of Corollary 3.3.1, for every k in $\{0,1,...,2^N - 1\}$ there exists a divisor D_k such that

$$(1-\kappa)^N y^{N-k/2^N} \leq D_k \leq y^{N-k/2^N},$$

We define $d := d_l D_k$, we have

$$(1-\kappa)^N \sqrt{n} \leq d \leq y^{1/2^N} \sqrt{n},$$

By using the values of N in (3.3.1) and κ in (3.3.2), we have

$$e^{-\eta/2}\sqrt{n} \leq d \leq e^{\eta/2}\sqrt{n}.$$

Applying the Taylor expansion for exponential functions, gives

$$\left(1-\eta+\frac{\eta^2}{2}+O(\eta^3)\right)^{1/2}\sqrt{n}\le d\le\left(1+\eta+\frac{\eta^2}{2}+O(\eta^3)\right)^{1/2}\sqrt{n}. \tag{3.3.33}$$

By using the assumption $n\le(1-\eta)x$ in the upper bound and lower bound above, we obtain

$$d\le\left(1-\frac{\eta^2}{2}+O(\eta^3)\right)^{1/2}\sqrt{x}\le\sqrt{x},$$

and

$$\frac{n}{d}\le\left(1+\eta+\frac{\eta^2}{2}+O(\eta^3)\right)^{1/2}\sqrt{n}\le\left(1-\frac{\eta^2}{2}+O(\eta^3)\right)^{1/2}\sqrt{x}\le\sqrt{x}.$$

Thus, we can write $n\in S((1-\eta)x,y)$ as the product of two divisors less than $\sqrt{x}$, and we can deduce that

$$\Psi\left((1-\eta)x,y\right)\le A(x,y)\le\Psi(x,y),$$

By using (3.2.10), we have

$$\frac{\Psi\left((1-\eta)x,y\right)}{\Psi(x,y)}=(1-\eta)^{\alpha}\left\{1+O\left(\frac{1}{u}+\frac{\log y}{y}\right)\right\}\to 1 \quad \text{as} \quad x,y\to\infty,$$

this finishes the proof.

□

3.4. Proof of Theorem3.1.3

In this section, we shall study the behaviour of $A(x,y)$ for large values of y. When y takes values very close to x, then the set of $y-$smooth integers contains integers having large prime factors. As we explained in the heuristic argument, one can expect that $A(x,y)=o(\Psi(x,y))$. To show this assertion, we recall the idea of Erdős used to prove the multiplication table problem for integers up to x.

We start our argument by giving an upper bound for $A^*(x)$, defined by

$$A^*(x):=\#\left\{ab: a,b\le\sqrt{x} \quad \text{and} \quad (a,b)=1\right\}. \tag{3.4.1}$$

We shall find an upper bound of $A^*(x)$ by considering the number of prime factors of a and b. We first define

$$\pi_k(x):=\#\{n\le x:\omega(n)=k\}$$

Therefore,

$$\begin{aligned} A^*(x) &\le \sum_k \min\left\{\pi_k(x), \sum_{j=1}^{k-1} \pi_j(\sqrt{x})\pi_{k-j}(\sqrt{x})\right\} \\ &\le \sum_k \min\left\{\frac{cx}{\log x}\frac{(\log_2 x)^{k-1}}{(k-1)!}, \sum_{j=1}^{k-1} \frac{c\sqrt{x}}{\log\sqrt{x}}\frac{(\log_2\sqrt{x})^{j-1}}{(j-1)!}\frac{c\sqrt{x}}{\log\sqrt{x}}\frac{(\log_2\sqrt{x})^{k-j-1}}{(k-j-1)!}\right\}, \end{aligned} \tag{3.4.2}$$

where in the last inequality, we used the well-known result of Hardy and Ramanujan that states there are absolute constants C and c such that

$$\pi_k(x) \le \frac{cx}{\log x}\frac{(\log_2 x + C)^{k-1}}{(k-1)!} \quad \text{for} \quad k = 0,1,2,.. \quad \text{and} \quad x \ge 2. \tag{3.4.3}$$

By simplifying the upper bound in (3.4.2) and using Stirling's formula

$$n! \sim n^{n+\frac{1}{2}}e^{-n}$$

we obtain

$$\begin{aligned} A^*(x) &\le \sum_k \min\left\{\frac{cx}{\log x}\frac{(\log_2 x)^{k-1}}{(k-1)!}, \frac{4c^2x}{(\log x)^2}\sum_{j=0}^{k-2}\frac{1}{(k-2)!}\binom{k-2}{j}(\log_2\sqrt{x})^{k-2}\right\} \\ &= \sum_k \min\left\{\frac{cx}{\log x}\frac{(\log_2 x)^{k-1}}{(k-1)!}, \frac{4c^2x}{(\log x)^2}\frac{(2\log_2\sqrt{x})^{k-2}}{(k-2)!}\right\} \\ &= \sum_{k\le\frac{\log_2 x}{\log 2}} \frac{4c^2x}{(\log x)^2}\frac{(2\log_2\sqrt{x})^{k-2}}{(k-2)!} + \sum_{k>\frac{\log_2 x}{\log 2}} \frac{cx}{\log x}\frac{(\log_2 x)^{k-1}}{(k-1)!} \\ &\ll \frac{x}{(\log x)^{1-\frac{1+\log\log 2}{\log 2}}(\log_2 x)^{1/2}} \to 0 \quad \text{as} \quad x \to \infty. \end{aligned} \tag{3.4.4}$$

We shall get the same upper bound for $A(x)$. Let $n \le x$ and there are a and b less than $\sqrt{x}$ such that $n = ab$. If $(a,b) = 1$ then n is counted by $A(x)$, and if $(a,b) = d > 1$ then we can write n as $n = a'b'd^2$ such that $(a',b') = 1$. So, $\frac{n}{d^2} \le \frac{x}{d^2}$, and $\frac{n}{d^2}$ will be counted by $A(\frac{x}{d^2})$. Therefore,

$$A(x) \le \sum_{d\le\sqrt{x}} A^*(\frac{x}{d^2}) \ll A^*(x)$$

By (3.4.4), we get

$$A(x) \ll \frac{x}{(\log x)^{1-\frac{1+\log\log 2}{\log 2}}(\log_2 x)^{1/2}}.$$

Thus,

$$A(x) = o(x) \quad \text{as} \quad x \to \infty.$$

Motivated by Erdős' idea for the multiplication table of integers up to x, we apply a similar method to find an upper bound for $A(x,y)$.
The first step of proof is to study the following function which plays a crucial role in this section. Let

$$N_k(x,y,z) := \#\{n \in S(x,y) : \Omega_z(n) = k\},$$

where $\Omega_z(n)$ is the truncated version of $\Omega(n)$, only counting divisibility by primes not exceeding z with their multiplicities. In other words

$$\Omega_z(n) := \sum_{\substack{p^v || n \\ p \leq z}} v.$$

In the following lemma, by using induction on k, we shall find an upper bound of type (3.4.3) for $N_k(x,y,z)$. The reason of applying truncation is to sieve out prime factors exceeding some power of y which are the cause of big error terms as k increases in each step of induction. The upper bound of $N_k(x,y,z)$ leads us to generalize Erdős' idea for $y-$smooth integers in a certain range of y.

Lemma 3.4.1. *Let $u \leq (C-\epsilon) \log\log y$, where C is a positive constant and $\epsilon > 0$ is arbitrarily small. Set the parameter z such that*

$$\log\log z \ll u.$$

Then, there are constants A and B such that the inequality

$$N_k(x,y,z) \leq \frac{A\Psi(x,y)}{\log z} \frac{(\log\log z + B)^k}{k!} \tag{3.4.5}$$

holds for every integer $k > 0$.

PROOF. When $k = 0$, by (3.2.15), evidently we have

$$N_0(x,y,z) = \theta(x,y,z) \le c\frac{\Psi(x,y)}{\log z},$$

where $c > 0$ is a constant. When $k = 1$, we can represent n as $n = pm$, where $p \le z$ and every prime factor q of m is between z and y, then using the definition of $\theta(x,y,z)$ we have

$$N_1(x,y,z) = \sum_{p\le z} \sum_{\substack{m\le x/p \\ q|m\Rightarrow z\le q\le y}} 1 = \sum_{p\le z} \theta(x/p,y,z).$$

By applying the estimate (3.2.10) and (3.2.15), there is constant c such that

$$N_1(x,y,z) \le \sum_{p\le z} \frac{c\Psi(x/p,y)}{\log z} = c\frac{\Psi(x,y)}{\log z} \sum_{p\le z} \frac{1}{p^\alpha} \left\{1 + O\left(\frac{1}{u}\right)\right\}.$$

For the last summand we have

$$\begin{aligned} \sum_{p\le z} \frac{1}{p^\alpha} &= \sum_{p\le z} \frac{1}{p} \left(p^{1-\alpha}\right) \\ &= \sum_{p\le z} \frac{1}{p} \left\{1 + O\left((1-\alpha)\log p\right)\right\}, \end{aligned} \tag{3.4.6}$$

since $(1-\alpha)\log p \le (1-\alpha)\log z$, and $(1-\alpha)\log z$ is bounded in our range (see (3.4.8)). Therefore,

$$\sum_{p\le z} \frac{1}{p^\alpha} = \log_2 z + O\left((1-\alpha)\log z\right), \tag{3.4.7}$$

By using the estimate of α in (3.2.6) and the upper bound of z, we get

$$(1-\alpha)\log z \ll \frac{\log u}{\log y}\log z \ll \frac{\log u}{\log_2 y} \ll \frac{\log_3 y}{\log_2 y}, \tag{3.4.8}$$

and we obtain

$$\sum_{p\le z} \frac{1}{p^\alpha} = \log_2 z + O\left(\frac{\log_3 y}{\log_2 y}\right). \tag{3.4.9}$$

Thus,

$$\sum_{p\le z} \frac{1}{p^\alpha} \left\{1 + O\left(\frac{1}{u}\right)\right\} = \log\log z + O(1), \tag{3.4.10}$$

since we have $\log\log z \ll u$.
Substituting (3.4.10) in the upper bound of $N_1(x,y,z)$, gives

$$N_1(x,y,z) \le \frac{c\Psi(x,y)}{\log z}\left(\log_2 z + O(1)\right).$$

We will show the lemma with $A = c$ and $B = O(1)$. We argue by induction: we assume that the estimate in (3.4.5) is true for any positive integer k, we now prove it for $n \in S(x,y)$ with $\Omega_z(n) = k+1$. There are $k+1$ ways to write n as $n = pm_1m_2$ such that $p \le z$ and $\Omega_z(m_1) = k$ and every prime factor of m_2 is greater than z. Then we have

$$\begin{aligned} N_{k+1}(x,y,z) &= \frac{1}{(k+1)}\sum_{p\le z}\sum_{\substack{m_1\in S(x/(p),y)\\ \Omega_z(m_1)=k\\ m_2\in S(x/(pm_1),y)\\ q|m_2\Rightarrow q>z}} 1 \le \frac{1}{(k+1)}\sum_{p\le z}\sum_{\substack{m_1\in S(x/(p),y)\\ \Omega_z(m_1)=k}} 1\\ &= \frac{1}{(k+1)}\sum_{p\le z} N_k(x/p,y,z)\end{aligned}$$

By the assumption for $\Omega_z(n) = k$ and (3.2.10), we get

$$\begin{aligned} N_{k+1}(x,y,z) &\le \frac{A(\log_2 z + B)^k}{\log z(k+1)!}\sum_{p\le z}\Psi(x/p,y)\\ &= \frac{A\Psi(x,y)}{\log z}\frac{(\log_2 z + B)^k}{(k+1)!}\sum_{p\le z}\frac{1}{p^\alpha}\left\{1 + O\left(\frac{1}{u}\right)\right\}. \end{aligned} \tag{3.4.11}$$

By applying the estimate in (3.4.10), we arrive at the following bound for $N_{k+1}(x,y,z)$

$$N_{k+1}(x,y,z) \le \frac{A\Psi(x,y)}{\log z}\frac{(\log_2 z + B)^{k+1}}{(k+1)!},$$

so we derived our desired result. □

Proof of Theorem 3.1.3. For a small $\epsilon > 0$, we set $u < \left(\frac{\lambda}{\log 2} - \epsilon\right)\log_2 y$, where λ is a fixed real number in the open interval $(1 - 2\log 2, 1 - \log 2)$.
We now set z satisfying

$$\log\log z = \frac{\log 2}{\lambda}u, \tag{3.4.12}$$

so the given ranges of u and z satisfy the conditions of Lemma 3.4.1.
By the definition of $A(x,y)$, we have the following evident bound of $A(x,y)$

$$A(x,y) \leq \sum_{k} \min \left\{ \sum_{\substack{n \in S(x,y) \\ \Omega_z(n)=k}} 1, \sum_{j=1}^{k-1} \sum_{\substack{a \in S(\sqrt{x},y) \\ \Omega_z(a)=j}} 1 \sum_{\substack{b \in S(\sqrt{x},y) \\ \Omega_z(b)=k-j}} 1 \right\}. \tag{3.4.13}$$

We set

$$L = \lfloor H \log_2 z \rfloor,$$

where

$$H := \frac{1-\lambda}{\log 2}.$$

We have $1 - 2\log 2 < \lambda < 1 - \log 2$. Thus, $1 < H < 2$.

By using (3.4.13), we write the following bound for $A(x,y)$

$$\begin{aligned} A(x,y) &\leq \# \{ n \in S(x,y) : \Omega_z(n) > L \} + \# \left\{ ab : a,b \in S(\sqrt{x},y), \Omega_z(a) + \Omega_z(b) \leq L \right\} \\ &= \sum_{k>L} N_k(x,y,z) + \sum_{k \leq L} \sum_{j=0}^{k} N_j(\sqrt{x},y,z) N_{k-j}(\sqrt{x},y,z). \end{aligned} \tag{3.4.14}$$

By applying Lemma 3.4.1, we have

$$\begin{aligned} A(x,y) &\ll \sum_{k>L} \frac{\Psi(x,y)}{\log z} \frac{(\log_2 z + c)^k}{k!} + \sum_{k \leq L} \sum_{j=0}^{k} \frac{\Psi^2(\sqrt{x},y)}{\log^2 z} \frac{(\log_2 z + c)^j}{j!} \frac{(\log_2 z + c)^{k-j}}{(k-j)!} \\ &= \sum_{k>L} \frac{\Psi(x,y)}{\log z} \frac{(\log_2 z + c)^k}{k!} + \sum_{k \leq L} \frac{\Psi^2(\sqrt{x},y)}{\log^2 z} \sum_{j=0}^{k} \frac{1}{k!} \binom{k}{j} (\log_2 z + c)^k \\ &= \sum_{k>L} \frac{\Psi(x,y)}{\log z} \frac{(\log_2 z + c)^k}{k!} + \sum_{k \leq L} \frac{\Psi^2(\sqrt{x},y)}{\log^2 z} \frac{(2\log_2 z + c)^k}{k!}. \end{aligned} \tag{3.4.15}$$

By applying the simple form of $\Psi(x,y)$ in Corollary 3.2.1, and using the assumption (3.4.12), we get

$$\frac{\Psi^2(\sqrt{x},y)}{\Psi(x,y)} \asymp (\log z)^{\lambda} \quad \text{as} \quad u,y \to \infty. \tag{3.4.16}$$

Thus,

$$A(x,y) \ll \frac{\Psi(x,y)}{\log z} \sum_{k>L} \frac{(\log_2 z + c)^k}{k!} + \frac{(\log z)^{\lambda}\Psi(x,y)}{\log^2 z} \sum_{k\le L} \frac{(2\log_2 z + c)^k}{k!}. \tag{3.4.17}$$

The maximum values of functions in the above summands (with respect to k) are attained at $k = \lfloor \log_2 z \rfloor$ and $k = \lfloor 2\log_2 z \rfloor$ respectively. We have $\log\log z < L < 2\log\log z$, so the function in the first summation in (3.4.17) in decreasing for $k > L$, and by using Stirling's formula $k! \sim k^{k+\frac{1}{2}} e^{-k}$, we have

$$\begin{aligned} \sum_{k>L} \frac{(\log_2 z)^k}{k!} &= \sum_{H\log_2 z<k\le e\log_2 z} \frac{(\log_2 z)^k}{k!} + \sum_{e\log_2 z<k\le 2e\log_2 z} \frac{(\log_2 z)^k}{k!} + \sum_{k>2e\log_2 z} \frac{(\log_2 z)^k}{k!} \\ &\ll (\log_2 z)\left(\left(\frac{e}{H}\right)^{H\log_2 z} + 1\right) \\ &\ll \frac{1}{(\log z)^{H\log H - H}}. \end{aligned} \tag{3.4.18}$$

The function in the second summation in (3.4.17) is increasing for $k \le L$, and we have

$$\sum_{k\le L} \frac{(2\log_2 z + c)^k}{k!} \ll (\log_2 z)\left(\frac{2e}{H}\right)^{H\log_2 z} = \frac{1}{(\log z)^{H\log H - H - H\log 2}} \tag{3.4.19}$$

Substituting the upper bounds obtained in (3.4.18) and (3.4.19) in (3.4.17), and using the definition of H, gives

$$A(x,y) \ll \frac{\Psi(x,y)}{(\log z)^{G(H)}},$$

where

$$G(H) := 1 + H\log H - H.$$

The function $G(H)$ is an increasing function in the interval $(1,2)$ with a zero at $H = 1$. Thus, for any arbitrary $1 - 2\log 2 < \lambda < 1 - \log 2$, we have

$$A(x,y) = o(\Psi(x,y)) \quad \text{as} \quad x,y \to \infty,$$

so we obtained our desired result. $\square$

Chapter 4

AN ERDŐS-KAC THEOREM FOR $Y-$SMOOTH AND $Y-$ULTRA-SMOOTH INTEGERS

4.1. Introduction

For an integer $n \geq 2$, let $\omega(n)$ denote the number of distinct prime divisors of n. In 1940, Erdős and Kac [13] in their celebrated work studied the distribution of $\omega(n)$ in the interval $[2,N]$. The theorem states that for any real number x, we have

$$\lim_{N\to\infty} \frac{1}{N}\#\left\{n \leq N : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \leq x\right\} = \Phi(x). \tag{4.1.1}$$

where $\Phi(x)$ is the *normal distribution function* defined by

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

There are several proofs of Erdős-Kac Theorem. For instance, it has been proved by Billingsley [5] and Granville and Soundararajan [19] using the method of moments and sieve theory. Different variations of this theorem have been considered by several authors. In the present note, we shall study the Erdős-Kac theorem for $y-$smooth numbers. Recall that

$$S(x,y) := \{n \leq x : P(n) \leq y\} \qquad x \geq y \geq 2,$$

is the set of $y-$smooth integers, where $P(n)$ is defined as the largest prime factor of n, with the convention $P(1) = 1$. Also, recall that we set

$$\Psi(x,y) := |S(x,y)| \qquad x \geq y \geq 2.$$

The main goal of this result is to prove an analogue of (4.1.1) with the set $S(x,y)$ in the range

$$u = o(\log\log y), \tag{4.1.2}$$

where, as always,

$$u := \frac{\log x}{\log y}.$$

Hildebrand [24], Alladi [2], and Hensley [22] have considered the distribution of prime divisors of $y-$smooth integers in different ranges of y.
Hensley proved an Erdős-Kac type theorem when u lies in the range

$$(\log y)^{1/3} \leq u \leq \frac{\sqrt{y}}{2\log y}.$$

By using different method Alladi obtained an analogue of the Erdős-Kac Theorem for the following range

$$u \leq \exp(\log y)^{3/5-\epsilon}.$$

Later, Hildebrand extended previous results to include the range

$$y \geq 3 \quad u \geq (\log y)^{20},$$

which is a completion of Alladi and Hensley's results.

Although (4.1.2) does not cover Alladi's, Hensley's and Hildebrand's ranges, our applied method is completely different and much easier than the methods used by previous authors. Our approach is based on the method of moments as Billinglsley used in [5]. We will introduce some approximately independent random variables, and by the Central Limit Theorem, we shall show that this random variables have a normal distribution, then by applying method of moments we get our desired result in (4.1.1).
The first step of the proof is to apply a truncation on number prime factors. This idea is from original proof of Erdős-Kac Theorem [13].

For a given real number y, set

$$\phi(y) := (\log\log y)^{\sqrt{\log\log\log y}},$$

then $y^{\frac{1}{\phi(y)}}$ is a function that helps us to sieve out all primes exceeding $y^{\frac{1}{\phi(y)}}$, and we will show the contribution of sieved primes is negligible in understanding the distribution of $\omega(n)$. Before stating the main result, we begin introducing some notation. Let $\omega(n)$ is the number of distinct prime divisors of a $y-$smooth number, namely

$$\omega(n) := \sum_{p\leq y} \mathbb{1}_{p|n}(n),$$

where $\mathbb{1}_{p|n}(n)$ is 1 and 0 according to the prime p divides n or not.
Let $\mu_\omega(x,y)$ be the mean value of $\omega(n)$, more formally

$$\mu_\omega(x,y) := \mathbb{E}_{n\in S(x,y)}[\omega(n)] = \frac{1}{\Psi(x,y)} \sum_{n\in S(x,y)} \omega(n),$$

and $\sigma^2_\omega(x,y)$ is the variance of $\omega(n)$, defined by

$$\sigma^2_\omega(x,y) := \mathbb{E}\left[(\omega(n) - \mu(x,y))^2\right].$$

Now we are ready to state the main theorem.

Theorem 4.1.1. *For any real number z, we have*

$$\frac{1}{\Psi(x,y)}\#\{n \in S(x,y) : \frac{\omega(n) - \log\log y}{\sqrt{\log\log y}} \leq z\} \to \Phi(z) \quad (y \to \infty) \tag{4.1.3}$$

holds in the range (4.1.2).

Theorem 4.1.1 is proved in Section 3. The proof relies on the method of moments and the estimate of $\Psi(x/d,y)/\Psi(x,y)$.

Let

$$U(x,y) := \{n \leq x : p^v||n \Rightarrow v \leq v_p\}$$

be the set of $y-$ultra-smooth integers, where

$$v_p := \left\lfloor \frac{\log y}{\log p} \right\rfloor.$$

We define

$$\Upsilon(x,y) := \Big|U(x,y)\Big|.$$

We also have the following theorem

Theorem 4.1.2. *For any real number z, we have*

$$\frac{1}{\Upsilon(x,y)}\#\{n \in U(x,y) : \frac{\omega(n) - \log\log y}{\sqrt{\log\log y}} \leq z\} \to \Phi(z) \quad (y \to \infty) \tag{4.1.4}$$

holds in the range (4.1.2).

The proof of Theorem 4.1.2 relies on the method of moments and the local behaviour of the function $\Upsilon(x,y)$. By recalling Theorem 1.5.8, for $u = o(\log\log y)$, we have

$$\frac{\Upsilon(x/d,y)}{\Upsilon(x,y)} = \frac{\Psi(x/d,y)}{\Psi(x,y)}\left\{1 + O\left(\frac{u\log 2u}{\sqrt{y}\log y}\right)\right\},$$

that is

$$\frac{\Upsilon(x/d,y)}{\Upsilon(x,y)} \sim \frac{\Psi(x/d,y)}{\Psi(x,y)} \quad \text{as} \quad y \to \infty.$$

Considering this relation between the local behaviour of $\Upsilon(x,y)$ and $\Psi(x,y)$ gives us a similar proof as Theorem 4.1.1, so we shall avoid proving this theorem.

4.2. Preliminaries

Here we briefly recall some standard facts from probability theory (See Feller [15] for more details) and we shall give a few important lemmas.

Remark 4.2.1. *If a random variable D_n converges to 0 in probability, particularly $\mathbb{E}\{|D_n|\} \to 0$, then a second random variable U_n (on the same probability space) tend to Φ in distribution if and only if $U_n + D_n \to \Phi$ in distribution.*

Remark 4.2.2. *If distribution function F_n satisfying $\int_{-\infty}^{\infty} x^k dF_n(x) \to \int_{-\infty}^{\infty} x^k d\Phi(x)$ as $n \to \infty$, for $k = 1, 2, ...$, then $F_n(x) \to \Phi(x)$ for each x.*

Remark 4.2.3. *If $F_n(x) \to \Phi(x)$ for each x, and if $\int_{-\infty}^{\infty} |x|^{k+\epsilon} dF_n(x)$ is bounded in n for some positive ϵ, then, $\int_{-\infty}^{\infty} x^k dF_n(x) \to \int_{-\infty}^{\infty} x^k d\Phi(x)$.*

Remark 4.2.4. *(A special case of the central limit theorem): If $X_1,X_2,\ldots$ are independent and uniformly bounded random variables with mean 0 and finite variance σ_i^2 and if $\sum \sigma_i^2$ diverges then the distribution of $\frac{\sum_{i=1}^{n} X_i}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{1/2}}$ converges to the normal distribution function.*

By recalling Theorem 1.5.1 for $d \leq y$ and $y \geq (\log x)^{1+\epsilon}$, we have

$$\Psi(x/d,y) = \frac{\Psi(x,y)}{d^\alpha}\left\{1 + O\left(\frac{1}{u_y} + \frac{\log d}{\log x}\right)\right\}, \tag{4.2.1}$$

where $u_y := u + \frac{\log y}{\log(u+2)}$ and $\alpha = \alpha(x,y)$ denotes the saddle point of the Perron's integral for $\Psi(x,y)$, which is the solution of the following equation

$$\sum_{p\leq y}\frac{\log p}{p^\alpha - 1} = \log x.$$

This function will play an important role in this work, so we briefly recall some fundamental facts about this function. By [9, Lemma3.1], for any $\epsilon > 0$, we have the following estimate for α

$$\alpha = 1 - \frac{\xi(u)}{\log y} + O\left(\frac{1}{L_\epsilon(y)} + \frac{1}{u(\log y)^2}\right) \qquad \text{if} \qquad y \geq (\log x)^{1+\epsilon}, \tag{4.2.2}$$

where $\xi(u)$ is a unique real non-zero root of the equation

$$e^\xi = 1 + u\xi,$$

and when $u \geq 3$, we have

$$\xi(u) = \log(u\log u) + O\left(\frac{\log\log u}{\log u}\right). \tag{4.2.3}$$

By [9, Lemma 4.1], we have the following important estimate

Lemma 4.2.1. *(De la Breteche, Tenenbaum) For any $x \geq y \geq 2$, uniformly we have*

$$\sum_{p\leq y}\frac{1}{p^\alpha} = \log\log y + \left\{1 + O\left(\frac{1}{\log y}\right)\right\}\frac{uy}{y + \log x}. \tag{4.2.4}$$

Here we use a particular case of Lemma 4.2.1. If the range of y is restricted to $\log x < y \leq x$, we get

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

thus,

$$\sum_{p\leq y}\frac{1}{p^\alpha} = \log\log y + u + O\left(\frac{u}{\log y}\right) \qquad y > \log x. \tag{4.2.5}$$

For $2 \le t \le y \le x$, we define

$$\omega_t(n) := \#\{p : p|n, p \le t\} = \sum_{p \le t} \mathbb{1}_{p|n}.$$

By using the saddle point method, Tenenbaum and de la Breteche in [8] obtained an estimate for the expectation and the variance of $\omega_t(n)$. First, we define

$$M(t) = M_{x,y}(y) := \sum_{p \le t} \frac{1}{p^\alpha}.$$

We state the following lemma from [8].

Lemma 4.2.2. *(Tenenbaum, de la Breteche) we have uniformly for* $2 \le t \le y \le x$

$$\mu_{\omega_t}(x,y) = M(t) + O(1). \tag{4.2.6}$$

We now study the expectation of $\omega(n)$, where $n \in S(x,y)$.

Lemma 4.2.3. *If* $u = o(\log \log y)$, *then we have*

$$\mu_\omega(x,y) = \log \log y + o(\log \log y).$$

Proof. Let $t = y$ in Lemma 4.2.2, then we have

$$\mu_\omega(x,y) = \sum_{p \le y} \frac{1}{p^\alpha} + O(1)$$

By using (4.2.5), we get

$$\mu_\omega(x,y) = \log \log y + u + O(1),$$

Now by letting $u = o(\log \log y)$, we have

$$\mu_\omega(x,y) = \log \log y + o(\log \log y),$$

and the proof is complete. □

Lemma 4.2.4. *If $u = o(\log\log y)$ and $t \le y^{1/\log u}$, then we have*

$$\sum_{p\le t} \frac{1}{p^\alpha} = \log\log t + O(1) \tag{4.2.7}$$

Proof. We have

$$\sum_{p\le t} \frac{1}{p^\alpha} = \sum_{p\le t} \frac{1}{pp^{\alpha-1}} = \sum_{p\le t} \frac{1}{p}\left\{1 + O\left((1-\alpha)\log p\right)\right\},$$

since $(1-\alpha)\log p$ is bounded. By the given estimate for α in (4.2.2) and using Mertens' estimate, we obtain

$$\begin{aligned}\sum_{p\le t} \frac{1}{p^\alpha} &= \sum_{p\le t} \frac{1}{p} + O\left(\frac{\xi(u)}{\log y}\sum_{p\le t}\frac{\log p}{p}\right)\\ &= \log\log t + O\left(\frac{\xi(u)}{\log y}\log t\right)\end{aligned} \tag{4.2.8}$$

By applying the estimate of $\xi(u)$ in (4.2.3), we get our desired result. □

Here we will introduce a truncated version of ω and in the following lemma and corollary we show that the contribution of large prime factors does not affect the expected value of number of prime factors of n and hence the distribution of $\omega(n)$, when u is small enough. We define

$$\omega_Y(n) := \sum_{p\le Y} \mathbb{1}_{p|n}(y), \tag{4.2.9}$$

where

$$Y := y^{\frac{1}{\phi(y)}}, \quad \text{and} \quad \phi(y) := (\log\log y)^{\sqrt{\log\log\log y}}.$$

Lemma 4.2.5. *If $u = o(\log\log y)$, then we have*

$$\sum_{p\le Y} \frac{1}{p^\alpha} = \log\log y + O\left((\log\log\log y)^{3/2}\right).$$

Proof. By Lemma 4.2.4, we have

$$\begin{aligned}\sum_{p\le Y} \frac{1}{p^\alpha} &= \log\log y - \log\phi(y) + O(1)\\ &= \log\log y + (\log\log\log y)^{3/2} + O(1),\end{aligned} \tag{4.2.10}$$

and we have our desired result. □

Now we define

$$\mu_{\omega_Y}(x,y) := \mathbb{E}\left[\omega_Y(n)\right].$$

In the following lemma we will show $\omega(n)$ can be replaced by $\omega_Y(n)$ in the statement of Theorem 4.1.1.

Lemma 4.2.6. *Let $h(n) := \omega(n) - \omega_Y(n)$, then we have*

$$\mathbb{P}\left(|h| \leq (\log\log y)^{1/4}\right) = 1 - o(1),$$

where $\mathbb{P}$ denotes the probability value.

Proof. We first find an estimate for $\mathbb{E}[h]$, we have

$$\mathbb{E}[h] = \mathbb{E}\left[\omega(n) - \omega_Y(n)\right] = \mu_\omega(x,y) - \mu_{\omega_Y}(x,y).$$

Using Lemma 4.2.3 and 4.2.5, we get

$$\mathbb{E}[h] \ll (\log\log\log y)^{3/2} \leq (\sqrt{\log\log y}). \tag{4.2.11}$$

For the variance of h, using (4.2.11), we get

$$\begin{aligned} \sigma_h^2(x,y) &:= \mathbb{E}\left[(h - \mathbb{E}[h])^2\right] \\ &= (\mathbb{E}[h])^2 \ll (\log\log\log y)^3. \end{aligned} \tag{4.2.12}$$

Now by Chebyshev's inequality and using (4.2.12), we have

$$\begin{aligned} \mathbb{E}\left(h \geq (\log\log y)^{1/4}\right) &\leq \mathbb{P}\left(\left|h - \mathbb{E}[h]\right| \geq (\log\log y)^{1/4}\right) \\ &\leq \frac{\sigma_h^2(x,y)}{(\log\log y)^{1/2}} = o(1), \end{aligned} \tag{4.2.13}$$

and we get our desired result. □

By the above Lemma and recalling Remark 4.2.1, the estimate in (4.1.4) is equivalent to the following

$$\frac{1}{\Psi(x,y)}\#\{n \in S(x,y) : \frac{\omega_Y(n) - \log\log y}{\sqrt{\log\log y}} \leq z\} \to \Phi(z) \quad (y \to \infty), \tag{4.2.14}$$

which we prove it in the next section.

4.3. Proof of Theorem 4.1.1

We begin this section by setting some random variables X_p on a probability space and one variable for each prime p, which satisfies

$$P(X_p = 1) = \frac{\Psi(\frac{x}{p}, y)}{\Psi(x,y)}, \quad \text{and} \quad P(X_p = 0) = 1 - \frac{\Psi(\frac{x}{p}, y)}{\Psi(x,y)} \tag{4.3.1}$$

The random variables X_p's are independent.

Now we define the partial sum S_Y as follows

$$S_Y := \sum_{p \leq Y} X_p,$$

where $Y = y^{1/\phi(y)}$.
By the definition of X_p's and the estimate in (4.2.1) and (4.2.5), we deduce that S_Y has a mean value and variance of the order $\log\log y$ in the range $u = o(\log\log y)$, this means that $\omega_Y(n)$ and S_Y have roughly the same variance and the same mean value.

In the following lemma we get an upper bound for the difference of jth moments of ω_Y and S_Y, where $j = 1,2,3,\dots$.

Lemma 4.3.1. *If $u = o(\log\log y)$, then for any positive integer j, we have*

$$A_j := \mathbb{E}_{n\in S(x,y)}[\omega_Y(n)^j] - \mathbb{E}[S_Y^j] \ll \frac{(\log\log y)^j}{u(\log\log y)^{\sqrt{\log\log\log y}}}.$$

Proof. By the definition of ω_Y and S_Y, we have

$$\mathbb{E}[\omega_Y^j(n)] = \frac{1}{\Psi(x,y)} \sum_{p_1 \dots p_j \leq Y} \sum_{n \in S(x,y)} \mathbb{1}_{p_1|n}(n) \dots \mathbb{1}_{p_j|n}(n),$$

and

$$\mathbb{E}[S_Y^j] = \sum_{p_1 \dots p_j \leq Y} \mathbb{E}\left[X_{p_1} \dots X_{p_j}\right].$$

So for the difference of jth moment, we have

$$\begin{aligned} A_j &= \sum_{p_1,\ldots,p_j \leq y^{\frac{1}{\phi(y)}}} \left(\frac{1}{\Psi(x,y)} \sum_{n \in S(x,y)} \mathbb{1}_{p_1|n}(n) \ldots \mathbb{1}_{p_j|n}(n) - \mathbb{E}[X_{p_1} \ldots X_{p_j}] \right) \\ &= \sum_{p_1,\ldots,p_j \leq y^{\frac{1}{\phi(y)}}} \left[\frac{\Psi(\frac{x}{p_1 \ldots p_j}, y)}{\Psi(x,y)} - \prod_{1 \leq i \leq j} \frac{\Psi(\frac{x}{p_i}, y)}{\Psi(x,y)} \right] \\ &= \sum_{p_1,\ldots,p_j \leq y^{\frac{1}{\phi(y)}}} \left[\frac{\Psi(\frac{x}{p_1,\ldots,p_j}, y)}{\Psi(x,y)} - \prod_{1 \leq i \leq j} \frac{\Psi(\frac{x}{p_i}, y)}{\Psi(x,y)} \right]. \end{aligned} \tag{4.3.2}$$

Without loss of generality we assume that p_i's are distinct, then by using the estimate (4.2.1), we have

$$\begin{aligned} A_j = &\sum_{p_1,\ldots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \ldots p_j)^\alpha} \left\{ 1 + O\left(\frac{1}{u_y} + \frac{\log p_1 \ldots p_j}{\log x} \right) \right\} \\ &- \sum_{p_1,\ldots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \ldots p_j)^\alpha} \prod_{i=1}^{j} \left\{ 1 + O\left(\frac{1}{u_y} + \frac{\log p_i}{\log x} \right) \right\}. \end{aligned}$$

The main terms in the above subtraction are the same and will be eliminated. Therefore,

$$\begin{aligned} A_j &\ll \sum_{p_1,\ldots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \ldots p_j)^\alpha} \left(\frac{1}{u_y} + \frac{\log p_1 \ldots p_j}{\log x} \right) \\ &\ll \sum_{p_1,\ldots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \ldots p_j)^\alpha} \left(\frac{1}{u_y} + \frac{\log y}{\phi(y) \log x} \right). \end{aligned} \tag{4.3.3}$$

If $u = o(\log \log y)$, then $u_y \geq \frac{\log y}{\log \log \log y}$. So we can ignore the term $\frac{1}{u_y}$. Thus,

$$A_j \ll \sum_{p_1,\ldots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \ldots p_j)^\alpha} \left(\frac{\log y}{\phi(y) \log x} \right).$$

We now use Lemma 4.2.5, and we get the following upper bound for each A_j

$$A_j \ll \frac{(\log \log y)^j}{u (\log \log y)^{\sqrt{\log \log \log y}}}. \tag{4.3.4}$$

□

Proof of Theorem 4.1.1. We start our proof by normalizing the random variable S_Y. Define

$$S := \frac{S_Y - \mu_{\omega_Y}(x,y)}{\sqrt{\sigma^2_{\omega_Y}(x,y)}}.$$

By recalling the central limit theorem, one can say that S has a normal distribution $\Phi(x)$, since X_p's are independent. We set

$$W := \frac{\omega_Y(n) - \mu_{\omega_Y}(x,y)}{\sqrt{\sigma^2_{\omega_Y}(x,y)}}.$$

By using the method of moments, we will show that the moments of W are very close to those corresponding sum S and they both converge to the kth moment of normal distribution for every positive integer k.
By the multinomial theorem, we have

$$\begin{aligned}\Delta^k &:= \mathbb{E}[(\omega_Y(n) - \mu_{\omega_Y}(x,y))^k] - \mathbb{E}[(S_Y - \mu_{\omega_Y(x,y)})^k] \\ &= \sum_{j=1}^{k} \binom{k}{j} (-\mu_{\omega_Y}(x,y))^{k-j} \left(\mathbb{E}[\omega_Y(n)^j] - \mathbb{E}[S_Y^j]\right).\end{aligned} \tag{4.3.5}$$

By combining the upper bound in (4.3.4) with (4.3.5), we arrive to the following estimate

$$\begin{aligned}\Delta^k &\ll \frac{1}{(\log\log y)^{\sqrt{\log\log\log y}}} \sum_{j=1}^{k} \binom{k}{j} (-\mu_{\omega_Y}(x,y))^{k-j} (\log\log y)^j \\ &= \frac{1}{u(\log\log y)^{\sqrt{\log\log\log y}}} \left(\log\log y + \mu_{\omega_Y}(x,y)\right)^k.\end{aligned} \tag{4.3.6}$$

Now using Lemma 4.2.3, we have

$$\Delta^k \ll \frac{(\log\log y)^k}{u(\log\log y)^{\sqrt{\log\log\log y}}}. \tag{4.3.7}$$

Thus,

$$\Delta^k \to 0 \quad \text{as} \quad x,y \to \infty.$$

We showed that the difference of kth moments goes to 0 for large values of y. By the remark (4.2.2), we conclude that two random variables S and W have a same distribution.

By Remark 4.2.4, the random variable S has a normal distribution. It remains to show that the moments of S are very close to those of the normal distribution.
By recalling Remark 4.2.3, we need to prove that the moment $E[S^k]$ are bounded in n when k increases.
In fact, we will show that for each $k \in \mathbb{N}$

$$\sup_n \left| \mathbb{E}\left(\frac{(S_Y - \mu_{\omega_Y}(x,y))^k}{(\sqrt{\sigma^2_{\omega_Y}(x,y)})^k} \right) \right| < \infty. \tag{4.3.8}$$

To complete the proof, we define the random variables $Y_p = X_p - \frac{\Psi(x/p,y)}{\Psi(x,y)}$, which are independent.
We have

$$\mathbb{E}\left((S_Y - \mu_{\omega_Y}(x,y))^k\right) = \sum_{j=1}^{k} \sum{}' \frac{k!}{k_1!...k_j!} \sum_{p_1...p_j \le y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_{p_1}^{k_1}]...\mathbb{E}[Y_{p_j}^{k_j}]. \tag{4.3.9}$$

Where $\sum'$ is over j-tuple $(k_1,..,k_j)$, where $k_1,\ldots,k_j$ are positive integers, and $k_1 + ... + k_j = k$.
By the definition of $Y_p's$, we have $\mathbb{E}[Y_{p_j}] = 0$.
To avoid zero terms, we can assume that $k_i > 1$ for each $1 \le i \le j$. Also we have $|Y_p| \le 1$. Thus,

$$\mathbb{E}[Y_p^{k_i}] \le \mathbb{E}[Y_p^2] \quad \forall k_i > 2.$$

Therefore, the value of inner sum in (4.3.9) is at most

$$\sum_{p_1...p_j \le y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_{p_1}^{k_1}]...\mathbb{E}[Y_{p_j}^{k_j}] \le \left(\sum_{p \le y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_p^2] \right)^j = \sigma^{2j}(x,y).$$

Each k_i is strictly greater than 1, and we have $k_1 + .. + k_j = k$, therefore $2j \le k$ and this implies that

$$\mathbb{E}\left(\frac{(S_Y - \mu_{\omega_Y}(x,y))^k}{\left(\sqrt{\sigma^2_{\omega_Y}(x,y)}\right)^k} \right) \le \sum_{j=1}^{k} \sum{}' \frac{k!}{k_1!...k_j!},$$

from which (4.3.8) follows.

We proved all necessary and sufficient conditions such that (4.2.14) and consequently (4.1.4) are true. □

Bibliography

[1] Krishnaswami Alladi. The Turán-Kubilius inequality for integers without large prime factors. *J. Reine Angew. Math.*, 335:180–196, 1982. ISSN 0075-4102. doi: 10.1515/crll.1982.335.180. URL http://dx.doi.org/10.1515/crll.1982.335.180.

[2] Krishnaswami Alladi. An Erdős-Kac theorem for integers without large prime factors. *Acta Arith.*, 49(1):81–105, 1987. ISSN 0065-1036.

[3] William D. Banks and David J. Covert. Sums and products with smooth numbers. *J. Number Theory*, 131(6):985–993, 2011. ISSN 0022-314X. doi: 10.1016/j.jnt.2010.11.001. URL http://dx.doi.org/10.1016/j.jnt.2010.11.001.

[4] A. S. Besicovitch. On the density of certain sequences of integers. *Math. Ann.*, 110(1): 336–341, 1935. ISSN 0025-5831. doi: 10.1007/BF01448032. URL http://dx.doi.org/10.1007/BF01448032.

[5] Patrick Billingsley. On the central limit theorem for the prime divisor functions. *Amer. Math. Monthly*, 76:132–139, 1969. ISSN 0002-9890.

[6] E. R. Canfield, Paul Erd˝ os, and Carl Pomerance. On a problem of Oppenheim concerning "factorisatio numerorum". *J. Number Theory*, 17(1):1–28, 1983. ISSN 0022-314X. doi: 10.1016/0022-314X(83)90002-1. URL http://dx.doi.org/10.1016/0022-314X(83)90002-1.

[7] N. G. de Bruijn. On the number of positive integers $\leq x$ and free of prime factors $> y$. *Nederl. Acad. Wetensch. Proc. Ser. A.*, 54:50–60, 1951.

[8] R. de la Bretèche and G. Tenenbaum. Entiers friables: inégalité de Turán-Kubilius et applications. *Invent. Math.*, 159(3):531–588, 2005. ISSN 0020-9910. doi: 10.1007/s00222-004-0379-y. URL http://dx.doi.org/10.1007/s00222-004-0379-y.

[9] Régis de la Bretèche and Gérald Tenenbaum. Propriétés statistiques des entiers friables. *Ramanujan J.*, 9(1-2):139–202, 2005. ISSN 1382-4090. doi: 10.1007/s11139-005-0832-6.

URL http://dx.doi.org/10.1007/s11139-005-0832-6.

[10] Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010. ISBN 978-0-521-76539-8. doi: 10.1017/CBO9780511779398. URL http://dx.doi.org/10.1017/CBO9780511779398.

[11] Veikko Ennola. On numbers with small prime divisors. *Ann. Acad. Sci. Fenn. Ser. A I No.*, 440:16, 1969.

[12] P. Erdős. An asymptotic in the theory of numbers.(russian). *Vestnik Leningrad. Univ.*, 15(13):41–49, 1960.

[13] P. Erdös and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, 62:738–742, 1940. ISSN 0002-9327.

[14] Paul Erdös. Some remarks on number theory. *Riveon Lematematika*, 9:45–48, 1955.

[15] William Feller. *An introduction to probability theory and its applications*, volume 1. The name of the publisher, 3 edition, 1968. ISBN 978-0471257080.

[16] Kevin Ford. Integers with a divisor in $(y, 2y]$. In *Anatomy of integers*, volume 46 of *CRM Proc. Lecture Notes*, pages 65–80. Amer. Math. Soc., Providence, RI, 2008.

[17] John B. Friedlander. Integers free from large and small primes. *Proc. London Math. Soc. (3)*, 33(3):565–576, 1976. ISSN 0024-6115.

[18] Andrew Granville. Smooth numbers: computational number theory and beyond. In *Algorithmic number theory: lattices, number fields, curves and cryptography*, volume 44 of *Math. Sci. Res. Inst. Publ.*, pages 267–323. Cambridge Univ. Press, Cambridge, 2008.

[19] Andrew Granville and K. Soundararajan. Sieving and the Erdős-Kac theorem. In *Equidistribution in number theory, an introduction*, volume 237 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 15–27. Springer, Dordrecht, 2007. doi: 10.1007/978-1-4020-5404-4_2. URL http://dx.doi.org/10.1007/978-1-4020-5404-4_2.

[20] H. Halberstam and H.-E. Richert. *Sieve methods*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.

[21] Richard R. Hall and Gérald Tenenbaum. *Divisors*, volume 90 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1988. ISBN 0-521-34056-X. doi: 10.1017/CBO9780511566004. URL http://dx.doi.org/10.1017/CBO9780511566004.

[22] Douglas Hensley. The distribution of $\Omega(n)$ among numbers with no large prime factors. In *Analytic number theory and Diophantine problems (Stillwater, OK, 1984)*, volume 70

of *Progr. Math.*, pages 247–281. Birkhäuser Boston, Boston, MA, 1987.

[23] Adolf Hildebrand. On the number of positive integers $\leq x$ and free of prime factors $> y$. *J. Number Theory*, 22(3):289–307, 1986. ISSN 0022-314X. doi: 10.1016/0022-314X(86)90013-2. URL http://dx.doi.org/10.1016/0022-314X(86)90013-2.

[24] Adolf Hildebrand. On the number of prime factors of integers without large prime divisors. *J. Number Theory*, 25(1):81–106, 1987. ISSN 0022-314X. doi: 10.1016/0022-314X(87)90017-5. URL http://dx.doi.org/10.1016/0022-314X(87)90017-5.

[25] Adolf Hildebrand and Gérald Tenenbaum. On integers free of large prime factors. *Trans. Amer. Math. Soc.*, 296(1):265–290, 1986. ISSN 0002-9947.

[26] Dimitris Koukoulopoulos. *Generalized and restricted multiplication tables of integers.* PhD thesis, University of Illinois at Urbana-Champaign, 2010.

[27] D. J. Newman. Simple analytic proof of the prime number theorem. *Amer. Math. Monthly*, 87(9):693–696, 1980. ISSN 0002-9890. doi: 10.2307/2321853. URL http://dx.doi.org/10.2307/2321853.

[28] Éric Saias. Sur le nombre des entiers sans grand facteur premier. *J. Number Theory*, 32(1):78–99, 1989. ISSN 0022-314X. doi: 10.1016/0022-314X(89)90099-1. URL http://dx.doi.org/10.1016/0022-314X(89)90099-1.

[29] Éric Saias. Entiers sans grand ni petit facteur premier. II. *Acta Arith.*, 63(4):287–312, 1993. ISSN 0065-1036.

[30] Éric Saias. Entiers sans grand ni petit facteur premier. III. *Acta Arith.*, 71(4):351–379, 1995. ISSN 0065-1036.

[31] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory.* Cambridge University Press, 1995. ISBN 978-0521412612.

[32] Gérald Tenenbaum. On ultrafriable integers. *Q. J. Math.*, 66(1):333–351, 2015. ISSN 0033-5606. doi: 10.1093/qmath/hau025. URL http://dx.doi.org/10.1093/qmath/hau025.

[33] Ti Zuo Xuan. The Turán-Kubilius inequality for integers free of large prime factors. II. *Acta Arith.*, 65(4):329–352, 1993. ISSN 0065-1036.